Detection accuracy of root fractures in cone-beam computed tomography images: a systematic review and meta-analysis

R. H. Ma, Z. P. Ge & G. Li
Department of Oral and Maxillofacial Radiology, Peking University School and Hospital of Stomatology, Beijing, China

Abstract

The aim of this review was to evaluate whether CBCT is reliable for the detection of root fractures in teeth without root fillings, and whether the voxel size has an impact on diagnostic accuracy. The studies published in PubMed, Web of Science, ScienceDirect, Cochrane Library, Embase, Scopus, CNKI and Wanfang up to May 2014 were the data source. Studies on nonroot filled teeth with the i-CAT (n = 8) and 3D Accuitomo CBCT (n = 5) units were eventually selected. In the studies on i-CAT, the pooled sensitivity was 0.83 and the pooled specificity was 0.91; in the 3D Accuitomo studies, the pooled sensitivity was 0.95 and pooled specificity was 0.96. The i-CAT group comprised 5 voxel size subgroups and the 3D Accuitomo group contained 2 subgroups. For the i-CAT group, there was a significant difference amongst the five subgroups (0.125, 0.2, 0.25, 0.3 and 0.4 mm; P = 0.000). Pairwise comparison revealed that 0.125 mm voxel subgroup was significantly different from those of 0.2, 0.25 and 0.3 mm voxel subgroups, but not from the 0.4 mm voxel subgroup. There were no significant differences amongst any other two subgroups (by α = 0.005). No significant difference was found between 0.08 mm and 0.125 mm voxel subgroups (P = 0.320) for the 3D Accuitomo group. The present review confirms the detection accuracy of root fractures in CBCT images, but does not support the concept that voxel size may play a role in improving the detection accuracy of root fractures in nonroot filled teeth.

Keywords: cone-beam computed tomography, meta-analysis, root fracture, systematic review.

Received 15 April 2015; accepted 19 June 2015

Introduction
Early detection of fractured roots is vital to prevent extensive damage to the supporting tissues (Tang et al. 2011). Although imaging is an important diagnostic adjunct to the clinical assessment of root fracture (Scarfe & Farman 2008), they can be overlooked when the X-ray beam does not pass along the fracture line (Avsever et al. 2014). Recently, 3D images using various computed tomography (CT) methods have been adopted to overcome the inherent disadvantages of conventional 2D radiographic methods, for example magnification, distortion and anatomic superimposition (Avsever et al. 2014). The introduction of cone-beam computed tomography (CBCT) specifically dedicated to imaging the maxillofacial region heralds a true paradigm shift from a 2D to a 3D approach to data acquisition and image reconstruction (Scarfe & Farman 2008). CBCT is capable of providing images with submillimetre resolution in a short scanning time (10–70 s), and radiation dosages.
are reportedly up to 15 times lower than those of conventional CT scans (Scarfe et al. 2006).

Numerous studies have been performed to identify the utility of a CBCT scan in the detection of root fractures both in vitro and in vivo. In these studies, the effect of voxel size on the detection accuracy of root fractures varies (Wenzel et al. 2009, Melo et al. 2010, Öner 2011, Junqueira et al. 2013, Amintavakoli 2014). Considering the fact that the smaller the voxel size used for capturing image, the larger the radiation dose exposed to the patient, a systematic review exclusively focused on the effect of voxel size on the detection accuracy of root fractures on CBCT images is necessary.

Method

Search strategy

Studies published in PubMed, Web of Science, ScienceDirect, Cochrane library, Embase, Scopus, CNKI and Wanfang up to May 2014 were searched. To avoid references omitted from electronically searching, a hand search for relevant references was also performed. The detailed search strategy is presented in Table 1.

