Stomatologija, Baltic Dental and Maxillofacial Journal, 10:36-43, 2008

Observer performance based on marginal bone tissue visibility in Scanora® panoramic radiography and posterior bitewing radiography

Deimante Ivanauskaite, Christina Lindh, Madeleine Rohlin

SUMMARY

Objectives. To evaluate image quality for marginal bone tissue assessment on panoramic radiographs taken with the Scanora dental programme and on posterior bitewing radiographs.

Methods. Panoramic and bitewing radiographs were taken of 96 patients. Six observers rated marginal bone level visibility as excellent, acceptable, or unacceptable. Five observers assessed image quality for detection of vertical bone defects and furcation involvements as acceptable or unacceptable. Observer agreement was calculated as overall agreement and kappa values.

Results. Image quality of 36% of the panoramic and 6% of the bitewing sites was rated unacceptable for marginal bone loss assessment in the maxillae while 8% of the panoramic and bitewing sites in the mandible were unacceptable. For detecting vertical bone defects, image quality was unacceptable at one-third of the maxillary sites and 5% of the mandibular sites on the panoramic radiographs. Detection of furcation involvement was acceptable at most sites on both types of radiographs. Kappa values for intra- and inter-observer agreement were higher for panoramic than for bitewing radiographs. The kappa value for marginal bone loss assessment by several observers was moderate (0.45) for panorama and fair (0.28) for bitewing radiography. Corresponding kappa values for detection of vertical bone defects were substantial (0.62) and fair (0.25).

Conclusion. Image quality as evaluated by visual grading analysis is adequate for marginal bone tissue assessment in mandibular molar and premolar regions and unacceptable in maxillary molar and premolar regions on panoramic radiographs taken with the Scanora technique compared to bitewing radiography.

Key words: alveolar bone loss; diagnostic imaging; observer variation; radiography, bitewing; radiography, panoramic.

INTRODUCTION

Diagnostic information on marginal bone tissue obtained in panoramic radiography has been compared with information in bitewing and periapical radiography in several studies. Concordance between panoramic and intraoral radiography has been found to range between 55% and 74% of the assessed sites [1, 2, 3, 4]. Results of studies on the diagnostic accuracy of panoramic radiography compared with a criterion standard that comprised probing during surgery varied. Whilst Åkesson et al (1992) found the accuracy of panoramic radiography to be comparable to that of intraoral radiography, Pepelassi et al (1997) found panoramic accuracy to be lower [5, 6]. One reason for these varying results may be differences in the panoramic equipment used and therefore in the image quality of the panoramic radiographs. Since these studies, the Scanora® multimodal system, with a smaller focal spot and narrower X-ray beam than other conventional panoramic equipment, has been increasingly introduced in the clinic. Molander et al (1995) found that the Scanora® dental programme with a magnification factor of 1.7 provided the best subjective image quality for dental diagnostics compared to other panoramic equipment, whereas Kaeppler et al (2000) found the Scanora® and Orthophos® jaw programmes to be comparable [7, 8].

Various methods for evaluating image quality have been used in film radiography [9]. Some meth-

[1]*Institute of Odontology, Faculty of Medicine, Vilnius University, Vilnius, Lithuania*
[2]*Faculty of Odontology, Malmö University, Malmö, Sweden*

Deimante Ivanauskaite[1] *– D.D.S., M.D.S., Assist. Prof.*
Christina Lindh[2] *– D.D.S., Odont. Dr., Assoc. Prof.in Oral and Maxillofacial Radiology*
Madeleine Rohlin[2] *– D.D.S., Odont. Dr., Professor in Oral and Maxillofacial Radiology, Hon. FRCS (Eng) Hon. Fellow Cardiff University*

Address correspondence to Deimante Ivanauskaite, Institute of Odontology, Medical Faculty, Vilnius University, Zalgirio 115, LT – 08217 Vilnius, Lithuania .
E-mail: deimante.ivanauskaite@gmail.com, deimantei@one.lt

ods focus on the physical characteristics of the imaging systems such as measurements of contrast, spatial resolution, and noise, whilst other methods include the human observer, an important link in the imaging chain. In visual grading analysis, one of these methods, the appearance of the whole image or parts of an image is evaluated visually. To our knowledge, no study has compared visibility levels of anatomic features on panoramic radiographs made with the Scanora® dental programme with visibility levels on intraoral radiographs. The risk is that observers would also vary in their analysis of image quality, as occurs in other judgemental tasks such as in the assessment of marginal bone level in panoramic radiography [10, 11, 12, 13].

