Workshop Report

Outcome Measures for Sjögren's Syndrome, April 10-11, 2003, Bethesda, Maryland, USA

STANLEY R. PILLEMER, JANINE SMITH, PHILIP C. FOX, and SIMON J. BOWMAN

The goal of the Outcome Measures for Sjögren's Syndrome Workshop was to develop consensus on a core set of outcome measures for clinical trials in Sjögren's syndrome (SS). Before the workshop, participants were sent an outline of issues relevant to the development of outcome measures in SS and were invited to submit abstracts on candidate outcome measures. Participants were asked to consider the quality of data for each measure in published studies [number of subjects included, reproducibility and validity, sensitivity to change, variability (standard deviation, standard error or other), power, number of subjects likely to be needed in clinical trials, and placebo response rate].

INTRODUCTION AND OVERVIEW

During the first part of this one and one-half day meeting, the workshop process was reviewed and overviews were presented on approaches to development of outcome measures as well as the prevalence of the features of SS. Workshop presenters (listed* below) then reviewed data on outcome measures relevant to clinical trials in SS. Presentations were structured broadly into 3 main areas: subjective, objective, and additional measures. After each presentation, voting machines recorded participant responses to questions about the data and outcome measures that had been addressed. The responses were viewed on a screen as histograms immediately after voting, and the results were discussed further.

Dr. Simon Bowman presented an overview of the framework used to develop outcome measures in multisystem rheumatic diseases such as systemic lupus erythematosus (SLE), beginning with the efforts of the European Community working group, the Copenhagen Model, and the SS Outcome Measures Workshop held in Oxford, UK, in

From the Gene Therapy and Therapeutics Branch, National Institute of Dental and Craniofacial Research, National Institutes of Health (NIH), Bethesda; MedImmune, Gaithersburg, MD; the National Eye Institute, NIH, Bethesda; Department of Oral Medicine, Carolinas Medical Center, Charlotte, NC; and the Rheumatology Department, University of Birmingham, Birmingham, UK.

Supported by The Sjögren's Syndrome Foundation; the National Institute of Dental and Craniofacial Research; the National Eye Institute; the Office of Research on Women's Health; Allergan, Inc.; Daiichi Pharmaceutical Corporation; and MGI Pharma, Inc.

S.R. Pillemer, MD, Gene Therapy and Therapeutics Branch, NIDCR, NIH, and Director, Clinical Development, Immunological and Inflammatory Diseases, MedImmune, Gaithersburg, MD; J. Smith, MD, National Eye Institute; P.C. Fox, DDS, Visiting Scientist, Carolinas Medical Center; S.J. Bowman, FRCP, Rheumatology Department, University of Birmingham.

WORKSHOP CO-CHAIRS: S.R. Pillemer, MD, Gene Therapy and Therapeutics Branch, NIDCR, NIH, and Director, Clinical Development, Immunological and Inflammatory Diseases, MedImmune, Gaithersburg, MD; and S.J. Bowman, FRCP, Rheumatology Department, University of Birmingham, Birmingham, UK.

