Consensus Training: An Effective Tool to Minimize Variations in Periodontal Diagnosis and Treatment Planning Among Dental Faculty and Students


Abstract: Considerable disagreements and variations exist in diagnosis and treatment planning of periodontal disease. Achieving high interrater periodontal diagnosis can prove challenging. The objectives of this study were to measure variations in periodontal diagnosis and treatment planning among predoctoral periodontics faculty members after consensus training and to compare such variation with those identified in third- and fourth-year dental students. Nine electronically stored case vignettes and survey instruments were made available to eighteen faculty members and twenty dental students under standardized conditions. A chi-square test was used to compare responses between groups, and multirater kappa tests were used to evaluate interrater agreement/reliability. Of the nine cases, only one differed between groups significantly in terms of treatment. Also, third-year students differed from fourth-year students on the diagnosis of aggressive periodontitis versus chronic periodontitis. Most respondents were able to distinguish clearly among diagnoses of chronic periodontitis, aggressive periodontitis, and gingivitis. This study established a baseline assessment of the current status of consensus after training. We will reassess variations after addressing the specific challenges identified. Programs designed and implemented to help decrease the variation in periodontal diagnosis and treatment planning among faculty members may ultimately translate into better agreement and better standardization of dental instruction.

Dr. John is Chairman, Department of Periodontics and Allied Dental Programs, Indiana University School of Dentistry; Dr. Lee is a Resident, Department of Periodontics and Allied Dental Programs, Indiana University School of Dentistry; Dr. Prakasam is Clinical Assistant Professor, Department of Periodontics and Allied Dental Programs, Indiana University School of Dentistry; Mr. Eckert is Biostatistician Supervisor, Department of Biostatistics, Indiana University School of Medicine; and Dr. Maupome is Professor, Department of Preventive and Community Dentistry, Indiana University School of Dentistry. Direct correspondence and requests for reprints to Dr. Vanchit John, Department of Periodontics and Allied Dental Programs, School of Dentistry, Indiana University, 1121 West Michigan Street, Indianapolis, IN 46202; 317-274-5124 phone; 317-274-1363 fax; vjohn@iu.edu.

Keywords: clinical education, clinical evaluation, dental education, diagnosis, educational methodology, periodontics, periodontal disease, consensus training, interrater reliability, calibration

Submitted for publication 4/12/12; accepted 8/20/12

Since considerable controversy exists in the diagnosis and treatment planning of periodontal disease,1 the American Academy of Periodontology (AAP) emphasizes the importance of diagnosing periodontal conditions based on clinical assessment.2 The pieces of information gathered during those assessments include 1) the patient’s medical, dental, social, and family history and list of medications taken; 2) periodontal charting; 3) interpretation of radiographic data, including extent and pattern of bone loss; 4) clinical oral examinations, including signs of inflammation (bleeding upon probing), extent and pattern of clinical attachment loss, and presence of plaque and calculus; and 5) miscellaneous signs and symptoms like pain and ulceration.2,3

The presence of multiple factors in the clinical decision making process involved in periodontal diagnosis and treatment planning requires a thoughtful interpretation and reasoning in order to arrive at a clinical diagnosis due to the complex pathophysiology of periodontal diseases. Most studies attempting to dissect the clinical decision making process pertain to medical specialties. These may be broadly grouped into those that are divided in stages leading to (or including at some point in time) a hypothetico-deductive approach at scrutinizing pieces of information4,5 and a separate group that postulates that clinicians recognize patterns of diagnostically meaningful arrays of information.6-8 Although it is not well established how information is collected, integrated, and made sense of in periodontal clinical reasoning, we may safely assume that determining a patient’s clinical status is achieved by collecting medical, dental, and social histories as well as intraoral and extraoral...
examination findings. Once a tentative diagnosis is established, additional data are collected in the form of periodontal charting, intraoral radiographs, and oral hygiene status. Based on the data collected, a final working diagnosis is then made.

