

## RELIABILITY AND VALIDITY OF PATIENT REPORTS FOR PHYSICAL THERAPY QUALITY ASSESSMENT: AN EMPIRICAL ANALYSIS REGARDING THE USE OF EXERCISES FOR NECK PAIN IN SPAIN

Francesc Medina-Mirapeix, PT, PhD<sup>1</sup>, Ana B. Meseguer-Henarejos, PT<sup>1</sup>,  
Joaquina Montilla-Herrador, PT<sup>1</sup>, Pilar Escolar-Reina, PT<sup>2</sup>, Sean M. Collins, PT, ScD, CCS<sup>3</sup>  
and Pedro J. Saturno-Hernández, MD, PhD<sup>2</sup>

From the Departments of <sup>1</sup>Physical Therapy School of Medicine, <sup>2</sup>Public Health and Preventive Medicine, School of Medicine, University of Murcia, Spain and <sup>3</sup>Physical Therapy, University of Massachusetts, Lowell, MA, USA

**Objective:** To determine usefulness and reliability of patient reports as a data source on the use of exercises for neck pain in primary healthcare.

**Design:** Survey research.

**Subjects:** A total of 176 patients with mechanical neck pain, surveyed about their physical therapy experience.

**Method:** A questionnaire was constructed on the selection of exercises, instructions for carrying them out and follow-up provided by the physical therapist. The same items reported by the surveyed patients were also reviewed in the physical therapy record. We analysed the reliability and validity of the questionnaire, and used the valid information from the survey to estimate a total of 8 indicators on the quality of care provided.

**Results:** Sensitivity of the items was high (>75%), suggesting that most of the recorded information could be provided by patients, and specificity was quite low, suggesting that they also provide relevant information that was not recorded. These results, particularly low specificity, were not homogeneous among items. Estimates for the indicators demonstrate room for improvement.

**Conclusion:** Patient reports could be useful as a complement to other sources of information for physical therapy quality assessment, and they can be reliable and valid substitutes for recorded data about specific aspects of the care provided.

**Key words:** neck pain, exercise, patient report, monitoring.

J Rehabil Med 2006; 38: 354–359

*Correspondence address:* Francesc Medina-Mirapeix, Departamento de Fisioterapia, Facultad de Medicina, Universidad de Murcia, Campus de Espinardo, ES-30100, Murcia, Spain. E-mail: mirapeix@um.es

Submitted March 4, 2005; accepted April 10, 2006

### INTRODUCTION

Healthcare quality assessment and improvement is increasingly considered an essential part of everyday activities in health centres. For this purpose, 2 complementary data sources are used primarily: medical records and information provided by patients (1, 2). For the latter, satisfaction surveys have been the

main strategy, aiming to gather data on patient opinions regarding the quality of healthcare received. More recently, the possibility has been raised of obtaining objective and reliable data about quality of healthcare from patient reports based on their experience (3).

The theoretical approach to patient reports has been discussed extensively (4). Its usefulness has been confirmed by empirical evidence (5) in specific aspects of medical care, but has not been demonstrated in physical therapy care.

Recommendations about exercise are one of the basic therapeutic resources in the treatment of neck pain (6–8). Large variability in the use of exercise was observed in a previous paper (9) on guidelines for neck pain treatment in primary healthcare. In general, the use of exercise for other conditions also demonstrates large variability in the amount, indications for and type of exercise (10). This variability is one of the arguments to improve the rational use of exercises. The purpose of this study is to determine the usefulness and reliability of patient reports as a data source to monitor and improve the quality of physical therapy care regarding the use of exercise for patients with mechanical neck pain in primary healthcare.

### METHODS

#### *Questionnaire design*

As a pilot experience on the use of patient reports for quality assessment in physical therapy, a questionnaire was constructed regarding the use of exercises for the treatment of neck pain. Issues related to the selection of exercises, instructions for carrying out the exercises and follow-up were identified by reviewing the scientific literature (11–16). Issues concerning the use of exercise that were important to patients were identified by implementing 3 focus groups with the participation of a total of 16 patients receiving treatment for mechanical neck pain. The final list of items to be included in the questionnaire was generated using patient statements and expressions obtained from analysis of the 3 focus groups.

