Evaluation of New Cone-beam Computed Tomographic Criteria for Radiographic Healing Evaluation after Apical Surgery: Assessment of Repeatability and Reproducibility

Thomas von Arx, Prof Dr med dent,* Simone F.M. Janner, Dr med dent,* Stefan Hänni, Dr med dent,‡ and Michael M. Bornstein, Prof Dr med dent* 

Abstract

Introduction: Conventional 2-dimensional radiography uses defined criteria for outcome assessment of apical surgery. However, these radiographic healing criteria are not applicable for 3-dimensional radiography. The present study evaluated the repeatability and reproducibility of new cone-beam computed tomographic (CBCT)-based healing criteria for the judgment of peri-apical healing 1 year after apical surgery. Methods: CBCT scans taken 1 year after apical surgery (61 roots of 54 teeth in 54 patients, mean age = 54.4 years) were evaluated by 3 blinded and calibrated observers using 4 different indices. Reformatted buccolingual CBCT sections through the longitudinal axis of the treated roots were analyzed. Radiographic healing was assessed at the resection plane (R index), within the apical area (A index), of the cortical plate (C index), and regarding a combined apical-cortical area (B index). All readings were performed twice to calculate the intraobserver agreement (repeatability). Second-time readings were used for analyzing the interobserver agreement (reproducibility). Various statistical tests (Cohen, kappa, Fisher, and Spearman) were performed to measure the intra- and interobserver concurrence, the variability of score ratios, and the correlation of indices. Results: For all indices, the rates of identical first- and second-time scores were always higher than 80% (intraobserver Cohen $\kappa$ values ranging from 0.793 to 0.963). The B index (94.0%) showed the highest intraobserver agreement. Regarding interobserver agreement, the highest rate was found for the B index (72.1%). The Fleiss’ $\kappa$ values for R and B indices exhibited substantial agreement (0.626 and 0.717, respectively), whereas the values for A and C indices showed moderate agreement (0.561 and 0.573, respectively). The Spearman correlation coefficients for R, A, C, and B indices all exhibited a moderate to very strong correlation with the highest correlation found between C and B indices ($\tau_c = 0.8069$). Conclusions: All indices showed an excellent intraobserver agreement (repeatability). With regard to interobserver agreement (reproducibility), the B index (healing of apical and cortical defects combined) and the R index (healing on the resection plane) showed substantial congruence and thus are to be recommended in future studies when using buccolingual CBCT sections for radiographic outcome assessment of apical surgery. (J Endod 2016;42:236–242)

Key Words

Apical surgery, cone-beam computed tomography, healing indices, radiographic healing, outcome, repeatability, reproducibility

The evolution of modern apical surgery is based on the enhancement of visualization (surgical microscope, endoscope), the introduction of microsurgical techniques and microinstruments, and the use of biocompatible and cement-inducing root-end filling materials (1–4).

The introduction of cone-beam computed tomographic (CBCT) imaging has been another milestone in modern endodontics (5). Although CBCT imaging has gained a wide reputation for diagnostics and treatment planning, CBCT scanning does not yet have the same impact on the assessment of endodontic treatment outcome (6). The latter is mainly because of economic aspects and concerns about radiation dose (7). In apical surgery, CBCT imaging has been shown to be an important tool for case assessment and treatment planning (8, 9). However, the actual benefit of CBCT imaging for treatment outcome and its cost-effectiveness (cost-benefit analysis) have not yet been elaborated in conjunction with conventional or surgical endodontics.

Recently, some clinical articles have documented the use of CBCT imaging for postoperative evaluation in apical surgery (10–12). These studies either compared periapical radiography (PA) and CBCT imaging in the outcome assessment of “radiographic healing” based on the absence or presence of radiolucencies in follow-up radiographs or measured the size of persistent lesions using CBCT imaging. Although some authors have suggested a new CBCT periapical index for the evaluation of endodontic lesions before treatment (13, 14), that index is not applicable and useful for CBCT assessment of the outcome of apical surgery.

Traditionally, the PA criteria for healing assessment of apical surgery are based on the work by Rud et al (15) and Molven et al (16). However, a recent study has shown that these 2-dimensional criteria may not be valid for the evaluation of 3-dimensional
A new experimental dog study has suggested scoring criteria for CBCT evaluation of teeth treated with apical surgery (17).

The objective of the present study was to evaluate these new CBCT-based criteria in a prospective clinical study for the radiographic outcome assessment 1 year after apical surgery.

Materials and Methods

The present study was approved by the Ethics Committee of the State of Bern, Switzerland (approval number KEK-BE 098/11) (18). Patients were enrolled provided that they agreed on having presurgical and 1-year follow-up PA and CBCT radiographs in conjunction with apical surgery. Patient information, recruitment, and treatment were performed according to the Declaration of Helsinki 2013.