Errors in the clinical reasoning process may occur due to inadequate knowledge, inaccurate data collection, and incorrect data interpretation. Interestingly, such errors are not uncommon. For example, Bener and Graber reported that the rate of diagnostic error in medicine may range from 5 to 15 percent. A dentist’s clinical decision making process can vary in any of the phases of clinical decision making as it does in medicine; therefore, differences in diagnosing and treatment planning can be expected. Bader and Shugars speculated that “disagreements [among clinicians] may be due to differing degrees of diagnostic thoroughness or strongly held personal opinions about appropriate treatment.” Other reports have also indicated that multiple interpretations of various data sets often lead to variations in diagnoses, which in turn lead to varying treatment plans being presented to patients. Disagreements and variations are also common in periodontal diagnosis and treatment planning, probably originating in the inclusion of multiple variables when establishing a diagnosis. Achieving consensus or even a high level of interrater agreement in periodontal diagnosis or treatment can be challenging; such a goal is generally elusive. Even periodontal epidemiologic studies are affected by inconsistencies in periodontitis case definition, mainly because of inherent instability in defining thresholds for clinical signs of inflammation. For example, a study by Costa et al. found that different periodontitis case definitions can create different prevalence and extent rates in the same periodontitis population.

The varying and, in some cases, selective use of diagnostic criteria can lead to variations in diagnosing and treatment planning in clinical practice and research. Many reports have documented inconsistencies and high variability in clinical decision making among dental faculty members. In fact, in restorative dentistry, studies have shown that there are variations and inaccuracies in diagnosing, treatment planning, and interpreting of radiographic data among dentists, dental faculty members, and dental students. In periodontics, Cosyn and De Bruyn reported that there can be a 13 to 50 percent variation in deciding when to perform periodontal surgery. Such variations have been reported anecdotally at the comprehensive care clinics of Indiana University School of Dentistry (IUSD).

Although complete elimination of such variation is probably unattainable, academic programs must emphasize the importance of future dental providers’ properly diagnosing and effectively treating patients in accordance with recommended diagnosis and treatment guidelines. Orsini et al. demonstrated that consensus training of experienced, noncalibrated examiners significantly improved their reliability and reproducibility in interpreting temporomandibular joint (TMJ) disk position and configuration in imaging studies. From this, we can speculate that training dental providers regarding periodontal decisions can have a positive effect on increasing a clinician’s reliability and reproducibility in diagnosis and treatment planning.

Similar to the Orsini et al. study, the IUSD Department of Periodontics and Allied Dental Programs has been conducting calibration sessions since 2003 to improve consensus and consistency of periodontal diagnoses and treatment plans. These regularly scheduled workshops are necessary due to high faculty turnover, created in part by an annual influx of new periodontal residents into comprehensive care clinic teaching as part-time instructors of dental students. To date, there have been few studies that have examined the effect of consensus training on both faculty members and dental students in the context of periodontal diagnosis and treatment planning. Lanning et al. found variations among preventive and periodontal faculty members and among periodontal graduate students in interpreting clinical findings, periodontal diagnoses, and treatment planning. As one component of the ongoing academic and patient care initiatives in the IUSD Department of Periodontics and Allied Dental Programs, we are starting to measure variations in periodontal diagnosis and treatment planning of clinical cases. The objectives of our study were 1) to measure variations in periodontal diagnosis and treatment planning of predoctoral periodontics faculty members (who underwent consensus training as part of their departmental calibration) and 2) to compare such variations with those of third- and fourth-year dental students.

**Methods**

The Indiana University-Purdue University Institutional Review Board (IRB) approved this study (1108006552). The study respondents were recruited from IUSD. Study respondents consisted of predoctoral periodontics faculty members and
dental students who volunteered to participate in the survey upon invitation from the student investigator. The predoctoral periodontics faculty group included full- and part-time periodontics faculty members and graduate periodontics residents (first, second, and third years) who were actively involved in the clinical education of dental students. The inclusion criteria for the predoctoral periodontics faculty group were to have a Doctor of Dental Surgery degree or equivalent, be involved in clinical/didactic teaching of periodontics for predoctoral dental students at IUSD, and have attended the graduate periodontics consensus training sessions. The dental student group consisted of third- and fourth-year students involved in providing clinical care. Participation in the study or refusal to participate did not affect dental student grades or progress in any manner; the student investigator was not involved in their clinical or didactic teaching to ensure that coercion was not a factor in students’ participation. No monetary compensation was given to students or faculty.

