Sleep Disturbance in Patients with Rheumatoid Arthritis: Evaluation by Medical Outcomes Study and Visual Analog Sleep Scales

FREDERICK WOLFE, KALEB MICHAUD, and TRACY LI

ABSTRACT. ***Objective.*** Except for some polysomnography studies, there have been no large quantitative studies of sleep disturbance (SD) in rheumatoid arthritis (RA). SD has taken on new importance with the observation that etanercept and infliximab reduce daytime sleepiness, and patient groups indicate that sleep is an important issue.

Methods. We evaluated 8676 patients with RA and a comparison group of 1364 subjects with nonfibromyalgia, noninflammatory disorders (NID) using the Medical Outcome Study (MOS) sleep questionnaire, including 2 MOS sleep problem indexes (SPI-I, SPI-II) and the MOS SD scale. In addition, patients completed a visual analog scale (VAS) sleep disturbance scale (SDS).

Results. The scales had similar mean values: SPI-I 35.4 (19.4), SPI-II 36.0 (19.1), SDS 35.0 (24.7), and VAS sleep 36.1 (29.7), and the values for the MOS scales exceeded population norms by 25% (VAS by 42%). In multivariable analyses SD was primarily determined by pain and mood. Patients receiving anti-tumor necrosis factor (TNF) did not have less abnormal sleep scores. SD was comparable in RA and NID. The VAS scale was more strongly associated with RA clinical variables than the MOS scales; however, the distributional characteristics of the scales differed, with the VAS scales capturing more extreme values. The standard error of the measurement (SEM), which is related to minimal (important) change, was SPI-I 9.0, SPI-II 7.3, SDS 9.6, and VAS sleep 10.4.

Conclusion. SD is increased in RA, and 25% to 42% of SD can be attributed to RA. SD is linked to pain, mood, and disease activity. SD is slightly greater in women and is less with increasing age. All scales appear to be valid in RA, with minimal differences in SEM. (First Release Sept 1 2006; J Rheumatol 2006;33:1942–51)

Key Indexing Terms:
SLEEP DISTURBANCE
DISEASE ACTIVITY
MEDICAL OUTCOMES STUDY
QUALITY OF LIFE

Rheumatoid arthritis (RA), like other pain conditions, chronic illnesses, and mood disorders[1-11], is associated with disturbed sleep[12-17], and patients with rheumatic disease indicate that sleep is an important concern to them[12]. In addition, sleep issues have taken on new importance to rheumatologists with the observation that etanercept and infliximab reduce daytime sleepiness[18,19]. While the fact that sleep is abnormal in persons with RA is no surprise, much is not known about RA and sleep disturbance in a quantitative way.

From the National Data Bank for Rheumatic Diseases, University of Kansas School of Medicine, Wichita, Kansas; Stanford University, Stanford, California; and Bristol-Myers Squibb, Princeton, New Jersey, USA.

Supported by a grant from Bristol-Myers-Squibb.

F. Wolfe, MD, National Data Bank for Rheumatic Diseases, University of Kansas School of Medicine; K. Michaud, MS, National Data Bank for Rheumatic Diseases, University of Kansas School of Medicine and Stanford University; T. Li, PhD, Outcomes Research, Bristol-Myers-Squibb.

Address reprint requests to Dr. F. Wolfe, National Data Bank for Rheumatic Diseases, Arthritis Research Center Foundation, 1035 N. Emporia, Suite 230, Wichita, KS 67214.
E-mail: fwolfe@arthritis-research.org

Accepted for publication May 5, 2006.

True sleep disturbance is assessed by polysomnography. However, polysomnography is complex and expensive. In addition, it is not suitable for rheumatic disease clinical trials and not indicated for routine patient care[20,21]. Such considerations have supported use of sleep logs/diaries and questionnaires that assess various aspects of sleep. However, studies of logs and diaries all report modest to poor correlations between subjective reports and objective findings[22]. Given these discordances, Sateia, writing about insomnia, suggests that such assessments "... may be better indicators of patient perception of sleep disturbance than they are reflective of true, quantitative sleep abnormalities. Nevertheless, these patient perceptions may well represent as valid an index of insomnia as objective assessments and may be more accurate than a single, global, and retrospective estimate of sleep pattern."

Sleep questionnaires typically address global estimates of sleep quality, specific sleep characteristics, and behaviors, symptoms and attitudes pertaining to sleep. Such questionnaires demonstrate high global test-retest correlations[22] and are the dominant methodology for assessing sleep quality in clinical trials, observational studies, and clinical care. Smith and Wegener recently reviewed questionnaires relevant to rheumatic disease patients[23], including the Medical Outcomes

Study (MOS) sleep questionnaire[24,25]. However, the simplest questionnaire is a visual analog scale (VAS) of global sleep quality[26-29]. A VAS scale is simple enough to use in rheumatic disease clinical care without increasing patient and staff burden when other non-sleep questionnaires are also being used.

We examined self-reported sleep quality in RA and the comparative results of 2 sleep questionnaires, the MOS scale and a VAS sleep scale. It is important to emphasize that these scales represent sleep perceptions rather than objective sleep disturbance, and that when we use terms of sleep disturbance, sleep problems, or sleep abnormality, we are referring to perceptions measured by questionnaires.

We undertook the current prospective study to answer the following questions. Quantitatively, what is the degree of sleep disturbance in RA and how much of the sleep abnormality can be attributed to RA? Which factors associate with sleep disturbance and to what degree? Does anti-TNF therapy alter sleep disturbance? What are psychometric properties of the questionnaires? What is the best clinical tool to evaluate sleep disturbance?

MATERIALS AND METHODS

Patient sample. Patients in this study were participants in the National Data Bank for Rheumatic Diseases (NDB) longitudinal study of rheumatic disease outcomes. Patients were recruited from the practices of 806 United States rheumatologists, and were followed with semiannual questionnaires[30-33]. Our report concerns the status of 8676 patients with RA who completed in 2002 study assessments that contained detailed sleep questionnaires. The maximum number of questionnaires that could have been completed was 2. Because of ongoing enrollment and dropouts, on average patients with RA in this study completed 1.5 semiannual questionnaires. For patients who completed more than one survey during this period a randomly selected survey was utilized for the main report. In addition, to have a comparison group, we also analyzed data from 1364 NDB patients with noninflammatory disorders [98.8% with osteoarthritis (OA) or back pain problems] who completed the questionnaire during the first half of 2002, after excluding patients diagnosed as having fibromyalgia. The specialized sleep questionnaires were administered to non-RA patients in only one assessment period. We included a noninflammatory disorder comparison group to examine whether sleep disturbance is greater in RA than in noninflammatory disorders. This information is of interest when considering the possible effects of anti-TNF therapy on sleep and fatigue.

