Practical Implications of Noncompliance in Randomized Clinical Trials for Temporomandibular Disorders

Coralyn W. Whitney, PhD
Research Assistant Professor
Department of Biostatistics
School of Public Health
Department of Dental Public Health Sciences
School of Dentistry

Samuel F. Dworkin, DDS, PhD
Professor
Department of Oral Medicine
School of Dentistry
Department of Psychiatry and Behavioral Sciences
School of Medicine

University of Washington
Seattle, Washington

Correspondence to:
Dr Coralyn W. Whitney
Department of Biostatistics
Box 358223
University of Washington
Seattle, Washington 98195–8223

Randomized clinical trials are recognized as providing the most rigorous evidence of treatment efficacy. For temporomandibular disorders, randomized clinical trials have been used to evaluate the efficacy of low-cost occlusal appliances or the adjunct use of cognitive behavioral interventions. However, noncompliance with treatment regimens and losses to follow up are common randomized clinical trial protocol violations that compromise the desired rigor of the trial. At times it is not clear to the investigator how to deal with these issues during the trial and at the data analysis phase. Often treatment efficacy is based on the compliant subjects, subjects who may no longer represent randomized groups or yield the desired "fair" estimate of treatment efficacy. This study focuses on management of compliance issues, the description and collection of data needed to obtain a more accurate assessment of treatment efficacy, and results particularly relevant to actual clinical practice and patient care decisions. These are applied to a randomized clinical trial evaluating the efficacy of a cognitive-behavioral intervention for temporomandibular disorders.
J OROFACIAL PAIN 1997;11:130–138.

key words: noncompliance, randomized clinical trials, temporomandibular disorders, intention-to-treat

A recent review of research on temporomandibular disorders (TMD) noted that less than 5% of the references regarding therapy were randomized controlled trials.[1] The review also noted that, as with medicine, the health care industry may demand increases in such evidence-based decisions for patient care in the field of dentistry. Indeed, randomized clinical trials (RCTs) are designed to provide the most rigorous and fairest assessment of a treatment's efficacy. However, this rigor arises from the extensive planning, resources, and management needed to randomly select from the target population and to adhere to study protocol. The adherence to protocol prevents biases from compromising the validity and generalizability of the study.[2–7] By minimizing the occurrence of these biases, the resultant research would be improved, and the validity and generalizability of reported findings to the patient population and the health care industry would be better assessed.[8] In the field of orofacial pain and TMD, researchers have recognized the need for such evidence-based decisions through increased evaluations of treatments based on sound scientific principles (eg, RCTs), thus enabling better and more effective treatments for patients to be identified and incorporated by clinicians into their practices (editorials).[9–11]

In reality, adherence to study protocol is a demanding effort, even in the best of studies, whether randomized or not. Protocol violations occur, and noncompliance is to be expected and planned for.[6] If such violations vary by treatment group assignment (eg, those who cannot tolerate a soft splint drop out of the study) or are correlated with the outcome or prognostic measures of interest[12] (eg, those experiencing no benefit discontinue treatment), the fairness of the assessment has been compromised. These violations do not necessarily invalidate the results of a study; however, depending on the degree of occurrence and how these are analyzed, they certainly have the potential to do so. The extent of these protocol violations needs to be described and considered when evaluating, interpreting, and reporting the results of a study.[13] Two ways to control factors influencing the final analysis and interpretation of such trials are *(1)* the report and management of noncompliance in the trial, and *(2)* the use of the intention-to-treat principle to analyze the data.

Noncompliance in Clinical Trials

Noncompliance is a catchall term that encompasses an array of possible research protocol violations that constitute a significant proportion of the potential sources of biases. Haynes[14] defined compliance as "the extent to which a person's behavior coincides with medical or health advice" (in terms of using splints, taking medications, following diets, or executing lifestyle changes). This definition is sufficiently broad to subsume the cascade of events that can threaten the integrity of an RCT. Specifically, failures in compliance, defined as noncompliance with study protocol, can begin at the level of *(1)* initial failure to begin participation after random assignment to a treatment or control condition; *(2)* nonadherence—the failure to adhere to study treatment protocol (eg, not taking required medications or not completing required behavioral interventions such as homework assignments or group tasks); *(3)* loss to follow up at any or all data collection points; and *(4)* violation of any other aspect of the study protocol.[12,13,15] Thus, the term *compliance* refers to satisfactory execution of all stages of the research protocol, in contrast to *adherence*, which is often more narrowly defined as satisfactory execution of specific study treatments (eg, using splint or taking prescribed medications) within the overall study protocol.

