

Fibromyalgia Syndrome Module at OMERACT 9: Domain Construct

PHILIP MEASE, LESLEY M. ARNOLD, ERNEST H. CHOY, DANIEL J. CLAUW, LESLIE J. CROFFORD, JENNIFER M. GLASS, SUSAN A. MARTIN, JESSICA MOREA, LEE SIMON, C. VIBEKE STRAND, DAVID A. WILLIAMS, on behalf of the OMERACT Fibromyalgia Working Group

ABSTRACT. The objective of the module was to (1) establish a core domain set for fibromyalgia (FM) assessment in clinical trials and practice, (2) review outcome measure performance characteristics, (3) discuss development of a responder index for assessment of FM in clinical trials, (4) review objective markers, (5) review the domain of cognitive dysfunction, and (6) establish a research agenda for outcomes research. Presentations at the module included: (1) Results of univariate and multivariate analysis of 10 FM clinical trials of 4 drugs, mapping key domains identified in previous patient focus group: Delphi exercises and a clinician/researcher Delphi exercise, and breakout discussions to vote on possible essential domains and reliable measures; (2) Updates regarding outcome measure status; (3) Update on objective markers to measure FM disease state; and (4) Review of the issue of cognitive dysfunction (dyscognition) in FM. Consensus was reached as follows: (1) Greater than 70% of OMERACT participants agreed that pain, tenderness, fatigue, patient global, multidimensional function and sleep disturbance domains should be measured in all FM clinical trials; dyscognition and depression should be measured in some trials; and stiffness, anxiety, functional imaging, and cerebrospinal fluid biomarkers were identified as domains of research interest. (2) FM domain outcome measures have generally proven to be reliable, discriminative, and feasible. More sophisticated and comprehensive measures are in development, as is a responder index for FM. (3) Increasing numbers of objective markers are being developed for FM assessment. (4) Cognitive dysfunction assessment by self-assessed and applied outcome measures is being developed. In conclusion, a multi-dimensional symptom core set is proposed for evaluation of FM in clinical trials. Research on improved measures of single domains and composite measures is ongoing. (J Rheumatol 2009;36:2318–29; doi:10.3899/jrheum.090367)

Key Indexing Terms:

FIBROMYALGIA OMERACT OUTCOME MEASURES CLINICAL TRIALS

Fibromyalgia (FM), also known as fibromyalgia syndrome, is characterized by chronic widespread pain and tenderness on physical examination, as defined by the 1990 American

College of Rheumatology (ACR) criterion¹. The FM criteria have been beneficial in identifying a more homogeneous group of individuals with chronic widespread pain upon

From Seattle Rheumatology Associates and Division of Rheumatology Research, Swedish Medical Center, Clinical Professor of Rheumatology, University of Washington, Seattle, WA; Department of Psychiatry, University of Cincinnati College of Medicine, Cincinnati, OH, USA; Sir Alfred Baring Garrod Clinical Trials Unit, Academic Department Rheumatology, King's College London, London, UK; University of Michigan, Ann Arbor, MI; Division of Rheumatology and Women's Health, University of Kentucky, Lexington, KY; Research Center for Group Dynamics and Department of Psychiatry, Division of Substance Abuse, University of Michigan; RTI-Health Solutions, Ann Arbor, MI, USA; Oregon Health and Science University, Portland, OR; Harvard Medical School and Beth Israel Deaconess Medical Center, Boston, MA; Division of Immunology/Rheumatology, Stanford University, Portola Valley, CA, USA.

L. Arnold, D. Clauw, L. Crofford, P. Mease, D.A. Williams were supported in part by Grant Number AR053207 from NIAMS/NIH. E. Choy was supported in part by an Integrated Clinical and Academic Centre grant from the Arthritis Research Campaign, UK. D.A. Williams was supported in part by Grant Number U01AR55069 from NIAMS/NIH. L. Arnold and L. Crofford were supported in part by Grant Number AR053207 from NIAMS/NIH.

P.J. Mease, MD, Seattle Rheumatology Associates, Chief, Division of

Rheumatology Research, Swedish Medical Center, Clinical Professor of Rheumatology, University of Washington; L.M. Arnold, MD, Department of Psychiatry, University of Cincinnati College of Medicine; E.H. Choy, MD, Director, Sir Alfred Baring Garrod Clinical Trials Unit, Academic Department Rheumatology, King's College London; D. Clauw, MD, Professor of Medicine and Psychiatry, University of Michigan; L. Crofford, MD, Gloria W. Singletary Professor of Internal Medicine, Chief, Division of Rheumatology & Women's Health, University of Kentucky; J.M. Glass, PhD, Assistant Research Scientist, Research Center for Group Dynamics, Research Assistant Professor, Department of Psychiatry, Division of Substance Abuse, University of Michigan; S.A. Martin, RTI-Health Solutions; J. Morea, MS, Oregon Health and Science University; L. Simon, MD, Associate Professor of Medicine, Harvard Medical School, Director of Rheumatology Clinical Research, Beth Israel Deaconess Medical Center; C.V. Strand, MD, Clinical Professor of Medicine, Adjunct, Division of Immunology/Rheumatology, Stanford University; D.A. Williams, PhD, Professor, Anaesthesiology, Medicine, Psychiatry, and Psychology, University of Michigan, Ann Arbor, Michigan, USA. E-mail: daveawms@med.umich.edu
Address correspondence to Dr. P.J. Mease, Seattle Rheumatology Associates, 1101 Madison St., Ste. 1000, Seattle, WA 98104. E-mail: pmease@philipmease.com.

which to conduct research aimed at better understanding FM. Currently, separate clinical diagnostic criteria for FM do not exist. Applying the ACR criteria in clinical practice may overemphasize the importance of tenderness (e.g., over sampling for women), the importance of peripheral as opposed to central factors, and distress (e.g., distress raises tenderness). Clinically, patients with FM often complain of other symptoms beyond pain. Additional symptoms include: fatigue, sleep disturbance, mood disturbance, cognitive dysfunction, and syndromes such as irritable bowel and bladder syndrome, and various forms of headache². Each patient with FM experiences a number of different symptoms to varying degrees, which may change over time and with treatment, thus constituting the need for continual assessment of the multidimensional nature of the condition. FM may occur on its own and also has been noted to be comorbid with rheumatoid arthritis, lupus, other chronic pain conditions, hypothyroidism, and infections such as Lyme disease or hepatitis C. FM is prevalent in at least 2% of the population, occurring more frequently in females than in males¹. Current research posits that FM results from disordered central pain and sensory processing. Disregulation of several neuropeptide and neurohormone networks have been identified, leading to a deficiency in pain inhibitory pathways and/or increase in facilitatory networks^{3,4}. The triggering and maintenance of FM appears to require both genetic disposition and environmental influences such as emotional or physical stressors or illness⁵.

Until the 1990s, there had been a paucity of well-controlled clinical trials of pharmacotherapy of FM. This was partly due to a lack of classification criteria and partly related to a poor understanding about pathophysiology, uncertainty about what core symptom domains could be reliably measured, a lack of objective markers of disease activity and severity, suboptimal confidence that measures could discriminate a therapeutic response, and perhaps a certain skepticism among some that the condition was legitimate. Stemming from the work of Moldofsky and Smythe on sleep disorders in FM⁶, studies with tricyclic antidepressants (TCA) were conducted and showed short-term benefit for pain and sleep in FM⁷. However, it was apparent that these agents were incomplete in their effectiveness and poorly tolerated.

