Correspondence

Prevalence of Rheumatoid Arthritis

To the Editor:

I read with great interest the epidemiological study of rheumatoid arthritis (RA) in Hungary[1]. However, 2 points need to be clarified. Firstly, the authors interviewed 10,000 inhabitants with a refusal rate of 21.5%. These 2150 persons are not seen in Figure 1. The last point is that if you survey 10,000 inhabitants and find 13 patients with RA the prevalence of RA will be simply 0.13% [10/5515 (0.18%) in women; 3/4485 (0.06%) in men] instead of 0.37% (0.48% and 0.23% in women and men, respectively) even if those who have refused are excluded.

SEBAHATTIN YURDAKUL, MD, Division of Rheumatology, Department of Medicine, Cerrahpasa Medical Faculty, University of Istanbul, Aksaray 34098, Turkey. E-mail: profsyurdakul@yahoo.com

REFERENCE
1. Kiss CG, Lovei C, Suto G, et al. Prevalence of rheumatoid arthritis in the South-Transdanubian region of Hungary based on a representative survey of 10,000 inhabitants. J Rheumatol 2005;32:1688-90.

Drs. Kiss and Czirjak reply

To the Editor:

Dr. Yurdakul raised 2 questions in connection with our recent epidemiological study of rheumatoid arthritis (RA)[1]. The first point was how we performed 10,000 successful interviews with a refusal rate of 21.5%. Our aim was to achieve altogether 10,000 successful interviews. The sample was representative, and randomly selected in several steps with regard to the region's settlement structure, age, and gender distribution. In case of an interview refusal, the following equally randomized, matching address was taken from the original sample ("dedicated additional" method). This procedure was performed until we completed the originally planned 10,000 successful, representative interviews. Because of this method we did not include data from those who refused the interview.

The second point was how we calculated the prevalence of RA. Out of the 632 individuals who declared joint symptoms that may have been caused by RA, only 471 cases provided an informed consent for further study, and finally 224 underwent clinical investigation. Among these individuals 13 patients with RA were identified. If we had the opportunity to investigate all the 632 individuals we would have identified more patients with RA. Assuming that the proportion of RA could be the same among the 408 individuals who did not participate in the clinical investigation, we would have identified another 24 cases with RA. Therefore if all the 632 individuals had participated in the clinical investigation we would have found around 37 patients with RA. This is the reason why the overall prevalence was 0.37% in our study.

CSABA G. KISS, MD; LASZLO CZIRJAK, MD, Department of Immunology and Rheumatology, Hungarian Brothers of St. John of God, University of Pecs Medical School, Irgalmasok u. 1, H-7621 Pecs, Hungary.

REFERENCE
1. Kiss CG, Lovei C, Suto G, et al. Prevalence of rheumatoid arthritis in the South-Transdanubian region of Hungary based on a representative survey of 10,000 inhabitants. J Rheumatol 2005;32:1688-90.

Is the Jadad Score the Proper Evaluation of Trials?

To the Editor:

Towheed[1] evaluated 3 published trials and an abstract describing another trial of Pennsaid, and noted that the Jadad score was a perfect 5 of 5 for each of the 3 published trials. This uncritical acceptance of the Jadad score led to corresponding uncritical acceptance of the trial results as well. For example, Towheed[1] seems convinced that the "excellent quality" of the trials, and their findings that Pennsaid is effective, prove that "Pennsaid deserves further consideration when the existing treatment guidelines for OA of the knee are updated." In fact, what needs updating is the method by which trial quality is evaluated, because there is no difficulty in finding serious methodological flaws in the very studies that earned such high praise. Given the space constraints, I will focus on only one study[2], which by itself could fill volumes with examples of what not to do in good clinical research.