In RCTs, it is important to recruit compliers and promote compliance. There are several factors that influence the extent of compliance:

1. Recruitment method: mail; telephone; in person; at clinic or office
2. Method of presentation: written materials; video; clinic personnel
3. Clinical setting: friendliness; organization; helpfulness; business
4. Treatments studied: medical; behavioral; combination
5. Duration of trial: weeks; months; years
6. Demands of trial: appliance use; medicines; office visits; outside-of-office visits or sessions; personal diaries; data collection; questionnaires
7. Belief system of participant: medical versus behavioral; personal value of study

Recruitment is optimum when conducted where potential subjects feel comfortable, are able to ask questions, and have their time schedules respected. Adequate explanation of the trial and thorough discussions of the trial's treatments and the requirements of the study are key, so that potential participants can make better assessments regarding their ability to comply. By default, longer clinical trials need to dedicate more time to maintenance of compliance as well as to the demands of the trial. Collecting "necessary" study data to answer the efficacy question is important. Physical examinations make the participants feel they are getting something tangible out of the study. However, extensive data collection can become burdensome to participants. Thus, when the trial is designed, choices may be available to maximize compliance without jeopardizing study integrity.

In studies of treatment efficacy, in spite of efforts to maximize compliance, investigators tend to exclude the noncompliers from the final analyses, basing conclusions on the compliers. Upon initial consideration, removing noncompliers from final data analyses may seem justified because, after all, only the remaining compliers receive the complete intervention of interest. However, with the array of aforementioned reasons for considering a subject as a noncomplier, such an analysis strategy limits the assessment of treatment efficacy and generalizations.

The RCT methodologists argue that once any subgroup (eg, noncompliers) has been excluded from the analysis, it cannot be assumed that the results based on the remaining subjects will yield a statistically or logically "fair" estimate of treatment efficacy, even if compliers and noncompliers appear similar on measured factors.[16,17] The reason is simple: the remaining subjects no longer represent "randomized" subgroups, and the statistical

foundation on which the RCT was formulated is violated.[18,19] As the foundation of the integrity of the RCT, randomization provides the basis for the assumption that the different treatment groups are similar at the beginning of the study on *(1)* the factors measured in the study (eg, outcome, palpation pain, and independent variables such as duration of symptoms); *(2)* those factors not measured, thus left to vary randomly, such as intelligence or vital signs; *(3)* unmeasurable factors, such as motivation for participating, which are also left to vary randomly across treatment groups. Under randomization, any observed differences at the beginning of the trial would be the result of chance alone, and not selection bias. Thus, exclusion of noncompliers from statistical analysis compromises the statistical validity of the trial, defeating the desired rigor and potentially limiting the acceptance of the evidence.

Obviously, if there are no follow-up data on some participants, these subjects cannot be included in the efficacy (ie, outcome) analysis. However, several questions then arise that need to be addressed, perhaps based on data available from the trial. Was the loss to follow up "random," or did it affect the treatment and control groups differentially? What could be said about those who attend all follow-up visits but do not adhere strictly to their treatment regimen? As Fleiss stated, "Is poor compliance (or participation) a surrogate for, or a consequence of, poor response to treatment?"[20] These questions need to be addressed in the assessment of the validity and generalizability of the results.

Intention-to-Treat Principle

The intention-to-treat principle, an established concept in clinical trial methodology, requires the inclusion of all subjects in analyses, regardless of adherence to treatment protocol.[21–23] The intention-to-treat principle requires a comparison of subjects intended for one treatment with those intended to receive another treatment.[18] According to the intention-to-treat principle, then, noncompliers are included in analyses according to the groups to which they were initially randomized—even if, through noncompliance, they did *not* receive the randomly assigned treatment. Although this may sound counterintuitive, such an approach obviates the need to define noncompliance, eliminating the temptation to either develop an "optimum" (in reality, post-hoc) definition of noncompliance that provides the best-case scenario results,[6,7,24] or, alternatively, trying to be conservative and providing the worst-case scenario results. Intention-to-treat analysis may actually better reflect how effective the treatment would be in clinical practice, where the treatment environment is not under the same degree of control as may be true for more ideal research conditions. Nevertheless, this principle maintains the statistical validity of the RCT by not allowing subgroups to be excluded from the analyses, exclusions that can introduce uncontrolled bias.