The participants individually responded to nine non-identifiable unique case records that were selected from previously documented presentations; a larger pool of cases were archived in a repository of case records/vignettes. Some of these presentations were previously used for consensus training exercises and/or for graduate periodontal education. The documented cases included the following information: medical history, dental history, intraoral photographs, radiographs (full mouth series), and periodontal charting that included probing depth (PD), clinical attachment loss (CAL), gingival recession, furcation involvement, and mobility. Nine patients, who were nineteen to eighty-one years of age presenting with varying degrees of periodontal diseases ranging from gingivitis to chronic and aggressive periodontitis, were selected.

In discussion among three of the authors (SL, SP, and VJ), clinical diagnoses for the nine cases were designated as indicated for reference purposes. These were as follows:

- Case 1: localized moderate chronic periodontitis
- Case 2: generalized mild chronic periodontitis (generalized mild with localized severe chronic periodontitis)
- Case 3: generalized aggressive periodontitis
- Case 4: localized severe chronic periodontitis (generalized moderate with localized severe chronic periodontitis)
- Case 5: gingivitis
- Case 6: localized aggressive periodontitis
- Case 7: generalized severe chronic periodontitis
- Case 8: generalized moderate chronic periodontitis
- Case 9: localized mild chronic periodontitis

The nine cases were made available to participants via a web-based survey (Indiana University SharePoint site, Indianapolis, IN). All data were deidentified, and the study was done in accordance with the IU IRB. No facial features were included, and no patient names or chart numbers were used in the case vignettes. Study participants independently responded to the survey through the website with their personal laptop computers at a designated lecture hall. After reviewing the cases, the participants answered a questionnaire with two demographic questions and selected one periodontal diagnosis and one treatment plan for each case from a list of multiple-choice answers. The first question inquired about the participant’s position/role at IUSD. The second question asked about the participants’ prior clinical experience and was designed to evaluate if that might impact the periodontal diagnosis and treatment planning. For each case, the questions were intended to evaluate the consensus amongst the respondents’ periodontal diagnosis and treatment plan based on a set number of predefined procedure names and quadrants. The participants were not allowed to review the cases or discuss them among each other throughout the study. A gold standard was not included in the study design.

Comparisons of the diagnosis and treatment planning between the two groups were performed using chi-square tests. Multirater kappa statistics were calculated to assess the agreement for the diagnosis and treatment responses in each of the two groups. Multirater kappa coefficients are statistical measures used to calculate the degree of agreement in classification against what can be expected by chance.3 The kappa and its standard error were used to calculate 95 percent confidence intervals for each of the two groups as well as compare the kappa coefficients between groups. A p-value of <0.05 was considered statistically significant.

**Results**

Eighteen predoctoral periodontics instructors participated in the survey: two full-time periodontal faculty members, five part-time periodontal faculty members, and four first-year, four second-year, and three third-year periodontal residents. Twenty dental students participated in the survey: ten third-year stu-
students and ten fourth-year students. As for prior dental experience, the group included one dental assistant, one dental hygienist, six general dentists, seven in another dental specialty, and twenty-three with no prior dental experience. Most of respondents with prior dental experience belonged to the faculty group.

**Distribution of Responses**

The responses of each individual case for diagnoses and treatment plans varied among predoctoral periodontal faculty and dental students (data not shown). In Cases 1, 2, 4, 7, and 9, when the severity and the extent of disease were considered, there appeared to exist a range of disagreement in the responses for periodontal diagnosis from the given choices, compared to the rest of the cases (data not shown). In Case 5, both faculty (94 percent) and dental students (95 percent) had high consensus in diagnosis for gingivitis as well as for the treatment plan of dental prophylaxis.