First, an unmasked trial was referred to as masked, and treated as masked. But masking means more than simply attempting to conceal treatment identities, it requires the success of this effort. The authors acknowledge the garlic taste of the active treatment. In addition, the differential rate of dry skin across treatment groups could certainly lead to unmasking. A block size of 6 is quite small in an unmasked trial with 3 treatment groups, and this has to be considered a methodological flaw that allows for prediction of upcoming allocations, and hence selection bias[3]. Was there selection bias in this trial? We do not know, because not only was selection bias not tested, but in fact even the baseline p values were suppressed. It is notable that many more patients with 2 bad knees ended up in the placebo group than in the active group. The worse of 2 knees will tend to be worse than a single bad knee, so this baseline imbalance represents an advantage for the active group, even if the p value exceeds 0.05. It may be argued that this was already taken care of, by considering Δ [change from baseline in the Western Ontario and McMaster University Osteoarthritis Index (WOMAC) subscale score for pain] as the primary outcome measure. This leads to our next methodological flaw. The WOMAC subscale score for pain is scored on a 5-point Likert scale, and hence is non-numeric data. What does Δ represent?

There has to be concern regarding the equating of all one-category shifts, for example. Is a change from 0 to 1 the same as a shift from 1 to 2, from 2 to 3, or from 3 to 4? This seems highly unlikely, and so a table is

needed showing how many patients in each treatment group shifted from each given baseline score to each given subsequent score, as in Berger, *et al*[4]. This most basic of data presentations is not provided, so the reader cannot produce a reasonable analysis (that the authors should have provided) that is not corrupted by the imposition of this artificial assumption that the 5 categories are spaced uniformly. Then, to make matters worse, some of these baseline scores were actually measured at Day 1, that is, subsequent to randomization. The potential for bias goes without saying[5,6]. To compensate, there was an unplanned increase in sample size, which essentially is an interim analysis with no penalty applied. Again, the potential for bias is clear to any beginning biostatistics student. Moreover, there were additional missing data, and these were imputed by carrying forward the last observation, with no mention of any sensitivity analyses. This is also quite problematic[7].

The primary analysis is the analysis of covariance (ANCOVA). The assumptions underlying the ANCOVA model include normality of residuals, equal variances, linearity, and independence. It is not likely that these assumptions can all be met, and when these assumptions are not met the ANCOVA may not be robust[8]. By not requiring such assumptions, a nonparametric analysis offers better robustness properties, and so should have been used instead. As it stands, the low p value rejects the combination of the null hypothesis and all assumptions, and hence may be attributable to the falsity of any of the assumptions instead of to the falsity of the null hypothesis. To make matters worse, the analysis labeled as intent-to-treat is based on a subset of the true intent-to-treat sample. That is, there is a post-randomization exclusion. The bias this can create is bad enough, but to call the analysis "intent-to-treat" is unconscionable.

The issues discussed include: (1) unmasking; (2) prediction of future allocations; (3) selection bias; (4) performing arithmetical operations on numbers assigned fairly arbitrarily to non-numeric categories; (5) failure to present the most meaningful data structures; (6) using post-randomization data as baseline data; (7) failing to apply a penalty for an unplanned interim analysis; (8) carrying forward the last observation without mentioning any sensitivity analyses; (9) using an analysis requiring so many unverifiable assumptions that it cannot be taken seriously in the context of an actual clinical trial; and (10) excluding from the analysis some post-randomized data. One can easily anticipate the responses of the authors when trying to defend their work. The study was masked, because the paper says that it was. This takes care of the first 3 issues. The categories are nearly equally spaced, the treatment did not yet have time to influence Day 1 data, the increase in sample size was not based on an attempt to get a nearly significant result to become significant, last observation carried forward (LOCF) and ANCOVA are industry standards, and only one randomized patient was excluded from the analysis called intent-to-treat.

In fact, one may be able to argue convincingly that any one of these issues cannot by itself invalidate the findings, or the conclusions based on the findings. But if any one bias can explain the results ("or" logic), or even if a combination of them can do the trick, then the conclusions are not supported. Supporting the conclusions therefore requires arguing that none of these 10 flaws materially affected the outcomes. Some of these arguments would be hard to support, but even if 10 solid arguments were provided, this still would not absolve the authors of their responsibility to conduct good research. Clearly, they did not, and this remains true even if it is found that the many flaws did not materially affect the research. The questions we are left with are (1) Is Pennsaid in fact effective for osteoarthritis of the knee? and (2) Is the Jadad score the way to evaluate trial quality? While deferring to those more knowledgeable than I am regarding the first question, I can offer an unqualified "No" to the second. It would take a much more comprehensive set of checks than the Jadad score offers to be able to replace critical thinking and evaluation with a checklist. Until such a comprehensive checklist is developed, peer review is needed to weed out flawed research. Clearly, peer review also failed in the case of this study.