Thus, data on protocol violations should be collected as a routine part of the research so that both participants and refusals can be described. Participants who did not comply with the treatment protocol, together with their reasons for noncompliance are reported.[15,23] At a minimum, partial follow-up assessments on the noncompliers should be collected.

To demonstrate how the intention-to-treat principle may be applied to assess the extent of noncompliance and its influence on the estimate of treatment efficacy, this principle, and related analyses, were applied to data resulting from an RCT evaluating the efficacy of a cognitive-behavioral intervention introduced early in the course of treatment for chronic TMD pain.[25] In particular, the following were studied: the potential for differential refusal rates; the impact of noncompliance through baseline comparisons of treatment groups and of compliers to noncompliers; and the impact of noncompliance on treatment efficacy by the inclusion of noncompliers in the analyses.

Materials and Methods

Study Design and Definitions

The design of the TMD clinical trial and the definition of terms and groups for analyses provide the foundation for demonstration of the applications of the principles and methods of dealing with noncompliance in an RCT.

Subjects were recruited for the RCT from patients who were experiencing pain and related symptoms of TMD, and who were seeking treatment at either the Temporomandibular Joint Clinic of Group Health Cooperative of Puget Sound, WA, or the Orofacial Pain and Dysfunction Clinic at the University of Washington, School of Dentistry, Seattle, WA. Those who agreed to participate were randomized either to the experimental condition, consisting of a brief, two-session cognitive-behavioral (CB) intervention just prior to beginning usual dental treatment for their TMD, or to the active control group, which received usual dental treatment alone (UT). The CB intervention was con-

ducted in groups and scheduled in the late afternoon or evening. Participants were interviewed and clinically examined at baseline, and at 3 months and at 1 year after intervention. Postintervention assessment occurred at the end of the second CB session for the experimental group and at a comparable time point for the control subjects. At that time, a self-administered follow-up questionnaire was completed by each subject of both groups. Additional details of the intervention, eligibility criteria, and study design are provided elsewhere.[25]

All eligible subjects were classified either as refused to participate ("refusals") or agreed to participate and were randomized into the study. For the present study, subjects who were randomized were further classified as compliers or noncompliers. Subjects who were randomized and completed their treatment protocol and all nonabbreviated follow ups are referred to as *compliers*. Randomized subjects not meeting these compliance criteria are referred to as *noncompliers*. Thus, for the experimental group, those who missed either or both CB intervention sessions are be classified as *noncompliers* and also have not completed a postintervention follow-up questionnaire. A participant not completing any of the follow-up visits is classified as a *noncomplier*. It is assumed that subjects in both treatment groups complied equally with their usual course of dental treatment.

Additional data were collected for the refusals. At the time of recruitment, those patients who refused to participate were asked why they refused, and they were asked to answer three questions regarding the TMD pain they had experienced in the previous 2 months: number of days of limited activity; number of days of facial pain; and average facial pain intensity. Reasons for refusal were categorized as inconvenience of time, inconvenience of location, or hard refusal (subjects who absolutely refused to give a reason for not participating). These data were used to assess generalizability.

All noncompliers were approached for 1-year follow up with an abbreviated version of the 1-year examination and interview to provide data on the primary variables of interest regarding treatment efficacy. The abbreviated version was used to maximize participation. The CB subjects who did not attend one or both CB sessions were also asked the reason for their nonparticipation. Subjects who completed either the complete or abbreviated 1-year follow up were available for analysis of treatment efficacy at 1 year. This group of subjects is referred to as the *intention-to-treat group*.

Primary variables of interest were demographics, duration (years) of symptoms, self-reported pain intensity and pain interference measures, unassisted and maximum assisted mandibular opening (millimeters), number of extraoral and intraoral muscles painful to clinical palpation, and psychologic variables of depression and somatization assessed by corresponding subscales of the Symptom Checklist-90-Revised (SCL-90-R).[26] Normative age- and sex-adjusted standardized SCL-90-R values, independently established in earlier studies,[27,28] were analyzed. Demographics consisted of age, gender, education, and income. Both pain intensity and interference were measured using 0-to-10 scales anchored at 0 ("no pain" or "no interference") and 10 ("as bad as it could be" or "unable to carry on activities"). Extraoral palpation pain scores were obtained by examination and could range from 0 to 20 painful sites, and intraoral palpation pain scores could range from 0 to 8.