PARTICIPANTS: Elaine Alexander, MD, PhD, Consultant, Arena Pharmaceuticals, Inc., San Diego, CA, USA; Karsten Asmussen, MD, PhD, Department of Rheumatology, Rikshospitalet, Copenhagen, Denmark; Jane Atkinson, DDS, Dean of Clinical Affairs, University of Maryland, Baltimore, MD; Bruce J. Baum, Chief, Gene Therapy and Therapeutics Branch, National Institute of Dental and Craniofacial Research, NIH, Bethesda, MD, USA; Arthur Bookman, MD, FRCPC, Division of Rheumatology, Western Hospital, Toronto, ON, Canada; Steven Carsons, MD, Division of Rheumatology, Winthrop University Hospital, Mineola, NY; Troy Daniels, MS, DDS, Dean for Academic Affairs, School of Dentistry, San Francisco, CA; Gary Foulks, MD, University of Pittsburgh School of Medicine, Department of Ophthalmology, Pittsburgh, PA; Philip C. Fox, DDS, Visiting Scientist, Carolinas Medical Center, President-Elect, Sjögren's Syndrome Foundation, Cabin John, MD; Arthur Grayzel, MD, MACR, Mamaroneck, NY; John Harley, MD, PhD, Arthritis and Immunology Laboratory, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA; Lennart Jacobsson, MD, PhD, Rheumatology, Lund University, Malmö University Hospital, Malmö, Sweden; Roland Jonsson, DMD, PhD, Professor, Broegelmann Research Laboratory, University of Bergen, Bergen, Norway; Stuart Kassan, MD, Clinical Professor of Medicine, University of Colorado Health Sciences Center, Denver, CO; Albert Kingman, PhD, Division of Population and Health Promotion Sciences, NIDCR, NIH, Bethesda, MD, USA; Haralampos Moutsopoulos, MD, FACP, Department of Pathophysiology, National University Medical School, Athens, Greece; Ann L. Parke, MB, Department of Rheumatology, University of Connecticut Health Center, Farmington, CT; Maureen Rischmueller, MD, Staff Consultant Rheumatologist, Rheumatology Department, The Queen Elizabeth Hospital, Woodville, SA, Australia; Oliver D. Schein, MD, MPH, Wilmer Ophthalmological Institute, Johns Hopkins Hospital, Baltimore, MD; Janine Smith, MD, National Eye Institute, NIH, Bethesda, MD; Nehad Solomon, MD, Division of Rheumatology, Winthrop University Hospital, Mineola, NY; Elke Theander, MD, Rheumatology, Lund University, Malmö University Hospital, Malmö, Sweden; Susan Vitale, PhD, National Eye Institute, Bethesda, MD; Claudio Vitali, MD, Chief, Department of Internal Medicine and Rheumatology, Ospedale "Villamarina" -Piombino, Livorno, Italy; Frederick Vivino, MD, The Penn Center for Primary Care; Presbyterian Medical Center, Philadelphia, PA.

Address reprint requests to Dr. S.R. Pillemer, Gene Therapy and Therapeutics Branch, National Institute of Dental and Craniofacial Research, 10 Center Drive, Rm IN113, MSC 1190, National Institutes of Health, Bethesda, Maryland 20892. E-mail: PillemerS@medimmune.com

which a framework for developing consensus on measures for clinical trials in SS was developed[1].

Dr. Stanley Pillemer presented an overview of the outcome measures developed for rheumatoid arthritis (RA) clinical trials, which focus more on the more frequently involved manifestations (e.g., joints). In both RA and SS, pulmonary and ocular disease, vasculitis, and neuropathy may occur. He suggested that the experience in RA may serve as a model for the development of outcome measures for SS, based on more prevalent features of SS, such as exocrine involvement and laboratory measures, rather than the comprehensive multisystem approach used in SLE.

Dr. Arthur Bookman presented the results of a cross-sectional evaluation of 323 consecutive patients that had been prescreened for objective evidence of dry eyes, evidence of dry mouth on inspection, or parotid swelling. Of these, 169 satisfied the American-European classification criteria for SS[2]. SS patients had more ocular damage, higher minor salivary gland biopsy scores, greater prevalence of lymphomas, and more frequent laboratory abnormalities. Subjective measures correlated poorly with the degree of salivary or lacrimal flow. It was acknowledged, however, that subjective distress is a critical variable for the patient and will likely remain an important consideration in efficacy trials for therapeutic interventions.

SUBJECTIVE MEASURES

Sicca symptoms

*Simon Bowman**: Sicca (dryness) features are the hallmark of SS and a major cause of disability and reduced quality of life (QOL) in patients with this condition. There is a growing realization that symptoms, rather than flow rates, matter more to patients and hence are important in clinical trials[3-6].

Most dry eye questionnaires were developed to screen populations in surveys. The most studied is the Ocular Surface Disease Severity Index (OSDI), which has been used in clinical trials and fulfills the majority of validation criteria in terms of discriminating dry eye patients and correlating with dry eye severity (see below)[3].

Ad hoc dry mouth measures have also been used in clinical trials[4]. The Xerostomia Inventory has been developed to assess oral dryness[5] and validated in terms of face, content, and criterion validity in individuals with symptomatic xerostomia. There has also been interest in developing more global oral health related quality of life (HRQOL) questionnaires.