Study selection and data extraction

The electronic search was carried out by two individuals (RHM and GZP). No language restriction was applied in the search. Two researchers (RHM and ZPG) independently removed duplicates and selected the studies on the basis of titles and abstracts using EndNote, version 17 (Thomson ResearchSoft, Stamford, CT, USA). Disagreements were resolved by discussion or referred to experts. Studies that were included were as follows: original in vivo or in vitro research, focused on root fracture and CBCT. The literature with the following conditions were excluded: sample size smaller than ten, animal trials, incomplete information about sensitivity/specificity or other important indicators. One study (Jakobson et al. 2014) on incomplete root fracture was also excluded because the most common root fracture clinically is complete fracture (Walton et al. 1984). Data characteristics from each selected study were extracted as follows: sample size, fracture type, the model of CBCT, parameters (include voxel size, field of view, tube voltage, tube current and time), reference test, evaluation criterion and score type, the number of true positive, false positive, false negative and true negative cases were recorded independently (Table S1).

Assessment of quality

QUADAS-2 (Quality Assessment of Studies of Diagnostic Accuracy-2) (Whiting et al. 2011) was used to assess the quality of the included studies. The assessments were estimated independently and checked by two researchers (RHM and ZPG). Disagreements were resolved by discussion or referred to experts. This quality assessment tool comprises 4 domains: patient selection, index test, reference

<table>
<thead>
<tr>
<th>Database</th>
<th>Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>PubMed (MEDLINE)</td>
<td>&quot;Cone-Beam Computed Tomography&quot;[Mesh] OR &quot;spiral cone-beam computed tomography&quot;[mesh] OR cone-beam ct OR cbct OR cone-beam AND ((&quot;Tooth Fractures/diagnosis&quot;[Mesh] AND &quot;Tooth Fractures/radiography&quot;[Mesh]) OR root fractur* OR tooth fracture* OR dental fracture*)</td>
</tr>
<tr>
<td>Web of Science</td>
<td>Subject: (&quot;dental fractur*&quot; OR &quot;root fractur*&quot; OR &quot;tooth fractur*&quot;) AND Subject: (&quot;cone-beam computed tomography&quot; OR cbct OR &quot;cone-beam ct&quot; OR cone-beam)</td>
</tr>
<tr>
<td>ScienceDirect</td>
<td>TITLE-ABSTR-KEY(&quot;cone-beam computed tomography&quot; OR &quot;cone-beam ct&quot; OR cbct) and TITLE-ABSTR-KEY(&quot;root fractur*&quot; OR &quot;dental fractur*&quot; OR &quot;tooth fractur*&quot;)</td>
</tr>
<tr>
<td>Cochrane Library</td>
<td>&quot;cone-beam computed tomography&quot; OR &quot;cbct&quot; OR &quot;cone-beam ct&quot; OR cone-beam:ti,ab,kw in Other Reviews, Trials and Methods Studies (Word variations have been searched) and &quot;tooth fractur*&quot; OR &quot;root fractur*&quot; OR &quot;dental fractur*&quot;:ti,ab,kw in Other Reviews, Trials and Methods Studies (Word variations have been searched)</td>
</tr>
<tr>
<td>Embase</td>
<td>&quot;cone-beam computed tomography&quot; OR &quot;cbct&quot; OR &quot;cone-beam ct&quot; OR cone-beam:ti,ab,kw in Other Reviews, Trials and Methods Studies (Word variations have been searched) and &quot;tooth fractur*&quot; OR &quot;root fractur*&quot; OR &quot;dental fractur*&quot;:ti,ab,kw in Other Reviews, Trials and Methods Studies (Word variations have been searched)</td>
</tr>
<tr>
<td>Scopus</td>
<td>(&quot;cone-beam computed tomography&quot; OR &quot;cone-beam ct&quot; OR cone-beam-or cbct) and (&quot;root fractur*&quot; OR &quot;dental fractur*&quot; OR &quot;tooth fractur*&quot;)</td>
</tr>
<tr>
<td>CNKI</td>
<td>Subject – cone beam CT or Subject – CBCT and Subject – root fracture (match exactly)</td>
</tr>
<tr>
<td>Wanfang</td>
<td>Subject:cone beam CT or CBCT and Subject:root fracture</td>
</tr>
</tbody>
</table>

standard, and flow and timing (Whiting et al. 2011). Each item is scored as ‘+’ (yes), ‘−’ (no) or '?' (unclear). This procedure was undertaken in Review Manager, version 5.2 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark).