Baseline between-group comparisons were analyzed using t and chi-square tests. One-year between-group comparisons of treatment efficacy were evaluated using analysis of covariance, with the baseline assessment of the outcome variable serving as a covariate. Treatment efficacy was calculated as the difference in adjusted mean treatment responses from the analysis of covariance and its corresponding 95% confidence interval. Tests were performed at an α level of .05.

Results

Less than half of the eligible subjects agreed to participate in the study (Table 1). Those agreeing to participate were randomized to either the CB or UT study groups (n = 95 and n = 90, respectively). Eventually, 46 of those randomized (25%) became noncompliers. At 1 year, more than half (n = 25) of the noncompliers agreed to complete the abbreviated 1-year evaluation, yielding an 89% (n = 164) follow up on the primary measures of interest. Baseline comparisons of the two treatment groups showed baseline between-group differences that depended on whether noncompliers were included, having implications for efficacy analysis.

Refusals

A majority of the refusals (70%) specified inconvenience of time and location of the CB intervention sessions as reasons for not participating. Of the 210 refusals, 56% agreed to answer the three TMD pain questions. Hard refusals were far more likely to not answer the three questions at all.

The participants and refusals differed somewhat on the three TMD questions. Compared to trial participants, refusals reported an average of 2.4 fewer days of pain (41.4 versus 43.8 days; $P = .33$), but both groups equally reported (52%) facial pain every day for the previous 2 months. For average facial pain intensity, the median response was 5 (on a 10-point scale) for both groups, but the mean level of pain in the refusals was half of a point higher ($P = .065$). The refusals more frequently reported no days of limited activity (84% versus 72%; $P = .02$).

Baseline Comparison of Compliers and Noncompliers

The compliers differed from the noncompliers on baseline assessments. Noncompliers had a more recent onset of facial pain (mean 3.2 versus 6.3 years; $P = .01$) and lower family income (69% versus 41% with income < $35,000, $P = .003$). Noncompliers were also more likely to have been assigned to the CB intervention group (31% of the CB group versus 19% of the UT group were noncompliers; $P = .05$). The major reason (76%) for noncompliance in the CB group was nonadherence to the treatment by not attending one or both CB sessions. The main reason given by the CB participants for nonattendance was "time conflict," but four indicated treatment intervention–related dissatisfaction. Following are the reasons, provided at the 1-year follow up, for not attending one or both sessions of CB intervention:

1. Time conflict (seven subjects)
2. Cannot remember (two subjects)
3. Did not wish to travel to that area of city (two subjects)
4. Attended one session—material not new (one subject)
5. Attended one session—disagreed with some material (one subject)
6. Emergency occurred (one subject)
7. Had the flu (one subject)
8. Sought health care elsewhere (one subject)
9. Frustrated with dentist (one subject)

Noncompliance in the UT group could only occur as loss to follow up.

Baseline Comparison of CB and UT Groups

The comparison of all randomized subjects (n = 185) revealed that the CB group had a statistically significantly higher overall level of pain intensity compared to the UT group (mean of 5.18 versus 4.61; $P = .05$). There was also a trend for the CB group to have greater restriction in maximum assisted opening ($P = .10$). By contrast, the CB group had significantly fewer painful extraoral muscles compared to UT (five versus seven, $P = .01$).

Baseline comparisons of the intention-to-treat subjects (n = 164; 139 compliers + 25 noncompliers with 1-year follow up) also demonstrated statistically significant differences between treatment groups, the difference in maximum assisted opening was statistically significant ($P = .03$).

Baseline comparisons of the treatment groups based on compliers only (n = 139) revealed that pain intensity and maximum assisted opening restrictions were higher for the CB group but not statistically significant (for both $P = .06$). However, in addition to fewer painful extraoral palpations in the CB group, they also had fewer painful intraoral palpations (3.0 versus 3.7, $P = .04$).