VANCE W. BERGER, PhD, National Cancer Institute and University of Maryland Baltimore County, Biometry Research Group, National Cancer Institute, Executive Plaza North, Suite 3131, 6130 Executive Boulevard, MSC 7354, Bethesda, Maryland 20892-7354, USA.
Address reprint requests to Dr. V.W. Berger. E-mail: vb78c@nih.gov

REFERENCES

1. Towheed TE. Pennsaid therapy for osteoarthritis of the knee: a systematic review and metaanalysis of randomized controlled trials. J Rheumatol 2006;33:567-73.
2. Bookman AM, Williams KSA, Shainhouse JZ. Effect of a topical diclofenac solution for relieving symptoms of primary osteoarthritis of the knee: a randomized controlled trial. CMAJ 2004;171:333-8.
3. Berger VW. Selection bias and covariate imbalances in randomized clinical trials. Chichester: John Wiley & Sons; 2005.
4. Berger VW, Zhou YY, Ivanova A, Tremmel L. Adjusting for ordinal covariates by inducing a partial ordering. Biometrical Journal 2004;46:48-55.
5. Voutilainen PE. Assessment of grouping variable should have been blind in trial of dementia [letter]. BMJ 2001;322:1491.
6. Berger VW. Valid adjustment for binary covariates of randomized binary comparisons. Biometrical Journal 2004;46:589-94.
7. Unnebrink K, Windeler J. Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. Stat Med 2001;20:3931-46.
8. Lachenbruch PA, Clements PJ. ANOVA, Kruskal-Wallis, normal scores, and unequal variance. Communications in Statistics — Theory and Methods 1991;20:107-26.

Dr. Towheed replies

To the Editor:

Dr. Berger's letter is a critical commentary on the pitfalls of using the Jadad score[1] as an instrument for measuring the quality of randomized controlled trials (RCT) evaluating Pennsaid in osteoarthritis (OA) of the knee[2]. In support of his position, he outlines 10 methodological problems in one of the included RCT[3]. This RCT received a Jadad score of 5, which is indicative of excellent quality.

Although it is generally agreed that trial quality ought to be investigated in systematic reviews, there is unfortunately no consensus on what methodology should be employed[4]. Indeed, there is as yet no gold standard by which to evaluate the internal validity or methodological quality of a RCT. None of the currently available scales for measuring the validity of trials can be recommended without reservation[5]. With this background, the selection of the Jadad checklist has relative merit since it uses a simple and easy to understand approach that incorporates the most important individual components of methodological quality. This includes randomization, blinding, and handling of patient attrition. Based on empirical evidence and theoretical considerations, randomization, blinding, and handling of patient attrition in the analysis should always be assessed when evaluating the quality of a RCT[4]. Allocation concealment is also a vital aspect of the trial's quality[6]. Inadequate concealment of treatment allocation and lack of double-blinding have been associated with an exaggeration of treatment effects[4,7]. For these reasons, there is ample justification for the selection of the Jadad scale and supplementing this checklist with an assessment of allocation concealment. It is a scale that is recommended by the Cochrane Musculoskeletal Group in the preparation of their Cochrane systematic reviews.

It must be acknowledged that the Jadad score is not a perfect instrument. For example, it places greater emphasis on the quality of reporting as opposed to the actual methodological quality of a trial. In addition, it does not assess allocation concealment. Despite these limitations, the Jadad

---

scale is the only instrument that has been constructed according to psychometric principles[4].

Berger outlines 10 "serious methodological flaws" in the RCT published by Bookman, *et al*[3]. Unfortunately, most of these items are a matter of opinion and debate, and in my opinion, of questionable validity and/or significance.