In parallel with increased understanding of the neuropathophysiology of FM, the more specifically targeted and better tolerated pharmaceutical agents of potential benefit for FM symptomatology became available. Examples of agents include serotonin and norepinephrine reuptake inhibitors (SNRI), which can augment the activity of these serotonin and norepinephrine, α 2- δ subunit modulators and that inhibit excitatory neuropeptides such as glutamate and substance P, and other neuromodulators as a means of diminishing pain and fatigue, improving sleep, and beneficially affecting other symptom domains of FM. Controlled

trials of several of these agents have been conducted utilizing a variety of measures and have demonstrated clinically meaningful improvements in pain, patient global impression of change, and function as compared to placebo. Two agents, pregabalin, an α 2- δ modulator, and duloxetine, an SNRI, have been approved for management of FM in the United States; and a third agent, milnacipran, an SNRI, has also recently been approved⁸⁻¹². However the approval process has seen a wide variety of outcome measures used, and approval has primarily been based on demonstration of efficacy in domains of pain, patient global impression of change, and total impact of FM as measured by the Fibromyalgia Impact Questionnaire (FIQ)¹³. There is a need for scientific validation of a core set of domains that more fully constitute FM syndrome for use in clinical trials. Performance characteristics of domain measures also need to be evaluated to assure clinicians, regulators, and the public about the soundness of our ability to evaluate therapies in FM and to provide guidance to developers of new therapies.

To this end, a group of clinician/researchers interested in FM gathered in 2004 to develop a workshop for Outcome Measures in Rheumatology Clinical Trials (OMERACT). The group included both academic and pharmaceutical-based researchers and focused on several areas. To gain a preliminary sense of the key domains needing to be assessed in FM, 23 clinician/researchers participated in a Delphi exercise based on a list of domains developed by the expert steering committee of the working group. Results of the exercise and of voting held at a workshop at OMERACT 7 are shown in Table 1¹⁴.

To better understand performance characteristics of measures of the key domains, a review of controlled clinical trials was conducted to determine effect sizes of the measures¹⁵. Pain and patient global measures appeared reliable and showed good effect sizes, but other domains such as

Table 1. Top 12 symptom domains ranked by median Delphi scores, and percentage of OMERACT 7 workshop attendees who voted domains as essential to assess in clinical trials of FM¹⁴.

Domain	Median Delphi Score*	OMERACT 7 Respondents (%)
Pain	16	100
Patient global	10	94
Fatigue	10	85
HRQOL	5	76
Sleep quality	8	70
Depression	5	65
Treatment side effects	5	58
Physical functioning	5	42
Clinician rated global	1	23
Dyscognition	2	21
Anxiety diagnosis	2	21
Tender point intensity	2	18

* Delphi exercise conducted before OMERACT 7 to prioritize domains.

sleep, fatigue, and function did not, raising questions about either the effectiveness of therapy for these domains or the quality of the measures. The group also reviewed more objective measures being explored in FM, e.g., functional magnetic resonance imaging (fMRI), and potential linkages with patient reported outcomes developed by the World Health Organization International Classification of Function (ICF) project and the Patient-Reported Outcomes Measurement Information System (PROMIS) network¹⁵. These projects represented broader and more in-depth attempts to characterize the full patient experience of disease, function, and health-related quality of life (HRQOL) impact of FM. The Delphi exercise concurred that the research agenda should continue each of these areas of work and explore in greater depth the patient perspective on outcomes domains relevant to FM.

In preparation for a second FM workshop in 2006 at OMERACT 8, the expanding working group, with the aid of MAPI Values, an independent research organization, conducted a series of patient focus groups to map the array of symptoms experienced and problems caused by FM¹⁶. Utilizing information gained in these discussions, a Delphi exercise was conducted among patients¹⁷. The key symptom areas identified as impacting the majority of the representative patients, although worded differently, showed considerable conceptual overlap with those identified in the clinician Delphi exercise, thus providing face validity to the 2 different exercises. In addition, an updated review of the performance of outcome measures used in more recent clinical trials, objective measure data¹⁸, and linkage work with the ICF and PROMIS FM-extension projects (D Williams) was reported. The research agenda included the need to determine which key domains, as identified by patients and clinician/researchers, represented the full core set of domains experienced in FM, and whether areas of domains overlap; also a preliminary analysis was necessary to develop a responder index for FM. Two FM OMERACT Steering Committee members (L. Arnold and L. Crofford) are co-primary investigators on a project funded by the US National Institutes of Health (NIH) to develop such a responder index, which was outlined at this meeting. The core domain construct work, which was subsequently completed and which is based on the research agenda formulated at OMERACT 8, is outlined below and by Choy, *et al* elsewhere in these proceedings¹⁹. In addition, a more complete understanding of the symptom domain of cognitive dysfunction (dyscognition) and appropriate measures for it were identified as a key subject for the research agenda. This history provides the foundation for the current report on proceedings of the FM Workshop presented at OMERACT 9.

Objectives. We sought to (1) establish a core domain set for the assessment of fibromyalgia (FM) in clinical trials and practice, (2) review the performance characteristics of outcome measures, including patient reported outcomes, cur-

rently being used to assess FM domains, (3) discuss development of a responder index for the assessment of FM in clinical trials, (4) review objective markers of FM, (5) review the domain of cognitive dysfunction in FM and its potential assessment in clinical trials and practice, and (6) establish a research agenda for further work to be done regarding FM outcomes research.

Module process. Since the OMERACT 8 meeting in May 2004, working group members met in regular teleconferences and in person at the ACR, EULAR and Myopain Society meetings. The working group, noted above, was constituted of clinicians/researchers, statisticians, pharmaceutical industry representatives, and patients from North America, Europe, and Australia. There were 4 subgroups (leaders): (1) Domain construct (Ernest Choy, Philip Mease, Lesley Arnold, Dan Clauw, Jennifer Glass, Susan Martin, David Williams), (2) Outcome measures/patient reported outcomes (PRO)/Responder index (David Williams, Susan Martin, Lesley Arnold), (3) Objective markers (Dan Clauw, Leslie Crofford, Jessica Morea), and (4) Cognition (Jennifer Glass). Liaison to the OMERACT executive committee was conducted by Lee Simon and Vibeke Strand. The group's fellow was Jessica Morea. Patient participants were Lynne Matallana, Kathy Longley, Michael Peterman, and Sharon Waldrop.

Methods and results by module subgroup. Domain construct. As noted above, the working group had previously conducted a clinician/researcher Delphi exercise and patient focus groups and Delphi to determine key domains considered important to assess in FM clinical trials (Tables 1 and 2). We analyzed FM trial data to determine how well these domains approximate the totality of the FM experience for a patient (content validity) and to what degree domains were overlapping versus independent. Patient global impression of change (PGIC) was used as a surrogate of overall improvement and was the dependent variable in multivariate regression analyses against which other domains were regressed. Outcome measures used in trials were mapped onto one or more of the following domains identified in the clinician/researcher Delphi: pain, patient global, fatigue, HRQOL, multidimensional function, sleep, depression, physical function, tenderness, dyscognition, anxiety, as well as stiffness, which had additionally been identified in the patient Delphi. Ten studies involving 4 pharmacological agents were analyzed: 2 serotonin and norepinephrine reuptake inhibitors (i.e., duloxetine and milnacipran), one α 2- δ modulator (i.e., pregabalin), and sodium oxybate (gamma hydroxybutyrate) — all of which have shown efficacy in FM clinical trials. Details of this study are summarized elsewhere in these proceedings¹⁹.

Univariate analysis showed that instruments that measured these various domains showed moderate to high correlation with PGIC; associations were highest with pain, fatigue, multidimensional function, physical function, and

Table 2. Comparison of OMERACT 7 voting and patient Delphi: key FM domains¹⁵.