All surgeries were performed using local anesthesia in an operating room by the same surgeon (19). All surgical steps were performed using a surgical microscope. A rigid endoscope (Hopkins Tele Otoscope; Karl Storz GmbH, Tuttingen, Germany) was used for intraoperative diagnostics during root-end management (resection, cavity preparation, and filling). All treated roots were filled with MTA (ProRoot Dentsply Tulsa Dental, Tulsa, OK).

Patients were recalled 1 year after surgery for healing assessment including a clinical examination, PA, and CBCT imaging. For the present analysis, only 1 tooth was randomly included (Quickcalcs; GraphPad Software Inc, La Jolla, CA) in patients who had apical surgery of multiple teeth.

The CBCT images were obtained with 3D Accuitomo 170 (J Morita Manufacturing Corp, Kyoto, Japan). The normal field of view was $4 \times 4$ cm or $6 \times 6$ cm for preoperative and $4 \times 4$ cm for postoperative CBCT imaging. The parameters of the recordings were 3.0 mA and 80 kV with an exposure time of 17.5 seconds. CBCT images were evaluated on a Dell 380 workstation (Dell SA, Geneva, Switzerland) and a 19-inch

### Table 1. Definition of R, A, and C Indices According to Chen et al (17)

<table>
<thead>
<tr>
<th>Index</th>
<th>Definition</th>
<th>Score 0</th>
<th>Score 1</th>
<th>Score 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>Resection plane (cut root face)</td>
<td>No bone deposition (no formation of PDL space)</td>
<td>Partial bone deposition (partial formation of PDL space)</td>
<td>Complete bone deposition (complete formation of PDL space of normal width)</td>
</tr>
<tr>
<td>A</td>
<td>Apical area (former bone defect)</td>
<td>No apparent bone formation</td>
<td>Partial bone formation</td>
<td>Complete bone formation</td>
</tr>
<tr>
<td>C</td>
<td>Cortical plate (access bone window)</td>
<td>Not re-established</td>
<td>Re-established, but concave</td>
<td>Re-established and flat</td>
</tr>
</tbody>
</table>

PDL, periodontal ligament.

Figure 1. A schematic illustration of the R index (bone deposition along the resection plane).

Figure 2. A schematic illustration of the A index (bone formation within the apical area).
on the R, A, and C indices (not involved in the radiographic assessment. The evaluation was based on the R, A, and C indices (Resection plane, Apical area, and Cortical plate) as suggested by Chen et al(17). The criteria of the scores are presented in Table 1 and Figures 1–3. Three observers were calibrated with 1-year follow-up CBCT images of teeth that had been previously treated with apical surgery and that did not belong to the present study cohort. No time limit was set for image evaluation, and the observers were allowed to use the software tools for image editing (size, brightness, and contrast). All images were evaluated twice with a minimum interval of 2 weeks to determine the intraobserver agreement (repeatability). Second-time readings were used for interobserver analysis (reproducibility) and healing judgment of the treatment.

Furthermore, a separate radiographic analysis was performed by the same 3 observers using a so-called B index. The latter was created to simplify the periapical radiographic assessment in combining the A and C indices of the initially applied indices. The B index was defined as follows:

- **B-score 2**: Complete hard tissue ("bony trabecular") fill of former lesion/osteotomy site and formation of an intact cortical plate in its anatomically correct shape
- **B score 1**: Any situation not attributable to B scores 0 or 2
- **B score 0**: Neither hard tissue fill of former lesion/osteotomy site nor formation of cortical plate

**Statistical Analysis**

For the intraobserver agreement, the Cohen kappa values were calculated for 2 ratings using squared weights (20). For the interobserver agreement between second-time readings of R, A, and C, and B indices, Fleiss’ kappa values were calculated. The Fisher exact test was used (confidence level of 95%) to check whether the ratios of agreement between the observers were significantly different from each other. Spearman correlation coefficients were calculated to analyze correlations between the indices. All tests were performed using the Internet-based R software package (R 3.0.3; www.r-project.org; The R Foundation, Vienna, Austria).

**Results**

Fifty-four teeth in the same number of patients fulfilled the criteria for CBCT assessment 1 year after apical surgery. At the time of surgery, patients had a mean age of 54.4 ± 10.2 years (range, 24–73 years) and

**TABLE 2.** Overview of Evaluated Teeth (N = 54) and Roots (N = 61)

<table>
<thead>
<tr>
<th></th>
<th>N teeth</th>
<th>N roots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maxilla</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central incisors</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Lateral incisors</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Canines</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1st premolars</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2nd premolars</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>1st molars</td>
<td>5</td>
<td>6 (5 x mesiobuccal, 1 x distobuccal)</td>
</tr>
<tr>
<td>Subtotal</td>
<td>34</td>
<td>35</td>
</tr>
<tr>
<td>Mandible</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canines</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2nd premolars</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1st molars</td>
<td>16</td>
<td>22 (15 x mesial, 7 x distal)</td>
</tr>
<tr>
<td>Subtotal</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>54</td>
<td>61</td>
</tr>
</tbody>
</table>

Eizo monitor with a resolution of 1280 × 1024 pixels (Eizo Nanao AD, Widenswil, Switzerland).

