Commentary

Incorporating Patient-Reported Outcomes in Periodontal Clinical Trials

Michael K. McGuire,* E. Todd Scheyer,* and Chad Gwaltney†

The authors review patient-reported outcome (PRO) metrics for dentistry, and in particular, periodontics. The PRO commentary for periodontics includes a review of split-mouth, randomized, controlled clinical trial results that specifically tracked pain at different sites over time after intervention and provided guidelines for peak pain time points and evidence for referred pain assessment when studying soft tissue augmentation procedures. Both the questions that are asked of patients and the timing of those questions are important study design considerations. The authors suggest PRO methodology for periodontal clinical trials that can be used to identify information important to patients and clinicians. J Periodontol 2014;85:1313-1319.

KEY WORDS
Dentist-patient relations; doctor-patient relations; outcome and process assessment (health care); patient-centered outcome research; patient cooperation.

As a measure of dental practices, the ultimate patient outcome is whether a patient preferentially returns to the dental offices, or better yet, refers a friend. In this regard, the patient’s opinions about treatments are fundamental to what dental professionals do. But in clinical research, patient-reported outcomes (PROs) — what the patient reports about his or her treatment experience — are fast becoming the metrics that will also help determine the services provided.

Intuitively, one knows that PROs such as quality of life should help find the best therapies, but the practice of effectively incorporating PROs in periodontal clinical studies is only now being refined. By definition, a PRO is “any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else.”1 This can include reports about specific symptoms, such as pain at a particular site in the mouth, or more general concepts, such as difficulty eating or, more generally, treatment satisfaction — always reported directly by the patient. However, creating valid PRO instruments and metrics requires more than simply removing clinicians and staff from the interpretation of patient opinions. The US Food and Drug Administration and the European Medicines Agency have provided position papers and guidelines for the development, implementation, and interpretation of PROs in clinical trials.2 Indeed, PRO instruments have served as the basis for regulatory approval and appeared in labeling for products across several therapeutic areas.3 In essence, PRO instruments should: 1) be free from error (be reliable); 2) measure what they are intended to measure (be valid); 3) be sensitive to changes in the patient’s condition (be able to detect treatment differences); and 4) be interpretable (be clinically meaningful).4 Although these goals might also sound easy to reach, effective PRO methodology is surprisingly difficult to develop.

PROs, by definition, are subjective, i.e., they involve the measurement of a wide spectrum of patient opinions and experiential responses to difficult-to-quantify endpoints, such as anxiety, pain, and satisfaction. Because many important patient experiences occur

* Private practice, Houston, TX.
† Department of Behavioral and Social Sciences, Brown University, Providence, RI.

doi: 10.1902/jop.2014.130693
soporadically (for example, sudden pain when chewing), they need to be recorded in near real time, as they are occurring in the patient’s natural environment. Also, patients may be reluctant to provide honest, unguarded answers, particularly when speaking with, and trying to please, their caregivers, so PROs must be administered and recorded in a way that minimizes these “demand characteristics.” For example, computerized assessments or questions administered by staff members who are not involved in the patient’s treatment may yield more accurate responses. Exacting item development procedures, usually including an iterative process of interviews with patients, and innovative trial designs must be implemented to best develop and administer clear and quantifiable PROs. This is critical in periodontal clinical trials, where patient experiences, such as pain, are central to evaluating new treatments.

**PROs in Dentistry and Periodontics**

There has been considerable PRO work in dentistry and the psychometric characteristics of various dental PRO instruments have been assessed. For example, Borges et al. examined the relationship among loss of periodontal support structures, masticatory function, and quality of life. Fardal and McCulloch and Kloostra et al. investigated the impact of pre-surgical anxiety, stress, and depression on periodontal and implant surgery and pain. McGrath et al., in reviewing the dental PRO literature, suggested more pretreatment assessments, but particularly more standardized PRO metrics to facilitate interstudy comparisons. In this regard, the Oral Health Impact Profile has been recommended as generally applicable to dentistry. However, standardized PRO instruments may not always generalize to the specifics of a given therapy. For example, O’Dowd et al. found patient satisfaction with periodontal therapy to be influenced by stigma and retrospective regret, a particular relationship that might be unique to some PRO measures within periodontics. From a regulatory perspective, it is clear that the items used in a PRO instrument should match the context of use in which the instrument will be used (e.g., patient population, type of trial, type of treatment). However, there are pluses and minuses associated with the use of generic and specific PRO instruments, and investigators should evaluate their options before launching a trial.