The final questionnaire comprises 4 sections with a total of 15 questions. The first 3 sections refer to aspects and objectives of the selection of the type of exercise, instructions for carrying out exercise and follow-up provided by the physical therapist (Table I). The fourth section includes demographic data, such as the patient age, gender, education and previous use of physical therapy that may have an effect on the information patients provide (17, 18). After the fourth section there are 2 open questions: one about difficulties in understanding the

Table I. Sections related to characteristics of the questionnaire

<i>Section I: Selection of exercises</i>	
1. Did your physical therapist (PT) teach you any type of exercise while in the primary care centre? (if your answer is no, skip to question 12)	Yes/No
2. Did your PT teach you stretching exercises?	Yes/No
3. Did your PT teach you strengthening exercises?*	Yes/No
4. Did your PT teach you exercises based in co-ordinated movements between your eyes and neck (for oculocervicokinetic re-education)?	Yes/No
<i>Section II: Instructions to practice the exercises</i>	
5. Did your PT give you any information on carrying out the exercises at home?	Yes/No
6. How many times did your PT tell you to do each exercise?*	
7. What did your PT tell you about possible adverse reactions when carrying out the exercises?*	
– You must not do the painful movements	
– If all the movements are painful you must stop doing them a few days.	
– Others:	
– None.	
8. Did your PT give you written instructions on how to carry out the exercises at home?	Yes/No
9. Did your PT tell you how you had to continue the exercise programme after discharge?	Yes/No
<i>Section III: Follow-up of exercises</i>	
10. Did your PT ask you if you had any difficulty or problem carrying out exercises at home?	Yes/No
11. Did your PT supervise and correct you the carrying out of exercises at the health centre?	Yes/No
<i>Section IV: Sociodemographics aspects</i>	
12. Age: _____	
13. Education level: _____	
14. Gender: Female/Male	
15. Previous use of physical therapy:	Yes/No
Others questions on comprehension level and relevant aspects of the healthcare received that were not mentioned in the questionnaire.	

Some questions were excluded on the final questionnaire because they \*were hard to understand or \*\*had low sensitivity.

questionnaire, and one on relevant aspects of the healthcare received that were not mentioned in the questionnaire and on which the patient would like to comment. Overall, there were 3 response formats: (i) yes/no, (ii) a list of various related options with the possibility of marking more than one option, and (iii) open.

#### Questionnaire distribution

The questionnaire was piloted at 5 physical therapy units in different primary healthcare centres (HC) of the National Health Service in Spain that volunteered to participate in the study. We anticipated a 70% response rate; therefore, to ensure a minimum of 25 questionnaires answered per HC, we sampled 35 neck pain patients per centre.

Over a 6 month period, at the end of the intervention, physical therapists in each unit gave the questionnaire to each patient receiving treatment for mechanical neck pain, between 18 and 79 years old, asking them to complete it as soon as possible upon returning home. Exclusion criteria were: post-traumatic neck pain (i.e. whiplash-associated disorders), people unable to read or write, and impairments of memory and/or cognitive behaviour associated with conditions such as Alzheimer's disease or senile dementia. Patients were asked to place the questionnaire in a box in the administrative area of the HC on the day the patient had to return for the medical appointment normally carried out after physical therapy treatment or for other reasons. Patients were assured that the physical therapist would not be able to identify individual answers. The questionnaires could also be returned by post directly to our University Department in a stamped addressed envelope.

A numeric code was printed in each questionnaire. When questionnaires were distributed, the physical therapist recorded the code and the patient's medical and physical therapy record (PR) number in order to link questionnaire and PR data and allow consistency studies, and enable reminders to be sent to the non-respondents.

The response to the questionnaire was encouraged by a reminder telephone call to those patients who did not return the questionnaire within the first 2 weeks. A second distribution was made by post for those individuals who had lost the first questionnaire.