For further analysis, reformatted CBCT scans in the buccolingual plane through the center of the root along its longitudinal axis were evaluated. The relevant section had been preselected by a surgeon who was not involved in the radiographic assessment. The evaluation was based on the R, A, and C indices (Resection plane, Apical area, and Cortical plate) as suggested by Chen et al (17). The criteria of the scores are presented in Table 1 and Figures 1–3. Three observers were calibrated with 1-year follow-up CBCT images of teeth that had been previously treated with apical surgery and that did not belong to the present study cohort. No time limit was set for image evaluation, and the observers were allowed to use the software tools for image editing (size, brightness, and contrast). All images were evaluated twice with a minimum interval of 2 weeks to determine the intraobserver agreement (repeatability). Second-time readings were used for interobserver analysis (reproducibility) and healing judgment of the treatment.

**TABLE 3.** Healing Rates per Index (N = 61)

<table>
<thead>
<tr>
<th>Index</th>
<th>Complete healing (%)</th>
<th>Partial healing (%)</th>
<th>No healing (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R (Resection plane)</td>
<td>54.1*†</td>
<td>29.5</td>
<td>16.4¹</td>
</tr>
<tr>
<td>A (Apical area)</td>
<td>47.5*</td>
<td>37.7</td>
<td>14.8¹</td>
</tr>
<tr>
<td>C (Cortical plate)</td>
<td>21.3*</td>
<td>32.8</td>
<td>45.9¹</td>
</tr>
<tr>
<td>B (one)</td>
<td>34.3*</td>
<td>45.9</td>
<td>19.7</td>
</tr>
</tbody>
</table>

*Statistically significant overall difference between R, A, and C values; Fisher exact test, P = .0003908.
†Statistically significant overall difference between R, A, and C values; Fisher exact test, P = .001033.
¹Statistically significant difference between R and B values; Fisher exact test, P = .045.
included 25 men (46.3%) and 29 women (53.7%). The total number of evaluated roots was 61 (Table 2).

With regard to healing classification, none of the cases had a study parameter with 3 different scores; hence, at least 2 observers always agreed on the same healing score per case and index. Based on a majority classification (same rating by at least 2 observers), the resection plane exhibited the highest rate of complete healing (54.1%), whereas rates of complete healing of the apical area (47.5%) and the cortical plate (21.3%) were lower. Complete healing with regard to the B index was 34.3% (Table 3).

### Repeatability (Intraobserver Agreement)

The rates of identical first- and second-time scores per study parameter and observer were always higher than 80% (Table 4). Scores never differed >1 when comparing first- and second-time ratings. On average (all observers), the B index (94.0%) showed the highest intraobserver agreement, which was significantly different from the R (\(\kappa = .0626\)) and A indices (\(\kappa = .561\)) but not the C index (\(\kappa = .573\)) (Table 5). The Fleiss’ kappa values (20) were as follows: <0, no agreement; 0–0.2, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–1, almost perfect agreement.

### Reproducibility (Interobserver Agreement)

For all study parameters, scores across the 3 observers never differed >1; hence, 2 observers always agreed on the same score for each case and index. The highest overall agreement (all 3 observers giving the same score) was found for the B index (72.1%), whereas the reproducibility regarding healing assessment of the apical area and the cortical plate (each 59.0%) had the lowest reproducibility (Table 5). The Spearman correlation coefficients for R, A, C, and B indices all exhibited a moderate to very strong correlation (Table 6). The highest correlation was found between C and B indices (\(r_s = 0.8069, very strong correlation\)), whereas the lowest correlation was observed between R and C indices (\(r_s = 0.5945, moderate correlation\)).

### Discussion

The present prospective clinical study evaluated the repeatability and reproducibility of new CBCT-based radiographic indices for assessing the 1-year healing outcome of apical surgery (Figs. 4–7).

