VILNIUS UNIVERSITY

Deimantė Ivanauskaitė

ALVEOLAR BONE LOSS IN RADIOGRAPHIC MODALITIES FOR DIAGNOSIS OF PERIODONTAL DISEASE

Doctoral dissertation
Biomedical sciences, Medicine (07 B)

Vilnius, 2011

The research project was carried out at Malmö University, Sweden with the support of grant ref nr 2419/1998 (380/123) from the Swedish Institute and Vilnius University, Lithuania during the period 1999-2011.

Scientific consultants:

prof. dr. Madeleine Rohlin (Dept. of Oral and Maxillofacial Radiology, Malmö University, Sweden)

prof. dr. Vytautė Pečiulienė (Vilnius University, Biomedical Sciences, Odontology – 08B).

CONTENT

ABBREVIATIONS

AB – alveolar bone

AC – alveolar crest

BL – bone level

CBCT – Cone Beam Computed Tomography

CEJ – cemento-enamel junction

CT – Computed Tomography

ICC – intra-class correlation coefficients

ICRP – International Commission on Radiological Protection

ICRU – International Commission on Radiation Units

MeSH – Medical Subject Heading used for indexing articles for PubMed

MRI – Magnetic Resonance Imaging

MSCT –Multislice Computed Tomography

NRPB – National Radiological Protection Board

PubMed – the U.S. National Institutes of Health (NIH) free digital archive of biomedical and life sciences journal literature.

QUADAS – Quality Assessment of Diagnostic Accuracy Studies

ROC – receiver operating characteristics

SBU – Statens beredning för medicinsk utvärdering (The Swedish Council on Health Technology Assessment)

SD – standard deviation

STARD – Standards for Reporting of Diagnostic Accuracy

TACT – Tuned-Aperture Computed Tomography

TMJ – Temporomandibular joint

1. INTRODUCTION

1.1 Chronic periodontitis with alveolar bone loss

1.1.1 Definitions

The most common diseases of the periodontal tissues are gingivitis and chronic periodontitis. According to a consensus report (1999), which was based on a World Workshop in Periodontitis, the universal features of *gingivitis* include “clinical signs of inflammation, signs and symptoms that are confined to the gingiva, reversibility of the diseases by removing the etiology, the presence of bacterial laden plaque to initiate and/or exacerbate the severity of the disease and a possible role as a precursor to attachment loss around teeth. Clinical signs of gingivitis must be associated with stable (*i.e.* nonchanging) attachment levels on a periodontium with no loss of attachment or alveolar bone or on a stable but reduced periodontium” (Mariotti 1999). Also, from the consensus report it was concluded that “if clinician is going to make a diagnosis of gingivitis on reduced but stable periodontium, it is necessary to longitudinally demonstrate that attachment loss is not occurring” (Caton & Greenwell 1999). Earlier, gingivitis and periodontitis were looked upon as expressions of the same disease with gingivitis always being followed by periodontitis. Presently, gingivitis without loss of supporting tissues is considered to be a condition separate from chronic periodontitis.

Chronic periodontitis is an inflammation of the periodontium with slow to moderate, progressive loss of the tooth supporting tissues. The classification of periodontitis has been revised at several consensus conferences. In the consensus report (1999), chronic periodontitis was defined as “An infectious disease resulting in inflammation within the supporting tissues of the teeth, progressive attachment, and bone loss. It is characterized by pocket formation and/or gingival recession” (Lindhe et al. 1999). At the workshop leading to the consensus report the term “Adult periodontitis” was replaced with “Chronic periodontitis” as epidemiological data and clinical experience suggest that the form of periodontitis commonly found in adults can also be seen in adolescents (Armitage 1999). Therefore it would be more accurate to adopt the term “Chronic Periodontitis” (Armitage 1999).

There is still a lack of uniformity in the definition of chronic periodontitis. Baelum and Lopez (2003) stated that “the past two decades have witnessed a large number of proposals for the classification of periodontitis” and that “periodontitis is a syndrome, the clinical manifestations of which may come in all sizes. Thereby, periodontitis has no diagnostic truth, just as there is no natural basis for a sharp distinction between health and disease or between different forms of periodontitis. Recognition of these facts and adoption of a nominalistic approach to the definition of periodontitis is needed to provide a rational framework for the development of a classification system that meets the needs of both clinicians and scientists”. According to Baelum and Lopez (2003), the periodontitis classification system proposed by van der Velden (2000) shows the way, but the number of diagnostic categories needed to be determined on the basis of documented differences regarding the consequences of the diagnoses. The definition and classification proposed by van der Velden (2000) is established on a combination of the extent of disease, the severity of disease, the diagnosis on the basis of clinical characteristics if applicable, and the diagnosis on the basis of the patient’s age. The severity of disease is based on the degree of bone loss or attachment loss as *(i)* minor: at all affected teeth, bone loss is ≤1/3 of the root length or attachment loss ≤3 mm, *(ii)* moderate: at ≥2 teeth, bone loss >1/3 and ≤1/2 of the root length or attachment loss 4–5 mm, and *(ii)* severe: at ≥2 teeth, bone loss is >1/2 of the root length or attachment loss ≥6 mm.

The debate on definitions and classification of periodontitis continues. Recently, a new classification of periodontal inflammation based on a "biologic systems model" was proposed (Offenbach et al. 2008). This model includes five categories that represent the transition from health to increasing severity of disease. Evaluations of periodontal disease burden are highly complicated by the ongoing discussion of a globally accepted case definition for periodontitis (Albandar 2007; Page & Eke 2007; Savage et al. 2009). From the systematic review of definitions of periodontitis and methods that have been used to identify periodontitis, Savage et al. (2009) concluded that studies of periodontal diseases are complicated by the diversity of definitions and measurements used to describe and quantify these diseases. To summarize, there is a lack of consensus as to a uniform definition and classification of chronic periodontitis. However, there is a common understanding that chronic periodontitis always includes alveolar bone loss to some degree.

1.1.2 Prevalence

Chronic periodontitis occurs with different prevalence in different countries and in different patient groups within the same country. A comprehensive summary of the prevalence of periodontal health in Europe based on data gathered before 2000 was presented by Sheiham and Netuveli (2002). According to Holtfreter et al. (2010) "only a small fraction of all European countries provided a comprehensive view on prevalence and extent of periodontal diseases" and in their study they mentioned publications from the United Kingdom (Kelly et al. 2000), Lithuania (Skudutytė et al. 2001), Switzerland (Menghini et al. 2002), Denmark (Krustrup & Petersen 2006), Norway (Skudutyte-Rysstad et al. 2007), Sweden (Hugoson et al. 2008), Finland (Suominen-Taipale et al. 2008), and Hungary (Hermann et al. 2009). Although a comparison of prevalence reported in published studies is complicated due to different definitions for periodontitis, methodological and recording disparities (Papapanou 1999; Albandar & Rams 2002; Kingman & Albandar 2002; Papapanou & Lindhe 2008), it seems reasonable to estimate and compare disease prevalence conservatively. According to the publication by Holtfreter et al. (2010), periodontal diseases were least prevalent in Sweden (Hugoson et al. 2008) and Switzerland (Menghini et al. 2002) and highest in Lithuania with a reported prevalence of 82 percent in 35–44-year-old subjects and 95 percent in 65–74-year-old subjects (Skudutytė et al. 2001). In 1999 – 2000 the prevalence of gingivitis and periodontitis among Lithuanian rural population aged 25-64 years ranged between 12 and 67 percent (mean 45.9%) for men and between 8 and 60 percent (mean 31.4%) for women (Globienė 2001). In a cohort of middle-aged Lithuanians, where mean age of the patients was 36.3 years, and they did not receive any periodontal therapy, the overall alveolar bone level was about 40 percent of the root length, with a wide range of 21 to 54 percent of the root length (Pūrienė 1997; Pūrienė et al. 2003) and furcation involvements in 6 percent of the teeth (Pūrienė et al. 2003).

The severity of changes of the periodontal tissues is said to increase with age. Results from epidemiological studies support this finding. In Lithuania, 12 percent of men and 8 percent of women in the age-group 25-34 years had gingivitis and periodontitis, while 43 to 67 percent of men and 21 to 57 percent of women in the age-groups older than 35 years had gingivitis and periodontitis (Globienė 2001). In the older 55-64 years age-group less than 5 percent of the persons were healthy (Globienė 2001). In Sweden, alveolar bone loss exceeding one third of the root length was seen in only 5 percent of 40-year-old individuals compared to 21 percent of 50-year-old and 33 percent of 70-year-old individuals (Norderyd & Huguson 1998). Alveolar bone loss exceeding two thirds of the root length was seen in only 2 percent of 40-year-old individuals but in 7 percent of individuals, who were 50 years and older (Norderyd & Huguson 1998).

Although the alveolar bone loss is increasing, the average progression is generally slow. From the age of 20 years, bone height is reduced over time corresponding to an annual loss of around 0.1 mm (Hugoson 2000). The progression of chronic periodontitis could, however, be more severe in some patient groups and depends on factors such as local factors in the patient's mouth, the general health condition, and life habits. Smoking procedures have an adverse effect on alveolar bone height and density even at an early age with low tobacco consumption (Rosa et al. 2008). The association between cigarette smoking and periodontitis was shown to be strong and there was a relationship with dose and duration of smoking (Moimaz et al. 2009).

Recent discussions have centred on whether chronic periodontitis may promote or aggravate the development of general diseases. In particular, the question has been raised as to whether individuals with chronic periodontitis are at increased risk of developing cardiovascular disease, diabetes mellitus, asthma or rheumatoid arthritis. Scientific evidence is contradictory as to whether individuals with chronic periodontitis are at increased risk of developing coronary heart disease. "Scientific evidence is lacking as to whether individuals with chronic periodontitis are at increased of developing diabetes mellitus, chronic obstructive pulmonary disease or rheumatoid arthritis" (SBU-report 169 2004).

1.1.3 Treatment

The goal of preventing and treating chronic periodontitis is to reduce dental biofilm accumulation and its ability to induce destruction of the alveolar bone. Besides the prevention based on self-performed oral hygiene measures, the common approach involves the use of various instruments to remove dental biofilm, so called mechanical infection control (scaling and root planning). From a systematic review on treatment outcomes of chronic periodontitis, it was concluded that mechanical infection control reduces probing pocket depth and improves probing attachment level (SBU-report 169, 2004). Furthermore, when combined with flap surgery deeper pockets are eliminated. Treating chronic periodontitis can prevent further alveolar bone loss but can generally not reverse what has already occurred. During the last 20 years, certain regenerative techniques have been used to restore lost periodontal tissues. Adjunctive therapy with enamel matrix derivative in angular bone defects results in improved alveolar bone level (SBU-report 169 2004).

In the 1980s, it became apparent that not everybody is equally susceptible to periodontal diseases and that severe periodontitis concentrates in a relatively small part of the population (Löe et al. 1986). The CPITN index (Community Periodontal Index of Treatment Needs (Ainamo et al. 1982) was constructed for measuring treatment needs. When applying the CPITN index for a population in Vilnius and Vilnius-region, in Lithuania, 6 to 47 percent of people in different age groups from 15 years old need professional oral hygiene while 16 percent of people from 25 to 34 year old and 25 to 62 percent of people from 34 years old need complex periodontal treatment (with periodontal surgery) (Mackevičienė et al. 1999). For randomly selected 65-72 years old inhabitants of Kaunas city, in Lithuania, 93.3 percent had periodontal pathology and required treatment, where 58.5 percent need professional dental hygiene procedures, one fifth of the subjects require not only professional oral hygiene procedures, but also complex periodontal treatment (Zūbienė et al. 2008).

1.2 Diagnostic methods

1.2.1 In general

Clinicians have to their disposal a variety of clinical information to serve them in their clinical decisions making. Some pieces of information derive from diagnostic examinations and simply talking and observing the patients can secure other information. The information gained from diagnostic examinations has different purposes. The examination can be used for evaluation of the oral health status, for judgements leading to a diagnosis or a differential diagnosis, for treatment planning, for the evaluation of treatment outcomes or for epidemiological studies. Depending on the purpose of the examination, the diagnostic method chosen and the design of the examination may vary. One and the same method can be used not only for the purpose of diagnosis (Greek terms "*dia*" = through, "*gnosis*" = knowledge) but also to make a prognosis ("*pro*" = before, "*gnosis*" = knowledge) *i.e.* to predict disease or disease development. A diagnostic/prognostic method, whether clinical, radiological or microbiological, is employed to raise or lower the assessed probability that a patient has a disease/no disease or has improved from the treatment. As no method is perfect it is important to know the inherent errors of an employed method. One way of doing this is to analyse the diagnostic efficacy of the method and take this into account when performing and interpreting the results of an examination.

1.2.2 How can we evaluate diagnostic methods?

The kinds of difficulties faced by clinicians when utilizing a diagnostic method are not homogenous. Blesser and Ozonoff (1972) presented a conceptual framework for the radiological method, which still is relevant and can be modelled for most diagnostic methods and for the evaluation of a diagnostic procedure. The model presented three distinct phases of the procedure: *(a)* the psychophysical phase, which includes the technical part of the methods and the physical part of the human nervous system, *(b)* the psychological phase representing the creation of meaningful patterns, which are influenced by the context and interpretative rules, and *(c)* the ontological phase consisting of the judgements, where we use learned criteria to classify the percept as clinically normal or abnormal. Disproportionate effort in research has gone into perfecting the technical parts of different diagnostic methods (Blesser & Ozonoff 1972). Although higher technical quality can contribute to higher diagnostic efficacy, there may be a point beyond which improvement no longer gain the patient and the treatment of the patient. In radiology, an image with good quality is an image that fulfils its diagnostic purpose (ICRU Report 54 1996, Tingberg 2000).

Fryback and Thornbury (1991) proposed a hierarchical conceptual model to evaluate imaging methods, the lowest level being the technical efficacy. Technical efficacy of an imaging method includes measurements of basic properties such as contrast, spatial resolution, and noise that are directly or indirectly determined. These properties are often studied using physical test phantoms. The second level is the diagnostic accuracy level, which includes the human observer and can be performed of images of either anthropomorphic phantoms or of images of patients. In visual grading analysis, which includes the observer, the appearance of the image and the visibility of anatomical structures are evaluated. Other methods, such as the receiver operating characteristics (ROC) analysis, comprise the diagnostic accuracy based on the detection of lesions. The ability of human observers to properly detect relevant image features and classify those accurately can then be described as all possible compromises between true positive and false positive decisions (Månsson 1994). These decisions are used to calculate sensitivity, specificity and predictive values, other expressions for the diagnostic accuracy of a diagnostic or prognostic method. The goal of a diagnostic method is to establish a connection between the physical characteristics of the method and the diagnostic outcome of the system for a given, clinically relevant task. To take this into account, the next levels of Fryback and Thornbury's model (1991) include the therapeutic impact and the patient outcomes of diagnostic methods.

1.2.2.1 Visual grading analysis

Various methods for evaluating image quality have been used in film radiography (Tingberg 2000). Some methods focus on the physical characteristics of the imaging systems such as measurements of contrast, spatial resolution, and noise, while 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 by observers. A special case of visual grading analysis is to compare the visibility of defined structures with the same structures in a reference image (Månsson 1994).

One of the earliest systems for visual grading analysis of radiographs in oral health care including human observers was the California Dental Association System. This system (Quality evaluation of dental care. Guidelines for the assessment of clinical quality and professional performance. Radiographs 1977), which was a quality system for all parts of oral health care, categorised the radiographs as follows:

- Excellent – provides necessary information
- Acceptable – with some defects which deviates from the ideal but still acceptable for diagnostic purpose
- Not acceptable – does not provide the necessary information.

Later, the Royal College of Radiologists and the National Radiological Protection Board (NRPB) proposed a system (1994), which is presented in Guidance Notes for Dental Practitioners on the Safe Use of X-Ray Equipment (2001) as follows:

● Excellent – no errors of patient preparation, exposure, positioning, processing or film handling

● Diagnostically acceptable – some errors of patient preparation, exposure, positioning, processing or film handling, but which do not detract from the diagnostic utility of the radiograph

● Unacceptable – errors patient preparation, exposure, positioning, processing or film handling, which render the radiograph diagnostically unacceptable.

Both systems, which are somewhat similar in number of grades and wording, have been applied in scientific studies of intraoral and panoramic radiography. A modification of the California system was used in the studies by Åkesson et al. (1989a); Åkesson 1991; Åkesson et al. (1992b), while the British system was used by Rushton et al. (1999) and Carmichael et al. (2000). As presented in Table 1.1, there is a spectrum of classification systems that have been applied in oral health care. There are examples of systems for relative grading, where images from two diagnostic methods are compared simultaneously and examples of systems for absolute grading, where methods are evaluated separately (Table 1.1).

Table 1.1 Visual grading analysis. Examples of systems with different categories and criteria implemented in oral health care and number of observers asked to assess the images. References are presented in order of publication year

Radiographic method	Categories/criteria for visual grading analysis	Observers (n)	Reference
Panoramic radiography	Each tooth site assessed as: *Excellent* – provides necessary information	3	Åkesson et al. (1989a; 1992b)

Radiographic method	**Categories/criteria for visual grading analysis**	**Observers (n)**	**Reference**
(film-based) and *Intraoral radiography (film-based)*	*Acceptable* – with some defects which deviates from ideal but still acceptable for purpose of marginal bone scoring *Unacceptable* – does not provide necessary information		modified from Quality Evaluation for Dental Care (1977)
Panoramic radiography (film-based) and Intraoral radiography (film-based)	Whole radiograph classified as: 1 = adequate 2 = marginal (with some technical defect but still acceptable for purpose of interpretation of the marginal bone level) 3 = inadequate	3	Åkesson et al. (1992b)
Panoramic radiography (film-based)	Visualization of diagnostically important structures scored as: 4 = fine details visualized, diagnosis definitely possible 3 = small detailed visualized, diagnosis probably possible 2 = only broad details seen, diagnosis doubtful 1 = significant structures not visible, no diagnosis possible	1	Molander et al. (1995b)
Panoramic radiography (film-based)	Clarity of 12 landmarks assessed as: + 2 = excellent + 1 = good 0 = satisfactory -1 = poor -2 = unacceptable	5 6	Wakoh et al. (1998; 2001)
Intraoral radiography (digital images)	Images graded on a scale, whether important structures for the diagnosis of caries, periodontal disease and periapical pathology rather than on the aesthetics of the image, were: 4 = optimally visualised 3 = adequately visualised 2 = visualised 1 = poorly visualised 0 = not visualized	8	Borg et al. (2000)
Panoramic radiography (film-based)	Confidence on a 4-point scale with which periapical tissues and crestal bone levels could be assessed:	2	Carmichael et al. (2000) modified from

Radiographic method	**Categories/criteria for visual grading analysis**	**Observers (n)**	**Reference**
	1 impossible to make a diagnosis (no fine detail on either side of the tooth seen) 2 probably non diagnostic (only broad outlines seen on one side) 3 probably diagnostic (some fine detail seen on both sides) 4 definitely diagnostic (fine detail seen on both sides, i.e. periodontal ligament space, interdental space and bone height).		Molander et al. (1995b)
Panoramic radiography (film-based)	Visibility of 7 anatomical features assessed as: 1 = very good 2 = good 3 = satisfactory 4 = incomplete 5 = poor	3	Kaeppler et al. (2000b)
Intraoral radiography (digital images)	Images graded for interpretation of 7 different features as: 0 = poor (image unusable for interpretation) 1 = acceptable (image useful for interpretation, but not perfect) 2 = excellent (image perfect for interpretation)	5	Kitagawa et al. (2000)
Panoramic radiography (digital images)	Visibility of 21 anatomical features and visibility of 30 pathological findings scored as: 1 = excellent 2 = more than adequately represented 3 = adequately represented 4 = barely adequately represented 5 = inadequate for diagnosis	10	Dannewitz et al. (2002)
Intraoral radiography (film-based)	Subjective evaluation of dried human mandibular segment containing premolar and molar teeth was performed as: 2 = good image 1 = adequate image 0 = image inadequate for diagnosis	5	Casanova and Haiter-Neto (2004)
Panoramic radiography (digital images)	Subjective image quality for caries, periapical pathology and marginal bone loss, visibility of mandibular canal,	5	Gijbels et al. (2004)

Radiographic method	Categories/criteria for visual grading analysis	Observers (n)	Reference
	condyles and anterior nasal spine evaluated from 1 “certainly impossible to evaluate” to 5 “certainly possible to evaluate”		
Panoramic radiography (film-based and digital images)	Visualization of diagnostically significant 7 anatomical structures: 5 = much better 4 = better 3 = equal 2 = worse 1 = much worse	10	Molander et al. (2004)
Panoramic radiography (film-based and digital)	Visibility of 11 anatomical structures assessed as: 1 = structure well visible 0 = structure partly visible -1 = structure not or hardly visible	5	Kaeppler et al. (2006)
Panoramic radiography (film-based) and Intraoral radiography (film-based)	Visibility grading analysis for scoring alveolar bone level: *Excellent* – provides necessary information for the assessment of alveolar bone level (good density, contrast, sharpness, resolution; right projection; no image distortion and overlapping) *Acceptable* – provides information for the assessment of alveolar level with some defect, which deviates from the ideal, but still acceptable *Unacceptable* – does not provide the necessary information for the assessment of alveolar bone level	6	Ivanauskaite et al. (2008) modified from Åkesson et al. (1989a; 1992b)
	Visibility grading analysis for detection of vertical bone defects and furcation involvements: A*cceptable* – provides information sufficient to assess tooth site for detection of vertical bone defect and alveolar bone between the roots for detection of furcation involvement *Unacceptable* – does not provide information sufficient to assess tooth site for detection of vertical bone defect and alveolar bone between roots for detection of furcation involvement	5	

Radiographic method	Categories/criteria for visual grading analysis	Observers (n)	Reference
Panoramic radiography (film-based) and *Intraoral radiography (film-based)*	Images graded for anatomical structures and pathological findings: 1 = well visible 0 = partly visible -1 = not or hardly visible	3	Peker et al. 2009

1.2.2.2 Diagnostic accuracy efficacy

When the diagnostic accuracy efficacy is analysed, the outcomes of a diagnostic method under study are compared with the outcomes of a reference standard, both obtained in individuals who are suspected of having the disorder of interest. All diagnostic methods have inherent errors. A perfect separation of those with disease from those without disease is therefore not possible. Some with disease will be diagnosed as healthy and some without disease will incorrectly be diagnosed as having the disease. Figure 1.1 illustrates a matrix that is valuable to collate outcomes of diagnostic methods. Individuals, teeth or sites with positive diagnostic findings, which in fact are diseased, are *True Positive (TP)* findings. Those with negative diagnostic findings, which in fact are healthy, are *True Negative (TN)* findings. The false diagnoses can either be *False Positive (FP)* findings when disease is considered to be present although the individuals, teeth or sites are healthy or *False Negative (FN)* findings when there is no diagnostic finding although disease is present. As said above the outcomes of the diagnostic method under study are compared with those of a reference standard. The reference standard is considered to be the best available method for establishing the presence or absence of the condition of interest. The reference standard can be a single method or a combination of methods, to establish the presence of the target condition. It can include laboratory tests, imaging tests, histology, but also dedicated clinical follow-up of subjects (STARD-document Bossuyt et al. 2003). In chronic periodontitis, it is difficult to obtain a reference standard as there is no clear-cut definition of the disease.

Figure 1.1 The probabilities of different outcomes in a population of 100 units (individuals, teeth or sites) of a theoretical diagnostic method under study in comparison with the outcomes of the reference standard. The figures are based on a disease prevalence of 21 percent (in bold) or 10 percent (in italic)

		REFERENCE STANDARD		
		Disease/Finding present +	Disease/Finding absent -	
OUTCOMES OF DIAGNOSTIC METHOD UNDER STUDY	+	True Positive (TP) **16** (*8*)	False Positive (FP) **16** (*18*)	All with a positive outcome **32** (*26*)
	-	False Negative (FN) **5** (*2*)	True Negative (TN) **63** (*72*)	All with a negative outcome **68** (*74*)
		All with Disease/Finding **21** (*10*)	All without Disease/Finding **79** (*90*)	The whole population **100** *(100)*

From the matrix presented in Figure 1.1 some expression used to present the diagnostic accuracy efficacy may be derived as defined by Weinstein et al. (1980) and calculated:

Sensitivity is equal to the true-positive rate *i.e.* the frequency of positive outcomes in those with disease. Based on the figures in the matrix it is:

- 16/21 = 76 percent or 0.76 for a disease prevalence of 21 percent
- 8/10 = 80 percent or 0.80 for a disease prevalence of 10 percent.

Specificity is equal to the true-negative rate *i.e*. the frequency of negative outcomes in those without disease. Based on the figures in the matrix it is:
- 63/79 = 80 percent or 0.80 for a disease prevalence of 21 percent
- 72/90 = 80 percent or 0.80 for a disease prevalence of 10 percent.

Sensitivity and specificity are the probabilities of outcomes given the presence or absence of disease. Clinically it may be more important to know the probabilities of disease given positive or negative outcomes of the method *i.e.* how often disease is present when the outcomes of the diagnostic method are positive and how often disease is absent when the outcomes are negative. This information is derived from the predictive values of the diagnostic methods, defined by Weinstein et al. (1980) as follows:

Predictive value positive is the frequency of disease in those with positive outcomes of the diagnostic method. Based on a disease prevalence of 21 percent (bold figures in Fig. 1.1) it is 16/32 = 50 percent or 0.50. But based on a disease prevalence of 10 percent (italic figures in Fig. 1.1) the predictive value positive is 8/26 = 31 percent or 0.31.

Predictive value negative is the frequency of nondiseased in those with negative outcomes of the diagnostic method. Based on a disease prevalence of 21 percent (bold figures in Fig.1.1) it is 63/68 = 93 percent or 0.93. But based on a disease prevalence of 10 percent (italic figures in Fig. 1.1) the predictive value negative is 72/74 = 97 percent or 0.97.

The disease prevalence makes a difference for predictive values. Based on similar values for sensitivity and specificity of the diagnostic method, the predictive value positive will decrease and the predictive value negative increase with lower prevalence. With higher prevalence the opposite will occur *i.e.* the predictive value positive will increase and the predictive value negative decrease. Other expressions for diagnostic accuracy efficacy are likelihood ratios, the diagnostic odds ratio, and the area under a receiver operator characteristic (ROC) curve.

1.2.2 3 Observer performance

Information obtained by a diagnostic method may be interpreted differently by different observers. Virtually every visual and tactile bit of information from an examination varies to some degree in its prominence from patient to patient. Observers may differ in their ability to detect findings and in their propensity to record them. Even observers, who can agree that they see the same features, may apply different perceptual thresholds in determining the presence or absence of disease. Thus, the reproducibility varies and limited reproducibility adversely affects the diagnostic accuracy efficacy. Therefore authors of scientific studies should, if possible, evaluate the reproducibility of a diagnostic methods implemented and report their procedure to do so (STARD-document Bossuyt et al. 2003). This is important not only in analysis of diagnostic methods for a diagnostic purpose, but also when a diagnostic method is used for planning and deciding an intervention or used to evaluate outcomes of an intervention. If the reproducibility of the diagnostic method used is poor, the reported outcomes of an interventional method could be questioned.

There are different sources for variability of a diagnostic method. Instrument variability concerns the amount of variation that arises during the management of the equipment or device. Other terms for this form of variation include imprecision, analytic methodological variation or analytical noise (error) (STARD-document Bossuyt et al. 2003). To achieve satisfactory instrument reproducibility in radiography one has to pay respect to basic properties, such as projection, density and contrast.

Observer variability can also arise when observers summarize their observations. This variability based on how observers perform sometimes is described as variation sometimes as agreement between readings. In the MeSH "Observer Variation" (http://www.ncbi.nlm.nih.gov/mesh) is defined as "The failure by the observer to measure or identify a phenomenon accurately, which results in an error. Sources for this may be due to the observer's missing an abnormality, or to faulty technique resulting in incorrect test measurement, or to misinterpretation of the data. Two varieties are inter-observer variation (the amount observers vary from one another when reporting on the same material) and intra-observer variation (the amount one observer varies between observations when reporting more than once on the same material)."

Table 1.2 Methods to evaluate observer performance of a diagnostic method and how the performance can be analysed and expressed

What is evaluated?	***How to analyse?***	***How is it expressed?***
One observer's (intra-observer) agreement or variation	The same observer repeats the same measurements or observations with a time interval	Agreement or variation (differences) between measurements or observations expressed as *e.g.* standard deviation, percent, kappa values, coefficients of variation
Several observers' (inter-observer) agreement or variation	Several observers perform the same measurements or observations	

As presented in Table 1.2, the observer performance of a method can be expressed as agreement or differences between measurements or observations. Depending on whether the observation is a dichotomy or a linear measurement, the observer performance can be calculated as percentage agreement between observations, kappa values, standard deviation of a measurement or a correlation between measurements (coefficient of correlation or variation). Generally, intra-observer agreement of a diagnostic method with a specific task is higher than the inter-observer agreement.

Agreement is also higher for observations performed with a short time interval as compared to a longer time interval.

1.2.3 Imaging methods for examination of the oral and maxillofacial region

Radiographs are prescribed for the diagnosis of disease, treatment planning, and follow-up care of patients with abnormalities of the oral and maxillofacial region. There are different radiographic methods used. Most common are bitewing and periapical radiographic examinations, which comprise intraoral radiography with the receptor placed in the oral cavity. Periapical and bitewing radiographs are indispensable adjuncts in the general dental practice to diagnose caries and bone lesions in the jaws. Most general dental practices are equipped with dental X-ray machines so that the dentists can perform intraoral radiography of their patients. A few practices, especially those for specialists in oral health care, are also equipped with panoramic machines.

For orthodontic treatment, panoramic and cephalometric examinations offer valuable information for diagnosis and treatment planning, in particular for adolescents. These radiographic examinations may offer information regarding the patient's condition that is not available through history and clinical examinations. Today, these extraoral radiographic examinations, as is the case with intraoral radiography, are more and more performed with digital receptors, which have replaced film-based systems.

In case of trauma in the maxillofacial region, panoramic radiographs are indispensable when assessing mandibular fractures. When dental injuries are suspected, intraoral radiography is the first option, as these radiographs provide greater diagnostic details than extraoral radiographs. However, if there is clinical evidence of fractures of the facial skeleton, it is probably more appropriate to refer the patient for a complete radiographic examination including Computed Tomography (CT) at the hospital, where treatment will be performed (Salvolini 2002). Also for patients in need of orthognatic surgery of the facial skeleton, a more comprehensive examination with CT is indicated.

For patients with temporomandibular joint (TMJ) symptoms, the panoramic radiograph is often used as a first choice imaging technique. To assess the temporomandibular disc in cases of internal derangements in which simple treatments have been unsuccessful, it may be useful to use Magnetic Resonance Imaging (MRI), which defines both hard and soft tissues and has therefore gradually replaced conventional tomography and CT in the examination of more severe temporomandibular joint disorders. MRI was shown to be the most accurate imaging method for diagnosing disc abnormalities (Liedberg et al. 1996) and to presents osseous changes of the TMJ (Tasaki & Westesson 1993). For radiographic examination of patients with widespread inflammatory lesions of the maxillofacial regions, such as osteomyelitis and phlegmonas, panoramic radiography is often the first choice prior to other radiographic examinations.

When more advanced dental treatment planning is needed, such as in extensive dental implant treatment, intraoral and panoramic radiographs need to be supplemented with more advanced imaging techniques. Recently, an impressive number of Cone Beam Computed Tomography (CBCT) scanners have been introduced for imaging of the maxillofacial region. With CBCT, images in three dimensions and with limited fields of view of the maxillofacial region are obtained. There are units with small fields of view (from 4 to 8 cm) to large ones (from 16 to 20 cm).

There are obvious advantages and disadvantages of CT and CBCT compared to intraoral and panoramic radiography. For the detection of small incremental changes in teeth and surrounding bone, the high resolution of intraoral radiography allows small details of teeth, *e.g*. caries lesions, and lesions of surrounding alveolar bone to be visualized. The diagnostic gain from applying CT and MRI compared to intraoral and panoramic radiography has not been shown to have clinical relevance for the treatment or other advantages for the patients in the general dental practice. Martínez Beneyto et al. (2007) summarized that "Although the radiological doses used by dentists are low individually, patients are often exposed to many repeat dental radiographic examinations". The use of Radiographic Referral Criteria has now become a legal requirement for all practitioners following the adoption of European Legislation. Recently the European Commission has published guidelines (Radiation Protection 136 2004) on radiation protection in dental radiology.