Bowman and coworkers recently developed the Sicca Symptoms Inventory (SSI) incorporating questions derived from patient interviews and published sicca symptom questions, and validated the SSI in terms of face and construct validity and reliability[6]. The sensitivity of these questionnaires in detecting improvements in clinical trials remains to be determined.

QOL instruments applied to the visual effects of eye disease and the use of questionnaires in dry eye diagnosis

*Anthony Bron, Angus Warwick Turner, Gary Foulks**: Interest grows in measuring the impact of dry eye symptoms on QOL in SS using psychometrically validated questionnaires. Mangione, *et al*[7] developed a 51-item HRQOL instrument, the National Eye Institute Visual Function Questionnaire (NEI-VFQ), based on the assessment of QOL indices in a large sample of patients with reduced vision from various causes. The instruments showed an internal consistency of between 0.66 and 0.94, with a subscale test-retest reliability of between 0.68 and 0.91. A reduced, 25-item instrument is also available (NEI-VFQ 25).

The OSDI is a 12-item questionnaire designed to assess the impact of dry eye symptoms on vision-related performance. The questionnaire items were generated from the comments of patients in clinical trials, suggestions of investigators, and material derived from other QOL instruments[3]. Its reliability and validity have been assessed in comparison to other questionnaires including the Medical Outcome Study Short Form-12 health survey (SF-12), the NEI-VFQ 25, and the McMonnies' questionnaire. The OSDI is accepted as reliable and valid for assessment of dry eye related symptoms. It has not, however, been tested in primary SS.

Assessment of fatigue in primary SS

*Elke Theander**, *Lennart Jacobsson*: A number of fatigue measurement instruments are available. They vary with respect to validity, reproducibility, responsiveness to change, and ease of use. A structured Medline literature search identified the visual analog scale (VAS) as the most commonly used fatigue measure in SS. It was used in 5 studies, 3 of which were intervention trials. The vitality score of the SF-36 was used in 4 SS studies as a measure of fatigue, but was not used in intervention trials. So although the SF-36 vitality score is widely used and is valid and discriminative, no data exist on its responsiveness in SS. The Chalder Fatigue Scale[8] was used in 3 studies on SS and in several others on SLE; none was an intervention trial; the Multidimensional Fatigue Instrument (MFI)[9] was used in 2 Dutch SS studies, but none was a longterm followup or intervention study. Other instruments, such as the Profile of Mood States, Nottingham Health Profile, Fatigue Severity Scale, Piper's fatigue instrument, and the Dutch Fatigue Scale, either have been used in only a single study or have not been used at all in SS. The Profile of Fatigue and Discomfort (PROFAD) is a newly developed instrument designed specifically for SS[10]. A Swedish version is being validated at the University of Lund, Malmö. The design of the PROFAD is a 16-item, 8-point scale, analyzing 6 different facets of fatigue, belonging to 2 fatigue domains (somatic and mental). Since it has been derived from specific symptoms of SS patients, it may be the most attractive to use in future studies together with SF-36 (vitality) and VAS.

HRQOL issues in SS patients

*Susan Racine, Ann Parke**: Data were presented from a study of 48 patients suggesting SS is associated with significantly higher levels of depressive symptoms, as based on the Center for Epidemiologic Studies Depression scale, than seen in 84 patients with early onset RA and 64 healthy control subjects. Dry mouth symptoms independently contributed to depressive symptoms among women with rheumatic diseases and were also common among early RA patients. Given that depressive symptoms influence QOL, this suggests that further studies of the relationships among depressive symptoms, dry mouth, and HRQOL are warranted.

OBJECTIVE MEASURES

Oral objective measures: sialometry in clinical trials

*Philip Fox**: Sialometry is the most widely used measure of salivary performance in clinical trials. When carefully performed by experienced investigators, the reproducibility is high but the result is technique-dependent.

Issues to consider include effects of concomitant medications on salivary function, unstimulated versus stimulated saliva collection, and whole saliva versus collection of saliva from individual glands. Whole unstimulated salivary flow is the simplest method and appeared to be the most reproducible for use in clinical trials. Collection of secretions from individual salivary glands is important when sialochemical studies are to be performed The level to which salivary function must increase to improve subjective dryness remains unclear. The use of percentage increase in saliva production as an outcome measure in populations with severe impairment of saliva production is problematic, since a very small absolute increment may represent a large percentage increase.