#### Physical therapy record data collection

After the time period for returning the questionnaire, the PR associated with each respondent was reviewed. Two members of the research team (unaware of questionnaire responses) performed this review. Information from the PR included data on the same aspects in the questionnaire related to the type of exercises, instructions for carrying out the exercises and follow-up of the exercises, as well as on those variables for consideration of a possible non-answer bias: age, gender, previous visits and education level.

#### Reliability and validity analysis

A questionnaire is reliable when differences in answers obtained correspond to true differences in the subject of investigation and not to differences in the interpretation of questions (19), which may occur if questions are confusing (20). For that reason, as a first step, reliability was indirectly studied by analysing how comprehensible the items were. Items with a higher number of non-specific answers or left unanswered were not well understood by the patients, possibly because they were not well formulated, resulting in a problem with reliability. We determined the non-answer (left blank) and non-specific answer frequencies for each question. We defined specific answers according to the format of the response: (i) in the closed questions, answers either of "yes" or of "no"; (ii) in open questions, that the response was relevant to the question; and (iii) in questions with various options, either by marking any of the options or by indicating "no". Additionally, we analysed the answers to the open question on problems with the understanding of any question.

In a second step, we studied the agreement (also called equivalence or reproducibility) (21) of the data-gathering methods (patient's report and PR). Reproducibility was analysed by calculating the percentage of agreement between information provided in the questionnaire and that available in the PR, for those questions that were understood by the patients (non-answer and non-specific answer rate <15%, and not pointed out as difficult to understand) and only for those items whose information was recorded in at least 20% of the PR consulted.

In order to determine the validity of the questionnaire as a substitute for recorded data, we calculated the sensitivity and specificity levels of those items that were reliable (non-answer and non-specific answer rate <15%) and whose information was recorded in at least 20% of the PR. PR data were used as a reference pattern (gold standard) (2, 3) and we considered "true-positive" cases in which both PR and patient report coincided when indicating that a particular aspect of the use of exercises was performed; whereas "true-negative" indicated cases in which both sources indicated the absence of the exercise aspect in question. When items were not recorded in at least 20% of the PR, the patient was considered the "gold standard" (4).

Furthermore, we analysed factors related to the sensitivity and specificity of each item using logistic regression models (for each question analysed) using the "enter" method (22). The dependent variables were the sensitivity and specificity (no = 0; yes = 1). Independent variables included: gender (male = 0; female = 1); age group (18–40 years = 0; 41–60 years = 1; over 60 years = 2); education (no education = 0; primary education = 1; secondary or university education = 2); and previous use of physical therapy (no = 0; yes = 1).

Indicators were developed on quality regarding the use of exercises from those answers that were valid (sensitivity >75%) and reliable (non-answer and non-specific answer rate <15%). We also built indicators from those answers that were not systematically recorded in the PR.

## RESULTS

### Response rates

The total number of respondents was 142 (80.7% response rate). Of all respondents, reminder telephone calls were necessary in 7.8%. The characteristics of the respondents are presented in Table II. This group presented no significant differences with respect to age or gender compared with those who did not answer.

### Reliability and validity analysis

The non-answer and non-specific answer rate were higher than 15% in 2 items of the questionnaire (number 6 and 7) and therefore they were not considered in the criterion validity analysis. Both items were within the "Instructions to practice the exercises" section: Did your physiotherapist give you any

Table II. Characteristics of the study population and differences between those who responded and those who did not respond. No significant differences were found

Variables	% Respondent (n = 142)	% Non-respondent (n = 34)
Age (years)		
18–40	24.1	17.6
41–60	50.4	67.6
61–79	25.5	14.7
Sex		
Men	17.0	14.7
Women	83.0	85.3
Educational level		
No education	15.1	
Primary education	61.1	
Secondary or University education	23.8	
Previous use of physical therapy		
No previous use	60.9	
Prior use	39.1	

information on the number of repetitions? (18.3%) and what did your physiotherapist tell you about possible adverse reactions when carrying out the exercises? (19%). In general, the questions in the survey were easy to understand. According to the answers to the open question on difficulties in the understanding of any item of the questionnaire, only 0.7% of patients had some minor problems in understanding some questions.