International Commission on radiological protection (ICRP) in 1991 published (ICRP 1990) and in 2007 revised (ICRP 2007) estimates of the radiosensitivity of tissues including those in the maxillofacial region. The effective dose (per the 1991 ICRP) in microsieverts (µSv) for one intraoral radiograph varied between 1-8.3 µSv (Dula et al. 2001; Gijbels et al. 2002; Martínez Beneyto et al. 2007). The dose range depended on the receptor and collimation used: phosphor storage receptor together with rectangular collimation resulted in a lower dose and the use of D-speed film together with round collimation in a higher dose. In the case of panoramic radiography with charge-coupled device, the effective doses (per the 2007 ICRP) were reported to range between 14.2 and 24.3 µSv; (Ludlow et al. 2008) depending on the panoramic machine used. Another study reported the effective dose (per the 1990 ICRP) to range between 8 and 38 µSv depending on the radiographic settings (Gavala et al. 2009). Corresponding effective doses, using standard imaging parameters, varied between 27 and 674 µSv (per the 1990 ICRP) and 27 and 674 µSv (per the 2007 ICRP) with the CBCT scanners and between 350 and 742 µSv (per the 1990 ICRP) and 685 and 1410 µSv (per the 2007 ICRP) with the MSCT (Multislice Computed Tomography) scanners (Suomalainen et al. 2009). Effective doses resulting from CBCT imaging were 1.3–53 times smaller than those from MSCT imaging (Suomalainen et al. 2009). Lofthag-Hansen et al. (2008) reported even lower doses (11-17 μSv) of CBCT for three commonly used examinations in dental radiology. The ALARA (As Low As Reasonably Achievable) principle should be followed in balance with the image quality needed for diagnosing different entities localized to the oral and maxillofacial region. All exposures to X-rays should be clinically justified and each exposure should be expected to give the patient a positive net benefit.

1.3 Examination of the periodontal tissues

Periodontal status should be evaluated for all patients, who search oral health care not only for patients searching periodontal treatment. The methods used in practice are the clinical methods probing of the periodontal pocket to record bleeding, to measure the pocket depth or periodontal attachment level, and the radiological methods to assess the alveolar bone.

1.3.1 Clinical methods for the examination of the periodontal tissues

Probing a periodontal pocket yields either a positive or a negative result regarding bleeding. The positive finding *"bleeding on probing"* is associated with an inflammation, which has been documented in histopathological studies (Greenstein et al. 1981; Cooper et al. 1983; de Souza et al. 2003). There was a significantly larger infiltration of inflammation cells when bleeding on probing occurred (Greenstein et al. 1981; Davenport et al. 1982; Cooper et al. 1983). In a study with histological analysis of tissue biopsies, sensitivity was found to be 0.91, the specificity 0.77, and the predictive values ranged between 0.80 and 0.90 indicating a high accuracy (de Souza et al. 2003).

Furthermore, the absence of bleeding on probing of periodontal pockets that had previously been inflamed indicates that treatment had led to an improvement in periodontal health (Lang et al. 1986; Chaves et al. 1990; Lang et al. 1990; Haffajee et al. 1991).

Healthy periodontal pockets are rarely deeper than 3 mm. There is heterogeneity in the literature when using pocket probing depth as a threshold for periodontitis. In a systematic review (Savage et al. 2009), it was concluded that the minimum threshold defining periodontitis was a depth of 3 mm and the maximum threshold a depth of 6 mm. *Measuring depth of periodontal pockets* overestimates the actual depth of the pocket in the presence of inflammation by 0.1-0.8 mm and underestimates it in normal periodontal tissue (SBU-report 169 2004). The examiner measurement errors are about 1 mm (SBU-report 169, 2004). The measurements can be performed with the aid of a manual periodontal probe or with an electronic probe. The results of a systematic review showed that the reliability of the electronic probe is not better than that of the manual probe (SBU-report 169, 2004).

1.3.2 Radiographic methods for the examination of the periodontal tissues

Radiographic examinations play an integral role in the diagnostics of periodontal tissue to detect alveolar bone loss, as well as in the choice of treatment and in follow-up examinations (Hirschmann 1987; Jeffcoat et al. 1995; Brägger 1996; Rushton et al. 1996; Tugnait et al. 2000a). There is a lack of consensus regarding what constitutes a normal alveolar bone level and bone loss, respectively. In adolescents, 18-year olds with clinically healthy gingiva the radiographic distance between the cemento-enamel junction (CEJ) and the alveolar crest ranged between 0-2 mm when measured in bitewing radiographs (Källestål & Matsson 1989). Benn (1990) presented “A review of the reliability of radiographic measurements in estimating alveolar bone changes" and reported that the distance, which comprise the limit for bone loss "varies in the literature from greater than 1.0 mm (Lennon & Davies 1974; Hugoson & Rylander 1982; Mann et al. 1985), greater than 1.5 mm (Davies et al. 1978), greater than 2.0 mm (Hoover et al. 1981; Kronauer et al. 1986) to over 3.0 mm (Blankenstein et al. 1978; Latcham et al. 1983)". Hausmann et al. (1991) concluded from their study of sites with no attachment loss that radiographic measurements of the CEJ-alveolar crest (AC) distance ranged between 0.4 and 1.9 mm (95% confidence limits) in bitewing radiographs. This means that different thresholds to distinguish between no disease and disease have been proposed.

Alveolar bone level and alveolar bone loss can be assessed in radiographs as relative or absolute measurements (Table 1.3). In relative measurements the alveolar bone level and alveolar bone loss is expressed relative to the root length or the tooth length by score or percentage. In absolute measurements the bone loss is mostly expressed in millimetres. The tooth or root length varies between different teeth and between individuals; for example, the tooth length of a maxillary canine could be between 23.1-28.9 millimetres as compared to that of a mandibular central incisor being between 19.6-23.4 millimetres (Ingle & Bakland 2002). The alveolar bone loss of 7 millimetres would probably mean different prognosis for these two teeth. When the bone loss is expressed relative to the root length, one takes the anatomical variations between teeth and individuals into account. Relative measurements made in radiographs also compensate for the enlargements of different radiographic methods and of different anatomical regions imaged by the same method. Different rulers, as presented in Table 1.3, have been applied for relative measurements to adjust for differences in projection geometry and to avoid measurements errors. With a Schei ruler (Schei et al. 1959), the bone loss is assessed in relation to root length, while the modified Schei ruler

(Engelberger et al. 1963) and Björn rulers (Björn & Homberg 1966; Björn et al. 1969) assess the bone loss in relation to the tooth length. A prerequisite for assessment with these rulers is that the entire root or tooth is imaged in the radiograph. However, this condition is not always fulfilled for example in bitewing radiography from which the alveolar bone level is often assessed clinically. For this reason, Håkansson et al. (1981) designed a ruler to be used in radiographs with partial reproduction of the teeth. To avoid measurement error of absolute measurements, different tools have been used, such as a calibrated ruler or the image of an additional object (Table 1.3). In some studies absolute and relative measurements were combined using computer-based analysis (Table 1.3). The use of computer-based analysis is time saving, because it is easy to mark referents points and then achieve the measurements directly on the computer screen.

Table 1.3 Methods to assess the alveolar bone level and alveolar bone loss in radiography (CEJ = cemento-enamel junction)

Method (Original reference)	Type of measurement	Examples of studies where the method was applied
	Relative measurements	
Ruler according to (Marshall-Day & Shourie 1949)	Ruler with 10 divisions to assess bone loss in relation to maximum bone height	Marshall-Day & Shourie (1949); Marshall-Day et al. (1955)
Schei ruler (Schei et al. 1959)	Ruler with 10 divisions with divergent lines to assess bone loss in relation to root length 1 mm apical to CEJ	Schei et al. (1959); Pepelassi & Diamanti-Kipioti (1997); Lanning et al. (2006a, b)
Modified Schei ruler (Engelberger et al. 1963)	Ruler with 10 degrees with divergent lines to assess bone loss in relation to tooth length	Engelberger et al. (1963); Adriaens et al. (1982)
Björn rulers		
(Björn and	Ruler with 6 divisions with divergent lines to	Björn & Holmberg

Method (Original reference)	Type of measurement	Examples of studies where the method was applied
Holmberg 1966)	assess bone loss in relation to tooth length	(1966); Ahlqwist et al. (1986); Michalowicz et al. (1991); Hugoson et al. (2000); Laurell et al. (2003)
(Björn et al. 1969)	Ruler with 20 divisions with divergent lines to assess bone loss in relation to tooth length	Björn, et al. (1969); Hugoson et al. (2000); Laurell et al. (2003)
(Björn 1974)	Ruler with 10 degrees with divergent lines to assess bone height in relation to tooth length	Björn (1974); Kaimenyi & Ashley (1988); Åkesson et al. (1989a); Rohlin et al. (1989)
Håkansson ruler (Håkansson et al. 1981)	Ruler with divergent lines to assess bone height in score 4 to 10 in radiographs with partial reproduction of the teeth	Håkansson et al. (1981); Åkesson et al. (1989a, b); Balčikonytė (2006) Ivanauskaite et al. (2006)
Ruler to measure in digitized images (Coelho 2010)	Ruler with indices with divergent lines to assess bone loss in relation to root length	Teeuw et al. (2009)
	Absolute measurements in mm	
Ruler compensating for vertical enlargement	Distance between CEJ and alveolar crest	Molander et al. (1991)
Digimatic caliper and splint with steel bolls to measure enlargement	Distance between cusp tip and most apical level of marginal bone	Åkesson et al. (1992a)
Periodontal probe	Distance between CEJ and alveolar crest	Pepelassi & Diamanti-Kipioti (1997); Pepelassi et al. (2000)
Computerized program	Distance between CEJ and marginal bone level	Persson et al. (2003)
	Combined absolute and relative measurements	

Method (Original reference)	Type of measurement	Examples of studies where the method was applied
Computer-based analysis	Linear measurement of distance CEJ to bone level in mm (A) Proportion of A relative to root length and/or tooth length	Hausman et al. (1992) Walsh et al. (1997); Eickholz et al. (1998); Eickholz & Hausmann (2000); Kim et al. (2002); Persson et al. (2003); Pūrienė et al. (2003); Kim et al. (2008)

1.3.2.1 Intraoral and panoramic radiography

Intraoral radiography with bitewing and periapical radiographs has been the routine method for the examination of the alveolar bone (Björn et al. 1969; Green et al. 1978; Håkansson et al. 1981; Selikowitz et al. 1981; Albandar et al. 1985; Salonen et al. 1991). As reported by Albandar et al. (1985), there was no significant difference between mean alveolar bone loss measured in periapical radiography as compared to bitewing radiography. Despite the widespread use, intraoral radiography has shortcomings for assessing alveolar bone loss. There are different methodological errors as presented in detail by Benn (1990) and Brägger (2005). Horizontal and vertical alignment errors exist in all intraoral radiographic techniques due to geometric parameters. Variations of the X-ray beam in the vertical plane will change the image of the alveolar level in relation to the CEJ. "If the physiological distance of CEJ to alveolar crest is set 1.5 mm and geometric projection errors of ±0.5 mm produce a range of 1.00 to 2.00 mm CEJ-crest distances, then a number of areas will be incorrectly classified as having suffered bone loss, while others will have "gained" 0.5 mm" (Benn 1990). Horizontal angulations in combination with changes of vertical angulations have a considerable influence on the image of the CEJ and the AC, which "may be of clinical significance, and a critical attitude to alveolar bone loss measures on radiographs using the radiographic CEJ as reference points is recommended" (Sewerin et al. 1987). Reproducibility of the images of the reference points can be improved by the use of intraoral positioning devices (Rushton et al. 1994; Potter et al. 1995; Rushton et al. 1995).

There are some clinical studies on the accuracy of assessment of alveolar bone loss in intraoral radiography. Intraoral radiography mostly underestimates the actual loss of the alveolar bone as compared to surgical measurements in patients with severe periodontitis (Suomi et al. 1968; Renvert et al. 1981; Hämmerle et al. 1990; Åkesson et al. 1992a; Eickholz et al. 1998; Eickholz & Hausmann 2000; Pepelassi et al. 2000). One reason for the underestimation could be that vertical and particularly horizontal angulations differ between the central beam and the orthoradial projection (Eickholz et al. 1998).

In addition to intra-oral radiography, *panoramic radiography* has been shown to be a useful adjunct for the examination of the alveolar bone and for the diagnosis of periodontal disease (Ehrlich et al. 1977; Horton et al. 1977; Douglass et al. 1986; Rohlin et al. 1989; Molander et al. 1995a; Flint et al. 1998). In a survey of radiographic practices for periodontal diseases in the UK and Irish dental teaching hospital it was found that all hospitals used panoramic and specific periapical radiographs as one of their radiographic regimes. Most respondents most frequently took panoramic and selected periapical radiographs (Tugnait et al. 2000b).

Panoramic radiography has compared favourable with intra-oral radiography concerning the assessment of the alveolar bone level both for clinical application (Gröndahl et al. 1971; Åkesson et al. 1989a, b; Rohlin et al. 1989; Molander et al. 1991) and for epidemiological studies (see for example Ahlqwist et al. 1986). The overall concordance between panoramic and intraoral radiography has been found to range between 50 (Gröndahl et al. 1971) and 70 percent (Åkesson et al. 1989a) of the assessed sites. Ahlqwist et al. (1986) found a very high concordance (correlation coefficient 0.96) between panoramic radiographs and intraoral full mouth surveys when assessing the alveolar bone level with a five-graded ruler. However, the concordance between intraoral and panoramic radiography varies between different regions of the jaws (Gröndahl et al. 1971; Adriaens et al. (1982); Åkesson et al. 1989a,b; Rohlin et al. 1989; Molander et al. 1991). Results of studies on the diagnostic accuracy of panoramic radiography compared with reference standard that comprised probing during surgery varied. While Åkesson et al. (1992a) found the accuracy of panoramic radiography to be comparable to that of intraoral radiography Pepelassi and Diamanti-Kipioti (1997) found the accuracy of panoramic radiography to be lower. One reason for the deviating results of different studies may be differences in the panoramic equipment used and therefore in the image quality of the panoramic radiographs. For the detection of vertical bone defect and furcation involvement intra-oral radiography has been shown to be superior to panoramic radiography (Åkesson et al. 1989a, b; Rohlin et al. 1989; Molander et al. 1991; Pepelassi & Diamanti-Kipioti 1997; Pepelassi et al. 2000).

Since these studies were performed, the Scanora® system (Sorodex, Helsinki, Finland) was developed for radiographic examinations of the dental and maxillofacial region (Tammisalo et al. 1992) as a multimodal X- ray unit that combines panoramic dental radiography with spiral tomography. Its development has renewed the interest in the application of tomography of the dental and maxillofacial region as well as the application of panoramic radiography for dental diagnostics. In a series of studies, Tapio Tammisalo and co-workers analysed the efficacy of the tomographic modes of the

Scanora® system that comprise detailed narrow beam technique and detailed zonography.

The diagnostic performance of the detailed narrow-beam technique was compared with periapical radiography for detecting periodontal pathology (marginal widening of periodontal membrane space, crestal erosion, vertical bone loss, furcation involvement or calculus) and found to be a good radiographic examination for periodontal disease, and an acceptable alternative to periapical radiography (Tammisalo et al 1994). Also for the detection of periapical bone lesions detailed narrow-beam radiography performed as well as periapical radiography. (Tammisalo et al. 1993). The diagnostic accuracy of detailed zonography using the Scanora multimodal X-ray system was also compared with that of periapical radiography for the detection of periodontal and periapical lesions (Tammisalo et al. 1995a, b). The conclusion was that zonography performs as well as periapical radiography in the detection of periodontal disease (Tammisalo et al. 1995b). Diagnostic accuracy in detailed tomography for periapical and periodontal lesion were compared with periapical radiography and conclusion was that both methods performed equally well for the overall diagnosis of periapical and periodontal lesions (Tammisalo et al. 1996.)

For panoramic radiography, Scanora® has two programs for the examination of the jaws. One program, the jaw program, has a magnification factor 1.3 as most panoramic machines utilize. The other one, the dental program, has a magnification factor 1.7. Molander et al. (1995b) showed that the Scanora® dental program provides the best image quality and argued that better image quality is due to the provision of a rotating anode and, an X-ray tube with a focal spot 0.3 x 0.3mm². In other panoramic machines it is either 0.5 x 0.5mm² or 0.6 x 0.6mm². The X-ray beam is narrower than in other panoramic machines. The subjective image quality of panoramic radiograph using Scanora® jaw program was comparable with other panoramic equipment for dental diagnosis (Molander et al. 1995b). These results were in line with those of Kaeppler et al. (2000b), who found the visualisation of seven anatomical structures in radiographs of Scanora® jaw program and Orthophos Plus® (program P1) to be comparable.

1.4 Systematic reviews

Scientific assessment in offer the greatest benefits for patients while utilizing resources in the most efficient way (SBU 2010a). Many routine methods for different interventions are ineffective. Some newer methods are widely used, even though their benefits, risks, and costs have never been critically evaluated. At the same time, there are methods that should be used on a much broader scale – methods shown by scientific assessment to be both beneficial and cost effective (SBU 2010b). Health care aims to identify interventions that.

A systematic review is a systematic retrieval of results from previous scientific studies and a way to identify which methods of health care work and which do not (Cochrane Library). Scientific studies are assessed in a systematic way. The results of studies meeting pre-established criteria and unbiased are combined to produce more reliable results. This combination of results as is the case with the searches of scientific studies has to be described in a reproducible manner. A systematic review uses predefined explicit methodology (Cochrane Library) and is therefore considered to be an original study. The methods used include steps to minimize bias in all parts of the process: inclusion criteria are established for the quality and relevance of the studies, available research findings addressing the important issues are systematically searched in computerized databases, research report is carefully reviewed and evaluated, and results from the selected studies are scrutinized, and used to form the body of evidence. Thus, the methodology of systematic reviews is different to other reviews.

Systematic reviews of diagnostic methods are powerful tools for producing measures of the diagnostic performance for a patient group of interest. By summarizing available evidence and explain differences among scientific studies, systematic reviews can help clinicians in their decision-making. By identifying knowledge gaps, systematic reviews can also help researchers to formulate relevant research questions for the future.

2. AIMS

2.1 Relevance of theme

In dental practice radiographic modalities are used in addition to clinical methods to gain information about the patients. The information gained from diagnostic examinations has different purposes. The examination can be used for evaluation of the oral health status, for judgements leading to a diagnosis or a differential diagnosis, for treatment planning, for the evaluation of treatment outcomes or for epidemiological studies. In the examination of the periodontium, which is comprised of the gingiva, alveolar bone, periodontal ligament, and cementum, the radiographic examination is an important modality to assess the status of the alveolar bone. The diagnosis periodontitis is based on a finding of alveolar bone loss. Changes of the alveolar bone can be assessed by different radiographic modalities, such as intraoral radiography (bitewing and periapical radiography) and panoramic radiography. Results of studies on the diagnostic performance of panoramic radiography to assess the alveolar bone deviate. Since the studies on panoramic radiography were performed, the Scanora® multimodal system was developed and is used world-wide for radiographic examinations of the dental and maxillofacial region. The dental program of this radiographic machine has a magnification factor of 1.7 and can be expected to provide a better image quality and thereby a better diagnostic performance than other panoramic modalities. However, the diagnostic performance of the dental program of Scanora® has not been analysed for the assessment of the alveolar bone, neither for clinical use nor for epidemiological studies.

One way to assess the diagnostic performance is to perform a visual grading analysis, which has not been made for panoramic radiography with the Scanora® dental program as compared with visibility levels of posterior bitewing radiography, which is frequently used for the assessment of alveolar bone. If the visibility is comparable for the radiographic modalities it may relevant to utilize panoramic radiography as this is more comfortable for the patients. Furthermore, observer performance of visual grading analysis has not been studied for panoramic radiography at all. It could be that observers will vary in their analysis of visibility of images, as occurs in other judgemental tasks.

Other diagnostic properties, such as the correspondence of findings between panoramic radiography using Scanora® dental program and posterior bitewing radiography, is also of interest.

In oral health care, however, systematic reviews have been more focused on preventive methods and treatment methods. There is no systematic review on panoramic radiography in the assessment of alveolar bone loss for diagnosis of periodontal disease.

Hence, analysis of alveolar bone loss in radiographic modalities and systematic review on could be helpful to suggest the more applicable radiographic methods for diagnosis of periodontal disease.

2.2 Purpose

The purpose of the present study was to examine diagnostic properties of panoramic radiography for the assessment of alveolar bone loss (alveolar bone level, detection of vertical bone defect and furcation involvement) for the diagnosis of periodontal diseases as compare to posterior bitewing radiography.

2.3 Tasks

1. To perform visual grading analysis for the assessment of alveolar bone level, detection of vertical bone defect and furcation involvement in panoramic radiography and compare with that of posterior bitewing radiography.

2. To evaluate the observer performance of visual grading analysis for the assessment of alveolar bone level, detection of vertical bone defect and furcation involvement in panoramic radiography and in posterior bitewing radiography.

3. To compare panoramic radiography and posterior bitewing radiography for the assessment of alveolar bone level, detection of vertical bone defect and furcation involvement.

4. To evaluate the observer performance for the assessment of alveolar bone level, detection of vertical bone defect and furcation involvement in panoramic radiography and in posterior bitewing radiography.

5. To evaluate evidence of the scientific literature on panoramic radiography for visual grading analysis, for diagnostic accuracy, and observer performance in the assessment of alveolar bone loss.

2.4 Statements to be defended

1. Observers will vary in their analysis of visibility of radiographic images, as occurs in other judgemental tasks. The results of visual grading analysis will influence the correspondence between results of assessing alveolar bone loss in panoramic radiography and those of posterior bitewing radiography.

2. For the assessment of alveolar bone loss, the results obtained in panoramic radiography are comparable with those in posterior bitewing radiography.

3. The systematic retrieval and synthesis of results from previous studies on panoramic radiography for diagnosis of periodontal disease will reveal that that the studies are performed with different design and on different samples. Important knowledge gaps will be identified, in particular on the diagnostic accuracy efficacy as such studies are difficult to perform.

3. MATERIAL AND METHODS

Ethical consideration: The Ethics Committee of Lund University, Sweden, approved the study on the 8th of December 1999, Protocol LU 628-99.

3.1 Panoramic radiography and posterior bitewing radiography

3.1.1 Patients

Ninety-six consecutive patients with a residual dentition of both anterior and posterior teeth in the upper and lower jaws, who were referred to the Department of Oral and Maxillofacial Radiology, Faculty of Odontology, Malmö University, Sweden for radiographic examination of teeth and surrounding bone, were selected. The patients were attending the Faculty of Odontology for comprehensive oral health care. The age and sex distribution of the patients is shown in Table 3.1. Forty-four patients were men with a mean age of 49 years (range 21 to 78 years) and 52 were women with a mean age of 48 years (range 20 to 85 years).

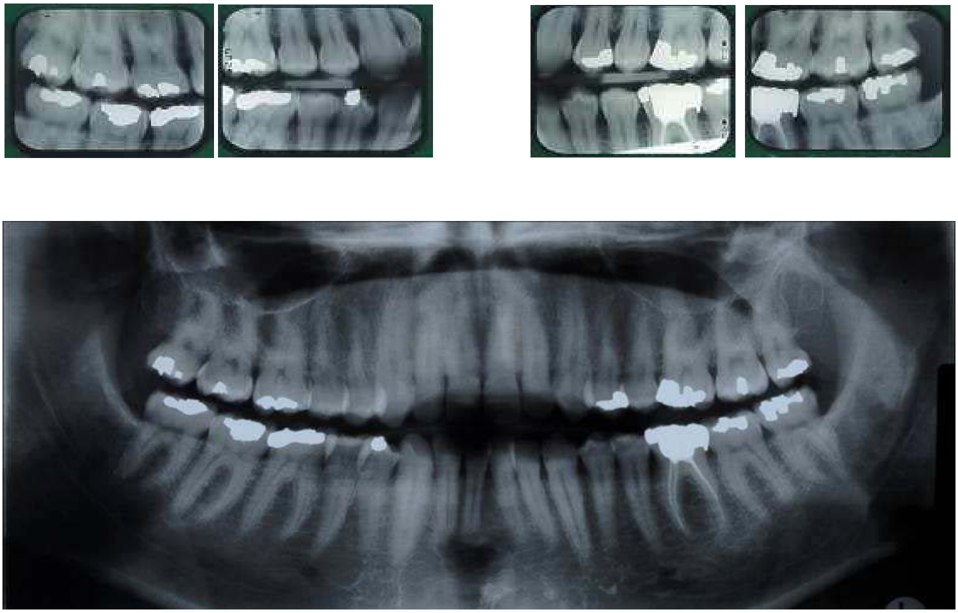
Table 3.1 Sex and age distribution of the patients

Sex	*Age-groups*							
	20-29	*30-39*	*40-49*	*50-59*	*60-69*	*70-79*	*80-89*	*Total*
Men	4	7	12	13	2	6	0	44
Women	7	8	15	12	3	3	4	52
Total	11	15	27	25	5	9	4	96

3.1.2 Radiographic examinations

For each patient a radiology technologist (dental assistant educated in radiographic techniques) performed panoramic radiography and a dentist and specialist in periodontology with experience of 1 year in dental and maxillofacial radiology (D.I., the author of the study) performed posterior bitewing radiography. An example of radiographs of a patient is presented in Figure 3.1.

Figure 3.1 Posterior bitewing and panoramic radiographs of a patient



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). The Scanora® dental panoramic program 003, which has a magnification factor of 1.7, was used. The voltage settings were programmed according to the preset settings 4/2 (66 kV, 10 mA, 15 s), 4/3 (66 kV, 13 mA, 15 s), 4/4 (66 kV, 15 mA, 15 s), 4/5 (66 kV, 20 mA, 15 s), 4/6 (66 kV, 20 mA, 19 s), 4/7 (66 kV, 20 mA, 23 s) or 5/5 (70 kV, 16 mA, 16 s). The vertical angulation of the tube was constant -5°. The films were processed in an automatic processor (Curix HT-33OU, AGFA, Belgium) with the developer G138 I (AGFA, Belgium) and a developing time of 2 minutes at temperature 32° C. In the following text on "Material and methods" this method will be referred to as panoramic radiography.

Posterior bitewing radiography was performed with a Heliodent 70 X-ray unit (Siemens, Erlangen, Germany), which operated at 70 kV and 7 mA. The bitewing radiographs were taken using Kwik-bite film-holder (Hawe-Neos Dental, Gentilino, Switzerland) for horizontal posterior bitewing radiographs and with Take-All film-holder (Wijkström, Menton, France) 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 +10°. Ektaspeed Plus films (Eastman Kodak Co., Rochester, N.Y., USA) were used and the exposure time was 0.32-0.64 s. The films were processed in an X-ray film automatic processor (XR 24 Nova, Dürr Dental, Bietigheim, Germany) with the developer Kodac Readymatic (Eastman Kodak Co., Rochester, N.Y., USA) and a developing time of 6 minutes at temperature 28º C. The bitewing films were mounted in opaque frames (Trollhätteplast AB, Trollhättan, Sweden). In the following text on "Material and methods" this method will be referred to as bitewing radiography.

The overall image quality of panoramic and bitewing radiographs was assessed by the author of the study as adequate or poor. Radiographs of poor quality in terms of film placement, projection, centring, density, contrast or sharpness were retaken before the radiographic examination of each patient was considered completed. For ethical reasons, only one retake was made to minimise the radiation dose to the patient.

3.1.3 Observations and observers

Observers were asked to participate in the study and make assessments of the panoramic and bitewing radiographs. Prior to the assessment, the observers had a joint discussion and calibration on how to assess the radiographs. The assessment criteria were discussed among the observers; the categories were specified and then written down by the author (D. I.) in a protocol "*Observer Instructions*" (Appendix 1). In order to ascertain that the criteria were appropriate and that the "*Observer Instructions*" (Appendix 1) was comprehensible, a number of radiographs were assessed by some of the observers prior to the real assessment took place. The protocol *"Assessment of radiographs"* covered some parts: *(i)* alveolar bone level, *(ii)* vertical bone defect, and *(iii)* furcation involvement (in Appendix 2). Prior to the scoring of the alveolar bone level and the detection of vertical bone defect or furcation involvement, the observers had to assess the image quality by visual grading analysis. Tooth sites or teeth presenting unacceptable image quality for the task were not scored or assessed. During all observations, the protocol "*Observer Instructions*" (Appendix 1) was available as a reference that the observer could follow step by step. Each observer made the observations independently. A light box (15x30 cm^2) with fixed intensity was used and placed in a quiet room with dimmed background lighting. The maximum observation time was one hour for both radiographic methods. The observers used a magnification viewer (DAB Dental, Sweden) with 2 times magnification when possible. Observations in panoramic radiographs and bitewing radiographs were made at a minimum interval of one week.

The observers were not informed about the periodontal status when it comes to the alveolar bone level. Neither was the observers informed about the number of tooth sites with a vertical bone defect and the number of teeth with a furcation involvement prior to their assessment.

Six observers with varying experience in dental and maxillofacial radiology (mean experience: 13 years; experience of each observer: 1, 3, 7, 16, 19, and 30 years) were asked to make the assessments. Three observers were specialists in dental and maxillofacial radiology with experience of 16, 19, and 30 years in dental and maxillofacial radiology. One observer had experience of 1 year in dental and maxillofacial radiology and was specialist in periodontology (author of the study). Two other observers were dentists, one with experience of 3 years and one with experience of 7 years in dental and maxillofacial radiology and completing specialist trainee programme in dental and maxillofacial radiology.

The six observers assessed the image quality for scoring of the alveolar bone level and scored the alveolar bone level, *i.e*. completed the first part of the protocol *"Assessment of radiographs"* (in Appendix 2). Five of the six observers (mean experience in dental and maxillofacial radiology: 15 years; experience of each observer: 3, 7, 16, 19, and, 30 years) completed the second and third parts of the protocol *"Assessment of radiographs"* (in Appendix 2). The observers assessed the image quality for the detection of vertical bone defects and of furcation involvement. When acceptable, they then recorded whether a vertical bone defect or a furcation involvement was present.

Three observers (mean experience in dental and maxillofacial radiology: 14 years; experience of each observer: 7, 16, and 19 years) made a second observation by completing all parts of protocol *"Assessment of radiographs"* (in Appendix 2) after four weeks to enable calculations of intra-observer agreement.

3.1.4 Teeth and sites for assessment

3.1.4.1 Teeth and sites for assessment of alveolar bone loss

Radiographs of the 96 patients were randomly divided in six groups (each group comprised 16 patients). The tooth sites to be assessed for visual grading analysis and for the assessment of alveolar bone loss by score were selected from the distal site of the second molar to the distal site of the canine and were divided in rotation into six sites. Table 3.2 presents the selection of actual tooth sites within patients' six groups.

The same six tooth sites within the same patient-group were assessed. For example, for the 16 patients of group 1, the sites from 17 distally to 16 distally and from 37 distally to 36 distally were assessed. Since not all patients had a full dentition, out of 576 tooth sites (96 patients x 6 sites), 499 tooth sites (245 in the upper jaw and 254 in the lower jaw) were imaged and available for assessment. The number of available tooth sites is presented within the shadowed areas in Table 3. 2.

Each observer evaluated the image quality of available tooth sites. When image quality was unacceptable, the alveolar bone loss of that site was not scored. In the randomly divided patient groups, the number of the tooth sites with different degrees of bone loss was unknown to the observers. Table 3.3 presents the distribution of scores of the tooth sites as assessed by one of the observers in panoramic radiography. Most tooth sites were scored as either 5 or 6, which corresponds to a bone loss of one third or less of the root length.