With respect to his first criticism (1, unmasking), this trial was indeed adequately double-blinded. Subjects should have been blinded to the garlic-like taste or odor related to the DMSO component of Pennsaid since DMSO was present in all 3 study solutions (in the Pennsaid as well as in the 2 placebo solutions). The 3 study solutions were identical clear colorless liquids in opaque bottles. Drying of the skin could potentially have unmasked the blinding, but the differences in the prevalence of this adverse effect were not extreme and one cannot assume that subjects were aware *a priori* that Pennsaid is more likely to cause this reaction than is a placebo solution. The only way to assess whether subjects were unblinded would have been to simply ask subjects at the end of the trial what group they were in. I don't believe this was carried out in the trial.

With respect to Berger's second criticism (2, lack of allocation concealment), this is also questionable. As clearly described in the Methods of the paper, the authors took the necessary steps to ensure that the randomization scheme was concealed and it does not appear reasonable to believe that the investigators could have somehow guessed the treatment assignments. Whether a block size of 6 could have compromised allocation concealment is also a matter of conjecture without any direct evidence to support it.

With respect to Berger's third criticism (3, selection bias), this is also questionable. There were no statistically significant baseline differences in terms of the number of subjects with bilateral knee involvement versus unilateral knee involvement. Furthermore, the baseline WOMAC pain and pain on walking scores are very similar for all 3 groups. The change scores were used in the outcome analysis. Finally, in the ANCOVA, the baseline WOMAC and number of knees treated were entered as covariates. Only one knee was used in the efficacy analysis and this was selected *a priori* as being the most symptomatic knee. His statement that "the worse of 2 knees will tend to be worse than a single bad knee" is purely conjecture and not based on any evidence that I am aware of.

With respect to Berger's fourth criticism (4, WOMAC is a non-numeric variable which requires a non-parametric statistical test), it is to be noted that the overall WOMAC pain score is not simply scored from 0 to 4. The pain subscale comprises 5 questions, each of which is scored on a scale of 0 to 4. Thus, the range of possible values for the overall WOMAC pain score is from 0 to 20. The physical function subscale has a range of scores from 0 to 68. This larger range of scores may well have been normally (or approximately normally) distributed, allowing a parametric test to be used. Without knowing the actual frequency distribution of the outcome variables, one cannot make a definite statement as to whether a parametric or non-parametric statistical test should have been used, but I would tend to give the authors the benefit of the doubt that they indeed selected the correct test.

With respect to Berger's fifth criticism (5, failure to present meaningful data structures), the *a priori* outcome selected in this study was to compare changes in the outcome variables between the 3 groups. The authors did not select *a priori* the percentage of patients changing categories in the WOMAC subscales as their method for outcome assessment. In doing so, they selected the most common and widely accepted format used for analyzing outcomes with the WOMAC.

With respect to Berger's sixth criticism (6, post-randomization data used as baseline data), he is correct in identifying this as a limitation. Indeed, the authors acknowledged this as well in the paper. Of note though, a sensitivity analysis was done by excluding data for those patients in whom the Day 1 scores were substituted for the baseline scores and this did not reveal any relevant differences that would change the conclusions of the paper.

With respect to Berger's seventh criticism (7, unplanned interim analysis), it is not clear how he equates the act of increasing sample size with an unplanned interim analysis. My understanding was that the decision to increase sample size was done to overcome the limitations of using Day 1 scores as baseline in some subjects and prior to the assessment of the study outcomes. There is no evidence that it was done for the purposes of making an unplanned interim analysis.

With respect to Berger's eighth criticism (8, last observation carried forward without a sensitivity analysis), this is a commonly used approach in the imputation of missing continuous data. The trial by Bookman, *et al* had a reasonably small percentage of withdrawals (16%). Rarely does one see a sensitivity analysis presented in the published report of trials in OA. In part, this may be because there are no imputation strategies that can be widely accepted for every possible situation. Approaches to imputing missing continuous data and their evaluation by sensitivity analyses are not currently at a stage of development that investigators can practically apply, and there is a lack of published guidelines in this area as to what constitutes a gold standard.

With respect to Berger's ninth criticism (9, ANCOVA), this is also speculative and without knowing the actual frequency distributions of the outcome variables, one cannot comment on this. However, the ANCOVA does allow adjustment for baseline differences, and one has to assume that the WOMAC scores were normally or approximately normally distributed. Regression diagnostics could have been carried out to check the normality assumption.

