Review

Assessment instruments for patients with fibromyalgia: properties, applications and interpretation

F. Salaffi[1], P. Sarzi-Puttini[2], A. Ciapetti[1], F. Atzeni[2]

[1]*Rheumatology Department, Polytechnic University of Marche, Ancona, Italy;*
[2]*Rheumatology Unit, L. Sacco University Hospital, Milan, Italy.*

Fausto Salaffi, MD
Piercarlo Sarzi Puttini, MD
Alessandro Ciapetti, MD
Fabiola Atzeni, MD

Please address correspondence to:
Fausto Salaffi, MD,
Clinica Reumatologica,
Università Politecnica delle Marche,
c/o Ospedale "A. Murri",
Via dei Colli 52, 60035 Jesi, Ancona, Italy.
E-mail: fsalaff@tin.it

Received on November 26, 2009; accepted in revised form on December 10, 2009.

Clin Exp Rheumatol 2009: 27 (Suppl. 56): S92-S105.

Key words: Fibromyalgia, assessment instruments, pain, fatigue, disturbed sleep, health-related quality of life, clinimetric properties.

Competing interests: none declared.

ABSTRACT

A comprehensive assessment of the multiple symptom domains associated with fibromyalgia (FM) and the impact of FM on multidimensional aspects of function should form a routine part of the care of FM patients. Clinical trials and long-term clinical registries have used various outcome measures, but the key domains include pain, fatigue, disturbed sleep, physical functioning, emotional functioning, patient global ratings of satisfaction, and their health-related quality of life (HRQL). A number of measures have been "borrowed" from the fields of rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis and adapted to FM, and others are being developed specifically for FM. However, despite the burgeoning theoretical literature and the proliferation of instruments for measuring various health status domains, no unified approach has been developed and there is little agreement concerning the meaning of the results. There is, therefore, still a need for further consensus and the development of a core set of measures and response criteria, more refined measuring instruments, standardised assessor training, cross-cultural adaptations of health status questionnaires, electronic data capture, and the introduction of standardised quantitative measurements into routine clinical care. This article discusses the advantages and limitations of a selection of both newly developed and well-established and validated distress screening instruments that underlines the continuing challenge of assessing FM.

Introduction

Fibromyalgia (FM) is a chronic condition characterised by generalised pain with characteristic tender points upon physical examination that is often accompanied by a number of associated symptoms such as fatigue, sleep disturbances, psychological and cognitive alterations, headache, migraine, variable bowel habits, diffuse abdominal pain, and increased urinary frequency (1-3). It affects at least 2% of the general population in Italy, and more than 90% of the patients are female (4, 5). The societal importance of this condition is underlined by the fact that its economic consequences are as great as those related to chronic low back pain (6). FM is frequently associated with depression, anxiety, memory and concentration difficulties, and accompanied by other chronic painful disorders.

Its outcomes are not clear, but recent studies suggest that FM patients are characterised by an increased number of physician visits, a self-reported reduction in the ability to perform daily activities, a reduced health-related quality of life (HRQL), and an increased risk of qualifying for a disability pension (7-11).

It is difficult to evaluate the effects of FM therapy because of the many aspects of the syndrome, which also explains why it is usually treated with a wide range of treatments. Although some therapies have been tested in randomised controlled trials (RCTs), the lack of standardisation and outcome measures has prevented any clear evaluation of their effects. In an attempt to identify the appropriate outcome domains, a multidimensional set of core symptoms (12) has been proposed for use in clinical trials that includes pain, tenderness, patient global status, fatigue, HRQL, physical function, disturbed sleep, depression and anxiety, and dyscognition (cognitive dysfunction), and received a high level of consensus among the attendees of OMERACT 9 (Outcome Measures in Rheumatoid Arthritis Clinical Trials) (13).

However, given the multifaceted nature of FM and the new therapies currently being tested (14), further measures are needed in order to develop a reliable and valid composite patient-reported outcome (PRO) response measure that more accurately assesses treatment effects (12). The validity, reliability and responsiveness of PRO data in evaluating and monitoring patients with rheumatic conditions have been clearly documented (15-17), and this article reviews the literature concerning the clinimetric properties of PRO instruments, their advantages and their limitations.

Assessing pain

Chronic generalised pain is a core feature of FM (1-3, 12, 13), and its assessment involves: i) patient reports of typical pain; and ii) an evaluation of the hypersensitivity to palpation of specific tender points (TPs).

Patient reports of pain

The available instruments for the assessment of pain incude visual analogue scales (VAS), verbal descriptor scales (VDS), numerical rating scale (NRS), a daily pain diary, the McGill Pain Questionnaire (MPQ) and the Short-Form McGill Pain Questionnaire (SF-MPQ), the Wisconsin Brief Pain Questionnaire (BPQ), the Brief Pain Inventory (BPI), and the Multidisciplinary Pain Inventory (MPI) (1, 18-20).

The standard VAS is a 10 cm scale with a border on each side: to the left of the "0" mark appears the indication "no pain at all", and to the right of the "10" mark "pain as bad as it could be". A number of studies have shown that the data obtained using self-report VAS scales are reproducible, but one of their limitations is that they must be administered on paper or electronically (18-21); furthermore, caution is required when photocopying them as this can lead to significant changes in length (20, 22). VDSs use categories such as "none", "mild", "moderate", "severe" and "excruciating" to describe severity. An NRS ranging from 0 ("no pain") to 10 ("worst pain imaginable") is more practical than a VAS as it is easier for most people to understand, and does not require vision, dexterity, or paper and pen; another customary range for an NRS is 0–100. It is even possible to determine the intensity of pain accurately by means of a telephone or computerised telephone interview, with the NRS data given by the patient being recorded in a computer database by an operator or directly via the telephone keyboard.

All three measures closely correlate with each other, although the correlation between NRS and VAS is the closest (23-25). Clinical trials have shown that an NRS is more reliable than a VAS, especially in the case of less educated patients (26), a critical issue that has also been pointed out by Joyce *et al*. (27). Simplicity and ease of rating are overriding criteria for pain assessments in clinical settings, which explains the prevalent use of a simple 0–10 cm NRS (23, 26), although a daily diary has also been used and found to be a useful means of identifying how pain affects the everday living activities of individuals.

Pain diagrams or drawings

As widespread pain is one of the two FM classification criteria proposed by the *American College of Rheumatology* (ACR) (28), and widespread pain and/or the extent of pain has been the subject of many investigations, various simple pain diagrams or drawings have been validated. Two of these are the Regional Pain Scale (RPS) (29) and the Self-assessment Pain Scale (SAPS) (30). The RPS is a valid means of measuring the extent of pain that can be used to identify patients with FM, including those with concomitant rheumatoid arthritis (RA) and osteoarthritis (OA); furthermore, as it is disease independent, it works just as well in identifying the patients with severe RA or OA alone who are likely to make the greatest of the available resources (29). The SAPS considers 16 non-articular sites by asking patients to "indicate below the amount of pain and/or tenderness you have experienced in the last 7 days in each of the body areas", and has a series of site descriptions followed by four boxes labelled 0 = none, 1 = mild, 2 = moderate, and 3 = severe; the possible scores therefore range from 0 to 48 but, in order to integrate them into one scale, they have been transformed into a 0-10 scale (30).

Patient self-report questionnaires

The development of a clinical science of pain assessment using patient self-report questionnaires has led to the creation of numerous instruments for evaluating various types and subtypes of chronic pain conditions and their impact on function. These often provide information about both the quality and quantity of pain, and many of them also provide information concerning a patient's psychological and functional status. However, their length may limit patient acceptance, especially if administered during the painful experience.

The complete McGill Pain Questionnaire (MPQ) is one of most widely tested instruments, and can provide detailed information on the characteristics of pain in FM (31-33). However, it is complex (it includes 78 pain adjectives divided into the four major categories – sensory, affective, evaluative, and miscellaneous sensory) and takes 15-20 minutes to complete. It also includes questions concerning changes in pain over time, and classifies pain intensity as "mild", "discomforting", "distressing", "horrible" and "excruciating" (33-35). This makes it more difficult to administer in non-research clinical settings, and simpler measures – such as a VAS – have become more widely accepted. The Short-Form (SF)-MPQ is a 15-item self-report scale derived from the original MPQ (36) that contains three components. The Pain Rating Index (PRI) consists of 15 representative words on a 4-point Likert-type scale ranging from 0 (*none*) to 3 (*severe*). It includes 11 sensory (e.g. tender) and four affective (*e.g*. sickening) items, and there are two items measuring pain intensity. Overall pain is assessed by means of an NRS based on a 10 cm line that approximates ratings between 0 (*no pain*) and 10 (*unbearable pain*) (36). It includes the PRI of the standard MPQ and an NRS (32, 33, 35).

The Wisconsin Brief Pain Questionnaire (BPQ) is a self-administered instrument that assesses pain history, worse pain, usual pain and pain now (37) using a human figure that is shaded

to indicate pain, pain intensity, the relief obtained from medication, and ratings of pain interference (0 = not at all, 1 = a little bit, 2 = moderately, 3 = quite a bit, 4 = extremely).

The Brief Pain Inventory (BPI) was developed to provide information about pain intensity (the sensory dimension) and the extent to which pain interferes with function (the reactive dimension), and also asks questions about pain relief, pain quality, and the patient's perception of the cause of pain. It uses a 0-10 NRS for item rating because of its simplicity, lack of ambiguity, and ease of use for cross-linguistic pain measurement. As pain can vary during the day, it asks patients to rate their pain at the time of responding to the questionnaire (pain now), and also the worst, least and average pain over the previous week (37); ratings can also be made for the previous 24 hours. Evidence for the validity of the BPI comes from a number of studies involving patients with FM patients with other painful diseases (38, 39).

The Multidisciplinary Pain Inventory (MPI) is a 61-item questionnaire that provides a more generalised measure of chronic pain and its impact (40, 41). It is divided into three sections ("impact of pain on patient's life", "responses of others to patient's communication of pain", and "participation in common daily activities") and 13 scales measuring pain severity, interference, life control, affective distress, support, punishing responses, solicitous responses, distracting responses, household chores, outdoor work, activities away from home, social activities, and general activities. The responses are given using a 7-point numerical scale. The MPI has been shown to be reliable and valid for both chronic pain and FM (42).