Interestingly, for Case 6, the third-year students were divided on the diagnosis and selected either localized aggressive periodontitis (40 percent) or localized severe chronic periodontitis (50 percent), while all the fourth-year students chose localized aggressive periodontitis (100 percent). In the faculty group, 78 percent chose localized aggressive periodontitis (Table 1). More importantly, of the nine cases in the study, only for Case 6 were treatment selections significantly different among the various groups. For this case, students favored locally delivered antimicrobials (65 percent), while faculty members favored either four quadrants of scaling and root planing (33 percent) or systemic antimicrobials (44 percent). When we compared the treatment choice of the different student groups, choices were divided between two quadrants of scaling and root planing (50 percent) and locally delivered antimicrobials (40 percent) by the third-year students and primarily locally delivered antimicrobials by the fourth-year students (90 percent) (Table 1).

These differences in treatment planning between the faculty and student groups (p<0.001) and between third- and fourth-year dental students (p=0.030) were statistically significant. There were no other statistically significant differences between faculty and students or between third- and fourth-year students for any of the other cases.

**Distribution of Responses of Combined Categories**

When diagnostic choices were combined into three discrete categories ignoring the descriptive choices of severity and extent (i.e., into categories of chronic periodontitis, aggressive periodontitis, and gingivitis), more robust consensus was observed among the respondents (Figures 1-3). Specifically, for Cases 1, 4, and 8, 100 percent of the respondents (both faculty and dental students) selected chronic periodontitis as the diagnosis. For Case 2, only one dental student respondent picked a diagnosis other than chronic periodontitis (Figure 1). For Case 7, combining the categories resulted in more agreement on the diagnosis of chronic periodontitis (67 percent for faculty and 80 percent for students) (Figure 3). For Case 9, combining the categories consolidated

<table>
<thead>
<tr>
<th>Case 6 diagnosis</th>
<th>Faculty</th>
<th>Students</th>
<th>p-value</th>
<th>Students: 3rd Year</th>
<th>Students: 4th Year</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized aggressive periodontitis</td>
<td>6%</td>
<td>0</td>
<td>0.38</td>
<td>0</td>
<td>0</td>
<td>0.0138</td>
</tr>
<tr>
<td>Generalized severe chronic periodontitis</td>
<td>6%</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Localized aggressive periodontitis</td>
<td>78%</td>
<td>70%</td>
<td>40%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Localized moderate chronic periodontitis</td>
<td>0</td>
<td>5%</td>
<td>10%</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Localized severe chronic periodontitis</td>
<td>11%</td>
<td>25%</td>
<td>50%</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Case 6 treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 quadrant scaling and root planing</td>
<td>0</td>
<td>5%</td>
<td>0.0001</td>
<td>10%</td>
<td>0</td>
<td>0.0303</td>
</tr>
<tr>
<td>2 quadrant scaling and root planing</td>
<td>6%</td>
<td>25%</td>
<td></td>
<td>50%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3 quadrant scaling and root planing</td>
<td>6%</td>
<td>15%</td>
<td></td>
<td>0</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>4 quadrant scaling and root planing</td>
<td>33%</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Locally delivered antimicrobials</td>
<td>11%</td>
<td>65%</td>
<td>40%</td>
<td>90%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Systemic antimicrobials</td>
<td>44%</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Percentages do not total 100% due to rounding.*
Agreement Among Groups

Multirater kappa tests for agreement among faculty members and for agreement among students planning were combined into one category of “scaling and root planing,” similar consensus increases were noted among respondents, except for Case 6.