In 20 of the subjects who answered the questionnaire (14%), it was not possible to find his or her PR. For this reason the sample size of the comparative analysis between the information provided by the physical therapy record and that of the questionnaire was 122 people.

Some items (numbers 9, 10 and 11) were not registered systematically in the PR. Information about 6 items (out of 11) from the questionnaire was found in more of the 20% of the PR. Patient survey sensitivity as a data source of the use of exercises in physical therapy within the HC has been high (>75%) in 5 of the 6 analysed items. Only the item to assess strengthening exercises (number 3) showed low (50%) sensitivity (Table III). On the contrary, the specificity of the items has been quite low, especially when the patients were asked about carrying out the exercises at home (number 5) and about the teaching of exercises while being in the HC (number 1). In 2 items, selection of exercises for oculocervicokinetic re-education (number 4) and strengthening (number 3), the specificity has been higher than 72% (Table III). With the exceptions of the information provided for those 2 items, the survey questions are more sensitive than specific. Patients provide additional information to that already present in the PR. Data becomes quite meaningful regarding this last issue: particularly relevant is that in 25.3% of cases, stretching exercises that the patient refers to are not included in the PR. According to all logistic regression models analysed, the patient report sensitivity and the specificity were not predicted by any of the variables (gender, age, education, previous use of physical therapy) considered.

In the final version of the questionnaire we excluded 2 questions that were hard to understand (numbers 6 and 7) and one (number 3) because of its low sensitivity. In general terms, the information provided by patients has been reliable and valid enough to define 8 indicators (Table IV) to monitor aspects of 3 areas related to the use of exercise: (i) the selection of exercise, (ii) the instructions for carrying them out, and (iii) the follow-up for neck pain treatment in primary healthcare. In our environment, these indicators can provide opportunities to improve the quality in the use of exercises for neck pain treatment. This room for improvement was particularly important (Table IV) in the selection of exercises for oculocervicokinetic re-education (only 37.2%), in the difficulties or problems carrying out exercises at home (only 60%), and in the provision of written instructions (only 72.5% received).

Reproducibility, the consistency between sources, of the questionnaire and the PR was moderate (72–82%), and lower in some items (<60%). Consistency is low for written instructions (Table III). In the majority of cases, the lack of a high agreement (>90%) occurred due to a lack of information in the

Table III. Sensitivity and specificity levels of patient report and agreement with recorded data

Items	Sensitivity		Specificity		Agreement	
	n	% (95% CI)	n	% (95% CI)	n	%
Selection of any type of exercise	91	94 (81–97)	27	15 (2–28)	118	76
Selection of stretching exercises	74	92 (88–98)	29	20 (5–35)	103	72
Selection of strengthening exercises	62	50 (37–63)	41	73 (59–87)	103	59
Selection of exercises for oculocervicokinetic re-education	32	75 (60–90)	71	86 (78–94)	103	82
Indication for carrying out exercises at home	89	98 (95–100)	23	4 (0–12)	112	79
Provision of written instructions	58	79 (69–89)	44	33 (20–46)	102	57

Agreement: percentage of coincidences between the questionnaire and the physical therapy record.

PR compared with that provided by the patient. This fact resulted in the low specificity figures for most of the items, as described above.

## DISCUSSION

Patient survey response ratio was higher than that of other studies that have used the postal system as a questionnaire distribution method (6). We believe that a postal or telephone survey may be a more realistic procedure when considering implementing the survey on a regular basis. However, a telephone survey is more costly, and a postal survey may have a lower response rate. If the healthcare personnel are committed to quality, as were the ones who voluntarily participated in our study, the method we have used for distributing the questionnaires may be both feasible and advisable.

### Validity of patient reports

Content validity was ensured by the process of questionnaire development. Both literature and qualitative analysis (focus groups) were useful for ensuring that items reflecting quality in the use of exercise for neck pain treatment were included, in particular aspects related to selection, instructions for carrying out exercises and follow-up.