PRO instruments are essential for differentiating therapies that could be ranked not only differently but also divergently by patient and caregiver. Kokich et al. compared dentist and lay perceptions of dental esthetics and found that professionally evaluated outcomes tend to be more critical than lay-evaluated outcomes. In other fields of medicine, clinical gold standard therapies such as saphenous vein or autogenous bone transplant procedures might prove more effective when compared with less morbid and alternative synthetic, allogenic, or xenogeneic graft therapies, but the alternatives are often effective enough, and they may be preferred by patients. Likewise, in the authors’ experience, when considering only traditional clinical metrics like root coverage for alternatives to autogenous soft tissue harvest, the alternatives have not always measured up clinically to gold standard autogenous grafts, but the clinical results for the alternatives have been effective, and patients have tended to prefer them.

Certainly pain assessments ought to be a common PRO instrument in periodontics (how much discomfort patients experience as a result of a given therapy). In this sense, a study could focus on product or intervention safety, in addition to product efficacy, and reduction of pain might be an endpoint that differentiates interventions.

**PRO Examination of Soft Tissue Augmentation**

Although intuitively one would predict a marked difference between procedures that do or do not require palatal harvest, in the authors’ previous work with donor graft alternatives, pain differences have been difficult to detect owing to the split-mouth design, among other factors. Likewise, and also counterintuitively, a minority (25% to 30%) of the authors’ study patients have continued to prefer traditional autogenous graft therapy over alternative non-harvest therapies.

Before initiating a recent soft tissue augmentation study, the authors decided to look for more reliable and sensitive PRO metrics. After teaming up with a psychologist (CG) specializing in PROs, the authors conducted a pilot study examining free gingival graft (FGG) versus a collagen matrix (CM) graft substitute. The study had 13 participants: seven patients receiving FGG only and six within-patient, contralateral, split-mouth, comparison patients. The psychologist interviewed the patients pre- and post-surgery to better understand their pain and discomfort, attitudes, and preferences and to develop sensitive PRO items for use in a clinical trial that could be easily and universally understood.

The authors then designed and implemented a single-masked (examiner), randomized, controlled, split-mouth study in 30 participants with insufficient zones of keratinized tissue (<2 mm). The protocol and patient informed consent process were approved by an institutional review board (Western IRB) and complied with federal (Code of Federal Regulations Title 21, Part 56) and Health Insurance Portability and Accountability Act (HIPAA) requirements. The study used a within-patient treatment comparison designed to establish non-inferiority for the CM therapy against...
traditional FGG therapy. The study sample was derived from the population of patients (three males, 10 females; aged 18 to 71 years; mean: 47 years) who presented at the authors’ practice from November 2010 to March 2011. Patients who had previously undergone grafting procedures at or adjacent to the study sites were excluded, as a history with soft tissue augmentation procedures could bias patients’ reports about their experience with the treatments. Oral consent was obtained from the patient’s for this WIRB exempt study.

The CM and FGG test and control materials were placed in direct contact with the appropriate, randomly assigned wound bed and sutured in place into the papillary region on the mesial and distal aspects of the tooth. The FGG and CM sites were left uncovered, i.e., no wound dressing, as the pilot study indicated that dressings could mask pain, whereas the graft harvest site was covered with surgical dressing. The time at which wound dressings were lost was recorded. To promote valid pain/discomfort measures, pain medications were not prescribed or encouraged, although patients desiring pain medications were prescribed hydrocodone/acetaminophen, and medications taken were recorded.

Based on the results of the pilot study, before and immediately after surgery, scripted PRO questionnaires were administered verbatim by independent recorders, i.e., not the clinical investigators/surgeons or the clinical examiner. Anxiety, pain/discomfort, and treatment preference PRO questionnaire scripts were designed with the following endpoints in mind: 1) anxiety: prior to surgery, patients were asked about anxiety regarding the two procedures; 2) pain/discomfort: patients were asked about pain and discomfort at the two treatment sites and palatal harvest site, as well as the side of the palate that was not a surgical site, immediately after the procedure and daily from day 1 through day 10 (maximum and overall discomfort, discomfort when eating and avoidance of surgical sites when eating, medication episodes due to pain, and whether wound dressings at palatal harvest sites were in place); 3) preference: patients were asked their preferences for the FGG or CM procedures immediately after surgery, at week 1, and at 3- and 6-month visits; and 4) esthetics: patients were asked about the esthetics of the treatments at 3- and 6-month visits.