The aim of this study was to evaluate the visibility of anatomic features for the assessment of marginal bone tissue on panoramic radiographs made with the Scanora® dental programme and on posterior bitewing radiographs. Furthermore, we hypothesised that image quality will not influence concordance between Scanora® and bitewing radiographs in the assessment of marginal bone level.

MATERIALS AND METHODS

Patients

Ninety-six consecutive patients referred to the Department of Oral and Maxillofacial Radiology, Malmö University, Malmö, Sweden, for radiographic examination of teeth and surrounding bone tissue were examined. Forty-four patients were male (mean age 49 years, range 21–78 years) and 52 were female (mean age 48 years, range 20–85 years). Ivanauskaite et al (2006) describes the patients' age and sex distribution in more detail [4]. The Ethics Committee of Lund University, Lund, Sweden, approved the study.

Radiographic techniques

Panoramic radiography was performed with the Scanora® (Soredex, Helsinki, Finland) multimodal radiography system using the screen/film combination Lanex medium/T-mat G (Eastman Kodak Co., Rochester, N.Y., USA). Panoramic radiographs were made with the Scanora® dental programme 003 (magnification factor 1.7) and voltage settings 4/2 (66 kV, 10 mA, 15 s), 4/3 (66 kV, 13mA, 15 s), 4/4 (66 kV, 15 mA, 15 s), 4/5 (66 kV, 20 mA, 15 s), 4/6 (66 kV, 20mA, 19 s), 4/7 (66 kV, 20 mA, 23 s), and 5/5 (70 kV, 16 mA, 16 s). Vertical angulation of the tube was a constant -5°. Films were processed in an automatic processor (Curix HT-33OU, AGFA, Belgium) with a developing time of 2 min at 32°C. In the following text we will refer to this technique as Scanora®.

Posterior bitewing radiographs were made with a Heliodent 70 (Siemens, Erlangen, Germany) and settings of 70 kV, 7 mA. Bitewing radiographs were taken using Kwik-bite (Hawe-Neos Dental, Gentilino, Switzerland) film-holders for horizontal posterior bitewing radiographs and Take-All (Wijkström, Menton, France) film-holders or paper tabs for vertical posterior bitewing radiographs. The focus–skin distance was 20 cm and a rectangular collimator (30 x 40 mm^2) was used. The vertical angulation of the tube was kept constant at +10°. Ektaspeed Plus film (Eastman Kodak Co., Rochester, N.Y., USA) was used, and the exposure time was 0.32–0.64 s. Films were processed in an X-ray film automatic processor (XR 24 Nova, Dürr Dental, Bietigheim, Germany) with a developing time of 6 min at 28°C (±0.5°C).

Observers and observations

Six observers categorised the visibility of all approximal sites, which had been selected for marginal bone level assessment. When the image of a site was considered readable, marginal bone loss was measured using a ruler as described by Ivanauskaite et al (2006) [4]. The experience of the six observers in oral radiology varied between 1 and 30 years (mean 13 years).

Five of the observers assessed all sites from the canine to the distal site of the second molar for vertical bone defects and all molars and first upper premolars for furcation involvement. When the imaged site or tooth was considered readable, the presence or absence of vertical bone defects and furcation involvements was determined. The experience of these five observers in oral radiology varied between 2 and 30 years (mean 15 years).

Each observer assessed the bitewing and Scanora® radiographs independently at a 1-week interval. Before assessment, the observers jointly discussed the categories and specified assessment criteria. Categories and assessment criteria were then written down for reference during assessment. Three observers made a second assessment after 4 weeks to enable calculation of intra-observer agreement.

Teeth/Sites for assessment

To determine *marginal bone level visibility*, the radiographs of the 96 patients were randomly divided into six groups (16 patients/group). Six sites per patient in rotation by teeth were assessed. The selection of sites from the distal site of the canine to the distal site of the second molar of each group of patients was previously presented in detail by