Assessment of pain hypersensitivity - tender point (TP) assessment

Another critical pain parameter in FM is hyperalgesic responses to external stimulation. Tender point (TP) assessment is a demonstrably useful part of the official ACR criteria for a diagnosis of FM (28). The guidelines proposed by the ACR indicate that the examination should be carried out by applying, bilaterally, the same manual finger pressure with a force of 4 kg (until blanching of the fingernail bed) at nine anatomical sites: occiput, low cervical, trapezius, supraspinatus, second rib, lateral epicondyle, gluteal, greater trochanter and knee. A TP is considered "positive" when the patient reports pain during the examination (43), and the score is the total number of TPs. In addition to the tender point count, other assessments of intensity have been developed but it does not seem that their use has increased accuracy (43).

Another method of measure hyperalgesia is to use myalgic scores based on dolorimetry. These are pain thresholds based on the amount of force required to elicit pain at each of the 18 FM TPs. Digital and dolorimeter assessments are methodologically different (44) as the former requires palpation at a constant force, whereas the latter is based on the amount of force required to induce pain. Their scores are affected by the different tactile sensations and surface areas involved, and the two methods may actually assess different aspects of hyperalgesia.

Assessing fatigue

Many of the validated instruments for measuring fatigue have been used in FM patients, but there is still no consensus as to which should be preferentially used (45).

The multidimensional nature of fatigue underlines the challenge of its assessment in a research setting. Although it can be assessed monodimensionally (*e.g.* by an intensity measurement alone), or as a dichotomous variable (the presence or absence of a defined criterion), or by means of four- or five-point VDS or NRS (Table I), the simplicity of these approaches must be balanced against the missed opportunity to capture information concerning other dimensions, including qualitative differences that may distinguish clinically meaningful subtypes (46); however, these simple scales presumably provide a global measure of fatigue severity.

Another instrument that has been validated in a number of rheumatic conditions is the vitality scale (VT) of the Medical Outcome Study (MOS)-SF36 (47, 48), which explores fatigue and the related concept of energy level. Item responses are rated on a 6-point Likert scale from "all the time" to "none of the time", and the score can vary from 0 (the worst score) to 100 (the best) (47).

Multidimensional fatigue assessment

Multidimensional fatigue assessment captures more information about the characteristics or impact fatigue, such as the global quality of life and symptom distress. Efforts to measure multiple dimensions began thirty years ago in non-medically ill populations but, since then, many instruments of this type have been validated in populations with chronic diseases (49) (Table I), some of which complement the measurement of fatigue severity by providing information concerning other characteristics, while others measure

Table I. Monodimensional fatigue measurements.

Type	Score
4-point verbal rating scale	None, mild, moderate, severe
5-point verbal rating scale	None, mild, moderate, severe, very severe
11-point NRS	How severe has fatigue been, on average, during the past week on a "0 (no fatigue) – 10 (worst fatigue imaginable) scale"
4- point numerical scale	0 = none 1 = increased fatigue over baseline, but not altering normal activities 2 = moderate fatigue or fatigue causing difficulty in performing some activities 3 = severe fatigue or an inability to perform some activities 4 = bed-ridden
VAS	0 (no fatigue) – 10 (worst possible fatigue)

NRS: numerical rating scale; VAS: visual analogue scale.

Table II. Characteristics of the self-administered fatigue instruments.

Instrument	No. of items	Response format	Score range	Measures
FibroFatigue scale (57)	12	–	–	Impact of fatigue impact on specific types of functioning
MAF (58)	16	10-point RS (14 items) or multiple-choice (4 choices) responses (2 items)	1-50	Degree, severity, distress, impact on activities of daily living
MFI (59)	20	5-point RS	20-100	General fatigue, physical fatigue, reduced activity, reduced motivation, mental fatigue
FACIT-F (60)	13	5-point RS	0-52	Severity, role and social impact
FSS (50)	9	7-point RS	1-7	Severity, physical, mental and social impact

MAF: Multidimensional Assessment of Fatigue; MFI: Multidimensional Fatigue Inventory; FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue scale; FSS: Fatigue Severity Scale; RS: rating scale.

the impact of fatigue on different types of functioning (50-56).

Multidimensional fatigue questionnaires have advantages and disadvantages. One important advantage is that they make it possible to analyse and clarify the nature of a fatigue syndrome or evaluate its response to treatment; furthermore, the broader range of captured experiences can add to its validity or improve its sensitivity to clinical changes. However, the disadvantages must also be considered.

A variety of measures have proved to be useful in measuring fatigue in FM and other rheumatic diseases, including the Fibromyalgia and Chronic Fatigue Syndrome Rating Scale (the FibroFatigue scale) (57), the Multidimensional Assessment of Fatigue (MAF) (58), and the Multidimensional Fatigue Index (MFI) (59), which measures various types of fatigue including physical and emotional fatigue. Another measure that has been validated in a number of diseases is the Functional Assessment of Chronic Illness Therapy (FACIT-Fatigue) system (60), which can be customised to certain indications. Finally, the Fatigue Severity Scale (FSS) (50), which was originally developed to asses fatigue in multiple sclerosis and lupus, can also be used in FM (Table II).

Fibromyalgia and Chronic Fatigue Sindrome Rating Scale (FibroFatigue Scale)

The FibroFatigue scale (57) is an observer's rating scale whose 12 items measure pain, muscular tension, fatigue, concentration difficulties, failing memory, irritability, sadness, sleep disturbances, autonomic disturbances, irritable bowel, headache, and the subjective experience of infection. Its inter-rater reliability is excellent, and it has been shown to be reliable, valid, capable of monitoring symptom severity and changes during treatment in patients with chronic fatigue syndrome and FM, and effective in detecting and measuring functional disability and symptom severity in FM patients (61, 62).

Multidimensional Assessment of Fatigue (MAF)

The Multidimensional Assessment of Fatigue (MAF) scale (58) is a good means of measuring fatigue in chronic illness as it is easy to administer and score, relatively short, and assesses the subjective aspects of fatigue by means of 16 items that cover the four dimensions of fatigue severity, distress, degree of interference in activities of daily living, and timing. Fourteen items are rated using a 10-point numerical scale, and two by means of multiple-choice responses with four choices. A global fatigue index ranging from 1 (no fatigue) to 50 (severe fatigue) can be computed using 15 of the 16 items (58).

Multidimensional Fatigue Inventory (MFI)

The MFI is organised in five dimensions (general fatigue, physical fatigue, reduced activity, reduced motivation, mental fatigue), each based on four statements (59) with five possible responses to each statement ranging from "yes, that is true" to "no, that is not true". A global fatigue score combining the five dimensions ranges from 20 to 100, with higher scores indicating greater fatigue. The psychometric properties of the MFI have been well documented, and it has been frequently used in rheumatic disorders, including FM (63).

Functional Assessment of Chronic Illness Therapy-Fatigue scale (FACIT-Fatigue)

This has 13 items and a five-point Likert-type rating scale (0 = "not at all"; 4 = "very much"), and explores the severity of fatigue on a monodimensional basis (60). The total score is the sum of the individual items, and ranges from 0 (maximum fatigue) to 52 (no fatigue). It is widely used to measure cancer-related fatigue, and has also been used in primary Sjögren's syndrome (64) and RA (65).

Fatigue Severity Scale (FSS)

The Fatigue Severity Scale (FSS) (50) consists of nine items and has a 7-point response format. Sample questions include "I am easily fatigued" and "exercise brings on my fatigue." The initial validation study found that its internal consistency was high for specific disease groups and healthy controls: it clearly distinguished patients from controls and moderately correlated with a single-item visual analogue scale of fatigue intensity. In all of the patients, a clinical improvement in fatigue was associated with reductions in FSS scores. The scale is also practical as it is brief and easy to administer and score.

Please complete the following question upon awakening. Select the number that best describes the quality of your sleep during the past 24 hours.										
0	1	2	3	4	5	6	7	8	9	10
Best possible sleep										**Worst possible sleep**

Fig. 1. Sleep Quality Numerical Rating Scale.

Assessing sleep

FM patients frequently report disturbed sleep (1-3): estimates of the percentage experiencing some sleep problem range from 70-80% in the population used to establish the ACR criteria (1-3,66-76) to as high as 95% and 99% in two recent studies (77, 78). It has also been shown that the symptoms of disturbed sleep in FM predict increased pain levels and decreased physical functioning (77, 79-81), and so accurately assessing the changes in sleep associated with FM treatments is critically important.

Various dimensions of sleep have been assessed in FM trials, including quantity, quality, the ease of falling asleep, the frequency of waking, and feeling refreshed upon awakening. The quality of sleep can be assessed using a single-item measure (the Sleep Quality NRS) (81), which instructs patients to "select the number that best describes the quality of your sleep during the past 24 hours" (0 = "best possible sleep" and 10 "worst possible sleep") (Fig. 1) or multidimensional instruments (82,83). A number of multidimensional measures have proved to be useful in measuring disturbed sleep in rheumatic diseases, including the Medical Outcome Study Sleep Scale (MOS-SS) (84,85), the Pittsburgh Sleep Quality Index (PSQI) (86), the Pittsburgh Sleep Diary (PSD) (87), and the Insomnia Severity Index (ISI) (88), of which the MOS-SS may represent the best choice.

Medical Outcome Study Sleep Scale (MOS-SS)

The MOS-SS is a 12-item questionnaire designed to evaluate key constructs of sleep, with derived subscales for the domains of sleep disturbance (4 items), quantity of sleep (1 item), snoring (1 item), awakening short of breath or with headache (1 item), sleep adequacy (2 items), and somnolence (3 items) (84, 85). It is also possible to generate a 9-item Sleep Problems Index that assesses overall sleep problems and includes the four sleep disturbance and two sleep adequacy items, two of the somnolence items, and awakening short of breath/with headache; higher scores indicate greater sleep impairment, and this index is often used in clinical trials as an indication of sleep quality. The MOS-SS has been found to have positive psychometric properties in a broad range of patient populations, including patients with chronic pain conditions similar to FM (89, 90).

