

1 **TITLE: Development and validity of a questionnaire of Patients' Experiences in**
2 **Post-Acute Outpatient Physical Therapy settings (PEPAP-Q).**

3

4 Francesc Medina-Mirapeix PT PhD¹, María Elena del Baño Aledo PT PhD², Jacinto
5 Javier Martínez Payá PT PhD², María Carmen Lillo Tomás PT PhD³, Pilar Escolar
6 Reina PT PhD⁴.

7

8 ¹ Chairperson & Professor, Department of Physical therapy, University of Murcia,
9 Spain

10 ² Professor, Central Unit of Anatomy, Catholic University San Antonio, Spain

11 ³ Professor, Department of Physical therapy, Miguel Hernandez University of Elche,
12 Spain

13 ⁴ Professor, Department of Physical therapy, University of Murcia, Spain

14

15

16 Correspondence:

17 María Elena del Baño Aledo. Central Unit of Anatomy, Catholic University San
18 Antonio. Campus de Los Jerónimos s/n. Guadalupe, 30107 (Murcia) Spain. E-mail:
19 mbano@pdi.ucam.edu Telephone: (+34) 968 27 88 00

20

21

1 **ABSTRACT**

2 **Background:** Patient feedback surveys are increasingly seen as a key component of
3 healthcare quality monitoring and improvement. **Objective:** To describe the
4 development and initial psychometric evaluation of a fixed-length questionnaire of
5 Patients' Experiences undergoing physical therapy treatment in a Post-Acute
6 **Outpatients** settings (PEPAP-Q). **Design:** An instrument development study with
7 validity and reliability testing. **Methods:** A total of 465 patients from three
8 rehabilitation centres for musculoskeletal conditions completed the questionnaire. A
9 cognitive pretest was applied to the draft version (n=94), and a revised version was
10 tested for test-retest reliability (n=90). Analyses to evaluate variance and non-response
11 rates to items, the factor structure of the questionnaire and the metric properties of
12 multi-item scales were conducted. **Results:** Exploratory factor analyses yielded
13 evidence for a seven-factor structure to the PEPAP-Q, with three factors that may be
14 conceptually viewed as professionals' attitudes and behaviour (providing information
15 and education, sensitive manners to patients' changes and emotional support) and four
16 factors conceptually reflecting organizational environment (attendance duration,
17 interruptions during care delivery, waiting times and patient safety). Item scale
18 correlations ranged from 0.70 to 0.93. The percentage of scaling success was 100% for
19 all the scales. Cronbach's alpha coefficient ranged from 0.70 to 0.87. Intraclass
20 correlation coefficients ranged from 0.57 to 0.80 (median=0.68). **Limitations:**
21 Generalization to other patients is not known. **Conclusion:** The PEPAP-Q is a reliable
22 test-retest and the scales have internal consistency and convergent and discriminant
23 validity. All the scales are distinct and unidimensional.

24 **Key words:** Quality of Health Care, Physical Therapy Department, Outpatients.

1 **Word Count:** 238 (Abstract) and 3874 (Introduction, Methods, Results, Discussion)
2

1 INTRODUCTION

2 Many clinical settings with physical therapy services routinely ask people for feedback
3 about the care that they have received.¹⁻³ Such questions have primarily elicited
4 information on satisfaction with services.⁴⁻⁷ However, health service research has
5 recently criticized the concept of patient satisfaction because of its inherent sources of
6 bias.^{8,9} It is widely recognised, for example, that patients remain reluctant to express
7 low satisfaction about their care and, as a consequence, patient satisfaction
8 questionnaires provide an optimistic picture of performance.¹⁰ Patients can even
9 describe high levels of satisfaction at the same time as describing experiences that are
10 suboptimal.⁹

11 Concern about the problems with patient satisfaction surveys has led to an emphasis on
12 measuring patients' experiences rather than satisfaction.¹¹ Patient satisfaction and
13 experience are closely linked, but have distinct meanings and ways to be measured.
14 'Experience' is related to things that happened during care and the extent to which
15 people's needs were met, whereas 'satisfaction' is related more to how people feel about
16 those things.¹² While patient satisfaction questionnaires ask respondents to rate their
17 care using general evaluation categories (e.g., excellent), patient experience
18 questionnaires ask respondents to say whether or not certain processes or events
19 occurred during a particular visit or a care episode.¹³

20 **A recent review illustrated that a significant number of generic tools have been**
21 **developed to assess patients' experiences in primary care and hospital settings.**¹²

22 Some examples are the Picker Patient experience Questionnaire (PPE),⁸ the NHS
23 National Adult Inpatient Survey¹⁴, the Patient Experience Questionnaire (PEQ) in
24 primary health care¹⁵ and the Generic Short Patient Experiences Questionnaire (GS-

1 PEQ)¹⁶. Two of these generic tools have been used to capture the experiences of
2 outpatient attending physical therapy services (GS-PEQ¹⁶ and PPE⁸).

3 The GS-PEQ, created by Sjetne et al in 2011,¹⁶ includes 10 items asking about outcome,
4 clinician services, user involvement, incorrect treatment, information, organization and
5 accessibility. **GS-PEQ has been applied to physical therapy services, but it was not**
6 **specifically designed for this context. This, plus its brevity, means that the**
7 **questionnaire results are not specific enough for small units within an**
8 **organization.**¹⁷ The PPE was adapted to capture the experiences of people with specific
9 musculoskeletal disorders (The Picker MSD questionnaire)¹⁸ and used in clinics with
10 multidisciplinary services. The Picker MSD questionnaire is a 30-item instrument
11 grouped in seven scales that were identified in a previous qualitative study as the
12 aspects that matter to patients with musculoskeletal disorders: accessibility, information,
13 respect, emotional support, coordination, continuity and overall impression. Content
14 validity of these scales was based on qualitative analysis of two focus groups but no
15 additional analysis was carried out to evaluate the dimensionality of each scale.¹⁸

16 Despite the great amount of active research and scientific publications in the field of
17 patients' experience and satisfaction^{8,14-18} very few **context-specific** questionnaires
18 measure patient experience in physical therapy services. Some exceptions are the
19 perceived quality questionnaire in primary physical therapy care (PQ-PPC)¹⁹ and the
20 client-centred rehabilitation questionnaire²⁰ (CCRQ) **for inpatient care**. The PQ-PPC
21 measures 4 dimensions (waiting times, organization, professional competence and
22 information). **The CCRQ measures 7 dimensions of client-centred rehabilitation in**
23 **inpatient units (decision-making, information, outcome evaluation, family**
24 **involvement, emotional support, physical comfort and continuity) identified based**
25 **on previous works from the Picker Institute⁸ and the Wascana group²¹, and focus**