OMERACT 7 Voting [†]		Patient Delphi [‡]		
Domain	%	Domain/item	Mean	%
Pain	100	Pain or physical discomfort	6.9	95
Patient global	94	Joints aching or pain	5.7	90
Fatigue	85	Lack of energy or fatigue	5.5	96
HRQOL	76	Impact on sleep [e.g., difficulty falling asleep, staying asleep, or getting up in the morning]	5.3	92
Multidimensional function	75	Problems with attention or concentration [e.g., difficulty concentrating on things, difficulty thinking, “fibro-fog”]	4.7	91
Sleep	70	Stiffness	4.2	91
Depression	65	Disorganized thinking [e.g., difficulty expressing yourself, difficulty answering questions quickly, or difficulty making plans]	3.6	85
Treatment side effects	58	Difficulty moving, walking, or exercising	3.5	86
Physical function	42	Having to push yourself to do things	3.1	83
Clinical global	23	Impact on ability to make plans, accomplish goals, or complete tasks	3.0	79
Tender point intensity	21	Feeling tender where touched	3.0	77
Dyscognition	21	Depression (e.g., disappointed, sad, resigned, or unmotivated)	3.0	74
Anxiety	21	Impacted/limited in doing normal daily life and household activities	2.8	82
		Memory problems	2.6	81

[†] FM domains ranked as most important in clinician-investigator Delphi exercise performed prior to OMERACT 7. Percent column shows percentage of OMERACT 7 attendees who agreed these domains were essential to assess in FM clinical trials. [‡] Mean scores [points assigned out of 100 possible] for the top 14 domains identified by patients as important in FM. Percentage column reflects percentage of patients who felt domains should be assessed. HRQOL: health-related quality of life.

stiffness; and only moderate with depression, anxiety, and dyscognition. It should be recalled that in a majority of these trials, patients with major depressive disorder had been excluded, resulting in a lower effect size of change scores since baseline depression scores were low. In addition, only one trial utilized a measure of self-assessed cognition, partly because of uncertainty about how best to approach and assess this domain.

Multivariate analysis showed moderate to high values of R², with studies having more non-overlapping domains demonstrating higher values, suggesting that if key domains are not assessed, the variance accounted for in PGIC will be diminished. Pain, fatigue, physical function, multidimensional function, and depression were retained as separate domains in trials of all 4 compounds. Tenderness was retained as a domain separate from pain in all 3 trials in which it was assessed, suggesting that it is a sign of allodynia and/or hyperalgesia separate from the subjective impression of pain. Sleep was retained in 2 out of 3 clinical trial groups; stiffness, assessed by a single question in the FIQ, in 2 out of 4; and dyscognition in none, the latter presumably related to either non- or insufficient assessment.

The domain construct was discussed in breakout sessions, taking into account clinician and patient Delphi exercises and data analysis, and as aided by the presence of patient participants. Voting, by audience response methodology, on the construct was done on 2 occasions: at the time of the module and in the plenary recap at the end of the meeting, when further clarification on key issues was offered.

Inner core set (domains to be assessed in all clinical trials of FM). There was little debate about whether core issues

Table 3. OMERACT 9 FM Module domain construct voting to determine which domains were essential to assess in all FM clinical trials (core set), or in some FM trials or electively as a research item. Endorsement by at least 70% of attendees is desirable for inclusion of a domain as a core set measure.

Domain	Core Set %	Some Trials — Research %
Pain	98	2
Fatigue	84	16
Multidimensional function	76	24
Patient global	73	27
Tenderness	61 (72) [†]	39
Sleep disturbance	58 (77) [†]	42
Cognitive dysfunction	39	61
Depression	35	65
Stiffness	25	75
Adverse events	23	77
Anxiety	18	82
CFS Biomarkers	4	96
Functional imaging	4	96

[†] In the initial voting on core domains, more than 50% endorsed tenderness (n voting = 102) and sleep disturbance (n voting = 107) as inner core set domains. The OMERACT process provides the opportunity for consensus building in an iterative manner throughout the course of the meeting. A final vote (n voting = 121) on sleep and tenderness, after further presentation and discussion, resulted in both domains reaching the OMERACT requirement for inner core set inclusion.

such as pain, fatigue, and patient global should be measured in all FM trials as relevant domains for the “inner” core set (Table 3). However, there was considerable discussion about other domains. One issue was the separation and overlap of the concepts of multidimensional function, physical func-

tion, and quality of life. The 2 principal instruments currently used for measuring these domains are the Medical Outcome Study Short Form Survey-36 (SF-36) and FIQ, which include both function and HRQOL questions. Work is under way to develop more sophisticated instruments that more comprehensively measure these domains through the PROMIS and PROMIS FM-extension project, and/or linkage with the ICF methodology. Since the current measures are primarily considered to be optimal instruments to assess the concept of multidimensional function, it was voted (63%), until more optimal HRQOL instruments are available, to subsume these concepts under the phrase “multidimensional function,” which was voted to be a core domain item, keeping open the possibility of separating out HRQOL as more sensitive and specific instruments are developed. Tenderness separated from pain in the multivariate analysis and was considered by more than 60% in initial voting and 70% in revised voting to be in the core set as an essential domain to measure in all trials. Sleep disturbance has long been considered an important part of the FM experience, and was so endorsed in the clinician and patient Delphi exercises. However, in the data analysis, it did not correlate highly with PGIC and was somewhat insensitive to change. More careful analysis of the instruments used to assess sleep demonstrated that some subscales performed well and others, e.g., “snoring” in the Medical Outcomes Study (MOS) sleep scale, did not. Thus, the poor correlation with

PGIC could have been due to dilution of quality of the scale by assessments that were irrelevant to FM patients. It was agreed that there should be a focus on development and testing of more relevant measurements of sleep in FM and use of more sensitive subscales of existing measures. Thus, with further discussion, it was voted to include sleep disturbance in the inner core (Table 3).

Outer core set (domains to be assessed in some but not all FM trials – second circle). Some domains were shown to be core domains in FM by the multivariate analysis but not considered by the majority of OMERACT attendees to be necessary to assess in all clinical trials in a development program. Depression was retained in the multivariate analysis as a core domain in FM and was voted, by 65%, that it should be assessed in some trials; but only 35% felt it should be assessed in all trials. Thus, in Figure 1 depression is listed in the second circle. Cognitive dysfunction, or dyscognition, was noted to be an important domain by patients versus less important as rated by clinicians/researchers in previous Delphi exercises. However, full understanding of depression as a domain and how best to assess it in FM trials is still uncertain and is an active research issue. Given its importance as a domain, 38% felt it should be in the core set and 45% thought that it should be measured in some trials. Thus, dyscognition was moved to the second circle (Figure 1).