Demographic and disease status variables. NDB participants completed semiannual, detailed 28-page questionnaires about all aspects of their illness. At each assessment, demographic variables were recorded including sex, age, ethnic origin, education level, current marital status, and medical history. Functional assessment measures included the Stanford Health Assessment Questionnaire (HAQ) functional disability index[34], the HAQ-II, a shortened, modified version of the HAQ with similar scaling but superior psychometric properties[35], and the MOS Short Form-36 (SF-36), from which the physical component summary score (PCS) and mental component score (MCS) were calculated[36,37]. We also collected VAS scales for pain, fatigue, sleep disturbance, and global severity[38]. We formed the Patient Activity Scale (PAS) by multiplying the HAQ by 3.33 and then dividing the sum of the VAS for pain, VAS global, and HAQ by 3[39]. This yields a 0–10 scale. The PAS is a general measure of RA disease activity. Severity groups based on the PAS were formed from the 4 quartiles of the PAS. Quartile 1 values indicate low disease activity. Quartile 4 values indicate very high disease activity; quartiles 2 and 3 occupy intermediate positions. Ranges within the quartiles are: Q1 0.0–1.9, Q2 1.9–3.8, Q3 3.8–5.6, and Q4 5.6–10.

The effect of comorbidity is assessed by a comorbidity score, which is the sum of 11 present or past comorbid conditions reported by the patient. Conditions include cancer, stroke, fracture, renal, neurologic, endocrine, gastrointestinal, cardiovascular, pulmonary, genitourinary, and psychiatric problems.

To assess quality of life (QOL) by utilities, we administered the EuroQol[40-43], utilizing the newly developed US tariffs[44] and a VAS-based QOL scale utilizing linear transformation[45]. The EuroQol contains 5 questions, 3 of which are about function, one about pain, and one about psychological status. The VAS QOL scale was anchored at one end with "death" and at the other end with "perfect health." It was transformed using the algorithm 0.44*(VAS QOL/100) + 0.49[45]. VAS QOL scales have higher utilities than multi-item scales such as the EuroQol, and are closer in that respect to values obtained by the standard gamble (SG) technique[45]. Raw VAS scores are proportionally lower than true utility scores. Transformations alter the metric of the raw VAS scale so that its values and distribution closely represent values obtained by the SG[45]. SG utility scores reflect the patient's assessment and valuation of his or her own health status[46-48]. Specifically, the SG methodology elicits preferences for a given health state by asking the subject to compare life in a given health state with a gamble between 2 outcomes, often a certain amount of life in perfect health and immediate painless death. The probability in the gamble is varied until the subject is indifferent between the certain health state and the gamble between perfect health and death[49].

The MOS Sleep questionnaire (Table 1) is a 12-item self-report questionnaire that yields 2 sleep problem indexes (SPI-I, 6 items, and SPI-II, 9 items) and 6 scale scores: the sleep disturbance scale (SDS, 4 items), daytime somnolence (3 items), snoring (1 item), awakening short of breath or with headache (1 item), and quantity of sleep (1 item). Answers refer to a retrospective assessment over the past 4 weeks. The quantity of sleep is scored as the average number of hours slept per night. Sleep is considered to be optimal (Optimal sleep scale) if the reported hours slept is 7 or 8, and not optimal otherwise. Except for the quantity of sleep measure, the MOS sleep scales and indexes are scored on a 0–100 possible range, higher scores indicating more sleep problems. [The MOS Sleep measure is available at http://www.rand.org/health/surveys/_tools/mos/mos_sleep_survey.pdf]. In this study we emphasized results of the 2 indexes — SPI-II, a 9-item comprehensive index, and the SPI-I, the shorter 6-item index — and the SDS.

In addition, we used a VAS that inquired, "How much of a problem has sleep (i.e., resting at night) been for you IN THE PAST WEEK? Place an X in the box below that best describes how much of a problem sleep has been for you on a scale of 0–100." The scale was anchored with the following descriptors, "Sleep is no problem" and "sleep is a major problem." Higher scores indicate more problems.

Statistical analyses. To obtain information regarding the minimal importance difference (MID) of the sleep variables (Table 3), we first determined Chronbach's alpha. Alpha represents a measurement of internal consistency. Alphas > 0.70 are thought to indicate acceptable internal validity. We calculated the standard error of the measurement (SEM) according to the formula[50,51]

$$SEM = SD^* \sqrt{1-reliability}$$

where SD is the standard deviation of the sleep variable under study and reliability is Cronbach's alpha. We also calculated the minimal detectable change (also called the reliable change or the smallest real difference) according to the formula[52]

$$MDC = SEM * 1.96 * \sqrt{2}$$

As alpha reliability is not available for VAS scales, we estimated it by obtaining the relative reliability of the VAS sleep scale compared to the SLP-II. This was done by comparing the test-retest reliability of the SPI-II to the test-retest reliability of the sleep VAS scale in 2338 patients who completed 2 questionnaire assessments 6 months apart and whose overall health had not changed. We defined the group without health change as those who indicated on their second SF-36 questionnaire that their health was unchanged compared with 6 months ago and whose Time 2 minus Time 1 PCS score was < 5 (< 1/2 of the

Table 1. Question concepts and associated sleep scales of the MOS sleep scale.

Questions	SDS	SNR	SOB	ADQ	SOM	SPI-I	SPI-II
Length of time to fall asleep	•						•
Hours slept							
Sleep was restless or tense	•						•
Enough sleep to feel rested				•		•	•
Awaken short of breath or with headache			•			•	•
Drowsy or sleepy during day					•		•
Trouble falling asleep	•					•	•
Trouble getting back to sleep	•					•	•
Trouble staying awake during day					•	•	•
Snoring		•					
Napping during day					•		
Get amount of sleep needed				•		•	•

SDS: Sleep disturbance scale; SNR: Snoring; SOB: Sleep short of breath or headache; ADQ: Sleep adequacy; SOM: Sleep somnolence; SPI-I: Sleep problems index I; SPI-II: Sleep problems index II.