Pittsburgh Sleep Quality Index (PSQI)

The Pittsburgh Sleep Quality Index (PSQI) retrospectively measures sleep quality and disturbances (86). It discriminates good and poor sleepers, and provides a brief and clinically useful assessment of multiple sleep disturbances. Its 19 items generate seven component scores, the sum of which (range 0-21) yields a global measure of sleep quality, with higher scores indicating poorer sleep (>5 indicates sleep disturbance). The components assess a broad range of domains associated with sleep quality, including the duration of sleep, sleep latency, the frequency and severity of specific sleep-related problems, and the perceived impact of poor sleep on daytime functioning. The questionnaire is perhaps the most widely used general measure of sleep, and its strengths lie in its coverage of multiple dimensions of sleep quality, its flexibility as a brief clinical tool, and its demonstrated validity and usefulness in chronic pain research and in patients with FM.

Pittsburgh Sleep Diary (PSD)

The Pittsburgh Sleep Diary (PSD) is used to quantify subjectively reported sleep and wake behaviours (87), and is divided into two daily questionnaires completed at "bedtime" and "wake time", with the timing and duration of various daytime and sleep-wake parameters and activities being completed by the participant. The bedtime component consists of six general items: the timing of meals; the consumption of caffeine, alcohol and tobacco products; the use of medications; and the timing and duration of exercise and nap periods. The daytime component gathers data on bedtime, "lights out" time, sleep latency, final wake time, method of final awakening, the frequency of nightly awakenings, wake after sleep onset time, the reasons for nightly awakenings, sleep quality, mood on final wakening, and alertness on final wakening. In addition to the categorical and frequency data generated by the bedtime questionnaire, the daytime questionnaire makes it possible to calculate standard continuity parameters.

Insomnia Severity Index (ISI)

The Insomnia Severity Index (ISI) (88) is a self-report instrument that measures an individual's perception of insomnia. It has seven items and a total score that ranges from 0 to 28: according to the recommended score interpretation guidelines, 0–7 indicates "no clinically significant insomnia", 8–14 "sub-threshold insomnia", 15–21 "clinical insomnia (moderate severity)", and 22–28 "clinical insomnia (severe)". The cut-off level of 14 has optimal sensitivity (94%) and specificity (94%) in distinguishing a group of adults diagnosed with primary insomnia from those without.

Psychological and behavioural assessment

Psychological and behavioural evaluations of FM patients can provide useful information concerning factors that may affect their pain and dysfunction, and give an idea of the impact of pain, fatigue and other symptoms on their psychological health (1-3, 91, 92). Anxiety and depression are major factors affecting a patient's quality of life, and the associated symptoms (inability to concentrate, loss of motivation,

Table III. Definitions and characteristics of screening instruments.

Screening instruments	Items	Time required (minutes)	Advantages	Disadvantages
Ultra-short	1–4	<2	• Very likely to be used in busy clinics • Sensitivity can be high • Low-to-moderate specificity • Inexpensive	• Can only assess one domain • Unsuitable for research
Short	5–20	2-10	• Moderately likely to be used in busy clinics • Probably highly sensitive, moderate-to-high specificity • Can assess multiple domains • May be suitable for research, needs to be tested	• Some cost in scoring
Long	21–50	>10	• Specificity and sensitivity can be high • Can assess multiple domains • Excellent for research	• Routine use unlikely unless automated • Potentially costly scoring (can be minimised by automation)

disturbed sleep, fatigue, pessimistic mood) may affect their response to treatment (14) and rehabilitation programmes (93).

Psychological assessment instruments come in varying lengths and formats (94), and one important factor is their length, defined as the number of questions or items they contain. The term "screening instrument" usually refers to a particularly short test whereas, although longer tests are more expensive to administer, they are sometimes needed to reach acceptable levels of reliability and validity. Table III shows the definitions and characteristics of screening instruments by length, as well as their advantages and disadvantages.

Ultra-short forms are typically limited to one psychological domain, such as depression or anxiety, and are the easiest to use in routine care settings. They usually consist of only one question, take only 1-2 minutes to complete, and require no scoring. Table IV shows the most frequently used questions for depression. A combination of one depression question, a one-question interview, a Distress Thermometer (DT) and an 11-point NRS creates a further ultra-short questionnaire that can be used in everyday practice (Table V). Ultra-short screening instruments have a potential economic advantage because of their brevity and the need for fewer staff resources to administer them. However, although they may be successfully used in busy daily practice, a recent meta-analysis (95) has shown that they are not very accurate in detecting depression in primary care and should only be used to rule out a diagnosis.

Among the short instruments (*i.e.* those with 5–20 items), the Zung Self-rating Depression Scale (ZSDS) (96), the Center for Epidemiologic Studies – Depression Scale (97), the Hospital Anxiety and Depression Scale (98), and the Hamilton Rating Scale for Depression (HRS-D) (99-101) all have adequate psychometric properties. The Somatic Symptoms Checklist (SSC) (102) and the Illness Attitudes Scale (IAS) (103) are less frequently used for FM patients (Table V). The long instruments (*i.e.* those with 21–50 items) include the Beck Depression Inventory (104), the Four-Dimensional Symptom Questionnaire (4DSQ) (106), the Symptom Checklist (SCL-90) (107), and the Rotterdam Symptom Checklist (108) (Table V).

Table IV. Simple verbal questions for depression used as an ultra-short measure.

- 'Are you depressed?'
- 'Are you depressed OR "Have you lost interest?'
- 'Are you depressed?' OR 'Have you experienced a loss of interest in things or activities that you would normally enjoy?'
- 'Over the past couple of weeks, have you been feeling unhappy or depressed?'

Zung Self-rating Depression Scale (ZSDS)

The Zung Self-rating Depression Scale (ZSDS) consists of 10 positively worded items and 10 negatively worded items asking about symptoms of depression (96), and has been found to be a reliable and valid means of measuring depressive symptoms in a number of studies (109-112). ZSDS scores are used to define four categories of severity: within the normal range or no significant psychopathology (<40); the presence of minimal to mild depression (40-47); moderate to marked depression (48-55); and the presence of severe to extreme depression (≥56).

Center for Epidemiologic Studies Depression Scale (CES-D)

The Center for Epidemiologic Studies Depression Scale (CES-D) has 20 items and has been validated in mixed samples of cancer patients and reference groups of healthy control subjects (97). Each item is assessed on a 4-point scale that addresses the frequency of the occurrence of each symptom (0 = none of the time, 3 = all of the time). A cut-off score of 19 is commonly used to indicate a need for a further assessment of depression in patients experiencing pain. Various studies of the scale's sensitivity and specificity have shown that it has very good psychometric properties (113, 114).

Hospital Anxiety and Depression Scale (HADS)

The Hospital Anxiety and Depression Scale (HADS) (98) examines the levels

Table V. Screening instruments for psychological and behavioural assessments.

Screening instruments	No. of items	Validity	Reliability	Generalisable
Ultra-short (1-4 items)				
Depression question	1	Moderate	–	No
Anxiety question	1	Moderate	–	No
One-question interview	1	Moderate	–	Yes
Combination of one depression question	2	Moderate	Moderate	No
Distress Thermometer (DT)	1	Moderate	Moderate	Yes
11-point numerical rating scale	1	Moderate	–	No
Short (5-20 items)				
Zung Self-rating Depression Scale (ZSDS) [96]	20	High	High	Yes
Center for Epidemiologic Studies – Depression Scale (CES-DS) [97]	20	High	High	Yes
Hospital Anxiety and Depression Scale [98]	14	Moderate	High	Yes
Hamilton Rating Scale for Depression (HAM-D) [100]	17	Moderate	Moderate	Yes
Somatic Symptoms Checklist (SSC) [102]	7	Moderate	Moderate	Yes
Illness Attitudes Scales (IAS) [103]	17	Moderate	Moderate	Yes
Long (21-50 items)				
Beck Depression Inventory [104]	21	High	High	Yes
Four-Dimensional Symptom Questionnaire (4DSQ) [106]	50	Moderate	High	Yes
Symptom Checklist (SCL-90) [107]	90	Moderate	Moderate	Yes
Rotterdam Symptom Checklist [108]	30	Moderate	Moderate	Yes

of anxiety and depression in the previous week. It consists of seven items for anxiety (HADS-A) and seven for depression (HADS-D) that are each self-rated on a four-point scale scored 0–3; higher scores are associated with a greater probability of a depressive or anxiety disorder. The depression scale (7 items, score range 0-21) mainly measures anhedonia, which is considered to be the central characteristic of major depressive disorder; the anxiety scale (7 items, score range 0–21) mainly measures symptoms of generalised anxiety disorder. The scale as a whole and each subscale has adequate internal consistency and is sensitive to change (115, 116).

Hamilton Rating Scale for Depression (HAM-D)

The Hamilton Rating Scale for Depression (HAM-D) (100,101) is probably the most widely used observer-rated rating scale doe depressive symptoms. The original scale had 21 items, but Hamilton suggested scoring only the initial 17 because the last four either occurred infrequently or described only aspects of the illness. The items are ranked 0–4 (when severity is quantifiable) or 0–2 (when they measure symptoms that are more difficult to assess reliably), with the highest scores indicating the greatest severity (100). The range for the 17-item scale is 0–50.

Somatic Symptoms Checklist (SSC)

The Somatic Symptoms Checklist (SSC) (102) was originally designed and validated as a screening test for somatisation disorder. It contains six items (and an additional item for females regarding menstrual cramps) in the form of questions (*e.g.* "have you ever had trouble breathing?") requiring a yes/no answer, and the scores are summed to provide the total number of reported somatic symptoms.

Illness Attitudes Scales (IAS)

The Illness Attitudes Scales (IAS) (103) consists of two subscales: health anxiety and illness behaviour. The first contains 11 items (*e.g.* 'are you worried that you may get a serious illness in the future?') scored on a five-point scale (0–4) with total scores ranging from 0 to 44; the second contains six items (*e.g.* 'how often do you see a doctor?') also scored on a five-point scale from 0 ('no') to 4 ('most of the time'), with total scores ranging from 0 to 24.