Treatment Efficacy

Evaluation of treatment efficacy at 1 year, based only on compliers, showed the CB group to have statistically lower mean pain intensity than the UT group (Table 2). The confidence intervals for both

Table 1 Final Study Disposition of Eligible Subjects by Treatment Group

Disposition	n	Treatment group CB	UT
Met eligibility criteria	395		
Refused to participate	210		
Randomized	185	95	90
Compliers	139	66	73
Noncompliers	46	29	17
With 1-year follow-up	25	17	8
Without 1-year follow-up	21	12	9

Table 2 Treatment Differences* and Confidence Intervals (CI) at 1-Year Follow-up

Outcome	Compliers (n = 139)	Intention-to-treat (n = 164)
Pain intensity	$\bar{X}_{UT} - \bar{X}_{CB} = 0.71$ CI = (0.004, 1.42) $P = .05$	$\bar{X}_{UT} - \bar{X}_{CB} = 0.63$ CI = (−0.02, 1.28) $P = .06$
Pain interference	$\bar{X}_{UT} - \bar{X}_{CB} = 0.77$ CI = (0.08, 1.46) $P = .03$	$\bar{X}_{UT} - \bar{X}_{CB} = 0.51$ CI = (−0.14, 1.16) $P = .14$

*Difference in the mean outcome response between UT and CB groups.

pain measures indicated a range of potential differences, from a little more than no difference, up to about 1.4 points, which is about twice the mean difference. When the same analyses were conducted using the intention-to-treat group, the effect sizes were smaller and not statistically different, despite the increased sample size. For pain interference, the impact was greater with a between-group difference of 0.5 points based on the intention-to-treat groups, as compared to 0.77 points based on compliers only. None of the other primary variables of interest showed statistically significant differences between groups using either the compliers or intention-to-treat groups.[22]

Discussion

Although randomized clinical trials are accepted as the gold standard for evaluating biomedical and behavioral interventions and for assessing the size of treatment effects, it is well accepted that RCTs typically involve complex procedures and are susceptible to many potential difficulties. Using our own recently reported RCT of a behavioral intervention as a case in point, the succession of analyses reported in the present study indicates the usefulness of examining in detail the observed departures from planned study protocols and the potential influence on treatment effects when subjects differentially refuse to participate or are noncompliant (ie, drop out of the trial or do not adhere to study protocols). Systematically examining each stage of subject participation in an RCT, from the identification of potential subjects through recruitment, randomization and follow up, provides critical insight into the important kinds of problems particular RCT designs may involve and allows more precise estimates of bias and the generalizability of findings. The systematic implementation of the approaches and methods gleaned from RCT methodology, the most important of which have been demonstrated in the present study, may encourage clinical investigators to present comparable compliance and intention-to-treat analyses, thus allowing a common set of statistical rules for interpreting treatment outcome studies conducted as RCTs.

In this trial evaluating a cognitive-behavioral intervention for TMD, more than half of the eligible subjects refused to participate. They cited inconvenient times or locations as their reasons and, because the intervention was conducted in groups and participants were recruited from two widely separated clinical facilities, the study protocol did impose logistic obstacles to participation. Given that the greater majority of TMD patients were women (88%), other responsibilities may have taken priority over participation. However, if patients actually declined because they did not perceive the intervention as relevant to their own treatment goals and/or they questioned the credibility of the treatments offered, we would be left with a biased sample of participants whose treatment orientation reflected the perceived expectations of the experimenters.

Once a subject joined the study, noncompliance depended on the treatment group to which that subject was randomized, with a greater proportion occurring in the CB group (31% noncompliance versus 19% in the UT group). Noncompliers tended to have lower incomes as well as pain of more recent onset. Perhaps subjects who had pain of a more recent onset were less motivated to go through the effort of attending groups, or a lower income presented them with more obstacles to participation (eg, child care costs).

It may be useful to inquire further concerning the possibility that subjects who have recent onsets of pain still hold the prevalent biomedical model. We have been encouraged by initial attempts undertaken in the present intervention to obtain detailed information regarding biomedical versus behavioral explanatory models that patients hold[29] to analyze whether patients who drop out of behavioral interventions or therapy tend to be those who hold a biomedical model for their chronic pain as opposed to a behavioral model. These issues of compliance may be especially relevant to behavioral RCTs, since it is generally agreed that chronic pain patients are likely to enter such RCTs with a bias favoring a biomedical explanatory model for their pain condition.