Table 2. Characteristics of 8676 patients with RA.

Variable	Mean (SD)
Demographics	
Age, yrs	61.1 (12.6)
Sex, % male	22.6
White, not of Hispanic origin, %	92.7
Education, yrs	
0–8, %	2.1
8–11, %	6.7
12, %	36.1
13–15, %	26.3
≥ 16, %	28.8
Clinical scales	
HAQ, 0–3	1.1 (0.7)
HAQ II, 0–3	1.0 (0.7)
Pain, 0–10	3.7 (2.7)
Fatigue, 0–10	4.3 (2.8)
Physical component score, SF-36	32.2 (10.2)
Mental component score, SF-36	44.6 (13.7)
EuroQol, US, 0–1	0.73 (0.17)
VAS-QOL, linear transformation, 0–1	0.79 (0.08)

PCS standard deviation). Because 6 months is too long for true test-retest reliability, we calculated the relative reliability as the ratio of VAS/SPI-II test-retest reliability. To obtain the reliability measure for the VAS we multiplied the ratio (0.814) and SPI-II alpha reliability. In very large samples alpha and test-retest reliability converge[53].

To determine the extent of sleep abnormality increase in RA compared with the general population, we estimated the level of sleep abnormality in the general population using data from the SF-36 to determine expected health at population norms (Table 4). We did this because population data with the age, sex, and educational attainment of the RA cohort are not available. We estimated population sleep scores for each patient as the score attained had his or her general health been at the expected level for age and sex in the general population, and compared the predicted sleep score with the measured sleep score. Specifically, in separate analyses we regressed SDS, SPI-I, SPI-II, and the VAS sleep scale on PCS, MCS, age, and sex using multiple linear regression, and predicted the sleep variable value that would be expected in each of 14 age and sex categories had the PCS and MCS been at the population mean for the age and sex category[54]. To determine the normative PCS and MCS values for the age and sex groups, we used published population norms[54]. This method has been shown to be valid, using external validation[55]. The difference score (Table 4) represents the difference between predicted and expected values. The "current sleep score" in Table 4 has values that are slightly different from the values in Table 3 because 441 patients (5.1%) had missing SF-36 values and were excluded from the analyses.

Table 3. Psychometric properties of the MOS and VAS sleep scales.

Sleep Scales	Mean (SD)	Floor, %	Ceiling, %	Alpha Reliability*	SEM†	MDC††
Medical Outcomes Study Sleep Module						
Sleep Disturbance Scale, 0–100	35.0 (24.7)	6.7	0.9	0.84	9.6	26.7
Snoring scale, 0–100	33.5 (31.1)	29.7	6.7	NA		
Short of breath scale, 0–100	14.3 (24.8)	67.4	1.2	NA		
Sleep adequacy, 0–100	50.5 (26.3)	5.2	3.4	0.70	14.4	39.8
Somnolence scale, 0–100	35.4 (22.6)	4.8	0.9	0.75	10.9	30.1
Sleep Problems Index I, 0–100	35.4 (19.4)	1.5	0.3	0.77	9.0	25.0
Sleep Problems Index II, 0–100	36.0 (19.1)	0.8	0.0	0.85	7.3	20.2
Sleep quantity, h	6.8 (1.4)	0.2	0.3	NA		
Optimal sleep, %§	51.2			NA		
Sleep disturbance VAS 0–100	36.1 (29.7)	15.2	2.1	NA	10.4¶	28.9¶

* Cronbach's alpha. † SEM = SD * $\sqrt{1 - \text{alpha}}$. †† MDC = SEM *1.96 *$\sqrt{2}$. § Defined as having 7 to 8 hours of sleep. ¶ Adjusted SEM (see Materials and Methods). NA: not applicable.

Table 4. Sleep abnormalities in patients with RA and adjustment to the general population.

Sleep Scale	Current Sleep Score, mean (SD)	"Normal" Sleep Score, mean (SD)	Difference in Sleep Score, mean (SD)
VAS Sleep Disturbance, 0–100	35.9 (29.6)	20.9 (2.6)	15.0 (29.5)
MOS Sleep Problems Index II, 0–100	35.2 (18.7)	26.2 (3.1)	9.0 (18.3)
MOS Sleep Problems Index I, 0–100	34.4 (18.6)	26.2 (3.2)	8.3 (18.3)
MOS Sleep Disturbance Scale, 0–100	34.3 (24.5)	24.4 (2.7)	9.9 (24.3)

Comparisons between groups utilized linear regression or t tests, as appropriate. Figures 2B and 2D were drawn utilizing localized polynomial regression[56]. All analyses were performed using Stata version 9.0[56].

RESULTS

Sleep problems in RA. Demographic and RA variables are shown in Table 2, and sleep related variables in Table 3. The results of the 4 major sleep scales were SPI-I 35.4 (19.4), SPI-II 36.0 (19.1), SDS 35.0 (24.7), and VAS sleep 36.1 (29.7). The distributions of these scales are shown in Figure 1. The SDS and the SLP-II reliabilities were 0.84 and 0.85. The SEM was lowest for the SPI-II, indicating it was best able to detect change in this sample. SEM values for the 3 other general scales were similar: SDS (9.6), SLP-I (9.0), and VAS (10.4). The VAS scale had the greatest floor effect (15.2%) and ceiling effect (2.1%).

Optimal sleep, defined as 7 to 8 hours, was found in 51.2%, with the average hours slept being 6.8 (SD 1.4). The number of hours slept, as measured by the optimal sleep scale, increased slightly with age (Figure 2), and hours slept translated into difference in quality of life (Figure 2A). Sleep problems as measured by SPI-II decreased with age (Figure 2D) and were associated with the number of comorbid conditions (Figure 2C). In addition, women had more abnormal SPI-II scores, 36.3 (18.8) vs 31.6 (18.1) for men ($t = 9.8$, $p \leq 0.001$), but the number of hours slept did not differ between the sexes, 6.8 (1.4) for women vs 6.8 (1.3) for men ($t = 0.2$, $p = 0.814$).