Beck Depression Inventory (BDI)

The Beck Depression Inventory (BDI) (104) is a 21-question multiple-choice self-report questionnaire that is one of the most widely used by healthcare professionals and researchers for measuring the severity of depression in a variety of settings. It was designed for adults and is composed of items relating to symptoms of depression such as hopelessness and irritability; cognition such as guilt and feelings of being punished; and physical symptoms such as fatigue, weight loss and lack of interest in sex. A cut-off score of >9 is used to indicate at least minimal symptoms of depression. The 13-item BDI–Short Form is also widely used, although it has a low level of inter-rater reliability and is only moderately specific (117).

Four-Dimensional Symptom Questionnaire (4DSQ)

The Four-Dimensional Symptom Questionnaire (4DSQ) is a 50-item self-rating questionnaire that measures "distress", "depression", "anxiety" and "somatisation" (106) by assessing the psychological and psychosomatic symptoms experienced during the previous seven days. The distress scale (16 items, score range 0–32) measures the symptoms of general psychological distress, which is conceptualised as the most general and most basic expression of human psychological suffering; the depression scale (6 items, score range 0–12) measures severe anhedonia and depressive cognitions (including suicidal ideation) as symptoms that are considered to be characteristic of depressive disorder; the anxiety scale (12 items, score range 0–24) measures irrational fears, panic and avoidance, which are characteristic features of most anxiety disorders; and

Table VI. Characteristics of selected generic instruments.

Instrument*	No. of items	No. of levels	Administration°	Scoring options#	Time required (minutes)
SF-36	36	3–6	S, I, P	Pr, SS	10–15
SIP	136	2	S, I, P	Pr, SS, SI	20–30
NHP	38	2	S, I	Pr	10–15
EuroQoL	6	3	S, I	SI	7–10

*SF-36: Medical Outcomes Study 36-Item Short-Form Health Survey; SIP: Sickness Impact Profile; NPH: Nottingham Health Profile; EuroQol: European Quality of Life Questionnaire; °S: self-administered; I: interviewer; P: proxy. #Pr: profile; SS: summary scores; SI: single index.
From: *Franchignoni F. & Salaffi F.* (130).

the somatisation scale (16 items, score range 0–32) measures a range of "psychosomatic" symptoms characteristic of bodily distress and somatoform disorders. Higher scores on all four scales indicate the presence of more symptoms. Two cut-off points are recommended to divide low, moderate and high scores.

Symptom Checklist (SCL-90)

The Symptom Checklist (SCL-90) is used to assess psychological distress and consists of eight dimensions (anxiety, agoraphobia, depression, somatic symptoms, distrust and interpersonal sensitivity, anger, hostility and sleeping disorders) designed to provide an overview of a patient's symptoms and their intensity at a specific time (107). The total SCL-90 score reflects general psychoneurotiscism or psychological distress, by the Global Severity Index can be used as a summary test. The SCL-90 has 90 items and can be completed in just 12–15 minutes.

Rotterdam Symptom Checklist (RSCL)

The Rotterdam Symptom Checklist (RSCL) is a 30-item questionnaire that has been extensively used in clinical trials (108). Although some studies have found that it has a four- or five-factor structure, it has also been suggested that it has a two-factor psychological and composite somatic structure (118). The psychological subscale has proved to be stabile across sub-samples to have a high degree of internal consistency (119, 120).

Assessing health-related quality of life (HRQL) and function

Assessing chronic pain and its impact on physical, emotional and social functions requires multidimensional qualitative and HRQL instruments (121, 122) as it has been shown that measuring HRQL is a key aspect of screening for disability and improving patient/clinician communications. A distinction is drawn between generic and specific measures of physical function and health status (123-130): the first provide a broad picture of health status across a range of conditions, whereas the second are more sensitive to the disorder under consideration and therefore more likely to reflect clinically important changes.

Generic measures

Generic measures, which are commonly developed for descriptive epidemiological or social science research applications, provide a profile of scores for the different components of health status and HRQL, or operational definitions of various constructs summarised by a single index value (130-134). The most widely used are the Medical Outcomes Study (MOS) 36-Item Short-Form Health Survey (SF-36) (47, 48), the Sickness Impact Profile (SIP) (135, 136), and the Nottingham Health Profile (NHP) (137-139) (Table VI).

Medical Outcomes Study (MOS) 36-Item Short-Form Health Survey (SF-36)

The SF-36 is a generic health questionnaire divided into eight scales that measure a different function domains and aspects of well-being (47, 48): 1) Physical functioning (10 items), or the extent to which health limits activities such as self-care, walking, climbing stairs, bending, lifting, and other moderate and vigorous activities; 2) Social functioning (2 items), or the extent to which physical health or emotional problems interfere with normal social activities; 3) Physical role functioning (4 items), or the extent to which physical health interferes with work or other daily activities; 4) Emotional role functioning (3 items), or the extent to which emotional problems interfere with work or other daily activities; 5) Mental well-being (5 items), or general mental health, including depression, anxiety, behavioural-emotional control, and general positive affect; 6) Vitality (4 items), whether one feels energetic and full of pep or tired and worn out; 7) Bodily pain (5 items), which includes the intensity of pain and its effect on normal work inside and outside the house; and 8) General health perceptions (5 items), a personal evaluation of health that includes current health, health outlook, and resistance to illness. The SF-36 also includes a single-item measure of health transition that is not used to score any multi-item scales. The eight scales, which are weighted on the basis of a normative algorithm, are scored from 0 to 100, with higher scores reflecting a better quality of life (48).

Subsequent algorithms have also been developed to calculate two psychometrically based summary measures, the Physical Component Summary Scale Score (PCS) and the Mental Component Summary Scale Score (MCS) which provide greater precision, reduce the number of statistical comparisons needed, and eliminate the floor and ceiling effects noted in several of the sub-scales (48).

It has been reported that, in comparison with healthy populations, FM patients are significantly impaired in all eight domains (125, 140). The SF-36 questionnaire takes about 15 minutes to complete, although most elderly patients prefer a standard interview to the self-administered approach.

The SF-36 was later used to develop the SF-12 (141), which measures the same health status concepts but provides only one score for the PCS and MCS summary measurements (140, 141), although there description is the same as that of the SF-36 PCS and MCS scores.

Sickness Impact Profile (SIP)

The Sickness Impact Profile (SIP) contains 136 items grouped into 12 dimensions of daily activity (ambulation, body care and movement, mobility, social interaction, emotional behaviour, alertness, communication, home management, recreation and pastimes, sleep and rest, eating, and work) (135, 136), and asks respondents check those that apply to them at the time of the interview. Each item is weighted on the basis of the relative severity of dysfunction implied by each statement. The scores for each dimension are summed and expressed as a percentage of the maximum possible score. Three summary scores are also calculated: the total score (includes all domains), a physical score (ambulation, body care and movement, and mobility), and a psychosocial score (social interaction, emotional behaviour, alertness, and communication) (135, 136). Higher scores reflect greater dysfunction. The SIP can be administered by an interviewer or self-administered but, although it is easy to administer and score, it is relatively time-consuming as it takes approximately 30 minutes to complete (135).

Nottingham Health Profile (NHP)

The Nottingham Health Profile (NHP) is a primary healthcare instrument that is intended to provide a brief indication of a patient's perceived emotional, social and physical health problems (137, 138). It originally consisted of two parts, but only part I is now used: it contains 38 items that can be grouped into six domains (physical mobility, pain, sleep, social isolation, emotional reactions, and energy level), with each question being weighted on the basis of severity. The questions were selected from statements generated in large surveys of people randomly selected from the general population, and respondents are required to answer "yes" or "no" to each. Scores range from 0 (no problems or limitations) to 100 (all problems are present). There is no summary score. The sum of all of the weighted values in a given domain represents a continuum between 0 (best health) and 100 (worst health) (137-139).

None of the above generic measures captures the individual value that a given respondent may assign to a particular health state, and two individuals may rate the same state differently depending on the value they assign to a symptom or impairment, and their willingness to accept trade-offs between benefits and risks.

In the context of HRQL evaluations, preference-based (or utility) measures are specifically designed to assess the value or desirability of a particular health status/outcome (142, 143). They provide a final score on a 0–1 scale where 0 is the worst possible imaginable state (or death) and 1 is perfect health. As the ratings can be elicited from different groups of individuals, such as patients, health professionals or the general public, that can be used as quality of life adjustment weights to calculate, for example, quality-adjusted life years and similar measures that can then be used in economic evaluations (142-144).

There are two main approaches to measuring HRQL. The first is to classify patients into categories on the basis of their responses to questions about their functional status (preference classification systems), and combining these categories or dimensions leads to descriptions of their overall health. One such instrument is the European Quality of Life Measure (EuroQol) (144, 145), a self-administered questionnaire used to measure health outcomes (145) that provides a simple descriptive profile and a single index value for health status that can be used for clinical and economic evaluations of health care, as well as in population health surveys. It covers five dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), each of which is divided into three levels (no problems, some or moderate problems, extreme problems), thus generating a total of 243 theoretically possible health states. The EuroQol is self-completed by respondents and ideally suited for use in postal surveys, clinics and face-to-face interviews. It is cognitively simple, and takes only a few minutes to complete (145).

The second approach to utility measurement is to ask patients to assign a value to their overall health directly. The most widely used techniques are rating scales (RS), time trade-offs (TTO) and the standard gamble (SG) technique (130, 142).

Disease-specific measures

Disease-specific measures are designed to assess specific diagnostic groups or patient populations, often with the goal of measuring responsiveness to treatment or "clinically important" changes. One obvious disadvantage of some of them is that they do not allow comparative judgements of the outcomes of different treatments in patients with different health problems, for example for resource allocation studies (130, 131, 133, 134), although this can be overcome by combining the use of disease-specific and generic measures. There are a number of broad disease-specific measures, such as the Fibromyalgia Impact Questionnaire (FIQ) (146,147) or the Revised Fibromyalgia Impact Questionnaire (FIQR) (148), the Arthritis Impact Measurement Scales 2 (AIMS2) (149), and the Health Assessment Questionnaire (HAQ) (150), which cover general aspects of functional status together with specific references to states or changes of particular concern to the target population.

