

R099

D. Sebring^{1,*}, P. Jonasson¹, T. Kvist¹, K. Buhlin² & H. Lund³

¹*Department of Endodontology, Institute of Odontology, Sahlgrenska Academy, Gothenburg*, ²*Department of Dental Medicine, Division of Periodontology, Karolinska Institutet, Huddinge* & ³*Department of Oral and Maxillofacial Radiology, Institute of Odontology, Sahlgrenska Academy, Gothenburg, Sweden*

Calibration for increased reliability in assessment of endodontic variables in panoramic radiographs

Aim To evaluate the influence of calibration on examiner variation in assessment of endodontic variables in panoramic radiographs.

Methodology Panoramic radiographs of 1593 individuals, representing patients with recent experience of a myocardial infarction ($n = 797$) and matched controls ($n = 796$) were obtained from a database. Variables to be assessed were; number of remaining teeth, number of root-filled teeth, number of teeth with periapical bone lesions and DMFT-score. A sample of 100 panoramic radiographs was randomly selected for this reliability study, being the first part of a case-control study. Two observers, one endodontist and one radiologist, did two separate assessments of the sample. The same examiners later assessed any disagreement cases jointly and the results were considered as the gold standard. Three other observers, one general dentist and two under-graduate students, did three separate assessments of the sample. The first assessment was done without any prior training, the second assessment was preceded by a calibration towards the gold standard and the third assessment was performed with the sample randomly hidden in the complete material of 1593 panoramic radiographs. Agreement was calculated and presented as Weighted Cohen's Kappa (κ).

Results Only the variable periapical lesions showed any significant disagreement between the observers. The endodontist and the radiologist showed an inter-agreement of $\kappa = 0.53$ in their initial assessments. The first assessment by the general dentist and the under-graduate students revealed κ ranging from 0.22 to 0.60 when compared to the gold standard. Following calibration in the second assessment agreement increased for all observers to κ between 0.59 and 0.80. The third assessment revealed still acceptable agreement with the gold standard ($\kappa = 0.54$ to 0.75). However, all observers exhibited a tendency in the direction towards the first assessment.

Conclusions Calibration towards a predetermined gold standard leads to an increased reliability in the assessment of apical lesions in panoramic radiographs; however, the long-term effect of such an intervention is uncertain.