Chapter 5. Recommendations for Future Research

Clinical Issues Concerning Antimicrobials

Our recommendations for further research on the role of antimicrobials as adjuncts in the treatment of chronic periodontitis reflect several concerns. First, some issues remain for the specific key question of this evidence report, which dealt with the *added* effectiveness of particular antimicrobials when they are used as an adjunct to scaling and root planing (SRP), including whether any antimicrobials warrant further investigation in this regard. Second is the design and analysis of any future studies of this question, as the limitations of the existing literature are not trivial. Some of those limitations may relate to the actual reporting of the trials or other studies reviewed here, not the underlying design and conduct of the investigations per se. Moreover, the research reviewed in this evidence report provides a fairly broad range of expected effects of adjunctive antimicrobial use, but those effects remain far smaller in magnitude than the benefits achieved by SRP alone. A third topic, therefore, centers on the issue of what size difference between SRP alone and SRP with an adjunctive antimicrobial has clinical significance or relates to outcomes of particular meaning to patients and their dentists or periodontists.

We are limiting our consideration of future research directions or priorities to which antimicrobials, if any, warrant further examination in the context of use as adjuncts to SRP. We did not review literature relating to, for instance, antimicrobial use as an alternative to SRP, and we did not include every possible antimicrobial, in every possible modality, in this evidence report. Therefore, we do not comment further on any potential for new research in those areas. We also do not comment on research focused solely on the effectiveness of SRP per se, which appears to be well grounded in robust evidence accumulated over the years, as that was not a key question for this systematic review.

Types of Antimicrobials

Of all the medications we did review in this evidence report, three would seem to have had sufficient promise as SRP adjuncts to justify continued investigation: tetracycline, minocycline and perhaps chlorhexidine and metronidazole. We base this conclusion on those results that seem to show that these pharmaceuticals, in either local or systemic form, conferred at least some extra benefit that was statistically significant when used in conjunction with SRP. The main outcomes in which this benefit occurred tended to be reductions in probing depth (PD) or gains in clinical attachment level (CAL), not in reductions in the presence of bacterial agents (specifically spirochetes).

With respect to tetracycline, the evidence for effectiveness of the drug applied locally, measured as reductions in PD, appeared to be fairly consistently statistically significant in the literature we reviewed. Further investigations of locally administered minocycline and doxycycline, both tetracycline-like antibiotics, may also provide better insights into their utility as SRP adjuncts. Remaining issues include the magnitude of the PD reductions and how long those reductions persist. As for adjunctive chlorhexidine, the evidence appears to substantiate statistically significant improvements in terms of reduction in PD and gain in CAL; however, the improvements over SRP alone are very modest, and how long they persist without continued treatment also remains to be established. The third antimicrobial with sufficient evidence

appears to show fairly consistent, statistically significant differences between treatment and control groups in reducing PDs is metronidazole; by extension, metronidazole in combination with amoxicillin seems to produce similarly encouraging results.

Thus, we would be comfortable in encouraging additional research to document more clearly whether these positive directions are real for these particular antimicrobials, the size of the improvements, and the time periods over which such improvements last. Narrower questions involve whether such results are similar across patients with different initial PDs or other clinical characteristics or whether positive results tend to be observed more in patients with more severe chronic periodontitis (e.g., initial PDs of 6 mm or more).

By and large, too few studies provided enough information to permit a thorough review of the possible impacts of antimicrobials, *in the specific role of adjunctive therapy*, on the presence (or absence or elimination) of pathogenic bacterial species. On balance, we judge the important goals of SRP with or without adjunctive therapy to be improvements in clinical measures related to possible bone or tooth loss and in patient-oriented outcomes. Nonetheless, chronic periodontitis is an infectious disease or inflammatory process for which the putative causative organisms most important for initiating or sustaining the disease have not been definitively identified. Therefore, continuing to investigate what organisms are most important in chronic periodontitis and the effects of adjunctive antimicrobial use on them may still be an important step in well-designed future studies.

The remaining antimicrobials reviewed in this report might warrant additional research, but it would have to be designed, in the first instance, to establish whether they can be expected to deliver consistently statistically significant added benefits over SRP alone. All in all, we would recommend that the dental research community and funding agencies put higher priority on clarifying the impacts of the three main antimicrobials noted earlier, rather than continuing to mount research on agents that have not, to date, shown as much promise.

Clinical Significance of Potential Benefits

A critical gap in the evidence base assembled so far concerns what clinical meaning to attach to differences in PDs, CALs, or other measures between what is achieved with SRP alone and what is achieved with SRP and adjunctive antimicrobials. Much of this literature commented on "before and after" measures of PD, CAL, and the like *within* treatment and control groups; the studies often did not give their own results about the net differences *between* treatment and control groups at the close of the follow-up period (that is, the "differences between the differences"). Where those data were available, or where we could calculate them, we determined that these net differences were often relatively small, at least on average across patients with different baseline levels of PD.