Our results showed that between-treatment-group comparisons of baseline measures differed depending on which participants were included in the assessment. Such differences could indicate that the "random assignment" of participants to the different treatment groups has been compromised, potentially undermining the validity and rigor of the study. It is this randomization that guarantees, on average, that groups are balanced on important prognostic factors and that any imbalances that do occur are the result of chance alone, and not selection or self-selection bias. Even if there were no observed differences, such baseline comparisons may not constitute a necessary and sufficient basis for the conclusions concerning the effects of noncompliance.

The most difficult notion to accept may be the contraintuitive requirement associated with intention-to treat analyses—that even those subjects who did not receive the benefit of the full treatment, by virtue of their noncompliance/nonadherence, must nevertheless be included in such analyses. However, intention-to-treat analyses maintain the integrity of the RCT in a readily replicated and standardized manner. Results of related RCTs may be more validly compared using this analysis, since at least one set of outcome analyses uses the same statistical rules. The intention-to-treat results may indeed represent an attenuation of the true treatment efficacy, as was seen here for two of the dependent variables analyzed. For both pain intensity and interference, treatment differences under the intention-to-treat analyses shifted from statistical significance to statistical nonsignificance, in spite of the increased sample size and power.

In theory, noncompliance may actually exaggerate or attenuate the true benefit, but in reality the direction of bias is unknown to the investigator. For example, in an RCT to assess types of occlusal appliances for TMD pain, if patients do not use their occlusal plate for the prescribed amount of time but remain in the trial, the outcome results may underestimate the true efficacy of the appliance. However, if those dissatisfied with their appliance drop out of the trial, leaving behind only subjects who are satisfied with their appliance, results may represent an exaggeration or an overestimate of the actual treatment efficacy for the appliance. Those who are more satisfied may maintain a more positive attitude, which will in turn influence their self-report of pain and interference. Thus, a trial could show a treatment to be efficacious, even if half the participants dropped out; if they dropped out because the treatment was intolerable, that is valuable information to be considered in evaluation of treatment efficacy. In addition, given the cyclical nature of TMD symptoms in the absence of treatment, and depending on the nature and timing of the treatments, one treatment may be less adhered to than another.[30] Thus, without additional data on treatment compliance and follow up of noncompliers, and in the absence of intention-to-treat analyses, the decision of whether the treatment is beneficial relies on only the investigator's bias in selecting which analyses are reported.

The aim of efficacy studies is to maximize the likelihood of discovering the true effect size when a new intervention—for example, a new drug or a new cognitive-behavioral intervention—is evaluated in a clinical trial. Such RCTs are designed to retain as many subjects as possible, and deliberate research strategies are introduced to maximize compliance. This is in contrast to effectiveness studies, which focus more pointedly on the impact of treatment in the general population and seek to recruit noncompliers, describing all consequences of treating a disease in a certain way and determining whether the treatment works under the usual clinical practice. Thus, intention-to-treat analyses may provide an estimate of the efficacy of an intervention different from the one based on compliers—perhaps a more realistic estimate of how effective the intervention would be when introduced into widespread practice, where noncompliance is human nature. If an efficacy trial does not yield results necessary to institute the treatment in practice, an effectiveness trial most likely would not be done.

Consideration of these issues of refusal, noncompliance, and intention-to-treat have very practical implications for the design of more robust RCTs in the area of orofacial pain. First, it is extremely important to maximize compliance among those who agree to participate. Compliance is needed to obtain the best estimate of the benefit or effects of the treatment under consideration. Investigators want and need everyone who participates to actually adhere to and comply with the intended treatment protocol. These are requirements of an efficacy trial. Obviously, if subjects do not comply, it is difficult to determine if the treatment actually works, even in a highly structured research setting. The noncompliance may, in fact, be a strong indicator of the unacceptability of the treatment to patients. To induce greater compliance, these issues can be discussed directly and frankly with potential research subjects; we have observed that participants are more likely to comply if they think they can gain some benefit from their participation. Finally, we observed that less attention to compliance is generally directed to usual treatment control groups. The assumption is that usual treatment conditions are associated with good compliance (taking prescribed medications, keeping appointments, etc). In fact, this assumption is rarely fulfilled, as routinely evidenced by missed appointments and failure to take medication as scheduled.