### Material

Although this study had a prospective design, patients could not be enrolled consecutively because some of the patients declined to have a CBCT scan taken 1 year after apical surgery. The number of evaluated teeth (\(N = 54\)) is comparable with the number (\(N = 52\)) in the study by Christiansen et al (10) but larger than the number (\(N = 11\)) in the study by Tanomaru-Filho et al (11); both studies also used CBCT imaging for radiographic outcome evaluation after apical surgery but with different objectives compared with the present study.
Methodology

According to current guidelines, CBCT imaging should not be used for postsurgical and follow-up evaluation in apical surgery, mostly because of radiation issues (7, 21, 22). However, previous studies have shown that conventional PA provides a different (ie, most often a better and hence an overestimated) radiographic judgment of periapical healing compared with CBCT imaging (10, 12). Hence, the present study was designed and approved by the institutional review board to provide further insight with regard to repeatability and reproducibility of CBCT-based healing criteria. Although histology remains the gold standard and reference for healing assessment after apical surgery, surgical harvesting of periapical tissue at the 1-year follow-up was judged “nonethical” and discarded.

The rationale for using CBCT imaging (instead of histology) is based on a recent experimental dog study reporting that the results obtained from CBCT imaging corresponded to the histologic results 6 months after apical surgery (17). Similar to Chen et al’s study parameters (17), the present investigation determined the healing on the resected root-end surface as well as the healing of the apical area and the cortical plate. From a clinical perspective, the healing on the resected root-end surface is the most critical because bone formation along the cut root face indicates that the root-end filling is bioinert, bioactive, and bacteria tight. However, such minute tissue healing is not visible on PA, and healing on the cut root face may only be inferred from the reformation of a periodontal ligament space.

The rationale for using buccolingual CBCT scans was mainly a practical one because mesiodistal CBCT scans of roots with an ectopic position (maxillary incisors, canines, buccal roots of first premolars) would be located outside the “bony housing,” with the apical area located external to the surrounding cortices (Fig. 8).

Healing Outcome

The highest percentage of complete healing (54.1%) was observed for the R index and the lowest (21.3%) for the C index. Cases with partial healing on the resection plane showed that bone is “creeping” over the cut root face from buccal and lingual aspects, eventually forming a bony bridge at the resection level (Fig. 5). Re-establishment of the cortical plate apparently takes longer compared with R and A indices because it is the furthest away from the concentric growth pattern of new bone formation and because there is no hard tissue surface (access window) for new bone deposition. Hence, cortical reformation has to await the infill (A index) of the bony crypt up to the level of the buccal cortex.

With regard to the B index, only 34.3% of the evaluated cases were considered completely healed in the present study; in other words,
65.7% still presented radiolucency in the apical region or incomplete cortical plate healing 1 year after apical surgery. Similarly, Christiansen et al (10) reported that in 69% of the cases a defect was detected on both the coronal and sagittal CBCT images 1 year after apical surgery. These figures indicate that radiographically detectable complete hard tissue healing may take well beyond the usual 1-year follow-up period. In contrast, clinical long-term studies have shown that only a few cases considered healed after 1 year will fail thereafter (23–26).

Repeatability
The assessment of the intraobserver first- and second-time readings showed that irrespective of the study parameter, there was substantial or almost perfect agreement (kappa values ranging from 0.793 to 0.963) among the 3 observers. This emphasizes the usefulness of the chosen indices. Because the B index showed higher agreement than the A and C indices among 3 observers and also combines apical and cortical healing judgment, thus simplifying the rating procedure, it is suggested to use the B index in future CBCT studies regarding radiographic healing evaluation after apical surgery.

Reproducibility
The present study only accepted a rigorous interobserver agreement (ie, scores had to be identical for all observers). Often studies including multiple observers with different ratings try to reach a consensus rating, with the risk that 1 observer has to give in. For interobserver agreement, the B index again showed a higher value (72.1%) than the A and C indices (each 59%). Also, the R index (67.2%) had a higher interobserver agreement than the A and C indices. When using buccolingual CBCT images for healing judgment after apical surgery, the R and B indices are recommended for clinical practice and future research based on our results.

Unfortunately, the agreement rates obtained in this study cannot be compared with data from other studies because such data have not been reported. In Christiansen et al’s study (10), the authors remarked that there was considerable variation between the observers’ detection of a remaining defect using CBCT imaging, but the actual interobserver agreement rate was not provided.

Conclusion
To our knowledge, this is the first clinical study to apply new CBCT-based criteria for radiographic healing evaluation 1 year after apical surgery. All study parameters showed excellent intraobserver agreement (repeatability). With regard to interobserver agreement (reproducibility), the B index (healing of apical and cortical defects combined) and the R index (healing on the resection plane) showed substantial congruence and thus are to be recommended in future studies when using buccolingual CBCT sections for radiographic outcome assessment of apical surgery.

Acknowledgments
The authors thank Anja MIthlemann, Institute of Mathematical Statistics and Actuarial Science, University of Bern, Switzerland, for her assistance regarding the statistical analysis and Bernadette Rauyler, Medical Illustrator, School of Dental Medicine, University of Bern, Switzerland, for the schematic illustrations.

The authors deny any conflicts of interest related to this study.

References
Clinical Research


