Sleep Medicine 10 (2009) 766–770

Contents lists available at ScienceDirect

Sleep Medicine

journal homepage: www.elsevier.com/locate/sleep

Original Article

Measurement properties of the Medical Outcomes Study Sleep Scale in patients with fibromyalgia

Joseph C. Cappelleri [a], Andrew G. Bushmakin [a], Anne M. McDermott [b], Ellen Dukes [c], Alesia Sadosky [c,*], Charles D. Petrie [a], Susan Martin [d]

[a] Pfizer Inc., Global Research and Development, 50 Pequot Avenue, New London, CT 06320, UK
[b] Outcomes Research Consultant, 13104 Riviera Terrace, Silver Spring, MD 20904, USA
[c] Pfizer Inc., Global Outcomes Research, 235 East 42nd Street, New York, NY 10017, USA
[d] Pfizer Inc., Global Outcomes Research, 2800 Plymouth Road, Ann Arbor, MI 48105, USA

ARTICLE INFO

Article history:
Received 3 April 2008
Received in revised form 16 July 2008
Accepted 24 September 2008
Available online 29 January 2009

Keywords:
Fibromyalgia
Sleep problems
Insomnia
Medical Outcomes Study
MOS Sleep Scale

ABSTRACT

Objective: Sleep problems are a common symptom of fibromyalgia (FM). The objective of this study was to evaluate the Medical Outcomes Study (MOS) Sleep Scale as a measure of FM-related sleep problems.
Methods: Analyses were based on data from the 1056 and 1077 studies, two randomized, double-blind, placebo-controlled trials of pregabalin for adults with FM. MOS Sleep Scale scores of study patients were compared with United States normative scores using a one-sample *Z* test. Subscale structure of the MOS was evaluated by confirmatory factor analyses, internal consistency was evaluated using Cronbach's alpha reliability coefficients. Estimated clinically important differences (CID) in MOS Sleep Disturbance subscale scores were evaluated using mixed-effects models of change in subscale scores as a function of the Patient Global Impression of Change (PGIC).
Results: 1056 and 1077 included 748 and 745 patients, respectively. Most patients were female (1056: 94.4%, 1077: 94.5%) and white (1056: 90.2%, 1077: 91.0%). Mean ages were 48.8 years (1056) and 50.1 years (1077). Baseline MOS Sleep Scale scores were statistically ($P < 0.001$) and substantially poorer than general population values. The MOS subscale structure was confirmed in both studies at each assessment except at baseline in the 1056 study. Cronbach's alpha coefficients were acceptable, at least 0.70, for all multi-item scales at baseline and end-of-study assessments in both studies, with the exception of the Sleep Adequacy subscale at baseline. The estimated CID for the MOS Sleep Disturbance subscale was 7.9.
Conclusions: The MOS Sleep Scale is an appropriate measure of FM-related sleep problems. These analyses provide the foundation for further use and evaluation of the MOS Sleep Scale in FM patients.

1. Introduction

The American College of Rheumatology (ACR) defines fibromyalgia syndrome (FM) using two criteria: (1) chronic widespread pain and (2) pain upon digital palpation in at least 11 of 18 defined tender point sites [1,2]. In addition to musculoskeletal pain, FM is associated with symptoms such as fatigue, sleep disturbance, and morning stiffness [3]. Further, patients report they experience depression and cognitive difficulties, and that FM interferes with their usual activities [4]. Assessing the effectiveness of new therapies therefore requires accurate assessment of a multi-dimensional array of symptoms and problems.

This paper focuses on the measurement of sleep problems in patients with FM. Disturbed sleep is consistently ranked by patients as a highly bothersome symptom of the disease. For example, Bennett et al. [5] conducted a web-based survey of over 2000 patients with FM and found that the self-reported intensity of non-restorative sleep was ranked third, behind symptoms of morning stiffness and fatigue. Mease et al. [4] conducted a patient symptom ranking and found that sleep ranked high among the most important symptoms (following pain, aching joints or pain, and lack of energy or fatigue).

To evaluate FM-related sleep problems and the impact of treatment on these symptoms, the clinical program for pregabalin (Lyrica®) included a well-established standardized (generic) patient survey—the Medical Outcomes Study (MOS) Sleep Scale [6]—as a secondary efficacy endpoint. The MOS Sleep Scale includes 12 questions about initiation and maintenance of sleep, respiratory problems during sleep, amount of sleep, perceived adequacy of sleep, and daytime somnolence [6]. It was developed and initially tested in a large sample of individuals with chronic illnesses. It has been further validated in a nationally representative sample of adults in the United States (US) [7]

* Corresponding author. Tel.: +1 212 733 9491; fax: +1 646 441 4757.
E-mail address: alesia.sadosky@pfizer.com (A. Sadosky).

1389-9457/$ - see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.sleep.2008.09.004

and in other pain populations including postherpetic neuropathic pain [7], neuropathic pain of broad etiology [8], and rheumatoid arthritis [9]. The objective of this study was to evaluate the measurement properties of the MOS Sleep Scale in a specific population of patients with FM.

2. Methods

2.1. Studies and subjects

This paper reports the psychometric properties of the MOS Sleep Scale using data from two clinical trials of pregabalin conducted in the US: 1077 [10] and 1056 [11]. The study designs for these trials have been described elsewhere [10,11]. The studies were randomized, double-blind, and placebo-controlled clinical trials of three doses of pregabalin (300, 450, and 600 mg/day). Patients were 18 years of age or older with FM as defined by the ACR criteria [1,2].