Summary

It seems imperative that we continue to strive to maximize treatment compliance and to minimize loss to follow up. Descriptive statistics should be

presented to describe the levels of compliance and participation and explore reasons for nonparticipation. Follow up on noncompliers can be accomplished and is important in the final efficacy analyses. As a research strategy, whenever possible, results should be reported according to the two types of efficacy analyses presented here: *(1)* analyses based on adherence to the intention-to-treat principle; and *(2)* analyses of compliers based on those who received the interventions and completed the study. The more data that can be collected and presented on protocol deviations, the better picture we get of the potential sources of bias, allowing readers to evaluate for themselves how these may affect the final conclusions. Such an approach would certainly strengthen, not weaken, the evidence and generalizability of the trial and facilitate the reduction in the number of studies yielding inconsistent results, providing a clearer picture to health care providers and their patients.

Acknowledgments

This study was supported by National Institute of Dental Research Grant No. P01 DE-08773.

Reference

1. Antczak-Bouckoms AA. Epidemiology of research for temporomandibular disorders. J Orofacial Pain 1995;9: 226–234.
2. Antczak AA, Tang J, Chalmers TC. Quality assessment of randomized control trials in dental research. I. Methods. J Periodontal Res 1986;21:305–314.
3. Antczak AA, Tang J, Chalmers TC. Quality assessment of randomized control trials in dental research. II. Results: Periodontal research. J Periodontal Res 1986;21:315–321.
4. Antczak-Bouckoms AA, Chalmers TC. The importance of design and analysis of clinical trials. J Oral Implantol 1988;14:36–42.
5. Turk DC, Rudy TE. Neglected factors in chronic pain treatment outcome studies: Referral patterns, failure to enter treatment, and attrition. Pain 1990;43:7–25.
6. Turk DC, Rudy TE. Neglected topics in the treatment of chronic pain patients: Relapse, noncompliance, and adherence enhancement. Pain 1991;44:5–28.
7. Turk DC, Rudy TE, Sorkin BA. Neglected topics in chronic pain treatment outcome studies: Determination of success. Pain 1993;53:3–16.
8. Jasanoff S. Innovation and integrity in biomedical research. Acad Med 1993;68(suppl 9):S91–S95.
9. De Laat A. Facts and fallacies. . . [editorial]. J Craniomandib Disord Facial Oral Pain 1991;5:5.
10. McNeill C. Flight from science and reason [editorial]. J Orofacial Pain 1995;9:213.
11. Okeson J. Challenging the fringe [editorial]. J Craniomand Pract 1995;13:139–141.
12. Haynes RB, Dantes R. Patient compliance and the conduct and interpretation of therapeutic trials. Control Clin Trials 1987;8:12–19.
13. Vander Stichele R. Measurement of patient compliance and the interpretation of randomized clinical trials. Eur J Clin Pharmacol 1991;41:27–35.
14. Haynes RB. Introduction. In: Haynes RB, Taylor DW, Sackett DL (eds). Compliance in Health Care. Baltimore: Johns Hopkins University Press, 1979:1–7.
15. Pullar T, Kumar S, Feely M. Compliance in clinical trials. Ann Rheum Dis 1989;48:871–875.
16. Armitage P. The analysis of data from clinical trials. Statistician 1979;28:171–183.
17. Detre K, Peduzzi P. The problem of attributing deaths of non-adherers: The VA coronary bypass experience. Control Clin Trials 1982;3:355–364.
18. Tsiatis A. Methodological issues in AIDS clinical trials. Intent-to-treat analysis. J Acquir Immune Defic Syndr 1990;3(suppl 2):S120–S123.
19. Peduzzi P, Detre K, Wittes J, Holford T. Intent-to-treat analysis and the problems of crossovers. An example from the Veterans Administration coronary bypass surgery study. J Thorac Cardiovasc Surg 1991;101:481–487.
20. Fleiss J. General design issues in efficacy, equivalency and superiority trials. J Periodontal Res 1992;27(special issue):306–313.
21. Lee YJ, Ellenberg JH, Hirtz DG, Nelson KB. Analysis of clinical trials by treatment received: Is it really an option? Stat Med 1991;10:1595–1605.
22. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Br J Cancer 1976;34:585–612.
23. Pocock SJ. Clinical Trials: A Practical Approach. New York: Wiley & Sons, 1983:176–186.
24. Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. N Engl J Med 1979; 301:1410–1412.
25. Dworkin SF, Turner JA, Wilson L, Massoth D, Whitney C, Huggins KH, et al. Brief group cognitive-behavioral intervention for TMD. Pain 1994;59:175–187.
26. Derogatis LR. SCL-90-R: Administration, Scoring, and Procedures Manual-II for the Revised Version. Towson, MD: Clinical Psychometrics Research, 1983.
27. Dworkin SF, Von Korff MR, LeResche L. Multiple pains and psychiatric disturbance: An epidemiological investigation. Arch Gen Psychiatry 1990;47:239–244.
28. Von Korff M, Dworkin SF, LeResche L, Kruger A. An epidemiological comparison of pain complaints. Pain 1988;32:173–183.
29. Massoth D, Dworkin SF, Whitney CW, Harrison RG, Wilson L, Turner J. Patient explanatory models for temporomandibular disorders (TMD) [abstract]. In: Gebhart GF, Hammond DL, Jensen TS (eds). Proceedings of the VIIth World Congress on Pain: Progress in Pain Research and Management, Vol 2. Seattle, WA: International Association for the Study of Pain, 1994.
30. Whitney CW, Von Korff M. The magnitude of regression to the mean in before-after treatment comparisons of chronic pain. Pain 1992;50:281–285.