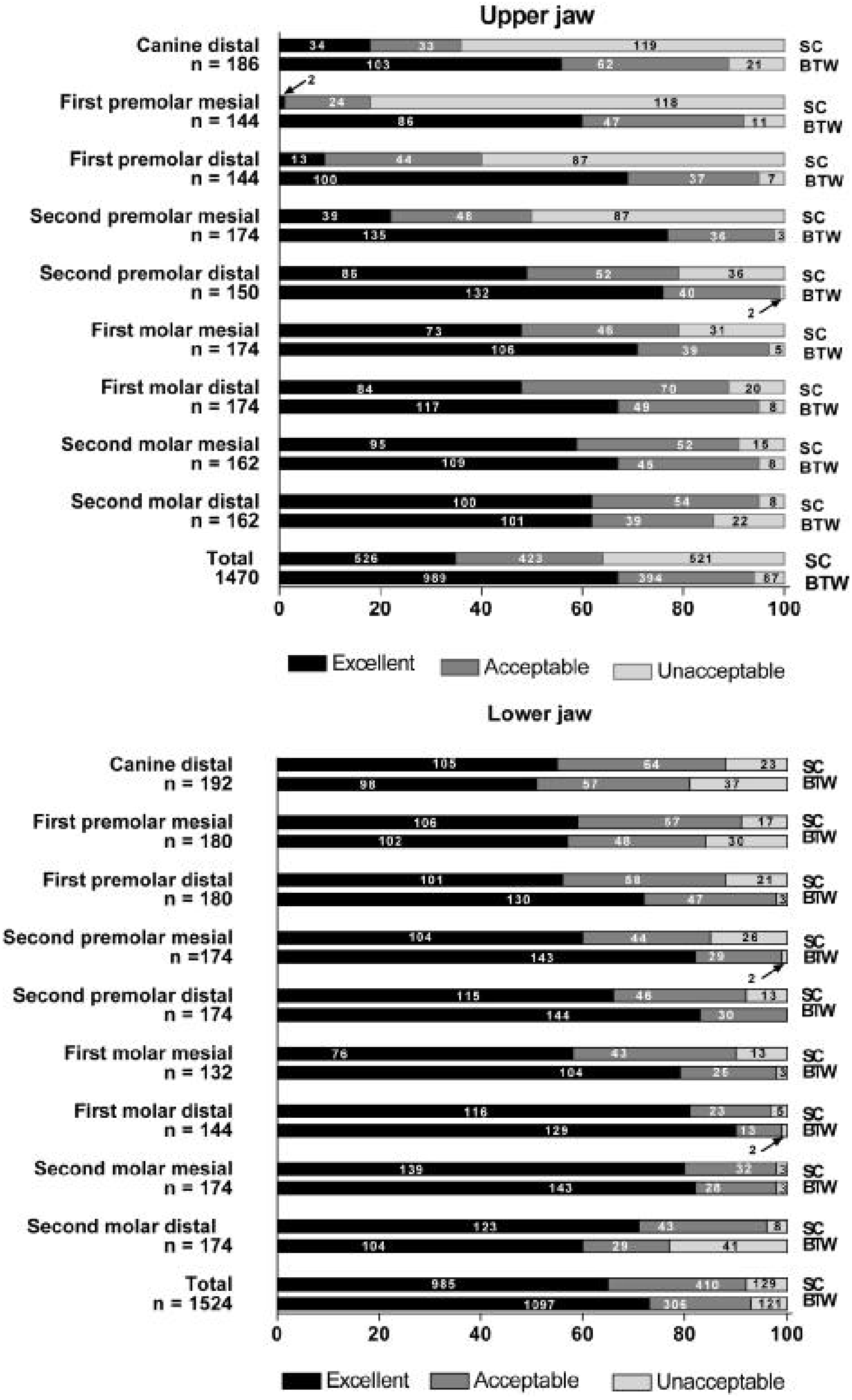

Fig. Percent distribution of visual grading analysis scores (excellent, acceptable, unacceptable) of marginal bone level visibility on panoramic radiographs taken with the Scanora dental programme (SC) and on posterior bitewing radiographs (BTW) by tooth site and jaw for six observers. Numbers of sites per rating are indicated within the bars.

Ivanauskaite et al (2006) [4]. Because not all patients had a full dentition, only 499 (245 maxillary and 254 mandibular sites) of 576 sites (96 patients x 6 sites) were available for assessment (Figure). Ivanauskaite et al (2006) described the distribution of marginal level ratings of the sites [4].

Visibility of vertical bone defects was assessed for all approximal sites from the distal site of the canine to the distal site of the second molar (1435 maxillary and 1450 mandibular sites) in all 96 patients. Table 1 presents the distribution of sites. *Visibility of furcation involvement* was assessed in 580 imaged molars (303 maxillary and 281 mandibular molars) and 164 maxillary first premolars.

Image quality and visual grading analysis

Overall radiographic image quality was assessed by one of the authors (DI). When image quality was poor due to faulty film placement, projection geometry, centring, density, contrast, or sharpness, radiographs were retaken before the radiographic examination of each patient was considered complete. For ethical reasons, only one retake was made to minimise the radiation dose to the patient.