Table 3.2 Patient groups and number of tooth sites in shadowed areas available for visual grading analysis and for scoring the alveolar bone loss: d – distal tooth sites; m – mesial tooth sites

Patient Group	*Upper jaw, tooth*																	
	17		16		15		14		13	23	24		25		26		27	
	d	m	d	m	d	m	d	m	d	d	m	d	m	d	m	d	m	d
1	15	15	14															
2				13	14	14												
3							11	11	15									
4										16	13	13						
5													15	15	12			
6																15	12	12
Patient Group	*Lower jaw, tooth*																	
	47		46		45		44		43	33	34		35		36		37	
	d	m	d	m	d	m	d	m	d	d	m	d	m	d	m	d	m	d
1																13	14	14
2													15	15	12			
3										16	14	14						
4							16	16	16									
5				10	14	14												
6	15	15	11															

Table 3.3 Distribution of the alveolar bone loss scores of assessed tooth sites. The assessment was made by one observer in panoramic radiography. Score 4 indicates normal level of the alveolar bone and score 5 – 10 indicated equidistant levels of alveolar bone loss

			Upper jaw			
Score	13/23	14/24	15/25	16/26	17/27	Total
4	1	1		2		4
5	7	12	24	24	26	93
6	8	6	17	18	19	68
7	1	3	6	7		17
8	1				1	2
9						
10						
Total	18	22	47	51	46	184
			Lower jaw			
Score	33/43	34/44	35/45	36/46	37/47	Total
4	3	3	4	3	1	14
5	18	34	33	31	40	156
6	9	15	13	9	14	60
7	1	4	1	1	1	8
8			1		1	2
9			1			1
10						
Total	31	56	53	44	57	241

3.1.4.2 Teeth and sites for detection of vertical bone defect and furcation involvement

For visual grading analysis for the detection of *vertical bone defects*, all imaged proximal tooth sites from the distal site of the second molar to the distal site of the canine (1435 sites in the upper jaw and 1450 sites in the lower jaw) of all 96 patients were assessed. Table 3.4 presents the distribution of the tooth sites. When image quality was unacceptable, the presence of a vertical bone defect could not be assessed.

Table 3.4 Number of tooth sites available for visual grading analysis for detection of vertical bone defect: d – distal tooth sites; m – mesial tooth sites

	Second molar		First molar		Second premolar		First premolar		Canine	Total
	d	m	d	m	d	m	d	m	d	
Upper jaw	150	150	153	153	157	157	164	164	187	1435
Lower jaw	152	152	129	129	169	169	180	180	190	1450
Total	302	302	282	282	326	326	344	344	377	2885

The vertical bone defects were not assessed clinically or surgically.

For visual grading analysis for the detection of *furcation involvements*, all imaged teeth with a root furcation were assessed. In total 584 molars (303 in the upper jaw and 281 in the lower jaw) and 164 first premolars in the upper jaw of the 96 patients were assessed. The furcation involvements were not assessed clinically or surgically.

3.1.5 Visual grading analysis

Visual grading analysis of the alveolar bone included assessment of subjective image quality in panoramic and bitewing radiography for the assessment of the bone loss by score and for the detection of vertical bone defect and furcation involvement.

3.1.5.1 Visual grading analysis for assessment of alveolar bone loss

Image quality of tooth site concerning the assessment of the alveolar bone loss by score was assessed. Three categories originally proposed by the California

Dental Association (1977) and modified by Åkesson et al. (1989b, 1992b) were further modified for this study on the visibility of alveolar bone:

- Excellent – provides necessary information for the assessment of bone loss (good density, contrast, sharpness, resolution; right projection; no image distortion and overlapping)

- Acceptable – provides information for the assessment of bone loss with some defect, which deviates from the ideal, but still acceptable

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

The procedure for observers' assessment and the categories for visual grading analysis of the alveolar bone level are described in protocol *"Observer Instructions*" (Appendix 1). Examples of images presenting different categories of image quality were available during the assessments. Examples are presented in Figures 3.2 and 3.3 Tooth sites assessed as "Excellent" and "Acceptable" proceeded to the assessment of bone loss by score.

Figure 3.2 A panoramic radiograph representing Excellent, Acceptable, and Unacceptable image quality. For the assessment of the alveolar bone loss the tooth sites 17 distally and mesial, 16 distally, 27 distally represent Excellent image quality, 27 mesially represents Acceptable image quality, and 23 distally, 24 and 25 mesially and distally, 26 mesially represent Unacceptable image quality

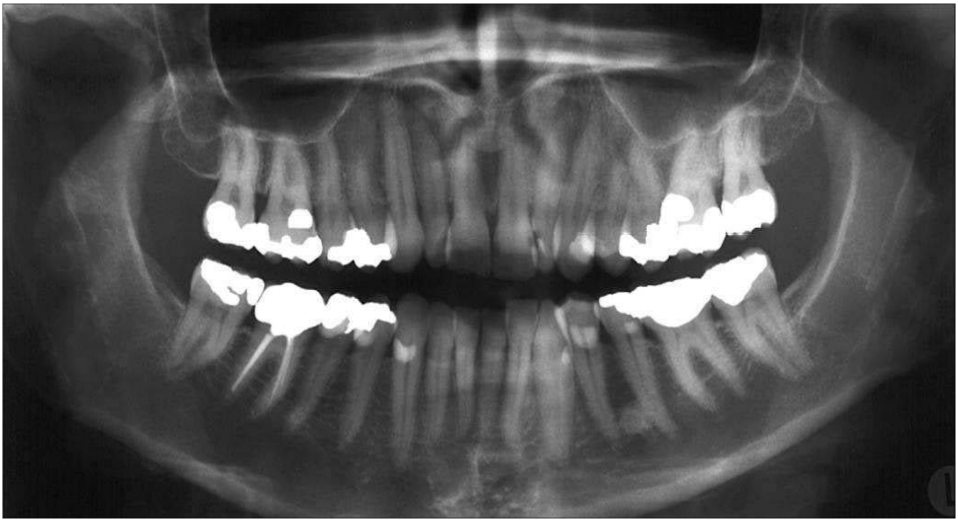
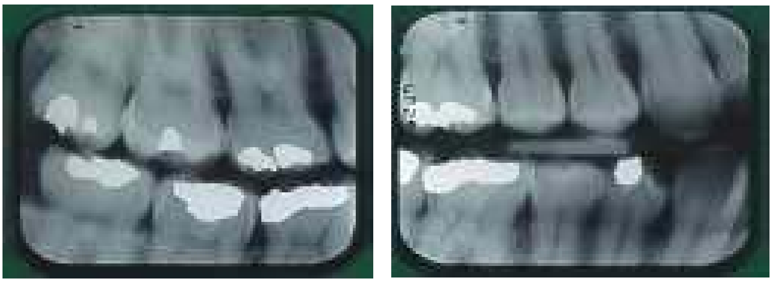


Figure 3.3 Posterior bitewing radiographs of the right side of the upper and lower jaws representing Excellent and Acceptable image quality in all tooth sites. For the assessment of alveolar bone loss the tooth sites 16 distally and mesially, 15 distally and mesially represent Excellent image quality and 17 distally and 47 distally represent Acceptable image quality



3.1.5.2 Visual grading analysis for detection of vertical bone defect and furcation involvement

Two categories were used in grading the visibility of alveolar bone for detection of vertical bone defect and furcation involvement:

- Acceptable – provides information sufficient to assess tooth site for detection of vertical bone defect and the bone between the roots for detection of furcation involvement

- Unacceptable – does not provide information sufficient to assess tooth site for detection of vertical bone defect and the bone between the roots for detection of furcation involvement.

The procedure for observers' assessment and the categories for visual grading analysis of the tooth sites and tooth furcation are described in protocol *"Observer Instructions"* (Appendix 1). Tooth sites and tooth furcation assessed as "Acceptable" proceeded to the detection of vertical bone defect and furcation involvement.

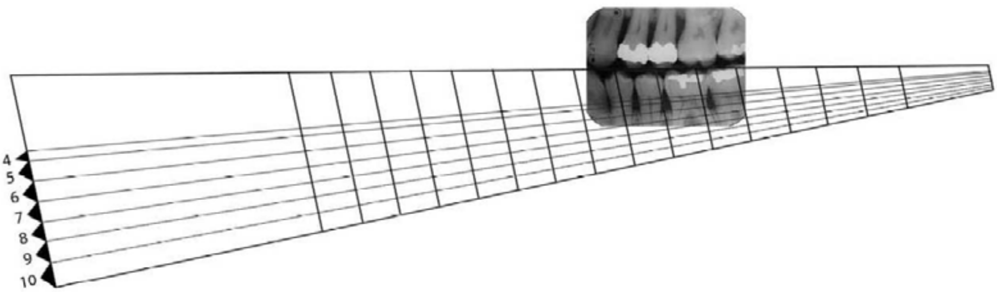
3.1.6 Assessment of alveolar bone loss

Assessment of alveolar bone loss included assessment of the bone level by score and the detection of vertical bone defect and furcation involvement in panoramic and bitewing radiography.

3.1.6.1 Assessment of bone level

The bone level was scored using a ruler designed by Håkansson et al. (1981) (Figure 3.4). A vertical line of the ruler was placed parallel to the longitudinal axis of the tooth to be assessed. The reference points were the tip or incisal edge of the crown of the tooth, the cemento-enamel junction (CEJ) of the tooth site and the alveolar bone crest of the tooth site, where the periodontal ligament space was considered to have a radiographically normal width. The last reference point was read as a score on the ruler. When the CEJ was not clearly visible or missing because of restoration of the crown, the CEJ was imagined. If the bone level was presented as two edges of the alveolar bone crest, the more apical edge was used. The assessment was expressed by scores from 4 to 10, where score 4 indicated a normal bone level and scores from 5 to 10 indicated equidistant levels of bone loss. The higher score the more apical the bone level was equal to more bone loss (Figure 3.4). In protocol *"Observer Instructions"* (Appendix 1) the procedure for the assessment of the bone level by score is described.

Figure 3.4 The ruler, which was used in this study to score a bone loss, placed on a bitewing radiograph. The ruler was originally designed by Håkansson et al. (1981) to assess the bone height in radiographs with partial reproduction of the teeth (either part of the crown or part of the root is missing)



3.1.6.2 Detection of vertical bone defect and furcation involvement

The alveolar bone of each tooth site was assessed to detect a vertical bone defect. A *vertical bone defect* was considered present when the bone presented an angular radiolucency, a circumscribed alveolar bone pocket, one-, two-, three-walled bone defects or two edges of the alveolar bone crest adjacent to the root surface. In protocol *"Observer Instructions"* (Appendix 1) the procedure for the assessment is described.

The alveolar bone between the tooth roots was assessed for the detection of *furcation involvement.* Furcation involvement was assessed as present or absent and was considered present when there was an obvious radiolucency between the tooth roots in the bone. In protocol *"Observer Instructions"* (Appendix 1) the procedure for the assessment is described.

3.1.7 Analysis

Panoramic and bitewing radiography was analyzed regarding visual grading analysis and regarding the assessment of the alveolar bone. Observer performance, expressed as intra-observer and inter-observer agreement, for the evaluation of image quality by visual grading analysis and for the assessment of alveolar bone was also analyzed.

3.1.7.1 Visual grading analysis for assessment of alveolar bone loss

Image quality was expressed as number of tooth sites and as percent of total number of assessed tooth sites in panoramic and bitewing radiography evaluated as excellent, acceptable or unacceptable image quality for the assessment of the alveolar bone level by score. For the detection of vertical bone defect and for furcation involvement, image quality was expressed as number and as percent of sites and tooth, respectively, evaluated as acceptable or unacceptable.

3.1.7.2 Assessment of alveolar bone level

For the assessment of the alveolar bone level by score, the agreement between panoramic and bitewing radiography was expressed as number and as percent of sites that presented identical scores in both radiographic methods of total number assessed. Agreement between panoramic and bitewing radiography was also calculated for sites, where there was a difference of one score between the methods.

3.1.7.3 Detection vertical bone defect and furcation involvement

For the detection of vertical bone defect and furcation involvement, the agreement between panoramic and bitewing radiography was expressed as number and as percent of sites/teeth presenting the same assessment (presence or absence) in both radiographic methods. Sensitivity *i.e.* sites/teeth with positive test results among all sites/teeth with vertical bone defect/furcation involvement was also calculated for panoramic and bitewing radiography, respectively. The results of the assessments of panoramic radiography together with bitewing radiography were used as the consensus radiographic standard.

3.1.7.4 Visual grading analysis and concordance between assessment of alveolar bone loss

To evaluate the influence of image quality on the concordance between assessments made in panoramic and bitewing radiography, the number of pairs where the sites were assessed to be excellent and presented the same score for alveolar bone loss in both radiographic methods was first calculated. Then, the number of pairs, where one site, irrespective of in panoramic or in bitewing radiography, was assessed to be acceptable, was calculated. The concordance between the radiographic methods was expressed as overall agreement and kappa values.

3.1.7.5 Observer performance

3.1.7.5.1 Intra-observer performance

Intra-observer performance was expressed as overall agreement in percent and as kappa value as described by Cohen (1960). The calculations were based on three observers' readings of visual grading analysis and of assessment of the alveolar bone loss.

3.1.7.5.2 Inter-observer performance

Inter-observer performance was expressed as kappa value for several observers as described by Fleiss (1971) for visual grading analysis in the assessment of the bone level and detection of vertical bone defects as well as for the assessment of alveolar bone loss. Additionally, for visual grading analysis and for the assessment of bone level by scores, inter-observer agreement for pairs of observers was calculated as *(i)* overall agreement in percent, *(ii)* kappa value, and *(iii)* Cohen's weighted kappa values. Furthermore, inter-observer agreement for pairs of observers was calculated as *(i)* overall agreement in percent and *(ii)* Cohen's weighted kappa values for visual grading analysis for the detection of vertical bone defects and of furcation involvement.

3.1.7.5.3 Interpretation of kappa values

The six-point scale as proposed by Landis and Koch (1977) for agreement of categorical data was used to interpret the kappa values. Values less than zero were termed poor agreement, 0.00 to 0.20 slight, 0.21 to 0.40 fair, 0.41 to 0.60 moderate, 0.60 to 0.80 substantial, and values higher than 0.81 indicated almost perfect agreement.

3.2 Evaluation of evidence on panoramic radiography by means of a systematic review

Evaluation of evidence was made by means of a systematic review. In order to achieve a systematic approach, the literature review was adapted according to Goodman (1996) and comprised the following steps: (*i*) problem specification, (*ii*) formulation of a plan for the literature search, (*iii*) literature search and retrieval of publications, and (*iv*) data extraction, interpretation of data, and evaluation of evidence from the literature retrieved.

3.2.1 Problem specifications and definitions of terms

For evaluation of the evidence in the scientific literature on panoramic radiography for the assessment of alveolar bone loss in periodontal diseases, the problem specifications were:

• What is the evidence for visual grading analysis?

• What is the evidence for diagnostic accuracy?

The following terms were chosen and defined:

In the Medical Subject Headings (MeSH) data base (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=mesh):

- "Radiography, Panoramic" was defined as "Extraoral body-section radiography depicting an entire maxilla, or both maxilla and mandible, on a single film". Year introduced: 1974 (http://www.ncbi.nlm.nih.gov/mesh: Radiography, Panoramic).

- "Periodontal Diseases" was defined as "Pathological processes involving the PERIODONTIUM including the gum (GINGIVA), the alveolar bone (ALVEOLAR PROCESS), the DENTAL CEMENTUM, and the PERIODONTAL LIGAMENT". Year introduced: 1965 (http://www.ncbi.nlm.nih.gov/mesh: Periodontal Diseases).

Visual grading analysis and diagnostic accuracy were not found as terms in MeSH. In this systematic review these terms were defined as follows:

- Visual grading analysis:

- image quality
- measurability/unmeasurability

- readability/non-readability
- interpretability/uninterpretability.

- Diagnostic accuracy:
  - concordance/correlation between panoramic radiography and other diagnostic methods
  - yield of normal and abnormal diagnoses in a case series
  - percentage of correct diagnoses
  - sensitivity, specificity, and predictive values
  - measures of ROC curve height or area under the curve
  - observer performance, observer agreement and “Observer variation”, (http://www.ncbi.nlm.nih.gov/mesh: Observer variation)

3.2.2 Formulation of a plan for the literature search

The literature source was the database Medline® and the search tool used was PubMed® (Entrez retrieval system, NCBI at the NLM, USA). The searches were made with (Tables 3.5 – 3.9) and without MeSH-terms (Tables 3.10 and 3.11). Firstly, the searches with MeSH-terms were performed with MeSH-terms automatically retrieved from the MeSH data base (Tables 3.5 – 3.7) and secondly with the MeSH-terms written by hand (Tables 3.8 and 3.9). Limits are presented in the Tables 3.5 – 3.11 and differed only in that sense that the search included publications with abstracts (Tables 3.5, 3.7, 3.8, and 3.10) or were without abstracts (Tables 3.6, 3.9, and 3.11). In one search (Table 3.7) the limits were activated in the third gate of the search (#3), while in the other searches (Tables 3.5, 3.6, 3.8 – 3.11), the limits were activated after each search term (#1 and #2).

Table 3.5 Search strategy with MeSH-terms automatically retrieved from the MeSH data base, limits activated, and number of retrieved publication. In this search the limit "only items with abstract" was activated

Search term	Retrieved publications (n)
#1 Radiography, Panoramic [MeSH]	1 478
#2 Periodontal Diseases [MeSH]	16 876
#3 #1 AND #2	278

Limits Activated: only items with abstracts, Humans, English, All Adult: 19+ years
Publication Date from 1974/01/01 to 2010/06/30.
Database (PubMed®) search performed on 2010/08/16

Table 3.6 Search strategy with MeSH-terms automatically retrieved from the MeSH data base, limits activated, and number of retrieved publication. In this search the limit "only items with abstract" was not activated

Search term	Retrieved publications (n)
#1 Radiography, Panoramic [MeSH]	1 836
#2 Periodontal Diseases [MeSH]	20 679
#3 #1 AND #2	320

Limits Activated: Humans, English, All Adult: 19+ years,
Publication Date from 1974/01/01 to 2010/06/30.
Database (PubMed®) search performed on August 16, 2010

Table 3.7 Search strategy with MeSH-terms automatically retrieved from the MeSH data base, limits activated, and number of retrieved publication. In this search the limits were activated in gate #3 and also "only items with abstract" was activated

Search term	Retrieved publications (n)
#1 Radiography, Panoramic [MeSH]	4 591
#2 Periodontal Diseases [MeSH]	61 046
#3 #1 AND #2	278

Limits Activated: only items with abstracts, Humans, English, All Adult: 19+ years,
Publication Date from 1974/01/01 to 2010/06/30.
Database (PubMed®) search performed on August 16, 2010

Table 3.8 Search strategy with MeSH-terms written by hand, limits activated, and number of retrieved publication. In this search the limit “only items with abstract“ was activated

Search term	Retrieved publications (n)
#1 Radiography, Panoramic [MeSH]	1 478
#2 Periodontal Diseases [MeSH]	16 876
#3 #1 AND #2	278

Limits Activated: only items with abstracts, Humans, English, All Adult: 19+ years,
Publication Date from 1974/01/01 to 2010/06/30
Database (PubMed®) search performed on August 16, 2010

Table 3.9 Search strategy with MeSH-terms written by hand, limits activated, and number of retrieved publication. In this search the limit “only items with abstract“ was not activated

Search term	Retrieved publications (n)
#1 Radiography, Panoramic [MeSH]	1 836
#2 Periodontal Diseases [MeSH]	20 679
#3 #1 AND #2	320

Limits Activated: Humans, English, All Adult: 19+ years,
Publication Date from 1974/01/01 to 2010/06/30.
Database (PubMed®) search performed on August 16, 2010

Table 3.10 Search strategy with search terms written by hand and not labelled [MeSH], limits activated, and number of retrieved publication. In this search the limit “only items with abstract“ was activated

Search term	Retrieved publications (n)
#1 Radiography, Panoramic	1 706
#2 Periodontal Diseases	17 253
#3 #1 AND #2	347

Limits Activated: only items with abstracts, Humans, English, All Adult: 19+ years,
Publication Date from 1974/01/01 to 2010/06/30.
Database (PubMed®) search performed on August 16, 2010

Table 3.11 Search strategy with search terms written by hand and not labelled [MeSH], limits activated, and number of retrieved publication. In this search the limit "only items with abstract" was not activated

Search term	Retrieved publications (n)
#1 Radiography, Panoramic	2 065
#2 Periodontal Diseases	21 069
#3 #1 AND #2	389

Limits Activated: only items with abstracts, Humans, English, All Adult: 19+ years, *Publication Date* from 1974/01/01 to 2010/06/30.
Database (PubMed®) search performed on August 16, 2010

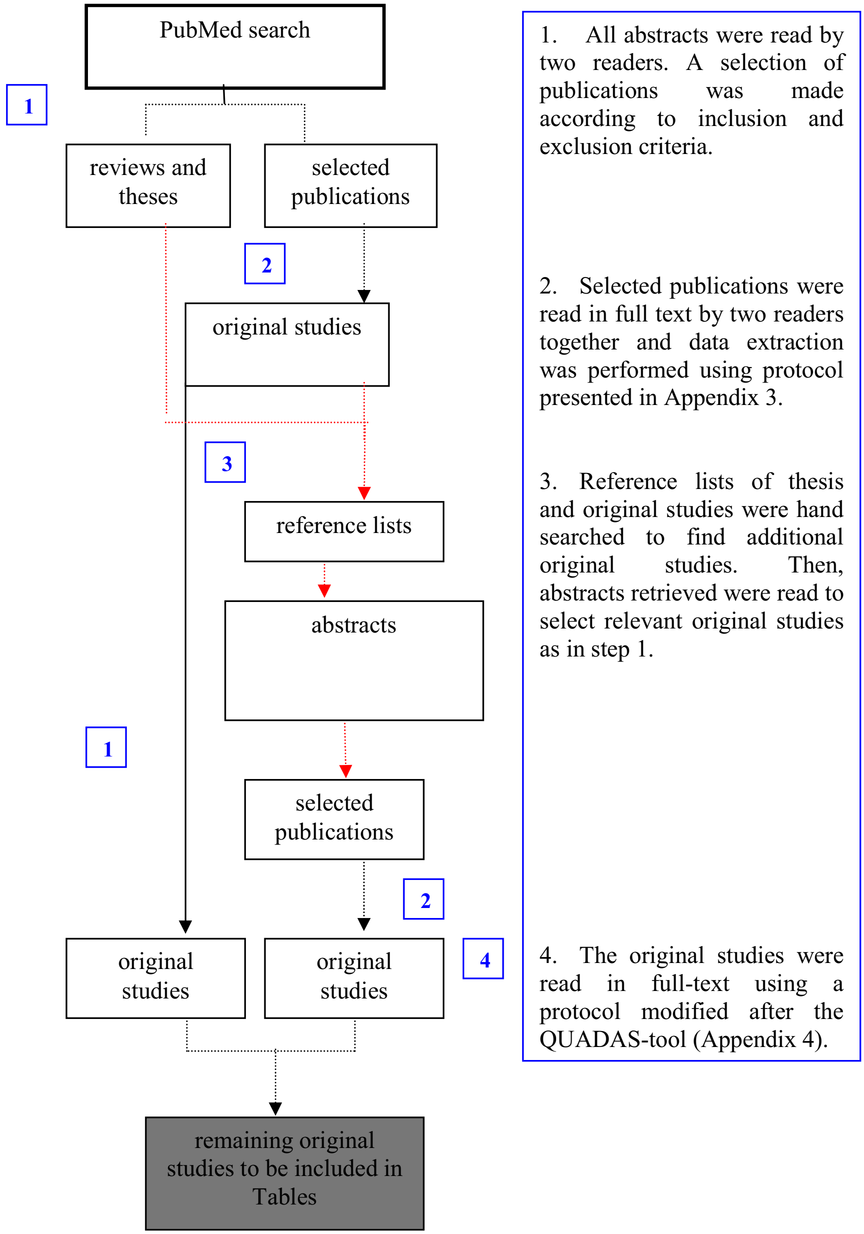
The Cochrane Database of Systematic Reviews (Cochrane Library) was searched using the search term *panoramic radiography.*

3.2.3 Literature search and retrieval of publications

The search in the PubMed database presented in Table 3.5 was selected as the most relevant search strategy to evaluate the evidence of the formulated problem specifications. The process to select retrieved publications is presented in Figure 3.5. Two readers (mean experience of systematic reviews 9.5 years; experience of each reader 4, 15 years) read the titles and abstracts of the retrieved publications. Original studies were included. Studies that reported *(i)* decision making, *(ii)* case reports, and *(iii)* other radiographic methods such as tomography, CBCT, tuned-aperture computed tomography (TACT), CT and MSCT were excluded. Decision to include a publication was made when an abstract was considered by at least one reader to be relevant. If considered relevant, the publication was ordered in full-text and read with the aid of a Protocol *"Inclusion or exclusion of publications in the systematic review"* (Appendix 3) in order to determine inclusion or exclusion of the publication. When at least one reader considered the full-text publication relevant, the publication was included.

The second step of the search was to hand search the reference list of original studies that were found to be relevant in the first step (Figure 3.5). The reference lists of theses and review articles listed in the PubMed-search were also hand searched. Titles were searched that contained *(i)* the term *panoramic radiography* in combination with *(ii)* words suggesting *periodontal diseases* or *image quality*. No publication date limits were specified in this step. Book chapters and reviews were excluded because the focus of the review was original studies. Decision to include publication was made when a selected abstract was considered by at least one of two readers to be relevant. The abstracts were read and when included the full-text publication was ordered and read with the aid of the pre-established protocol *"Inclusion or exclusion of publications in the systematic review"* (Appendix 3) in order to determine inclusion or exclusion of the publication.

Figure 3.5 Flow diagram modified after Ribeiro-Rotta et al. (2007) presenting the selection of publications in different steps of the process



3.2.4 Data extraction, interpretation of data, and evaluation of evidence

Two readers independently read the publications, extracted data and interpreted the data with the aid of Protocol *"Interpretation of publications in the systematic review"* (Appendix 4), which was constructed according to the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool (Whiting et al. 2003). The protocol was designed also on the basis of literature about the critical appraisal of studies on diagnostic methods (Jaeschke et al. 1994a, 1994b). In situations when the interpretations of data were ambiguous, joint discussions between the readers were held.

Evaluation of evidence was based on the study design, individual quality items of the Protocol *"Interpretation of publications in the systematic review"* (Appendix 4), the direction and magnitude of results of the included studies *i.e.* similarity in reported results. Where sufficient data were available, disease prevalence, sensitivity, and specificity of panoramic radiography were calculated as presented in Figure 1.1. Sensitivity and specificity was calculated based on the consensus radiographic standard, which consisted of the combined readings of panoramic and intraoral radiography.

4. RESULTS

4.1 Panoramic radiography and posterior bitewing radiography

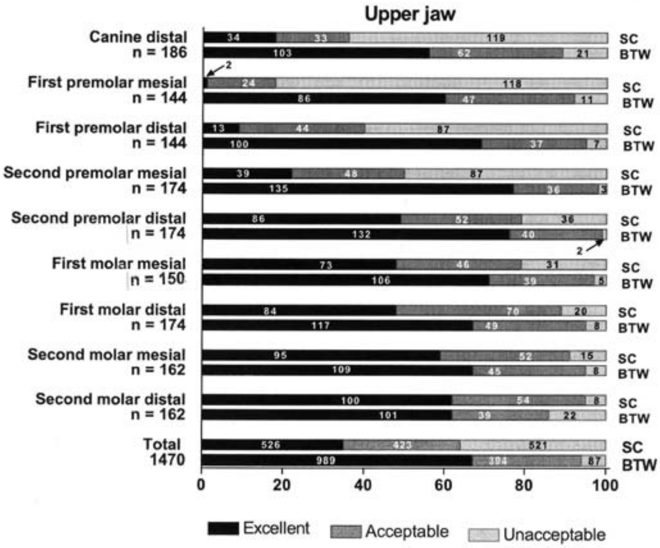
In this part of Results, panoramic radiography using Scanora® dental panoramic program will be referred to as panoramic radiography and posterior bitewing radiography as bitewing radiography.

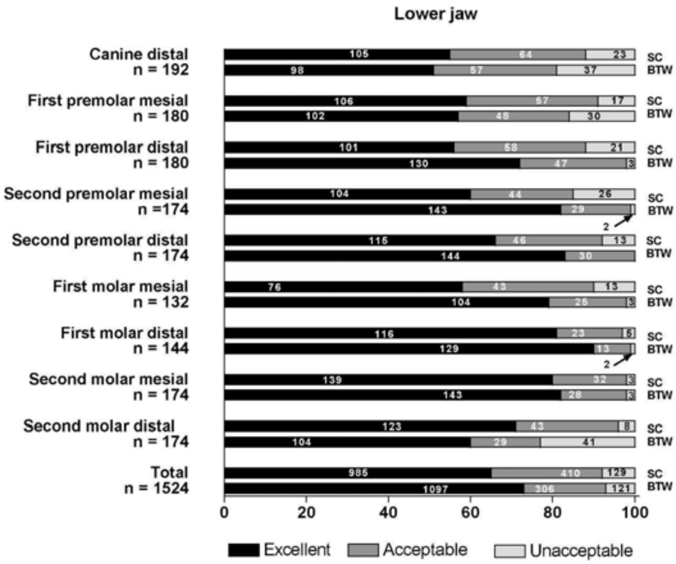
4.1.1 Visual grading analysis

4.1.1.1 Visual grading analysis for assessment of alveolar bone level

Six observers scored 499 sites (245 in the upper jaw and 254 in the lower jaw) resulting in 2 994 assessments (1 470 assessments in the upper jaw and 1 524 assessments in the lower jaw) in each radiographic method. Figure 4.1 and Table 4.1 present the distribution of the categories *excellent, acceptable*, *unacceptable* in panoramic radiography and bitewing radiography.

Figure 4.1 Distribution for visual grading analysis. Categories (excellent, acceptable or unacceptable) for grading alveolar bone level visibility in panoramic radiography taken with the Scanora® dental programme (SC) and in posterior bitewing radiography (BTW) by tooth site for six observers. Numbers of sites per category are presented within the bars





As presented in Table 4.1, 65% of the sites in the upper jaw were categorized as *excellent* and *acceptable* in panoramic radiography as compared to 94% in bitewing radiography. In the lower jaw, the frequency of sites being *excellent* and *acceptable* was similar in panoramic radiography and bitewing radiography (92%).

In panoramic radiography, 36% of assessments in the upper jaw and 65% in the lower jaw were categorized as *excellent* (Table 4.1). Corresponding figures in bitewing radiography were 67% and 72%, respectively. In panoramic radiography, 48%–62% of the assessments in the upper jaw distally to the second premolar were categorized as *excellent*. But the frequency of assessments being *excellent* was low in the distal sites of the canine (18%), the mesial sites of second premolar (22%), and first premolar (1%–9%) regions. In bitewing radiography, assessments categorized as *excellent* ranged between 56%–77% in the upper jaw. The frequency of assessments categorized as *excellent* was similar for the radiographic methods for the distal site of the second molar (62%). In the lower jaw, 55%–81% of the assessments were categorized as *excellent* in panoramic radiography with the highest frequency for the distal site of the first molar (81%) and the mesial site of the second molar (80%). Corresponding figures in bitewing radiography were 51%–90% with the highest frequency for the distal site of the first molar (90%). The frequency of assessments categorized as *excellent* was comparable in the radiographic methods in the distal site of the canine, the mesial site of the first premolar and the mesial site of the second molar in the lower jaw.

As Table 4.1 presents, 29% of assessments in the upper jaw and 27% in the lower jaw were categorized as *acceptable* in panoramic radiography. Corresponding figures in bitewing radiography were 27% and 20%, respectively. Thus, the overall results on sites in the upper and lower jaws categorized as *acceptable* in the radiographic methods were comparable. In panoramic radiography, sites the of the upper jaw *i.e* the distal site of the second molar to the distal site of the first premolar, 28–40% of the assessments were categorized as *acceptable* with a higher frequency compared to bitewing radiography (21-28%). But the frequency of assessments categorized as *acceptable* was low in panoramic radiography in the mesial site of the first premolar (17%) and in the distal site of the canine (18%). In corresponding sites, bitewing radiography was assessed acceptable in 32% and 33%, respectively. In the lower jaw, the assessments categorized as *acceptable* were higher in panoramic radiography and ranged between 16-33%. Corresponding figures in bitewing radiography were 9-30%.