To what extent does RA contribute to sleep problems? To answer this question we first determined actual sleep difficulty scores and then estimated sleep difficulty scores under the assumption that each study participant had an SF-36 PCS and

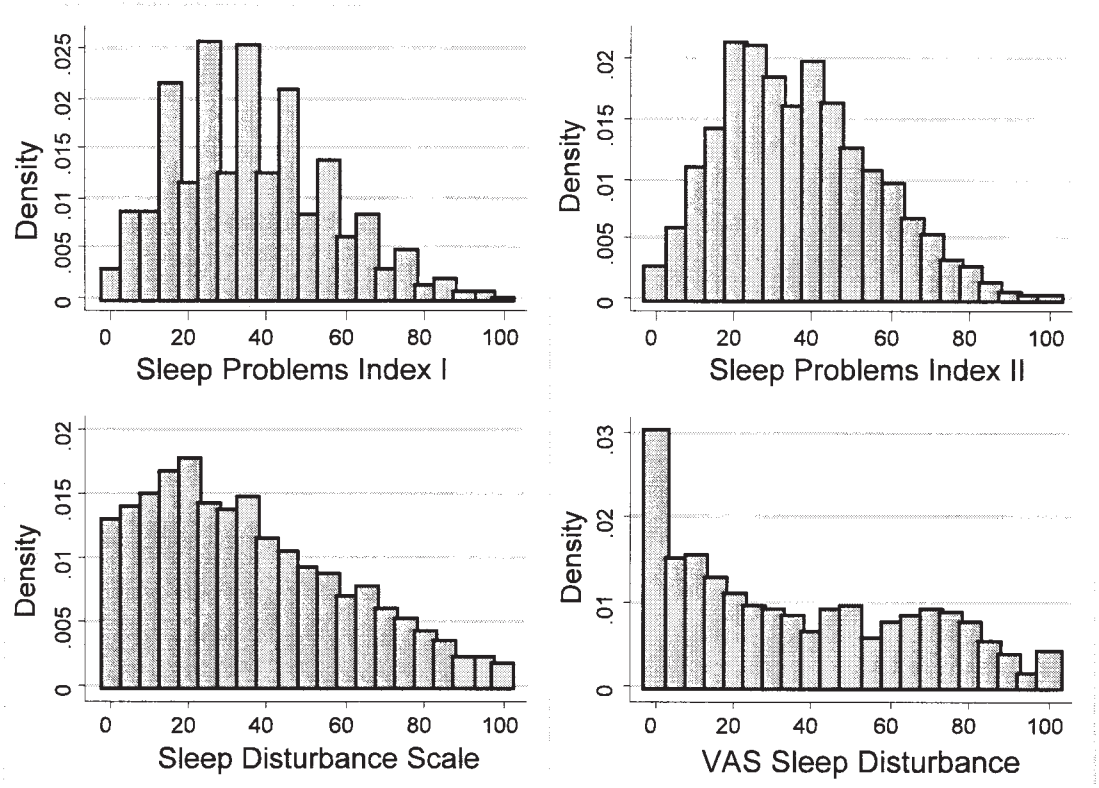

Figure 1. Histograms of the MOS Sleep Problems Index I (SPI-I, upper left), MOS SPI-II (upper right), MOS sleep disturbance scale (lower left), and VAS sleep scale (lower right).

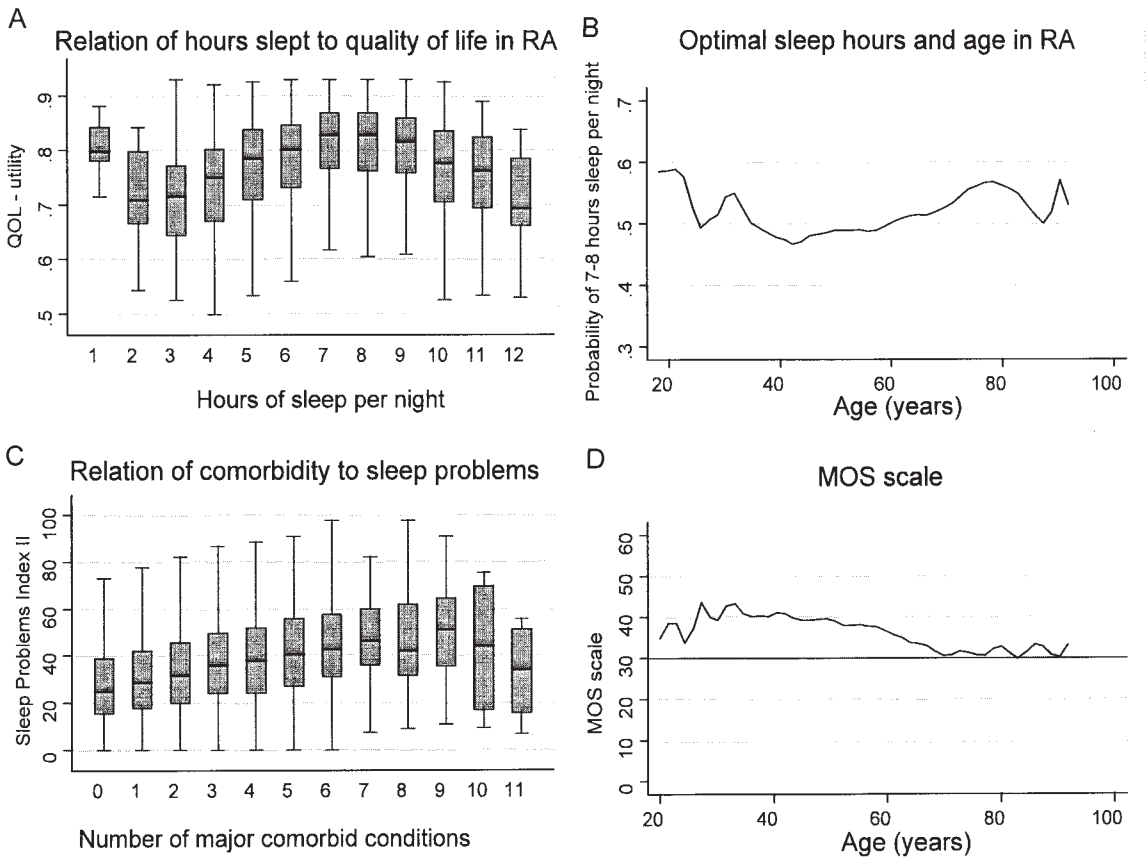

Figure 2. A. Relation of number of hours slept to the linearly transformed VAS quality of life scale. Quality of life is at its maximum among persons sleeping 7–8 hours per night. B. The relation between optimal sleep (defined as 7 or 8 h) and patient age. The connected line is determined from localized polynomial regression. Optimal sleep increases slightly with age. C. Relation of the number of comorbid conditions to the SPI-II. Sleep is increasingly impaired as the number of comorbid conditions increases. D. The relation of age to problems with sleep. Sleep problems are shown to decrease with age.