Three categories originally proposed by the California Dental Association (1977) and modified by Åkesson et al. (1992) were further modified for this study for use in grading marginal bone level visibility [14, 11]:

- Excellent – provides necessary information for the assessment of marginal bone level (good density, contrast, sharpness, resolution; right projection; no image distortion and overlapping).
- Acceptable – provides information for the assessment of marginal bone level with some defect, which deviates from the ideal, but still acceptable.

- Unacceptable – does not provide the necessary information for the assessment of marginal bone level .

Visibility to allow detection of vertical bone defects and of furcation involvement was assessed as:

- Acceptable – provides information sufficient to assess the presence or absence of a vertical bone defect or furcation involvement.
- Unacceptable – does not provide information sufficient to assess the presence or absence of a vertical bone defect or furcation involvement.

Analysis

Intra-observer agreement was calculated for three observers and expressed as percent overall agreement and as Cohen's kappa [15]. Inter-observer agreement for several observers was calculated and expressed as Fleiss' kappa [16]. Additionally, inter-observer agreement for pairs of observers was calculated as percent overall agreement, weighted kappa values (Cohen's), or Cohen's kappa [15]. The six-point scale proposed by Landis and Koch (1977) was used to interpret kappa values [17]. Values less than zero were termed poor agreement, 0.00–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.60–0.80 substantial, and higher than 0.81 almost perfect agreement.

To evaluate the influence of image quality on the concordance between marginal bone loss ratings, the number of site pairs (Scanora® and bitewing images of the same site in the same patient) where the site was rated excellent and had the same marginal bone loss score on both radiographs was first calculated. The number of site pairs was then determined where the site on one of the radiographs (Scanora® or bitewing) was assessed to be acceptable. Concordance was expressed as overall agreement and kappa values.

RESULTS

Visibility grading for assessment of the marginal bone level

Figure presents the distribution of marginal bone level scores in the visual grading analysis. On Scanora® radiographs, 48%–62% of the sites distal to the maxillary second premolar were rated excellent. But the frequency of maxillary sites rated excellent was low in the canine (18%) and first premolar (1%–9%) regions. In the mandible, 55%–81% of the Scanora® sites were rated excellent.

In the maxillary region, only 6% of the bitewing sites were rated unacceptable compared to 36% of the Scanora® sites (Figure). In the mandible, the frequency of unacceptable sites on Scanora® (8%) and bitewing (7%) radiographs was similar. Most unacceptable sites occurred at the maxillary and mandibular canines (bitewing and Scanora®), in the maxillary premolar region (Scanora®), and at mesial mandibular premolar and distal mandibular and maxillary second molar sites (bitewing).

Visual grading analysis of vertical bone defects

Table 1 lists the frequencies of visual grading scores for vertical bone defect detection. In Scanora® radiographs, one-third of the maxillary sites were unacceptable, with the highest frequency (25%–66%) of unacceptability at canine and premolar sites. Overall, few bitewing sites were unacceptable. The highest frequency (4%–6%) of unacceptable bitewing sites occurred at the distal sites of mandibular and maxillary second molars, the distal sites of the mandibular canines, and the mesial sites of the mandibular first premolars.

Visual grading analysis of furcation involvement

Of 3740 assessments of furcation involvement,

Table 1. Visual grading analysis scores (acceptable or unacceptable) of image quality for vertical bone defect detection at sites on Scanora® and posterior bitewing radiographs (n – total of all assessments made by five observers, SC – Scanora® [panoramic radiograph taken using the Scanora® dental programme], BTW – bitewing, SC+BTW – site pair [Scanora® and bitewing images of the same site in the same patient]).

	Maxillae				Tooth site	Mandible				
n	Acceptable	Unacceptable				Acceptable	Unacceptable			n
	SC+BTW (%)	SC (%)	BTW (%)	SC+BTW (%)		SC+BTW (%)	SC (%)	BTW (%)	SC+BTW (%)	
935	38	58	1	3	Caninedistal	88	5	6	1	950
820	32	66	0.4	2	First premolarmesial	89	5	5	1	900
820	52	47	0.4	1	First premolardistal	92	7	0.4	0.3	900
785	50	48	1	1	Second premolarmesial	92	7	0.2	0.4	845
785	74	25	1	0.1	Second premolardistal	93	7	0	0.1	845
765	75	25	0.4	0.1	First molarmesial	91	9	0.2	0.2	645
765	89	9	1	1	First molardistal	99	1	0.2	0	645
750	90	8	1	1	Second molarmesial	99	1	0	0	760
750	89	5	6	0.4	Second molardistal	96	0.3	4	0	760
7175	**65**	**32**	**1**	**1**	**Total**	**93**	**5**	**2**	**0.4**	**7250**

only 59 (2%) were rated unacceptable. Of these, 22 were on Scanora® and 37 on bitewing radiographs. One observer made 37 of the 59 unacceptable ratings.