Research agenda (domains that may or may not be included in FM trials – outer circle). Several domains were high-

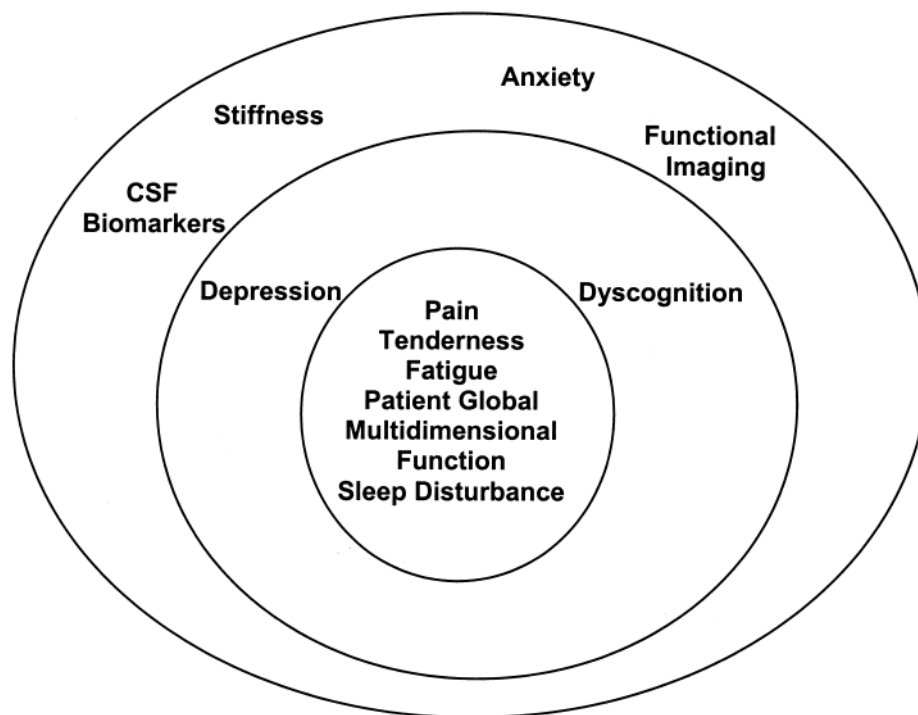


Figure 1. Hierarchy of domains for fibromyalgia. The inner circle includes the core set of domains to be assessed in all clinical trials of FM. The second concentric circle includes the outer core set of domains to be assessed in some but not all FM trials. The outermost circle includes the domains on the research agenda that may or may not be included in FM trials.

lighted in discussions as being of potential interest to further explore; these are listed in the third circle. Stiffness has been identified by patients as an important symptom domain. In multivariate analysis it did not separate out in all trials as a domain distinct from pain, although it was only assessed with a single question in the FIQ. Thus, it is part of the research agenda (outer circle). Functional imaging and cerebrospinal fluid biomarkers are examples of potential objective markers that may be important and discriminative, although not currently feasible for all trials. These were, therefore, listed in the research agenda. Because anxiety was considered to be an essential part of the core set by just 18%, it was placed in the outer circle.

In previous FM workshops, adverse events (AE) were listed as an important domain to assess in trials. Since AE are naturally assessed in clinical trials, it was felt to be unnecessary to list as a symptom in the core set.

Outcome measures/PRO. The outcomes measurement (OM) committee within the FM Working Group of OMERACT works to identify PRO that best assess the domains of most relevance to individuals with FM. The work of this group is informed by ongoing initiatives either within or outside OMERACT, and at OMERACT 9 this group presented data in the following areas: (1) Building evidence supporting the valid use of existing PRO specifically for FM; (2) Developing responder indices based upon existing PRO; (3) Further refinement of the domain definitions of relevance specifically for FM; (4) Developing new and next generation outcomes measures specific to FM; and (5) Integrating the guidance of regulatory bodies to the work of improved outcomes measurement in FM.

Studies supporting the valid use of existing PRO in FM. Many of the outcomes measures currently used in FM research were developed and validated for use with other medical conditions. Thus many indices used to assess domains of relevance were “adopted” from other conditions. Adopting instruments is neither uncommon nor inappropriate when exploring a relatively new and poorly understood condition such as FM. For example, a research definition for FM did not exist prior to 1990 and until the recent work within OMERACT, there was no consensus regarding the clinical domains of relevance to this condition. A lack of basic foundations in the understanding of this condition, not to mention insufficient interest, time and funding, precluded developing assessment instruments more specific to FM. Borrowing and adopting assessment instruments has facilitated basic exploration of the nature and impact of FM, and represents a methodological advance over previous unstandardized methods of inquiry.

As interest and understanding of FM matures, the need for greater rigor in assessment methods also advances. It is plausible to suspect that “adopted” instruments are suitable for use in FM; however, support of this suitability needs to be based upon performance within individuals with FM

rather than upon assumptions of equivalence. Several studies are currently underway examining the performance characteristics of validated instruments in other conditions for use in studies with FM. An example of one such effort is the ongoing work of Strand and colleagues on use of the SF-36 in FM. Importantly, as with other rheumatic diseases, the SF-36 represents a generic measure of health-related quality of life (HRQOL) that meets the OMERACT filter in rheumatoid arthritis, osteoarthritis, systemic lupus erythematosus, and FM, and may be well suited for use with other disease specific instruments, once developed.

The SF-36 is a brief, well established, self-administered patient questionnaire for assessment of HRQOL that can also be viewed as a measure of multidimensional function, including “participation”²⁰. The SF-36 measures 8 domains of health status: physical functioning, role limitations because of physical problems, bodily pain, general health perceptions, energy/vitality, social functioning, role limitations due to emotional problems, and mental health. A summary score for physical functional status (physical component score, PCS) can be calculated by combining and weighting the various individual scales²¹. Individual or group domain and summary scores can be compared to national norms for the US and other populations, or contrasted for various medical conditions²².

To date, the SF-36 has been used in over 70 studies involving individuals with FM; including randomized controlled trials (RCT) of tramadol, gabapentin, pregabalin, duloxetine, and milnacipran. The domains of coverage within the SF-36 map nicely with the domains identified as being relevant in the aforementioned Delphi studies. Domain scores have been consistently observed to improve in studies where active treatment arms can be compared to placebo; supporting the SF-36 as being responsive to change in individuals with FM when change is expected to occur. To date, there is much evidence supporting the use of the SF-36 as an index of multidimensional function as it satisfies the OMERACT filter for FM. Of interest, data from both RCT and longitudinal observational studies demonstrate remarkably similar decrements in baseline domain and PCS and mental component summary scores compared with age/gender and population matched normative data. Trials of gabapentin, pregabalin, duloxetine, and milnacipran have demonstrated treatment associated mean improvements in summary and domain scores that are remarkably similar and well exceed minimum clinically important differences (MCID)^{9,23-26}.

Responder indices using combined domains. Responder indices have become popular for identifying treatment successes in illnesses where improvement needs to occur across multiple domains. Such responder indices have a history of use in FM. However, consistent with the work on relevant domains, there has not been consensus regarding composition of these indices. For example, Simms, *et al*^{27,28} reported on the use of an index requiring improvement on 4 out of

6 criteria defined as 50% improvement in pain, sleep fatigue, patient global and physician global, dolorimeter improvements and improved myalgic score. These criteria were later used in RCTs of amitriptyline²⁹. This initial response index for FM was a first attempt beyond assessment of pain and tenderness in clinical trials. However, the Simms criteria were not as sensitive as would be desired in part because physical function was not included. A second attempt at a responder index for FM was the work of Dunkl *et al*³⁰, requiring improvements in 3 of 4 measures including FIQ, pain intensity, tender point count, and pain intensity.

Clinical trials of new compounds for FM have also used responder indices as primary efficacy endpoints. For example, in RCTs of milnacipran, a responder index required participants to report $\geq 30\%$ improvement in pain intensity, a patient global change of “moderately improved or much improved” and ≥ 6 points improvement in SF-36 PCS score¹². Clinical trials of sodium oxybate used a different responder index: $\geq 20\%$ improvement in pain intensity, $\geq 20\%$ improvement in FIQ, and a patient global assessment of “much better or very much better”³¹.

Thus, to date, most responder indices have been rationally derived, based upon what investigators or regulatory bodies deemed to represent improvement in the context of a clinical trial involving FM. Given consensus regarding relevant domains is only just evolving, most responder indices have not benefited from a data driven development process. Arnold, Crofford, and colleagues are currently working on the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) sponsored project to develop a responder index for FM based on both consensus and empirical data for eventual use in FM RCT. This project begins with the consensually derived domains for FM, links existing assessment instruments to each domain, evaluates each measure for 5 types of validity in FM, and evaluates the performance of each instrument as a member of a composite index. The project also establishes consensus among clinicians regarding criteria for improvement in FM, tests the consensually derived criteria with empirical data, and identifies which definitions of improvement result in fewest placebo improvements. This project is ongoing and will inform the efforts of the working group in subsequent sessions of OMERACT.