With respect to Berger's tenth criticism (10, excluding post-randomization data), only one subject out of 248 was excluded from the efficacy analysis. This is not perfect, but reasonable, given the relatively large number of evaluated subjects. It is unlikely that the exclusion of one subject would have materially affected the study's conclusions.

In summary, there is ample justification for using the Jadad checklist to evaluate the quality of a RCT, and supplementing this with an assessment of allocation concealment. It is not often that one will read a published RCT that is truly a masterpiece of methodological perfection and rigor in all the various aspects of trial reporting and conduct. Bookman, *et al*'s study[3] is a reasonably robust and methodologically sound trial that is likely to be associated with a low degree of bias.

TANVEER E. TOWHEED, MD, Department of Medicine, Queen's University, Etherington Hall, Room 2066, Kingston, Ontario K7L 3N6, Canada. E-mail: tt5@post.queensu.ca

REFERENCES

1. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Control Clin Trials 1996;17:1-12.
2. Towheed TE. Pennsaid therapy for osteoarthritis of the knee: a systematic review and metaanalysis of randomized controlled trials. J Rheumatol 2006;33:567-73.
3. Bookman AM, Williams KSA, Shainhouse JZ. Effect of topical diclofenac solution for relieving symptoms of primary osteoarthritis of the knee: a randomized controlled trial. CMAJ 2004;171:333-8.
4. Juni P, Altman DG, Egger M. Assessing the quality of randomized controlled trials. In: Egger M, Smith GD, Altman DG, editors. Systematic reviews in health care. Meta-analysis in context. London: BMJ Publishing Group; 2001:87-109.
5. Higgins JPT, Green S, editors. Assessment of study quality. Cochrane handbook for systematic reviews of interventions 4.2.5 [Updated May 2005]; Section 6. In: The Cochrane Library, Issue 3. Chichester: John Wiley & Sons, Ltd.; 2005.
6. Schulz KF. Subverting randomization in controlled trials. JAMA 1995;274:1456-8.
7. Schulz KF, Chalmers I, Hayes RJ, Altman D. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408-12.

Antibodies Against α-Fodrin Are Associated with Dry Eyes and Mouth in the General Population

To the Editor:

Establishing the diagnosis of Sjögren's syndrome (SS) is difficult, since there are other common causes of reduced tear and saliva production. Further, complaints of dry eyes and mouth are common in the population, may be associated with depression, and barely correlate with objective test results[1,2]. Autoantibodies against Ro/SSA and La/SSB are used as markers of SS[3]. A large study has, however, failed to show an association of these autoantibodies with dry eyes and mouth in the general population[2]. Antibodies against α-fodrin were first described in a mouse model of SS[4]. There is some dispute on the prevalence of these antibodies in patients with SS, ranging from approximately 20% to 98% in various studies, the average being between 50% and 80%[5]. We studied whether antibodies against α-fodrin were associated with objective signs of dry eyes and mouth in the general population.

Participants of our study approved by the local ethical committee were recruited on a summer festival of Hannover Medical School, visited mostly by personnel and students. We advertised our study with the slogan "400 ml of beer for 4 ml of blood," since participants were rewarded with beer (or alternatively banana juice). After giving informed and written consent, participants had to answer a questionnaire on subjective signs of dry eyes and mouth[6]. Schirmer's (normal > 5 mm/5 min on both eyes) and Saxon's test (normal at least 3 g/2 min)[7] were performed to measure tear and saliva production. Blood was drawn for measurement of antibodies against α-fodrin, Ro, and La (Aesku.Diagnostics, Wendelsheim, Germany). Fisher's exact test was used to calculate associations of presence of autoantibodies with reduction of tear and saliva production.