It is difficult to define a clinically meaningful level of improvement in saliva production. Most studies have therefore analyzed data in terms of statistically significant differences in saliva production, between treatment groups comparing baseline to postintervention levels.

Sialochemistry in clinical trials

*Jane Atkinson**: Elevated sodium and chloride ion concentrations are typical in SS and have potential diagnostic value[11]. Activated lymphocytes in the salivary glands produce autoantibodies that can be measured in saliva, including IgA and IgM-rheumatoid factor, anti-SSA, and anti-SSB. Other inflammatory proteins that are increased in SS saliva include lactoferrin, lysozyme, the cell surface protein α_2-microglobulin, and cytokines such as interleukin 6 (IL-6), IL-10 and tumor necrosis factor-α (TNF-α). One study has suggested that salivary IL-6 levels may be a useful measure in clinical trials, and 2 studies have suggested that salivary IgM concentrations may decline with treatment (bromhexine and N-acetylcysteine). While there is currently insufficient evidence that sialochemical measures are useful as outcome measures in clinical trials, further investigation is warranted.

Salivary scintigraphy as an outcome measure in SS

*George Hermann, Frederick Vivino**: Dynamic imaging after intravenous injection of $^{99}mTcO_4^-$ allows estimation of functioning salivary gland uptake and secretion into the oral cavity. Semiquantitative assessment of salivary gland dysfunction using scintigraphy is useful in the diagnosis and classification of SS and potentially also as an outcome measure.

Data were presented on 413 visitors to a xerostomia clinic and 30 healthy, non-xerostomia controls (unstimulated salivary flow > 0.3 ml/min) who underwent stimulated salivary scintigraphy. In all, 187 patients met modified American-European criteria for SS[2]. Controls showed vigorous uptake[12]. In all, 136 SS patients showed uptake or secretory dysfunction (sensitivity 73%) characterized by one of 3 patterns: (I) deficient uptake with deficient or indeterminate secretion (n = 95, 70%), (2) deficient uptake but normal secretion (n = 15, 11%), or (3) normal uptake but deficient secretion (n = 25, 19%). Unevaluable secretion due to lack of uptake occurred in 63 studies.

Sixty-one patients were then empirically treated with pilocarpine (10-30 mg/day) or the Salitron^{TM} electrostimulation system (minimum 6 V tid) for 3 or more months[13]. The 28 responding patients increased their unstimulated salivary flow rates by a mean of 0.3 ± 0.26 ml/min compared with the mean change of 0.02 ± 0.05 ml/min of the 33 nonresponders (Mann-Whitney $p < 0.001$). Mean scintigraphic scores of nonresponders significantly exceeded those of responders for all 3 scintigraphic parametric strata: uptake (10.6 vs 7.3, $p < 0.03$), release (11.9 vs 7.6, $p < 0.001$), and combined uptake/release (22.5 vs 15.3, $p < 0.001$). Receiver operating characteristic analysis produced areas under the curve of 0.73 for uptake, 0.77 for release, and 0.75 for combined uptake/release scores. In 6 initial treatment responders, subjective xerostomia worsened over time periods of 21-84 months despite adequate therapy, and mean combined uptake/release scintigraphy scores increased from 15.2 to 22.7 in followup scans. These preliminary data suggest that scintigraphic assessment may serve as an outcome measure in future clinical trials and as a surrogate marker of glandular disease progression in longterm studies.

Ophthalmic outcome measures in clinical studies of treatment for SS

*Janine Smith**: Articles on ophthalmic outcome measures used in clinical trials in patients with SS reported in English in the past 10 years were identified through a literature review using PubMed search engine. Twenty-seven clinical trials were identified: 15 were randomized and placebo-controlled and 12 were open-label and/or included no control

group. Of the clinical studies reviewed, 5 randomized, placebo-controlled clinical trials of treatments of SS showed a statistically significant treatment effect over placebo for an ocular outcome measure: bromhexine, cevimeline, gamma-linolenic acid supplementation (Efamol), pilocarpine, and N-acetyl-cysteine. However, only the trials of the secretagogues[3,14] had a clearly defined ophthalmic primary outcome measure, which was in each case a global assessment of ocular dryness.