MCS score at the expected population norm for age and sex (see Materials and Methods). Table 4 shows the sleep scores for the 4 key scales, SPI-I, SPI-II, SDS, and the VAS sleep scale. The mean values for the scales range from 34.3 to 35.9.

The estimated "normal" or general population scores range from 24.4 to 26.2 for the 3 MOS scales, and is 20.9 for the VAS scale. These data are consistent with published population data[24,25]. Using expected versus observed scores (Table 4), we then estimated the percentage of the total sleep scores that is associated with having RA. For the VAS scale, this value is 41.8%; for the SPI-I and -II and the SDS the values are 24.1%, 25.6%, and 28.9%.

RA factors related to sleep problems in RA. As might be expected from the association of RA and sleep abnormality shown in Table 5, the sleep abnormality scales are also correlated with clinical variables, such as HAQ, global severity, and pain (Table 5). The summary score of patient activity measures (PAS), as expected, has the highest correlations, with correlation coefficients ranging from 0.451 (SDS) to 0.560 (VAS scale). The associations with clinical variables are more clearly demonstrated in Figures 3A and 3B. Hierarchical multivariable models of sleep abnormality scores (Table 6) show that for the SPI-II, SPI-I, SDS, and VAS sleep scale, pain and depression explain almost all of the explainable variance, with respective R-squares for the VAS and SPI-II of 0.334 and 0.314. When HAQ is added as a fourth blocked variable, the additional R-square increases are 0.007 and 0.008, respectively.

What is the relationship of anti-TNF therapy to sleep disturbance? We tested the hypothesis that sleep abnormality scores are lower in patients receiving anti-TNF therapy by using multivariable regression in which we regressed the sleep scales on anti-TNF therapy, adjusting for age, sex, HAQ, and pain. Patients using anti-TNF therapy had minimally higher SDS scores: 1.03 (95% CI 0.08 to 2.0; $p = 0.034$). There was no significant association between anti-TNF therapy use and sleep abnormalities using SPI II, SLP-I, daytime somnolence, and the VAS sleep scale in the 4 separate regression analyses ($p > 0.200$).

Is sleep disturbance increased in RA compared with other noninflammatory disorders? After excluding patients with fibromyalgia, 1364 patients in the NDB with OA back disor-

Table 5. Correlation of sleep scales with clinical variables.

Variable	MOS Sleep Index II	MOS Sleep Index I	MOS Sleep Disturbance Scale	VAS Sleep Disturbance
MOS Sleep Problems Index II (0–100)	1.000	0.973	0.896	0.735
MOS Sleep Problems Index I (0–100)	0.973	1.000	0.827	0.707
MOS Sleep Disturbance Scale (0–100)	0.896	0.827	1.000	0.731
VAS Sleep disturbance (0–100)	0.735	0.707	0.731	1.000
Mental component score (SF-36)	–0.607	–0.588	–0.497	–0.512
Anxiety (0–10)	0.570	0.588	0.477	0.488
Fatigue (0–10)	0.561	0.541	0.427	0.550
Depression (0–10)	0.529	0.516	0.444	0.449
Patient Activity Scale (0–10)	0.514	0.487	0.451	0.560
EuroQol (0–1)	–0.499	–0.476	–0.434	–0.484
Patient global severity (0–10)	0.474	0.453	0.408	0.526
Pain (0–10)	0.457	0.434	0.400	0.515
HAQ II (0–3)	0.430	0.403	0.378	0.422
VAS QOL (linear transformation)	–0.417	–0.401	–0.343	–0.393
HAQ (0–3)	0.402	0.376	0.361	0.407
Physical component score (SF-36)	–0.397	–0.371	–0.336	–0.404

All correlations are significant at $p < 0.001$.

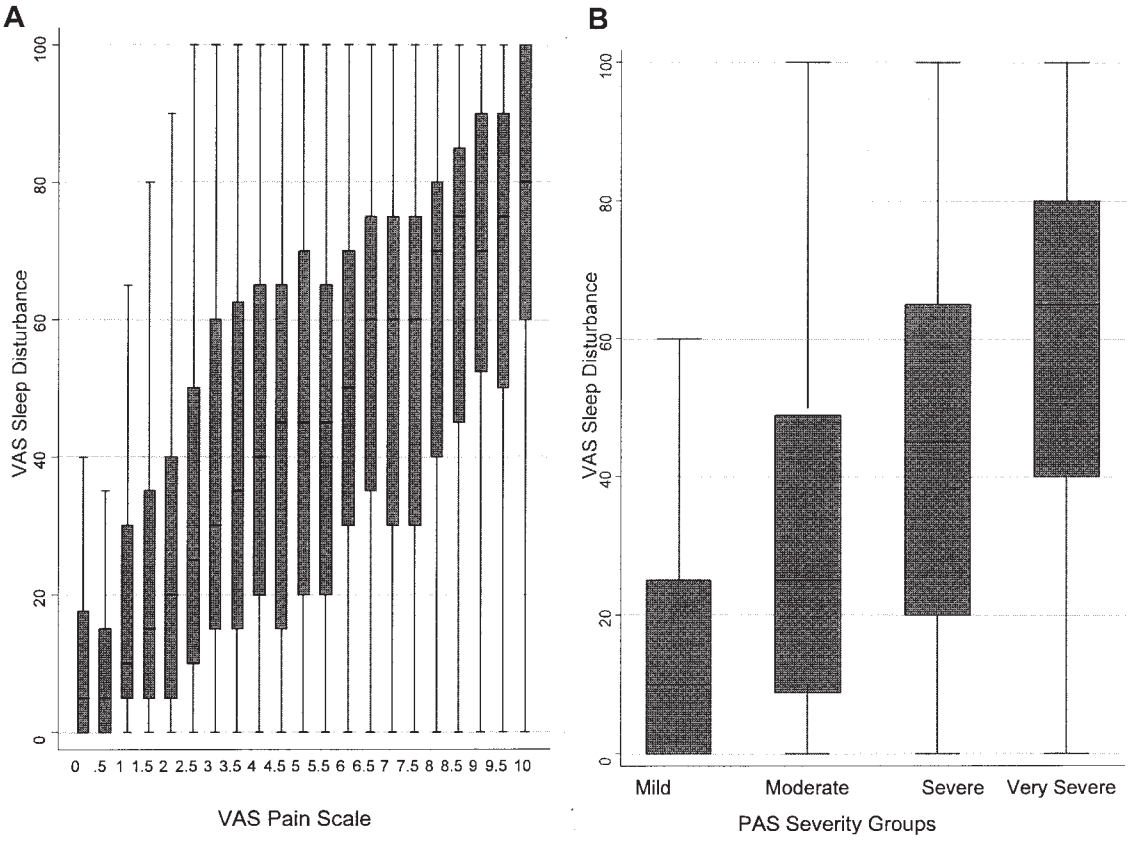

Figure 3. The relation between (A) VAS pain and (B) disease severity groups and VAS sleep disturbance scale.