Data synthesis and analysis

To exclusively study the effect of voxel size on root fractures with a meta-analysis, there must be more than two studies performed on the same CBCT units. Thus, the studies on the most popular CBCT units, i-CAT and 3D Accuitomo, were chosen. As the influence of root fillings (gutta-percha or metallic post) remains unclear (Costa et al. 2012, 2014, Khedmat et al. 2012, Moudi et al. 2014), the related studies were excluded. As a result, only the studies performed using i-CAT and/or 3D Accuitomo CBCT units with data relating to nonendodontically treated teeth and valid information on voxel size were included (Table S1).

Using the random-effects model for the two groups of studies, the pooled sensitivity (SEN), pooled specificity (SPE), pooled positive likelihood ratio (+LR) and pooled negative likelihood ratio (−LR) were calculated. Threshold analysis for each group (group i-CAT and group 3D Accuitomo) was also carried out to ensure that there were no threshold effects so that the focus would be on analysing other heterogeneities. For the group which may represent substantial heterogeneity (50% < I² < 90%) (Higgins & Green 2011), a meta-regression analysis was conducted to determine the source of heterogeneity.

Then, the correlation between voxel size and diagnostic accuracy was analysed. The i-CAT group contained five voxel size subgroups and the 3D Accuitomo group two subgroups. The average diagnostic accuracy for each subgroup was calculated and the statistical difference amongst subgroups was determined using chi-square tests. For the i-CAT group, a Bonferroni correction at α = 0.005 [α = α/(k(k−1)/2); α = 0.05, k = number of groups = 5] (Armstrong 2014) was used to decrease the risk of a type I error in pairwise comparisons. The computational process was carried out using Meta-DiSc, version 1.4 (http://www.hrc.es/investigacion/metadisc.html) and SPSS, version 13.0 (SPSS Inc, Chicago, IL, USA).

Results

In total, 513 studies were found with only 244 studies remaining after removing duplicates. Of these, 169 were excluded because their themes were not root fracture or they did not focus on the use of CBCT. Of the 75 remaining studies, 39 were excluded for various reasons such as sample size smaller than 10 teeth and duplicate content. In the remaining 36 studies, 12 were finally selected in accordance with the inclusion criteria (Fig. 1).

The final meta-analysis included 7 studies (Wenzel et al. 2009, Melo et al. 2010, 2013, Özer 2011, Junqueira et al. 2013, da Silveira et al. 2013, Nascimento et al. 2014) on the i-CAT and 4 studies (Kambaroğlu et al. 2009, 2013, Ning et al. 2011, Avsever et al. 2014) on the 3D Accuitomo along with one study that included both the i-CAT and 3D-Accuitomo CBCTs (Brady et al. 2014). The overall methodological quality of the studies in the i-CAT CBCT was better than that of the studies in the 3D Accuitomo CBCT (Fig. 2). The i-CAT CBCT studies included a total of 372 teeth, in which 178 teeth were fractured. In the 3D Accuitomo CBCT studies, a total of 237 teeth were included amongst which 118 teeth were fractured.

Pooled statistics

Figure 3 presents the forest plots of sensitivity (true positive/(true positive + false negative)), specificity (true negative/(true negative + false positive)), pooled SEN, pooled SPE and the I² value. I² statistically describes the percentage of variation due to differences amongst studies of the total variation (Zhang et al. 2014). The premise to analyse and discuss the nonthreshold effective heterogeneity is the absence of threshold effect. For the i-CAT group, the spearman correlation coefficient was −0.471 and P-value was 0.066. The corresponding values for the 3D Accuitomo group were 0.061 and 0.936. These results indicate that no threshold effect exists in both groups (by α = 0.05). For the i-CAT group, the pooled SEN was 0.83 (95% confidence interval (95%CI): 0.78 to 0.86) with I² = 62.0% and the pooled SPE being 0.91 (95%CI: 0.87 to 0.93) with I² = 63.9%, and for the 3D Accuitomo group, the pooled SEN was 0.95 (95%CI: 0.90 to 0.98) with I² = 0.0% and the pooled SPE being 0.96 (95%CI: 0.92 to 0.99) with I² = 0.0%. For the i-CAT group, the type of root fracture and voxel size was added into the meta-regression analysis, but neither of these factors caused heterogeneity (P-value was 0.27 for type and 0.55 for voxel). For the i-CAT group, the +LR was 8.36 and the −LR was 0.18. The corresponding values for the 3D Accuitomo group were 15.44 and 0.07, respectively.
Chi-square test