Various ophthalmic outcome measures have been included in clinical trials of SS. The most common measures were: Schirmer test (n = 20), surface vital dye staining (n = 18), assessment of symptoms of dry eye (n = 17), tear breakup time (n = 10), tear lysozyme (n = 5), artificial tear use (n = 3), conjunctival impression cytology (n = 3), and tear lactoferrin (n = 1).

In future studies, the mechanism of action of the therapeutic modality and the responsiveness of the outcome measure should be considered in the study design process. The development of composite measures of ocular surface disease could improve the ability to measure clinically relevant changes in disease activity and response to treatment. There is a need for the identification and standardization of valid, reproducible, responsive ophthalmic outcome measures for use in clinical trials of SS.

Estimating ocular surface damage by grading ocular surface staining

*Anthony Brown, Angus Warwick Turner, Gary Foulks**: Van Bijsterveld (1969) investigated the sensitivity and specificity of 3 diagnostic tests for dry eye — Schirmer I, Rose Bengal, and lysozyme lysis — in 550 controls and 43 patients with sicca syndrome. Grading of surface damage used a scale of 0-3 for each of 3 ocular surface compartments in the interpalpebral aperture, giving a scale range of 0-9. With a staining intensity score limit of 3.5, the probability of misclassification of a patient was 5%, and of a person from the control group, 4% (sensitivity 95%; specificity 96%).

The Oxford grading scheme (1984) quantifies ocular surface damage in patients with dry eyes by comparison of surface staining with grading charts simulating the distribution of surface damage usually encountered in dry eye. The maximum score for each panel is 5, and therefore the range of grades is from 0 to 15.

Although the reliability depends on the consistency with which the observer compares their findings with a grading chart, a study of intra- and interobserver repeatability in Oxford suggested that the scoring system has good reproducibility. (Hardman-Lea; Association for Eye Research meeting 1986). The National Eye Institute (NEI)/Industry scheme was devised at the NEI/Industry meeting of 1994[16] and has been used in clinical trials. The grading charts depict the cornea and conjunctiva, and the observer is instructed to grade the cornea and the 2 exposed conjunctival regions. Grading is conducted in 5 corneal zones and 4 conjunctival zones. Each zone is graded from 0 to 3, based on reference to 4 boxes in which the number density of the portrayed dot clusters is scaled (one box is empty). The range of grades is from 0 to 33 (15 + 9 + 9). Further work is needed to ensure standardization of use in clinical trials.

Photographic digitization of the staining pattern, if validated, would offer a major advantage not only for clinical trial use, but also in rheumatology clinics, where an ophthalmologist able to accurately grade surface staining may not be available.

Identification of predictive biomarkers in SS clinical trials: a metaanalytic approach

Nehad Soloman, Mickel Khlat, Anang Modi, Martin Feuerman, Steven Carsons*: The literature of the past 40 years was reviewed comprehensively to compile all SS clinical trials that included clinical and serological outcomes. Studies included for analysis required at least 10 subjects, commonly available biomarkers, and the use of a criteria set for diagnosis. Eight studies met these requirements. Statistical analysis was carried out on both serologic and clinical outcomes using paired t tests derived from weighted means and standard deviations at baseline and endpoint, as well as by formal metaanalysis utilizing standard effect sizes. Biomarkers studied included IgG, IgM, IgA, SSA, SSB, RF IgG, RF IgM, erythrocyte sedimentation rate, and C-reactive protein. Clinical outcomes included both subjective and objective measures. A statistically significant decrease in IgG and SSB post-treatment with various modalities was noted. IgG was evaluated in all patients (n = 1.64, p = 0.0013) SSB decreased post-treatment and had the largest effect size (–0.61) despite the small number of patients studied (n = 29, p = 0.0158). Out of the total number of patients studied (n = 164), a small subset of patients (n = 38) were treated with hydroxychloroquine and showed a statistically significant decrease post-treatment only in IgG (p = 0.02). Only one hydroxychloroquine study examined SSB (n = 10, $p < 0.05$). Of the objective measures studied, salivary flow had a significant improvement post-treatment (n = 109, p = 0.0001). All subjective measures of outcome displayed a decrease in the percentage of patients with symptoms post-treatment. Dry eyes (81% vs 69%, p = 0.08), dry mouth (64% vs 46%, p = 0.006), and arthralgias (80% vs 43%, p = 0.001). Pearson's correlation coefficient showed an inverse linear relationship between IgG and salivary flow (r = –0.58) that trended towards significance (p = 0.07). Of the biomarkers studied, therefore, total quantitative immunoglobulins, in particular IgG, may serve as a good measure of outcome in SS. SSB may be promising as well, but larger studies are needed to assess the specificity of these markers in SS.