1 **groups with clients.** These instruments reflect substantial variations in the dimensions
2 of care that matter to patients in different contexts. Some authors have indicated that
3 useful questionnaires should explore patients' experiences in each specific context,
4 across well-defined domains.²⁰

5 Both the PQ-PPC and the CCRQ were developed using the same methodology of Picker
6 MSD and, as consequence the structure and content validity of their scales is supported
7 just by a conceptual approach defined by qualitative analysis of focus groups, while
8 their unidimensionality was not specifically examined. The notion of unidimensionality
9 in measurement theory refers to a condition in which a set of indicators share only one
10 underlying factor.²² The assessment of unidimensionality has been presented as a top
11 priority in any outcome instruments used in rehabilitation.^{23,24} **In fact, when items do**
12 **not fit a common unidimensional scale, all calculations using a total score, with the**
13 **exception of statistical correlations with other quantitative criteria (predictive**
14 **validity), become uninterpretable.**²⁵ Yet, little evidence about the dimensionality of
15 patients' experience scales is available.

16 Our research group developed a questionnaire on patients' experiences in primary care
17 using a **qualitative approach** to develop the scales.¹⁹ The outpatient setting studied
18 provides physical therapy to people with chronic clinical conditions, but who usually
19 continue working and have a low level of dependency. The next phase of our research,
20 which is the focus of the present paper, examined the use of **a psychometric**
21 **methodology to develop scales for measuring** patients' experiences with physical
22 therapy services in a **post-acute, outpatient setting.** **We selected this context because**
23 **it remains unexplored by existing patient experience questionnaires. In Spain,**
24 **there are outpatient settings that solely provide post-acute care to patients after an**
25 **acute care admission and others that see ambulatory patients also. We selected the**

1 **first setting because all the patients are receiving early rehabilitation and are in a**
2 **dependent situation with problems of pain, mobility and functionality.**^{23,24,26}

3 In this article, therefore, we describe the development and initial psychometric
4 evaluation of a fixed-length questionnaire of Patients' Experiences undergoing a
5 physical therapy treatment in Post-Acute **Outpatients** settings (PEPAP-Q). The specific
6 aims of psychometric evaluation are to examine: (1) the variance and non-response rates
7 to items, (2) the factor structure of the questionnaire, (3) the unidimensionality of each
8 scale, (4) test-retest reliability and (5) construct validity.

9

1 **METHODS**

2 The instrument development was conducted in three phases. In Phase 1, nine domains
3 of patient experiences were identified based on a literature review, focus groups with
4 patients, and review by experts. In Phase 2, several items for each domain were
5 developed and tested for clarity and relevance by conducting cognitive pretesting with
6 rehabilitation patients and physical therapists. In Phase 3, the psychometric properties of
7 the tool were examined using a self-administered survey of patients discharged from
8 three physical therapy facilities. In this paper, we look at item reduction, factor structure
9 and the unidimensionality of scales, test-retest reliability, and construct validity testing.

10

11 **Phase 1: Domains of patients' experiences in physical therapy settings**

12 Several models have been proposed to identify what specific aspects of health care are
13 relevant to patients.^{27,28} According to the research evidence, most instruments consider
14 three domains as the highest priority for measuring patient experience: **the**
15 **characteristics of interaction (patient-professional relationship), elements of the**
16 **organization of the service (waiting times, facilities, security...) and overall**
17 **assessments (general satisfaction, perceived quality).**¹²

18 In order to better understand the importance given to experiences of physical therapy
19 from the client's perspective, **nine focus groups were set up with adults involving**
20 **musculoskeletal disorders who were receiving post-acute physical therapy service**
21 **in an outpatients setting.** A semistructured interview topic guide formed from a
22 literature review was used. Additional questions were included according to themes that
23 started to emerge from the initial focus groups. Data analysis was undertaken using
24 grounded theory.

1 The results of this qualitative phase are reported elsewhere.^{29,30} In summary, nine
2 domains grouped in the following two areas were identified: professionals' attitudes and
3 behaviour, and organizational environment. In the professionals' attitudes and
4 behaviour area, domains focused on: 'providing information and education', 'sensitive
5 manners to patients' changes', 'emotional support', 'friendly and respectful
6 communication' and 'technical expertise'. In the organizational environment area,
7 domains focused on: 'attendance duration', 'interruptions during care delivery',
8 'waiting times in the sequence of treatment' and 'patient safety'.

9

10 **Phase 2: Development of items**

11 The frequency with which certain processes and events occurred during the course of
12 physical therapy treatment was investigated. For this, several candidate items were
13 written for each domain. The precise wording of items was based on patients'
14 comments of the qualitative study and modified through a process of discussion and
15 consensus among the members of the study team. The response format used was a 5-
16 point frequency Likert scale. Response options ranged from "never" to "always". An
17 initial pool of 30 items was generated, distributed among nine domains: 'information'
18 (4 items), 'sensitive' (3 items), 'support' (5 items), 'friendliness' (4 items), 'expertise'
19 (3 items), 'attendance duration' (3 items), 'interruptions' (3 items), 'waiting times' (3
20 items) and 'safety' (2 items).

21 The questionnaire was pretested for clarity and adequacy of content with seven
22 professionals (mean care provision experience: 10 years) and patients from the
23 participant settings. A total of 94 patients with musculoskeletal disorders were

1 randomly selected from those receiving post-acute physical therapy. Both professionals
2 and patients were asked to report on the relevance and understanding of each item.

3 A question was taken to have poor face validity if less than 80% of either patients or
4 professionals rated the item as being comprehensible and adequate. Thus, items in the
5 technical expertise domain were considered to be unsuitable because many physical
6 therapists thought that patients are not capable of evaluating professional competence.
7 For this reason, the three items of 'technical expertise' were removed from the
8 questionnaire.