During the baseline phase, study patients had to have an average daily diary pain score of at least 4 (within the last 7 days) on a numeric rating scale ranging from 0 (“no pain”) to 10 (“worst possible pain”). Further, study patients had to have a score of at least 40 mm on the 100 mm visual analogue scale (VAS) of the Short-Form McGill Pain Questionnaire [12] at the screening and baseline (randomization) study visits. In the 1077 study, patients with a 30% or greater reduction in the VAS from the screening to the randomization study visits (a single-blind placebo run-in period) were discontinued; this criterion in the 1077 study was intended to exclude potential placebo responders. In both studies, the primary efficacy measure was endpoint mean pain score as defined as the mean of the last 7 pain diary entries while the patient was on study medication.

2.2. The MOS Sleep Scale

The MOS Sleep Scale [6] was included as a secondary efficacy variable. It was completed by patients at the baseline and end-of-treatment visits (week 13 in the 1056 study and week 14 in the 1077 study) in both studies and at interim visits (weeks 5 and 9) in the 1056 study. For 10 of the 12 MOS Sleep Scale questions, patients were asked to report how often each particular sleep symptom or problem was applicable to them on a six-point categorical scale ranging from “all of the time” to “none of the time.” The question about time to fall asleep used a five-point categorical response scale ranging from “0 to 15 min” to “more than 60 min.” Quantity of sleep was reported by patients as the average number of hours they slept each night. In the 1056 study, patients were asked to respond to the MOS Sleep Scale questions based on their experience during the past 4 weeks. In the 1077 *Study*, a 1 week recall was used in response to FDA guidelines [13].

Patients’ responses to the 12 questions were aggregated into the MOS Sleep Scale subscale scores and the 9-item Sleep Problems Index score (see Table 1 in online supplementary material) according to the developers’ scoring algorithm [6]. Higher scores indicated more sleep problems (e.g., higher Sleep Disturbance subscale scores indicated more problems initiating or maintaining sleep) except for the Sleep Adequacy subscale where higher scores reflected more adequate sleep.

Psychometric analyses on the MOS Sleep Scale were obtained from all available study patients across all treatment groups. Treatment-related effects on the MOS Sleep Scale have been reported previously [10,11]. In the 1077 study, the MOS Sleep Disturbance subscale was designated *a priori* as the primary MOS Sleep Scale variable for the treatment effect analysis based on the relevance of its content for FM. Therefore, estimated clinically important differences (CID) reported here focused only on the MOS Sleep Disturbance subscale.

2.3. Statistical analyses

The MOS Sleep Scale scores of study patients were compared with scores obtained from a nationally representative sample in the US [7] by computing 95% confidence intervals for the study baseline mean scores and determining whether or not the normative scores of the scales were within their corresponding 95% confidence intervals from the studies. A one-sample Z test for the mean was also conducted to test whether the means of the scales from each of the two trials differed from the corresponding normative means, which was assumed fixed or constant.

A confirmatory factor analysis was performed to evaluate the structure of the MOS Sleep Scale based on three multi-item subscales and four single-item subscales specifically in patients with FM (Table 1 in online supplementary material). Bentler’s comparative fit index was used to assess the fit of the model and values greater than 0.90 were considered to represent adequate fit [14]. Further, standardized factor loadings of at least 0.40 indicated that items on a multi-item scale belonged to their hypothesized subscales [15].

To assess the internal consistency and validity of the multi-item MOS Sleep Scale, we computed item-to-total correlations corrected for overlap (with the item in question removed from the total subscale score) and Cronbach’s alpha reliability coefficients. Item-to-total correlations of at least 0.40 [16] and Cronbach’s alpha reliability coefficients of at least 0.70 [17] were considered acceptable measurement properties.

To estimate CID, we examined differences from baseline in MOS Sleep Disturbance subscale scores as a function of the Patient Global Impression of Change (PGIC) rating. The PGIC is a categorical scale with 7 levels: very much improved, much improved, minimally improved, no change, minimally worse, much worse, or very much worse. It is a patient rating of the study treatment completed at interim visits and at endpoint to assess the overall change from baseline. The estimated CID for each study was based on a mixed-effects model of change from baseline in MOS Sleep Disturbance subscale scores expressed as a function of discrete levels of PGIC [18,19–22]. The model was a longitudinal (first-order autoregressive) repeated measures model in the 1056 study (at weeks 5, 9, 13, or end of study visit) and a cross-sectional regression model in the 1077 study (at week 14 or end of study visit). The models provided average estimates of differences in the MOS Sleep Disturbance subscale score that corresponded to one-category differences on the PGIC.

3. Results

3.1. Descriptive analysis

The 1056 and 1077 studies included 748 and 745 patients, respectively. Most patients were female (94.4% in the 1056 study and 94.5% in the 1077 study) and white (90.2% in the 1056 study and 91.0% in the 1077 study) (Table 1). In the 1056 study, the mean age of patients was 48.8 years and the average duration of FM was 9 years. In the 1077 study, the mean age of patients was 50.1 years and the average duration of FM was 10 years. In both studies, baseline mean pain scores were approximately 7 on a scale from 0 (“no pain”) to 10 (“worst possible pain”).

Consistent with the clinical profile of FM, patients’ baseline MOS Sleep Scale scores were statistically ($P < 0.001$) and substantially poorer than general population values (Table 2). Patients with FM reported sleeping an average of 5.4 and 5.6 h per night