For the i-CAT group, the number of experimental subgroups for voxel size subgroups of 0.125 mm, 0.2 mm, 0.25 mm, 0.3 mm and 0.4 mm was 5, 5, 2, 3 and 2, respectively. The average value of diagnostic accuracy was computed for each subgroup. A significant difference was found amongst the five voxel size subgroups ($P = 0.000$). Pairwise comparison for every two subgroups is shown in Table 2. The diagnostic accuracy of the 0.125 mm voxel size subgroup was significantly different from those of 0.2 mm, 0.25 mm and 0.3 mm subgroups, respectively, but not from the 0.4 mm voxel subgroup. No significant difference was found amongst any other two subgroups (by $\alpha' = 0.005$). For the 3D Accuitomo group, there were only 2 voxel size subgroups: 0.08 mm and 0.125 mm. The number of experimental groups for both the voxel size of 0.08 mm and 0.125 was three. Chi-square tests were also conducted for these two subgroups with no significant difference between the 0.08 mm and 0.125 mm voxel sizes ($P = 0.320$).

Discussion

Originally, the pooled outcomes for all of the related studies was derived (36 studies/86 groups of data), and it was noted that $I^2$ was 85.5% for pooled sensitivity and 81.7% for pooled specificity. These results implied that the heterogeneity amongst studies was considerable ($75% < I^2 < 100\%$) (Higgins & Green 2011). Furthermore, the pooled outcome which was
Figure 2  Evaluation of methodological quality of the studies included. (a) i-CAT CBCT studies; (b) 3D Accuitomo CBCT studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Risk of bias</th>
<th>Applicability concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient selection</td>
<td>Index test</td>
</tr>
<tr>
<td>Brady 2014</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>da Silveira 2013</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Junqueira 2013</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Melo 2010</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Melo 2013</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Nascimento 2014</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Özer 2011</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Wenzel 2009</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Risk of bias</th>
<th>Applicability concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient selection</td>
<td>Index test</td>
</tr>
<tr>
<td>Avsever 2014</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Brady 2014</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Kamburoğlu 2009</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Kamburoğlu 2013</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ning 2011</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

- High risk
- Unclear risk
- Low risk
Figure 3 Forest plot of the studies included. (a) i-CAT CBCT studies; (b) 3D Accuitomo CBCT studies.
obtained in this case was neither scientific nor reasonable. In other words, it is inappropriate to conduct a meta-analysis for all 36 studies together. When subgroups-analysis of these studies was taken into consideration, the sources of inconsistency were so large that the result could not be interpreted. This may be due to the different CBCT machine used for study (Hassan et al. 2010). Therefore, two of the most representative CBCT units in which the number of conducted studies was enough to perform a meta-analysis were chosen. Eight studies on the i-CAT CBCT and five studies on the 3D Accuitomo CBCT were finally included.

In the i-CAT group, the pooled SEN was 0.83 and the pooled SPE was 0.91. The corresponding values in the studies on the 3D Accuitomo were 0.95 and 0.96, respectively. These data indicate that the detection accuracy of root fractures in untreated teeth on CBCT images is sufficient for clinical purpose.