Table 6. Hierarchical multivariable analysis of sleep scale predictors. Blocks of variables are entered sequentially. The upper set of models enters depression, then pain; the lower set enters pain, then depression. The differences between blocks in all models are significant at $p < 0.001$.

Block	R-Square (VAS)	R-Square Change (VAS)	R-Square (SPI-II)	R-Square Change (SPI-II)
Age, sex, education level, minority status	0.032		0.056	
Depression	0.208	0.176	0.263	0.208
Pain	0.334	0.126	0.314	0.051
HAQ	0.341	0.007	0.322	0.008
Age, sex, education level, minority status	0.032		0.056	
Pain	0.277	0.245	0.210	0.154
Depression	0.334	0.057	0.314	0.104
HAQ	0.341	0.007	0.322	0.008

SPI: Sleep Problems Index.

ders were evaluated with the sleep scales. After adjustment for age and sex, SPI-II scores were not significantly lower in RA (coefficient –0.7, 95% CI –1.8 to –0.2; p = 0.228); SPI-I scores were slightly lower in RA (–1.3, 95% CI –2.5 to –0.2; p = 0.024), SDS scores were not lower in RA (0.4, 95% CI –1.1 to 1.8; p = 0.476); and the VAS sleep scale scores were higher in RA (1.9, 95% CI 0.2 to 3.6; p = 0.028). Overall, these differences were small, and no consistent pattern of difference between RA and OA/back pain patients was identified.

Scale differences. The distribution curves of the 4 scales differ. SPI-I and II are relatively normally distributed (Figure 1), while SDS and VAS sleep disturbance scales are skewed to the left. The VAS scale is better correlated with clinical variables than the other scales, followed by the SDS (Table 5). As noted above and shown in Table 4, the MOS scales and the VAS sleep scale differ in the relative contribution of RA to the sleep abnormality, and this is a reflection of the distributional differences shown in Figure 1.

Use in research and clinical practice. To facilitate the use of sleep scales in research and clinical practice and to allow comparisons among scales, we present Table 7 to serve as a conversion template.

Table 7. Sleep problems index II and sleep disturbance scale values at levels of the VAS sleep scale.

VAS Sleep Scale	MOS Sleep Problems Index II	MOS Sleep Disturbance Scale
0	16.8	10.0
5	20.4	15.4
10	24.4	19.4
15	26.8	23.7
20	29.8	27.4
25	32.1	30.8
30	34.1	33.1
35	35.9	34.5
40	37.4	36.8
45	39.3	40.1
50	41.6	43.5
55	42.8	43.6
60	46.4	48.6
65	47.3	49.8
70	50.0	54.1
75	51.1	56.3
80	54.6	59.6
85	57.1	64.0
90	60.0	66.4
95	66.3	75.1
100	67.8	76.4

DISCUSSION

Using data from the Medical Outcomes Study (n = 3445) in a general population, Spritzer and Hays (1988) noted a score of 29.2 (SD 18.0) for the 9-item SPI-II, and a score of 29.2 (SD 23.4) for the SDS[25]. More recently (2005), Hays, *et al* reported mean scores in the general population of 25.8 for the SPI-II and 24.5 for the SDS[24]. Their results are similar to those we obtained by our SF-36 PCS and MCS based estimation (26.2 and 24.4) for the 2 respective scales, and offer additional support for the validity of this methodology[55].

Using the MOS scales, about 25% of the sleep disturbance score is associated with RA. As shown in Table 4, the differences between the estimated population values and the observed values were 8.3 to 9.9 for the MOS scales, or about 24% to 29% of the respective scores. Although the VAS sleep scale has a similar mean value [36.0 (SD 29.6)], the estimated normative value (20.9) was lower than what was seen with the MOS scale, and the percentage difference was increased to 41.8%. Overall, these data indicate that chronic illness plays an important role in disturbed sleep. It should be noted, however, that there was no difference in sleep scores between RA and OA patients, indicating that it is not RA per se that causes the problem. Pain and depression may be part of the common pathway to sleep disturbance, as shown in Table 6.

The VAS scale differs in a number of ways from the MOS scales. As might be expected, the standard deviation of the

VAS scale is greater (Table 3). In addition, Figure 1 shows that the SPI-II is relatively normally distributed and the VAS sleep scale is skewed to the left and has a dip at its center. The SDS resembles the VAS sleep scale but has fewer patients at the floor. The distribution of the VAS scale seen here, with the dip, is typical of VAS scales, as patients endorse more extreme values than are obtained with the multi-item scales. This is also shown in Table 7, where values of the VAS sleep scale are much higher than values of the MOS scales above the scale means. There are also fewer patients at the floor in the other MOS scales than with the VAS scale. The VAS scale was more highly correlated with clinical measures (Table 5) than were the MOS scales. Interestingly, VAS scales often outperform multi-item scales in their ability to detect change and in their interaction with clinical variables. However, they are often not as good as measurement tools.