Figure 1. Distribution of responses for diagnosis and treatment plans for cases 1-3

Note: The diagnosis bar graph represents three discrete diagnoses after combining the responses on the extent for aggressive periodontitis and extent and severity for chronic periodontitis. The treatment plan bar graph represents the five discrete treatment plan options after combining all the scaling and root planing responses into one group. The responses are represented as percentages of the total number of respondents.
When the diagnoses were combined into three categories (aggressive periodontitis, chronic periodontitis, and gingivitis), there were no statistically significant differences in the distributions of the diagnoses for any of the cases. After we combined the categories, kappa coefficients improved and ranged from 0.58 to 0.69. The agreement among faculty were relatively low, with all kappa coefficients ranging between 0.35 and 0.54 (Table 2). Faculty and students had similar kappa coefficients for diagnosis, but treatment agreement was lower for faculty than for students. Agreement among fourth-year students was better than agreement among third-year students for both diagnosis and treatment.

Figure 2. Distribution of responses for diagnosis and treatment plans for cases 4-6

Note: The diagnosis bar graph represents three discrete diagnoses after combining the responses on the extent for aggressive periodontitis and extent and severity for chronic periodontitis. The treatment plan graph represents the five discrete treatment plan options after combining all the scaling and root planing responses into one group. The responses are represented as percentages of the total number of respondents.
and students in terms of diagnosis and treatment planning, especially in the context of periodontics training for predoctoral dental students. While specifically targeted consensus training programs for faculty involved in predoctoral teaching may be taking place in a number of institutions, evidence of the efficacy of such programs is lacking. The present findings from our study are one attempt to try and

members was significantly higher than the agreement between students (Table 2).

Discussion

A limited number of studies have examined consensus in an academic setting among faculty

Figure 3. Distribution of responses for diagnosis and treatment plans for cases 7-9

Note: The diagnosis bar graph represents three discrete diagnoses after combining the responses on the extent for aggressive periodontitis and extent and severity for chronic periodontitis. The treatment plan graph represents the five discrete treatment plan options after combining all the scaling and root planing responses into one group. The responses are represented as percentages of the total number of respondents.
quantitatively evaluate the results of one approach to design standardization.

One of the few studies on this topic was published by Lanning et al.\textsuperscript{21} They used three web-based cases to examine variations that existed among clinical instructors (periodontists, periodontal graduate students, dental hygienists, and general dentists, for a total of twenty-seven respondents) through a questionnaire that surveyed variations in radiographic interpretation, periodontal diagnosis, and treatment planning. Their study illustrated considerable variations in periodontal diagnosis and treatment planning among clinical instructors; it also showed that periodontal graduate students had greater consistency in periodontal diagnosis than other groups that participated in the study. Lanning et al. attributed the differences between instructors and graduate students to the explanation that graduate students had indirect consensus training sessions through seminars and case conferences, while clinical faculty members did not have such opportunities to meet and discuss cases—aided by the overall emphasis on accepted diagnostic and treatment criteria.

No conclusion could be drawn from previous dental experience in our study primarily due to the relatively small sample size of subjects with prior experience as well as diversity in their experiences. Despite this, we can speculate that previous experiences may not be as significant as in the Lanning et al. study,\textsuperscript{21} suggesting that training and clinical experiences did not influence clinical instructors’ responses.

As could reasonably be expected in terms of progressive refinement of clinical skills and thinking, the third-year dental students had the lowest agreement in diagnosis, but they were not significantly lower than faculty members’ agreement. This was anticipated due to their lack of clinical exposure and hence a more limited integration between their didactic knowledge and clinical application, which have been previously reported to contribute to variability.\textsuperscript{12} Surprisingly, predoctoral periodontics faculty members showed the lowest agreement in treatment planning, but in general there were low agreements among all groups. These variations primarily stemmed from the selection of varying number of quadrants of scaling and root planing; complementarily, it can be argued that the differences can be ascribed to different deeply rooted opinions based on clinical judgment and experiences.\textsuperscript{13}

One of the areas in which we noted variations in survey responses was the extent and severity of disease. The categories of mild, moderate, and severe attachment loss, as described in the AAP classification of disease, is often used without careful analysis of the clinical and radiographic findings. This is an area that we feel will need more instruction and calibration among the faculty and students.