Fibromyalgia Impact Questionnaire (FIQ)

The Fibromyalgia Impact Questionnaire (FIQ) (146, 147) is a 10-item, self-administered, disease-specific assessment and outcome instrument developed to measure the components of health status that are believed to be most affected by FM. The first item contains 11 questions related to physical functioning, each of which is rated using a 4-point Likert-type scale; items 2 and 3 ask the patient to mark the number of days they felt well and the number of days they were unable to work (including housework) because of FM symptoms; and items 4-10 are horizontal linear scales marked in 10 increments for the rating of working difficulties, pain, fatigue, morning tiredness, stiffness, anxiety and depression. Each of the 10 items has a maximum score of 10, and so the maximum possible total score is

100. The scoring is complicated by the need to reverse scores in one question and use constants to convert the first 13 questions to a standardised 0–10 scale. The average FM patient scores about 50, and severely affected patients usually 70+. The FIQ takes approximately five minutes to complete, and has been extensively used as an outcome measure in FM-related studies (151). It appears to be a sensitive measure of changes related to symptoms and disability, and makes it possible to distinguish FM from some other health problems involving chronic pain (30).

Revised Fibromyalgia Impact Questionnaire (FIQR)

The Revised Fibromyalgia Impact Questionnaire (FIQR) attempts to address the limitations of the FIQ while retaining the essential properties of the original (148). It has 21 individual questions framed in the context of the previous seven days, all of which are based on an 11-point NRS, with 10 being "worst". It is divided into three linked sets of domains: a) "function" contains nine questions; b) "overall impact" has two, but they now relate to the overall impact of FM on functioning and overall symptom severity; and c) "symptoms" contains 10 questions, four of which are new and relate to memory, tenderness, balance and environmental sensitivity (loud noises, bright lights, odours, and cold temperatures). The scoring is much simpler than that of the FIQ: the function score (range 0-90) is divided by three, the overall impact score (range 0–20) is unchanged, and the symptoms score (range 0–100) is divided by two, and the total score is the sum of the three modified domains. The weighting is different insofar as 30% of the total score is ascribed to "function" (as opposed to 10% in the FIQ), 50% to "symptoms" (as opposed to 70% in the FIQ), while "overall impact" remains the same at 20%, as does the maximum total score of 100. The FIQR takes approximately half as long to complete as the FIQ (148).

Arthritis Impact Measurement Scale 2 (AIMS2)

The Arthritis Impact Measurement Scale 2 (AIMS2), a widely used disease-specific measure with a broad scope that is used to assess functional limitations and disability, has two versions, AIMS2 (78 items) and AIMS2 SF (26 items) (149), both of which are designed to assess the severity of arthritic pain and the extent to which it affects health (152,153). The respondents are asked to consider the areas of mobility, walking and bending, hand and finger function, arm function, self-care, household tasks, social activity, family support, arthritic pain, work, level of tension, and mood over the previous month and, for each area, rate their degree of satisfaction, the impact of the disease, and where they would like to see improvements. Finally, they are asked to summarise their current, future and overall perceptions of health, and to describe any existing medical problems that affect it.

Health Assessment Questionnaire (HAQ)

The most widely used form of the Stanford Health Assessment Questionnaire (HAQ) is a 20-item, self-administered questionnaire that examines difficulties in performing eight daily living activities (dressing and grooming, rising, eating, walking, hygiene, reach, grip, and outside activities) (150). For each item, the patients are asked to rate the level of difficulty over the previous week on a 4-point scale ranging from 0 (no difficulty) to 3 (unable to perform). The final HAQ score is the average of the eight category scores; it ranges from 0 to 3, with the highest score representing the greatest disability.

Various modifications have been made to the HAQ for RA: the Multidimensional HAQ (MDHAQ) keeps one question from each of the eight categories, thus reducing the number of items to eight, and its score is calculated as the mean of the scores for each activity. The MDHAQ includes 10 activities of daily living (ADLs), eight derived from the HAQ and two additional complex ADLs: walking two miles and participating in sports and games (154). The MDHAQ also includes VAS's to assess pain, fatigue and global status, and a listing of 57 symptoms. To analyse the quantitative scores for pain, fatigue, functional disability, and the number of symptoms on a review of systems (including the ratios of scores for pain to physical function and fatigue to physical function), and to study further how these scores can help to identify patients with FM, DeWalt *et al*. analysed 78 consecutive patients with FM over a two-year period, using 149 patients with RA as a "control" group. The results demonstrated that the FM patients had significantly higher pain:physical function and fatigue:physical function ratios, and reported a significantly larger number of symptoms (155).

Measures of overall health status

The number of TPs (a surrogate for diffuse pain) does not fully capture the essence of FM syndrome, in which accompanying fatigue is often severe and nearly always present, but the Symptom Intensity Scale (SIS) (156) and Fibromyalgia Assessment Status (FAS) (30) are accurate surrogate composite measures. Unlike instruments intended for a particular disease such as the Disease Activity Score (DAS) (157, 158), which measures disease activity only in RA, SIS (156) or FAS scores (30) can be used as a measure of global health status (or disease severity).

The SIS questionnaire consists of two parts: a list of 19 anatomical areas concerning which patients are asked whether they feel pain (the total number of "yes" answers being the RPS score), and a VAS for fatigue (156). The SIS score can be used to identify and quantify FM simply from the information supplied. As the continuous SIS score closely correlates with the patient' perceived pain and general health, it is ideal for outpatient evaluations and complements a complete patient history and physical examination by measuring biopsychosocial factors.

The FAS index is a short and easy to complete self-administered instrument that combines a set of questions relating to non-articular pain (SAPS range 0–10), fatigue (range 0–10) and the quality of sleep (range 0–10), thus providing a single composite measure of disease severity ranging from 0-10 (30). The final score is calculated by adding the three sub-scores and dividing the

result by three. All three measures are printed on one side of one page for rapid review, and scored by a health professional without the need for a ruler, calculator, computer or Website. Our data suggest that it is a reliable, valid and responsive disease-specific composite measure for assessing treatment effects in patients with FM, and is suitable for use in clinical trials and everyday clinical practice (30).

Conclusions

This paper reviews the most widely used and studied instruments for assessing FM patients. Only those with detailed psychometric examinations have been described; many others have not been considered because they lack extensive statistical analyses (sometimes because they are too recent) or were simply judged to be inferior. In addition to considering the general relative merits of the instruments indicated in this paper, it is essential to examine each one carefully in order to ensure that whichever is selected is appropriate to the specific purpose and requirements of a study. The strengths and weaknesses of the various instruments may vary depending on the population and the reason for use, and so any final decision needs to be context specific. For evaluative applications, outcome measurement procedures should meet each of three major criteria: validity, reliability, and responsiveness. The first two are important for any measurement, but responsiveness (sensitivity to change) is the quintessential requirement of any procedure aimed at evaluating change after effective treatment (159, 160). Furthermore, some pragmatic issues are also important (161-164), including interpretability (the measures should provide results that are easily understood by others), acceptability (how acceptable it is for respondents in terms of response rate, completion time, cultural applicability, and so on), and feasibility (ease of administration and processing: *i.e*. the amount of effort and disruption it causes to staff and clinical care activities including, for example, the professional expertise required to apply or interpret it, and the presence of a clear instruction manual). Finally, there should be concrete evidence concerning the usefulness of these measures in improving our understanding of the complex relationships between interventions, clinical and context variables, and outcomes.