9 The instrument finally distributed for evaluation comprised 27 items of patients'
10 experiences with physical therapy. **Two overall quality evaluations, one item about**
11 **treatment adherence and 4 questions related to patients' characteristics were also**
12 **included. Patients' overall evaluations were assessed using two additional items.**³¹
13 **The first – referred to as the 'satisfaction item' – asked patients to rate their**
14 **satisfaction with care on a 10-point scale ranging from 1 (minimum) to 10**
15 **(maximum). The second – referred to as 'perceived service quality item' – asked**
16 **patients to rate their perceptions of service quality on a 5-point scale (with**
17 **response categories of poor, fair, good, very good, and excellent). Adherence to**
18 **prescribed exercise (if applicable) was measured using a frequency-based response**
19 **scale (never, seldom, often, almost always, and always) adapted from the**
20 **adherence scale of Sluijs et al.**³²

21

22 **Phase 3: Pilot study and psychometric testing**

23 In Phase 3 psychometric testing of the questionnaire was carried out by surveying
24 subjects who were receiving post-acute physical therapy **in outpatients** services from 3

1 post-acute centers in Barcelona, Madrid and Seville, Spain. During a six-month period,
2 a consecutive sample of eligible patients was identified from the patients' register and
3 recruited by the attending physiotherapist who assessed the eligibility criteria. The
4 subjects were included if they were more than 18 years old and receiving physical
5 therapy care. They had to speak, read and understand Spanish. Patients were excluded if
6 they presented any cognitive deficit.

7 The study was approved by the Bioethics Committee of the University of Murcia and an
8 informed consent form was obtained from all participants. The questionnaire was
9 administered during patients' clinical appointments at the end of the course of the
10 specific episode of physical therapy. The questionnaire and an introductory letter were
11 handled by the attending physical therapist. Patients were asked to place the
12 questionnaire in a box in the administrative area of the center within the following 1-2
13 days. If necessary, non-respondents received a reminder letter handed to them by the
14 physical therapist after 1 week and another letter 2 weeks later. Patients were assured
15 that the professionals would not be able to identify individual answers. To test-retest
16 reliability a subsample of patients completed the questionnaire a second time within 5
17 days of their initial response.

18

19 **Data analysis**

20 Descriptive statistics were calculated to assess the demographic and clinical
21 characteristics of the study participants. **We conducted a series of analyses to assess**
22 **item reduction, the factor structure of the questionnaire and the metric properties**
23 **of multi-item scales. Finally, the underlying structure of these scales was examined.**

24

1 *Item reduction and instrument structure*

2 **Variance and non-response rates to items were evaluated for item reduction. As**
3 **described elsewhere³³ items with high rates of non-response ($\geq 10\%$) and SD less**
4 **than 0.60 were removed to avoid items with poor understanding and/or variability.**

5 Exploratory factor analysis (EFA) was performed to identify latent factors that could be
6 responsible for the covariation of the data. Principal component analysis and Varimax
7 rotation were used for the initial extraction of factors. Items with loadings of 0.50 or
8 higher were retained and items with factor loadings higher than 0.40 or more than 1
9 factor were removed.²⁴ To examine the stability of this structure, we also performed
10 supplementary EFAs in two subgroups: patients with a high level of satisfaction **with**
11 **care** (above the mean) and unsatisfied ones (below the mean). The Kaiser-Meyer-Olkin
12 (KMO) value (preferably > 0.60) and Bartlett's test of sphericity (preferably significant)
13 were used to assess the **sampling adequacy** for factorisation.

14 Multitrait scaling analysis uses a correlation matrix of all items and scales to test the
15 extent to which items converge and diverge from other scales. **Scales were scored**
16 **using a summative method, as described elsewhere.³¹ Each item was coded as a**
17 **dichotomous score, indicating either presence or absence of a problem. A problem**
18 **is defined as an aspect of healthcare that could, from the patients' perspective, be**
19 **improved. Thus, an 'always' response was considered an absence of problem in**
20 **direct items. We assumed that any other response categories could show at least**
21 **one problematic event during the episode of care. For each respondent, a score for**
22 **each of the scales of the aforementioned instrument were created. Each scale was**
23 **scored from 0 (no problem) to 100 (all items coded as a problem).** Correlations
24 between each item and its hypothesized scale were calculated and corrected for overlap

1 by not including the item in the scale. A correlation of 0.40 or higher was considered
2 satisfactory; items with a correlation of less than 0.40 were removed from further
3 analyses.³⁴ Scaling success rates were computed for each scale as the percentage of
4 items within a scale which correlated more highly or significantly more highly with
5 their hypothesized scale than with the other scales. An item correlated significantly
6 more highly with its own scale if the correlation between this item and its hypothesized
7 scale was more than two standard errors higher than its correlation with other scales.³⁵
8 In addition, we calculated internal consistency reliability for each hypothesized domain
9 using the Cronbach alpha coefficient (considered acceptable if α is greater than 0.7).

10

11 *Reliability and validity*

12 The reliability of the PEPAP-Q scales was assessed using test-retest methods. Intraclass
13 correlation coefficients (ICC) were calculated for test-retest reliability using a two-way
14 mixed effect, single-rater model (judged excellent if > 0.75).

15 Several construct validation strategies were designed in which we hypothesized various
16 associations and evaluated the extent to which our data corresponded to the hypotheses.

17 Known-groups validity was examined by defining two groups: one subgroup of patients
18 that rated the overall quality item as “very good” or “excellent” and the subgroup that
19 evaluated it worse. We hypothesized that the problem scores in all the scales would be
20 lower in patients with high evaluations of service quality and higher in those with poor
21 quality perception. A Student t-test was used to test for a difference in the mean scale
22 scores between these two groups. Also, the mean scores of the response data to the
23 questionnaire across men and women were compared using a Student t-test for

1 independent samples. Our hypothesis was that there would be no significant difference
2 in the mean scores between these two groups.

3 Convergent validity was tested by calculating correlations between scales and patient
4 satisfaction, while correlations with the treatment adherence item were used for testing
5 discriminative validity. *A priori* hypotheses for these correlations were as follows: (1)
6 the scales scores would correlate strongly (> 0.30) with patient satisfaction; and (2)
7 none of the scales scores would correlate significantly with the treatment adherence
8 item. The Spearman correlation coefficient was used in these analyses. A p-value of less
9 than 0.05 was considered statistically significant. All analyses were performed using the
10 SPSS statistical software program (SPSS v.15).