Laboratory measures for longterm outcomes, including lymphoma and mortality

*Haralampos Moutsopoulos and Steve Carsons**: The risk of lymphoma development in patients with primary SS is 40 times higher than in the general population. Three sequential studies from Greece clearly substantiated the increased incidence of lymphoma and its association with increased mortality in patients with primary SS; they also specify adverse predictors of long-term outcomes. The presence of parotid gland enlargement, palpable purpura, low C4 levels, and mixed monoclonal cryoglobulinemia at first visit were shown to be adverse prognostic factors that distinguish patients at high risk for the development of lymphoproliferative disorders.

Increased mortality is observed in patients with primary SS compared to the general population (mortality ratio: 1.15). One to 5 deaths of patients with primary SS are attributed to lymphoma. The presence of the adverse prognostic factors (palpable purpura, low C4 levels) is strongly correlated with the increased mortality. By contrast, the mortality rate of patients with primary SS is identical to that of the general population after the exclusion of high-risk patients with SS. The data from the Greek studies revealed that simple clinical and laboratory variables can be excellent predictors of the adverse long-term outcomes in SS[17].

Labial salivary gland (LSG) histopathology as a measure of therapeutic or disease outcome in patients with SS

*Troy Daniels**: LSG histopathology has routinely been used since the 1970s as a diagnostic criterion for the salivary component of SS. However, use of this tissue as an outcome measure has been very infrequent, mainly because it requires repetition of an invasive procedure. There have been 2 studies in which therapeutic outcomes of SS have been assessed with repeated LSG biopsies. In the first retrospective study, second biopsy specimens from 3 of 5 untreated patients and from 4 of 5 patients treated with prednisone, chlorambucil, or x-ray exhibited increased lymphocytic infiltration, while 2 of 4 patients treated with cyclophosphamide exhibited decreased infiltration[18]. In a second randomized, double-blind study, no significant changes were found in second biopsy specimens from any of 22 patients receiving prednisone or piroxicam[19]. Other studies that have assessed the progress of SS over time in patients not receiving treatment have shown progressive salivary lymphocytic infiltration on repeated LSG biopsy in a substantial portion of patients. However, all the studies have used different diagnostic criteria to establish the presence of SS in the study patients and different means of histologically assessing the LSG biopsy specimens.

ADDITIONAL OUTCOME MEASURES

Lower urinary tract symptoms and sleep disturbance in primary SS

*Maureen Rischmueller**: Rischmueller and coworkers recently demonstrated in a controlled, cross-sectional study that urological symptoms (primarily urgency) and daytime sleepiness are more severe in female patients with primary SS; as well, a trend towards increased fatigue in SS patients was noted[20]. This study used the American Urology Symptom Index-7 to quantify urological symptoms, the Epworth Sleepiness Scale (ESS) to assess daytime somnolence, and the FACIT-F scale to assess fatigue. These results might suggest that urological symptoms are an under-recognized feature of SS and that fatigue may in some cases be secondary to an underlying sleep disorder leading to daytime sleepiness.

Increased daytime sleepiness in SS might result from repeated waking due to ocular or oral discomfort, or possibly from obstructive sleep apnea associated with altered surface tension in dry upper airway mucosal surfaces. Alternatively, both urological symptoms and sleep disturbance might reflect muscarinic receptor/autonomic dysfunction. The Autonomic Symptom Profile is a recently developed, self-administered, 169-item questionnaire concerning different aspects of autonomic symptoms that shows promise in assessing autonomic symptoms in clinical trials and epidemiological studies, and may be of value in SS.