Although sleep problems are common in the general population, pain conditions, chronic illness, and mood disorders are associated with increased sleep disturbance[1-11]. In the RA patients under study, sleep problems were clearly related to clinical activity, as shown in Figure 3. Therefore it would be expected that treatments that improve RA clinical status will reduce sleep disturbance. How much change in the sleep scores is important? There are no data regarding MID for the MOS or VAS sleep scales. One method of determining MID is to anchor the sleep scales to other quantitative scales. However, this method requires clinical trial data and is also problematic because sleep is highly dependent on improvement in pain. Wyrwich, *et al* have made a strong case from the review of multiple studies that the MID is best approximated by the SEM[50,51,57]. The SEM does not account for multiple administrations, nor does it include a cutoff at the 95th percentile. This has led others to suggest that minimal detectable change or reliable change is a better metric (Table 3)[52]. The issues are complex and involve different ways of conceptualizing change, as well as differences between patient and group change[53,58]. However, Wyrwich's SEM proposal has strong support from multiple studies. In addition, SEM is close to one-half a standard deviation change, which is close to the average MID[57]. Using the SEM as an approximation of MID, Table 3 indicates that the SEM is smallest for the SPI-II (7.3). However, the values for the other general scales are only slightly increased: SDS (9.6), SPI-I (9.0), and VAS 10.4.

Which scales should be used? If one believes that measurement of sleep has a role in the clinic, as the authors do, then the VAS sleep scale is the only appropriate choice for the clinic because of the difficulty in administering and scoring the MOS questionnaires, particularly when other questionnaire assessments are also being made. Indeed, the first author administered the VAS sleep scale 6752 times in his clinical practice to assess sleep problems for patient care purposes. In addition, there is little difference between SEM values between the VAS and other scales.

The choice is less clear for clinical trials. The SPI-II has slightly greater reliability and a smaller SEM than the other MOS scales. However, the longer SPI-II did not perform as well as the SDS in a clinical trial[24]. This might be expected, as the SPI-II contained items that are unrelated to sleep disturbance (e.g., snoring). Even so, the differences in performance of the MOS scales were slight in that study. It is likely that any of the 3 MOS scales or the VAS scale will perform satisfactorily in clinical trials. However, data from additional trials is required to give more definite answers. Given the general excellent performance of VAS scales in RCT, we suspect the VAS will perform well. Table 7 provides the ability to translate VAS results to MOS scores.

It is of interest that we found that patients who were receiving anti-TNF therapy did not have lower daytime somnolence or overall sleep quality scores. These data are at variance with the recent observation that etanercept and infliximab reduce daytime sleepiness[18,19], although the methodology of the studies is considerably different.

In summary, sleep disturbance is increased in RA, and 25% to 42% of sleep disturbance scores can be attributed to RA. Sleep disturbance is linked to pain, mood, and disease activity. Patients receiving anti-TNF therapy do not have less sleep abnormality. Sleep disturbance is slightly greater in women and is less with increasing age. In addition, sleep disturbance scores are similar in RA and noninflammatory rheumatic disorders. The MOS SPI-II and VAS sleep scales were more strongly correlated with clinical and utility variables than the other scales. The SEM, which is related to minimal (important) change, was SPI-I 9.0, SPI-II 7.3, SDS 9.6, and VAS sleep 10.4. All scales appear to be valid in RA, with minimal differences in SEM.

REFERENCES

1. de Weerd A, de Haas S, Otte A, et al. Subjective sleep disturbance in patients with partial epilepsy: a questionnaire-based study on prevalence and impact on quality of life. Epilepsia 2004;45:1397-404.
2. Wells RD, Day RC, Carney RM, Freedland KE, Duntley SP. Depression predicts self-reported sleep quality in patients with obstructive sleep apnea. Psychosom Med 2004;66:692-7.
3. Anderson KO, Getto CJ, Mendoza TR, et al. Fatigue and sleep disturbance in patients with cancer, patients with clinical depression, and community-dwelling adults. J Pain Symptom Manage 2003;25:307-18.
4. Coyne KS, Zhou Z, Bhattacharyya SK, Thompson CL, Dhawan R, Versi E. The prevalence of nocturia and its effect on health-related quality of life and sleep in a community sample in the USA. BJU Int 2003;92:948-54.
5. Moldofsky H. Sleep and pain. Sleep Med Rev 2001;5:385-96.
6. Newman AB, Enright PL, Manolio TA, Haponik EF, Wahl PW. Sleep disturbance, psychosocial correlates, and cardiovascular disease in 5201 older adults: the Cardiovascular Health Study. J Am Geriatr Soc 1997;45:1-7.
7. Nicassio PM, Wallston KA. Longitudinal relationships among pain, sleep problems, and depression in rheumatoid arthritis. J Abnorm Psychol 1992;101:514-20.
8. Menefee LA, Cohen MJ, Anderson WR, Doghramji K, Frank ED, Lee H. Sleep disturbance and nonmalignant chronic pain: a comprehensive review of the literature. Pain Med 2000;1:156-72.
9. Menefee LA, Frank ED, Doghramji K, et al. Self-reported sleep