References

1. MEASE P: Fibromyalgia syndrome: review of clinical presentation, pathogenesis, outcome measures, and treatment. *J Rheumatol* 2005; 32 (Suppl. 75): 6-21.
2. BELT NK, KRONHOLM E, KAUPPI MJ: Sleep problems in fibromyalgia and rheumatoid arthritis compared with the general population. *Clin Exp Rheumatol* 2009; 27: 35-41.
3. ORELLANA C, CASADO E, MASIP M, GALISTEO C, GRATACÓS J, LARROSA M: Sexual dysfunction in fibromyalgia patients. *Clin Exp Rheumatol* 2008; 26: 663-6.
4. WOLFE F, ROSS K, ANDERSON J, RUSSELL IJ, HEBERT L: The prevalence and characteristics of fibromyalgia in the general population. *Arthritis Rheum* 1995; 38: 19-28.
5. SALAFFI F, DE ANGELIS R, GRASSI W; MARCHE PAIN PREVALENCE; INVESTIGATION GROUP (MAPPING) STUDY: Prevalence of musculoskeletal conditions in an Italian population sample: results of a regional community-based study. I. The MAPPING study. *Clin Exp Rheumatol* 2005; 23: 819-28.
6. BOONEN A, VAN DEN HEUVEL R, VAN TUBERGEN A *et al*.: Large differences in cost of illness and wellbeing between patients with fibromyalgia, chronic low back pain, or ankylosing spondylitis. *Ann Rheum Dis* 2005; 64: 396-402.
7. KASSAM A, PATTEN SB: Major depression, fibromyalgia and labour force participation: a population-based cross-sectional study. *BMC Musculoskelet Disord* 2006; 7: 4.
8. LUNDBERG G, GERDLE B: Tender point scores and their relations to signs of mobility, symptoms, and disability in female home care personnel and the prevalence of fibromyalgia syndrome. *J Rheumatol* 2002; 29: 603-13.
9. BRUUSGAARD D, EVENSEN AR, BJERKEDAL T: Fibromyalgia – a new cause for disability pension. *Scand J Soc Med* 1993; 21: 116-9.
10. PICAVET HS, HOEYMANS N: Health-related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DMC3 study. *Ann Rheum Dis* 2004; 63: 723-9.
11. SALAFFI F, DE ANGELIS R, STANCATI A, GRASSI W. MARCHE PAIN; PREVALENCE INVESTIGATION GROUP (MAPPING) STUDY: Health-related quality of life in multiple musculoskeletal conditions: a cross sectional population based epidemiological study. II. The MAPPING study. *Clin Exp Rheumatol* 2005; 23: 829-89.
12. MEASE P, ARNOLD LM, CHOY EH *et al*.: Fibromyalgia syndrome module at OMERACT 9: domain construct. *J Rheumatol* 2009; 36: 2318-29.
13. CHOY EH, ARNOLD LM, CLAUW DJ *et al*.: Content and criterion validity of the preliminary core dataset for clinical trials in fibromyalgia syndrome. *J Rheumatol* 2009; 36: 2330-4.
14. SARZI-PUTTINI P, TORTA R, MARINANGELI F *et al*.: Italian Fibromyalgia Network. Fibromyalgia syndrome: the pharmacological treatment options. *Reumatismo* 2008; 60 (Suppl. 1): 50-8.
15. PINCUS T, MACLEAN R, YAZICI Y, HARRINGTON JT: Quantitative measurement of patient status in the regular care of patients with rheumatic diseases over 25 years as a continuous quality improvement activity, rather than traditional research. *Clin Exp Rheumatol* 2007; 23 (Suppl. 47): S69-S81.
16. LEVY G, CHEETHAM C, CHEATWOOD A, BURCHETTE R: Validation of patient-reported joint counts in rheumatoid arthritis and the role of training. *J Rheumatol* 2007; 34: 1261-5.
17. SOKKA T, HAUGEBERG G. PINCUS T: Assessment of quality of rheumatoid arthritis care requires joint count and/or patient questionnaire data not found in a usual medical record: examples from studies of premature mortality, changes in clinical status between 1985 and 2000, and a QUEST-RA global perspective. *Clin Exp Rheumatol* 2007; 47: S86-S97.
18. KREMER E, ATKINSON JH, IGNELZI RJ: Measurement of pain: patient preference does not confound pain measurement. *Pain* 1981; 10: 241-8.
19. DIXON JS: Agreement between horizontal and vertical visual analogue scales. *Br J Rheumatol* 1986; 25: 415-6.
20. JENSEN MP, KAROLY P, BRAVER S: The measurement of clinical pain intensity: a comparison of six methods. *Pain* 1986; 27: 117-26.
21. GUYATT GH, TOWNSEND M, BERMAN LB, KELLER JL: A comparison of Likert and visual analogue scales for measuring change in function. *J Chronic Dis* 1987; 40: 1129-33.
22. SNOW S, KIRWAN JR: Visual analogue scales: a source of error. *Ann Rheum Dis* 1988; 47: 526.
23. PRICE DD, BUSH FM, LONG S, HARKINS SW: A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. *Pain* 1994; 56: 217-26.
24. SYRJALA KL: The measurement of pain. *In*: YARBO CH, MCGUIRE DB, editors. Cancer Pain: Nursing Management. Orlando, FL: Grune & Stratton; 1987. p. 133-50.
25. SALAFFI F, STANCATI A, SILVESTRI CA, CIAPETTI A, GRASSI W: Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain* 2004; 8: 283-91.
26. FERRAZ MB, QUARESMA MR, AQUINO LRL, ATRA E, TUGWELL P, GOLDMITH CH: Reliability of pain scales in the assessment of literate and illiterate patients with rheumatoid arthritis. *J Rheumatol* 1990; 17: 1022-4.
27. JOYCE CRB, ZUTSHI DW, HRUBES V, MASON RM: Comparison of fixed internal and visual analogue scales for rating chronic pain. *Eur J Clin Pharmacol* 1975; 8: 415-20.

28. WOLFE F, SMYTHE HA, YUNUS MB *et al.*: The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis Rheum* 1990; 33: 160-72.
29. WOLFE F: Pain extent and diagnosis: development and validation of the Regional Pain Scale in 12,799 patients with rheumatic disease. *J Rheumatol* 2003; 30: 369-78.
30. SALAFFI F, SARZI-PUTTINI P, GIROLIMETTI R, GASPARINI S, ATZENI F, GRASSI W: Development and validation of the self-administered Fibromyalgia Assessment Status: a disease-specific composite measure for evaluating treatment effect. *Arthritis Res Ther* 2009; 11: R125.
31. MELZACK R, TORGERSON WS: On the language of pain. *Anesthesiology* 1971; 34: 50-9.
32. MELZACK R, KATZ J, JEANS ME: The role of compensation in chronic pain: Analysis using a new method of scoring the McGill Pain Questionnaire. *Pain* 1985; 23: 101-12.
33. MELZACK R: The McGill Pain questionnaire: major properties and scoring methods. *Pain* 1985; 1: 85-92.
34. MELZACK R, KATZ J: Pain assessment in adult patients. *In*: *Wall and Melzack's Textbook of Pain*, MCMAHON SB, KOLTZENBURG M, eds. (2006) 5th Edn. London: Elsevier. 291-304.
35. NOLLI M, SALAFFI F, FERRACCIOLI GF, CAMERINI S, ULERI G, UGGERI E: Analysis of variables of Melzack's semantic questionnaire in patients with osteoarthrosis. *Minerva Anestesiol* 1988; 54: 215-22.
36. MELZACK R: The short-form McGill Pain Questionnaire. *Pain* 1987; 30: 191-7.
37. DAUT RL, CLEELAND CS, FLANERY RC: Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. *Pain* 1983; 17197-210.
38. ARNOLD LM, HUDSON JI, WANG F *et al.*: Comparisons of the efficacy of duloxetine for the treatment of fibromyalgia in patients with versus without major depressive disorder. *Clin J Pain* 2009; 25: 461-8.
39. ARNOLD LM, PRITCHETT YL, D'SOUZA DN, KAJDASZ DK, IYENGAR S, WERNICKE JF: Duloxetine for the treatment of fibromyalgia in women: pooled results from two randomized placebo-controlled clinical trials. *J Womens Health (Larchmt)* 2007; 16: 1145-56.
40. TURK DC: Customizing treatment for chronic pain patients: who, what, and why. *Clin J Pain* 1990; 6: 255-70.
41. BERGSTRÖM G, JENSEN IB, BODIN L, LINTON SJ, NYGREN AL, CARLSSON SG: Reliability and factor structure of the Multidimensional Pain Inventory–Swedish Language Version (MPI-S). *Pain* 1998; 75: 101-10.
42. TURK DC, OKIFUJI A, SINCLAIR JD, STARZ TW: Differential responses by psychosocial subgroups of fibromyalgia syndrome patients to an interdisciplinary treatment. *Arthritis Care Res* 1998; 11: 397-404.
43. OKIFUJI A, TURK DC, SINCLAIR JD, STARZ TW, MARCUS DA: A standardized manual tender pointsurvey. I. Development and determination of a threshold point for the identification of positive tender points in fibromyalgia syndrome. *J Rheumatol* 1997; 24: 377-83.
44. HARRIS RE, GRACELY RH, MCLEAN SA *et al.*: Comparison of clinical and evoked pain measures in fibromyalgia. J *Pain* 2006; 7: 521-7.
45. NEUBERGER GB: Measures of fatigue. *Arthritis Care Res* 2003; 49: S175-S183.
46. BELZA BL, HENKE CJ, YELIN EH, EPSTEIN WV, GILLISS CL: Correlates of fatigue in older adults with rheumatoid arthritis. *Nurs Res* 1993; 42: 93-9.
47. WARE JE JR, SHERBOURNE CD: The MOS 36-item short form health survey (SF-36). 1. Conceptual frame-work and item selection. *Med Care* 1992; 30: 473-81.
48. WARE JE, KOSINSKI M, KELLER SD: SF-36 physical and mental health summary scales: a user's manual. Boston: The Health Institute, New England Medical Centre; 1994.
49. HJOLLUND NH, ANDERSEN JH, BECH P: Assessment of fatigue in chronic disease: a bibliographic study of fatigue measurement scales. *Health Qual Life Outcomes* 2007; 5: 12-7.
50. KRUPP LB, LAROCCA NG, MUIR-NASH J, STEINBERG AD: The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989; 46: 1121-3.
51. WYSENBEEK AJ, LEIBOVICI L, WEINBERGER A, GUEDJ D: Fatigue in systemic lupus erythematosus: prevalence and relation to disease expression. *Br J Rheumatol* 1993; 32: 633-5.
52. WHITE PD, GROVER SA, KANGRO HO, THOMAS JM, AMESS J, CLARE AW: The validity and reliability of the fatigue syndrome that follows glandular fever. *Psychol Med* 1995; 25: 917-24.
53. HANN DM, JACOBSEN PB, AZZARELLO P *et al.*: Measurement of fatigue in cancer patients: development and validation of the Fatigue Symptom Inventory. *Qual Life Res* 1998; 7: 301-10.
54. OMDAL R, WATERLOO K, KOLDINGSNES W, HUSBY G, MELLGREN S: Fatigue in patients with systemic lupus erythematosus: the psychosocial aspects. *J Rheumatol* 2003; 30: 283-7.
55. DALTROY LH, ROBB-NICHOLSON C, IVERSEN MD, WRIGHT EA, LIANG MH: Effectiveness of minimally supervised home aerobic training in patients with systemic rheumatic disease. *Br J Rheumatol* 1995; 34: 1064-9.
56. GILES I, ISENBERG D: Fatigue in primary Sjögren's syndrome: is there a link with the fibromyalgia syndrome? *Ann Rheum Dis* 2000; 59: 875-8.
57. ZACHRISSON O, REGLAND B, JAHRESKOG M, KRON M, GOTTFRIES CG: A rating scale for fibromyalgia and chronic fatigue syndrome (the FibroFatigue scale). *J Psychosom Res* 2002; 52: 501-9.
58. TACK BB: Fatigue in rheumatoid arthritis: conditions, strategies, and consequences. *Arthritis Care Res* 1990; 3: 65-70.
59. SMETS EM, GARSSEN B, BONKE B, DE HAES JC: The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* 1995;39: 315-25.
60. YELLEN SB, CELLA DF, WEBSTER K, BLENDOWSKI C, KAPLAN E: Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACIT) measurement system. *J Pain Symptom Management* 1997; 13: 63-74.
61. SARZI-PUTTINI P, ATZENI F, TURIEL M *et al.*: The Italian version of the Fibrofatigue Scale, a reliable tool for the evaluation of fibromyalgia symptoms. J *Psychosom Res* 2004; 56: 213-6.
62. ABLIN JN, ODES L, NEUMANN L, BUSKILA D: The Hebrew version of the FibroFatigue scale: validation of a questionnaire for assessment of fibromyalgia and chronic fatigue syndrome. *Rheumatol Int* 2009 Sep 25. [Epub ahead of print]
63. VARNI JW, BURWINKLE TM, LIMBERS CA, SZER IS: The PedsQL as a patient-reported outcome in children and adolescents with fibromyalgia: an analysis of OMERACT domains. *Health Qual Life Outcomes* 2007 12; 5: 9.
64. SEGAL B, BOWMAN SJ, FOX PC *et al.*: Primary Sjögren's Syndrome: health experiences and predictors of health quality among patients in the United States. *Health Qual Life Outcomes* 2009 27; 7: 46.
65. KEYSTONE E, BURMESTER GR, FURIE R *et al.*: Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. *Arthritis Rheum* 2008 15; 59: 785-93.
66. CASSISI G, SARZI-PUTTINI P, ALCIATI A *et al.*: Italian Fibromyalgia Network. Rheumatism. Symptoms and signs in fibromyalgia syndrome. *Reumatismo* 2008; 60 (Suppl. 1): 15-24.
67. HENRIKSSON CM: Living with continuous muscular pain – patient perspectives. Part II. Strategies for daily life. *Scand J Caring Sci* 1995; 9: 77-86.
68. HALLBERG LRM, CARLSSON SG: Psychosocial vulnerability and maintaining forces related to fibromyalgia. *Scand J Caring Sci* 1998; 12: 95-103.
69. HELLSTROM O, BULLINGTON J, KARLSSON G, LINDQVIST P, MATTSSON B: A phenomenological study of fibromyalgia. Patient perspectives. *Scand J Prim Health Care* 1999; 17: 11-6.
70. HALLBERG LRM, CARLSSON SG: Coping with fibromyalgia. A qualitative study. *Scand J Caring Sci* 2000; 14: 29-36.
71. RAYMOND MC, BROWN JB: Experience of fibromyalgia. Qualitative study. *Can Fam Physician* 2000; 46: 1100-6.
72. BERNARD A, PRINCE A, EDSALL P: Quality of life issues for fibromyalgia patients. *Arthritis Care Res* 2000; 13: 42-50.
73. ASBRING P, NARVANEN AL: Women's experiences of stigma in relation to chronic fatigue syndrome and fibromyalgia. *Qual Health Res* 2002; 12: 148-60.
74. STURGE-JACOBS M: The experience of living with fibromyalgia: confronting an invisible disability. *Res Theor Nurs Prat Int J* 2002; 16: 19-31.
75. CUNNINGHAM MM, JILLINGS C: Individu-