11

12 **RESULTS**

13 During a six-month period 520 subjects were identified. The data of four participants
14 were not taken into account because they were not able to understand the instructions
15 for the questionnaire. Out of 516 patients, 355 patients responded initially and another
16 110 patients responded after reminders. In total, 465 completed the questionnaire
17 (90%). A total of 94 patients participated in the test-retest. The demographics of the
18 participants and the test-retest subsample are shown in Table 1. Of the respondents,
19 71.3% were male, while the length of in-patient hospital stays ranged from 0 – 30 days
20 (mean 9 days). The health problems involved surgical recovery from lower back injury
21 (7.2%), upper limb fracture (30%), lower limb fracture (39.3%), shoulder injury
22 (12.5%) and knee injury (11%). Respondents and non-respondents were not compared
23 for assessing non-response bias because of the high response rate (90%).

1 A first selection of items was made from the descriptive response distribution for each
2 item. The four items of the ‘friendly communication’ scale did not satisfy the cut-off
3 criterion of variance > 0.60, so they were deleted. No additional items were removed
4 because of poor item-scale correlation or factor loadings higher than 0.40 on more than
5 one factor. The definitive version of the questionnaire is reproduced in Appendix A.

6

7 **Instrument structure**

8 A factor analysis restricted to the 23 items of patients’ experiences confirmed a 7-
9 dimensional structure; the 7 factors explained 68.6% of the total variance, the Kaiser
10 statistic was 0.83 and the Bartlett statistic was 5018.03 ($p<0.01$). Table 2 shows the
11 factors and the items that loaded onto them significantly. Furthermore, the same 7-
12 factor structure was obtained in ‘satisfied’ and ‘unsatisfied’ patients (data not shown).
13 The 7 factors explained 70.0% of the total variance in the satisfied group (n=319) and
14 72.5% in the unsatisfied group (n=138). The Kaiser and Bartlett statistic were 0.75 and
15 3134.16 respectively in the satisfied group ($p<0.01$), and 0.79 and 1563.27 in the
16 unsatisfied group ($p<0.01$).

17 Multitrait scaling analysis generally supported the scaling of items into the hypothesized
18 scales (Table 3). Item scale correlations ranged from 0.70 to 0.93. The percentage of
19 scaling success was 100% for all the scales. Cronbach’s alpha coefficient ranged from
20 0.70 to 0.87 and exceeded 0.7 for all the scales.

21

22

23

1 **Reliability and validity**

2 Intraclass correlation coefficients ranged from 0.57 to 0.80 (median =0.68). As
3 expected, mean scores between subgroups of patients with a high evaluation of service
4 quality and those with poor quality perceptions differed significantly, except for the
5 ‘patient safety’ scale. The difference in mean scores was found to range from 17.13 to
6 43.21 points, with lower scores for the first group. As hypothesized, there was no
7 significant difference in the scale scores between men and women’s subgroups.

8 As hypothesized, all the correlations between scales and patient satisfaction were good.
9 The recommended level of 0.3 was achieved for all scales except for three dimensions:
10 ‘duration of attendance’, ‘waiting times in the sequence of treatment’ and ‘patient
11 safety’. As the three scales had high face validity, important aspects of care and good
12 reliability and known-group validity, we decided to retain them in the final instrument.

13 As expected, the obtained correlations with the treatment adherence item were lower
14 and not significant for almost all scales, which reinforces convergent validity.

15

16 **Factor structure of PEPAP-Q scales**

17 The EFA of the scales identified two factors with Eigen values of 2.484 and 1.401,
18 which explained 31.05% and 17.50% of total variance, respectively (Table 5).
19 Correlations for three scales (‘providing information and education’, ‘emotional
20 support’, ‘friendly and respectful communication’) were high on the first factor and low
21 on the second. Conversely, correlations for four other scales (‘attendance duration’,
22 ‘interruptions during care delivery’, ‘waiting times in the sequence of treatment’ and
23 ‘patient safety’) were high on the second factor and low on the first one.

1 **DISCUSSION**

2 This study presents preliminary evidence of the validity and reliability of the subscales
3 of the fixed-length PEPAP questionnaire. Seven scores pertaining to distinct dimensions
4 of patient experience in **post-acute, outpatient physical therapy settings** can be
5 computed: ‘providing information and education’, ‘sensitive manners to patients’
6 changes’, ‘emotional support’, ‘attendance duration’, ‘interruptions during care
7 delivery’, ‘waiting times in the sequence of treatment’ and ‘patient safety’. In our
8 sample, all scores had strong item-scale correlations, excellent item scaling success,
9 good internal consistency (Cronbach’s alpha coefficient over 0.7) and acceptable test-
10 retest reliability. The stability of the scales was supported by a stable factorial structure
11 among patients with high evaluations of service quality as well as those with poor
12 quality perceptions.

13 Cott et al (2006) used a qualitative approach to develop the CCRQ and they found that
14 most of the scales were moderately to strongly intercorrelated, so these scales did not
15 identify unique dimensions of patients’ experiences.²⁰ Therefore, our approach to
16 developing the PEPAP-Q, based on a combination of qualitative research and multitrait
17 scaling analysis, appears to be more promising than previous approaches used for
18 patients’ experiences questionnaires. Moreover, the PEPAP-Q can be used both to
19 monitor service performance and to inform on quality improvement efforts in **post-**
20 **acute, outpatient physical therapy settings**, and can be considered better than generic
21 instruments (e.g., GS-PEQ), which do not reflect what really matters to patients in a
22 specific context.¹⁶

23 The initial analysis undertaken in this study to explore the high-order factor structure of
24 7 dimensions of PEPAP-Q suggests that two major conceptual domains underlie these

1 dimensions. One factor or domain had high loadings for scales that reflect experiences
2 related to professionals' attitudes and behaviour. A second domain showed an affinity
3 for scales that reflect aspects of the organizational environment. The findings of other
4 authors and our previous studies support this idea. For example, Cott et al (2006) also
5 identified 7 domains that were important components of client-centered rehabilitation
6 from the client's perspective and were related to client-physiotherapist interaction and
7 organizational aspects of care such as continuity and physical comfort.²⁰ Some
8 differences from our instrument in the domains were due to the specific characteristics
9 of inpatient services²⁰ or primary health care.¹⁹