In total, 168 participants (107 women, 61 men) were recruited. Average age was 38 (range 18 to 76) years. Objective reduction of tear production alone was present in 29%, of saliva production alone in 2%, and of both tear and saliva production in 2% of participants. Subjective symptoms of both dry eyes and dry mouth were present in 14% of participants. IgA antibodies against α-fodrin were observed in 5%, IgG antibodies against α-fodrin in 3%, IgA and/or IgG antibodies against α-fodrin in 7%, IgG antibodies against Ro in 1/168 and against La in none of the participants. Both pathologic Schirmer's as well as Saxon's test but not complaints of dry eyes and mouth correlated with age (Figure 1). IgA antibodies against α-fodrin were present in 2% of the participants with neither dry eyes nor dry mouth, but in 3 of 4 participants with both dry eyes and dry mouth (Figure 2) (p = 0.0002 for comparison with normal participants without dry eyes or dry mouth). IgG antibodies against α-fodrin were present in 2% of the participants with neither dry eyes nor dry mouth, but in 2 out of 4 participants with both dry eyes and dry mouth (p = 0.005). The one participant with antibodies against Ro had dry eyes only and did not complain of sicca symptoms. There was no correlation of antibodies against α-fodrin or Ro with subjective complaints of dry eyes or dry mouth or a combination of both. However, complaints of dry eyes and of dry mouth were associated with each other (p = 0.0045; data not shown).

Antibodies against α-fodrin are the first laboratory markers that are associated with both dry eyes and mouth in the general population. However, since in our study salivary gland biopsies could not be taken, none of the 3 participants with a combination of dry eyes and mouth and antibodies against α-fodrin would have fulfilled American/European con-
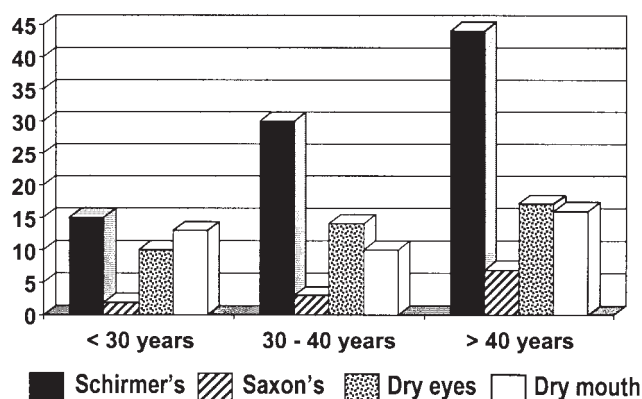

Figure 1. Prevalence (%) of pathologic results for Schirmer's and Saxon's test, subjective complaints of dry eyes and dry mouth in participants stratified by age.
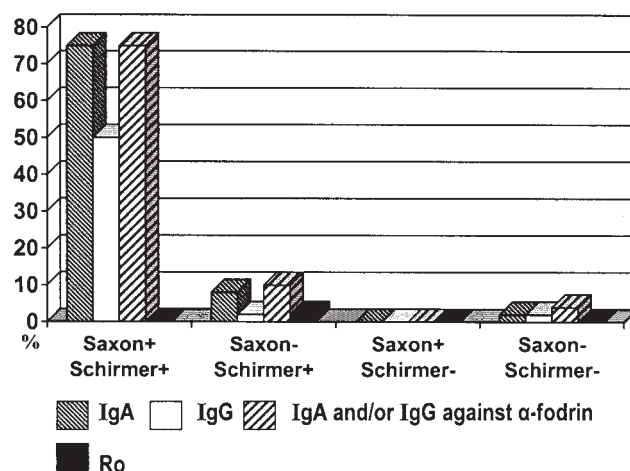

Figure 2. Prevalence of IgA, IgG, IgA and/or IgG antibodies against α-fodrin, and Ro antibodies in participants with both dry eyes and mouth (Saxon+ Schirmer+, n = 4), only dry eyes (Saxon– Schirmer+, n = 48), only dry mouth (Saxon+ Schirmer–, n = 4), and neither dry eyes nor dry mouth (normal, n = 112).

sensus criteria for classification of SS[3]. As described in other studies[2], subjective complaints of dry eyes and mouth correlated neither with objective test results nor with any of the autoantibodies studied. They probably are symptoms of fibromyalgia and depression rather than of SS[8]. Differences in reliance on these subjective parameters in classification of SS, that to our point of view will decrease the specificity of diagnostics, may also explain different results of prevalence of antibodies against α-fodrin in SS obtained in various studies. For diagnostics of SS, simple questions are not sufficient, but objective tests for sicca syndrome should always be performed. The crucial issue will be to define whether patients with sicca syndrome and α-fodrin antibodies but without antibodies against Ro and La suffer from SS. If this turns out to be the case, prevalence of SS would be higher than anticipated at present and α-fodrin antibodies would have to be determined in diagnostics of SS in addition to Ro and La antibodies.