Refinements in domain definition studies (Item-level refinements). Identification of consensually derived domains of relevance for FM is an important first step in gaining a better understanding of what needs to be assessed in FM. However, studies that attempt to validate adopted measures for use in FM must rely on several assumptions. First, instruments purporting to measure a given domain (e.g., fatigue) will in fact measure those facets of fatigue that are relevant to individuals with FM. Second the domain names (e.g., fatigue) have shared meaning for individuals with FM, clinicians, and other medical populations from which exist-

ing measures may have been adopted. Early investigations into this area of inquiry suggest that neither assumption holds completely.

Identifying concepts contained within existing domain measures. Perhaps the largest body of work in this area comes from the investigators associated with The International Classification of Functioning, Disability, and Health project (ICF) within the World Health Organization (ICF-WHO). The ICF developed a domain categorization coding system that identifies the relevant domains of functional status assessment for medical illnesses in general³². This large system can be broken down into core sets for specific illnesses. Currently the core set closest to FM is the “chronic widespread pain core set (CWP).” CWP affects between 5 and 15% of the population and includes FM as an extreme subset³³⁻³⁶. When used, this coding system helps to identify relevant domains of functional limitations for different diseases/conditions and then provides a code (much like the ICD10) that identifies the area of functioning affected by the condition.

Various standardized instruments used to assess domains in FM have been examined and items within each instrument have been mapped to specific categories (subdomains) within each broader domain (e.g., fatigue can be subcategorized into physical, mental, motivation, etc.). One recent study found that out of 42 RCT in FM, 27 different questionnaires were used to assess FM. From the 27 different questionnaires, 1138 distinct health-related concepts could be identified based upon items. These concepts were linked to 113 ICF categories. Each questionnaire differed greatly from the other with regard to the specific subdomain categories covered and the relative importance paid to the broader ICF domains of body structure, body function, activities/participation, and environmental factors. The least well covered broad domain for all existing questionnaires was environmental factors³⁷.

A second manuscript explored differences in the ICF categories that were represented in PRO commonly used in FM research that purportedly assess the same construct. This manuscript applied ICF linkages to common indices of pain, fatigue, sleep function, and affect. In each case the domains were indexed by assessment tools that varied substantially, depending upon which assessment tool was chosen. Thus, quite disparate conclusions might be found for a given construct, based on which assessment instrument was used and which specific facets of the construct the instrument and its scales emphasize³⁸.

That different instruments emphasize different facets of constructs is not always limiting. As we learn more about how patients with FM define and think about the various domains of relevance, we will be better able to match our assessment instruments to the way individuals with FM use these terms resulting in improved assessment ability with increased sensitivity in our measures of outcomes. That different instruments assess different facets of domains is also

reason not to limit by decree which assessment tools must be used for FM, as the choice of instrument might be best driven by which facets of the domain a given intervention hopes to address.

Identifying FM-specific definitions to the domains of relevance. Efforts to learn about how patients with FM think about the domains of relevance are currently in progress. Methodologies typically start with a consensually derived generic definition of the domain (e.g., ICF definitions or definitions from the NIH Roadmap PROMIS project) that are then agreed to or modified by focus groups of individuals with FM.

One such study recently presented at an NIH PROMIS conference found generally good agreement among patients with FM with generic definitions of pain, fatigue, negative mood, and physical functioning. For each domain, however, insufficient depth of impact was expressed as a concern of the definition. For example, individuals with FM reported that most existing definitions of fatigue focused on simply being tired and failed to capture the profound unrecoverable and disabling exhaustion that accompanies FM³⁹.

Development of new outcomes measures specific to FM. Perhaps the largest scale project aimed at developing new highly sensitive FM-specific measures for the domains of relevance to FM is the NIH/NIAMS sponsored project "FM-Specific extension of the PROMIS network." PROMIS is an NIH Roadmap initiative that is building a next generation Patient-Reported Outcomes Measurement Information System (PROMIS). In the development of PROMIS, each domain is defined generally, and then patient reported outcome measures are developed and linked to those specific domain definitions. PROMIS, still under development, is to be a publicly available user-friendly computerized adaptive testing (CAT) system that for efficient generic measurement of PRO outcomes (PRO) across a wide range of chronic diseases and dimensions³⁹. Although costly and time-consuming to develop and maintain, a national public resource of this nature will be of benefit as the system will be able to assess multiple domains using fewer items (i.e., less patient burden) with greater precision (i.e., increased power for clinical trials with fewer subjects).

PROMIS was established for the general assessment of chronic illnesses, and as might be expected, many of the domains identified in PROMIS are of relevance to FM, such as pain, fatigue, negative mood, and physical function. Several domains identified in the OMERACT Delphi exercises, however, were not included in the first iteration of PROMIS, such as sleep disturbance, dyscognition, stiffness, and tenderness. Williams and colleagues are currently participating in a cooperative agreement with the US NIH/NIAMS to develop a FM-specific extension of PROMIS. The goals of this initiative include: (1) Determining whether PROMIS definitions of pain, fatigue,

physical function, and negative mood hold up or require modification for patients with FM; (2) Developing new definitions for sleep disturbance, dyscognition, stiffness, and tenderness for FM, (3) Developing new item banks for new domains and/or supplementing existing banks with FM-specific items; (4) Performing large-scale field testing following the methods of the larger PROMIS initiative thus facilitating the development of FM-specific calibrations for existing and new item banks, and (5) Developing static short forms and CAT assessments specific to domains of relevance for FM. These new item banks and calibrations will be merged within the context of the larger PROMIS roadmap initiative.

Regulatory considerations concerning the use of PRO for FM. Many of the domain assessment tools currently in use for FM were developed in academia to explore and gain a better understanding of FM. With broader interest and new treatments for FM, researchers and PRO developers must become aware of not only methods of test development but also guidelines of regulatory bodies such as the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) if the assessment device is to be used in a clinical trial for eventual product approval. One such regulatory body, the FDA, released valuable draft regulatory guidance for (1) the use of existing measurement tools, (2) the development of new measurement tools, and (3) the transition of tools from one medium to another (e.g., paper to electronic formats)⁴⁰. Of particular importance in the draft guidance is the documentation of patient input during the PRO instrument development process, both in the identification of the domains of importance in any particular disease area as well as at the item-level development and evaluation. Understanding the current PRO instrument requirements from the perspective of regulatory bodies, can guide decisions related to choice of currently available instruments versus development of new instruments.

OBJECTIVE MARKERS

Participants at OMERACT 9 were presented an update on the current understanding of the underlying pathophysiology of FM and the biomarkers that relate to these pathophysiological processes. Researchers and clinicians now view FM as a common pain syndrome characterized by primarily central, non-nociceptive pain; as well, a variety of aberrant pain and sensory processing pathways have been identified that can lead to pain or sensory amplification.

All potential biomarkers that have been identified to date in FM are related in some way to this central amplification. Current biomarkers under study include but are not limited to experimental or evoked pain testing (EPT); MRI imaging (sometimes during EPT); and levels of neurotransmitters in cerebrospinal fluid, including substance P, glutamate, serotonin, and norepinephrine; muscle biopsy; polysomnography; cytokines; and sensory testing.

The objective biomarker breakout session focused on 3 main issues: (1) the “objectivity” of biomarkers, (2) whether a marker belongs in the core domain of outcomes that must be measured in a clinical trial or requires further study before becoming a core domain, and (3) application of the OMERACT filter of truth, discrimination and feasibility to specific biomarkers.