10 Tests of known-groups indicated that these scales initially discriminated on the basis of
11 patients' perceived quality, which suggests that this instrument is reasonably well suited
12 to the population of these **post-acute, outpatient physical therapy settings.**
13 **Discrimination by sex of the respondent was not observed. Previous studies³⁶ have**
14 **found that men and women have different perceptions about specific areas of care,**
15 **such as ambient condition of the facilities (temperature, cleanliness...), but these**
16 **differences were not assessed in our study because the questionnaire did not**
17 **include items about the physical environment of the setting.**

18 In the evaluation of convergent validity, the professional aspect appeared more relevant
19 than organizational aspects in influencing satisfaction and perceived quality. These
20 results are consistent with those of other authors who reported that being treated with
21 respect by health care providers and being involved in treatment decisions are strongly
22 linked to patient satisfaction⁴⁻⁷. In contrast, and as expected, in the evaluation of
23 divergent validity, low and non-significant correlations of PEPAP-Q scales with the
24 adherence to prescribed exercise were observed.

1 **The aim of this study was to develop a context-specific instrument of patients'**
2 **experiences in post-acute, outpatient physical therapy settings. We selected this**
3 **context because it is a service in high demand that remains unexplored by existing**
4 **patient experience questionnaires developed for rehabilitation services.^{19,20}**
5 **Moreover, post-acute, outpatient physical therapy settings have several**
6 **characteristics that differ from other physical therapy services; for instance,**
7 **patients are in a dependent situation, with important problems of mobility and**
8 **functionality, so the importance of feeling physically secure is highly considered by**
9 **them.**

10

11 **Study limitations**

12 Despite what we consider the generally good psychometric properties of this
13 instrument, several limitations should be noted. First, the instrument was developed in
14 Spanish, and the English translation included in this article has not been revalidated.
15 Those interested in using the questionnaire in a non-Spanish translation should conduct
16 a formal adaptation of the instrument that achieves conceptual and semantic equivalence
17 between the original version and the translated version of the questionnaire to assess the
18 usefulness of the cross-cultural adaptation.^{37,38} We recommend following the
19 "guidelines for the process of cross-cultural adaptation of self-report measures".³⁹

20 **Second, the instrument is limited to post-acute, outpatient settings and patients**
21 **with musculoskeletal conditions, so the appropriateness of the proposed**
22 **instrument to different settings is not known. Nevertheless, evidence in patient**
23 **satisfaction literature⁴⁰ suggests that there are differences in the management of**
24 **patients with acute, post-acute or chronic musculoskeletal conditions.**

1 **Third, PEPAP-Q does not address technical aspects of care. Nevertheless, we**
2 **decided to drop these questions because the physical therapists consulted thought**
3 **that patients' assessment of technical expertise could be biased by achieved**
4 **outcomes at the moment of their participation in the survey. Moreover, some**
5 **authors⁴¹ also argue that patients' assessments by questionnaires are not reliable**
6 **for assessing the technical quality of care.**

7

8 **Implications for practice and research**

9 The potential applications of this study are numerous. From a strategic standpoint, the
10 performance of each organizational unit across these dimensions can be tracked. From a
11 competitive standpoint, the identified factors can be used to compare rehabilitation
12 services with competitors' offerings. To identify and resolve problems perceived by
13 patients is very important in health services, due to the deep impact that even a small
14 percentage of them has on service outcomes.⁴²

15 The PEPAP-Q allows for analysis at several levels of **post-acute, outpatient physical**
16 **therapy settings**. For example, a health manager interested in an overview of the
17 service can use the global measures to determine an overall evaluation, while, if the
18 objective is a deeper analysis of patients' experiences, each scale score can be used to
19 identify specific aspects of the service to improve.

20 Although the current results provide good psychometric support to the questionnaire, it
21 is important to recognize the scope of this study and to consider future research
22 directions - for example, to explore possible associations between patients' experiences
23 and other outcome measurements, such as clinical results, or to explore the relative
24 importance of patients' experiences in the overall quality measurements. Further

1 validity testing of the PEPAP-Q must be conducted in order to assess its predictive
2 validity. In our study we measured patients' satisfaction levels during treatment, so the
3 predictive validity of the questionnaire remains unknown.

4 In conclusion, The PEPAP-Q was developed, pre-tested and refined as a result of an
5 extensive consultation with patients and professionals. The instrument has strong test-
6 retest reliability and the scales all have internal consistency and convergent and
7 discriminative construct validity. All these properties support the value of the PEPAP-Q
8 for use in quality improvement tasks as well as in research.

9

10 **Conflicts of interest**

11 There are no conflicts of interest.

12

13 **Source of funding**

14 This study was supported by the research grant from the Ministry of Health and
15 Consumer Affairs (PI060836) Spain.

16

17

1 **REFERENCES**

- 2 1. Cleary P, Edgman-Levitan S. Health care quality: incorporating consumer
3 perspective. *JAMA* 1997;278:1608-12.
- 4 2. Vanti C, Monticone M, Ceron D, Bonetti F, Piccarreta R, Pillastrini P, et al. Italian
5 Version of the Physical Therapy Patient Satisfaction Questionnaire: Cross-Cultural
6 Adaptation and Psychometric Properties. *Phys Ther* 2013; 93(7): 911-922.
- 7 3. Curry A, Sinclair E. Assessing the quality of physiotherapy services using Servqual.
8 *Int J Health Care Qual Assur* 2002; 15: 197–205.
- 9 4. Keith RA. Patient satisfaction and rehabilitation services. *Arch Phys Med Rehabil*
10 1998; 79: 1122–1128.
- 11 5. Goldstein NS, Elliott SD, Guccione AA. The development of an instrument to
12 measure satisfaction with physical therapy. *Physical therapy* 2000;80(9):853-863.
- 13 6. Beattie PF, Pinto MB, Nelson MK, Nelson R. Patient satisfaction with outpatient
14 physical therapy: instrument validation. *Physical therapy* 2002;82(6):557-565.
- 15 7. Beattie PF, Nelson RM, Lis A. Spanish-language version of the MedRisk
16 Instrument for Measuring Patient Satisfaction With Physical Therapy Care (MRPS):
17 preliminary validation. *Phys Ther.* 2007;87:793–800.
- 18 8. Jenkinson C, Coulter A, Bruster S et al. Patient’s experiences and satisfaction with
19 health care: Results of a questionnaire study of specific aspects of care. *Qual Saf*
20 *Health Care* 2002;11:335-339.
- 21 9. Medina-Mirapeix F, Jimeno-Serrano FJ, Escolar-Reina P, del Baño-Aledo ME. Is
22 patient satisfaction and perceived service quality with musculoskeletal rehabilitation
23 determined by patient experiences? *Clin Rehabil* 2012;27(6): 555-564.
- 24 10. Fitzpatrick R. Capturing what matters to patient when they evaluate their hospital
25 care. *Qual Saf Health Care* 2002;11:306.