Regarding criterion validity, as we found some items recorded in PR while others were not, we used different “gold standard” for each of them. On the one hand, PR data was used as a “gold standard”. Professional record or direct observations are sources for assessing quality of technical aspects in medical and physical therapy care. Nevertheless, PR is traditionally used due to its low cost in contrast with observation. Drawbacks have been identified in the PR method in relation to its limited role with interpersonal aspects, its availability and legibility of records. The PR or medical record are the sources more often used for assessing quality of care. As one of our objectives was to determine the validity of the questionnaire as a substitute for recorded data for assessing quality in HC, PR data itself was used as “gold standard”. Using PR data as the gold standard, we found that information provided by patients in 5 of the items analysed can be a valid and useful substitute to PR for assessing service quality in HC. The high sensitivity observed in the majority of items analysed within patients’ reports means that if the data were included in the PR, the patient would mention them in the survey. The fact that the specificity levels were generally lower could be due to a possible under-registration of many of the activities in PR within the physical therapy clinics, a feature identified as a problem by other authors (5).

Items not registered in more than 20% of PR provide information about interpersonal aspects of care between

Table IV. Indicators built from patient report and their frequency

Indicators	Frequency	
	Absolute	% (95% CI)
<i>I. Selection of exercises</i>		
% patients that provide information on . . .		
1. . . any type of exercise (n=138)	128	92.8 (88.5–97.1)
2. . . stretching exercises (n=121)	107	88.4 (82.7–94.1)
3. . . exercises for oculocervicokinetic re-education (n=121)	45	37.2 (28.6–45.8)
<i>II. Indications on for carrying out exercises</i>		
% patients that provide information on . . .		
4. . . carrying out exercises at home (n=130)	128	98.5 (96.4–100)
5. . . written instructions on how to carry out exercises at home (n=131)	95	72.5 (64.9–80.1)
6. . . duration of the exercise programme after discharge (n=131)	113	86.3 (80.4–92.2)
<i>III. Follow-up of exercises</i>		
% patients that provide information on . . .		
7. . . potential difficulties or problems in carrying out exercises at home (n=125)	75	60 (51.4–68.6)
8. . . supervision the carrying out of exercises (n=122)	112	91.8 (87–96.6)

physical therapists and patients (information about potential difficulties or problems in carrying out exercises at home, supervision and duration of the exercise program after the discharge). In those items, the criterion validity could not be analysed. For this type of data, the patient rather than the record may be considered the “gold-standard”. Records rarely capture information about the interpersonal aspects of care (4) and an alternative source of data is needed. We believe that the patient is best qualified to inform about interpersonal aspects of care, and that it would be very difficult to question the validity of the patients’ perspective on interpersonal features (4). In summary, this study, as well as previous research in fields other than physical therapy, suggests that patients can provide relevant information to assess quality that is not usually available from traditional recorded data sources.

#### *Reliability of patient reports*

In accordance with Nunnally & Bernstein (19) and Mangione (20), we accepted that a questionnaire is reliable when the differences in answers obtained correspond to true differences in the subject of the investigation and not to differences in the interpretation of the questions, which may occur if the questions are confusing. To assess this, we considered that questions pointed out as difficult to understand by respondents, or with a higher number of non-specific answers, and/or that were not answered, were not well understood by the patients and therefore these questions may not be reliable. The observed non-specific answer rate suggests that the items were well understood, with the exceptions of 2 items. However, these 2 items could be affected because they were open and included multiple response options. Previous studies on self-administered questionnaires have reported that the probability of a non-response in open questions is 2–3 times higher than in closed questions (23). On the other hand, the relative ease with which the questionnaire was understood may be due to the fact that a significant proportion of individuals who responded had a higher educational level than those who did not respond. This possible bias has already been pointed out as a common observable fact in postal surveys (24), but it could not be assessed in our study, due to the higher response rate that left only a few non-respondent cases to compare and to the relative under-registration of the educational level in the PR.