ECLAM-SS: a modified version of the ECLAM to evaluate disease activity in patients with SS

Claudio Vitali and Nicoletta Del Papa*. The European Consensus Lupus Activity Measurement Scale (ECLAM) has been developed by a European consensus group as a reliable and sensitive method of measuring disease activity in SLE[21]. A modified version, the ECLAM-SS, is now proposed, for use in primary SS. As well as deleting some items that are not relevant to SS, additional items on major salivary gland swelling and cryoglobulinemia have been added to improve its potential validity and sensitivity. Large multicenter studies are now needed to validate the ECLAM-SS in assessing disease activity in SS.

Outcome measures in primary SS — experience from randomized clinical trials (review)

*Karsten Asmussen**: A search of Medline and the Cochrane Controlled Trial Register from 1966 to 2002 identified 34 randomized controlled trials testing the efficacy of 15 drugs on more than 2000 patients with SS. Over 100 different measures were reported for assessing SS. The most frequent related to the sicca features and included symptom scores, specific complaints, global exocrine scores, exocrine flow rates, and signs. The majority of these showed sensitivity to change. Very few studies addressed non-exocrine disease manifestations. Some studies used markers like immunoglobulin levels, autoantibodies, cytology, and serial labial gland biopsy. In many trials, response to placebo was sig-

nificant, averaging 10-25% for symptomatic measures and glandular output.

In summary, a growing number of clinical trials have put focus on SS as a treatable disorder. A diversity in assessment tools in these trials has hampered comparison across studies. Consensus on a core set of criteria for monitoring this important autoimmune disorder is needed.

VOTING RESULTS: PROPOSED OUTCOME MEASURES

The results of the voting were displayed after each presentation, leading to considerable discussion. During the evening of the first day results were summarized, and outcome measures were ranked from those with the highest to those with the lowest support. The Workshop Chairs decided a priori that measures supported by more than 50% of the participants were to be put to a vote on the morning of the second day as candidates for a core set of outcome measures. Measures with > 6% to 50% support would be voted on for inclusion as secondary outcome measures; there would also be a discussion of the merits of measures that received less than 6% of participant support.

More than 90% of participants voted in favor of the listed items as a whole for each of the 3 categories (Table 1). During deliberations, participants strongly expressed the view that the outcome measures selected should serve to enhance further data collection on relevant outcomes in clinical trials of SS. It was evident that considerable information was available in some areas relevant to SS outcomes, for example, many studies on outcome measures pertaining to dry eyes, that could not adequately be presented in the time allocated during the workshop. In other areas, information was insufficient to determine whether certain outcomes could be useful or should be ruled out.

The group as a whole was disparate and included ophthalmologists, dentists, rheumatologists, statisticians, basic scientists, and clinicians. This raised concerns of the validity of voting in specialized areas outside of their own expertise. Participants therefore recommended that working groups in ophthalmology, dentistry, rheumatology, and other areas be established to advise which would be the most valid and useful measures in those areas. In addition, the participants recommended that outcome measures selected should, as far as possible, be included in all future clinical trials in SS in order to allow for an iterative process as the field develops.

A broader bibliography for this workshop is available online at: www.sjogrens.org/research/outcomemeasures.html

REFERENCES

1. Bowman S, Pillemer S, Jonsson V, et al. Revisiting Sjögren's syndrome in the new millennium: Perspectives on assessment and outcome measures. Report of a workshop held on 23 March 2000 at Oxford, UK. Rheumatology Oxford 2001;40:1180-8.
2. Vitale 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.
3. Schiffman RM, Christianson MD, Jacobsen G, Hirsch JD, Reis BL. Reliability and validity of the Ocular Surface Disease Index. Arch Ophthalmol 2000;118:615-21.
4. Vivino FB, Ai-Hashimi I, Khan Z, et al. Pilocarpine tablets for the treatment of dry mouth and dry eye symptoms in patients with Sjogren's syndrome. Arch Intern Med 1999;159:174-81.
5. Thomson WM, Chalmers JM, Spencer AJ, Williams SM. The Xerostomia Inventory: a multi-item approach to measuring dry mouth. Comm Dent Health 1999;16:12-7.
6. Bowman AJ, Booth DA, Blatts RG, Field A, Rostron J and the UK Sjögren's Interest Group. Validation of the Sicca Symptoms Inventory for clinical studies of Sjögren's syndrome. J Rheumatol 2003;30:1259-66.
7. Mangione CM, Lee PP, Pitts J, Gutierrez P, Berry S, Hays RD. Psychometric properties of the National Eye Institute Visual Function Questionnaire. Arch Ophthalmol 1998;116:1496-504.
8. Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale. J Psychosom Res 1993;37:147-53.
9. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res 1995;39:315-25.
10. Bowman AJ, Booth DA, Platts RG and the UK Sjögren's Interest Group. Measurement of fatigue and discomfort in primary Sjögren's syndrome using a new questionnaire tool. Rheumatology Oxford 2004;43:758-64.