In panoramic radiography, 35% of assessments in the upper jaw and 8% in the lower jaw were categorized as *unacceptable* (Table 4.1). Corresponding figures in bitewing radiography were 6% in the upper jaw and 8% in the lower jaw. Thus, the overall results for sites in the lower jaw categorized as *unacceptable* were similar for the radiographic methods. Most sites in the upper jaw that were categorized as *unacceptable* in panoramic radiography occurred in the distal site of the canines (64%), in the sites of the first premolar, the mesial site (82%) and the distal site (60%), respectively and the mesial site of the second premolar (50%). In the lower jaw, the frequency of *unacceptable* sites in panoramic radiography was low in the distal site of the first molar (3%), the mesial site of the second molar (2%) and comparable with bitewing radiography. Bitewing radiography presented higher frequency of assessments categorized as *unacceptable* in the distal site of the canine (19% versus 12%) and in the mesial site of the first premolar (17% versus 9%). Fewer sites, *i.e.* from the distal site of the first premolar to the mesial site of the first molar, were categorized as *unacceptable* in bitewing radiography (0-2%) than in panoramic radiography (7-15%). While, in bitewing radiography the frequency of unacceptable sites for the second molar distally was higher (23%) than in panoramic radiography (4%).

Table 4.1 Visual grading analysis. Categories (excellent, acceptable or unacceptable) for grading alveolar bone level visibility in panoramic radiography using Scanora® dental panoramic program (Scanora) and in posterior bitewing radiography (Bitewing) expressed as percent (%) of total number (n) of assessments made by six observers. E – excellent; A – acceptable; U – unacceptable

Upper jaw								***Lower jaw***						
	Scanora %			**Bitewing %**				**Scanora %**			**Bitewing %**			
n	**E**	**A**	**U**	**E**	**A**	**U**	***Tooth site***	**E**	**A**	**U**	**E**	**A**	**U**	**n**
186	18	18	64	56	33	11	*Canine* Distal	55	33	12	51	30	19	192
144	1	17	82	60	32	8	First premolar Mesial	59	32	9	57	26	17	180
144	9	31	60	69	26	5	First premolar Distal	56	32	12	72	26	2	180
174	22	28	50	77	21	2	Second premolar Mesial	60	25	15	82	17	1	174
174	49	30	21	76	23	1	Second premolar Distal	66	27	7	83	17	0	174
150	49	30	21	71	26	3	First molar Mesial	58	32	10	79	19	2	132
174	48	40	12	67	28	5	First molar Distal	81	16	3	90	9	1	144
162	59	32	9	67	28	5	Second molar Mesial	80	18	2	82	16	2	174
162	62	33	5	62	24	14	Second molar Distal	71	25	4	60	17	23	174
1470	**36**	**29**	**35**	**67**	**27**	**6**	**Total**	**65**	**27**	**8**	**72**	**20**	**8**	**1524**

4.1.1.2 Visual grading analysis for detection of vertical bone defects and furcation involvement

Five observers categorized the visibility for the detection of *vertical bone defect* in 2 885 teeth sites (1 435 sites in the upper and 1 450 sites in the lower jaw) as acceptable or unacceptable resulting in 7 175 assessment in the upper jaw and 7 250 assessments in the lower jaw in each radiographic method. Table 4.2 presents the frequencies of categories for grading vertical bone defect visibility in different tooth groups. In the upper jaw, two-thirds (65%) of the sites were categorized as *acceptable* in both radiographic methods. In panoramic radiography, one-third of the upper jaw sites were categorized as *unacceptable*, with the highest frequency for the mesial site of the first premolar (66%) and the distal site of the canine (58%). Overall, few sites in bitewing radiography were categorized as *unacceptable*. In the lower jaw, most sites were categorized as *acceptable* in both radiographic methods (93%). The frequency of sites in the lower jaw categorized as *unacceptable* ranged between 0.3% (the distal site of the second molar) and 9% (the mesial site of the first molar) in panoramic radiography. Corresponding figures in bitewing radiography were 0% (the mesial site of the second molar and the distal site of the second premolar) and 6% (the distal site of the canine).

Five observes categorized the visibility for the *detection of furcation involvement* in 748 teeth (467 in the upper jaw and 281 in the lower jaw) as *acceptable* or *unacceptable* resulting in 3740 assessment in each radiographic method. For the detection of furcation involvement, most teeth were categorized as *acceptable* in both radiographic methods. Only 59 assessments (2%) were categorized as *unacceptable*. Of these, 22 assessments were on panoramic radiography and 37 in bitewing radiography. One observer made 37 of the 59 unacceptable ratings.

Table 4.2 Visual grading analysis. Categories (acceptable or unacceptable) for grading vertical bone defect visibility in panoramic radiography using Scanora® dental panoramic program (SC) and in posterior bitewing radiography (BTW), and sites assessed in both methods (SC+BTW) expressed as percent (%) of total number (n) of assessment made by five observers: A – acceptable; U – unacceptable

	Upper jaw					*Lower jaw*				
	A	U				A	U			
	SC + BTW	*SC*	*BTW*	*SC + BTW*		*SC + BTW*	*SC*	*BTW*	*SC + BTW*	
n	%	%	%	%	Tooth site	%	%	%	%	n
935	38	58	1	3	Canine Distal	88	5	6	1	950
820	32	66	0.4	2	First premolar Mesial	89	5	5	1	900
820	52	47	0.4	1	First premolar Distal	92	7	0.4	0.3	900
785	50	48	1	1	Second premolar Mesial	92	7	0.2	0.4	845
785	74	25	1	0.1	Second premolar Distal	93	7	0	0.1	845
765	75	25	0.4	0.1	First molar Mesial	91	9	0.2	0.2	645
765	89	9	1	1	First molar Distal	99	1	0.2	0	645
750	90	8	1	1	Second molar Mesial	99	1	0	0	760
750	89	5	6	0.4	Second molar Distal	6	0.3	4	0	760
7175	**65**	**32**	**1**	**1**	**Total**	**93**	**5**	**2**	**0.4**	**7250**

4.1.1.3 Observer performance of visual grading analysis in the assessment of the alveolar bone loss

The hypothesis was verified as the observers varied both within themselves and between each other in their analyses of visibility of images. The intra-observer agreement was higher than the inter-observer agreement.

4.1.1.3.1 Intra-observer performance

Table 4.3 presents the intra-observer agreement for visual grading analysis in the assessment of the *alveolar bone loss by scores*. The overall intra-observer agreement of three observes was comparable for panoramic radiography (range 76–87%) and bitewing radiography (range 78–91%). Expressed as mean kappa value of three observers, intra-observer agreement for panoramic radiography (0.66) presented substantial agreement whilst bitewing radiography (0.49) presented moderate agreement. Two observers presented high intra-observer agreement when expressed in percentage, but the kappa values presented only moderate agreement for bitewing radiography.

Table 4.3 Intra-observer agreement for visual grading analysis (categories: excellent, acceptable or unacceptable) for alveolar bone level visibility in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). Values expressed in overall agreement (%) and κ – Cohen's kappa

Observer	**Scanora**		**Bitewing**	
	%	κ	%	κ
1	76	0.59	78	0.56
2	87	0.75	89	0.49
3	82	0.63	91	0.42
Mean	**82**	**0.66**	**86**	**0.49**

Overall agreement of three observers for the visual grading analysis for detection of *vertical bone defects* (Table 4.4) was high for both methods. The mean of three observers, intra-observer agreement in panoramic radiography was 94% and 98% in bitewing radiography. Mean kappa value for panoramic radiography, 0.83 (almost perfect agreement), was higher than that for bitewing radiography, 0.43 (moderate agreement).

Table 4.4. Intra-observer agreement for visual grading analysis categories (acceptable or unacceptable) for vertical bone defect detection visibility in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). overall agreement (%) and κ – Cohen's kappa

Observer	**Scanora**		**Bitewing**	
	%	*κ*	*%*	*κ*
2	93	0.79	98	0.34
3	93	0.81	98	0.62
4	95	0.88	98	0.32
Mean	**95**	**0.83**	**98**	**0.43**

Overall agreement of three observers for visual grading analysis for detection of *furcation involvement* detection was high for both methods (93%–98%).

4.1.1.3.2 Inter-observer performance

There was a large variation in the number of sites assessed by the six observers. Table 4.5 presents inter-observer agreement for several observers. Inter-observer agreement for visual grading analysis in the assessment of the alveolar bone level for six observers was moderate for panoramic radiography (κ = 0.45) and fair for bitewing radiography (κ = 0.28). Kappa values for both methods 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).

Table 4.5 Inter-observer agreement of several observers for visual grading analysis in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). Six observers graded alveolar bone level visibility (categories: excellent, acceptable or unacceptable). Five observers graded vertical bone defect visibility (categories: acceptable or unacceptable). K – Fleiss' kappa

	Alveolar bone level		*Vertical bone defect*	
	Scanora	**Bitewing**	**Scanora**	**Bitewing**
	K	K	K	κ
Several observers	0.45	0.28	0.62	0.25

Overall inter-observer agreement for *alveolar bone level* visibility for pairs of observers (Table 4.6) was comparable for panoramic (43–90%) and bitewing radiography (45–93%), while weighted kappa values were higher for panoramic radiography.

Weighted kappa values for pairs of observers varied substantially, 0.31-0.63 for panoramic radiography and 0.16-0.43 for bitewing radiography. Kappa values of pairs of observers (Table 4.6) ranged between 0.19 and 0.53 for panoramic radiography and 0.09 and 0.39 for bitewing radiography. The lowest weighted kappa and kappa values for pairs of observers often included observer 6, who used the category acceptable (panoramic and bitewing radiography) and unacceptable (panoramic radiography) more frequently than the other observers. For the other five observers, the lowest weighted kappa values for pair of observers was 0.47 for panoramic radiography and 0.23 for bitewing radiography.

Table 4.6 Inter-observer agreement of pairs of observers for visual grading analysis in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). Six observers graded alveolar bone level visibility (categories: excellent, acceptable, or unacceptable). Agreement was expressed as overall agreement (%), Cohen's weighted kappa value (κw), and Cohen's kappa value (κ)

Pair of observers	*Alveolar bone level*					
	Scanora			**Bitewing**		
	%	κw	κ	%	κw	κ
1/2	89	0.61	0.53	89	0.41	0.38
1/3	90	0.59	0.50	90	0.32	0.27
1/4	85	0.53	0.45	85	0.43	0.39
1/5	62	0.47	0.37	65	0.32	0.27
1/6	43	0.33	0.19	55	0.32	0.24
2/3	91	0.63	0.52	93	0.40	0.35
2/4	87	0.61	0.53	85	0.36	0.33
2/5	65	0.55	0.44	74	0.35	0.27
2/6	46	0.39	0.26	45	0.17	0.09
3/4	85	0.51	0.39	84	0.23	0.18
3/5	67	0.56	0.47	72	0.27	0.19
3/6	44	0.35	0.23	45	0.16	0.09
4/5	65	0.56	0.44	66	0.36	0.29
4/6	52	0.45	0.33	54	0.34	0.24
5/6	45	0.31	0.20	51	0.24	0.18

Weighted kappa values for pairs of observers varied substantially, 0.31-0.63 for panoramic and 0.16-0.43 for bitewing radiography. Kappa values of pairs of observers (Table 4.6) ranged between 0.19 and 0.53 for panoramic and 0.09 and 0.39 for bitewing radiography. The lowest weighted kappa and kappa values for pairs of observers often included observer 6, who used the category acceptable more frequently than the other observers.

Inter-observer agreement of five observers for *vertical bone defect* visibility was substantial for panoramic radiography (κ = 0.62) and fair for bitewing (κ = 0.25) radiography (Table 4.5). Table 4.7 presents the inter-observer agreement of pairs of observers. Overall agreement of pairs of observers was high for panoramic radiography (range 83–91%) and very high for bitewing radiography (range 96–98%). Corresponding weighted kappa values were moderate or substantial for panoramic radiography (range 0.51–0.76) but poor or slight for bitewing radiography (range 0.11–0.38).

Table 4.7 Inter-observer agreement of pairs of observers for visual grading analysis in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). Five observers graded vertical bone defects visibility (categories: acceptable or unacceptable). Agreement was expressed as overall agreement (%), and Cohen's kappa value (κ)

Pairs of observers	*Vertical bone defect*			
	Scanora		**Bitewing**	
	%	κ	%	κ
2/3	91	0.73	97	0.19
2/4	88	0.67	98	0.13
2/5	90	0.60	97	0.11
2/6	90	0.63	96	0.15
3/4	91	0.76	97	0.27
3/5	86	0.53	96	0.26
3/6	88	0.62	96	0.40
4/5	83	0.49	97	0.11
4/6	86	0.60	96	0.23
5/6	88	0.51	96	0.38

Inter-observer agreement of five observers for *furcation involvement* visibility was 97% (κ = 0.02) for panoramic radiography and 96% (κ = 0.06) for bitewing radiography. Agreement for pairs of observers was high for panoramic radiography (range 97–98%) and bitewing (range 96–100%) radiography (Table 4.8). However, the kappa values were very low for both methods.

Table 4.8 Inter-observer agreement of pairs of observers for visual grading analysis in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). Five observers graded furcation involvement visibility (categories: acceptable or unacceptable). Agreement was expressed as overall agreement (%), and Cohen's kappa value (κ)

Pair of observers	*Furcation involvement*			
	Scanora		**Bitewing**	
	%	κ	%	κ
2/3	97	-0.0123	96	-0.0026
2/4	99	0	96	0.06
2/5	99	0	96	0.05
2/6	99	0.2	96	0.1
3/4	98	0	100	-0.0013
3/5	98	0	100	-0.0018
3/6	100	-0.0047	99	-0.002
4/5	to few rating categories		100	-0.0018
4/6	100	0	99	-0.002
5/6	100	0	100	-0.4

4.1.2 Assessment of the alveolar bone loss

4.1.2.1 Assessment of alveolar bone level

Out of 2 994 possible assessments (499 available sites x 6 observers), 647 in the upper jaw and 374 assessments in the lower jaw were categorized as unacceptable in visual grading analysis in one or both radiographic methods. Remaining assessments (823 in the upper and 1150 in the lower jaw) that were categorized excellent or acceptable proceeded for assessment of alveolar bone level by score.

Table 4.9 presents the assessment of alveolar bone level by score. There was an agreement between panoramic and bitewing radiography in 56% of the sites of the upper jaw and 58% of the sites of the lower jaw. The highest agreement (60-68%)

was found in the premolar region of the upper jaw and in the molar region of the lower jaw (55-66%). The mean agreement between the two methods, if a difference of one score was allowed, was 95% of the sites in the upper jaw and 94 % in the lower jaw (range: 88-100%). Agreement between panoramic radiography and bitewing radiography for alveolar bone level assessment by score was not influenced by the categories used in visual grading analysis, *excellent* or *acceptable*. Thus, the agreement was 57% ($\kappa = 0.32$) independently whether both methods in the site pair were categorized as *excellent* or when one of the methods in the site pair was categorized as *acceptable*.

As presented in Table 4.9 disagreement between panoramic and bitewing radiography was found in two ways. Panoramic radiography presented a lower score (less bone loss) as compared to bitewing radiography in 16% of the sites of the upper jaw and in 17% of the sites of the lower jaw. On the other hand, larger bone loss was found in panoramic radiography in more sites (28% of the sites of the upper jaw and 25% of the sites in the lower jaw). Disagreement between the radiographic methods was found in the upper jaw particularly in the distal site of canine (56%) and in the molar sites (46-48%). Panoramic radiography presented higher scores in this region except for the mesial site of the first molar. In the lower jaw, the disagreement was found particularly in the distal site of the canine (49%) and the mesial site of the first premolar (49%), where bitewing radiography presented higher scores for the distal site of the canine and panoramic radiography for the mesial site of the first premolar (Table 4.9).

Table 4.9 Assessment of the alveolar bone level by score. Agreement and disagreement between panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing) by tooth site. Values were expressed as percent (%) of number of sites (n), which were assessed in both methods by six observers altogether

		Agreement %	*Disagreement %*	
Tooth site	**n**	**Scanora equal to Bitewing**	**Less bone loss in Scanora**	**Larger bone loss in Scanora**
		Upper jaw		
13/23 d	50	44	14	42
14/24 m	24	62	21	17
14/24 d	50	60	20	20
15/25 m	80	68	16	16
15/25 d	118	64	20	16
16/26 m	100	52	30	18
16/26 d	138	54	11	35
17/27 m	133	53	10	37
17/27 d	130	54	13	33
Total	**823**			
Mean		**56**	**16**	**28**
		Lower jaw		
33/43 d	93	50	35	14
34/44 m	131	50	8	41
34/44 d	143	66	23	11
35/45 m	128	54	17	29
35/45 d	140	55	21	24
36/46 m	104	63	15	22
36/46 d	126	66	18	16
37/47 m	159	60	10	30
37/47 d	126	55	11	34
Total	**1150**			
Mean		**58**	**17**	**25**

4.1.2.2 Detection of vertical bone defect and furcation involvement

Out of 14 425 possible assessments for the detection of *vertical bone defect* (2 885 available sites x 5 observers) 2 558 assessments in the upper and 529 in the lower jaw were categorized as *unacceptable* in visual grading analysis in one or both radiographic methods. Remaining assessments (4614 in the upper jaw and 6 724 in the lower jaw) that were categorized as *acceptabl*e proceeded for detection of vertical bone defect.

Table 4.10 presents the results of the assessment for the detection of vertical bone defect. There was an agreement between panoramic and bitewing radiography in 89% of sites of the upper jaw and in 94% of the sites in the lower jaw. The agreement between the methods ranged between 86% (for the mesial site of the second molar) and 94% (for the distal site of the canine) in the upper jaw. In the lower jaw, the agreement ranged between 89% (for the mesial site of the second molar) and 97% (for the distal site of the canine) in lower jaw. The majority of agreements were recordings of no vertical bone defect in either radiographic method. The agreement of a positive finding, *i.e.* a vertical bone defect, in the upper jaw ranged between 0.2% (for the distal site of the canine) and 6 % (for the mesial sites of the premolars and the second molar). In the lower jaw, the highest agreement of the presence of a vertical bone defect (6%) between the methods was found in the mesial site of the second molar. A vertical bone defect was detected in only panoramic radiography in average 5% of the assessments (range: 1-11%). The corresponding figure for the bitewing radiography was 3% (range: 1-10%). In the molar region in particular, more vertical bone defects were visible in panoramic radiography. The sensitivity to detect a vertical bone defect was 0.73 in panoramic radiography and 0.53 in bitewing radiography.

Table 4.10 Detection of vertical bone defect. Agreement and disagreement between panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing) by tooth site. Values are expressed as percent (%) of number of sites (n), which were assessed in both methods by five observers altogether

		Agreement %		*Disagreement %*	
Tooth site	*n*	*No defect* **Scanora = Bitewing**	*Defect* **Scanora = Bitewing**	*Defect visible only in* **Scanora**	**Bitewing**
		Upper jaw			
13/23 d	52	94	0.2	5	1
14/24 m	262	81	6	3	10
14/24 d	426	92	1	4	3
15/25 m	392	84	6	5	5
15/25 d	581	89	2	5	4
16/26 m	573	83	3	8	6
16/26 d	684	82	5	11	2
17/27 m	679	80	6	10	4
17/27 d	665	86	3	7	4
Total	**4614**				
Mean		**85**	**4**	**7**	**4**
		Lower jaw			
33/43 d	832	96	1	2	1
34/44 m	800	94	2	2	2
34/44 d	830	95	2	1	2
35/45 m	778	89	2	7	2
35/45 d	783	93	1	3	3
36/46 m	584	87	3	7	3
36/46 d	637	90	1	6	3
37/47 m	754	83	6	7	4
37/47 d	726	85	5	7	3
Total	**6724**				
Mean		**91**	**3**	**4**	**2**

Out of 3 740 possible assessments for the detection of *furcation involvement* (748 available teeth x 5 observers) 59 assessments were categorized as unacceptable in visual grading analysis in one or both radiographic methods. Remaining assessments (2 277 in the upper jaw and 1 404 in the lower jaw) that were categorized as acceptable proceeded for the detection of furcation involvement.

Table 4.11 Detection of furcation involvement. Agreement and disagreement between panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing) by tooth. Values are expressed as percent (%) of number teeth (n), which were assessed in both radiographic methods by five observers altogether

		Agreement %		*Disagreement %*	
Tooth	*n*	*No furcation involvement*	*Furcation involvement*	*Furcation involvement visible only in*	
		Scanora = Bitewing	**Scanora = Bitewing**	**Scanora**	**Bitewing**
		Upper jaw			
14/24	800	96	1	2	1
16/26	751	83	10	4	3
17/27	726	89	5	5	1
Total	**2277**				
Mean		**89**	**5**	**4**	**2**
		Lower jaw			
36/46	645	75	12	5	8
37/47	759	87	7	5	1
Total	**1413**				
Mean		**81**	**10**	**5**	**4**

Table 4.11 presents the results of the detection of furcation involvement. The overall agreement of the panoramic and bitewing radiography for the upper jaw was 94% and for the lower jaws 91%. The agreement between the radiographic methods ranged between 87% for the first molar of the lower jaw and 97% for the first premolar of the upper jaw. The majority of agreement involved recordings of no furcation involvement. The agreement of positive findings *i.e.* furcation

involvements ranged between 1% for the first premolar of upper jaw and 12% for the first molar of lower jaw. In 4% (range: 2-5%) of the teeth, a furcation involvement was only detected in panoramic radiography. The corresponding figures for bitewing radiography were 3% of the teeth (range: 1-8%). Sensitivity was 0.80 for panoramic radiography and 0.70 for bitewing radiography.

4.1.2.3 Observer performance for assessment of alveolar bone loss

4.1.2.3.1 Intra-observer performance

Table 4.12 presents the intra-observer agreement for the assessment of *alveolar bone level* by score in panoramic and bitewing radiography. The mean overall agreement of three observers in panoramic radiography was 67% (72%, 69%, 59%) and in bitewing radiography 66% (76%, 66%, 55%). Corresponding kappa values for panoramic radiography was 0.43 (0.46, 0.45, 0.37) and for bitewing radiography 0.43 (0.55, 0.43, 0.31.) The observer, who had the highest value for bitewing radiography, also presented the highest kappa value for panoramic radiography. Correspondingly, the same observer presented the lowest kappa values in both methods.

Table 4.13 presents intra-observer performance for the detection of *vertical bone defect*. The mean overall agreement of three observers was 95% in panoramic radiography as well as in bitewing radiography. Corresponding kappa values were also identical for panoramic and bitewing radiography (0.58). The observers with the lowest and highest kappa values differed between the methods.

Table 4.12 Intra-observer agreement for assessment of the alveolar bone level by score in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). Values expressed in overall agreement (%) and kappa value (κ)

Observer	Scanora		Bitewing	
	%	κ	%	κ
1	72	0.46	76	0.55
2	69	0.45	66	0.43
3	59	0.37	55	0.31
Mean	**67**	**0.43**	**66**	**0.43**

Table 4.13. Intra-observer agreement for detection of vertical bone defect in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). Values expressed in overall agreement (%) and kappa value (κ)

Observer	Scanora		Bitewing	
	%	κ	%	κ
2	93	0.57	94	0.62
3	92	0.62	92	0.52
4	99	0.56	99	0.61
Mean	**95**	**0.58**	**95**	**0.58**

Table 4.14 presents intra-observer performance for the detection of *furcation involvement*. The mean overall agreement of three observers was 96% in panoramic radiography and 97% in bitewing radiography. Corresponding kappa values were 0.72 for panoramic and 0.76 for bitewing radiography. The observers with the lowest and highest kappa values differed between the methods.

Table 4.14 Intra-observer agreement for detection of furcation involvement in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). Values expressed in overall agreement (%) and kappa value (κ)

Observer	**Scanora**		**Bitewing**	
	%	κ	%	κ
2	98	0.79	97	0.72
3	93	0.65	96	0.77
4	97	0.73	97	0.78
Mean	**96**	**0.72**	**97**	**0.76**

4.1.2.3.2 Inter-observer performance

Tables 4.15 and 4.16 present the inter-observer agreement for the assessment of *alveolar bone level*. The kappa values of six observers for panoramic radiography (0.28) were comparable with that of bitewing radiography (0.29), which can be interpreted as fair agreement. Inter-observer agreement of pair of observers varied widely (Table 4.16). Agreement ranged between 0.08 (slight agreement) and 0.58 (moderate agreement) in panoramic radiography and 0.04 (slight agreement) and 0.43 (moderate agreement) in bitewing radiography.

Table 4.15 Inter-observer agreement of several observers for assessment of alveolar bone in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing) expressed as kappa value (κ) for several observers

	Alveolar bone level	Vertical bone defect	Furcation involvement
Number of observers	6	5	5
	κ	κ	κ
Scanora	0.28	0.35	0.53
Bitewing	0.29	0.38	0.65

Inter-observer agreement for the detection of *vertical bone defect* is presented in Tables 4.15 and 4.16. Kappa values of five observers were comparable for the methods, 0.38 (fair agreement) for panoramic radiography and 0.35 (fair agreement) for bitewing radiography (Table 4.15). Inter-observer agreement of pair of observers (Table 4.16) ranged between 0.22 (fair agreement) and 0.47 (moderate agreement) in panoramic radiography and 0.27 (fair agreement) and 0.44 (moderate agreement) in bitewing radiography.

Table 4.16 Interobserver agreement of pair of observers for assessment of alveolar bone loss in panoramic radiography using Scanora® dental panoramic program (Scanora) and posterior bitewing radiography (Bitewing). Six observers scored the alveolar bone level and five observers assessed vertical bone defect and furcation involvement. Values expressed in kappa value (κ)

	Alveolar bone level		*Vertical bone defect*		*Furcation involvement*	
Pair of observers	**Scanora**	**Bitewing**	**Scanora**	**Bitewing**	**Scanora**	**Bitewing**
	κ	κ	κ	κ	κ	κ
1/2	0.31	0.31				
1/3	0.23	0.34				
1/4	0.38	0.39				
1/5	0.40	0.31				
1/6	0.25	0.21				
2/3	0.17	0.37	0.42	0.42	0.67	0.70
2/4	0.33	0.38	0.31	0.27	0.77	0.65
2/5	0.45	0.38	0.34	0.42	0.65	0.71
2/6	0.31	0.04	0.43	0.44	0.39	0.54
3/4	0.08	0.43	0.25	0.29	0.68	0.71
3/5	0.58	0.40	0.40	0.43	0.69	0.76
3/6	0.36	0.22	0.47	0.44	0.42	0.59
4/5	0.20	0.41	0.22	0.27	0.62	0.72
4/6	0.38	0.14	0.37	0.42	0.33	0.54
5/6	0.14	0.04	0.37	0.37	0.57	0.67

Inter-observer agreement for the detection of *furcation involvement* is presented in Tables 4.15 and 4.16 kappa values of five observers were 0.53 (moderate agreement) for panoramic radiography and 0.65 (substantial agreement) for bitewing radiography (Table 4.15). Inter-observer agreement of pair of observers (Table 4.16) ranged between 0.33 (fair agreement) and 0.77 (substantial agreement) in panoramic radiography and 0.54 (moderate agreement) and 0.76 (substantial agreement) in bitewing radiography.

4.2 Evaluation of evidence on panoramic radiography for the assessment of alveolar bone loss

4.2.1 Literature search and retrieval of publications

In the first step, the PubMed search yielded 278 titles and abstracts on panoramic radiography and alveolar bone (Table 3.5). No review was identified in the Cochrane Library. Figure 4.2 depicts a flow diagram of the selection process for studies relevant to the systematic review. Reading of 278 titles and abstracts resulted in the inclusion of 25 publications (24 original studies and one thesis), which were ordered in full-text. The agreement between the two readers for all retrieved abstracts was 98% (272/278) and for the included publications 76% (19/25). After, the 25 publications were read by the two readers together with the aid of the protocol *"Inclusion or exclusion of publications in the systematic review"* (Appendix 3), 19 publications remained (Figure 4.2). The reasons for exclusion are presented in Table 4.17.

The second step of the search, *i.e*. the hand search of the reference lists of the 19 included original studies and one thesis, yielded 12 additional titles of publications (Figure 4.2). The abstracts of 8 publications were found and 3 publications which did not present any abstract were ordered in full-text, 1 publication was not retrieved. The agreement between the two readers for the selected abstracts was 67% (6/9). After 11 publications were read by the two readers together with the aid of the protocol *"Inclusion or exclusion of publications in the systematic review"* (Appendix 3), 7 publications remained.

All remaining 26 full-text original studies (Figure 4.2) were read with the aid of the protocol based on the QUADAS-tool *"Interpretation of publications in the systematic review"*. Included 17 publications are listed in Tables 4.18-4.22 and 9 excluded publications with reasons for exclusion in Table 4.17.