Observer agreement

Table 2 presents intra-observer agreement for the visual grading analysis. Overall agreement in marginal bone level assessment was comparable for Scanora® (range 76%–87%) and bitewing (range 78%–91%) radiography. Kappa values were higher for Scanora® than for bitewing radiography.

Overall agreement in vertical bone defect detection was very high for Scanora® (93%–95%) *and* bitewing (98%) radiography. Kappa values for

Table 2. Intra-observer agreement for visual grading analysis in panoramic radiography (Scanora® dental programme) and posterior bitewing radiography for marginal bone level assessment (scores: excellent, acceptable, or unacceptable) and vertical bone defect detection (scores: acceptable or unacceptable).

Observer	Marginal bone level				Vertical bone defects			
	Scanora®		Bitewing		Scanora®		Bitewing	
	%	κ	%	κ	%	κ	%	κ
1	76	0.59	78	0.56				
2	87	0.75	89	0.49	93	0.79	98	0.34
3	82	0.63	91	0.42	93	0.81	98	0.62
4					95	0.88	98	0.32

κ = Cohen's kappa

Table 3. Inter-observer agreement of several observers and pairs of observers for visual grading analysis in panoramic radiography (Scanora® dental programme) and posterior bitewing radiography. Six observers rated visibility of marginal bone level (scores: excellent, acceptable, or unacceptable). Five observers rated visibility of vertical bone defects (scores: acceptable or unacceptable).

	Marginal bone level				Vertical bone defects			
	Scanora®		Bitewing		Scanora®		Bitewing	
	κ		κ		κ		κ	
Several observers	**0.45**		**0.28**		**0.62**		**0.25**	
Observer pair	**%**	**κw**	**%**	**κw**	**%**	**κ**	**%**	**κ**
1/2	89	0.61	89	0.41				
1/3	90	0.59	90	0.32				
1/4	85	0.53	85	0.43				
1/5	62	0.47	65	0.32				
1/6	43	0.33	55	0.32				
2/3	91	0.63	93	0.40	91	0.73	97	0.19
2/4	87	0.61	85	0.36	88	0.67	98	0.13
2/5	65	0.55	74	0.35	90	0.60	97	0.11
2/6	46	0.39	45	0.17	90	0.63	96	0.15
3/4	85	0.51	84	0.23	91	0.76	97	0.27
3/5	67	0.56	72	0.27	86	0.53	96	0.26
3/6	44	0.35	45	0.16	88	0.62	96	0.40
4/5	65	056	66	0.36	83	0.49	97	0.11
4/6	52	0.45	54	0.34	86	0.60	96	0.23
5/6	45	0.31	51	0.24	88	0.51	96	0.38

κ = Fleiss' kappa (several observers), Cohen's kappa (observer pair)
κw = Cohen's weighted kappa (observer pair)

Scanora® were higher than for bitewing radiography. Overall agreement in furcation involvement detection was very high for both methods (93%–98%).

Table 3 presents inter-observer agreement in visual grading analysis. Inter-observer agreement for six observers in marginal bone level assessment was moderate for Scanora® (κ=0.45) and fair for bitewing (κ=0.28) radiography; kappa indices for both techniques were lower for the category acceptable (κ=0.26 and 0.22, respectively) than for excellent (κ=0.52 and 0.31, respectively) and unacceptable (κ=0.64 and 0.46, respectively). Overall inter-observer agreement for pairs of observers was comparable for Scanora® and bitewing radiography, whilst weighted kappa indices for inter-observer agreement were higher for Scanora®. Weighted kappa values for pairs of observers vary substantially, 0.31–0.63 for Scanora® and 0.16–0.43 for bitewing radiography. The lowest weighted kappa values for pairs of observers often included observer 6, who used the category acceptable (Scanora® and bitewing images) and unacceptable (Scanora® images) more frequently than the other observers.

Inter-observer agreement of five observers for the detection of vertical bone defects was substantial for Scanora (κ=0.62) but only fair for bitewing (κ=0.25) radiography (Table 3). Overall agreement for pairs of observers was high for Scanora® (range 83%–91%) and very high for bitewing (range 96%–98%) radiography. Corresponding kappa values were moderate or substantial for Scanora® (range 0.51–0.76) but poor or slight for bitewing radiography (range 0.11–0.38). Inter-observer agreement of five observers for detection of furcation involvements was 97% (κ=0.02) for Scanora® and 96% (κ=0.06) for bitewing radiography. Agreement for pairs of observers was high for Scanora® (range 97%–98%) *and* bitewing (range 96%–100%) radiography.