No conclusion could be drawn from previous dental experience in our study primarily due to the relatively small sample size of subjects with prior experience as well as diversity in their experiences. Despite this, we can speculate that previous experiences may not be as significant as in the Lanning et al. study,\textsuperscript{21} suggesting that training and clinical experiences did not influence clinical instructors’ responses.

As could reasonably be expected in terms of progressive refinement of clinical skills and thinking, the third-year dental students had the lowest agreement in diagnosis, but they were not significantly lower than faculty members’ agreement. This was anticipated due to their lack of clinical exposure and hence a more limited integration between their didactic knowledge and clinical application, which have been previously reported to contribute to variability.\textsuperscript{12} Surprisingly, predoctoral periodontics faculty members showed the lowest agreement in treatment planning, but in general there were low agreements among all groups. These variations primarily stemmed from the selection of varying number of quadrants of scaling and root planing; complementarily, it can be argued that the differences can be ascribed to different deeply rooted opinions based on clinical judgment and experiences.\textsuperscript{13}

One of the areas in which we noted variations in survey responses was the extent and severity of disease. The categories of mild, moderate, and severe attachment loss, as described in the AAP classification of disease, is often used without careful analysis of the clinical and radiographic findings. This is an area that we feel will need more instruction and calibration among the faculty and students.

Case 4 showcases how respondents may have placed different emphasis on the extent and severity in their diagnostic rationales (as proposed by Bader and Shugars in that disagreement occurs because of differences in attention to details of diagnosis as well as strongly held personal beliefs\textsuperscript{13}). Case 4 posed a

### Table 2. Summary of kappa statistics among the different groups of study participants

<table>
<thead>
<tr>
<th></th>
<th>Kappa</th>
<th>SE</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td>0.40</td>
<td>0.01</td>
<td>0.38 0.42</td>
</tr>
<tr>
<td>Faculty, excluding part-time</td>
<td>0.39</td>
<td>0.02</td>
<td>0.36 0.42</td>
</tr>
<tr>
<td>Students</td>
<td>0.40</td>
<td>0.01</td>
<td>0.38 0.42</td>
</tr>
<tr>
<td>Students: 3rd year</td>
<td>0.35</td>
<td>0.02</td>
<td>0.31 0.39</td>
</tr>
<tr>
<td>Students: 4th year</td>
<td>0.49</td>
<td>0.02</td>
<td>0.44 0.53</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td>0.37</td>
<td>0.01</td>
<td>0.34 0.40</td>
</tr>
<tr>
<td>Faculty, excluding part-time</td>
<td>0.37</td>
<td>0.02</td>
<td>0.33 0.41</td>
</tr>
<tr>
<td>Students</td>
<td>0.42</td>
<td>0.01</td>
<td>0.40 0.45</td>
</tr>
<tr>
<td>Students: 3rd year</td>
<td>0.39</td>
<td>0.02</td>
<td>0.34 0.43</td>
</tr>
<tr>
<td>Students: 4th year</td>
<td>0.54</td>
<td>0.03</td>
<td>0.48 0.59</td>
</tr>
<tr>
<td>Diagnosis: combined categories</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td>0.67</td>
<td>0.02</td>
<td>0.64 0.71</td>
</tr>
<tr>
<td>Faculty, excluding part-time</td>
<td>0.61</td>
<td>0.02</td>
<td>0.58 0.65</td>
</tr>
<tr>
<td>Students</td>
<td>0.58</td>
<td>0.04</td>
<td>0.51 0.65</td>
</tr>
<tr>
<td>Students: 3rd year</td>
<td>0.69</td>
<td>0.04</td>
<td>0.62 0.76</td>
</tr>
<tr>
<td>Students: 4th year</td>
<td>0.67</td>
<td>0.03</td>
<td>0.61 0.72</td>
</tr>
</tbody>
</table>
clinical picture of generalized moderate with localized severe chronic periodontitis. The majority of the faculty group (83 percent) arrived at one of two diagnoses: generalized moderate chronic periodontitis (44 percent) or localized severe chronic periodontitis (39 percent). Dental students (60 percent) selected generalized moderate chronic periodontitis (50 percent) or localized severe chronic periodontitis (10 percent). Interestingly, 30 percent of the dental students chose generalized severe chronic periodontitis, reflecting a gap in their understanding of the concept of extent of involvement. In the same vein, the interpretation of Case 2 was particularly challenging because it posed a clinical picture of generalized mild chronic periodontitis with localized severe chronic periodontitis. A majority of the participants (39 percent faculty, 55 percent dental students) chose generalized moderate chronic periodontitis, which lies between the two ends of the spectrum between generalized mild and localized severe chronic periodontitis.