Neurotransmitters and muscle biopsy were the only markers designated as totally objective, but no single marker was designated as a core domain. When applying the OMERACT filter, some markers were considered more useful in research than in clinical practice. For example, polysomnography was considered truthful and discriminating but might only be feasible in a clinical trial where the investigational intervention aims to improve sleep, and may not be feasible in clinical practice. In the case of biomarker neural imaging, some participants rated it 7 out of 10 in terms of truth but rated it low on feasibility due to cost, and evidence was considered insufficient to assign a score on the discrimination scale.

The goal of OMERACT is to place a set of disease markers through a filter of truth, feasibility, and discrimination to achieve a succinct and practical set of outcomes to measure change in health status. With numerous available markers measured in so many different ways, it is impossible to compare the efficacy of potential treatments. In preparation for OMERACT, and during the workshop, it became evident that there were too many biomarkers with too little evidence to support existence of a core set that would pass through the OMERACT filter. As such, the biomarker research agenda focused on a single class of biomarker that has the most support for feasibility, truth, and discrimination: EPT.

Encompassing multiple techniques, including tender point intensity, pressure pain thresholds, and heat/cold thresholds EPT is emerging as a promising evidence-based biomarker. The goal is to quantify the experience of pain objectively and to demonstrate that FM is related to aberrations in central, rather than peripheral, pain processing. The presence of hyperalgesia (increased pain in response to normally painful stimuli) and allodynia (pain in response to non-painful stimuli) implicate central pain mechanisms and are measured by EPT.

Research shows that some methods of EPT are correlated with reports of clinical pain in patients with FM. For example, Geisser, *et al* found that dolorimetry and pressure thresholds were associated with clinical pain, but heat stimuli were not⁴¹. Particularly, the use of the multiple random staircase (MRS) method for delivering pressure stimuli has been shown to be associated with patients’ reports of clinical pain. MRS uses an interactive software system to determine low, medium, and high pain intensity thresholds for each subject based on their response to random stimuli. Harris, *et al* compared MRS to other evoked pain measures and found that it was the only “objective” technique that tracked with improvement during the course of treatment⁴².

Such findings may indicate that experimental pain testing, and MRS specifically, correspond to a patient’s clinical condition, rendering this type of testing a potential biomarker of disease status, progression, and improvement. In addition, MRS is not subject to bias in terms of variation between clinicians, or to fluctuations within individual clinicians, as with tender point counts, and is not associated with patient distress^{43,44}. With both dolorimetry and tender point count, the patient is aware of when the stimulation is forthcoming⁴⁵, and such techniques have been shown to be influenced by patient distress⁴⁶. EPT in FM yields a measure of objective pain that correlates with clinical pain, is less subject to bias, underscores the central pain mechanisms in FM, and is less invasive than other biomarkers (e.g., collecting cerebrospinal fluid; CSF).

COGNITION

Although pain and fatigue are hallmark symptoms of FM, many patients find that problems with cognitive function (dyscognition) are just as troublesome^{5,15,47,48}. A small but growing body of literature supports the presence of dyscognition in FM⁴⁹. In this section the current state of knowledge about dyscognition is reviewed.

Measurement of dyscognition can be divided into 2 categories: self-report of cognitive difficulties, and performance-based measures of cognition; most reports are performance-based⁴⁹. About one dozen studies have been published that use either standardized neuropsychological tests or non-standardized but common measures from cognitive science. Although these studies have used a variety of measures, a pattern has begun to emerge where deficits are seen in 4 separate cognitive systems. Most notably, problems with verbal working memory have been consistently reported. Working memory refers to a memory system that combines short-term storage (on the order of seconds) with other mental operations such as retrieving knowledge from semantic memory and deleting or adding items. Working memory is an important construct in cognition as it functions as basic skill. Results from 4 different measures of working memory, the Paced Auditory Serial Attention Test (PASAT)⁵⁰⁻⁵², the Reading Span Test⁵³, the Everyday Test of Attention⁵⁴, and Consonant Trigrams⁵² have all found impairment on this crucial cognitive system.

Related to working memory are attention and executive control. Attention is the ability to maintain focus on a specific item, task, or location. Executive control involves the many processes used to maintain focus, such as ignoring irrelevant items, suppressing responses not consistent with a goal, and planning. The results from the PASAT and the Test of Everyday Attention point to a problem with executive control of attention in FM. Ongoing work indicates greater memory impairment in FM patients when they have distraction^{52,55,56}. An important point is that most standardized neuropsychological tests are conducted without distraction.

Deficits are also seen in memory systems with longer duration. Episodic memory refers to the ability to remember a specific episode. Many of our memory tasks fall into this category, such as remembering a list of items to buy at the grocery store. Patients with FM perform more poorly than controls on word list tasks⁵³ as well as standardized tests of memory^{51,52,57,58}.

The final area where deficits have been reported is in semantic memory, particularly the ability to access semantic memory. Semantic memory refers to our knowledge of facts. It is separate from episodic memory (e.g., you may remember that there are 12 inches in a foot, but not remember when you learned this fact). Patients anecdotally report word finding problems, and there are reports of decreased performance on both verbal fluency tasks^{53,57} and on vocabulary tests⁵³.

There are now a number of computerized neuropsychological batteries (e.g., CANTAB, COGSTATE). Computerized batteries would help the ease of testing, data collection, and interpretation across clinical trials and other studies. To date, there is only one report that used a computerized battery, the Automated Neuropsychological Assessment Metrics⁵⁹. Unfortunately, this battery did not yield any differences between patients with FM and controls, perhaps due to the lack of distraction and working memory tests. Future work will be needed to assess the utility of other computerized neuropsychological batteries in FM research.

Self-report of cognitive function is an important addition to performance-based measures because it can be influenced by many factors, including effort required for performance, stress regarding performance, and depression. There is a surprising paucity of studies using self-report instruments of dyscognition in FM, although several studies include 1 or 2 items about memory or concentration. An exception is a study of memory beliefs in FM with the Metamemory in Adulthood Questionnaire, used frequently for studying memory in older adults⁵⁵. FM patients reported lower memory capacity, more memory deterioration, low self-efficacy over memory, higher anxiety about memory performance, and more strategy use to support memory than in age and education matched controls. Among FM patients, performance on a memory task was correlated with perceived memory capacity. Further work using other well-validated, self-report measures of cognitive function would be very helpful in clinical trials, since self-report measures are easy to administer and fulfill the need for patient reported outcomes.

To summarize, existing data support dyscognition as a salient symptom, and objective cognitive impairments can be demonstrated in patients with FM. This will be important in future clinical trials, but the field is not yet at the point where we can recommend outcome measures that should be included in all trials. In addition, during breakout discussions, 3 important areas that have not been well studied were identified: (1) There was considerable concern about how

other conditions (e.g., depression, anxiety, fatigue, and medications) could influence dyscognition; (2) Some aspects of dyscognition described by patients have not been well studied, in particular the idea of mental exhaustion and feelings of dissociation; and (3) There was a good deal of discussion about the frequent lack of correspondence between objective cognitive testing and self-report of dyscognition. The group noted that self-report may also include other noncognitive aspects. For example, someone with cognitive losses compared to pre-illness state may still perform well when their pre-illness state was above average.

CONCLUSION

FM is a condition characterized by chronic widespread pain, excessive tenderness, and a number of associated symptoms such as fatigue, sleep disturbance, mood disorder, and cognitive dysfunction with associated impairment of function and HRQOL. The symptom complex is caused by dysregulation of central sensory processing systems. Evidence points to genetic, environmental, and concomitant disease state factors in its etiology. As therapies are developed that not only address pain, but also other symptom domains, clinicians, regulatory agencies, patients, and others need to know the relative contribution of these various domains to the disease experience of the patient and how best to measure them in a reliable and feasible manner in clinical trials.