Visibility grading and concordance between techniques in marginal bone level assessment

Image quality did not influence agreement between Scanora® and bitewing radiography for marginal bone level assessment. Agreement was 57% (κ=0.32) when both techniques in the site pair were rated excellent and when one of the techniques in the site pair was rated acceptable.

DISCUSSION

Methods and materials

Methods for evaluating image quality in diagnostic procedures can be divided into a few major groups based on measuring principle and type of result. In a hierarchical conceptual model that Fryback and Thornbury (1991) proposed as an organising structure for evaluating diagnostic methods, the lowest level was technical efficacy [18]. Technical efficacy of an imaging method includes measurements of basic properties such as contrast, spatial resolution, and noise that are directly or indirectly determined. But the ultimate goal of an imaging method is to establish a connection between the physical characteristics of the imaging system and the diagnostic outcome of the system for a given, clinically relevant task. This emphasises that an image should have good quality in relation to its diagnostic purpose [19].

Moreover, an asymmetry between physical characteristics and diagnostic outcome may exist in that higher technical efficacy does not guarantee an improvement in diagnostic outcome. There are numerous radiographic techniques in which the sacrifice of physical parameters of quality improves diagnostic accuracy [20]. For example, resolution is lower in panoramic radiography than in intraoral radiography. So a higher level of efficacy evaluation, such as the one proposed by Fryback and Thornbury (1991), is necessary [18]. At this next level, diagnostic efficacy, the image must be interpreted by an observer in an attempt to make a diagnosis. A pragmatic approach to this level is visual grading analysis, which is based on

Table 4. Examples of categories/criteria for visual grading analysis implemented in panoramic radiography and number of observers asked to grade visibility

Radiographic technique	Categories/criteria for visual grading analysis	Number of observers	Reference
Panoramic radiography	*Excellent* - provides necessary information *Acceptable* - with some defects but still acceptable for diagnostic purpose *Unacceptable* - does not provide necessary information	3	Åkesson et al. (1992) modified from Quality Evaluation for Dental Care (1977) [11, 14]
	Clarity of 12 landmarks: + 2 = excellent + 1 = good 0 = satisfactory -1 = poor -2 = acceptable	5	Wakoh et al. (1998) [21]
	Visibility of 21 anatomical features: 1 = excellent 2 = more than adequately represented 3 = adequately represented 4 = barely adequately represented 5 = inadequate for diagnosis	10	Dannewitz et al. (2002) [22]
	Visibility of 11 anatomical structures: 1 = structure well visible 0 = structure partly visible -1 = structure not or hardly visible	5	Kaeppler et al. (2006) [23]
Scanora®	Visualisation of diagnostically important structures: 4 = fine details visualised, diagnosis definitely possible 3 = small detailed visualised, diagnosis probably possible 2 = only broad details seen, diagnosis doubtful 1 = significant structures not visible, no diagnosis possible	1	Molander et al. (1995) [7]
	Visibility of 7 anatomical features: 1 = very good 2 = good 3 = satisfactory 4 = incomplete 5 = poor	3	Kaeppler et al. (2000) [8]
	Visibility for assessment of marginal bone level: excellent, acceptable, unacceptable	6	Present study for assessment of marginal bone level: modified from Åkesson et al. (1992) [11]
	Visibility for assessment of vertical bone defects and furcation involvements: acceptable, unacceptable	5	
Comparison film-based and digital images	Images graded on a scale where important structures were visualised: 5 - much better 4 - better 3 - equal 2 - worse 1 - much worse	10	Molander et al. (2004) [24]

the visibility of certain anatomical structures linked to a diagnostic task. Thus, this method takes into account technical factors *and* observer-dependent factors.

There is a spectrum of classification systems for visual grading analysis. Visual grading analysis can be performed as either relative or absolute grading. In relative grading, images from two diagnostic methods are compared simultaneously, whereas in absolute grading, the two methods are evaluated separately. Because we were interested in comparing Scanora® and bitewing radiography, not only for visual grading analysis but also for the assessment of marginal bone tissue, we used absolute grading in this study. A special case of visual grading analysis is the use of image criteria with various levels of visibility of defined structures [9]. Table 4 presents examples of classification systems for visual grading analysis with different descriptors that have been implemented in oral and maxillofacial radiography. Three to five grades were usually chosen, and 3–21 anatomic features to assess visibility. But only in Åkesson et al's study (1989, 1992) and the present study was visual grading analysis directly linked to a diagnostic task, that is, to score marginal bone loss or to identify vertical bone defects and furcation involvements [11,12].