als' descriptions of living with fibromyalgia. *Clin Nurs Res* 2006; 15: 258-73.
76. MARTIN S, CHANDRAN A, ZOGRAFOS L, ZLATEVA G: Evaluation of the impact of fibromyalgia on patients' sleep and the content validity of two sleep scales. *Health Qual Life Outcomes* 2009; 10; 7: 64.
77. MEASE PJ, ARNOLD LM, CROFFORD LJ *et al.*: Identifying the clinical domains of fibromyalgia: Contributions from the clinician and patient Delphi exercises. *Arthritis Rheum* 2008; 59: 952-60.
78. BIGATTI SM, HERNANDEZ AM, CRONAN TA, RAND KL: Sleep disturbances in fibromyalgia syndrome: Relationship to pain and depression. *Arthritis Rheum* 2008; 59: 961-7.
79. RAPHAEL KG, JANAL MN, NAYAK S, SCHWARTZ JE, GALLAGHER RM: Psychiatric comorbidities in a community sample of women with fibromyalgia. *Pain* 2006; 124: 117-25.
80. ZAUTRA AJ, FASMAN R, PARISH BP, DAVIS MC: Daily fatigue in women with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Pain* 2007; 128: 128-35.
81. MARTIN S, CHANDRAN A, LAURIE ZOGRAFOS L, ZLATEVA G: Evaluation of the impact of fibromyalgia on patients' sleep and the content validity of two sleep scales. *Health Quality of Life Outcomes* 2009, 7: 64.
82. LOMELI HAP-O, I, TALERO-GUTIERREZ C, MORENO CB *et al.*: Sleep evaluation scales and questionnaires: a review. *Actas Esp Psiquiatr* 2008; 36: 50-9.
83. COLE JC, DUBOIS D, KOSINSKI M: Use of patient-reported sleep measures in clinical trials of pain treatment: a literature review and synthesis of current sleep measures and a conceptual model of sleep disturbance in pain. *Clin Ther* 2007; 19 (Suppl.): 2580-8.
84. HAYS R, STEWART A: Sleep measures. *In* Measuring Functioning and Well-Being: the Medical Outcomes Study Approach Edited by: STEWART A, WARE J. DURHAM: Duke University Press; 1992:235-59.
85. HAYS RD, MARTIN SA, SESTI AM, SPRITZER KL: Psychometric properties of the Medical Outcomes Study Sleep measure. *Sleep Med* 2005, 6: 41-4.
86. BUYSSE DJ, REYNOLDS CF, MONK TH, BERMAN SR, KUPFER DJ: The Pittsburg sleep quality index: a new instrument for psychiatric practise and research. *Psychiatry Res* 1989; 28: 193-213.
87. HAYTHORNWAITE JA, HEGEL MT, KERNS RD: Development of a sleep diary for chronic pain patients. *J Pain Symptom Management* 1991; 6:65-72.
88. BATIEN CH, VALLIE'RES A, MORIN CM: Validation of the Insomnia Severity Index as an outcome for insomnia research. *Sleep Med Rev* 2001; 2: 297-307.
89. RUSSELL IJ, CROFFORD LJ, LEON T *et al.*: The effects of pregabalin on sleep disturbance symptoms among individuals with fibromyalgia syndrome. *Sleep Med* 2009; 10: 604-10.
90. REJAS J, RIBERA MV, RUIZ M, MASRRAMON X: Psychometric properties of the MOS (Medical Outcomes Study) Sleep Scale in patients with neuropathic pain. *Eur J Pain* 2007, 11: 329-40.
91. WOLFE F, MICHAUD K: Predicting depression in rheumatoid arthritis: the signal importance of pain extent and fatigue, and comorbidity. *Arthritis Rheum* 2009 15; 61: 667-73.
92. PAGANO T, MATSUTANI LA, GONNCALVES FERRIERA EA, MARQUES AP, BRAGANCA FERRIERA CA: Assessment of anxyity and quality of life in fibromyalgia patients. *Sao Paulo Med J* 2004; 122: 252-8.
93. HASSETT AL, GEVIRTZ RN: Nonpharmacologic treatment for fibromyalgia: patient education, cognitive-behavioral therapy, relaxation techniques, and complementary and alternative medicine. *Rheum Dis Clin North Am* 2009; 35: 393-407.
94. ASBERG M, MONTGOMERY SA: A Comprehensive Psychopathological Rating Scale. *Acta Psychiat Scand* 1978; 271: 5-9.
95. MITCHELL AJ, COYNE JC: Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. *Br J Gen Pract* 2007; 57: 144-51.
96. ZUNG WWK: From art to science. *Arch Gen Psychiatry* 1973; 29: 328-37.
97. RADLOFF L: The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977; 1: 385-401.
98. ZIGMOND AS, SNAITH RP: The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983; 67: 361-70.
99. HAMILTON M: The assessment of anxiety states by rating. *Br J Med Psychol* 1959; 32: 50-5.
100. HAMILTON M: Rating depressive patients. *J Clin Psychiatry* 1980; 41: 2124-7.
101. RAMOS BRIEVA JA, CORDERO-VILLAFAFILA A: A new validation of the Hamilton Rating Scale for Depression. *J Psychiatr Res* 1988; 22: 21-8.
102. OTHMER E, DESOUZA C: A screening test for somatization disorder (hysteria). *Am J Psychiatry* 1985; 142: 1146-9.
103. KELLNER R: Abridged manual of the illness attitude scale. New Mexico: Department of Psychiatry, School of Medicine, University of New Mexico, 1987.
104. BECK AT, WARD CH, MENDELSON M, MOCK J, ERBAUGH J: An inventory for measuring depression. *Arch Gen Psychiatry* 1961; 4: 561-71.
105. VODERMAIER A, LINDEN W, SIU C: Screening for Emotional Distress in Cancer Patients: A Systematic Review of Assessment Instruments *J National Cancer Inst* 2009; 101: 1464-88.
106. TERLUIN B, VAN MARWIJK HWJ, ADÈR HJ *et al.*: The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006; 6: 34.
107. DEROGATIS L: SCL-90-R Administration, Scoring and Practice Manual (2000) Baltimore, MD: Clinical Psychiatric Research.
108. DE HAES JC, VAN KNIPPENBERG FC, NEIJT JP: Measuring psychological and physical distress in cancer patients: structure and application of the Rotterdam Symptom Checklist. *Br J Cancer* 1990; 62: 1034-8.
109. BARRETT J, HURST MW, DISCALA C, ROSE RM: Prevalence of depression over a 12-month period in a non-patient population. *Arch Gen Psychiatry* 1978; 35: 741-4.
110. BIGGS JT, WYLIE LT, ZIEGLER VE: Validity of the Zung Self-rating Depression Scale. *Br J Psychiatry* 1978; 32: 381-5.
111. GABRYS JB, PETERS K: Reliability, discriminant and predictive validity of the Zung Self-rating Depression Scale. *Psycholog Reports* 1985; 57: 1091-6.
112. AGRELL B, DEHLIN O: Comparison of six depressive rating scales in geriatric stroke patients. *Stroke* 1989; 20: 1990-4.
113. ERCOLANI M, TROMBINI G, CHATTAT R *et al.*: Fibromyalgic syndrome: depression and abnormal illness behavior. Multicenter investigation. *Psychother Psychosom* 1994; 61: 178-86.
114. GIESECKE T, WILLIAMS DA, HARRIS RE *et al.*: Sub-grouping of fibromyalgia patients on the basis of pressure-pain thresholds and psychological factors. *Arthritis Rheum* 2003; 48: 2916-22.
115. HERRMANN C: International experiences with the Hospital Anxiety and Depression Scale. A review of validation data and clinical results. *J Psychosom Res* 1997; 42: 17-41.
116. BJELLAND I, DAHL AA, HAUG TT, NECKELMANN D: The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* 2002; 52: 69-77.
117. BECK AT, RIAL WY, RICKELS K: Short form of depression inventory: cross-validation. *Psychol Rep* 1974; 34: 1184-6.
118. TCHEN N, SOUBEYRAN P, EGHBALI H *et al.*: Quality of life in patients with aggressive non Hodgkin's lymphoma. Validation of the medical outcomes study short form 20 and the Rotterdam Symptom Checklist in older patients. *Crit Rev Oncol Hematol* 2002; 43: 219-26.
119. PACI E: Assessment of validity and clinical application of an Italian version of the Rotterdam Symptom Checklist. *Qual Life Res* 1992; 1: 129-34.
120. WATSON M, LAW M, MAGUIRE GP *et al.*: Further development of a quality of life measure for cancer patients: The Rotterdam Symptom Checklist (revised). *Psychooncology* 1992; 1: 35-44.
121. TESTA MA, SIMONSON DC: Assessment of quality of life outcomes. *N Engl J Med* 1996; 28: 835-40.
122. ARNOLD LM, CROFFORD LJ, MEASE PJ *et al.*:Patient perspectives on the impact of fibromyalgia. *Patient Educ Couns* 2008; 73: 114-20.
123. MCHORNEY CA: Health status assessment methods for adults: Past accomplishments and future challenges. *Ann Rev Public Health* 1999; 20: 309-35.
124. VELIKOVA G, BOOTH L, SMITH AB *et al.*: Measuring quality of life in routine oncology practice improves communication and patient well-being: A randomized controlled trial. *J Clin Oncology* 2004; 22: 714-24.
125. SALAFFI F, GIROLIMETTI R, SARZI PUTTINI P, ATZENI F, GASPARINI S, GRASSI W: Health-related quality of life in fibromyal-