Figure 4.2 Flow diagram modified after Ribeiro-Rotta et al. (2007) presenting the selection process and number of included publications in different steps of the process

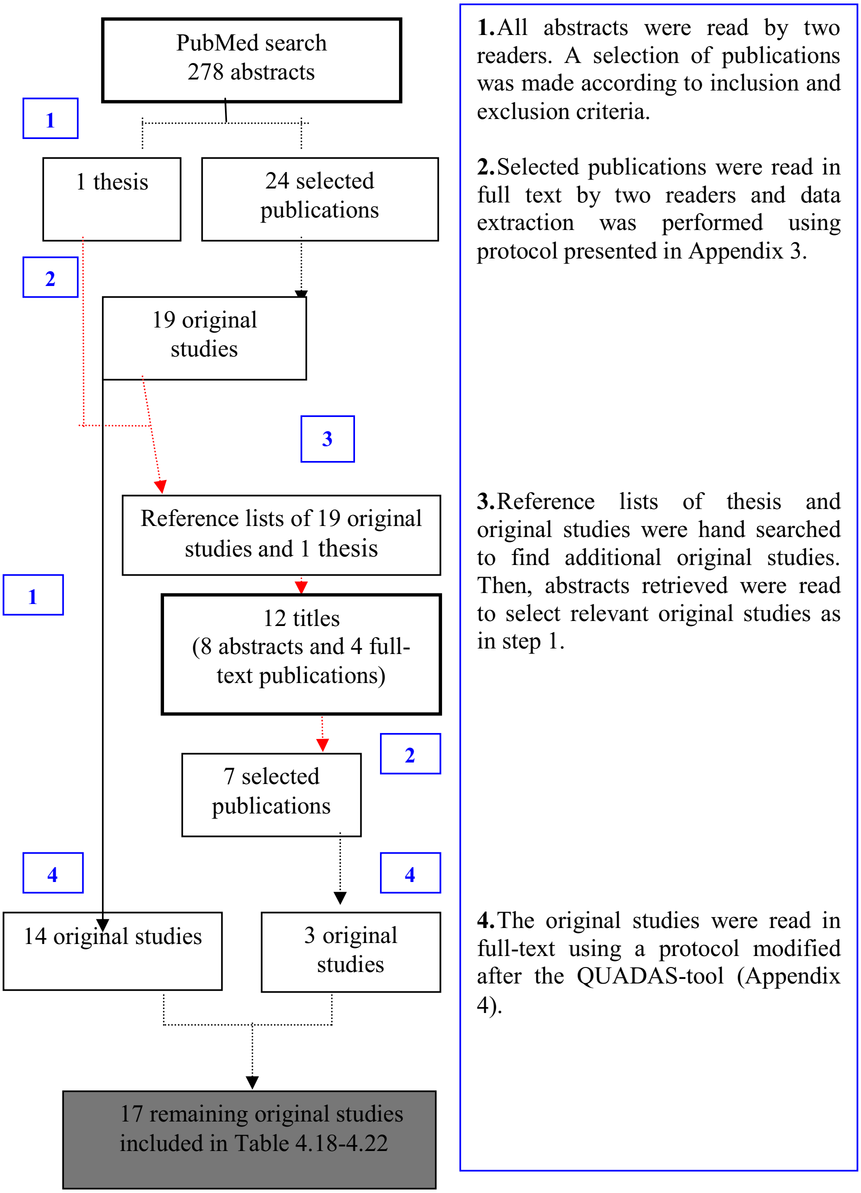


Table 4.17 Publications read in full-text and reasons for exclusion of publications. References in grey fields were included and presented in Tables 4.18 – 4.22

No.	*Publications retrieved in PubMed search and read in full-text 1st author, year*	*Reason for exclusion according to protocol "Inclusion or exclusion of publications in the systematic review (Appendix 2)*	*Reason for exclusion according to protocol based on the QUADAS-tool "Interpretation of publications in the systematic review" (Appendix 4)*
1.	Ivanauskaite et al. 2008	-	-
2.	Ivanauskaite et al. 2006	-	-
3.	Jenkins et al. 2005	-	-
4.	Zechner et al. 2003	-	No information on observers; unclear analysis
5.	Persson et al. 2003	-	-
6.	Rushton et al. 2002b	Panoramic radiography not compared to other diagnostic methods; study of selection criteria	-
7.	Rushton et al. 2002a	Panoramic radiography not compared to other diagnostic methods; study of selection criteria	-
8.	Pepelassi et al. 2000	-	-
9.	Flint et al. 1998	-	Findings in marginal and periapical bone could not be analysed separately; number of teeth/sites not presented
10.	Pepelassi & Diamanto-Kipioti 1997	-	-
11.	Bolin et al. 1996	Analysis of alveolar bone height not of alveolar bone loss	
12.	Molander et al. 1995a	-	-
13.	Molander et al. 1995b	-	-
14.	Åkesson et al. 1992b (in PubMed 1993)	-	-
15.	Schulte et al. 1992	-	Evaluation of Periotest method not of panoramic radiography
16.	Åkesson et al. 1992a	-	-
17.	Molander et al. 1991	-	-
18.	Åkesson 1991	No original study (Thesis)	
19.	Soikkonen et al. 1990	Study on cadaver	-

No.	*Publications retrieved in PubMed search and read in full-text 1st author, year*	*Reason for exclusion according to protocol "Inclusion or exclusion of publications in the systematic review (Appendix 2)*	*Reason for exclusion according to protocol based on the QUADAS-tool "Interpretation of publications in the systematic review" (Appendix 4)*
20.	Åkesson et al. 1989b	-	-
21.	Akesson et. al 1989a	-	-
22.	Kaimenyi & Ashley 1988	-	-
23.	Douglass et al. 1986	-	Panoramic radiography not described
24.	Valachovic et al. 1986	-	Sample and panoramic radiography not described
25.	Ehrlich et al. 1977	No separate analysis of alveolar bone loss	-
No.	*Publications retrieved in hand search and read in full-text 1st author, year*	*Reason for exclusion according to protocol "Inclusion or exclusion of publications in the systematic review (Appendix 2)*	*Reason for exclusion according to protocol based on the QUADAS-tool "Interpretation of publications in the systematic review" (Appendix 4)*
1.	Molander et al. 2004	-	-
2.	Dannewitz et al. 2002	No separate analysis of alveolar bone loss	-
3.	Kaeppler et al. 2000a	No analysis of alveolar bone loss	-
4.	Walsh et al. 1997	-	Number of sites not presented
5.	Rohlin et al. 1989	-	-
6.	Adriaens et al. 1992	-	Analytical methods inadequately described
7.	Stenström et al. 1982	-	-
8.	Gröndahl et al. 1971	-	Number of sites not presented
9.	Pfeifer & Dean 1969	No abstract, full-text publication not retrieved	-
10	Uotila & Wolf 1968	No analysis of diagnostic accuracy	-
11.	Ainamo & Tammisalo1968	-	Sample and examination methods not described to permit replication

No.	*Publications retrieved in PubMed search and read in full-text 1st author, year*	*Reason for exclusion according to protocol "Inclusion or exclusion of publications in the systematic review (Appendix 2)*	*Reason for exclusion according to protocol based on the QUADAS-tool "Interpretation of publications in the systematic review" (Appendix 4)*
12.	Ainamo & Tammisalo 1967	Study on skull	-

4.2.2 Evidence from studies of visual grading analysis

For evaluation of the evidence in the scientific literature on panoramic radiography for the assessment of alveolar bone loss in periodontal diseases, the problem specifications were:

• What is the evidence for visual grading analysis?

Nine studies were included on visual grading analysis for the assessment of alveolar bone loss in panoramic radiography, (Table 4.18). The reported measurability concerning the assessment of the *alveolar bone level* varied between the studies, from about 99% (Pepelassi & Diamanti-Kipioti 1997) to about 50% of the sites (Åkesson et al. 1989a). The number of sites possible to assess varied among tooth groups, being higher in the lower jaw as compared to the upper jaw (Stenström et al. 1981; Åkesson et al. 1989a; Pepelassi & Diamanti-Kipioti 1997; Jenkins et al. 2005; Ivanauskaite et al. 2008). Failure to identify the CEJ, overlap and blurring were often the cause for uninterpretability (Stenström et al. 1981; Kaimenyi & Ashley 1988; Persson et al. 2003). Less sites were unmeasurable when the bone loss was assessed by a ruler than as a distance between the CEJ and alveolar bone (Kaimennyi & Ashley 1988). Considering different panoramic machines, Molander et al. (1995b) concluded that Scanora® dental panoramic program resulted in higher image quality as compared to other machines. On average, visualisation of the alveolar bone level in the upper and lower jaw was equal for film based and digital panoramic radiography (Molander et al. 2004).

Four studies (Åkesson et al. 1989a; Pepelassi & Diamanti-Kipioti 1997; Persson et al. 2003; Ivanauskaite et al. 2008) presented comparisons between panoramic and intraoral radiography concerning the visibility for the assessment of the alveolar bone level. According to Persson et al. (2003) there were more non-readable sites in both jaws in panoramic radiography (16.7%) as compared to periapical radiography (7.9%). The results presented by Åkesson et al. (1989a) and Pepelassi and Diamanti-Kipioti (1997) indicated that the percentage of unacceptable sites in the upper and the lower jaws, respectively was comparable for panoramic and periapical radiography varying between 0.2-1.9% (Pepelassi & Diamanti-Kipioti 1997) and 15-29% (Åkesson et al. 1989a). In two studies (Åkesson et al. 1989a; Ivanauskaite et al. 2008), the visibility in panoramic and bitewing radiography was compared. Even though the percentage of unacceptable sites differed between the studies, the results indicated the same tendency. In the lower jaw, the percentage of sites with unacceptable image quality was comparable for panoramic and bitewing radiography, while panoramic radiography presented more unacceptable sites than bitewing radiography in the upper jaw.

The image quality for the detection of *vertical bone defect* and *furcation involvement* was analyzed in only one study (Ivanauskaite et al. 2008). In the lower jaw, above 90% of the sites presented acceptable image quality for the detection of vertical bone defect in both in panoramic and bitewing radiography. In the upper jaw, however, only two thirds of the sites were acceptable in panoramic radiography compared to 98% in bitewing radiography. Most teeth were assessed to present acceptable image quality for the detection of a furcation involvement.

Table 4. 18 Visual grading analysis/Measurability of panoramic radiography in assessment of alveolar bone

1st author Year	Panoramic machine Number of observers	Sample	Measurement unit	Statistical method	Reported results	Comments
Åkesson 1989a	• Orthopantomograph Model OP5 • 5 observers, each one evaluated radiographs of 20 patients	• 100 patients (mean age 38 yrs; 21-77) • sites for comparison of panoramic - *vs* periapical 5144 - *vs* bitewing 3182	image quality assessed as: • excellent • acceptable • unacceptable when bone loss scored with ruler by: - Björn et al (1969) - Håkansson et al. (1981)	• percentage of sites with quality categories - panoramic *vs* periapical - panoramic *vs* posterior bitewing • x^2- test to test differences between methods ($p \leq 0.05$)	• unacceptable - panoramic *vs* periapical: upper jaw 27% *vs* 29% lower jaw 20% *vs* 15% - panoramic *vs* posterior bitewing: - upper jaw 59% *vs* 29% - lower jaw 37% *vs* 27% • higher quality of intraoral than panoramic radiography	• patients with different severity of periodontitis
Ivanauskaite 2008	• Scanora dental panoramic program • 6 observers for bone loss • 5 observers for vertical bone defects and furcation involvement	• 96 patients (mean age 48.5 yrs; 20-85) • sites - for bone level: upper jaw 245 lower jaw 254 - for bone defect: upper jaw 1435 lower jaw 1446 • teeth for furcation involvement: 774 for comparison of panoramic *vs* posterior bitewing	• assessment of image quality as: - excellent - acceptable - unacceptable for bone level scored with ruler by Håkansson et al. (1981) • assessment of image quality as: - acceptable -unacceptable for detection of vertical bone defect and furcation involvement	• percentage of sites with quality categories for bone level and vertical defect • percentage of teeth with quality categories for furcation involvement	• excellent + acceptable image quality for bone level panoramic *vs* bitewing: upper jaw 65% *vs* 94% lower jaw 92% *vs* 92% • acceptable image quality for vertical bone defect in upper jaw 65% *vs* 98% lower jaw 93% *vs* 98% • for furcation involvement only 2% were rated unacceptable	• patients with varying degrees of bone loss

1st author Year	Panoramic machine Number of observers	Sample	Measurement unit	Statistical method	Reported results	Comments
Jenkins 2005	• Orthophos • 4 observers read all radiographs	• 50 patients (mean age 45 yrs±10; 27-73) with advanced periodontitis • 951 teeth	image quality assessed as • adequate • inadequate when mesial/distal bone margins clearly visible together with apex	• percentage of different tooth types with adequate image quality	• adequate image quality - upper jaw: 9.2% (first premolar) - 71% (first incisor) - lower jaw: 58.6% (canine) - 94.3% (second molar)	• aim of study was to analyse radiation doses • number of sites for different tooth types not presented

Table 4. 19 Accuracy of panoramic radiography in assessment of alveolar bone loss with criterion standard obtained during surgery

1st author Year	Radiographic method(s) Number of observers	Sample	Measurement unit	Criterion standard	Statistical method	Reported results	Comments
Åkesson 1992a	• Orthopantomograph Model OP5 • periapical and posterior bitewing radiography • 5 observers assessed all radiographs	• 23 patients with severe periodontitis (mean age 49 yrs; 27-73) • 137 sites in upper and 100 in lower jaw - 176 sites for panoramic *vs* bitewing - 213 sites for panoramic *vs* periapical	• bone loss: distance (mm) between occlusal level of buccal cusp and most apical level of marginal bone	• probing with extended probe (tip of 0.6 mm) • vertical distance in mm between occlusal level of buccal cusp and most apical level of marginal bone	• mean underestimation of radiographic measurements expressed in mm and in percent of distance obtained as criterion standard	• mean underestimation in panoramic *vs* bitewing *vs* periapical: - 3.3 *vs* 3.0 *vs* 2.3 mm - upper jaw 18% *vs* 22% *vs* 13% of distance - lower jaw 24% *vs* 17% *vs* 14% of distance	• splint used to obtain reference points • enlargement in radiographs calculated
Kaimenyi 1988	• Panelipse • 1 observer	• 9 patients (30-40 yrs) • 61 sites	• bone loss: linear measurement of distance CEJ-AB (mm)	• direct measurements during surgery	• difference in mm between surgical and radiographic measurements	• difference within: - 1 mm: 46% - 2 mm: 74% - 3 mm: 89% of sites	• enlargement in radiographs not described
Pepelassi 1997	• Ortho Ceph 10 • periapical radiography • 1 observer	• 100 patients with moderate to severe periodontitis (18-75 yrs) • 2536 teeth • 5072 sites (1966 sites with small osseous destructions)	• osseous destruction assessed as present • bone loss: - linear measurement of distance CEJ-AC (mm) - Schei ruler (1959)	• linear measurement in mm with calibrated probe	• agreement in % for detection of small destructions • difference between radiographic and surgical measurements in mm	• panoramic detected 0.8% and periapical 2.9% of small osseous destructions • mean bone loss (mm) during surgery 5.60 in panoramic 5.21 in periapical 5.59	• enlargement in radiographs not described

1st author Year	Radiographic method(s) Number of observers	Sample	Measurement unit	Criterion standard	Statistical method	Reported results	Comments
Pepelassi 2000	• Ortho Ceph 10 • periapical radiography • 1 observer	• 100 patients with moderate to severe periodontitis (18-75 yrs) • 2536 teeth, 1049 teeth with bone defect • 1234 defects	• bone defect assessed as present	• linear measurement with calibrated probe • impression to define dimensions and volume of bone defect	• detection of osseous defect in % of all defects • regression analysis to assess effects of factors on detection of defect	• defect depth overestimated by radiography • detection of defects panoramic *vs* periapical 21 *vs.* 62% • detection in panoramic dependant only on buccolingual width of defect	• enlargement in radiographs not described; may result in overestimation of defect depth

AB=alveolar bone; AC=alveolar crest; CEJ=cemento-enamel junction

Table 4. 20 Comparison between panoramic radiography and other radiographic method(s) in assessment of alveolar bone

1st author Year	Panoramic machine Number of observers	Sample	Measurement unit	Method for comparison	Statistical method	Reported results	Comments
Åkesson 1989b	• Orthopantomograph Model OP5 • 5 observers, each one evaluated radiographs of 20 patients	• 100 patients (mean age 38 yrs; 21-77) • bone loss of canine distally to second molar distally: - upper jaw 594 sites (285 with bone defect) - lower jaw 902 sites (218 with bone defect) • furcation involvement: 151 teeth	• bone loss scored with ruler by Håkansson et al. (1981); score 4 indicates normal BL and score 10 total loss of marginal bone • vertical bone defects and furcation involvements (presence)	• 4 posterior bitewing radiographs	• concordance between techniques in percentage of: - sites (bone loss, vertical bone defects) - teeth (furcation involvement)	• concordance on - bone loss in: upper jaw: 69% lower jaw: 68% - bone defects in: upper jaw 42% lower jaw 53% - furcation involvement: 65%	• patients with different severity of periodontitis
Ivanauskaite 2006	• Scanora panoramic dental program • 6 observers for bone loss • 5 observers for vertical bone defect and furcation involvement	• 96 patients (mean age 48.5 yrs; 20-85) • bone loss of canine distally to second molar distally: - upper jaw 245 sites; - lower jaw 254 sites • bone defect: - upper jaw: 1435 sites	• bone loss scored with ruler by Håkansson et al. (1981); score 4 indicates normal BL and score 10 total bone loss • bone defect and furcation involvement (present or absent)	• 4 posterior bitewing radiographs	• agreement between techniques in percentage of sites (bone loss and bone defects) or teeth (furcation involvement)	• agreement on - bone loss: identical scores in upper jaw 56% lower jaw 58%; - bone defects in: upper jaw 89% lower jaw 94% - furcation involvement 93%	• patients with varying degrees of bone loss

1st author Year	Panoramic machine Number of observers	Sample	Measurement unit	Method for comparison	Statistical method	Reported results	Comments
		- lower jaw: 1446 sites • furcation involvement: 774 teeth					
Molander 1991	• Orthopantomograph Model OP5 • 2 observers each read half of radiographs	• 200 patients (mean age 42.8 yrs±15.3) • 8968 sites in whole dentition	• bone loss measured as distance between CEJ and marginal bone level in mm • angular bony defects and furcation involvement (present or absent)	• 16 periapical + 4 bitewing radiographs	• agreement between techniques in percentage of sites • difference in mm at most	• agreement on - bone loss 55% (48-77), agreement decreased with increasing bone loss - angular bony defects 40% (33-46) - furcation involvement molars 61% upper premolars 12%	• ruler for panoramic radiography with enlargement of 1.3
Molander 1995a	• Orthopantomograph ® Model OP 5 • 3 observers	• 40 patients (mean age 42.1 yrs; range 17-80; SD=16.5) • median number of teeth per patient 24 (4-28) • disease prevalence 59.8%	• marginal bone loss - osteolysis of crestal bone layer - distance CEJ-AC larger than 1.5 mm	• panoramic radiography + full mouth survey (16 periapical + 4 posterior bitewing radiographs)	• sensitivity • specificity	• sensitivity around 80% • specificity around 35%	• threshold for bone loss was low
Persson 2003	• PM 2002 CC Proline • 2 observers, each one read half of	• 292 subjects (mean age 55.5±12.6 yrs) • 21 462 sites in panoramic and 11	• distance CEJ-BL in mm • proportional distance between	• intraoral radiographs	• intra-class coefficient correlation (ICC) between	• comparison panoramic and intraoral - distance CEJ-BL: ICC lowest for tooth	• questionable that different observers assessed

1st author Year	Panoramic machine Number of observers	Sample	Measurement unit	Method for comparison	Statistical method	Reported results	Comments
	panoramic radiographs 1 another observer read all intraoral radiographs	395 in intraoral radiography	CEJ-BL relative to root length (CEJ-BL/root length)		readings from intraoral and panoramic radiographs • differences between readings in panoramic and intraoral	37 (0.80), highest for tooth 22 (0.89) - mean difference between readings 0.01-0.71 mm	methods under comparison • from figures (plot diagram) half sites presented no bone loss
Rohlin 1989	• Orthopantomograph Model OP5 • 5 observers, each one read radiographs of 20 patients	• 100 patients (age 20-79 yrs) • sites: - upper jaw 1530 (368 with bone defect) - lower jaw 1873 (311 with bone defect) • with furcation involvement 161 teeth	• bone loss scored with ruler Björn et al. (1969); score 4 indicates normal BL and 10 total loss of marginal bone • angular bone defect and furcation involvement (presence)	• 14 periapical radiographs	• concordance between techniques in percentage of: - sites (bone loss, vertical bone defect) - teeth (furcation involvement)	• concordance on - bone loss in: upper jaw: 66%, lower jaw: 74%, - bone defects in upper jaw: 46% lower jaw: 40% - furcation involvement upper jaw: 64% lower jaw: 70%	• patients with different severity of periodontitis

AC=alveolar crest; BL=bone level; CEJ=cemento-enamel junction

Table 4. 21 Observer performance of panoramic radiography for visual grading analysis

1st author Year	Panoramic machine Number of observers	Sample	Measurement unit	Intra-observer agreement Percent	Intra-observer agreement Kappa	Inter-observer agreement Percent	Inter-observer agreement Kappa	Comments
Åkesson 1989a	• Orthopantomograph Model OP5 • 5 observers	• 3 patients • 144 sites for comparison between panoramic and periapical radiography • 100 sites for comparison between panoramic and bitewing radiography	• assessment of image quality as: - excellent - acceptable - unacceptable when bone loss was scored with ruler by: - Björn et al. (1969) - Håkansson et al. (1981) • observer agreement expressed as - overall - kappa	• ruler by Björn et al. (1969) 76% 81% 85% 85% 85% • ruler by Håkansson et al. (1981) 77% 77% 80% 83% 97%	 0.54 0.57 0.75 0.75 0.76 0.61 0.67 0.67 0.70 0.94			• calculated mean kappa for ruler by: - Björn et al. (1969) 0.67 for panoramic *vs* 0.68 for periapical - Håkansson et al. (1981) 0.72 for panoramic *vs* 0. 67 for bitewing
Åkesson 1992b	• 3 samples - two samples with: Orthopantomograph Model OP5 - one sample: different panoramic machines from different clinics • 3 observers	• 3 patients • 150 sites	• assessment of image quality as: - excellent - acceptable - unacceptable when bone loss was scored with ruler by: - Björn et al. (1969) - Håkansson et al. (1981)	• ruler by Björn et al. (1969) 75% 78% 86% • ruler by Håkansson et al. (1981) 84% 85% 91%	 0.62 0.66 0.77 0.67 0.73 0.77			

1st author Year	Panoramic machine Number of observers	Sample	Measurement unit	Intra-observer agreement Percent Kappa	Inter-observer agreement Percent Kappa	Comments
Ivanauskaite 2008	• Scanora dental panoramic program • 6 observers for bone loss • 5 observers for vertical bone defects and furcation involvement	96 patients (mean age 48.5 yrs; 20-85) • bone loss of canine distally to second molar distally: - upper jaw 245 sites; - lower jaw 254 sites • bone defect: - upper jaw: 1435 sites - lower jaw: 1446 sites • furcation involvement: 774 teeth	• assessment of image quality as: - excellent - acceptable - unacceptable when bone loss scored with ruler by Håkansson et al. (1981) • assessment of image quality as: - acceptable - unacceptable for detection of vertical bone defect and furcation involvement	Panoramic *vs* bitewing • bone loss assessed with ruler by Håkansson et al. (1981) 76 *vs* 78% 0.59 *vs* 0.56 87 *vs* *89*% 0.75 *vs* 0.49 82 *vs* 91% 0.63 *vs* 0.42 • bone defect 93 *vs* 98% 0.79 *vs* 0.34 93 *vs* 98% 0.81 *vs* 0.62 95 *vs* 98% 0.88 *vs* 0.32	Panoramic *vs* bitewing • agreement of 6 observers bone loss 0.45 *vs* 0.28 bone defect 0.62 *vs* 0.25 furcation involvement 97% *vs* 96% 0.02 *vs* 0.06 • agreement for pair of observers bone loss 43-91% *vs* 45-93% 0.31-0.61 *vs* 0.16-0.43 bone defect 0.51-0.73 *vs* 0.11-0.40 furcation involvement 97-98% *vs* 96-100%	• overall agreement for detection of furcation involvement for both methods 93%-98% • agreement for pair of observers accounted by Cohen's kappa

Table 4. 22 Observer performance of panoramic radiography in assessment of alveolar bone loss

1st author Year	Panoramic machine Number of observers	Sample	Measurement unit	Intra-observer agreement Percent Kappa	Inter-observer agreement Percent Kappa	Comments
Åkesson 1989b	•Orthopantomograph Model OP5 • 5 observers	• 3 patients • 100 sites in posterior regions	• bone loss scored with ruler by Håkansson et al. (1981) • observer agreement expressed as - overall - kappa	• panoramic *vs* posterior bitewing - bone loss 66 *vs* 74% 0.46 *vs* 0.55 71 *vs* 76% 0.52 *vs* 0.61 76 *vs* 77% 0.60 *vs* 0.62 77 *vs* 83% 0.60 *vs* 0.71 82 *vs* 90% 0.65 *vs* 0.82		• only sites considered readable included
Ivanauskaite 2006	• Scanora dental panoramic program • intraobserver: 3 observers • interobserver: 5 observers	• 96 patients • bone loss assessments of canine distally to second molar distally: - upper jaw 245 sites - lower jaw 254 sites • bone defect assessments: - upper jaw 1435 sites - lower jaw 1446 sites • furcation involvement assessments 744 teeth	• bone loss scored with ruler by Håkansson et al. (1981); score 4 indicates normal] 10 total bone loss • bone defect and furcation involvement (present or absent) • observer agreement expressed as - overall - kappa	• panoramic *vs* bitewing - bone loss 72 *vs* 76% 0.46 *vs* 0.55 69 *vs* 66% 0.45 *vs* 0.43 59 *vs* 55% 0.37 *vs* 0.31 - bone defect 93 *vs* 94% 0.57 *vs* 0.62 92 *vs* 92% 0.62 *vs* 0.52 99 *vs* 99% 0.56 *vs* 0.61 - furcation involvement 98 *vs* 97% 0.79 *vs* 0.72 93 *vs* 96% 0.65 *vs* 0.77 97 *vs* 97% 0.73 *vs* 0.78	• panoramic *vs* bitewing - agreement for several observers expressed as kappa: bone loss 0.28 *vs* 0.29 bone defect 0.35 *vs* 0.38 furcation involvement 0.53 *vs* 0.65 - agreement for pair of observers bone loss 0.08-0.58 *vs* 0.04-0.43	• patients with varying degree of bone loss • only sites considered readable included • observer agreement of panoramic comparable to that of posterior bitewing radiography • one observer had lowest kappa values for bone loss in both methods

1st author Year	Panoramic machine Number of observers	Sample	Measurement unit	Intra-observer agreement Percent Kappa	Inter-observer agreement Percent Kappa	Comments
Rohlin 1989	• Orthopantomograph Model OP5 • 5 observers	• 3 patients • 144 sites	• bone loss scored with ru Björn et al.; (1969) score indicates normal BL and bone loss • observer agreement exp - overall - kappa	• panoramic *vs* periapical - bone loss 78 *vs* 78% 0.63 *vs* 0.62 79 *vs* 81% 0.66 *vs* 0.70 80 *vs* 82% 0.68 *vs* 0.71 85 *vs* 84% 0.77 *vs* 0.74 86 *vs* 85% 0.78 *vs* 0.74		• only sites consi readable readable include • calculated mean appa: 0.70 same for panoramic and periapical
1st author Year	**Panoramic machine Number of observers**	**Sample**	**Measurement unit**	**Intra-observer performance Difference between two repeated measurements in mm Mean SD**	**Inter-observer performance Underestimation compared to "true value" (mm) for 5 observers**	**Comments**
Åkesson 1992a	•Orthopantomograph Model OP5 • intraobserver: 2 observers • interobserver: 5 observers	• 23 patients with severe periodontitis (mean age 49 yrs; 27-73) • 137 sites in upper and 100 in lower jaw - 176 sites for panoramic *vs* bitewing - 213 sites for panoramic *vs* periapical	• radiographic measurement was compared to probing with extended probe during surgery ("true value") • vertical distance was measured in mm between occlusal level of buccal cusp and most apical level of marginal bone	• panoramic Observer 1: 0.02 1.6 Observer 2: 0.3 1.6 • bitewing Observer 1: 0.2 0.7 Observer 2: 0.4 0.9 • periapical Observer 1: 0.2 1.1 Observer 2: 0.2 1.5	• panoramic 1.8, 3.1, 3.4, 3.5,4.8 • bitewing 2.1, 2.8, 3.2, 3.3, 3.8 • periapical 1.4, 2.0, 2.0, 2.5, 3.4	Inter-observer performance: values based on graphs in Fig. 3 in the study by Åkesson et al. (1992a)

BL=bone level, SD – standard deviation

4.2.3 Evidence from studies of assessment of alveolar bone loss

For evaluation of the evidence in the scientific literature on panoramic radiography for the assessment of alveolar bone loss in periodontal diseases, the problem specifications were:

• What is the evidence for diagnostic accuracy?

There were four studies on the *diagnostic accuracy of panoramic radiography in which panoramic radiography was compared to a criterion standard obtained during/after surgery* (Table 4.19). The results of three included studies indicated that panoramic radiography underestimates the degree of bone loss (Kaimenyi & Ashley 1988; Åkesson et al. 1992a; Peppelassi & Diamanti-Kipioti 1997), while the results of one study (Pepelassi et al. 2000) presented the depth of osseous defects to be overestimated by panoramic radiography. In the study by Pepelassi et al. (2000), similar to two other studies (Kaimenyi & Ashley 1988; Pepelassi & Diamanti-Kipioti 1997) the enlargement factor of radiography was not taken into account. When bone loss was minor, the difference between measurements in panoramic radiography and during surgery was small (Pepelassi & Diamanti-Kipioti 1997). In cases with severe periodontitis, however, the mean difference between distances measured in panoramic radiography and during surgery were larger, corresponding to 18% of the distance in the upper jaw and 24% in the lower jaw (Åkesson et al. 1992a). Panoramic radiography presented lower accuracy compared to bitewing and to periapical radiography (Åkesson et al. 1992a; Pepelassi & Diamanti-Kipioti 1997; Pepelassi et al. 2000). Mean underestimation in panoramic radiography was not only influenced by the severity of bone loss, but also on the tooth group (range 2.5-5.4 mm) and the observer assessing the radiographs (Åkesson et al. 1992a). In the study by Åkesson et al. (1992a) being the only among the five included studies to be based on more than one observer, the observer with the highest accuracy presented a mean underestimation of 1.9 mm and the observer with the lowest accuracy a mean underestimation of 4.9 mm. Mean depth of defects measured during surgery was 3.1 mm, measured in panoramic radiography 4.1 mm, and in periapical radiography 3.4 mm (Pepelassi et al. 2000). When measurements were performed with the aid of a Schei ruler, there was no significant difference between detection of osseous bone defects in panoramic and periapical radiography (Pepelassi & Diamanti-Kipioti 1997). Only 21% of bone defects were detected in panoramic as compared to 62% in periapical radiography (Pepelassi et al. 2000). No study on furcation involvement was identified where panoramic radiography was compared to a reference standard obtained during surgery.

Six studies presented *comparison between panoramic and other radiographic methods* (Table 4. 20): two studies with comparison between panoramic and posterior bitewing radiography (Åkesson et al. 1989b; Ivanauskaite et al. 2006), two studies with comparison between panoramic and periapical radiography (Rohlin et al. 1989; Persson et al. 2003), and two studies with comparison between panoramic and combined periapical and bitewing radiography (Molander et al. 1991, Molander et al. 1995a). *Bone loss* was assessed with a ruler in three studies (Åkesson et al. 1989b, Rohlin et al. 1989, Ivanauskaite et al. 2006). In two studies (Molander et al.1991; Persson et al. 2003), bone loss was measured in millimetres, but only in one of these studies (Molander et al. 1991) there was a description of the enlargement of panoramic radiography. In four studies (Åkesson et al. 1989b; Rohlin et al. 1989; Persson et al. 2003; Ivanauskaite et al. 2006) only sites evaluated as readable was compared while in two studies (Molander et al. 1991; Molander et al. 1995a) the comparison between the radiographic methods included all sites. When the bone loss was scored with a ruler, there was a complete agreement between panoramic and intraoral radiography in 56-69 % of the sites in the upper jaw and in 58-74% in the lower jaw (Åkesson et al. 1989b; Rohlin et al. 1989; Ivanauskaite et al. 2006). There was a very high agreement if a difference of ± 1 score was accepted, being 95% for the agreement between panoramic and posterior bitewing radiography (Ivanauskaite et al. 2006) and 97% in the upper jaw and 99% in the lower jaw for the agreement between panoramic and periapical radiography (Rohlin et al. 1989). Panoramic radiography scored less bone loss than bitewing radiography in 28% of sites in the upper jaw and 25% of the sites in the lower jaw (Ivanauskaite et al. 2006). On the other hand, panoramic radiography scored more bone loss in 28% of the sites in the upper jaw and 25% of the sites in the lower jaw (Ivanauskaite et al. 2006). The concordance between panoramic and posterior bitewing radiography (Åkesson et al. 1989b) and between panoramic and periapical radiography (Rohlin et al. 1989) was highest for the lowest score (no bone loss). For linear measurements, Persson et al. (2003) reported low mean differences between readings in panoramic and intraoral radiographs (0.00 – 0.06 mm). Molander et al. (1991) reported differences at most of ± 1 mm in 88% and of ± 2 mm in 96% of sites. The sensitivity to detect alveolar bone loss was high (80%) and the specificity low (35%) (Molander et al. 1995a).

There were four studies (Åkesson et al. 1989b; Rohlin et al. 1989; Molander et al. 1991; Ivanauskaite et al. 2006) on the concordance between panoramic and intraoral radiography on the detection of *vertical bone defect*. The agreement ranged from 46 to 89% of the sites in the upper jaw and from 40 to 94% of the sites in the lower jaw (Åkesson et al. 1989b; Rohlin et al. 1989; Ivanauskaite et al. 2006). Some bone defects were only detected in panoramic radiography and some only in intraoral radiography. Åkesson et al. (1989b) found that 21% of the vertical bone defects in both jaw were detected only in panoramic radiography. On the other hand, 37% of the bone defects in the upper jaw and 26% in the lower jaw were detected only in bitewing radiography (Åkesson et al. 1989b). Bone defects were detected only in panoramic radiography in 21% and only on periapical radiography in 36% (Rohlin et al. 1989). Sensitivity to detect bone defects in panoramic radiography was higher (0.73) than that of bitewing radiography (0.53) (Ivanauskaite et al. 2006) based on a criterion radiographic standard.

There were four studies (Åkesson et al.1989b; Rohlin et al. 1989; Molander et al. 1991; Ivanauskaite et al. 2006) on the concordance between panoramic and intraoral radiography on the detection of *furcation involvement*. The concordance between panoramic and intraoral radiography ranged between 64 and 93% of all assessed teeth (Åkesson et al. 1989b; Rohlin et al. 1989; Ivanauskaite et al. 2006), but was low for the upper premolars (12%) (Molander et al. 1991). Calculated sensitivity to detect furcation involvement in panoramic radiography was comparable (0.80) to that of bitewing radiography (0.70) (Ivanauskaite et al. 2006). Åkesson et al. (1989b) found that 19% of the furcation involvements were detected only in panoramic radiography and 16% only in bitewing radiography. Similar results were obtained for the comparison between panoramic and periapical radiography. Furcation involvements were detected only in panoramic radiography in 16% and only in periapical radiography in 15% (Rohlin et al. 1989).