The excerpts from scripted PRO questionnaires illustrate the design and tone of the evaluations conducted by telephone or at recall visits by third-party recorders (see Figs. 1 through 3).

**PRO FINDINGS**

Baseline clinical measurements for contralateral test sites were comparable, with no significant differences, except that baseline wound beds and corresponding graft measures were larger for the CM therapy, as specified by study protocol. Surgery and postoperative sequelae were uneventful, with normal healing observed at both CM and FGG sites. No unanticipated adverse events were recorded.

The overall pain burden for the two therapies, combined and averaged with their corresponding harvest and non-harvest sites, is shown in Figure 4. The average maximum pain reported was significantly different over postoperative days 1 to 3 (CM = 1.48 ± 0.87, FGG = 2.24 ± 1.59, P = 0.003) and postoperative days 1 to 10 (CM = 0.97 ± 0.63, FGG = 1.53 ± 1.23, P = 0.004).

Peak pain measures diminished (and were not significantly different between the CM and FGG procedures) after ≈1 week. When evaluating peak pain for patients (while eating, with and without wound dressing at the palatal harvest site, and with and without out pain medications), comparisons between the two therapies were similar to that depicted in Figure 5, and pain measures were not significantly amplified by any of the conditions considered. Understanding pain in the context of this type of split-mouth design is complex and requires an assessment of different sites in the mouth, to clearly understand the impact of treatments. An overall pain score for
DAILY INTERVIEW SCRIPT (first 10 days following surgery):

Hello, __________ As you know, we are talking to patients each day after their surgeries to discuss how they are doing. I am going to ask you a few questions now about any pain and discomfort that you may have experienced as a result of the surgeries on both sides of your mouth. As a reminder, on the RIGHT side of your mouth – the same side as your RIGHT hand - you had (tissue taken from the roof of your mouth and placed on your gum/CM placed on your gum). On the LEFT side of your mouth – the same side as your LEFT hand – you had (tissue taken from the roof of your mouth and placed on your gum/CM placed on your gum).

A. To start with, please think about the RIGHT [LEFT] side of your mouth along the gingival where you had the Gingival graft [CM] procedure:
   1. On a scale of 0 to 10, where 0 = No discomfort at all and 10 = As much discomfort as you could imagine, what is the maximum amount of discomfort that you felt along the right [left] gumline since you woke up this morning? ______________
   2. Repeat for opposite side/procedure.

B. Now, please think about the roof of your mouth on your RIGHT [LEFT] side where you had (no procedure at all/ tissue removed for the graft):
   1. On a scale of 0 to 10, where 0 = No discomfort at all and 10 = As much discomfort as you could imagine, what is the maximum amount of discomfort that you felt on the right [left] side of the roof of your mouth since you woke up this morning? ______________
   2. Repeat for opposite side/procedure.

C. Now I want to ask you about the dressing that was placed on each of the surgical sites in your mouth.
   1. Is the dressing still attached at the gumline on your RIGHT [LEFT] side? Yes/No
   2. Repeat for opposite side.

D. Now, I want to ask you about any medications that you took today.
   1. Since you woke up this morning, how many times did you take the (name of anti-inflammatory medication)? ______________
   2. Did you take any pain medications? Yes/ No (If Yes, ask #3. If No, continue to section G.)
   3. For what reason did you take the pain medication?
      Pain at gumline on right side? Yes/No
      Pain at gumline on left side? Yes/No
      Pain at roof of mouth on right side? Yes/No
      Pain at roof of mouth on left side? Yes/No

G. Finally, I want to understand if you have avoided any part of your mouth when you were eating today.
   1. Did you chew food on the LEFT [RIGHT] side of your mouth to avoid pain or discomfort on your RIGHT [LEFT] side? Yes/No
   2. Repeat for opposite side.