The primary objective of the OMERACT 9 FM module to achieve relative consensus on a domain construct for FM clinical trials was accomplished through: (1) Review of work presented in previous OMERACT workshops (clinician/researcher Delphi, patient focus group and Delphi exercises); (2) Presentation of a study in which the key clinical domains identified in these exercises were mapped against the patient global impression of change noted in 10 FM pharmacologic studies to determine the degree to which key domains both constituted the global patient experience of FM and were not completely overlapping; (3) Presentation of the current status of outcome measures, objective biomarkers, and understanding about disease state; (4) Discussion of the above in breakout groups; and (5) A voting process. Figure 1 demonstrates the outcome of this process. Domains considered essential to measure in all FM clinical trials include pain, tenderness, fatigue, patient global, multidimensional function, and sleep disturbance. Domains considered important to measure at some point in a clinical development program, but not essential to measure in all clinical trials, are depression and cognitive dysfunction, also known as dyscognition. Domains that are of research interest and considered elective to measure at this time, include stiffness, anxiety, and objective markers such as functional imaging, e.g., fMRI, and cerebrospinal fluid biomarkers. It is well recognized that this domain construct is a "work in progress." For example, it is recognized that there are important elements of HRQOL that are not neces-

sarily subsumed under the concept of “multidimensional function,” yet the best instruments that we currently have available, the SF-36 and FIQ, to measure these domains are primarily measures of function. Further, whereas both clinical experience and emerging research suggest that cognitive dysfunction is an important clinical domain in FM, optimal assessment methods are still in development; thus, whether ultimately cognitive dysfunction will be considered a more essential domain to measure in all trials is uncertain. As new and more sophisticated instruments become available to more completely measure the totality of patient experience vis-à-vis these domains, and as we gain a more full understanding of the disease process and better ways to measure impact of therapeutic intervention, this framework is expected to evolve.

As in previous OMERACT meetings, an update was provided on the outcome measures used in FM trials and the current status of objective markers of FM disease state. The quality of their performance was discussed, and areas needing improvement were reviewed, particularly assessment of sleep, mood disturbance, tenderness, stiffness, multidimensional function and HRQOL.

Since the majority of outcome measures are PRO, it is important that they fulfill the standards of evidence being developed by regulatory agencies. The working group reported on several projects underway, which will be more fully reviewed in future OMERACT meetings as part of the group’s research agenda: linkages with existing disease assessment networks such as PROMIS and the ICF and the development of an FM responder index. These will be developed in the context of the OMERACT filter of truth (forms of validation), discrimination, and feasibility. Objective markers of FM disease state continue to be developed, such as cerebral spinal fluid biomarkers and functional imaging. The relationship of these markers to disease state and their ability to reflect change in response to treatment remains on the research agenda.

Special focus was placed on the domain of cognitive dysfunction in FM during the module. This domain is ranked highly by patients in terms of disease impact, and understanding about this problem is emerging. There have been fledgling attempts to measure change of this domain via self-assessment questionnaires. There are a number of more objective and potentially feasible applied measures, e.g., computer based cognition assessment methods, which are beginning to be studied in FM clinical trials and will be reviewed at future OMERACT meetings.

ACKNOWLEDGEMENT

We thank the support from Nooshine Dayani, Qu Peng, and Robert Palmer from Forest Laboratories Inc, Chinglin Lai, Yanping Zheng and Diane Guinta from Jazz Pharmaceuticals Inc., Daniel Kajdasz and Amy Chappell from Eli Lilly & Co., and Gergana Zlateva and Emir Birol from Pfizer Inc. *OMERACT FM working group*: Robert Allen, Dennis Ang, Lesley Arnold, Annelise Boonen, Daniel Buskila, Larry Bradley, Alarcos Cieza, Ernest Choy, Dan Clauw, Leslie Crofford, Brian Cuffel, Michael Gauthier,

Michael Gendreau, Jennifer Glass, Don Goldenberg, Richard Gracely, Diane Guinta, Kim Jones, Chinglin Lai, Geoff Littlejohn, Yves Mainguy, Susan Martin, Lynne Matallana, Philip Mease, Jamal Mikdashi, Jessica Morea, Robert Palmer, Daniel Radecki, I. Jon Russell, Stuart Silverman, Lee Simon, Michael Spaeth, Tanja Stamm, Raj Tummala, Olivier Vitton, Brian Walitt, David Williams, Madelaine Wohlriech, and Gergana Zlateva.

REFERENCES

1. 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.
2. Mease P. Fibromyalgia syndrome: Review of clinical presentation, pathogenesis, outcome measures, and treatment. *J Rheumatol* 2005;32:6-21.
3. Staud R. The neurobiology of chronic musculoskeletal pain (including chronic regional pain). In: Wallace DJ, Clauw DJ, editors. *Fibromyalgia and other central pain syndromes*. Philadelphia: Lippincott Williams & Wilkins;2005:45-62.
4. Clauw DJ. Fibromyalgia: Update on mechanisms and management. *J Clin Rheum* 2007;13:102-9.
5. Bennett RM, Jones J, Turk DC, Russell IJ, Matallana L. An internet survey of 2,596 people with fibromyalgia. *BMC Musculoskel Disorders* 2007;9:27-38.
6. Moldofsky H, Scarisbrick P, England R, Smythe HA. Musculoskeletal symptoms and non-REM sleep disturbance in patients with “fibrositis syndrome” and healthy subjects. *Psychosom Med* 1975;37:341-5.
7. Arnold LM, Keck Jr PE, Welge JA. Antidepressant treatment of fibromyalgia. A meta-analysis and review. *Psychosomatics* 2000;41:104-13.
8. Mease PJ, Russell IJ, Arnold LM, et al. A randomized, double-blind, placebo-controlled, phase III trial of pregabalin in the treatment of patients with fibromyalgia. *J Rheumatol* 2008;35:502-14.
9. Arnold LM, Russell IJ, Diri EW, et al. A 14-week, randomised, double-blind, placebo-controlled monotherapy trial of pregabalin for pain in patients with fibromyalgia. *J Pain* 2008;9:792-805.
10. Russell IJ, Mease PJ, Smith TR, et al. Efficacy and safety of duloxetine for treatment of fibromyalgia in patients with or without major depressive disorder: Results from a 6-month, randomised, double-blind, placebo-controlled, fixed-dose trial. *Pain* 2008;136:432-44.
11. Mease PJ, Clauw DJ, Gendreau RM, Rao SG, Kranzler J, Chen W, Palmer RH. The efficacy and safety of milnacipran for treatment of fibromyalgia. a randomized, double-blind, placebo-controlled trial. *J Rheumatol* 2009;36:398-409.
12. Clauw DJ, Mease P, Palmer RH, Gendreau RM, Wang Y. Milnacipran for the treatment of fibromyalgia in adults: a 15-week, multicenter, randomized, double-blind, placebo-controlled, multiple-dose clinical trial. *Clin Ther* 2008;30:1988-2004.
13. Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. *J Rheumatol* 1991;18:728-33.
14. Mease PJ, Clauw DJ, Arnold LM, et al. Fibromyalgia syndrome. *J Rheumatol* 2005;32:2270-7.
15. Mease P, Arnold LM, Bennett R, et al. Fibromyalgia syndrome. *J Rheumatol* 2007;34:1415-52.
16. Arnold LM, Crofford LJ, Mease PJ, et al. Patient perspectives on the impact of fibromyalgia. *Patient Edu Couns* 2008, Epub.
17. Mease PJ, Arnold LM, Crofford LJ, et al. Identifying the clinical domains of fibromyalgia: Contributions from clinical and patients delphi exercises. *Arthritis Care Res* 2008;59:952-60.
18. Dadabhoy D, Crofford LJ, Spaeth M, Russell IJ, Clauw DJ. Evidence-based biomarkers for fibromyalgia syndrome. *Arthritis*