4.2.4 Evidence from studies of observer performance

Evidence on *observer performance* expressed as overall percent and as kappa values in panoramic radiography is presented in Tables 4.21 and 4.22.

Observer performance of panoramic radiography in *visual grading analysis* (Table 4.21) when assessing alveolar bone loss was reported in three studies (Åkesson et al. 1989a; Åkesson et al 1992b; Ivanauskaite et al. 2008). In these studies, rulers were used for the assessment of the *bone loss*. Kappa values of intra-observer agreement ranged between 0.54 and 0.94, most observers presenting a value above 0.60 (Åkesson et al. 1989a; Åkesson et al. 1992b: Ivanauskaite et al. 2008). The intra-observer agreement for visual grading analysis concerning the alveolar bone assessment was compared for panoramic and bitewing radiography in only one study (Ivanauskaite et al. 2008). The overall agreement was comparable (panoramic radiography: range 76–87% versus bitewing radiography: range 78–91%), but Kappa values were higher for panoramic radiography. Inter-observer agreement for visual grading analysis concerning the bone loss assessment was studied in only one publication (Ivanauskaite et al. 2008). Agreement for six observers was moderate for panoramic radiography (κ=0.45) and fair for bitewing radiography (κ=0.28). For visual grading analysis concerning detection of *vertical bone defect* only one study elucidated observer performance (Ivanauskaite et al. 2008). Kappa values for intra-observer agreement for panoramic radiography were higher than for bitewing radiography. Inter-observer agreement of five observers was substantial for panoramic radiography (κ=0.62) but only fair for bitewing radiography (κ=0.25) (Table 4.21). Inter-observer agreement of five observers for visual grading analysis concerning detection of *furcation involvement* was 97% (κ=0.02) for panoramic radiography and 96% (κ=0.06) for bitewing radiography (Ivanauskaite et al. 2008).

Four studies (Åkesson et al. 1989b; Rohlin et al. 1989; Åkesson et al. 1992a; Ivanauskaite et al. 2006) presented the observer performance for *assessment of alveolar bone loss.* Table 4.22 shows three studies (Åkesson et al. 1989b; Rohlin et al. 1989; Ivanauskaite et al. 2006), which present intra-observer agreement as overall percent and kappa values when assessing the *bone loss* with a ruler. The Kappa values, which varied between the studies as well as within one and the same study, ranged between 0.37 and 0.78 in panoramic radiography.

Corresponding values of intraoral radiography were 0.31 and 0.82. When measuring the bone loss in millimetres, the differences between two repeated measurements were small (mean 0.02 – 0.3 mm; SD 1.6) and comparable to bitewing radiography (mean 0.2-0.4 mm SD 0.7-0.9) and periapical radiography (mean 0.2 mm; SD 1.2-1.5) (Åkesson et al. 1992a). For the detection of *bone defect* and *furcation involvement*, intra-observer agreement was higher than for the assessment of the bone level/loss (Ivanauskaite et al. 2006).

Inter-observer performance was presented in two studies (Åkesson et al. 1992a; Ivanauskaite et al. 2006). Table 4.22 presents inter-observer agreement as overall percent and kappa values of three studies when assessing alveolar bone with a ruler and one study where the measurements were compared to those obtained during surgery. The agreement was comparable for panoramic and bitewing radiography, 0.28 and 0.29, respectively. Inter-observer agreement for pair of observers when assessing alveolar bone level with a ruler varied widely for panoramic radiography (kappa 0.08-0.58) as well as for bitewing radiography (kappa 0.04-0.43). One of the observers deviated from the other observers, independently of radiographic method. This is in line with the results of inter-observer performance when assessing alveolar bone loss in millimetres (Åkesson et al. 1992a). The five observers varied substantially and systematically. The mean underestimation of bone loss was highest for panoramic, bitewing and periapical radiography for one observer and lowest for all methods for another observer. The mean underestimation compared to measurements performed during surgery varied between 1.8 and 4.8 mm in panoramic radiography, between 2.0 and 3.6 mm in bitewing radiography, and 1 and 3.2 mm in periapical radiography. For the detection of bone defect and furcation involvement, inter-observer agreement of 5 observers was comparable for panoramic radiography and bitewing radiography (kappa 0.35 and 0.38, respectively). For the detection of furcation involvement, agreement for bitewing radiography was somewhat higher (kappa 0.53 and 0.65, respectively) (Ivanauskaite et al. 2006).

5. DISCUSSION

5.1 Methodological considerations

5.1.1 Radiographic methods

When performing scientific studies including radiographic methods, there are several factors to consider from an ethical point of view. One factor, which is most important, is the radiation dosage to the patients. In this study, patients were referred for radiographic examination prescribing panoramic and posterior bitewing radiography. Thus, patients received no additional radiation dose due to their participation in this study. Even so it was important to apply for and to have the ethical approval. It is important to consider an ethical application and to receive an approval when a radiographic examination is included in a scientific study in oral health care, also when the radiographic examination is employed for evaluation of treatment outcomes.

Scanora® dental panoramic program was chosen as the panoramic method to be examined in the studies of patients. This was due to that it was a new technique for panoramic radiography and there were few studies in the scientific literature on this equipment. Furthermore, the results of one study indicated that the best subjective image quality was achieved with Scanora® dental panoramic program as compared with seven other panoramic machines and programs for dental diagnostics (Molander et al. 1995b). According to Molander at al. (1995b), this is due to the provision of the smaller focal spot and narrower X-ray beam compared to other panoramic machines. Panoramic radiography using Scanora® dental panoramic program has a magnification factor 1.7, which can influence the visual grading analysis for the alveolar bone loss.

The examinations in this study with Scanora® dental panoramic program and posterior bitewing radiography were performed of the same patients at the same time. Even if most changes of the alveolar bone are chronic and slowly progressing, it is optimal to perform both examinations at the same time to ascertain that a delay did not lead to worsening or improvement of the condition (STARD-document Bossuyt et al. 2003). As posterior bitewing radiography was prescribed for the patients included in this study, we did not have access to intraoral radiographic examination of the whole dentition. From a scientific viewpoint it would have been an advantage to be able to present comparisons of the diagnostic properties of Scanora® dental panoramic program and intraoral radiography of the incisors and the mesial site of canine as well.

Vertical angulations were kept constant, -5° in panoramic radiography and +10° in bitewing radiography. As the X-ray beam of panoramic radiography originates from the opposite side as compared with the X-ray beam of bitewing radiography, the difference in vertical angulation is negligible. For obvious reasons the horizontal angulation was different in panoramic and bitewing radiography, particularly in the premolar region. This influenced the results of the visual grading analysis, the assessment of the alveolar bone level by scores, the detection of vertical bone defect and furcation involvement. In panoramic radiography, one can influence the horizontal angulation to a limited extent by optimizing the patient position in the radiographic apparatus. In bitewing radiography, the horizontal angulation can be shifted manually and thereby an optimal orthoradial projection can be obtained more easily.

5.1.2 Sample

It is important to know the described sampling schema of patients, teeth, and sites assessed, since it may be helpful in judging the generalisability of the scientific findings. According to the QUADAS document (Whiting et al. 2003), it is important to obtain a sample with patients with different demographic and clinical features, because a study of the diagnostic accuracy may have limited clinical applicability if the spectrum of tested patients is not similar to the patients on whom the test will be used in practice. The spectrum of the patients participating in the present studies would be acceptable taking factors such as disease prevalence and severity, age and sex of the patients into account. The age and sex distribution of the examined sample was presented in Table 3.1. The distribution of alveolar bone scores of assessed sites in Table 3.3 describes the disease prevalence and severity. A more precise description of the disease prevalence and severity could have included a combination of different diagnostic examinations, such as bleeding on probing and probing the depth of periodontal pockets.

Also, all relevant information regarding how participants were selected to the study has to be provided (STARD-document Bossuyt et al. 2003). The patients taking part in this study had been referred to the clinic for radiographic examination. They were to have not only periodontal treatment but to have oral health care in general. Thus, radiography was part of a general examination of the oral health status. The consecutive referred patients of this study allowed a compilation of a sample that may be representative of the general population consuming oral health care in Scandinavia. However, if the sample had been compiled in Lithuania, a sample with a more extensive need of oral health care would probably have been examined.

In the sample of this study, patients were of different age-groups between 20- and 89-year-old, with 40-49 years and 50-59 years were more strongly represented than the other age-groups. Less number of patients was from the younger age-groups (20- to 39-year-old) and even less number of patients was from the older age-groups (60- to 89-year-old). Although the patients' age was not homogenously distributed within the sample of the present study, the sample represented an adequate spectrum of tested patients, as periodontal diseases are more prevalent in patients 40-59 years as compared to younger age-groups. It has been shown in Lithuania (Globienė 2001) as well as in Sweden (Norderyd & Huguson 1998) that the prevalence of periodontal disease is different in different age-groups of patient. Results of cross-sectional studies suggest that the prevalence of bone loss increases in the 40-59-age-group (Norderyd & Huguson 1998). In the age-group 40 years, 5 percent had alveolar bone loss exceeding one third of the root length around the majority of their teeth while the corresponding figure in the age-group 50 years was 21 percent (Huguson et al. 1998; Norderyd & Huguson 1998). Thus, the likelihood of findings that will influence patient management will be higher in individuals of these age-groups than in those of younger age-groups, also as in age-group over 60-year-old, because in these patients it could be less number of teeth.

In the sample of the present study, the scores corresponded mostly to alveolar bone loss of one third or less of the root length. There were also scores, which presented normal alveolar bone level and bone loss more than half of the roots length. About 15 percent of the sites of the present study presented a vertical bone defect and around 10 percent of the teeth a furcation involvement. The frequency of findings or disease will influence the concordance between diagnostic methods as well as other diagnostic properties such as the predictive values. The prevalence of vertical bone defects and furcation involvements will result in a higher concordance in such a sample as compared with a sample with a higher frequency of positive findings. On the other hand, the setting from which the sample was collected could be considered a primary care setting that represented a population with a relative low frequency of vertical bone defects and furcation involvements. In a secondary care setting, such as in a specialist clinic of periodontology, the spectrum of disease as well as the range and relative frequency of changes of the bone would be higher.

Furthermore, is important to collect a sample with a wide age distribution and a distribution close to equal between the sexes in the case of panoramic radiography where the patients' anatomy influences image quality. For the evaluation of diagnostic accuracy, not just the number of patients is important, but also the number of teeth, sites, and of assessments. In this study of 96 patients, 499 tooth sites out of 2885 tooth sites in the posterior and distal canine region were chosen for the visual grading analysis and for the assessment of alveolar bone loss. The methods of randomly dividing the patients into groups resulted in a representative sample in this study with an even distribution between jaws (upper and lower), sites (right and left), and regions (molar, premolar, and distal canine). As the image quality of panoramic radiography varies between the upper and lower jaws as well as between the anterior and posterior parts of the jaws, it was also important to make a selection of sites. Because the patients of this study were consecutive and not searching periodontal treatment but oral health care in general, the prevalence and severity of periodontal changes of the sample was not known beforehand. For the visual grading analysis concerning detection of the vertical bone defects and for the detection of vertical bone defects, all possible 2885 tooth sites from the canine distally to the second molar distally of the 96 consecutive patients were assessed. For the visual grading analysis concerning detection of furcation involvement and the detection of furcation involvement, all possible 748 teeth with a furcation were assessed. The number of observers (six for the visual grading analysis for the assessment of alveolar bone loss by score and five for visual grading analysis for the detection of the vertical bone defects and furcation involvements) markedly increased the number of assessments.

5.1.3 Observers

As diagnostic methods are applied by numerous examiners or observers, it is important that evaluations of all diagnostic methods include multiple observers. Otherwise, it would be difficult to determine whether the study findings were the result of one or few observer's skills in applying the method(s) under examination. Variability in the processing and reading of a diagnostic examination will affect measures of diagnostic properties. Observer performance as an important diagnostic property anticipates a study design with several observers. It has been proposed that a number of observers larger than six have little consequence on the results when a reasonably large sample is examined (Swets & Picket 1982; Hintze et al. 2003). As it was considered that the sample was reasonably large, six observers were asked to participate in this study in the visual grading analysis concerning the assessment of alveolar bone loss and in the scoring of the bone level. Five observers participated in the visual grading analysis for the detection of the vertical bone defect and furcation involvement and for the detection of these two findings.

If the amount of observers' training is presented in a study, it can help readers of the publication to judge whether similar results are attainable in their own settings with possibly less experienced observers (STARD-document Bossuyt et al. 2003). In this study, the amount of observers' training in the field of dental and maxillofacial radiology ranged from 1 to 30 years and the mean was 13 to 15 years. Three were specialists and three were not specialists in dental and maxillofacial radiology. One of those not being specialist in radiology was specialist in periodontology. Thus, the observers' training varied and included a range of expertise.

5.1.4 Reading conditions

The effect of viewing conditions on the visual detection of radiographic details has been studied for example by the method of perceptibility curves (Welander et al. 1983). Extraneous light and improper masking of radiographs reduce the amount of information available to the observer (Welander et al. 1983). In Sweden, where the assessments of radiographs were made, the Swedish Standards Institute (SIS) sets the standard (SS-EN 12464-1) for ambient light conditions in a dental practise to about 1000 lux and in a radiology department for reading images to 300 lux. There are different results of lightening conditions in different studies. According to Cederberg et al. (1998) different lighting conditions did not appear to affect the diagnostic performance of intraoral radiography. On the other hand, in a study about viewing conditions on the perceptibility of radiographic details, Welander et al. (1983) concluded that room light should be dimmed so that reflected light from the radiographs and its surroundings is minimized. In line with this, it was demonstrated that reducing ambient light from 1000 lux to less than 50 lux significantly increased the accuracy when diagnosing approximal carious lesion on a monitor (Hellén-Halme at al. 2008). There are no controversial opinions about that radiographs should be mounted in light-masking frames and that a viewing box should be used (Welander et al. 1983; Espelid 1987; Patel et al. 2000) and extraneous light should be eliminated (Welander et al. 1983). The importance of magnification when it comes to interpretation of intraoral radiographs has been stressed, but Espelid (1987) could not confirm that 2x magnification always provided improved viewing conditions. In the UK, the routine use of "ideal viewing conditions; have been recommended in a set of influential guidelines" (Patel et al. 2000). It is important to read radiographs in dimmed background lighting and to put them on a light box with fixed intensity. In the present studies (Ivanauskaite et al. 2006; 2008), a light box (15 x 30 cm^2) with fixed intensity was used and placed in a quiet room with dimmed background lighting. The observers used a magnification viewer with 2 times magnification, when possible.

Not only the professional background and expertise but also prior training and calibration will influence the observer's assessments. A strategy to improve the diagnostic outcomes and observer agreement is to support the observers in the diagnostic process. One way to do this is to discuss criteria to be implemented in order that the observers learn together, internalize the criteria and experience the criteria themselves prior to the actual readings. In order to ascertain that the criteria were appropriate in this study, the alveolar bone loss of several of patients was assessed by the observers prior to the data collection. Providing observers with written diagnostic cues and reference images, an approach previously adopted by, for example Larsen et al. (1977), Rohlin and Petersson (1989) and Cholitgul et al. (1990), is one type of support. Measures such as reference images are important, as oral calibration may only result in an initial effect (Poulsen et al.1980; Reit 1987).

The knowledge of results of one diagnostic method can influence the assessment in the diagnostic method to be compared. Such knowledge is likely to increase the agreement between the methods. To avoid diagnostic review bias in this study, there was an interval between the readings of panoramic radiography and bitewing radiography. Also, it is important to balance the number of radiographs and time asked from the observer to assess the radiographs. If the observers have a limited time and get impatient they can bias the result of the study. Therefore the maximum observation time was one hour in this study for both radiographic methods.

5.1.5 Visual grading analysis

Methods for evaluating properties of a method can be divided into a few major groups based on measuring principle and type of result as described in the “Introduction”. In the hierarchical conceptual model that Fryback and Thornbury (1991) proposed as an organising structure for evaluating diagnostic methods, the lowest level was the technical efficacy. But the 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.

An asymmetry between physical characteristics and diagnostic outcome may exist in that higher technical efficacy of a diagnostic method 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 (Blesser & Ozonoff 1977). Kullendorff et al. (1996) demonstrated that although intraoral radiography with Ektaspeed film presented superior high-contrast resolution compared to direct digital radiography, there was no statistical difference in diagnostic accuracy measured as the area under a ROC-curve for the detection of periapical bone lesions. Another example are the results by Dannewitz et al. (2002) that demonstrated that panoramic radiographs taken at reduced milliamperes of approximately 50 percent had inferior image quality as assessed by means of the visibility of 21 anatomical features, but there was no difference when scoring 30 pathological findings. That is why a higher level of efficacy evaluation, such as the ones proposed by Fryback and Thornbury (1991), is necessary to evaluate. At the next level, diagnostic accuracy efficacy, the image must be interpreted by observers in an attempt to make a diagnosis. A pragmatic approach to this level is visual grading analysis, which is based on the visibility of certain anatomical structures linked to a diagnostic task. This method takes into account technical factors *and* observer-dependent factors.

There is a spectrum of classification systems for visual grading analysis. Table 1.1 in the Introduction presents examples of classification systems for visual grading analysis with different descriptors that have been implemented in oral and maxillofacial radiography. In relative grading, images from two diagnostic methods are compared simultaneously, (see for example Molander et al. 2004) whereas in absolute grading, the two methods are evaluated separately (see for example Åkesson et al. 1989a; 1992b; Ivanauskaite et. al. 2008). Because we were interested to compare Scanora® panoramic radiography and bitewing radiography, not only for visual grading analysis but also for the assessment of bone loss, we applied 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 (Tingberg 2000). The diagnostic task can be different in different methods but also with the same radiographic method such as bitewing radiography, which has two different tasks: to diagnose caries and to diagnose changes of the alveolar bone. Since image quality is task-dependant, it is possible for one and the same image to be of both high and low quality, depending on the task. Thus, it is important to link the visual grading analysis to a well-defined diagnostic task. But only in the studies by Åkesson et al. (1989a; 1992b) and in the present study (Ivanauskaite et. al. 2008), the visual grading analysis was directly linked to a diagnostic task, *i.e.*, to score alveolar bone loss or to detect vertical bone defects and furcation involvements.

5.1.6 Assessment of alveolar bone loss

Disproportional effort in research has gone into perfecting the images system compared with the psychological and otological parts of radiological process

(Blesser & Ozonoff 1972; Fryback & Thornbury 1991; Mileman & Kievit 1992). Diagnostic accuracy is influenced by errors originating from the decision-making process involving the intellectual process of recognising the range of normality and classifying the kinds of abnormality (Blesser & Ozonoff 1972). Measurements of the observer agreement are very useful in determining the extent to which the inaccuracy of an imaging system is due to decision-making errors (Swets & Picket 1972). When observer agreement is assessed the truth is not consider. That is way observer agreement serves only as an upper limit of accuracy (Swets & Picket 1972). In the assessment of diagnostic outcome expressed as sensitivity, specificity or predictive value for positive and negative examinations, the observers' reports are scored against the truth. A strategy to improve the diagnostic outcome and the observer agreement is to support the observers in the diagnostic process. Providing observers with written diagnostic cues and reference images, an approach previously adopted by, for example Larsen et al. (1977); Rohlin and Petersson (1989); Cholitgul et al. (1990), is one type of support.

The imaging process is embedded in the clinical process, whereby the clinician uses the information from the imaging process for clinical decision-making. High diagnostic accuracy of an imaging test does not guarantee it will in fact contribute to improved management of patients. Radiographic examinations with no impact on the therapy can not be expected to benefit the patients, but can still be valuable to the clinician by means of reassurance of the treatment plan. More concern should be directed towards studies of the relationship between the diagnostic and therapeutic processes.

For the assessment of the alveolar bone loss relative measurements rather than absolute measurements was chosen in this study. Relative measurements with the aid of a ruler will account for the enlargement of different radiographic methods and of different jaws/teeth imaged by one and the same radiographic method. The ruler designed by Håkansson et al. (1981) was utilized for the assessment of the alveolar bone loss by scores in the studies of Scanora® panoramic radiography as compared to bitewing radiography (Ivanauskaite et al. 2006; 2008). This ruler was designed to be used in radiographs with partial reproduction of the teeth such as bitewing radiographs. As a ruler take the enlargement of the image into account, it was applied for measurements in both panoramic and bitewing radiography, despite different magnifications in the radiographic methods. The scoring of the alveolar bone level, where the cemento-enamel junction (CEJ) was not clearly visible or missing because of restorations or if the bone level presented two edges of the alveolar bone crest, could have influenced the score, the agreement between the radiographic methods and the observer performance. As compared to the rulers using the crown tip and the root apex as reference points (rulers by Schei et al. 1959; Björn et al. 1969), the ruler by Håkansson et al. (1981) has the disadvantage that the CEJ has to be identified.

In the studies by Ivanauskaite et al. (2006; 2008), there were two bitewing radiographs of each side (of premolars and of molars) of each patient, which could result in that the same site or tooth was visible in two bitewing radiographs. The sites to be assessed were not marked in the radiographs. This means that a site or a tooth may have been assessed in different bitewing radiographs (when present in both molar and premolar bitewing radiographs) by different observers and at the repeated observations. This likely decreased the observer agreement for the assessment of the bone loss, the detection of vertical bone defect and furcation involvement.

5.1.7 Systematic review

When the searches in the databases were performed, some limits were not set such as *(i)* author, *(ii)* journal, *(iii)* gender or *(iv)* on subset of journal groups or topics and any default tag terms. Different searches yielded different number of publications, which imply that it is important to describe the search as precise as possible. When possible, MeSH-terms should be implemented in PubMed searches as these terms are computerized and adapted for PubMed data base. Unfortunately the search terms and the Cochrane Library are incongruent. The time frame was set from January 1, 1974 as panoramic radiography was defined as MeSH term in 1974. Studies performed prior to 1974 were assumed to be found in the hand search. However, the hand search of the reference lists is not that precise as the searches in the data bases. When searching in the PubMed data base, publications with search terms not only in the title but in the abstract are retrieved. Although both types *i.e.* the search in the data base and the hand search are dependent of the search terms implemented, the results of the hand search are more influenced as the search terms have to be included in the title. For example, the publication *"Rotational panoramic radiography in epidemiological studies of dental health. Comparison between panoramic radiographs and intraoral full mouth surveys"* (Ahlqwist et al. 1986) was not retrieved even if it would have been relevant in the systematic review. Reference lists of reviews and theses could be beneficial to retrieve additional publication in systematic reviews.

In order to facilitate and control the process of the systematic review, two protocols were implemented: one to include or exclude publications and one for interpretation of data. The protocol for interpretation of data was constructed according to the QUADAS-tool (Whiting et al. 2003). As this protocol was designed for diagnostic accuracy studies elucidating sensitivity, specificity, and predictive values, it was only partly applicable for the interpretation of publications elucidating image quality, concordance between methods, and observer performance.

5.1.8 Statistical methods

When possible and appropriate the values for sensitivity, specificity, and predictive values were calculated. Synonyms for test sensitivity and test specificity are the true-positive rate and true-negative rate, respectively. A test with a high true-positive (and low false-negative) rate is said to be a sensitive test; a test with a low false-positive (and high true-negative) rate is said to be a specific test. The prerequisite for such calculations is that the assessments are expressed as dichotomies. However, in all included studies, except for the one by Molander et al. (1995a), the alveolar bone loss was assessed as a score in relation to the length tooth or root length or as a linear distance in millimeters in relation to the cemento-enamel junction (CEJ). In the studies, where the assessment of the vertical bone defect or furcation involvement was performed, the assessment was mostly expressed as a dichotomy (present or not present). Yet, sensitivity and specificity were only calculated in few studies (Ahlqwist et al. 1986; Molander et al. 1995a; Ivanauskaite et al. 2006).

Observer agreement or observer variation can be expressed as overall agreement, or as kappa, or weighted kappa (Altman 1991). Other authors use intra-class correlation coefficients (ICC), for example when measurements are made in millimetres. In the studies by Ivanauskaite et al. (2006; 2008), the agreement between observers was expressed as both overall agreement and kappa values, which take into account agreement that can be expected to occur by chance (Altman 1991). When it comes to the kappa value, there is no value that can be regarded universally as indicating good agreement. The proposal by Landis and Koch (1977) is an established system, which has been applied for the interpretation of kappa values in several studies of different diagnostic purposes (see for example Panmekiate et al. 1994; Hellén-Halme et al. 2008; Ivanauskaite et al. 2008; Sogur et al. 2009; Kamburoğlu et al. 2010). Comparisons can be made to other studies within the same field. Therefore, this system was applied to interpret values from 0.00 indicating poor agreement to values higher than 0.81 indicating almost perfect agreement.

5.2 Discussion of Results

5.2.1 Visual grading analysis for assessment of alveolar bone loss

A diagnostic test produces uninterpretable results with varying frequencies depending on the test but also on the sample examined and the clinical environment where the test is performed. These problems are reported to a limited extent and in some studies the uninterpretable results are simply removed from the analysis. This may lead to a biased analysis of the outcomes of the diagnostic test. If the uninterpretable results occur randomly and are not related to the true disease status of the individual, *i.e*. alveolar bone loss in this study, these should however, not have an effect on test performance (Whiting et al. 2003). It is important to report the uninterpretable results and their cause so that the impact of these results can be analysed.

It is difficult to compare the results of the reviewed studies on image quality in panoramic radiography, because researchers have used different panoramic machines, receptors, number of observers, and criteria for the assessment. The image quality of different panoramic machines was compared in only one study for the assessment of the alveolar bone (Molander et al. 1995b). Scanora® dental program presented significantly higher mean scores by visual grading analysis than other panoramic machines, while there were no differences between the mean scores of other panoramic machines and programs. This is in accord with the results by Kaeppler et al. (2000a) who presented an evaluation of anatomical structures by Scanora® jaw program and Orthophos Plus®. Both machines were given good ratings and the differences were very small suggesting that both machines are suitable for clinical use (Kaeppler et al. 2000a). In both studies (Molander et al. 1995b; Kaeppler et al. 2000a), mean scores or median scores of the ratings and differences were calculated. The calculation of mean values is, however, questionable as the scores were based on an ordinal scale.

In the reviewed studies on visual image analysis, film/screen combination were used as receptors and in one study (Molander et al. 2004) two storage plate systems were additionally used. In three studies (Kaimenyi & Ashley 1988; Pepelassi & Diamanti-Kipioti 1997; Jenkins et al. 2005), the type of film and screen combination was not described. In the others studies (Stenström et al. 1981; Åkesson et al. 1989a; Molander et al. 1995b; Persson et al. 2003; Molander et al. 2004, Ivanauskaite et al. 2008) different film/screen combination was used. As different film/screen combinations may result in differences in physical parameters, such as resolution, one should be cautious when comparing the results of the different studies. Molander et al. (2004) found when comparing film and storage plate systems for the visualization of the alveolar bone that the receptor systems were equal in about 80 percent of the patients.

The number of observers varied from one observer (Kaimenyi & Ashley 1988; Pepelassi & Diamanti-Kipioti 1997), two observers (Stenstöm et al. 1981), four observers (Jenkins et al. 2005), five or six observers (Åkesson et al. 1989a; Ivanauskaite et al. 2008), seven observers (Molander et al. 1995b) to ten observers (Molander et al. 2004). In one study (Persson et al. 2003), the number of observers assessing the non-readability of the panoramic radiography was unclear.

Criteria to assess image quality for the assessment the alveolar bone loss varied. In most studies (Stenström et al. 1981; Kaimenyi & Ashley 1988; Pepelassi & Diamanti-Kipioti 1997; Persson et al. 2003; Jenkins et al. 2005) two grades were used, while a 3-grade scale was implemented in two studies (Åkesson et al. 1989a; Ivanauskaite et al. 2008). Only in three studies (Åkesson et al. 1989a; Jenkins et al. 2005; Ivanauskaite et al. 2008), criteria for the assessment of image quality were presented, a prerequisite to permit replication of the visual grading analysis. The other authors reported whether the sites were uninterpretable (Stenström et al. 1981), unmeasurable (Kaimenyi & Ashley 1988; Pepelassi & Diamanti-Kipioti 1997) or non-readable (Persson et al. 2003) without any description of the criteria. In two other studies, visualization of the alveolar bone was assessed using a 5-point scale (Molander et al 2004) or a 4-point scale for visualization of the crestal bone (Molander et al. 1995b).

The number of sites with acceptable/adequate and unacceptable/inadequate image quality varied between the reviewed studies. The lowest frequency of unacceptable sites was found to be only about 1 percent (Pepelassi & Diamanti-Kipioti 1997) and the highest 48 percent of the sites (Åkesson et al. 1989a). Differences in sample may partly explain the differences, as only patients with severe periodontitis were examined in the study by Pepelassi and Diamanti-Kipioti (1997) and patients with different severity in the study by Åkesson et al. (1989a). The frequency of mandibular sites rated unacceptable in panoramic radiographs with Scanora® dental program was lower (8%) in the study by Ivanauskaite et al. (2008) than Åkesson et al. (1989a) reported for radiographs taken with Orthopantomograph® Model OP5 (20%–37%). As the radiographs were taken in the same department and similar criteria for the assessment of image quality were applied, the results indicate that Scanora® dental program presented higher image quality than Orthopantomograph® Model OP5. This was the case also for the maxillary sites as the frequency of unacceptable sites were lower with Scanora® dental program (Ivanauskaite et al. 2008) (36%) than in the study of Orthopantomograph® Model OP5 (Åkesson et al. 1989a) (average 42%). The results of Molander et al. (1995) underpin the assumption that the image quality of radiographs taken with the Scanora® dental program (score 3.05) was higher than of radiographs taken with Orthopantomograph® Model OP5 (score 2.69) as well as with other panoramic machines and programs.

The image quality varied in different parts of the dentition, an acceptable panoramic image being obtained more frequently in the lower jaw where panoramic and intraoral radiography was comparable (Stenström et al. 1981; Åkesson et al. 1989a; Pepelassi & Diamanti-Kipioti 1997; Jenkins et al. 2005; Ivanauskaite et al. 2008). Even if Scanora® dental program presented significantly higher image quality in visual grading analysis than other panoramic machines and programs, more than one-third of the sites were rated unacceptable in the maxilla as compared to only 6 percent in bitewing radiography (Ivanauskaite et al. 2008). Yet, there are reasons to believe that the overall image quality of panoramic radiographs of studies included in this review are more optimal than of those performed in general practice. About one third of panoramic radiographs performed in general dental practice was found to be of inadequate quality as compared to 11-14 percent in a department of oral and maxillofacial radiology (Åkesson et al. 1992b).

In only three sites, the image quality for assessment alveolar bone was superior in Scanora® dental panoramic program as compared to posterior bitewing radiography (Ivanauskaite et al. 2008). These sites – all in the lower jaw – were the canine distally, the first premolar mesially, and the second molar distally. A probable explanation for this result is that these sites were not imaged in the bitewing radiographs, despite the overall assessment of image quality (which allowed one retake in cases of poor quality) that was made before the assessments at site level begun. This further underpins that the evaluation of image quality should be directly linked to the diagnostic task.

Unmeasurability was often caused by overlap of teeth or failure to identify the cemento-enamel junction (CEJ) (Pepelassi & Diamanti-Kipioti 1997; Persson et al. 2003; Jenkins et al. 2005). Kaimenyi and Ashley (1988) found that when the ruler by Björn et al. (1969) was used to score the alveolar bone loss, 15 percent of the sites were unmeasurable as compared to 26 percent when the linear distance between the CEJ and the alveolar bone was measured in millimetre for the same sample. It could therefore be discussed whether the CEJ should be used as a reference point in panoramic radiography for example in epidemiological studies. Not only are the difficulties to identify the CEJ a disadvantage of linear measurements in millimetres, but also, the different enlargements of different parts of the panoramic image are in favour of a ruler when assessing the alveolar bone loss in epidemiological studies.