3-A Nd 6-MONTH ESTHETICS & PREFERENCE

Think about the appearance, or look of each treated site [allow patient to utilize a mirror and remind him/her that we are interested in the left and right sides].

Considering only the appearance of the two study sites, which one do you prefer?

0 Patient's Right
0 Patient's Left

Figure 2.
Daily interview script.

Figure 3.
Script for 3- and 6-month esthetics and preference questionnaire.

different sides of the mouth will mask true differences.

Peak pain at treatment sites, when harvest/palatal sites were excluded, was not significantly different when averaged over days 1 to 3 (CM 2.87 ± 1.71, FGG 2.70 ± 1.83, P = 0.54) or days 1 to 10 (CM 1.92 ± 1.27, FGG 1.61 ± 1.28, P = 0.15). However, daily pain was slightly greater for CM from days 8 to 10, perhaps because of the larger recipient bed and the suturing technique used (Fig. 5).

Peak pain at palatal harvest sites was significantly different (P < 0.001) from non-harvest sites when evaluated over days 1 to 3 (CM 0.09 ± 0.28, FGG 1.79 ± 1.66) and days 1 to 10 (CM 0.03 ± 0.10, FGG 1.44 ± 1.32) (see Fig. 6). Pain measures were highest over the first week, then diminished thereafter (see Fig. 6 for daily evaluations). Peak palatal pain experiences were higher for patients whose wound dressings were not in place during postoperative days 1 to 3 (CM 0.40 ± 0.97, FGG 2.97 ± 2.03, P = 0.005).

Patients recorded pain at non-harvest sites until day 6, but no later, and so received recorded values of 0. The FGG treatment site was reported as significantly more painful than the palatal harvest site for postoperative days 1 and 2, but no statistical difference could be discerned for the following postoperative days.

Evaluation of pretreatment anxiety and maximum reported pain scores for FGG patients produced a positive relationship of 0.25, i.e., higher anxiety scores correlated with higher pain scores; however, this relationship was fairly weak and not statistically significant. Patient anxiety regarding the two procedures did not differ significantly.

Three of the thirty patients requested and used pain
medication the first day after surgery, with only one patient continuing to use the prescription on subsequent days. The reason given for using pain medication was equally attributed to CM and FGG therapies. Likewise, there was no statistically significant difference between therapies for patients who reported avoiding chewing on one or the other sides of their mouths ($P = 0.82$ for days 1 to 3 and $P > 0.99$ for days 1 to 10).

Immediately after the surgery, patients overwhelmingly preferred the CM procedure (90%), but by 3 and 6 months post-surgery, preference for the CM therapy had dropped to 60% of patients. Patients who preferred the CM therapy reported their reason as avoiding a harvest site, but patients who preferred the FGG therapy reported their reasons as bleeding or discomfort at the CM site or nervousness about an “unknown” material. Timing matters: as the memory/experience of the harvest site faded over time, other factors become more salient, and preferences changed.

More than two-thirds of patients (70%) preferred the appearance of their CM sites.

**DISCUSSION**

The importance of and the difficulties in identifying PROs in periodontics have been recognized for more than a decade. In 2003, a systematic review of surgical therapies for the treatment of gingival recession commissioned by the American Academy of Periodontology concluded that most studies at that time lacked standardized measures for PROs and suggested that future studies should be designed primarily to investigate patient-oriented outcomes such as esthetics, hypersensitivity, morbidities, and overall satisfaction. Unfortunately, little has been accomplished over the last 10 years to standardize methods of evaluating PROs in periodontal trials, and few, if any, studies have been dedicated to PROs.

This study indicates that discomfort was reduced with the CM therapy. A significant pain difference was detectable, but only when the total pain burden for both harvest and non-harvest sites was considered. Measuring and separating pain, not only at treatment and harvest sites but also at non-harvest sites, may be key to detecting differences in future periodontal studies, especially those involving autogenous tissue therapies.