quality and quality of life for individuals with chronic pain conditions. Clin J Pain 2000;16:290-7.
10. Roberts RE, Shema SJ, Kaplan GA. Prospective data on sleep complaints and associated risk factors in an older cohort. Psychosom Med 1999;61:188-96.
11. Houssien DA, McKenna SP, Scott DL. The Nottingham Health Profile as a measure of disease activity and outcome in rheumatoid arthritis. Br J Rheumatol 1997;36:69-73.
12. Kirwan J, Heiberg T, Hewlett S, et al. Outcomes from the Patient Perspective Workshop at OMERACT 6. J Rheumatol 2003;30:868-72.
13. Uutela T, Hakala M, Kautiainen H. Validity of the Nottingham Health Profile in a Finnish out-patient population with rheumatoid arthritis. Rheumatology Oxford 2003;42:841-5.
14. Drewes AM, Nielsen KD, Hansen B, Taagholt SJ, Bjerregard K, Svendsen L. A longitudinal study of clinical symptoms and sleep parameters in rheumatoid arthritis. Rheumatology Oxford 2000;39:1287-9.
15. Hirsch M, Carlander B, Verge M, et al. Objective and subjective sleep disturbances in patients with rheumatoid arthritis. A reappraisal. Arthritis Rheum 1994;37:41-9.
16. Devins GM, Edworthy SM, Paul LC, et al. Restless sleep, illness intrusiveness, and depressive symptoms in three chronic illness conditions: rheumatoid arthritis, end-stage renal disease, and multiple sclerosis. J Psychosom Res 1993;37:163-70.
17. Mahowald MW, Mahowald ML, Bundlie SR, Ytterberg SR. Sleep fragmentation in rheumatoid arthritis. Arthritis Rheum 1989;32:974-83.
18. Vgontzas AN, Zoumakis E, Lin HM, Bixler EO, Trakada G, Chrousos GP. Marked decrease in sleepiness in patients with sleep apnea by etanercept, a tumor necrosis factor-alpha antagonist. J Clin Endocrinol Metab 2004;89:4409-13.
19. Zamarron C, Maceiras F, Mera A, Gomez-Reino JJ. Effect of the first infliximab infusion on sleep and alertness in patients with active rheumatoid arthritis. Ann Rheum Dis 2004;63:88-90.
20. Kushida CA, Littner MR, Morgenthaler T, et al. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. Sleep 2005;28:499-521.
21. Littner M, Hirshkowitz M, Kramer M, et al. Practice parameters for using polysomnography to evaluate insomnia: an update. Sleep 2003;26:754-60.
22. Sateia MJ, Doghramji K, Hauri PJ, Morin CM. Evaluation of chronic insomnia. An American Academy of Sleep Medicine review. Sleep 2000;23:243-308.
23. Smith MT, Wegener ST. Measures of sleep. Arthritis Rheum 2003;49 Suppl:S184-S196.
24. Hays RD, Martin SA, Sesti AM, Spritzer KL. Psychometric properties of the Medical Outcomes Study Sleep measure. Sleep Med 2005;6:41-4.
25. Spritzer KL, Hays RD. MOS Sleep Scale: A manual for use and scoring, version 1.0. 1.0. 2003 [pamphlet]. Available at: http://gim.med.ucla.edu/FacultyPages/Hays/sleep.htm. Los Angeles, CA; 2003.
26. Pincus T, Sokka T. Quantitative measures for assessing rheumatoid arthritis in clinical trials and clinical care. Best Pract Res Clin Rheumatol 2003;17:753-81.
27. Moul DE, Nofzinger EA, Pilkonis PA, Houck PR, Miewald JM, Buysse DJ. Symptom reports in severe chronic insomnia. Sleep 2002;25:553-63.
28. Guilleminault C, Palombini L, Poyares D, Chowdhuri S. Chronic insomnia, premenopausal women and sleep disordered breathing: part 2. Comparison of nondrug treatment trials in normal breathing and UARS post menopausal women complaining of chronic insomnia. J Psychosom Res 2002;53:617-23.
29. Zisapel N, Nir T. Determination of the minimal clinically significant difference on a patient visual analog sleep quality scale. J Sleep Res 2003;12:291-8.
30. Wolfe F, Michaud K. Severe rheumatoid arthritis (RA), worse outcomes, comorbid illness, and sociodemographic disadvantage characterize RA patients with fibromyalgia. J Rheumatol 2004;31:695-700.
31. Wolfe F, Michaud K, Anderson J, Urbansky K. Tuberculosis infection in patients with rheumatoid arthritis and the effect of infliximab therapy. Arthritis Rheum 2004;50:372-9.
32. Wolfe F, Michaud K. Heart failure in rheumatoid arthritis: rates, predictors, and the effect of anti-tumor necrosis factor therapy. Am J Med 2004;116:305-11.
33. Michaud K, Messer J, Choi HK, Wolfe F. Direct medical costs and their predictors in persons with rheumatoid arthritis: a 3 year study of 7,527 patients. Arthritis Rheum 2003;48:2750-62.
34. Fries JF, Spitz PW, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.
35. Wolfe F, Michaud K, Pincus T. HAQ-II: Development and validation of a revised version of the Health Assessment Questionnaire (HAQ). Arthritis Rheum 2004;50:3296-305.
36. Mchorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32:40-66.
37. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). 1. Conceptual framework and item selection. Med Care 1992;30:473-83.
38. Wolfe F. Fatigue assessments in rheumatoid arthritis: comparative performance of visual analog scales and longer fatigue questionnaires in 7760 patients. J Rheumatol 2004;31:1896-02.
39. Wolfe F, Michaud K, Pincus T. A composite disease activity scale for clinical practice, observational studies and clinical trials: the Patient Activity Scale (PAS/PAS-II). J Rheumatol 2005;32:2410-5.
40. 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.
41. 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;3:551-9.
42. Wolfe F, Hawley DJ. Measurement of the quality of life in rheumatic disorders using the EuroQol. Br J Rheumatol 1997;36:786-93.
43. EuroQol Group. EuroQol — A new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
44. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005;43:203-20.
45. Mrus JM, Yi MS, Freedberg KA, et al. Utilities derived from visual analog scale scores in patients with HIV/AIDS. Med Decis Making 2003;23:414-21.
46. Russell AS, Conner-Spady B, Mintz A, Maksymowych WP. The responsiveness of generic health status measures as assessed in patients with rheumatoid arthritis receiving infliximab. J Rheumatol 2003;30:941-7.
47. Maetzel A, Strand V, Tugwell P, Wells G, Bombardier C. Cost effectiveness of adding leflunomide to a 5-year strategy of conventional disease-modifying antirheumatic drugs in patients with rheumatoid arthritis. Arthritis Rheum 2002;47:655-61.
48. Suarez-Almazor ME, Conner-Spady B. Rating of arthritis health states by patients, physicians, and the general public. Implications for cost-utility analyses. J Rheumatol 2001;28:648-56.
49. Tufts-New England Medical Center. The CEA Registry. Institute for Clinical Research and Health Policy Studies [Internet. Accessed June 25, 2006.] Available from: http://www.tufts-nemc.org/cearegistry/glossary.html
50. Wyrwich KW. Minimal important difference thresholds and the

standard error of measurement: is there a connection? J Biopharm Stat 2004;14:97-110.
51. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. J Clin Epidemiol 1999;52:861-73.
52. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. J Clin Epidemiol 2004;57:1008-18.
53. Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. J Clin Epidemiol 2001;54:1204-17.
54. Ware JEJ, Kosinski M. The SF-36® physical and mental health summary scales: A manual for users of Version 1, 2nd edition. 1st ed. Lincoln: QualityMetric, Inc.; 2001.
55. Wolfe F, Michaud K, Choi HK, Williams R. Household income and earnings loss among 6396 persons with rheumatoid arthritis. J Rheumatol In press 2005.
56. Stata Corporation. Stata Statistical Software: Release 9.0. College Station, TX: Stata Corporation; 2003.
57. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003;41:582-92.
58. Liang MH. Longitudinal construct validity — Establishment of clinical meaning in patient evaluative instruments. Med Care 2000;38:84-90.