Thus, even in the face of statistical significance, the dental field is left without a good sense for the clinical significance of these comparatively small net improvements. One problem is that large samples can produce statistically significant results that have little, if any, clinical significance or relevance for the typical practice of periodontology. For that reason, statistical significance should never be the sole criterion by which to interpret these research results. By extension, the dental community must consider clinical factors as well as have an appreciation of the value of these net changes in terms that relate to outcomes valued by patients (e.g., appearance, functioning, or pain).

Therefore, we recommend that, in future studies of these medications as adjuncts to SRP, more attention be given to what levels of improvement should be considered clinically significant. Such information is needed to help guide changes in actual dental practice. One useful step for researchers is to attempt to reach some consensus on what extent or range of expected improvement in PD or CAL should be the goal of the adjunctive treatment. In so doing, in conjunction with newer studies as suggested above, dental researchers might then be able to narrow the field of eligible antimicrobial agents even further, providing a better knowledge base for options in dental practice.

Other Research Questions

Our evidence report did not deal with issues of costs or cost-effectiveness of antimicrobials as adjuncts to SRP, partly for reasons of time and resource availability and partly for lack of solid evidence on effectiveness in the first place. The first priority, as suggested above, is to understand the marginal benefits of adjunctival medications over SRP alone and which ones provide clinically meaningful marginal benefits. At that point, however, questions of the marginal costs of those medications comes into play and, from that, questions of the relative cost-effectiveness of different medications become important. We would recommend, as future research begins to answer the first-order clinical questions, that data be collected to address the economic ramifications of the use of antimicrobials as adjuncts to SRP.

Some experts in the field noted the paucity of information on so-called patient-oriented outcomes in this research base. We would agree that more work needs to be done, once the clinical significance of the current measures of periodontal health is clarified, on correlating these with health status or quality-of-life measures that matter to patients. These might include domains involving pain, eating and nutrition, concerns about appearance, impacts on social interaction, as well as effects of the disease and treatment options on usual daily activities (e.g., days lost from work) and on their out-of-pocket costs of care.

At the outset, we aimed to use reduction or elimination of bacterial causative agents as an important outcome variable, but this proved problematic because of the variety of species that appeared in this literature, the variety of ways changes in the presence or absence of these species were reported, and the fact that commonly reported species are not considered by some experts as comprehensive enough. Thus, as a sidelight to research on the specific issues of adjunctive antimicrobial therapy might be further studies that focus on clarifying the broad range of bacterial agents culpable in chronic periodontitis and their relative significance in this disease process, the effectiveness of therapy in eliminating or at least suppressing these pathogens, and correlating results about specific bacterial species with results relating to changes in clinical measures such as PD or CAL.

Improving Study Design and Conduct

Chapter 3 noted many of the difficulties we encountered in identifying appropriate research articles that would meet our *a priori* inclusion criteria and then in reviewing the included material in any coherent and systematic way. We may thus have omitted some relevant literature from this report, but even more important is the likelihood that some of the research that we did include could not be fully used or was open to incorrect interpretations because of poor reporting practices, confusing study designs, underpowered studies, and poorly conducted investigations.

The reporting practices may be the easiest to correct in the future (even though analysis and reporting of the data from these types of studies seem to have improved in recent years). Several authoritative statements from international groups provide clear instructions on appropriate ways to report on systematic reviews (QUORUM[105]), randomized controlled trials (CONSORT[106]), and observational studies (MOOSE[107]). Authors and journal editors alike should take heed of these guidelines as a critical step in improving this literature overall.

Other guidance can come from the growing movement to grade the quality of individual articles that are included in reviews such as this one to begin with. Among the critical work now available is a lengthy report on systems to grade the quality of studies (i.e., articles) and rate the strength of evidence from the RTI-University of North Carolina Evidence-based Practice Center[39] and related methods of the US Preventive Services Task Force.[108]

Study design issues, such as randomization, allocation concealment, blinding, and similar elements, must be given more careful consideration. Randomization should be a standard for all trials in this area. Use of placebo controls, and not simply variation in treatment arms, will be another useful step for trials attempting to establish the efficacy of a given medication. Every effort should be made to blind (mask) all parties (subjects, treatment providers, and outcome examiners) to the group (treatment arm[s], control) to which the patients belong. This is necessary to reduce the possibility of bias, which past research suggests typically exaggerates the effect of the treatment over what is experienced in the control group. In keeping with the reporting standards noted above, investigators should report clearly on randomization, control groups, and level of blinding achieved in their studies.