Visibility for the detection of vertical bone defects and furcation involvements in panoramic radiography was only analyzed by Ivanauskaite et al. (2008). The results demonstrated that Scanora® dental panoramic program is suitable for the detection of vertical bone defects in all sites investigated, except for the canine and premolar regions in the upper jaw. For the detection of furcation involvements, visibility was similar in Scanora® dental panoramic program and posterior bitewing radiography. These results are important when Scanora® is to be used to study the prevalence of vertical bone defects and furcation involvements in epidemiological studies.

5.2.2 Assessment of alveolar bone loss

5.2.2.1 Assessment of alveolar bone level

The fact that no meta-analysis could be performed was due to that the studies on the assessment of the alveolar bone level in panoramic radiography differed in design and samples being analysed. In some studies, it was not possible to analyse the design of the study and the execution of panoramic radiography was not described in sufficient detail to permit replication. As is emphasized in QUADAS (Whiting et al. 2003) and relevant for all diagnostic accuracy studies independently whether it is a radiographic method or methods for clinical examinations and laboratory analysis, a sufficient description of the execution of the index test and the reference standard is important. "Firstly, variation in measures of diagnostic accuracy can sometimes be traced back to differences in the execution of the index test or reference standard. Secondly, a clear and detailed description (or citations) is needed to implement a certain test in another setting." When tests/methods are executed in different ways in different studies, it could have an impact on the results.

One factor that is of importance is the enlargement in radiographs. This factor was not presented in three out of four reviewed studies on the *accuracy of measurement of alveolar bone loss compared to surgical measurements* (Kaimenyi & Ashley 1988; Pepelassi & Diamanti-Kipioti 1997; Pepelassi et al. 2000). Furthermore, it was not reported whether the enlargement was accounted for when the measurements of the bone loss were performed in the radiographs. As the enlargement factor varies within one and same image for different regions, the enlargement factor that the manufacturer of the panoramic machine presents could be questioned. The enlargement was evaluated to be 27 percent in the upper jaw and 26 percent in the lower jaw in the fifth study analyzing the accuracy of measurement of bone loss in panoramic radiography with Orthopantomograph Model OP5 compared to surgical measurements (Åkesson et al. 1992a). In bitewing radiography, the enlargement was 9 percent in the upper jaw and differed significantly from that of the lower jaw, being 4 percent (Åkesson et al. 1992a). There was also a difference between the enlargement of the upper jaw (8 percent) and that of the lower jaw (5 percent) in periapical radiography (Åkesson et al. 1992a). In another study of panoramic radiography, it was shown that the enlargement of vertical linear distance was lower for the premolars in the lower jaw (13-15 percent) than for the second premolar and first molar in the upper jaw (17-28 percent) (Thanyakarn et al. 1992). The results by Yitschaky et al. (2004) were in accord with the results by Thanyakarn et al. (1992) demonstrating that the vertical magnification of the premolars in the upper jaw was higher than that of the lower jaw. Thus, absolute linear measurements demand that the enlargements of images are calculated with the aid of some indicator such as grids or steel balls, which are placed in the object plane. In one study on panoramic radiography, where the measurements of alveolar bone loss in millimetres was compared with the criterion reference obtained during surgery, steel balls were used to assess the enlargement factor of radiographs (Åkesson et al. 1992a). The advantage of choosing a ruler is that the ruler is designed to take the enlargement of radiography and enlargement of different regions and teeth into consideration. In a study on accuracy, it would, however, be difficult to compare the radiographic assessment of bone loss made by a ruler with linear measurements made during surgery.

The examined sample differed between the studies of accuracy, in panoramic radiography being only 9 patients in one study (Kaimenyi & Ashley 1988), 21 patients in a second study (Åkesson et al. 1992a) and the same 100 patients in two other studies (Peppelassi & Diamanti-Kipioti 1997; Pepelassi et al. 2000). The number of observers also differed between the studies of accuracy. In only one study (Åkesson et al. 1992a), there were several observers while only one observer made the measurements of the bone loss in the other three studies (Kaimenyi & Ashley 1988; Pepelassi & Diamanti-Kipioti 1997; Pepelassi et al. 2000). As is discussed above and will be discussed below, the number of observers will influence the accuracy of a diagnostic method and the generalisability of the study results.

The overall results of the systematic review on studies on *panoramic radiography with a reference standard obtained during surgery* (Kaimenyi & Ashley 1988; Åkesson et al. 1992a; Pepelassi & Diamanti-Kipioti 1997; Pepelassi et al. 2000) were that panoramic radiography underestimates the degree of bone loss. In cases with severe periodontitis, the mean difference between distances measured in panoramic radiography and during surgery varied between 13-32 percent of the distance, corresponding to 18 percent of the distance in the upper jaw and 24 percent in the lower jaw (Åkesson et al. 1992a). Also in bitewing and periapical radiography, there was an underestimation but panoramic radiography presented a slightly lower mean accuracy than bitewing radiography and lower accuracy than periapical radiography (Åkesson et al. 1992a). The average underestimation of periapical radiography was 13-14 percent of the true distance in one study (Åkesson et al. 1992a), while the corresponding value was reported to be 21 percent of the surgical measurements in another study (Suomi et al. 1968). Corresponding values for underestimation in bitewing radiography was 17 percent in the upper jaw and 22 percent in the lower jaw (Åkesson et al. 1992a). The relative magnitudes of the underestimation reported by Hämmerle et al. (1990) and Renvert et al. (1981) for intraoral radiography were not possible to calculate and therefore not possible to compare with the results of other studies on panoramic radiography. As there are significant differences between the radiographic techniques when assessing bone loss as a linear distance, the techniques should not be used interchangeably in epidemiological studies for sites where linear measurements are needed. The radiographic underestimation depended on the degree of bone loss (Pepelassi & Diamanti-Kipioti 1997): in initial periodontal disease bone loss was underestimated radiographically, in moderate disease the measurements were relatively accurate, and in severe disease bone loss was radiographically overestimated according to Pepelassi and Diamanti-Kipioti (1997). The underestimation also depended on the type of tooth site, with the highest mean deviation between measurements made in panoramic radiography and surgery found in the mandibular anterior teeth (Åkesson et al. 1992). The explanation for the radiographic underestimation of bone loss might be related to that severe bone loss extends as a two-wall defect. In the lower jaw, the remaining buccal or lingual cortical wall is thick and might obscure the image of the most apical bone level. As the reference standard was obtained during surgery, the result of underestimation of bone loss is relevant only for patients in need of surgery.

Sensitivity and specificity using the full mouth survey with 20 intraoral radiographs together with panoramic radiography as the reference standard to detect bone loss was calculated only in one study (Molander et al. 1995a). The sensitivity was high (80%) but the specificity low (35%). One reason for this could be that the threshold for bone loss was set at a distance of only 1.5 millimetres between CEJ and the alveolar bone crest (Molander et al. 1995a).

The reported concordance between assessment of bone loss in panoramic and intraoral radiography is rather high in the reviewed studies. Identical scores of the bone loss in panoramic and intraoral radiography ranged from about 55 percent (Molander et al. 1991; Ivanauskaite et al. 2006) to about 70 percent of the sites (Åkesson et al. 1989b; Rohlin et al. 1989). The different outcomes might partly be due to different designs of the studies. In the studies by Molander et al. (1991) all sites were assessed, while uininterpretable sites were not included in the studies by Åkesson et al. (1989b), Rohlin et al. (1989) and Ivanauskaite et al. (2006). The higher concordance can be reached because measurements were performed only in sites judged as acceptable or excellent. The different results of the study by Ivanauskaite et al. (2006) as compared to the studies by Åkesson et al. (1989b) and Rohlin et al. (1989) may depend on that six observers scored all sites in the first study (Ivanauskaite et al. 2006) whereas each of five observers assessed only one fifth of the material in the two other studies (Åkesson et al. 1989b; Rohlin et al. 1989). Various observers arrive at different results. If the analysis had been based on the observers with the highest agreement, the concordance between the methods would have been 65% of the sites in the study by Ivanauskaite et al. (2006). However, if the observer with the lowest agreement was the only observer in a study, agreement would have been found in only 50% of the sites. The agreement between the methods is very high (94%-95% of the sites) if a difference of one score or 1 millimeter was allowed (Kaimenyi & Ashley 1988; Åkesson et al. 1989b; Rohlin et al. 1989; Ahlqwist et al. 1991; Walsh et al. 1997). For some sites, the largest amount of bone loss was recorded in panoramic radiography (Åkesson et al. 1989b; Rohlin et al. 1989; Ivanauskaite et al. 2006) and for other sites in intraoral radiography (Molander et al. 1991). This should be taken into account when using radiography in epidemiological studies of bone loss. As identified in a systematic review of definitions used in epidemiological studies of periodontal diseases (Savage et al. 2009), only two of the included studies (Norderyd & Huguson 1998; Laurell et al. 2003) applied radiography. In one of these (Norderyd & Huguson 1998), panoramic radiography together with six bitewing radiographs and in the other study (Laurell et al. 2003) full-mouth radiography was used to measure inter-proximal bone height.

The high concordance between panoramic and intraoral radiography can motivate another strategy than the standard recommendation for the initial examination of a patient with chronic periodontitis that for many years was a full-mouth periodontal probing complemented by a full-mouth set of intraoral radiographs. Mol (2004) pointed out in a review that modern panoramic X-ray units are able to produce image layers that resemble the outline of the jaw. However, errors in patient positioning can result in suboptimal projections and limit comparisons with follow-up images. Less image detail, blurred or over-projected structures may compromise the picture of the alveolar bone, especially in the frontal region of the mouth. When one has access to a panoramic machine, a panoramic radiograph could be selected after the initial clinical examination and then supplemented by a limited number of intraoral radiographs, depending on the severity and distribution of increased probing pocket depths, furcation involvements or various non-periodontal findings. Molander et al. (1995a) found that on average 5.1 intraoral radiographs were sufficient to supplement the panoramic radiograph for a comprehensive oral diagnosis. Of these 3.1 contained information different from the panoramic radiograph but 2.0 did not. The most requested radiographs were the bitewing radiographs and periapical radiographs of the canines and incisors. Given that the radiographic examination is preceded by a careful clinical examination, it is possible to select intraoral radiographs more precisely as shown also for the diagnosis of periapical lesions by panoramic radiography (Rohlin & Åkerblom 1992). This emphasizes the importance that a thorough clinical examination should precede every radiographic examination. Panoramic radiographs should not be performed routinely as a screening tool (see for example Rushton et al. 2002a; Rushton et al. 2002b; Jenkins et al. 2005).

5.2.2.2 Detection of vertical bone defect and furcation involvement

The results on *vertical bone defect* from the systematic review demonstrated that panoramic radiography using Ortho Ceph 10 underestimates the number of osseous defects compared to surgical findings (Pepelassi & Diamanti-Kipioti

1997; Pepelassi et al. 2000). Periapical radiography was more accurate detecting 63 percent compared to panoramic radiography detecting only 21 percent of the osseous defects. The deviation of radiographic measurements compared to surgery, as well as the difference between the two radiographic methods, depended on where the defect was located in the jaw, the tooth group and the degree of osseous destruction. The radiographic methods agreed most in the assessment of osseous defects in the severe periodontitis group and least in the initial periodontitis group (Pepelassi et al. 2000). Underestimation of vertical bone defect could be explained by methodological errors and measurement errors. The mean depth of the osseous defects was overestimated by panoramic radiography (Pepelassi et al. 2000), when the depth was small and large and so was the mesiodistal width of the bone defect as compared with surgery. This is difficult to explain, but could depend on that the enlargement of the panoramic radiograph was not compensated for. Results of the study by Pepelassi et al. (2000) also indicated that the ability to detect osseous defects in panoramic radiography was mainly depending on the buccolingual width of the defect. Thus, vertical bone defects of small depth and/or small buccolingual width are the most difficult lesions to detect radiographically.

The results of the study by Ivanauskaite et al. (2006) are to some extent contradictory to the results of Pepelassi et al. (2000) as panoramic radiography presented higher sensitivity than intraoral radiography. One of the differences between the two studies is that Pepelassi et al. (2000) used periapical radiography, while Ivanauskaite et al. (2006) used bitewing radiography. In patients with severe periodontitis it has previously been shown that periapical radiography is more accurate than bitewing radiography (Åkesson et al. 1992a), probably because the interpretation of the bone level is easier in periapical radiographs of sites with deep vertical bone defects compared to bitewing radiography, imaging the bone level more narrowly.

When it comes to concordance between panoramic and intraoral radiography the results of the systematic review reveal rather low concordance in three studies (Åkesson et al. 1989b; Rohlin et al. 1989; Molander et al. 1991) and high concordance in one study (Ivanauskaite et al. 2006). The estimates of a concordance of 42 to 46 percent in the upper jaw and 37 to 53 percent in the lower jaw (Åkesson et al. 1989b; Rohlin et al. 1989; Molander et al. 1991) were based only on the positive diagnoses of vertical bone defects. The high concordance of 89 percent in the upper jaw and 94 percent in the lower jaw (Ivanauskaite et al. 2006) was on the other hand based on all positive and negative diagnoses *i.e.* where the vertical bone defect was present or not. Different results could also be a reflection of improvements in panoramic radiography as Scanora® dental program presented a higher image quality as compared to Orthopantomograph® Model OP5 (Molander et al. 1995b) used in the previous studies on concordance (Åkesson et al. 1989b; Rohlin et al. 1989; Molander et al. 1991).

The results on *furcation involvement* from the systematic review demonstrated that the concordance between panoramic and intraoral radiography varied widely from about half of the teeth (42%; Molander et al. 1991) to two-thirds of the teeth (65-67%; Åkesson et al. 1989b; Rohlin et al. 1989) to almost all teeth (93%; Ivanauskaite et al. 2006). The contradictory results may be due to insufficient standardization of radiographic techniques but may mainly be due to how the concordance was calculated. The high concordance found in the study by Ivanauskaite et al. (2006) was based on all positive and negative diagnoses *i.e.* where a furcation involvement was present or not, while the concordance in the other studies (Åkesson et al. 1989b; Rohlin et al. 1989; Molander et al. 1991) was based only on the positive diagnoses of furcation involvements. The concordance for the detection of furcation involvement was lower in the upper jaw, particularly in the first premolar, where only 15 percent was found as compared to 60 percent in the molars (Molander et al. 1991). Due to the projection of the first premolar in panoramic radiography, the roots and the furcation areas are difficult to identify. The image is influenced not only by the projection geometry but also by the anatomic complexity.

5.2.3 Observer performance

Variability in the processing and reading of a diagnostic examination will affect measures of diagnostic properties. As a diagnostic method will be applied by numerous observers/examiners it is important to evaluate the method with the aid of several observers. Otherwise, it would be difficult to determine whether the findings were the result of one special observer's skill to apply the method under examination. The results of comparisons between diagnostic methods are also observer dependent. The concordance between two methods cannot override the intra-observer agreement of one separate method.

Yet, only one to two observers were asked to participate in most studies on the assessment of bone loss included in the systematic review of panoramic radiography (Kaimenyi & Ashley 1988, Molander et al. 1991; Pepelassi & Diamanti-Kipioti 1997: Pepelassi et al. 2000). In the recent studies by Ivanauskaite et al. (2006; 2008) and in three other studies (Åkesson et al. 1989a; Rohlin et al.1989; Åkesson et al. 1992a), five to six observers were asked to assess the radiographs. When assessing image quality in panoramic radiography, seven oral radiologists were asked to participate in one study (Molander et al. 1995b) and ten observers familiar with panoramic radiography in another study (Molander et al. 2004). Persson et al. (2003) calculated the intra-class correlation coefficients of panoramic and intraoral radiography including three observers. As two observers assessed half of the panoramic radiographs each and a third observer observed the intraoral radiographs, the design when it comes to study observer performance can be questioned.

The results of the systematic review demonstrate that observer performance vary within the same observer and between observers. *Intra-observer agreement* of the different observers in three studies on *visual grading analysis* (Åkesson et al. 1989a; Åkesson et al. 1992b; Ivanauskaite et al. 2008) ranged from moderate to almost perfect agreement, being substantial for most observers. The lowest kappa value for Scanora® dental panoramic program (κ = 0.59) presented by Ivanauskaite et al. (2008) was comparable to intra-observer ratings reported for the same machine (κ = 0.55) by Kaeppler et al. (2006) for assessing anatomic features. Most kappa values presented in the reviewed studies indicated substantial intra-observer agreement for the assessment of image quality in panoramic radiography and the values were higher than those of posterior bitewing radiography (Ivanauskaite et al. 2008).

In the reviewed studies, *intra-observer agreement* for the assessment of *bone loss* in panoramic radiography as expressed as kappa values (Åkesson et al.1989b; Rohlin et al. 1989; Ivanauskaite et al. 2006) ranged between moderate agreement and substantial agreement. Expressed as differences in millimetres between two repeated measurements (Åkesson et al. 1992a) observers are consistent in their assessment of bone loss. The observers' performance in panoramic radiography was comparable to that in intraoral radiography. As shown in the study by Åkesson et al. (1992a), where one observer had the highest intra-observer agreement and another observer the lowest intra-observer agreement, independent of radiographic method, single observers may influence mean intra-observer agreement of a method to a great extent. A fact to take into account when choosing observers and applying radiographic methods for evaluation of treatment outcomes. Another factor to consider is that the way to assess the bone loss can influence the observer's reliability. This was the case when two methods (a digital calliper and an image analysis program) were used to measure the alveolar bone level in children in bitewing radiographs (Pierro et al. 2008). According to the authors, the image program resulted in larger measurements (range 0.11-0.23 mm) compared to the digital caliper.

Inter-observer agreement was lower than intra-observer agreement, in accordance with other studies on observer performance. Inter-observer agreement varied substantially, even though the observers jointly discussed and wrote down assessment criteria in a meeting that was intended to serve as a calibration (Ivanauskaite et al. 2006; 2008). The agreement of one observer in particular with other observers was low, independently if the bone loss was assessed by score (Ivanauskaite et al. 2006) or as a linear distance (Åkesson et al. 1992a). Some weighted kappa values fell into the category poor according to Landis and Koch (1977). This strengthens the fact that variations in the interpretation of radiographs from different machines or made with different imaging techniques may depend more on observer variation than differences in visibility.

That not only the observer performance to assess the alveolar bone level by score but also to detect vertical bone defects and furcation involvements are presented by Ivanauskaite et al. (2006) strengthen the conclusions on panoramic radiography These findings affect the periodontal diagnosis of the tooth and are significant in treatment planning and prognosis. Moreover, when a diagnostic system is implemented, it is necessary to identify not only its technical factors but also other factors, such as observer performance, that influence the outcomes of the actual system. For the detection of bone defects and furcation involvements, intra- and inter-observer agreement was higher than for the assessment of the bone loss by score (Ivanauskaite et al. 2006).

Determining kappa values for observer agreement takes into account agreement that can be expected to occur by chance (Altman 1991). Two observers, or two observations, may emerge with low kappa values despite relatively high overall agreement. When two observers express binary ratings, as for the visibility for detection of vertical bone defects and furcation involvements (Ivanauskaite et al. 2008), 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 values (Feinstein & Cicchetti 1990). Such results often occurred in the visual analysis in the detection of vertical bone defects and furcation involvements in bitewing radiography, as there was an unbalance with a very high frequency of acceptable sites.

When the reproducibility of two different systems is to be compared, it is essential that the trade-off of the expected difference between the systems is pre-determined and the statistical power is adequate. A sample with adequate number of sites should be included and aligned with number of observers. According to Swets and Picket (1982), a number of observers larger than six will have little consequences for the results when a reasonably large sample is examined. When multiple observers have been asked to participate, most observers have been selected from among experts in a specialized field. In the studies by Ivanauskaite et al. (2006; 2008), the observers were selected to represent varying experiences in oral radiology. When assessing periapical pathology the comparison of three groups of observers, five oral radiologists, five endodontists, and five general dental practitioners showed no difference between their diagnostic accuracy when assessing panoramic radiographs (Rohlin et al. 1991). As the diagnostic task was different one can not generalize the results to the assessment of bone loss. Future research on panoramic radiography should include representative groups of observers to find out whether the results are valid for the general dental practitioners as well as for researchers in general.

6. CONCLUSIONS

1. For the assessment of alveolar bone level and the detection of vertical bone defect, the image quality in the upper jaw in the canine and premolar regions panoramic radiography presented a lower image quality than posterior bitewing radiography. For the detection of furcation involvement, the visibility of the radiographic methods was comparable.

2. Inter-observer agreement for several observers as well as for pairs of observers for visual grading analysis for the assessment of alveolar bone level, detection of vertical bone defect and furcation involvement was higher for panoramic radiography than for posterior bitewing radiography. The performance was observer-dependent. This will result in a large variation in assessments.

3. Provided that image quality is excellent or acceptable, panoramic radiography is comparable to posterior bitewing radiography for scoring the alveolar bone level. Agreement between the methods was not influenced by the categories used in visual grading analysis, *excellent* or *acceptable*. For the detection of vertical bone defect and furcation involvement, the sensitivity of panoramic radiography was higher than that of posterior bitewing radiography.

4. Intra-observer agreement as well as inter-observer agreement of several observers for the assessment of alveolar bone level, detection of vertical bone defect and furcation involvement was similar for panoramic radiography and posterior bitewing radiography. The inter-observer

agreement for pairs of observers varied widely for both methods as it was observer-dependent.

5. The systematic review revealed that panoramic radiography underestimated the degree of bone loss as was the case with bitewing and periapical radiography in patients with severe periodontitis. Panoramic radiography presented somewhat lower accuracy than intraoral radiography when surgical measurements comprised the standard reference. The intra-observer agreement seemed to be more influenced by the observers than by the radiographic method. There was no study on the accuracy of panoramic radiography to identify alveolar bone changes over time.

7. RECOMMENDATIONS

7.1 For clinical practice

1. To avoid additional radiation to the patient, a clinical examination should always precede the radiographic examination to identify changes in the periodontal tissue. Bleeding on probing and increased pocked depth are clinical findings that comprise selection criteria for the radiographic examination.

2. Panoramic radiography could be used for the assessment of alveolar bone loss. However, panoramic radiography needs to be supplemented with intraoral radiography when image quality is insufficient or information is needed to identify tissue changes over time.

3. For patients with healthy periodontal tissue or with clinical findings of gingivitis or minor periodontitis, only posterior bitewing radiography is recommended for assessment of the periodontal tissues and caries.

7.2 For research

1. For epidemiological studies, panoramic radiography could be recommended to assess alveolar bone loss. When presenting the prevalence of different entities, the underestimation of bone loss and of number of undetected vertical bone defects and furcation involvements, should be taken in to account.

2. For clinical studies where alveolar bone changes over time are to be assessed, intraoral radiography is recommended.

3. As assessments in panoramic and intraoral radiography are observer-dependant, the observer performance should be analysed and presented in all epidemiological and clinical scientific studies, where the results are based on assessments made in radiographs.

8. REFERENCES

1. Adriaens PA, De Boever J, Vande Velde F. Comparison of intra-oral long-cone paralleling radiographic surveys and orthopantomographs with special reference to the bone height. J Oral Rehabil 1982;9:355-65.

2. Ahlqwist M, Halling A, Hollender L. Rotational panoramic radiography in epidemiological studies of dental health. Comparison between panoramic radiographs and intraoral full mouth surveys. Swed Dent J 1986;10:73-84.

3. Ainamo J, Barmes D, Beagrie G, Cutress, Martin T, Sardo-Infirri J. Development of the World Health Organization (WHO) Community Periodontal Index of Treatment Needs (CPITN). Int Dent J 1982;32:281-91.

4. Ainamo J, Tammisalo EH. The orthopantomogram in quantitative assessment of marginal bone loss. Suom Hammaslääk Toim 1967;63:132-8.

5. Ainamo J, Tammisalo EH. Orthopantomographic diagnosis of advanced marginal bone lesions. Suom Hammaslääk Toim 1968;64:1-5.

6. Åkesson L, Rohlin M, Håkansson J. Marginal bone in periodontal disease: an evaluation of image quality in panoramic and intra-oral radiography. Dentomaxillofac Radiol 1989a;18:105-12.

7. 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 1989b;17:266-71.

8. Åkesson L. Panoramic radiography in the assessment of marginal bone level. Swed Dent J Suppl 1991;78:1-129.

9. 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 1992a;19:326-32.

10. 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 1992b;17:9-21. (in PubMed 1993)

11. Albandar JM, Abbas DK, Waerhaug M, Gjermo P. Comparison between standardized periapical and bitewing radiographs in assessing alveolar bone loss. Com Dent Oral Epidem 1985;3:222-5.

12. Albandar JM, Rams TE. Global epidemiology of periodontal diseases: an overview. Periodontology 2000 2002;29:7–10.

13. Albandar JM. Periodontal disease surveillance. J Periodontol 2007;7:1179-81.

14. Altman DG. Some common problems in medical research. In Altman DG. Practical statistics for medical research. 1st ed. London: Chapman & Hall 1991:396-439.

15. Armitage GC. Development of a classification system for periodontal diseases and conditions. Ann Periodontol 1999;4:1-6.

16. Baelum V, Lopez R. Defining and classifying periodontitis: need for a paradim shift? Eur J Oral Sci 2003;111:2-6.

17. Balčikonytė E. Osteoporozė ir periodonto patologija. Vilniaus universitetas, Daktaro disertacija Bopmedicinos mokslai, Odontologija (08B) 2006. (in Lithuanian)

18. Benn DK. A review of the realibility of radiographic measurements in estimating alveolar bone changes. J Clin Periodontol 1990;17:14-21.

19. Björn AL. Dental health in relation to age and dental care. Odontol Revy 1974;25;Suppl 29. Thesis Lund University, 1974.

20. Björn H, Holmberg K. Radiographic determination of periodontal bone destruction in epidemiological research. Odontol Revy 1966;17:232-50.

21. Björn H, Halling A, Thyberg H. Radiographic assessment of marginal bone loss. Odontol Revy 1969;20:165-79.

22. Blankenstein R, Murray JJ, Lind OP. Prevalence of chronic periodontitis in 13-15-year-old children. A radiographic study. J Clin Periodontol 1978;5:258-92.

23. Blesser B, Ozonoff D. A model for the radiologic process. Radiology 1972;103:515-21.

24. Bolin A, Eliasson S, von Beetzen M, Jansson L. Radiographic evaluation of mandibular posterior implant sites: correlation between panoramic and tomographic determinations. Clin Oral Implants Res 1996;4:354-9.

25. Borg E, Attaelmanan A, Gröndahl H-G. Subjective image quality of solid-state and photostimulable phosphor systems for digital intra-oral radiography. Dentomaxillofac Radiol 2000;29:70-7.

26. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HCW for the STARD group. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Radiology 2003;226:24-8. Available from: http://radiology.rsnajnls.org/cgi/search?sortspec=relevance&fulltext=STARD&author1=Bossuyt&pubdate_year=2003. (Accessed August 16, 2010)

27. Brägger U. Radiographic diagnosis of periodontal disease progression. Curr Opin Periodontol 1996;3:59-67.

28. Brägger U. Radiograpich parameters: biological significance and clinical use. Periodontology 2000 2005;39:73-90.

29. Carmichael FA, Hirschmann PH, Scaife B, Sheard L, Mackenzie A. A comparison of the diagnostic utility of two image receptors for panoramic radiography. Dentomaxillofac Radiol 2000;29:57-60.

30. Casanova MLS, Haiter-Neto F. Effects of developer depletion on image quality of Kodak Insight and Ektaspeed Plus films. Dentomaxillofac Radiol 2004;33:108-13.

31. Caton JG, Greenwell H. Consensus report: dental plaque-induced gingival diseases. Ann Periodontol 1999;4:18-9.

32. Cederberg RA, Frederiksen NL, Benson BW, Shulman JD. Effect of different background lighting conditions on diagnostic performance of digital images and film images. Dentomaxillofacial Radiol 1998;27:293-7.

33. Chaves ES, Caffesse RG, Morrison EC, stults DL. Diagnostic discrimination of bleeding on probing during maintenance periodontal therapy. Am J Dent 1990;3:167:70.

34. Cholitgul W, Petersson A, Rohlin M, Tanimoto K, Åkerman S. Diagnostic outcome and observer performance in sagittal tomography of the temporomandibular joint. Dentomaxillofacial Radiol 1990,19:1-6.

35. Coelho L. Dental image analyzer http://www.ieeta.pt/dia/ (Accessed August 16, 2010).

36. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20:37-46.

37. Cochrane Library. http://www.thecochranelibrary.com/view/0/index.html. (Accessed August 16, 2010)

38. Cooper PG, Caton JG, Polson AM. Cell populations associated with gingival bleeding. J Periodontol 1983;54:497-502.

39. 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.

40. Davenport RH Jr, Simpson DM, Hassell TM. Histometric comparison of active and inactive lesions of advanced periodontitis. J Periodontol 1982;53:285-95.

41. Davies PH, Downer MC, Lennon MA. Periodontal bone loss in English secondary school children. A longitudinal radiological study. J Clin Periodontol 1978;5:278-84.

42. de Souza PH, de Toledo BF, Rapp GE, Zuza EP, Neto CB, Mendes AJ. Reliability of bleeding and non-bleading on probing to gingival histological features. J Int Acad Periodontol 2003;5:71-6.

43. Douglass CW, Valachovic RW, Wijesinha A, Chauncey HH, Kapur KK, McNeil BJ. Clinical efficacy of dental radiography in the detection of dental caries and periodontal disease. Oral Surg Oral Med Oral Pathol 1986;62:330-9.

44. Dula K, Mini R, van der Stelt PF, Buser D. The radiographic assessment of implant patients: decision-making criteria. Int J Oral Maxillofac Implant 2001;16:80-9.

45. Ehrlich J, Weisgold A, Updegrave W. A comparative study of panorex and periapical radiographs for evaluating caries and periodontal disease. Refuat Hapeh Vehashinayim 1977;26:27-9.

46. Eickholz P, Kim TS, Benn DK, Staehle HJ. Validity of radiographic measurements of interproximal bone loss. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1998;85:99-106.

47. Eickholz P, Hausmann E. Accuracy of radiographic assessment of interproximal bone loss in intrabony defects using linear measurements. Eur J Oral Sci 2000;108:70-3.

48. Engelberger A, Rateitschak KH, Marthaler TM. Zur messung des parodontalen Knochenschwundes. Helv Odontol Acta 1963;7:34-8. (In German)

49. Espelid I. The influence of viewing conditions on observer performance in dental radiology. Acta Odontol Scand 1987:45;153-61.

50. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol 1990;43:543-9.

51. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378-82.

52. Flint DJ, Paunovich E, Moore WS, Wofford DT, Hermesh CB. A diagnostic comparison of panoramic and intra-oral radiographs. Oral Surg Oral Med Oral Pathol 1998;85:731-5.

53. Fryback D, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991;11:88-94.

54. Gavala S, Donata C, Tsiklakis K, Boziari A, Kamenopoulou V, Stamatakis HC. Radiation dose reduction in direct digital panoramic radiography. Eur J Radiol 2009;71:42-8.

55. Gijbels F, Jacobs R, Sanderink G, De Smet E, Nowak B, Van Dam J, Van Steenberghe D. A comparison of the effective dose from scanography with periapical radiography. Dentomaxillofacial Radiol 2002;31:159-63.

56. Gijbels F, Sanderink G, Pauwels H, Jacobs R. Subjective image quality of digital panoramic radiographs displayed on monitor and printed on various hardcopy media. Clin Oral Invest 2004;8:25-9.

57. Globienė J. Lietuvos rajonų gyventojų periodonto būklė. Stomatologija. 2001;2:14-6. (In Lithuanian)

58. Goodman C. Literature searching and evidence interpretation for assessing health care practices. Stockholm: The Swedish Council on Technology Assessment in Health Care, 1996.

59. Green ML, Green BL. The vertical periapical bite-wing: a better radiograph for periodontal diagnosis. Dent Hyg (Chic). 1978;52:477-9.

60. Greenstein G, Caton J, Polson AM. Histologic characteristics associated with bleeding after probing and visual signs of inflammation. J Periodontol 1981, 52:420-5.

61. Gröndahl HG, Jönsson E, Lindahl B. Diagnosis of marginal bone destruction with orthopantomography and intra-oral full mouth radiography. Sven Tandlak Tidskr 1971;64:439-46.