The variations among participants in terms of treatment planning are particularly thought-provoking. Most of the respondents varied on the number of quadrants of scaling and root planing that they would choose for each case. This probably reflects the participants’ use of rigid schema governing clinical judgments as opposed to actual evaluation of treatment needs in their decision making processes. In other words, although clinicians may treat the entire mouth, they plan to code the treatment needs differently (one to four quadrants) based on these schema. Faculty members tended to choose four quadrants over other options (one to three quadrants), possibly reflecting a more nuanced understanding of compensation for time. When we combined all the treatment plans of scaling and root planing into one group rather than based on the number of quadrants, less variation was observed (Figures 1-3).

When we looked at the use of antimicrobials as part of therapy, dental students tended to choose locally delivered antimicrobials as a treatment option for localized aggressive periodontitis: for example, Case 6 had 65 percent of the dental student group selecting this option. This was different in a statistically significant manner when compared to the treatment choices made by the faculty. Additionally, when we look at the initial diagnosis by the third-year dental students (i.e., a diagnosis of localized chronic periodontitis), it was significantly different from the fourth-year dental students. Thus, Case 6 was unique among nine cases that were presented to the group. Case 6 involved a twenty-five-year-old African American female patient with general good health and minimal gingival inflammation. She had radiographic evidence of vertical bone loss around the lower central incisors and most of her first molars (teeth #3, 19, and 30). Additionally, all first molars had clinical attachment loss greater than 6 mm. A typical diagnosis for this would be localized aggressive periodontitis although a question may remain whether it is an active disease process or not. The accepted protocol for treating a diagnosis of aggressive periodontitis involves the use of adjunctive antimicrobials combined with initial therapy as previous reports have shown scaling and root planing alone can have a less than optimal outcome. Our finding may reflect lack of understanding among dental students in the distinction between utilizing local versus systemic antimicrobials. Remarkably for this case, 44 percent of the faculty members chose systemic antimicrobials for treatment as opposed to scaling and root planing (55 percent) in the context of limitations of choices of treatment.

We may interpret such findings as reflecting that either this group of faculty members assumed that scaling and root planing would be done in conjunction with systemic antimicrobials or the use of systemic antimicrobials was perceived to be paramount in the treatment of aggressive periodontitis. Conversely, in Case 3 in which the majority of the faculty group (78 percent) selected a diagnosis of generalized aggressive periodontitis for treatment, these respondents picked scaling and root planing as opposed to antimicrobials. We suggest they selected scaling and root planing as more critical in their opinions of the two treatment choices. A complete understanding of these diagnostic pathways and treatment choices in clinical decisions ought to be addressed and explored more thoroughly in future studies.

One of the challenges in the study design was the necessity to limit the number of responses to one for each case, in terms of both diagnosis and treatment plan; the caveat could be that the respondents would pick the more advanced diagnosis and its relevant, most appropriate treatment. This was a necessary compromise because allowing multiple responses for each case would make interpretation of consensus statistically challenging. The choice of Cohen’s kappa as a means to quantitatively contrast performance made it necessary to attain precise categories; choices of diagnosis and management were therefore limited to one response per survey area.
Alternatively, open-ended responses might have led to greater richness of detail in narrative but would have made it a necessity to undertake a different type of analysis (e.g., qualitative approaches). We can speculate that this feature enhanced variations in our study because in reality most clinical diagnoses and treatments do not fit into completely discrete categories. Therefore, interpretation of the results should be done with caution, keeping in mind this limitation.