gia: a comparison with rheumatoid arthritis and healthy controls by using SF-36 health survey. *Clin Exp Rheumatol* (in press).
126. SALAFFI F, CAROTTI M, STEFANIA G, GASPARINI S, GRASSI W: The health-related quality of life in rheumatoid arthritis, ankylosing spondylitis, and psoriatic arthritis: a comparison with a selected sample of healthy people. *Health Quality of Life Outcomes* 2009; 7-25.
127. MAS AJ, CARMONA L, VALVERDE M, RIBAS B; EPISER STUDY GROUP: Prevalence and impact of fibromyalgia on function and quality of life in individuals from the general population: results from a nationwide study in Spain. *Clin Exp Rheumatol* 2008; 26: 519-26.
128. DWORKIN RH, TURK DC, FARRAR JT *et al.*: Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005; 113: 9-19.
129. MCGRATH PJ, WALCO GA, TURK DC *et al.*: Core outcome domains and measures for pediatric acute and chronic/recurrent pain clinical trials: PedIMMPACT recommendations. *J Pain* 2008; 9: 771-83.
130. FRANCHIGNONI F, SALAFFI F: Generic and specific measures for outcome assessment in orthopaedic and rheumatologic rehabilitation. *In*: *Assessment in Physical Medicine and Rehabilitation - Advances in Rehabilitation*. BARAT M, FRANCHIGNONI F Eds. 2004; 16: 45-77.
131. STEINER GL, NORMAN DR: Health measurement scales: a practical guide to their development and use, 2nd edn. Oxford: Oxford University Press 1996.
132. TORRANCE G: Utility approach to measuring health related quality of life. *J Chronic Dis* 1987; 40: 593-600.
133. GUYATT GH, FEENY DH, PATRICK DL: Measuring health-related quality of life. *Ann Intern Med* 1993; 118: 622-9.
134. GUYATT GH, VAN ZANTEN SJOV, FEENY DH, PATRICK DI: Measurement quality of life in clinical trials: a taxonomy and review. *Can Med Assoc* 1989; 140: 1441-8.
135. BERGNER M, BOBBITT RA, POLLARD WE *et al.*: The Sickness Impact Profile: validation of the health status measure. *Med Care* 1976; 14: 56-67.
136. BERGNER N, BOBBITT RA, CARTER WB, GILSON BS: The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981; 19: 787-805.
137. HUNT SM, MCKENNA SP, MCEWEN J: The Nottingham Health Profile: subjective health status and medical consultations. *Soc Sci Med* 1981; 15: 221-9.
138. HUNT SM, MCEVEN J, MCKENNA SP: Measuring health status: a new tool for clinicians and epidemiologists. *J Roy Coll Gen Pract* 1985; 35: 185-8.
139. HOUSSIEN DA, MCKENNA SP, SCOTT DL: The Nottingham Health Profile as a measure of disease activity and outcome in rheumatoid arthritis. *Br J Rheumatol* 1992; 36: 69-73.
140. HOFFMAN DL, DUKES EM: The health status burden of people with fibromyalgia: a review of studies that assessed health status with the SF-36 or the SF-12. *Int J Clin Pract* 2008; 1: 115-26.
141. WARE JE: A 12-item short form health survey (SF-12): construction of scales and preliminary tests of reliability and validity. *Med Care* 1996; 32: 220-33.
142. BENNETT KJ, TORRANCE GW: Measuring health state preferences and utilities: rating scales, time trade-off and standard gamble techniques. *In* SPILKER B (Ed): *Quality of Life and Pharmacoeconomics in Clinical Trials,* 2nd ed. Philadelphia: Lippincott-Raven 1996: 253-65.
143. FROBERG DG, KANE RL: Methodology for measuring health-state preferences – II: scaling methods. *J Clin Epidemiol* 1989; 42: 459-71.
144. WOLFE F, HAWLEY DJ: Measurement of the quality of life in rheumatic disorders using the EuroQol. *Br J Rheumatol* 1997; 36: 786-93.
145. HURST NP, KIND P, RUTA D, HUNTER M, STUBBINGS A: Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EUROQOL (EQ- 5D). *Br J Rheumatol* 1997; 36: 551-9.
146. BURKHARDT CS, CLARK SR, BENNETT RM: The Fibromyalgia Impact Questionnaire: development and validation. *J Rheumatol* 1991; 18: 728-33.
147. SARZI-PUTTINI P, ATZENI F, FIORINI T *et al.*: Validation of an Italian version of the Fibromyalgia Impact Questionnaire (FIQ-I). *Clin Exp Rheumatol* 2003; 21: 459-64.
148. BENNETT RM, FRIEND R, JONES KD, WARD R, HAN BK, ROSS RL: The Revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties. *Arthritis Res Ther* 2009; 11: 415.
149. MEENAN RF, MASON JH, ANDERSON JJ, GUCCIONE AA, KAZIS LE: AIMS2: The content and properties of a revised and expanded Arthritis Impact Measurement Scales health status questionnaire. *Arthritis Rheum* 1992; 35: 1-10.
150. FRIES JF, SPITZ PW, KRAINES RG, HOLMAN HR: Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23: 137-45.
151. BENNETT R: The Fibromyalgia Impact Questionnaire (FIQ): a review of its development, current version, operating characteristics and uses. *Clin Exp Rheumatol* 2005; 23 (Suppl. 39): S154-S162.
152. SALAFFI F, PIVA S, BARRECA C *et al.*: Validation of an Italian version of the Arthritis Impact Measurement Scales 2 (ITALIAN-AIMS2) for patients with osteoarthritis of the knee. Gonarthrosis and Quality of Life Assessment (GOQOLA) Study Group. *Rheumatology* 2000; 39: 720-7.
153. SALAFFI F, VACCARO CM, MANACORDA T, PARDINI L, MONTECUCCO CM: Sociodemographic aspects and therapeutic pathways in patients with rheumatoid arthritis in Italy. *Reumatismo* 2009; 61: 197-211.
154. WOLFE F: Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): analyses in 2491 rheumatoid arthritis patients following leflunomide initiation. *J Rheumatol* 2001; 28: 982-9.
155. DEWALT DA, REED GW, PINCUS T: Further clues to recognition of patients with fibromyalgia from a simple 2-page patient Multidimensional Health Assessment Questionnaire (MDHAQ). *Clin Exp Rheumatol* 2004; 22: 453-61.
156. WOLFE F, RASKER JJ: The Symptom Intensity Scale, fibromyalgia, and the meaning of fibromyalgia-like symptoms. *J Rheumatol* 2006; 33: 2291-9.
157. VAN DER HEIJDE DMFM, VAN'T HOF MA, VAN RIEL PLCM *et al.*: Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990; 49: 916-20.
158. PREVOO ML, VAN'T HOF MA, KUPER HH *et al.*: Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995; 38: 44-8.
159. SALAFFI F, STANCATI A, CAROTTI M: Responsiveness of health status measures and utility-based methods in patients with rheumatoid arthritis. *Clin Rheumatol* 2002; 21: 478-87.
160. SALAFFI F, STANCATI A, NERI R, GRASSI W, BOMBARDIERI S: Measuring functional disability in early rheumatoid arthritis: the validity, reliability and responsiveness of the Recent-Onset Arthritis Disability (ROAD) index. *Clin Exp Rheumatol* 2005; 23 (Suppl. 39): S31-42.
161. GUYATT GH, KIRSHNER B, JAESCHKE R: Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol* 1992; 45: 1341-5.
162. STUCKI G, DALTROY L, KATZ JN, JOHANNESSON M, LIANG MH: Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. *J Clin Epidemiol* 1996; 49: 711-7.
163. ANDRESEN EM: Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil* 2000; 81 (Suppl. 2): S15-20.
164. GUILLEMIN F, BOMBARDIER C, BEATON D: Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993; 46: 1417-32.