- 1 11. Cleary PD, Edgman-Levitan S, McMullen W, et al. The relationship between
2 reported problems and patient summary evaluations of hospital care. *QRB Qual Rev*
3 *Bull* 1992; 18: 53–59.
- 4 12. de Silva D. Evidence scan: Measuring patient experience. London: Health
5 Foundation, 2013.
- 6 13. Coulter A, Ellins J. Patient-focused interventions: A review of the evidence.
7 London: The Health Foundation; 2006.
- 8 14. De Courcy A, West E, Barron D. The National Adult Inpatient Survey conducted in
9 the English National Health Service from 2002 to 2009: how have the data been
10 used and what do we know as a result? *BMC Health Serv Res* 2012; 12: 71.
- 11 15. Steine S, Finset A, Laerum E. A new, brief questionnaire (PEQ) developed in
12 primary health care for measuring patients' experience of interaction, emotion and
13 consultation outcome. *Fam Pract* 2001; 18(4): 410-418.
- 14 16. Sjetne IS, Bjertnaes OA, Olsen RV, Iversen HH, Bukholm G. The Generic
15 Short Patient Experiences Questionnaire (GS-PEQ): identification of core items
16 from a survey in Norway. *BMC Health Services Research*. 2011;11(1):88-98.
- 17 17. Reeves R, Seccombe I. Do patient surveys work? The influence of a national survey
18 programme on local quality-improvement initiatives. *Qual Saf Health Care*
19 2008;17:437–441.
- 20 18. Jenkinson C, Coulter A, Gyll R, et al. Measuring the experiences of health care for
21 patients with musculoskeletal disorders (MSD): development of the Picker MSD
22 questionnaire. *Scand J Caring Sci* 2002; 16: 329–233.
- 23 19. Meseguer-Henarejos AB, Medina-Mirapeix F, Escolar-Reina P, Montilla-Herrador
24 J, Hernández-Cascales N, Hidalgo-García MC.. Calidad percibida en la fisioterapia

- 1 de atención primaria: situación y oportunidades de mejora en la atención a la
2 cervicalgia. *Rev Iberoam Fisioter Kinesiol* 2008; 11(2): 68–80.
- 3 20. Cott CA, Teare G, McGilton KS, Lineker S. Reliability and construct validity of the
4 Client-Centred Rehabilitation Questionnaire. *Dis Rehabil* 2006; 28(22):1387-1397.
- 5 21. Asmundson G, Jones S. The Wascana client-centred care survey: development and
6 psychometric evaluation. *Can J Quality Health Care* 1993;13:19 – 21.
- 7 22. McDonald RP. The Dimensionality of Tests and Items. *British Journal of*
8 *Mathematical and Statistical Psychology*. 1981;34:100-117.
- 9 23. Medina-Mirapeix F, Navarro-Pujalte E, Oliveira-Sousa SL, Escolar-Reina P,
10 Montilla-Herrador J, Valera-Garrido JF, Collins SM. Mobility activities
11 measurement for outpatient rehabilitation settings. *Arch Phys Med Rehabil*
12 2011;92:632-639.
- 13 24. Haley SM, Coster WJ, Andres PL, et al. Activity outcome measurement for
14 postacute care. *Med Care* 2004;42(Suppl 1):49-61.
- 15 25. Silverstein BS, Kilgore KM, Fisher WP, Harley JP, Harvey RF. Applying
16 psychometric criteria to functional assessment in medical rehabilitation: I. Exploring
17 unidimensionality. *Arch Phys Med Rehabil* 1991;72: 631–637.
- 18 26. Grill E, Müller M, Quittan M, Strobl R, Kostanjsek N, Stucki G. Brief ICF Core Set
19 for patients in geriatric post-acute rehabilitation facilities. *J Rehabil Med*
20 2011;43(2):139-144.
- 21 27. Scalise D. The patient experience. *Hosp Health Netw* 2003; 77(12): 41-48.
- 22 28. Kings College and the King's Fund. What matters to patients? Developing the
23 evidence base for measuring and improving patient experience. London: Kings
24 College and the King's Fund, 2011.

- 1 29. Medina-Mirapeix F, del Baño-Aledo ME, Oliveira-Sousa SL, Escolar-Reina P,
2 Collins SM. How the rehabilitation environment influences patient perception of
3 service quality: a qualitative study. *Arch Phys Med Rehabil* 2013;94(6):1112-7.
- 4 30. del Baño-Aledo ME, Medina-Mirapeix F, Escolar-Reina P, Montilla-Herrador J,
5 Collins SM. Relevant patient perceptions and experiences for evaluating quality of
6 interaction with physiotherapists during outpatient rehabilitation: a qualitative study.
7 *Physiotherapy* 2013. <http://dx.doi.org/10.1016/j.physio.2013.05.001>
- 8 31. Medina-Mirapeix F, Jimeno-Serrano FJ, del Baño-Aledo ME, Escolar-Reina P,
9 Montilla-Herrador J, Lomas-Vega R, Franco-Sierra MA. Outpatients' perceptions of
10 their experiences in musculoskeletal rehabilitation care. *Eur J Phys Rehabil Med*
11 2012; 48:1-8.
- 12 32. Sluijs EM, Kok GJ, van der Zee J. Correlates of exercise compliance and physical
13 therapy. *Phys Ther* 1993, 73:771-86.
- 14 33. Loblaw A, Bezjak A, Bunston T. Development and testing of a visit specific patient
15 satisfaction questionnaire: The Princess Margaret Hospital satisfaction with doctor
16 questionnaire. *J Clin Oncol* 1999;17(6):1931-1938.
- 17 34. Ware JE, Gandek B. Methods for testing data quality, scaling assumptions and
18 reliability: the IQOLA Project approach. *J Clin Epidemiol* 1998;51:945-952.
- 19 35. Green J, Fukuhara S, Shinzato T et al. Translation, cultural adaptation, and initial
20 reliability and multitrait testing of the Kidney Disease Quality of Life instrument for
21 use in Japan. *Quality of Life Research*. 2001;10:93-100.
- 22 36. Lam ETC, Zhang JJ, Jensen BE. Service Quality Assessment Scale (SQAS): An
23 instrument for evaluating service quality on health-fitness clubs. *Measurement in*
24 *physical education and exercise science*. 2005;9(2):79-111.