An additional limitation of our study was that the answer choices were limited to closed-ended replies as opposed to open-ended options. The feedback from the participants tended to indicate a desire to offer a combination of diagnoses with regard to the extent of involvement, i.e., generalized and localized. Similarly, some participants also indicated a desire to select multiple options for the treatment plan, e.g., combining scaling and root planing with antibiotics. Such feedback will guide future research designs. The number of examiners in our study did not allow for use of open-ended questions or for allowing combinations of diagnoses/treatment plans. The small cohort size may have contributed to the lack of significance obtained in some areas. We were particularly limited in the recruitment of faculty members as inclusion criteria allowed only those involved in teaching of clinical periodontics to dental students. We recruited 100 percent of eligible faculty members. Even with such a caveat, it is worth noting that our pool of examiners was larger (forty versus twenty-seven) than in a previous key study, as well as pointing out that we included a larger number of cases in the study (nine versus three).

An important consideration in the interpretation of the results is that although there were no other statistically significant differences between faculty and students or between third- and fourth-year students for other study cases, such a finding does not automatically indicate agreement. It simply illustrates that the overall distribution of responses was not significantly different between the groups.

A definitive conclusion can be made that most respondents (third- and fourth-year dental students, graduate periodontics residents, and full-time and part-time faculty members) were able to clearly distinguish between chronic periodontitis (95-100 percent), aggressive periodontitis (67-83 percent), and gingivitis (67-95 percent). In addition, in terms of treatment plans, most respondents were able to clearly distinguish between scaling and root planing (75-100 percent) and prophylaxis (72-95 percent).

The indications for the use of antimicrobials among our respondents were not as robust (55-65 percent) (Figures 1-3).

Setting aside the fact that different cases were used in our study and in the Lanning et al. report, some parallels can be drawn for purely illustrative purposes between the two studies based on consensus achieved. Compared to the Lanning et al. study (which involved only clinical instructors), there appeared to be greater consensus in our study group. For example, the diagnoses in their study were 67 percent for gingivitis and 96 percent for chronic periodontitis compared to similar percentages of respondents in our study. Only 17 percent of respondents in the Lanning et al. study selected aggressive periodontitis for the three cases, as opposed to 55-65 percent of respondents in our study. Similarly, for treatment options, scaling and root planing was selected by 83-92 percent of their respondents and prophylaxis was selected by 75 percent, which are similar to our respondents. This is remarkable considering that our respondents included third- and fourth-year dental students unlike the Lanning et al. study.

Conclusions

While consensus training among predoctoral periodontal educators may theoretically help decrease undue variation in periodontal diagnoses and treatment plans, the ultimate goal is to calibrate and standardize criteria leading to modeling better agreement for dental students. In our study, we found variations involving diagnoses of extent and severity of periodontal conditions, as well as variation in the extent and choices of treatment planning. Some degree of variation is to be expected because there are a number of subjective factors (radiographic interpretations, clinical gingival appearances, etc.) that a practitioner must integrate for arriving at a diagnosis and subsequently a treatment plan. Our study pointed out the need to continue to develop creative concepts of teaching to attain good calibration among clinical faculty members and dental students. By establishing a baseline to identify, characterize, and document the current variation levels, we are now positioned to make targeted improvements to certain diagnostic and management factors in our consensus training. Future research will determine additional outcomes of carefully delineated improvements and how such ongoing effort may ultimately continue to ensure better quality patient care.
Acknowledgments

The authors would like to acknowledge Dr. Steven Blanchard, Dr. Michael Kowolik, and Dr. Susan Zunt for their valuable comments and feedback in preparation of this manuscript. The authors would like to acknowledge Ms. Rossok for administrative support. In addition, we express our gratitude to the study respondents who participated in this project.

REFERENCES