- Res Ther 2008;10:211.
19. Choy EH, Arnold LM, Clauw D, et al. Content and criterion validity of the preliminary core dataset for clinical trials in fibromyalgia syndrome. *J Rheumatol* 2009;36:2335-42.
 20. Ware J, Kosinski M, Dewey J. How to score version 2 of the SF-36 health survey standard and acute forms. Lincoln: Quality Metric; 2000.
 21. Ware JE, Kosinski M. Interpreting SF-36 summary health measures: a response. *Qual Life Res* 2001;10:405-13.
 22. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220-33.
 23. 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;62:115-26.
 24. 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 randomised, placebo-controlled clinical trials. *J Womens Health* 2007;16:1145-56.
 25. Arnold LM, Goldenberg D, Stanford SB, Lalonde JK, Sandhu HS, Keck PE. Gabapentin in the treatment of fibromyalgia: a randomised, double-blind, placebo-controlled, multicenter trial. *Arthritis Rheum* 2007;56:1336-44.
 26. Strand V, Williams DA, Gendreau M, Palmer RH. Improvement in health related quality of life and multidimensional function with milnacipran treatment in fibromyalgia. 2008 (in progress).
 27. Simms RW, Felson DT, Goldenberg DL. Development of preliminary criteria for response to treatment in fibromyalgia syndrome. *J Rheumatol* 1991;18:1558-63.
 28. Simms RW, Felson DT, Goldenberg DL. Criteria for response to treatment in fibromyalgia [abstract]. *Arthritis Rheum* 1998;31:S100.
 29. Carette S, Bell MJ, Reynolds WJ, et al. Comparison of amitriptyline, cyclobenzaprine, and placebo in the treatment of fibromyalgia. A randomised, double blind clinical trial. *Arthritis Rheum* 1994;37:32-40.
 30. Dunkl PR, Taylor AG, McConnell GG, Alfano AP, Conaway MR. Responsiveness of fibromyalgia clinical trial outcome measures. *J Rheumatol* 2000;27:2683-91.
 31. Russell IJ, Perkins AT, Michalek JE. Sodium oxybate relieves pain and improves function in fibromyalgia syndrome: a randomized, double-blind, placebo-controlled, multicenter clinical trial. *J Rheumatol* 2009;60:299-309.
 32. World Health Organization. International classification of functioning, disability, and health. Geneva:WHO; 2001.
 33. 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.
 34. Croft P, Rigby AS, Boswell R, Schollum J, Silman A. The prevalence of chronic widespread pain in the general population. *J Rheumatol* 1993;20:710-713.
 35. Coster L, Kendall S, Gerdle B, Henriksson C, Henriksson KG, Bengtsson A. Chronic widespread musculoskeletal pain — a comparison of those who meet criteria for fibromyalgia and those who do not. *Eur J Pain* 2008;12:600-10.
 36. Offenbaecher M, Cieza A, Brockow T, et al. Are the contents of treatment outcomes in fibromyalgia trials represented in the International Classification of Functioning, Disability and Health? *Clin J Pain* 2007;23:691-701.
 37. Proding B, Cieza A, Williams DA, et al. Measuring health in patients with fibromyalgia: content comparison of questionnaires based on the International Classification of Functioning, Disability and Health. *Arthritis Rheum* 2008;59:650-8.
 38. Williams DA. A fibromyalgia specific extension of the PROMIS network. 2008. Presented at the Second PROMIS Conference: Improving measurement of patient reported outcomes — new tools and the science behind them. Bethesda, Maryland, 2008.
 39. Fries J, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. *Clin Exp Rheumatol* 2005;23:S53-7.
 40. US Department of Health and Human Services. Food and Drug Administration. CDER2002193, docket no. 2006D-0044. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims draft guidance. *Health Qual Life Outcomes* 2006;4:79.
 41. Geisser ME, Gracely RH, Giesecke T, Petzke F, Williams DA, Clauw DJ. The association between experimental and clinical pain measures among persons with fibromyalgia and chronic fatigue syndrome. *Eur J Pain* 2007;11:202-7.
 42. Harris RE, Gracely RH, McLean SA, et al. Comparison of clinical and evoked pain measures in fibromyalgia. *J Pain* 2006;7:521-7.
 43. Petzke F, Clauw DJ, Ambrose K, Khine A, Gracely RH. Increased pain sensitivity in fibromyalgia: effects of stimulus type and mode of presentation. *Pain* 2003;105:403-13.
 44. Petzke F, Gracely RH, Park KM, Ambrose K, Clauw DJ. What do tender points measure? Influence of distress on 4 measures of tenderness. *J Rheumatol* 2003;30:567-74.
 45. Gracely RH, Grant MA, Giesecke T. Evoked pain measures in fibromyalgia. *Best Pract Res Clin Rheumatol* 2003;17:593-609.
 46. Wolfe F. The relation between tender points and fibromyalgia symptom variables: evidence that fibromyalgia is not a discrete disorder in the clinic. *Ann Rheum Dis* 1997;56:268-71.
 47. Katz RS, Heard AR, Mills M, Leavitt F. The prevalence and clinical impact of reported cognitive difficulties (Fibrofog) in patients with rheumatic diseases with and without fibromyalgia. *J Clin Rheumatol* 2004;10:53-8.
 48. 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.
 49. Glass JM. Cognitive dysfunction in fibromyalgia and chronic fatigue syndrome: new trends and future directions. *Curr Rheumatol Rep* 2006;8:425-9.
 50. Sletvold H, Stiles TC, Landro NI. Information processing in primary fibromyalgia, major depression and healthy controls. *J Rheumatol* 1995; 22:137-42.
 51. Grace GM, Nielsen WR, Hopkins M, Berg MA. Concentration and memory deficits in patients with fibromyalgia syndrome. *J Clin Exp Neuropsychol* 1999;21:477-87.
 52. Leavitt F, Katz RS. Distraction as a key determinant of impaired memory in patients with fibromyalgia. *J Rheumatol* 2006;33:127-32.
 53. Park DC, Glass JM, Minear M, Crofford LJ. Cognitive function in fibromyalgia patients. *Arthritis Rheum* 2001;44:2125-33.
 54. Dick B, Eccleston C, Crombez G. Attentional functioning in fibromyalgia, rheumatoid arthritis, and musculoskeletal pain patients. *Arthritis Rheum* 2002;47:639-44.
 55. Glass JM, Park DC, Minear M, Crofford LJ. Memory beliefs and function in fibromyalgia patients. *J Psychosom Res* 2005;58:263-9.
 56. Cote KA, Moldofsky H. Sleep, daytime symptoms, and cognitive performance in patients with fibromyalgia. *J Rheumatol* 1997;24:2014-23.
 57. Landro NI, Stiles TC, Sletvold H. Memory functioning in patients with primary fibromyalgia and major depression and healthy controls. *J Psychosom Res* 1997;42:297-306.
 58. Grisart J, van der Linden M, Masquelier E. Controlled processes and automaticity in memory functioning in fibromyalgia patients: relation with emotional distress and hypervigilance. *J Clin Exp Neuropsychol* 2002;24:994-1009.
 59. Walitt B, Roebuck-Spencer T, Bleiberg J, Weinstein A. Automated neuropsychiatric measurements of information processing in fibromyalgia. *Rheumatol Int* 2008;28:561-6.