- 1 37. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related
2 quality of life measures: literature review and proposed guidelines. *J Clin*
3 *Epidemiol.* 1993;46:1417–1432.
- 4 38. Streiner DL, Norman GR. *Health Measurement Scales: A practical guide to their*
5 *development and use.* (4th edition) Oxford: Oxford University Press; 2008.
- 6 39. Beaton BD, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of
7 cross-cultural adaptation of self-report measures. *Spine.* 2000;25:3186–3191.
- 8 40. Hills R, Kitchen S. Satisfaction with outpatient physiotherapy: Satisfaction with
9 outpatient physiotherapy: A survey comparing the views of patients with acute and
10 chronic musculoskeletal conditions. *Physiother Theory Pract* 2007;23(1):21–36.
- 11 41. Grotle M, Garrat A, Løchting I, Kjekken I, Klokke M, Uhlig T, et al.
12 Development of the rehabilitation patient experiences questionnaire: data quality,
13 reliability and validity in patients with rheumatic diseases. *J Rehabil Med*
14 2009;41:576–81.
- 15 42. Coulter A. Can patients assess the quality of health care? Editorial. *BMJ.*
16 2006;333:1-2.
- 17
- 18

1 **Table 1.** Characteristics of patient sample (n=465) and test-retest subsample (n=94)

2

| VARIABLES | SAMPLE | SUBSAMPLE |
|---------------------------|---------------|------------------|
| | % | % |
| SEX | | |
| Men | 71.3 | 64.6 |
| Women | 28.7 | 35.4 |
| AGE | | |
| <30 years | 30.8 | 22.0 |
| 30-45 years | 40.5 | 41.5 |
| >45 years | 28.7 | 36.6 |
| EDUCATION LEVEL | | |
| No studies | 46.4 | 48.8 |
| Non-high school graduate | 32.8 | 34.1 |
| High school graduate | 20.8 | 17.1 |
| BODY PART INVOLVED | | |
| Lower extremity | 50.3 | 60.1 |
| Trunk | 7.2 | 4.3 |
| Upper extremity | 42.5 | 35.6 |

3

4

5

1 **Table 2.** Factor analysis of 23 items of patients' experiences with physical therapy (n= 465)
 2

| SCALE / items | FACTORS | | | | | | |
|---|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| EMOTIONAL SUPPORT | | | | | | | |
| My therapist is aware of my worries. | 0.783 | 0.286 | 0.018 | 0.090 | -0.048 | 0.118 | 0.076 |
| I get encouragement to address worries | 0.752 | 0.256 | 0.080 | 0.027 | 0.055 | 0.156 | 0.139 |
| My therapist is interested in my recovery | 0.716 | 0.254 | 0.070 | -0.044 | 0.098 | 0.148 | 0.078 |
| My therapist supervises the exercises | 0.705 | 0.101 | 0.056 | 0.096 | 0.257 | 0.101 | -0.075 |
| I have opportunities to talk about worries/doubts | 0.693 | 0.239 | 0.010 | 0.131 | 0.142 | 0.166 | -0.138 |
| PROVIDING INFORMATION AND EDUCATION | | | | | | | |
| I am told how to prevent complications | 0.358 | 0.789 | 0.015 | -0.002 | 0.025 | 0.021 | -0.012 |
| I get information about usefulness of the therapies | 0.243 | 0.782 | 0.009 | 0.017 | 0.090 | 0.065 | -0.051 |
| I get information about the prognosis. | 0.245 | 0.767 | 0.059 | -0.007 | 0.139 | 0.172 | -0.008 |
| My therapist answers my questions | 0.163 | 0.714 | 0.004 | 0.174 | 0.003 | 0.138 | 0.111 |
| DURATION OF ATTENDANCE | | | | | | | |
| My therapist does not monitor the treatment | 0.081 | 0.016 | 0.916 | 0.120 | 0.063 | 0.092 | 0.060 |
| My therapist cannot prevent risk situations | 0.016 | 0.054 | 0.885 | 0.103 | 0.049 | 0.052 | 0.058 |
| My therapist does not spend enough time with me | 0.080 | 0.033 | 0.778 | 0.232 | 0.167 | 0.115 | -0.034 |
| INTERRUPTIONS DURING DELIVERY OF CARE | | | | | | | |
| My therapist has to attend to other professionals | 0.055 | 0.022 | 0.123 | 0.869 | 0.164 | 0.049 | 0.021 |
| My therapist has to help other patients | 0.111 | 0.028 | 0.153 | 0.857 | 0.164 | 0.077 | 0.076 |
| My therapist has administrative demands | 0.050 | 0.114 | 0.174 | 0.759 | 0.154 | 0.069 | 0.133 |
| WAITING TIMES IN THE SEQUENCE OF TREATMENT | | | | | | | |
| The treatment tables are occupied | 0.103 | 0.016 | 0.088 | 0.128 | 0.826 | 0.046 | 0.101 |
| Other equipment is occupied | 0.111 | 0.121 | 0.011 | 0.204 | 0.825 | 0.132 | 0.071 |
| Boxes or devices are occupied | 0.156 | 0.085 | 0.175 | 0.153 | 0.765 | 0.086 | 0.033 |
| SENSITIVE MANNERS TO PATIENTS' CHANGE | | | | | | | |
| My therapist adapts treatment to pain | 0.143 | 0.131 | 0.124 | 0.089 | 0.146 | 0.838 | 0.025 |
| My therapist adapts treatment to functional status | 0.138 | 0.217 | 0.147 | 0.063 | 0.086 | 0.773 | 0.093 |
| My therapist adapts treatment to changes of mood | 0.320 | 0.034 | 0.004 | 0.055 | 0.043 | 0.738 | -0.090 |
| PATIENT SAFETY | | | | | | | |
| I experience dangerous situations | -0.036 | -0.004 | 0.004 | 0.051 | 0.050 | 0.054 | 0.873 |
| I have to ask other patients for help | 0.087 | 0.039 | 0.076 | 0.148 | 0.130 | -0.034 | 0.841 |
| Critical value: | 7.142 | 3.333 | 2.309 | 1.892 | 1.748 | 1.477 | 1.386 |
| % explanation of variance: | 23.81 | 11.11 | 7.70 | 6.31 | 5.83 | 4.92 | 4.62 |
| Kaiser-Meyer-Olkin: 0.828 Bartlett's test: 5018.027 p<0.01 | | | | | | | |