Resumen

Implicaciones prácticas del incumplimiento en los estudios clínicos al azar de los desórdenes temporomandibulares

Los estudios clínicos al azar son reconocidos debido a que proveen la evidencia mas rigurosa en cuanto a la eficacia del tratamiento. En el caso de los desórdenes temporomandibulares, los estudios clínicos al azar han sido utilizados para evaluar la eficacia de aparatos oclusales de bajo costo o el uso adjunto de intervenciones de comportamiento cognoscitivo. Sin embargo, el incumplimiento de los régimenes de tratamiento y la pérdida de pacientes durante el seguimiento, son violaciones de protocolo comunes en los estudios clínicos al azar, que comprometen la exactitud deseada en el estudio. A veces no es claro para el investigador, como lidiar con estos asuntos durante el tratamiento y la fase de análisis de la información. A menudo la eficacia del tratamiento se basa en las personas que cumplieron con éste, lo cual puede implicar que estas personas ya no representen grupos al azar necesariamente, o que produzcan el juicio "recto" deseado en cuanto a la eficacia del tratamiento. Este estudio se concentra en el manejo de los asuntos relacionados al cumplimiento, la descripción y recolección de la información necesaria para obtener una evaluación mas exacta de la eficacia del tratamiento, y de los resultados pertinentes particularmente al ejercicio real de la clínica y decisiones en cuanto al cuidado del paciente. Estos son aplicados a un estudio clínico al azar que evalúa la eficacia de las intervenciones de comportamiento cognoscitivo para los desórdenes temporomandibulares.

Zusammenfassung

Praktische Folgerungen der Noncompliance in zufälligen klinischen Versuchen für temporomandibuläre Erkrankungen

Zufällige klinische Versuche werden für die Lieferung der strengsten Beweise für die Behandlungswirksamkeit anerkannt. Für temporomandibuläre Erkrankungen werden zufällige klinische Versuche verwendet, um die Wirksamkeit von preiswerten okklusalen Schienen oder die zusätzliche Anwendung von kognitiven Verhaltenseinmischungen zu ermitteln. Jedoch stellen Nichtmitarbeit in Bezug auf Behandlungsvorschriften und Weiterfahrensverluste allgemeine Verletzungen des Ablaufs zufälliger klinischer Versuche dar, welche die gewünschte Strenge des Versuches vereiteln. Zur Zeit ist es dem Untersucher nicht klar, wie er diese Ergebnisse während des Versuches und zur Datenanalysenphase behandeln soll. Oft basiert die Wirksamkeit der Behandlung auf kooperativen Leuten, Leute, welche nicht länger zufällige Gruppen darstellen möchten oder die gewünschte „gerechte" Schätzung der Behandlungswirksamkeit liefern. Diese Studie zielt auf die Behandlung von Compliance-Ergebnissen, die Beschreibung und Sammlung der Daten, welche nötig sind, um eine genauere Beurteilung der Behandlungswirksamkeit zu erhalten, und sie ergibt besonders relevante Ergebnisse in Bezug zur aktuellen klinischen Gewohnheit und Entscheidungen zur Patientenbehandlung. Diese werden verwendet für einen zufälligen klinischen Versuch, um die Wirksamkeit eines Eingriffes in den Bewusstseinshintergrund bei temporomandibulären Erkrankungen herauszufinden.