In addition, the study sample should be large enough to have adequate power to detect a statistically significant and clinically desired difference. Many of the studies we reviewed had apparently reasonable effect sizes, but they were based on small samples with large variances and could not have reached statistical significance. The problem here may be two-fold: the size of the original samples and the possibility of attrition (especially for studies with very long follow-up periods) such that samples at completion of the study became too small to provide adequate power for the analyses. This latter issue may pose particular challenges for investigators who propose to carry out intention-to-treat analyses but have instead to rely on final data only on completers.

Study reports often were unclear as to the underlying denominators for results, sometimes reporting on persons enrolled but then presenting data on some other unit. Thus, researchers should make it clear what the unit of analysis is – persons, teeth, sites, or pockets – and on what basis their means and measures of variance have been calculated. Specifically, investigators should ensure that their reports specify the number of units on which the mean for each group has been calculated and the variance (either standard deviation or standard error). Without this or comparable information, they or others cannot easily include the results in meta-analyses.

In addition, when using split-mouth designs, analytic techniques that take into account the nesting of observations within subjects need to be used, and when tests of statistical significance between groups are performed on multiple groups, techniques that adjust for the true significance level need to be used and reported. It remains to be demonstrated whether split-mouth designs of local therapies can adequately control the contamination or spill-over effect to be able to measure the true difference between the test and control groups.

Researchers in this area need to establish what measures are most meaningful for reporting treatment effects. Reaching some consensus on core outcomes for studies would help immensely for future systematic reviews on these topics, because the sheer number of possible outcomes complicated our work. With the inputs from our technical expert panel and representatives of the sponsors of this review, we selected PD reduction and CAL gain as the targeted clinical outcomes for several reasons. They appeared fairly frequently and consistently in the literature over the period covered by our review. They are also meaningful measures for clinicians, who can take such measures themselves to monitor the effects of treatment on their own patients. In the more recent literature, however, we saw a move away from reporting outcomes in terms of these metrics to outcomes that are somewhat less easy to understand or to measure objectively and reliably. Among them are variables such as percentage changes in prognosis, shifts from one category of treatment to another (extraction or surgery to maintenance), and other measures involving time (e.g., period of noninfection, time to recurrence). Moreover, investigators would find that easily used and understood statistical techniques are more readily available for analysis of metric data than for the analysis of percentages. A consistent, agreed-upon set of "critical" outcome measures would foster better comparisons across research projects and with past research. If the field moves to some of these newer outcome variables, attention will need to be given to standardizing how they are defined and reported and developing ways to convey absolute results and variances.

Some observers have noted that this literature contains little about measurement error and how it might affect reported results. Among the concerns are ambiguities about the level of training of those doing the SRP, the extensiveness and thoroughness of the SRP, the level of training and standardization of persons collecting the clinical measures, inaccuracies in measuring PD or CAL (or level of pre-existing inflammation), reliability of measurements across multiple examiners, and similar factors subject to variability in assessment and reporting. Moreover, time devoted to SRP, which is now the best proxy for the thoroughness of SRP, is a relatively imprecise measure and does not, in any case, ensure that SRP treatment was comparable across studies, patients, teeth, surfaces, or sites. Among the suggestions for overcoming some of these problems, at least in research venues, is the use of fiber optic devices that permit visual inspection of root surfaces and determination of the thoroughness of subgingival calculus removal and the level of cleanliness and smoothness of the root surface. The idea is that teeth (or surfaces, etc.) would be considered eligible for entry into a trial only after they had met some basic standard of SRP success. Whether moving to such a direct measure of SRP performance in place of time spent on SRP would yield more reliable and valid results, given the presumed additional costs to the research project, is itself an empirical question.

Finally, investigators need to be clearer as to the underlying diagnoses for their subjects. This point concerns two sources of ambiguity for those involved with developing the evidence based on these questions. The first problem is the mix of terms different research teams used for what was apparently the same disease: sometimes periodontal disease, sometimes periodontitis, with several different adjectives (adult, chronic, severe, moderate, mild) used, sometimes alone and sometimes in combination. We made every effort to focus this review on chronic periodontitis in adults (and in particular to eliminate studies in which patients could have had refractory or aggressive periodontitis), but on occasion we needed to draw an inference as to whether chronic periodontitis was indeed the disorder in question. Greater standardization of disease descriptors and their definitions, at least for use by the research community, would be helpful.

The second issue concerned whether subjects were being treated for periodontitis for the first time or were being retreated for chronic periodontitis. This confusion reflected in part the unpredictable use of descriptors such as recurrent, persistent, or refractory; although refractory may have a generally well-understood meaning within the periodontal and dental research community, recurrent and persistent have less agreed-upon definitions or connotations. The current literature generally did not make clear whether persons receiving retreatment had unsuccessful earlier treatment or were simply being retreated after successful treatment at some time in the past (i.e., were on some form of maintenance schedule). If researchers are including "maintenance" patients in their trials, they should explain this decision. More generally, investigators need to be certain that they are including only the types of patients for whom positive results from the particular study would be applicable in everyday practice.