Other approaches to test question reliability could have been considered such as test-retest consistency and appraisal of the internal consistency of the questionnaire. Test-retest reliability analysis measures the stability or consistency of the answers people give to the same question when they are asked it at different points in time. This analysis assumes no real changes have occurred that would cause them to answer differently (19, 20). While the consistency of answers to the same questions over time may vary as a function of transient personal or situational factors (21), there can be no real change between these periods since the interventions that patients are reporting on have

already occurred. Since the questions (objective, occurrence or not) may be affected by recall bias in a test-retest analysis, the eventual results of such an analysis would not be conclusive due to the inherent problem of the method, therefore identifying a large enough sub-sample of patients that did not return to the HC may be cumbersome and not cost-effective. Internal consistency reliability analysis was not performed because questions were designed as independent, items not intended to measure the same concept or theoretical construct, but addressing different aspects of the process of care, and therefore are not necessarily correlated.

The information provided by patients is generally reliable, but the reproducibility of measures of the same concepts between patient report and PR data is only partial. The discrepancy between the information recorded in the PR and the patients may reflect problems of under-registration of the professional but also of communication between the professional and the patient (25, 26).

#### *Usefulness of the patient’s report for quality assessment*

Low specificity and agreement levels appear to be related to patient reports about activities that do not appear in the PR. This indicates that patients may be a better source of information than the PR for healthcare quality assessment, particularly in those areas related to information on the indications for carrying out exercise and follow-up. In addition, patients may be a more homogeneous and comparable data source among HC than PR. Recorded data may have problems that vary from one HC to another (illegibility, heterogeneity or even entirely missing PR) (27, 28), associated with number of relevant reliability, and therefore validity issues.

This study determined that patient reports can be a useful substitute for the PR in assessing service quality with several items: “selection of any type of exercise”, “selection of stretching exercises”, “selection of exercises for oculocervico-kinetic re-education”, “indications for carrying out exercises at home” and “provision of written instructions”. Furthermore, patient reports can provide additional information about items that are not usually included in the PR: “duration of the exercise programme after discharge”, “difficulties or problems in carrying out exercises at home” and “supervision the carrying out of exercises”.

In conclusion, this study demonstrates that patient reports are a useful and reliable data source to monitor and improve the quality of physical therapy care regarding the use of exercise for neck pain patients. The indicators constructed on the basis of the information provided by patients allow us to draw conclusions regarding the quality level of the assessed process of care. Therefore, the results of this study regarding indicators of compliance confirm that in our environment the use of exercise in patients with mechanical neck pain is a quality problem requiring intervention to improve. These interventions should be tailored to the specific aspects (e.g. information on potential

problems in carrying out exercises at home, selection of exercises for oculocervicokinetic re-education, etc.) that could be identified by patients' reports.

#### ACKNOWLEDGEMENTS

This study was supported by research grants from the Health and Social Politics Council of the Autonomous Community of Murcia (Project No. 000509) and the Ministry of Health and Consumers Affairs (PI030317) Spain. Members of the patients' report for physical therapy study group were: Ana M. Fernández, Nuria Hernández, Immaculada Esteban, and J. Fermín Valera-Garrido.