TORSTEN WITTE, MD; JANINE BIERWIRTH, MD; REINHOLD E. SCHMIDT, MD, Department of Clinical Immunology, Hannover Medical School, Hannover, Germany; TORSTEN MATTHIAS, MD, Aesku.Diagnostics, Wendelsheim, Germany.

REFERENCES

1. Schein OD, Hochberg MC, Munoz B, et al. Dry eye and dry mouth in the elderly: a population-based assessment. Arch Intern Med 1999;159:1359-63.
2. Hay EM, Thomas E, Pal B, Hajeer A, Chambers H, Silman AJ. Weak association between subjective symptoms or and objective testing for dry eyes and dry mouth: results from a population based study. Ann Rheum Dis 1998;57:20-4.
3. Vitali C, Bombardieri S, Jonsson R, et al. Classification criteria for Sjogren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. Ann Rheum Dis 2002;61:554-8.
4. Haneji N, Nakamura T, Takio K, et al. Identification of α-fodrin as a candidate autoantigen in primary Sjogren's syndrome. Science 1997;276:604-7.
5. Witte T. Antifodrin antibodies in Sjogren's syndrome: a review. Ann NY Acad Sci 2005;1051:235-9.
6. Vitali C, Bombardieri S, Moutsopoulos HM, et al. Assessment of the European classification criteria for Sjogren's syndrome in a series of clinically defined cases: results of a prospective multicentre study. The European Study Group on Diagnostic Criteria for Sjogren's Syndrome. Ann Rheum Dis 1996;55:116-21.
7. Kohler PF, Winter ME. A quantitative test for xerostomia. The Saxon test, an oral equivalent of the Schirmer test. Arthritis Rheum 1985;28:1128-32.
8. Rhodus NL, Fricton J, Carlson P, Messner R. Oral symptoms associated with fibromyalgia syndrome. J Rheumatol 2003;30:1841-5.

WHO-ILAR COPCORD Pilot Study in Tehran, Iran

To the Editor:

COPCORD (Community Oriented Program for Control Of Rheumatic Diseases) was created for the recognition, prevention, and control of rheumatic disorders. A small-scale COPCORD study was performed in a rural community in Iran in 1993[1]. The study was conducted on 2502 persons in Fasham district of Shemiranat (the northern suburb of Tehran). The population is Caucasian. Iran is a country with different ethnic groups: Caucasians, Turks, and Semites[2]. It was therefore necessary to have a new COPCORD study taking into account all the ethnic populations of Iran. A pilot study was designed by the Rheumatology Research Center at Tehran University for Medical Sciences, to test the feasibility of a large-scale project.

Five districts (north, south, east, west, and the center) were selected randomly in Tehran. The population of Tehran reflects the ethnic distribution of Iran (Caucasians 75.4%, Turks 22%, Semites 2.6%). The 5 clusters in the 5 districts were the Namjoo, Esma'il-Abad, Jahan Ara, Niroo-Hava'I, and Tehransar. With the help of the Iranian Post Office organization, houses were selected randomly for interviews. The average population per household (aged 15 yrs and over) is 2.7 persons in Tehran (Iran population census 1996). Interviewers were selected from Bachelor of Science certified nurses or nurse-midwives. They were trained by interviewing 20 subjects. The observed agreement was 0.96. The chance-expected agreement was 0.53. The kappa coefficient was 0.92 (standard error 0.11). The z status was 8.19. The one-tailed p value was < 0.0001.

Laboratory technicians were briefed one week before the pilot study. Rheumatologists were selected from among the rheumatology subspecialty fellows, and were briefed accordingly. Five team supervisors were selected from 28 candidates and trained. The original COPCORD questionnaire was translated from English to Farsi by a rheumatologist not working for the project. The Farsi version was then back-translated by another rheumatologist, unaware of the original version. The 2 versions did not differ significantly.