The patients taking part in this study had been referred to the clinic for radiographic examination of teeth and surrounding bone tissue, and the radiographs were part of a normal radiographic examination. Most patients were between 40 and 59 years. Yet, only 15% of the sites presented radiographically a vertical bone defect and around 10% of the teeth a furcation involvement [4]. These frequencies might have influenced the analysis and the kappa values for observer performance, particularly in the identification of vertical bone defects and furcation involvements.

Results

The frequency of mandibular sites rated unacceptable on Scanora® radiographs in this study was lower (8%) than what Åkesson et al (1992) reported for radiographs taken with Orthopantomograph Model OP5 in the same department (14%–15%) [12]. However, we found more maxillary sites to be unacceptable (36%) than did Åkesson et al (1992) (18%–24%) [12]. Although similar visual grading analysis methods were applied, the number of observers differed: we used six observers and Åkesson et al (1992) used three [12]. The number of observers may influence results in the number of unacceptable sites so that the number may increase with more observers. The results of Molander et al. (1995) underpin the assumption that the image quality of radiographs taken with the Scanora® dental programme (score 3.05) was higher than of radiographs taken with Orthopantomograph Model OP5 (score 2.69) [7]. Kaeppler et al (2000) compared radiographs taken with the Scanora® jaw programme and with Orthophos Plus for visualisation of anatomical features and found no significant difference in ratings [8].

In only three sites was image quality for marginal bone level assessment superior on Scanora® radiographs. These sites – all in the mandible – were distal to the canine, mesial to the first premolar, and distal to the second molar. The probable explanation for this is that these sites were not imaged on the bitewing radiographs, despite the overall assessment of image quality (which allowed one retake in cases of poor quality) that was made before assessments at site level were begun.

Visibility for detection of vertical bone defects and furcation involvements in panoramic radiography has, to our knowledge, not been analysed previously. Our results demonstrated that Scanora® is suitable for the detection of vertical bone defects at all sites investigated in this study, except for maxillary sites in the canine and premolar regions. For detection of furcation involvements, visibility was similar on Scanora® and bitewing radiographs. These results will be important when Scanora® is used to study prevalence of vertical bone defects and furcation involvements.

Observer agreement

Determining kappa values for degree of observer agreement takes into account agreement that can be expected to occur by chance alone [15]. But two observers, or two observations, may emerge with low kappa values despite relatively high overall agreement. So we chose to report overall agreement and kappa values. When two observers express binary ratings, as for visibility for detection of vertical bone defects and furcation involvements in this study, the results are arranged in a 2 x 2 table. If the horizontal and vertical marginal totals are symmetrically unbalanced, high overall agreement will be associated with low levels of kappa [25]. This often occurred in the detection of vertical bone defects and furcation involvements on bitewing radiographs, where the frequency of acceptable sites was very high, and somewhat explains the low corresponding kappa values (Table 2).

Intra-observer agreement expressed was high. The lowest kappa value for Scanora®, κ=0.59, was comparable to intra-observer ratings of κ=0.55 reported by Kaeppler et al (2006) for the detection of anatomical structures in panoramic radiography [23]. Radiographs from two panoramic systems were compared. Three observers graded image quality in five categories according to 14 anatomical features. Calculations of inter-observer agreement expressed as weighted kappas were similar to what we found in this study. Most weighted kappa values fell into the category poor according to Landis and Koch (1977) [17]. This strengthens the fact that variations in the interpretation of radiographs from different machines or made with different imaging techniques depend more on observer variation than differences in visibility. The results of this study, with one observer consistently reporting fewer sites that fulfilled the criteria, demonstrate the importance of including several observers in any analysis of diagnostic methods.

CONCLUSION

Scanora® panoramic radiography simplifies examination of marginal bone tissue compared to intraoral radiographs. Image quality of Scanora® radiographs, according to visual grading analysis, is adequate for assessing marginal bone level and vertical bone defects in mandibular, but not maxillary, molar and premolar regions compared to bitewing radiography. Visibility of furcation involvements on Scanora® radiographs was rated as high as on bitewing radiographs, which makes Scanora® panoramic radiography suitable for patients with severe marginal bone loss. The level of visibility had no influence on concordance between the two techniques. This result indicates that when a site on a Scanora radiograph is readable, the assessment of marginal bone level is reliable.