#### REFERENCES

1. Donabedian A. Evaluating the quality of medical care. *Milbank Q* 1996; 44: 166–169.
2. Hargraves JL, Heather R, Zapka J, Nerrenz D, Frazier H, Oray J, et al. Using patient reports to measure health care system performance. *Clin Perform Qual Health Care* 1993; 1: 208–213.
3. Brown JB, Adams ME. Patients as reliable reporters of medical care process: recall of ambulatory encounter events. *Med Care* 1992; 30: 400–411.
4. Davies AR, Ware JE. Involving consumers in quality of care assessment. *Health Aff* 1988; 7: 33–48.
5. Harlow SD, Linet MS. Agreement between questionnaire data and medical records: the evidence for accuracy of recall. *Am J Hygiene* 1989; 129: 233–247.
6. Aker PD, Gross AR, Goldsmith CH, Peloso P. Conservative management of mechanical neck pain: systematic overview and meta-analysis. *BMJ* 1996; 313: 1291–1296.
7. Storro S, Moen J, Svebak S. Effects on sick-leave of a multidisciplinary rehabilitation programme for chronic low back, neck or shoulder pain: comparison with usual treatment. *J Rehabil Med* 2004; 36: 12–16.
8. Kjellman G, Oberg B. A randomized clinical trial comparing general exercise, McKenzie treatment and a control group in patients with neck pain. *J Rehabil Med* 2002; 34: 183–190.
9. Saturno P, Medina F, Valera F, Montilla J, Escolar P, Gascón JJ. Validity and reliability of guidelines for neck pain treatment in primary health care. A nationwide empirical analysis in Spain. *Int J Qual Health Care* 2003; 15: 487–493.
10. Gahimer JE, Domholdt E. Amount of patient education in physical therapy practice and perceived effects. *Phys Ther* 1996; 76: 1089–1096.
11. Sluijs EM. A checklist to assess patient education in physical therapy practice: development and reliability. *Phys Ther* 1991; 71: 561–569.
12. Revel M, Minguet M, Gregoy P, Vaillant J, Manuel JL. Changes in cervicocephalic kinesthesia after a proprioceptive rehabilitation program in patients with neck pain: a randomised controlled study. *Arch Phys Med Rehabil* 1994; 75: 895–899.
13. Berg HE, Berggren G, Tesch PA. Dynamic neck strength training effect on pain and function. *Arch Phys Med Rehabil* 1994; 75: 661–665.
14. Randlov A, Ostergaard M, Manniche C, Kryger P, Jordan A, Heegaard S, et al. Intensive dynamic training for females with chronic neck/shoulder pain. A randomised controlled trial. *Clin Rehabil* 1998; 12: 200–210.
15. Levoska S, Keinänen-Kiukaanniemi S. Active or passive physiotherapy for occupational cervicobrachial disorders? A comparison of two treatment methods with a 1-year follow-up. *Arch Phys Med Rehabil* 1993; 74: 425–430.
16. Jordan A, Bendix T, Nielsen H, Hansen Rolsted F, Host K, Winkel A. Intensive training, physiotherapy or manipulation for patients with chronic neck pain. A prospective, single-blinded, randomised clinical trial. *Spine* 1998; 23: 311–319.
17. Hulka BS, Kupper LL, Cassel JC, Efrid RL. Medication use and misuse: physician-patient discrepancies. *J Chronic Dis* 1975; 28: 7–21.
18. Brown JB, Adams ME. Patients as reliable reporters of medical care process: recall of ambulatory encounter events. *Med Care* 1992; 30: 400–411.
19. Nunally JC, Bernstein IH. *Psychometric theory*, 3rd edn. New York: MacGraw-Hill; 1994.
20. Mangione TW. *Mail surveys: Improving the quality*. London: Sage; 1995.
21. Aday JA. *Designing and conducting health surveys*. San Francisco: Jossey-Bass; 1996.
22. Kleinbaum DG, Klein M. *Logistic regression. A self-learning text*, 2nd edn. New York: Springer Verlag; 2002.
23. Griffith LE, Cook DJ, Guyatt GH, Charles CA. Comparison of open and closed questionnaire formats in obtaining demographic information from Canadian general internist. *J Clin Epidemiol* 1999; 52: 997–1005.
24. Scott C. Research on mail surveys. *J R Stat Soc* 1961; 24: 143–205.
25. Gascón Cánovas JJ, Saturno Hernández PJ, Sánchez Sánchez JA. Utilidad del informe del usuario en la monitorización del uso racional de medicamentos. *Aten Primaria* 2001; 27: 395–402 (in Spanish).
26. Yasin S. Detecting and improving compliance. Is concordance the solution? *Aust Fam Physician* 1998; 27: 255–260.
27. Demlo LK, Campbell PM, Brown SS. Reability of information abstracted from patient's medical records. *Med Care* 1978; 16: 995–1005.
28. Payne B. The medical record as a basis for assessing physician competence. *Ann Intern Med* 1979; 94: 623–629.