It is generally acknowledged that if \( + \text{LR} > 10 \), the method could be used to diagnose the studied disease; if \( - \text{LR} < 0.1 \), the method could be used for the exclusion of the disease. However, if LR equals 1, the method has no diagnostic value (Zhang et al. 2014). From the results of the meta-analysis, the statistics of LRs for i-CAT studies or that of 3D Accuitomo was close to or greater than the threshold value. This shows clearly that CBCT is reliable for the diagnosis of root fractures in nonendodontically treated teeth.

Although a general statistical analysis indicates a significant difference amongst the i-CAT subgroups, no significant difference was found for the pairwise comparison between the voxel size subgroups of 0.125 mm and 0.4 mm. Furthermore, the \( P \)-value for the pairwise comparison of voxel size subgroups of 0.125 mm and 0.25 mm was 0.004, a value close to the border value of 0.005. This may be caused by a number of factors, such as sample size, parameter settings or subjective factors of researchers in the individual studies. In addition, different generations of the i-CAT CBCT units used in these studies along with the proprietary software used for processing and observing the images may also play a role. In these studies, at least four models of i-CAT CBCT units were involved (Table S1).

Considering the above and the results from the 3D Accuitomo, it cannot be concluded that the smaller the voxel size, the higher the diagnostic accuracy of dental root fracture is in nonroot filled teeth. This outcome is similar to the result from the study by Amintavakoli (2014), in which four different voxel sizes were used in the KODAK 9000 3D CBCT unit for the detection of both vertical and horizontal root fractures in vitro.

There was only one review (Long et al. 2014), which was based on a systematic method, focusing on the use of CBCT in diagnosis of root fractures. It summarized the in vivo findings of the available studies from January 1990 to April 2013. There were twelve investigations on diagnostic accuracy of in vivo root fracture using CBCT. The advantage of this review was that all of the included studies were in vivo, leading to a higher level of evidence. However, seven of twelve studies included in the review did not describe the details surrounding the blinding method. This will influence the accuracy and validity of sensitivity and specificity and introduce bias into the combined values, which may lead to limitations in methodological quality.

ALARA (the acronym for As Low As Reasonably Achievable) is a fundamental principle for diagnostic radiology (Farman 2005), and this is an important concept in practical applications. It is well known that the smaller the voxel size, the greater the dose exposed to patient (Ngan et al. 2003). The pooled statistics in the present review does not support the hypothesis that image quality is enhanced when the voxel size is smaller.

It is worth noting that voxel size is not equal to spatial resolution, which is one of parameters used to describe image quality. For easy comparison, voxel size is usually used for the calculation of theoretical spatial resolution. However, an effective spatial resolution available in CBCT images is affected by the two-dimensional detector, the three-dimensional reconstruction process, patient movement during the scan and various other parameters (Brüllmann & Schulze 2015), not by voxel size per se.

### Conclusion

The present review confirmed that CBCT images are accurate for the detection of root fractures in nonendodontically treated teeth. The diagnostic accuracy of

### Table 2

<table>
<thead>
<tr>
<th>P-values</th>
<th>0.2 mm</th>
<th>0.25 mm</th>
<th>0.3 mm</th>
<th>0.4 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.125 mm</td>
<td>0.000*</td>
<td>0.004</td>
<td>0.000</td>
<td>0.028</td>
</tr>
<tr>
<td>0.2 mm</td>
<td>0.521</td>
<td>0.649</td>
<td>0.239</td>
<td></td>
</tr>
<tr>
<td>0.25 mm</td>
<td>0.348</td>
<td>0.620</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3 mm</td>
<td>0.153</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant difference at \( P < 0.005 \).
root fractures in root filled teeth or teeth with posts still needs further investigation. Voxel size does not impact on the diagnostic accuracy of root fracture in nonroot filled teeth.

Acknowledgements

We would like to express our sincere appreciation for Dr. Yannan Sun (Department of Orthodontics, Peking University School and Hospital of Stomatology) in help with preparation of this systematic review. The authors deny any conflict of interests related to this study.

References


Supporting Information

Additional Supporting Information may be found in the online version of this article: Table S1. Details of the included and excluded studies.