The authors discovered that the side of the palate from which no harvest graft is taken is also experienced as painful by patients, i.e., pain appears to be referred. They also discovered that peak pain occurs over the first 7 days, and particularly the first 3 days, post-surgery. This is an important finding, since previous studies have tended to record pain at 7 days, when the opportunity to differentiate therapies or tease out particulars about pain findings may have passed. Finding this threshold time point required
Patients overwhelmingly preferred the esthetics of CM therapy, favoring its natural color and texture match with native tissue. Patients also tended to prefer the CM procedure over the FGG procedure, primarily because donor-site surgery was avoided, though the larger wound bed (and number of sutures) required for the CM therapy may have diminished patient preference over time. The authors discovered that it is important to record PROs early and as closely to real time as possible, because patients tend to forget the intensity of pain over time, and this recall error may serve to diminish treatment differences. Also, although the companion study to this commentary reports that keratinized tissue gains were superior for the FGG control therapy, the CM test therapy was sufficient, i.e., 29 of 30 CM cases achieved ≥2.0 mm keratinized tissue width. It is possible that we may never find alternative therapies that will have clinical outcomes superior to traditional gold standard therapies involving autogenous tissues. Reliance on PROs will therefore become more important for the future selection of alternative therapeutics.

Although this article focuses on the benefits and strengths of PROs, it is also important to consider their limitations when designing a trial. Patient responses may be influenced by knowing the nature of their treatments and by subtle cues from investigators. Patient views on broad outcomes such as treatment satisfaction can be affected by irrelevant factors, such as how much they liked the clinical staff. Although many of these factors are minimized in randomized controlled trials, it is important to carefully consider the design of the PRO items and the administration procedure, to avoid sources of bias and error. It is often best to include both biologic/anatomic and PRO endpoints in trials to provide a comprehensive picture of a treatment’s effects.

**CONCLUSIONS**

One of the key take-away messages from this project is that both the questions that are asked of the patient and the timing of those questions are important to consider during the design of a new study. In this project, the authors were able to understand the nature of the patient’s experience of test and control therapies only by asking multiple, specific questions about the severity of pain experienced at different sites in the mouth. This revealed that any pain advantage offered by a non-harvest therapy is due to the absence of tissue harvest and not reduced pain at the treatment site. Additionally, pain and satisfaction changed dramatically over time, suggesting that single measures of either construct may miss important treatment differences at other time points. PROs can have immense value, particularly as it becomes more difficult to demonstrate incremental benefits of new treatments on objective measures, but researchers must think carefully about how to develop and implement them in trials.

Researchers considering PRO measures should refer to the guidelines provided by recent literature and established instruments such as the Oral Health Impact Profile before designing and incorporating PRO methodologies in their studies. The authors conducted pilot study cognitive interviews that helped develop additional, meaningful PRO items, those that patients understand and are adequately sensitive for measuring meaningful outcomes. Pre-assessment of parameters such as anxiety and esthetics may be informative (certainly patients already satisfied with their appearance could be a special subgroup in esthetics studies), although in the case of this study, presurgical anxiety did not relate strongly with discomfort following surgery. The efforts to divide and individually analyze conditions such as “when eating” or “avoidance when eating,” “influence of medications,” “overall versus maximum pain,” and “wound dressing in place or not in place” did not change the general outcome regarding the timing and duration of pain or the overall differences between therapies, but these and other similar considerations may be useful in other studies.

The authors intend to follow the results of this study long term, as participants of any PRO study should be followed, to make sure that the alternative therapy, although generally less painful and
preferred by patients, continues to provide clinically effective outcomes. The authors hope their work has provided methodology and results that can be used in and compared with future PRO soft tissue augmentation evaluations.

ACKNOWLEDGMENTS

The authors thank Rebecca Garcia, RDH, Director of Clinical Research, Perio Health Clinical Research Center, Houston, Texas (PHCRC), for recording outcome measures and coordinating data, and Cindy Wainscott, CDA, PHCRC, for study administration. This study was supported by a grant from Geistlich Pharma AG, Wolhusen, Switzerland. Drs. McGuire and Scheyer have received lecture fees from the company. Chad Gwaltney is employed by ERT, Inc., Philadelphia, Pennsylvania, which develops PRO instruments for use in clinical trials and provides electronic solutions for capture of PRO data in the clinic and in the field.

REFERENCES


Correspondence: Dr. Michael K. McGuire, 3400 S. Gessner, #102, Houston, TX 77063. E-mail: mkimerio@swbell.net.

Submitted November 21, 2013; accepted for publication February 20, 2014.