3
4

1 **Table 3.** Summary of results for multitrait scaling analyses and test-retest reliability
 2

| SCALES | Item-scale correlation range (median) | % scaling success | Cronbach Alpha | ICC (95% CI) |
|--|--|--------------------------|-----------------------|---------------------|
| Emotional Support | 0.76-0.83 (0.79) | 100% | 0.84 | 0.80 (0.72-0.87) |
| Sensitive manners to patients changes | 0.79-0.86 (0.84) | 100% | 0.76 | 0.68 (0.55-0.78) |
| Providing Information and Education | 0.70-0.88 (0.85) | 100% | 0.84 | 0.79 (0.69-0.86) |
| Duration of attendance | 0.83-0.93 (0.90) | 100% | 0.87 | 0.62 (0.47-0.74) |
| Interruptions during delivery of care | 0.81-0.90 (0.89) | 100% | 0.84 | 0.75 (0.64-0.83) |
| Waiting times in the sequence of treatment | 0.77-0.85 (0.80) | 100% | 0.81 | 0.59 (0.43-0.71) |
| Patient Safety | 0.84-0.92 | 100% | 0.70 | 0.57 (0.41-0.70) |

3 ICC: Intraclass Correlation Coefficient in test-retest reliability
 4
 5
 6
 7

1 **Table 4.** Construct validity of the scales

2

| SCALES | Mean of problem score differences between groups (<i>p value</i>) | | Correlation with overall quality | Correlation with Satisfaction | Correlation with Adherence |
|---|--|----------------|--|-------------------------------------|----------------------------------|
| | High-Low quality perception | Men-Women | | | |
| Emotional Support | -40.93 (0.001) | -2.31 (0.795) | -0.54* | -0.55* | -0.27* |
| Sensitive manners to patients' changes | -39.20 (0.000) | -4.02 (0.648) | -0.49* | -0.45* | -0.16 |
| Providing Information and Education | -38.95 (0.000) | -10.26 (0.196) | -0.38* | -0.45* | -0.19* |
| Duration of attendance | -21.53 (0.011) | -6.91 (0.378) | -0.20* | -0.21* | -0.05 |
| Interruptions during delivery of care | -43.21 (0.000) | 6.69 (0.420) | -0.50* | -0.48* | -0.21* |
| Waiting times in the sequence of treatment | -30.46 (0.005) | -8.46 (0.326) | -0.24* | -0.26* | -0.13 |
| Patient Safety | -17.13 (0.115) | 5.62 (0.328) | -0.22* | -0.22* | -0.04 |

3

* Significant correlation at 0.05

4

1 **Table 5.** Factor analysis of scales (n= 465)

2

| SCALES | FACTORS | |
|--|----------------------------|--------------|
| | 1 | 2 |
| Emotional Support | 0.808 | 0.197 |
| Providing Information and Education | 0.743 | 0.093 |
| Sensitive manners to patients' changes | 0.649 | 0.236 |
| Interruptions during delivery of care | 0.161 | 0.719 |
| Duration of attendance | 0.037 | 0.647 |
| Waiting times in the sequence of treatment | 0.302 | 0.629 |
| Patient Safety | -0.101 | 0.565 |
| | Critical value: | 2.484 |
| | % explanation of variance: | 31.05 |
| Kaiser-Meyer-Olkin: 0.710 Bartlett's test: 547.368 p<0.01 | | |

3

4

1 **Appendix A. Patient Experiences with Post-Acute **outpatients** physical therapy Questionnaire:**
2 **Items scoring and subscales.**

- 3 1. I get information about the prognosis.
- 4 2. I get information about the usefulness of the therapies.
- 5 3. I am told how to prevent complications during treatment.
- 6 4. My therapist gives complete answers to my questions.
- 7 5. I get encouragement to address worries.
- 8 6. My therapist shows personal interest in my recovery.
- 9 7. My therapist is aware of my worries.
- 10 8. My therapist supervises my exercise practice.
- 11 9. I have opportunities to talk about worries/doubts.
- 12 10. My therapist adapts treatment to pain.
- 13 11. My therapist adapts treatment to changes of mood.
- 14 12. My therapist adapts treatment to changes in functional status.
- 15 13. The duration of attention by the physiotherapist is shorter than I expected.
- 16 14. Therapist seems to have a very limited time for observing my exercise practice.
- 17 15. My therapist doesn't stay with me to prevent risks during moving or therapies.
- 18 16. My treatment is interrupted because my therapist has to help other patients.
- 19 17. My treatment is interrupted because my therapist has to attend to other professionals.
- 20 18. My treatment is interrupted because my therapist has phone calls/administrative demands.
- 21 19. I have to wait during treatment because boxes or devices are occupied.
- 22 20. I have to wait during treatment because tables are occupied.
- 23 21. I have to wait during treatment because other equipment is occupied.
- 24 22. I have to ask other patients for help to prevent risks (e.g. falls).
- 25 23. I experience dangerous situations due to the lack of help from professionals.

26 Items: Item statements are presented in the order in which they appear in the questionnaire. However, the style of the
27 questionnaire is not reproduced here.

28 Scoring: 5-point Likert scale: 5 = *Always*; 4 = *Very often*; 3 = *Sometimes*; 2 = *Rarely*; 1 = *Never*

29 Reversed scoring on items: 13,14,15,16,17,18,19,20,21,22,23 (5 = 1, 4 = 2, 2 = 4, 1 = 5)

30 Items in each of the 7 Scales: Information: 1,2,3,4; Sensitive: 5,6,7,8,9; Support: 10,11,12; Attendance duration:
31 13,14,15; Interruptions: 16,17,18; Waiting times: 19,20,21; Safety: 22,23.

32
33