Table 1. Voting results on proposed components of a set of core and secondary outcome measures. Data are presented as the percentage of participants (see list) voting to include the component.

Core (> 50%)	Secondary (> 6–50%)	Additional data required (≤ 6%)
Sicca symptoms (oral, ocular)	Composite activity, 48%	Sleep disturbance
Sicca objective oral	C4, 39%	Labial salivary gland biopsy
Unstimulated whole SF	Stimulated SF, 39%	Individual gland SF
Sicca objective ocular	Composite damage, 37%	Corneal sensitivity
Dye score	Anti-La antibody, 36%	Impression cytology
Schirmers: ± anesthetic	Tear breakup time, 17%	Tear chemistry
Health related quality of life	C3, 15%	Lower urinary tract symptoms
Laboratory measures	Scintigraphy, 13%	
Total IgG level	Sialochemistry, 10%	
Fatigue		

SF: salivary flow.

11. Talk WW, Vissink A, Stegenga B, Bootsma H, Nieuw Amerongen AV, Kallenberg CG. Sialometry and sialochemistry: diagnostic tools for Sjogren's syndrome. Ann Rheum Dis 2001;60:1110-6.
12. Hermann GA, Vivino FB. Abnormal scintigraphic patterns of uptake and secretion in Sjogren's syndrome [abstract]. Arthritis Rheum 2000;43 Suppl:S306.
13. Vivino FB, Hermann GA. Scintigraphic predictors of response to treatment in Sjögren's syndrome [abstract]. Arthritis Rheum 2000;43 Suppl:S408.
14. Petrone D, Condemi JJ, Fife R, Gluck O, Cohen S, Dalgin P. A double-blind, randomized, placebo-controlled study of cevimeline in Sjögren's syndrome patients with xerostomia and keratoconjunctivitis sicca. Arthritis Rheum 2002;46:748-54.
15. Bron AJ. The Doyne Lecture. Reflections on the tears. Eye 1997;11:583-602.
16. Lemp MA. Report of the National Eye Institute/Industry workshop on Clinical Trials in Dry Eyes. CLAO J 1995;21:221-32.
17. Ioannidis JP, Vassiliou VA, Moutsopoulos HM. Long-term risk of mortality and lymphoproliferative disease and predictive classification of primary Sjogren's syndrome. Arthritis Rheum 2002;46:741-7.
18. Anderson LG, Cummings NA, Asofsky R, et al. Salivary gland immunoglobulin and rheumatoid factor synthesis in Sjogren's syndrome. Natural history and response to treatment. Am J Med 1972;53:456-63.
19. Fox PC, Datiles M, Atkinson JC, et al. Prednisone and piroxicam for treatment of primary Sjogren's syndrome. Clin Exp Rheumatol 1993;11:149-56.
20. Walker J, Gordon T, Lester S, et al. Increased severity of lower urinary tract symptoms and daytime somnolence in primary Sjögren's syndrome. J Rheumatol 2003;30:2406-12.
21. Vitali C, Bencivelli W, Isenberg DA, et al. Disease activity in systemic lupus erythematosus: report of the Consensus Study Group of the European Workshop for Rheumatology Research. II. Identification of the variables indicative of disease activity and their use in the development of an activity score. The European Consensus Study Group for Disease Activity in SLE. Clin Exp Rheumatol 1992;10:541-7.