ACKNOWLEDGEMENTS

This investigation was supported by grants from the Swedish Institute for Dr. Deimante Ivanauskaite's studies in the Department of Oral Radiology, Malmö University, Malmö, Sweden. We wish to thank Odont. Dr. Bertil Kinnby for assistance with the graph.

REFERENCES

1. Rohlin M, Åkesson L, Håkansson J, Håkansson H, Näsström K. Comparison between panoramic and periapical radiography in the diagnosis of periodontal bone loss. *Dentomaxillofac Radiol* 1989;18:72-6.
2. Åkesson L, Rohlin M, Håkansson J, Håkansson H, Näsström K. Comparison between panoramic and posterior bitewing radiography in the diagnosis of periodontal bone loss. *J Dent* 1989;17:266-71.
3. Molander B, Ahlqwist M, Gröndahl H-G, Hollender L. Agreement between panoramic and intra-oral radiography in the assessment of marginal bone height. *Dentomaxillofac Radiol* 1991;20:155-60.
4. Ivanauskaite D, Lindh C, Rangne K, Rohlin M. Comparison between Scanora^{N} panoramic radiography and bitewing radiography in the assessment of marginal bone tissue. *Stomatologija. Baltic Dent Maxillofac J* 2006;8:9-15.
5. Åkesson L, Håkansson J, Rohlin M. Comparison of panoramic and intraoral radiography and pocket probing for the measurement of the marginal bone level. *J Clin Periodontol* 1992;19:326-32.
6. Pepelassi EA, Diamanti-Kipioti A. Selection of the most accurate method of conventional radiography for the assessment of periodontal osseous destruction. *J Clin Periodontol* 1997;24:557-67.
7. Molander B, Ahlqwist M, Gröndahl HG. Image quality in panoramic radiography. *Dentomaxillofac Radiol* 1995;24:17-22.
8. Kaeppler G, Axmann-Krcmar D, Reuter I, Meyle J, Gómez-Romịn G. A clinical evaluation of some factors affecting image quality in panoramic radiography. *Dentomaxillofac Radiol* 2000;29:81-4.
9. Tingberg A. Quantifying the quality of medical x-ray images. An evaluation based on normal anatomy for lumbar spine and chest radiography [Thesis]. Lund University; 2000.
10. Kaimenyi JT, Ashley F. Assessment of bone loss in periodontitis from panoramic radiographs. *J Clin Periodontol* 1988;15:170-4.
11. Åkesson L, Rohlin M, Håkansson J. Marginal bone in periodontal disease: an evaluation of image quality in panoramic and intra-oral radiography. *Dentomaxillofac Radiol* 1989;18:105-12.
12. Åkesson L, Håkansson J, Rohlin M, Zöger B. An evaluation of image quality for the assessment of the marginal bone level in panoramic radiography. *Swed Dent J* 1992;17:9-21.
13. Walsh TF, Al-Hokail OS, Fosam EB. The relationship of bone loss observed on panoramic radiographs with clinical periodontal screening. *J Clin Periodontol* 1997;24:153-7.
14. Quality evaluation of dental care. Guidelines for the assessment of clinical quality and professional performance. Radiographs. Los Angeles: California Dental Association; 1977. p. 13-5.
15. Altman DG. Some common problems in medical research. In: Altman DG. Practical statistics for medical research. 1st ed. London: Chapman & Hall; 1991. p. 369-439.
16. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378-82.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
18. Fryback D, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11:88-94.
19. International Commission on Radiation Units and Measurements [ICRU]. Medical imaging – the assessment of image quality. ICRU Report 54. ICRU publications. Bethesda; 1996.
20. Blesser B, Ozonoff D. A model for the radiologic process. *Radiology* 1972;103:515-21.
21. Wakoh M, Farman AG, Kitagawa H, Nishikawa K, Kuroyanagi K. A sensitometric comparison of Kodak Ektavision and Fuji Super HR-S panoramic radiographic films. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1998;86:249-53.
22. Dannewitz B, Hassfeld S, Eickholz P, Mühling J. Effect of dose reduction in digital panoramic radiography on image quality. *Dentomaxillofac Radiol* 2002;31:50-5.
23. Kaeppler G, Dietz K, Reinert S. The effect of dose reduction on the detection of anatomical structures on panoramic radiographs. *Dentomaxillofac Radiol* 2006;35:271-7.
24. Molander B, Gröndahl H-G, Ekestubbe A. Quality of film-based and digital panoramic radiography. *Dentomaxillofac Radiol* 2004;33:32-6.
25. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990;43:543-549.

Received: 04 02 2008
Accepted for publishing: 28 03 2008