The pilot study started on October 3, 2003, and 5 interview teams participated. Each team consisted of a supervisor, 3 interviewers, one rheumatologist, one laboratory technician, and 2 drivers (2 cars). Teams were supervised by the administrative director (A. Tehrani) under the supervision of the project directors (F. Davatchi and A.R. Jamshidi). All data were entered into a computer. Five percent of interviews were subjected to quality control, and errors were found in less than 1% of these.

One hundred sixty-eight houses were visited. The completed interviews totalled 284 (response rate 60%). Subjects' ages ranged from 15 to 82.5 years; mean age was 39.2 (standard deviation 17.4). The male to female ratio was 0.87 to 1. The ethnic distribution was Caucasians 66.2%, Turks 32%, and Semites 1.8%.

One hundred twenty-nine patients (45.4%, 95% CI 5.8) complained of musculoskeletal disorders (MSD) during the past week. Among them, 26.5% had a recent antecedent of trauma. Abstracting data for traumatic complaints, the percentage of rheumatic complaints became 34.5%. The distribution was shoulder 18.3%, wrist 13.4%, hand 15.1%, hip 10.2%, knee 26.1%, ankle 12.7%, great toe 11.6%, cervical spine 13.7%, dorsal and lumbar spine 22.2%, and other 12.3%. Past complaint of MSD was 21.1% (95% CI 4.7). The distribution was shoulder 9%, wrist 8.4%, hand 6.5%, hip 1.9%, knee 21.3%, ankle 4.5%, great toe 1.9%, cervical spine 11.6%, dorsal and lumbar spine 18.1%, and other 9.7%. The incidence of past and present musculoskeletal complaint was 57.4% (95% CI 5.8). The present disability in activities of daily living (mild to severe) was 23.9% (95% CI 5). Rheumatologic diagnoses were degenerative joint disease (neck 0.7%, lumbar spine 0.7%, knee 9.8%, multiple joints 2.1%, other 2.4%), low back pain 2.8%, sciatica 0.35%, 'tennis elbow' 0.7%, shoulder tenosynovitis 0.7%, and other tendonitis/tenosynovitis 1.1%. No inflammatory disorder was detected.

The pilot study helped to finalize the plans for the larger COPCORD study itself, designed to evaluate 10,000 persons. The ethnic distribution in the evaluated population closely resembles the estimated figure for Iran[2]. The number of interviewed persons was too low to estimate rheumatologic disorders correctly, and thus our results must be interpreted with caution.

FEREYDOUN DAVATCHI, MD, Professor and Head; AHMAD-REZA JAMSHIDI, MD, Associate Professor; ARASH TEHRANI BANI-HASHEMI, MD, Rheumatology Research Center, Shariati Hospital, Tehran University for Medical Sciences, Tehran 14114, Iran; JOHN DARMAWAN, MD, PhD, WHO-ILAR COPCORD Coordinator.

REFERENCES

1. Forghanizadeh J, Abhari R, Piroozian M, et al. Prevalence of rheumatic disease in Fasham. J Iran University Medical Sciences 1995;3:132-42.
2. Davatchi F, Shahram F. Epidemiology of Behcet's disease in Iran. In: Bang D, Lee ES, Lee S, editors. Behcet's disease. Seoul, Korea: Design Mecca Publishing Co.; 2000:589-91.

Corrections

Hunter DJ, York M, Chaisson CE, Woods R, Niu J, Zhang Y. Recent diuretic use and the risk of recurrent gout attacks: The Online Case-Crossover Gout Study. J Rheumatol 2006; 33:1341-5. Authors' degrees should be as follows: C.E. Chaisson, MPH; R. Woods, MPH; Y. Zhang, DSc. We regret the error.

Asherson RA, Cervera R, Shepshelovich D, Shoenfeld Y. Nonthrombotic manifestations of the antiphospholipid syndrome: Away from thrombosis? J Rheumatol 2006;33:1038-44. Prof. R.A. Asherson's departmental appointments should be as follows: Professor of Immunology, Division of Immunology, School of Pathology, University of the Witwatersrand, Johannesburg, South Africa. We regret the error.