62. Guidance Notes for Dental Practitioners on the Safe Use of X-Ray Equipment (2001). http://www.hey.nhs.uk/hullrad/ppt/dentistry/2005/dental_guidance.pdf

63. Haffajee AD, Socransky SS, Lindhe J, Kent RL, Okamoto H, Yoneyama T. Clinical risk indicators for periodontal attachment loss. J Clin Periodontol 1991;18:117-25.

64. Håkansson J, Björn A-L, Jonsson BG. Assessment of the proximal periodontal bone height from radiographs with partial reproduction of the teeth. A methodology study. Swed Dent J 1981;5:147-53.

65. Hämmerle CH, Ingold HP, Lang NP. Evaluation of clinical and radiographic scoring methods before and after initial periodontal therapy. J Clin Periodontol 1990;17:255-63.

66. Hausmann E, Allen K, Clerehugh V. What alveolar crest level on a bite-wing radiographs represents bone loss? J Periodontol 1991;62:570-2.

67. Hausmann E, Allen K, Carpio Christersson LA, Clerehugh V. Computerized methodology for detection of alveolar bone loss from serial intraoral radiographs. J Periodontol 1992;63:657-62.

68. Hellén-Halme K, Petersson A, Warfvinge G, Nilson M. Effect of ambient light and monitor brightness and contrast settings on the detection of approximal caries in digital radiographs: an *in vitro* study. Dentomaxillofac Radiol 2008;37:380-4.

69. Hermann P, Gera I, Borbély J, Fejérdy P, Madléna M. Periodontal health of an adult population in Hungary: findings of a national survey. J Clin Periodontol 2009;36:449-57.

70. Hintze H, Frydenberg M, Wenzel A. Influence of number of surfaces and observers on statistical power in a multiobserver ROC radiographic caries detection study. Caries Res 2003;37:200-6.

71. Hirschmann PN. Radiographic interpretation of chronic periodontitis. Int Dent J 1987;37:3-9.

72. Holtfreter B, Kocher T, Hoffman T, Desvarieux M, Micheelis M. Prevalence of periodontal disease and treatment demands based on a German dental survey (DMS IV). J Clin Periodontol 2010;37:211-9.

73. Hoover JN, Ellegaard B, Attström R. Radiographic and clinical examination of periodontal status of first molars in 15-16-years-old Danish schoolchildren. Scand J Dent Res 1981;89:260-3.

74. Horton PS, Sippy FH, Kerber PE, Paule CL. Analysis of interpretation of full mouth and panoramic surveys. Oral Surg 1977; 44:468-75.

75. Hugoson A, Rylander H. Longitudinal study of periodontal status in individuals aged 15 years in 1973 and 20 years in 1978 in Jönköping, Sweden. Community Dent Oral Epidemiol 1982;10:37-42.

76. Hugoson A, Norderyd O, Slotte C, Thorstensson H. Distribution of periodontal disease in a Swedish adult population 1973, 1983 and 1993. J Clin Period 1998; 25:542-8.

77. Hugoson A, Laurell. A prospective longitudinal study on periodontal bone height changes in a Swedish population. J Clin Periodontol 2000;27:665-74.

78. Hugoson A, Sjödin B, Norderyd O. Trends over 30 years, 1973-2003, in the prevalence and severity of periodontal disease. J Clin Periodontol 2008;35:405-14.

79. ICRP Publication 60: 1990 recommendations of the International Commission on Radiological Protection. Ann ICRP. 1991;21:1-201.

80. ICRP Publication 103: The 2007 Recommendations of the International commission on Radiological Protection. Ann ICRP. 2007;37:1-332.

81. ICRU Report 54. Medical imaging – the assessment of image quality. ICRU publications. Bethesda: 1996.

82. Ingle JI, Van Himel T, Hawrish CE, Glickman GN, Serene T, Rosenberg PA, Buchanan LS, West JD, Ruddle CJ, Camp JH, Roane JB, Cecchini SCM. Endodontic cavity preparation. In Ingle JI, Bakland LK. Endodontics. 5th ed. London: BC Decker Inc Hamilton 2002:405-570.

83. Ivanauskaite D, Lindh C, Rangne K, Rohlin M. Comparison between Scanora® panoramic radiography and bitewing radiography in the assessment of marginal bone tissue. Stomatologija 2006;8:9-15.

84. Ivanauskaite D, Lindh C, Rohlin M. Observer performance based on visibility of the marginal bone tissue in Scanora® panoramic radiography and posterior bitewing radiography. Stomatologija 2008;10:36-43.

85. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-based medicine working group. JAMA 1994a;5:389-91.

86. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? Evidence-based medicine working group. JAMA 1994b;9:703-7.

87. Jeffcoat MK, Wang IC, Reddy MS. Radiographic diagnosis in periodontics. Periodontol 2000 1995;7:54-68.

88. Jenkins WM, Brocklebank LM, Winning SM, Wylupek M, Donaldson A, Strang RM. A comparison of two radiographic assessment protocols for patients with periodontal disease. Br Dent J 2005;198:565-9.

89. Kaimenyi JT, Ashley F. Assessment of bone loss in periodontitis from panoramic radiographs. J Clin Periodontol 1988;15:170-4.

90. Kaeppler G, Axmann-Kremar D, Reuter I, Meyle J, Gómez-Román G. A clinical evaluation of some factors affecting image quality in panoramic radiography. Dentomaxillofac Radiol 2000a;29:81-4.

91. Kaeppler G, Vogel A, Axmann-Kremar D. Intra-oral phosphor and conventional radiography in the assessment of alveolar bone structures. Dentomaxillofac Radiol 2000b;29:362-7.

92. 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.

93. Källestål C, Matsson. Criteria for assessment of interproximal bone loss on bite-wing radiographs in adolescence. J Clin Periodontol 1989;6:300-4.

94. Kamburoğlu K, Senel B, Yüksel SP, Özen T. A comparison of the diagnostic accuracy of *in vivo* and *in vitro* photostimulable phosphor digital images in the detection of occlusal caries lesion. Dentalmaxillofacial Radiol 2010;39:17-22.

95. Kelly M, Steele J, Nuttall N, Bradnock G, Morris J, Nunn J, Pine C, Pitts N, Treasure E, White D. Adult dental health survey oral health in the United Kingdom 1998. London: Office for National Statistics, 2000.

96. Kim TS, Benn DK, Eickholz P. Accuracy of computer-assisted radiographic measurement of interproximal bone loss in vertical bone defects. Oral Surg Oral Med oral Pathol Oral Radiol Endod 2002;94:379-87.

97. Kim TS, Obst C, Zehaczek S, Geen C. Detection of bone loss with different X-ray techniques in periodontal patients. J Periodontol 2008;79:1141-49.

98. Kingman A, Albandar J. Methodological aspects of epidemiological studies of periodontal diseases. Periodontology 2000 2002;29:11-30.

99. Kitagawa H, Farman AG, Scheetz JP, Brown WP, Lewis J, Benefiel M, Kuroyanagi K. Comparison of three intra-oral storage phosphor systems using subjective image quality. Dentomaxillofac Radiol 2000;29:272-6.

100. Kronauer E, Borsa G, Lang NP. Prevalence of incipient juvenile periodontitis at age 16 years in Switzerland. J Clin Periodontol 1986;13:103-8.

101. Krustrup U, Petersen EP. Periodontal conditions in 35-44 and 65-74 year-old adults in Denmark. Acta Odontol Scand 2006;64:65-73.

102. Kullendorff B, Nilsson M, Rohlin M. Diagnostic accuracy of direct digital dental radiography for the detection of periapical bon lesions. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1996;82:344-50.

103. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

104. Lang NP, Joss A, Orsanic T, Gusberti FA, Siegrist BE. Bleeding on probing. A predictor for the progression of periodontal disease? J Clin Periodontol 1986;13:590-6.

105. Lang NP, Adler R, Joss A, Nyman S. Absence of bleeding on probing. An indicator of periodontal stability. J Clin Periodontol 1990;17:714-21.

106. Lanning SK, Best AM, TempleHJ, Richards PS, Carey A, McCauley LK. Accuracy and consistency of radiographic interpretation among clinical instructors using two viewing systems. J Dent Educ 2006a;70:149-59.

107. Lanning SK, Best AM, TempleHJ, Richards PS, Carey A, McCauley LK. Accuracy and consistency of radiographic interpretation among clinical instructors in conjunction with a training program. J Dent Educ 2006b;70:545-57.

108. Larsen A, Dale K, Eek M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. Acta Radiol (Diagn) 1977;18:481-91.

109. Latcham LN, Powell RN, Jago JD, Seymour GJ, Aitken JF. A radiographic study of chronic periodontitis in 15 year old Queensland children. J Clin Periodontol 1983;10:37-45.

110. Laurell L, Romao C, Hugoson A. Longitudinal study on the distribution of proximal sites showing significant bone loss. J Clin Periodontol 2003;30:346-52.

111. Lennon MA, Davies RM. Prevalence and distribution of alveolar bone loss in a population of 15-year-old schoolchildren. J Clin Periodontol 1974:1;175-82.

112. Liedberg J, Panmekiate S, Petersson A, Rohlin M. Evidence-based evaluation of three imaging methods for the temporomandibular disc. Dentomaxillofacial Radiol 1996;25:234-41.

113. Lindhe J, Ranney R, Lamster I. Consensus report: chronic periodontitis. Ann Periodontol 1999;4:38.

114. Löe H, Anerud A, Boysen H, Morrison E. Natural history of periodontal disease in man. Rapid, moderate and no loos of attachment in Sri Lankan laborers 14 to 46 years of age. J Clin Periodontol 1986;13:431-40.

115. Lofthag-Hansen S, Thilander-Klang A, Ekestubbe A, Helmrot E, Gröndahl K. Calculating effective dose on a cone bean computed tomography device: 3D Accuitomo and 3D Accuitomo FPD. Dentomaxillofac Radiol 2008;37:72-9.

116. Ludlow JB, Davies-Ludlow LE, White SC. Patient risk related to common dental radiographic examination: the impact of 2007 International Commission on Radiological Protection recommendations regarding dose calculation. J Am Dent Assoc 2008;139:1237-43.

117. Mackevičienė G, Pūrienė A, Balčiūnienė I. Burnos higiena ir periodonto sveikatos būklė tarp Vilniaus ir Vilniaus krašto gyventojų. Medicina. 1999;4:41–4. (In Lithuanian)

118. Månsson LG. Evaluation of radiographic procedures. Investigation related to chest images. Thesis, Gothenburg University 1994. ISBN 91-628-1212-2.

119. Mann J, Pettigrew J, Beideman R, Green P, Ship I. Investigation of the relationship between clinically detected loss of attachment and radiographic changes in early periodontal disease. J Clin Periodontol 1985;12:247-53.

120. Mariotti A. Dental plaque-induced gingival diseases. Ann Periodontol 1999;4:7-19.

121. Marshall-Day CD, Schourie KL. A roentgenographic survey of periodontal disease in India. J Amer Dent Ass 1949;39:52.

122. Marshall-Day CD, Stephens RG, Quigley LF. Periodontal disease: prevalence and incidence. J Periodontol 1955;26:185-203.

123. Martínez Beneyto Y, Alcaráz Banos M, Pérez Lajarin L, Rushton VE. Clinical justification of dental radiology in adult patient: a review of the literature. Med Oral Patol Oral Cir Bucal 2007;12:E244-51.

124. Menghini G, Steiner M, Helfenstein U, Imfeld C, Brodowski D, Hoyer C, Hofmann B, Furrer R. Imfeld T. Schweiz Monatsschr Zanmed 2002;7:708-17. (In German)

125. Medical Subject Headings (MeSH) data base. http://www.ncbi.nlm.nih.gov/mesh

126. MeSH term “Periodontal Diseases”. http://www.ncbi.nlm.nih.gov/mesh: Periodontal Diseases

127. MeSH term “Observer Variation”. http://www.ncbi.nlm.nih.gov/mesh: Observer Variation

128. MeSH term “Radiography, Panoramic”. http://www.ncbi.nlm.nih.gov/mesh: Radiography, Panoramic

129. Mileman PA, Kievit J. Achieving efficacy in oral radiology-out of the woods, and into decision tree? Dentomaxillofacial Radiol 1992;21:115-7.

130. Michalowicz BS, Aeppli DP, Kuba RK, Bereuter JE, Conry JP, Segal NL, Bouchard TJ, Pihlstrom BL. A twin study of genetic variation in proportional radiographic alveolar bone height. J Dent Res 1991;70:1431-5.

131. Moimaz SA, Zina LG, Saliba O, Garbin CA. Smoking and periodontal disease: clinical evidence for an association.Oral Health Prev Dent 2009;7:369-76.

132. Mol A. Imaging methods in periodontology. Periodontol 2000 2004: 34: 34–48.

133. Molander B, Ahlqwist M, Gröndahl HG, Hollender L. Agreement between panoramic and intra-oral radiography in the assessment of marginal bone height. Dentomaxillofac Radiol 1991;20:155-60.

134. Molander B, Ahlqwist M, Gröndahl HG. Panoramic and restrictive intraoral radiography in comprehensive oral radiographic diagnosis. Eur J Oral Sci 1995a;103:191-8.

135. Molander B, Ahlqwist M, Gröndahl HG. Image quality in panoramic radiography. Dentomaxillofac Radiol 1995b;24:17-22.

136. Molander B, Gröndahl H-G, Ekestubbe A. Quality of film-based and digital panoramic radiography. Dentomaxillofac Radiol 2004;33:32-6.

137. Norderyd O, Huguson A. Tooth loss and periodontal bone level in individuals of Jönköping county. A comparison between two adult population living in the city and in the surrounding area. Swed Dent J 1998;4:165-74.

138. NRPB. Guidelines on radiology standards for primary dental care. Doc NRPB 1994; 5: No 3.

139. Offenbach S, Barros SP, Beck JD. Rethinking periodontal inflammation. J Periodontol 2008; 79:1577-84.

140. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;7:1387-99.

141. Panmekiate S, Petersson A, Rohlin M, Åkerman S. Observer performance and diagnostic outcome in diagnosis of the disk position of the temporamandibular joint using arthrography. Oral Surg Oral Med Oral Pathol 1994;78:255-63.

142. Papapanou P. Epidemiology of periodontal diseases: an update. Journal of the International Academy of Periodontology 1999;4:110-6.

143. Papapanou P, Lindhe J. Epidemiology of periodontal diseases. In Lindhe J, Lang NP, Karring T. (eds) Clinical periodontology and implant dentistry, 5th edition, pp. 129-79. Oxford: Oxford Blackwell Scientific Publication Ltd. 2008.

144. Patel N, Rushton VE. Macfarlane TV, Horner K. The influence of viewing conditions on radiological diagnosis of periapical inflammation. Br Dent J 2000;189:40-2.

145. Peker I, Toraman AM, Uslan G, Altunkaynak B. The comparisson of subjective image quality in conventional and digital panoramic radiography. Indian J Dent Res 2009;20:21-5.

146. 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.

147. Pepelassi EA, Tsiklakis K, Diamanti-Kipioti A. Radiographic detection and assessment of the periodontal endosseous defects. J Clin Periodontol 2000;27:224-30.

148. Persson RE, Tzannetou S, Feloutzis AG, Brägger U, Persson GR, Lang NP. Comparison between panoramic and intra-oral radiographs for the assessment of alveolar bone levels in a periodontal maintenance population. J Clin Periodontol 2003;30:833-9.

149. Pfeifer J, Dean J. The value of panoramic radiography in periodontal diagnosis. J Wis State Dent Soc 1969;45:3-7.

150. Pierro VSS, de Souza IPR, Luiz RR, Barcelos R, Moraes RS. Reliability of two methods for measurements of alveolar bone level in children. Dentomaxillofac Radiol 2008;37:34-9.

151. Potter BJ, Shrout MK, Harrell JC. Reproducibility of beam alignment using different bite-wing radiographic techniques. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1995;79:532-5.

152. Poulsen S, Bille J, Rugg-Gun AJ. Evaluation of a calibration trial to increase interexaminer reliability of radiographic diagnosis of approximal carious lesions. Community Dent Oral Epidemiol 1980;8:135-8.

153. Pūrienė A. Bone status in a group of lithuaninans with untreated periodontitis. Acta Med Lithuanica 1997;2:57-60.

154. Pūrienė A, Balčiūnienė I, Ivanauskaitė D. Oral status of patients with untreated periodontitis. A clinical-radiological analysis. Acta Med Lithuanica 2003;10:95-100.

155. Quality evaluation of dental care. Guidelines for the assessment of clinical quality and professional performance. Radiographs. Los Angeles: California Dent Assoc 1977:13-5.

156. Radiation protection 136. European guidelines on radiation protection in dental radiology. European commission. 2004.

157. Reit C. The influence of observer calibration on radiographic periapical diagnosis. Int End J 1987;20:75-81.

158. Renvert S, Baderstein A, Nilvéus R, Egelberg S. Healing after treatment of periodontal intraosseous defects. I. Comparative study of clinical methods. J Clin Periodontol 1981;8:387:99.

159. Ribeiro-Rotta RF, Lindh C, Rohlin M. Efficacy of clinical methods to assess jawbone tissue prior to and during endosseous dental implant placement: a systematic literature review. Int J Oral Maxillofac Implants 2007;22:289-300.

160. 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.

161. Rohlin M, Petersson A. Rheumatoid arthritis of the temporomandibular joint: radiologic evaluation based on standard reference films. Oral Surg Oral Med Oral Pathol 1989;67:594-9.

162. Rohlin M, Kullendorff B, Ahlqwist M, Stenström B. Observer performance in the assessment of periapical pathology: a comparison of panoramic with periapical radiography. Dentomaxillofac Radiol 1991;20:127-31.

163. Rohlin M, Åkerblom A. Individualized periapical radiography determined by clinical and panoramic examination. Dentomaxillofac Radiol 1992;21:135-41.

164. Rosa GM, Lucas GQ, Lucas ON. Cigarette smoking and alveolar bone in young adults: a study using digitized radiographs. J Periodontol 2008;79:232-44.

165. Rushton VE, Horner K. A comparative study of radiographic quality with periapical techniques in general dental practise. Dentomaxillofac Radiol 1994;23:37-45.

166. Rushton VE, Horner K. The impact of quality control on radiography in general dental practice. Br Dent J 1995:179:254–61.

167. Rushton VE, Horner K. The use of panoramic radiology in dental pactice. J Dent 1996;24:185-201.

168. Rushton VE, Horner K, Worthington HV. The quality of panoramic radiographs in a sample of general dental practices. Br Dent J 1999;26:630-3.

169. Rushton VE, Horner K, Worthington HV. Screening panoramic radiography of new adult patients: diagnostic yield when combined with bitewing radiography and identification of selection criteria. Br Dent J 2002a;192:275-9.

170. Rushton VE, Horner K, Worthington HV. Routine panoramic radiography of new adult patients in general dental practice: relevance of diagnostic yield to treatment and identification of radiographic selection criteria. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2002b;93:448-95.

171. Salonen LW, Frithiof L, Wouters FR, Hellden LB. Marginal bone height in an adult Swedish population. A radiographic cross-sectional epidemiologic study. J Clin Periodontol 1991;18:223-32.

172. Salvolini U. Traumatic injuries: imaging of facial injuries. Eur Radiol 2002;12:1253-61.

173. Savage A, Eaton KA, Moles DR, Needleman I. A systematic review of definitions of periodontitis and methods that have been used to identify this disease. J Clin Periodontol 2009;36:458-67.

174. SBU a. http://www.sbu.se/en/About-SBU/ (Accessed August 16, 2010)

175. SBU b. http://www.sbu.se/en/Assessment-and-Evidence/ (Accessed August 16, 2010)

176. SBU Report 169. Chronic periodontitis – prevention, diagnosis and treatment. A systematic review. Summary and conclusions. Stockholm: 2004.

177. Schei O, Waerhaug J, Lovdal A, Arno A. Alveolar bone loss as related to oral hygiene and age. J Periodontol 1959;30:7-16

178. Schulte W, d'Hoedt B, Lukas D, Maunz M, Steppeler M. Periotest for measuring periodontal characteristics – correlation with periodontal bone loss. J Periodont Res 1992;27:184-90.

179. Selikowitz HS, Sheiham A, Albert D, Willjams GM. Retrospective longitudinal study of the rate of alveolar bone loss in humans using bite-wing radiographs. J Clin Periodontol 1981;8:431-8.

180. Sewerin IB, Andersen V, Stoltze K. Influence of projection angles upon position of cementoenamel junction on radiographs. Scand J Dent Res 1987;95:74-81.

181. Sheiham A, Netuveli GS. Periodontal diseases in Europe. Periodontology 2000 2002;29:104-21.

182. SIS, Swedish standards Institute. Light and lighting – lighting of work places. Part 1: indoor work places. SS-EN 12464-1, 2003 (In English) http://www.sis.se/defaultmain.aspx?tabName=%40Search&SearchString=SS-EN+12464-1 (Accessed August 16, 2010)

183. Skudutytė R, Aleksejūnienė J, Eriksen HM. Periodontal conditions among Lithuanian adults. Acta Medica Lituanica 2001;1:57-62.

184. Skudutyte-Rysstad R, Eriksen HM, Hansen BF. Trends in periodontal health among 35-year-olds in Oslo, 1973-2003. J Clin Periodontol 2007;34:867-72.

185. Soikkonen K, Wolf J, Tenkanen M. Clinical and panoramic assessment of marginal bone loss. A cadaver study. Proc Finn Dent Soc 1990;86:137-41.

186. Sogur E, Baksi BG, Gröndahl H-G, Lomcali G, Sen BH. Detectability of chemically indused periapical lesions by limited cone beam computed tomography, intra-oraldigital and conventional film radiography. Dentomaxillofacial Radiol 2009;38:458-64.

187. Stenström B, Julin P, Lavstedt S. Comparison between panoramic radiographic techniques. Part II: Marginal bone level interpretability with Status-X® and Orthopantomogtraph®, Model OP 3. Dentomaxillofac Radiol 1982;11:37-45.

188. Suomalainen A, Kiljunen T, Käser Y, Peltola J, Kortesniemi M. Dosimetry and image quality of four dental cone beam computed tomography scanners compared with multislice computed tomography scanners. Dentomaxillofac Radiol 2009;38:367-78.

189. Suomi JD, Plumbo J, Barbano JP. A comparative study of radiographs and pocket measurements in periodontal disease evaluation. J Periodontol 1968;39:311-5.

190. Suominen-Taipale L, Nordblad A, Vehkalahti M, Aromaa A. Oral health in the Finish adult population. Health 2000 Survey, Helsinki: National Public Health Institute, 2008.

191. Swets JA, Picket RM. Evaluation of diagnostic systems: methods from signal detection theory. New Your: Academic Press 1982:1-79.

192. Tammisalo E, Hallikainen D, Kanerva H, Tammisalo T. Comprehensive oral X-ray diagnosis: Scanora multimodal radiography. A preliminary description. Dentomaxillofac Radiol 1992;21:9-15.

193. Tammisalo T, Luostarinen T, Vähätalo K, Tammisalo EH. Comparison of periapical and detailed narrow-beam radiography for diagnosis of periapical and detailed narrow-beam radiography for diagnosis of periapical bone lesions. Dentomaxillofacial Radiol 1993;22:183-7.

194. Tammisalo T, Vähätalo K, Luostarinen T, Leimola-Virtanen, Tammisalo EH. Comparison of periapical and detailed narrow-beam radiography for diagnosis of periodontal pathology. Dentomaxillofac Radiol 1994;23:97-101.

195. Tammisalo T, Luostarinen T, Vähätalo K, Tammisalo EH. A comparison of detailed zonography with periapical radiography for the detection of periapical lesions. Dentomaxillofac Radiol 1995a;24:114-20.

196. Tammisalo T, Luostarinen T, Vähätalo K, Rosberg K, Tammisalo EH. Radiographic detectability of periodontal diseases. A comparison of periapical radiography with detailed zonography. Dentomaxillofac Radiol 1995b;24:185-90.

197. Tammisalo T, Luostarinen T, Vähätalo K, Neva M. Detailed tomography of periapical and periodontal lesion. Diagnostic accuracy compared with periapical radiography. Dentomaxillofac Radiol 1996;25:89-96.

198. Tasaki MM, Westesson PL. Temporomandibular joint: diagnostic accuracy with sagittal and coronal MR imaging. Radiology 1993;186:723-9.

199. Teeuw WJ, Coelho L, Silva A, van der Palen CJ, Lessmann FG, van der Velden U, Loos BG. Validation of a dental image analyzer toll to measure alveolar bone loss in periodontitis patient. J Periodon Res 2009;44:94-102.

200. Thanyakarn C, Hansen K, Rohlin M, Åkesson. Measurements of tooth length in panoramic radiographs. 1: The use of indicators. Dentomaxillofac Radiol 1992;21:26-30.

201. 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. ISBN 91-628-4225-0.

202. Tugnait A, Clerehugh V, Hirshmann PN. The usefulness of radiographs in diagnosis and management of periodontal diseases: a review. J Dent 2000a;28:219-6.

203. Tugnait A, Clerehugh DV, Hirshmann PN. Survey of radiographic practices for periodontal disease in UK and Irish dental teaching hospitals. Dentomaxillofac Radiol 2000b;29:376-81.

204. Uotila E, Wolf J. orthopantomographic measurements of alveolar crest height in parodontitis profunda. Odontol Tidskr 1968;76:27-32.

205. Valachovic RW, Douglass CW, Reiskin AB, Chauncey HH, McNeil BJ. The use of panoramic radiography in the evaluation of asymptomatic adult dental patients. Oral Surg Oral Med Oral Pathol 1986;61:289-96.

206. Van der Velden U. Letter to editor. Diagnosis of periodontitis. J Clin Periodontol 2000;27:960-1.

207. 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.

208. Wakoh M, Nishikawa K, Kobayashi N, Farman AG, Kuroyanagi K. Sensitometric properties of Agfa Dentus OrthoLux, Agfa Dentus ST8G, and Kodak Ektavision panoramic radiographic film. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2001;91:244-51.

209. 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.

210. Weinstein MC, Fineberg HV. Clinical decision analysis. Philadelphia London Toronto: W.B. Saunders Company.1980:page 85.

211. Welander U, McDavid WD, Higgins NM, Morris CR. The effect of viewing conditions on the perceptibility of radiographic details. Oral Surg Oral Med Oral Pathol 1983;56:651-4.

212. Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PMM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003;3:25-37.

213. Yitschaky M, Haviv Y, Afarmian DJ, Abed Y, Redlich M. Prediction of premolar tooth lengths based on their panoramic radiographic lengths. Dentomaxillofac Radiol 2004;33:370-2.

214. Zechner W, Watzak G, Gahleitner A, Busenlechner D, Tepper G, Watzek G. Rotational panoramic versus intraoral rectangular radiographs for evaluation of peri-implant bone loss in the anterior atrophic mandible. Int J Oral Maxillofac Implants 2003;18:873-8.

215. Zūbienė J, Klumbienė J, Tamošiūnas A, Milčiuvienė S. Kauno miesto senyvo amžiaus žmonių periodonto ligų paplitimo ir burnos higienos būklės įvertinimas. Visuomenės sveikata 2008;6:2108-14. (in Lithuanian)

9. PUBLICATIONS ON THEME OF DOCTORAL DISSERTATION

1. *Comparison between Scanora® panoramic radiography and bitewing radiography in the assessment of the marginal bone tissue.*

Deimante Ivanauskaite, Christina Lindh, Klara Rangne, and Madeleine Rohlin.
Stomatologija 2006;8:9-15.

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

Deimante Ivanauskaite, Christina Lindh, and Madeleine Rohlin.
Stomatologija 2008;10:36-43.

10. PRESENTATIONS ON THEME OF DOCTORAL DISSERTATION AT CONFERENCES

1. *Panoramic Radiography Using Scanora and Posterior Bitewing Radiography for the Assessment of the Marginal Bone Tissue.*

Ivanauskaite D, Lindh C, Rangne K, Rohlin M.

13th International congress of DentoMaxilloFacial Radiology; Glasgow, UK; August 5-7, 2001.

2. *Assessment of Marginal Bone Tissue in Panoramic Radiographyh – A Systematic Review.*

Rangne K, Ivanauskaite D, Lindh C, Rohlin M.

8th European congress of Dentomaxillofacial radiology; Cracow, Poland; June 6-8, 2002.

3. *Subjective Image Quality in Panoramic Radiography Using Scanora® and Posterior Bitewing Radiography in the Assessment of Marginal Bone Tissue.*

Ivanauskaite D, Lindh C, Rohlin M.

9th European congress of Dentomaxillofacial radiology. Malmo, Sweden; June 17-19, 2004.

4. *Visual Grading Analysis and Assessment of Marginal Bone Tissue in Panoramic Radiography. A Systematic Review.*

Ivanauskaite D, Lindh C, Rohlin M (Sweden).

12th European congress of Dentomaxillofacial radiology; Istanbul, Turkey; June 2-5, 2010.

APPENDICIES

Appendix 1 Protocol *"Observer Instructions"*

Bitewing radiography

Radiographs should be assessed using a light box and a magnifying viewer under subdued lighting. The observer must take a pause after every hour of work.

ALVEOLAR BONE LEVEL

The **image quality for the assessment of the alveolar bone level** is assessed and the **alveolar bone level** is scored with the ruler at the mesial and distal tooth sites noted in the protocol *"Assessment of radiographs"*. The distal bone level of distal roots and the mesial bone level of mesial roots are assessed in multi-rooted teeth.

Image quality is assessed at each site as:

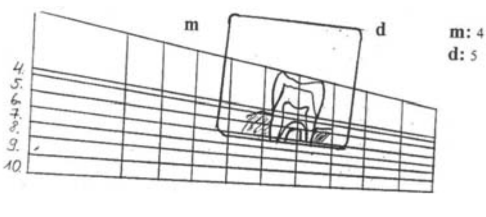
E – excellent	provides necessary information (clear appearance of the cemento-enamel junction (CEJ) with no overlapping, no filling or artificial crown obscuring the CEJ, clear appearance of the alveolar crest and the periodontal ligament space (PDL); good –appropriate density, contrast resolution; correct projection – no image distortion, no overlapping)
A – acceptable	some defects but still acceptable for the purpose of scoring the alveolar bone level
U – unacceptable	does not provide necessary information for the purpose of scoring the alveolar bone level

The reference points are:

- the cusp tip or incisial edge of each teeth
- the CEJ of the mesial and distal site of the tooth
- the alveolar bone at that level where the PDL is considered to have a normal width

How to use the ruler

1. The vertical lines of the ruler are placed parallel to the longitudinal axis of the tooth.
2. The ruler is placed with the coronal reference line over the cusp tip or incisial edge of the crown and the second reference line over the CEJ. Example:



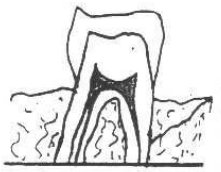
3. The reference line over the alveolar bone, where the PDL is considered to have a normal width, represents the alveolar bone level.
4. If the alveolar bone crest has two edges on the radiograph, the PDL is considered to have a normal width from the lowest.
5. If the CEJ is not clearly visible, the alveolar bone loss is scored from an imaginary point, where the CEJ would have been on the tooth.

Remarks

- If it is not possible to make an assessment because **image quality = U** (unacceptable), mark the score **alveolar bone level** with / in the protocol
- If it is possible to score *i.e.* **image quality = E (**excellent) or **= A** (acceptable) the **alveolar bone level** is scored **4, 5, 6, 7, 8, 9 or 10.**

VERTICAL BONE DEFECT

The **image quality for detection of vertical bone defect** and **detection of vertical bone defect** (vertical bone resorption, three-; two-; one-walled bone defects, see examples) is assessed at the mesial and distal tooth sites noted in the protocol. Examples:



FURCATION INVOLVEMENT

The **image quality for detection of furcation involvement** and **detection of furcation involvement** (a radiolucency between the roots, see example) is assessed for the teeth noted in the protocol. Example:

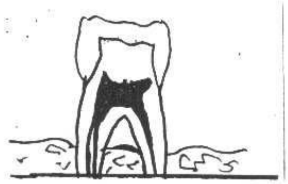


Image quality for **detection of vertical bone defect** or **furcation involvement** is assessed as:

A – acceptable	provides information sufficient to assess tooth sites for detection of vertical bone defect or the bone between the roots for detection of furcation involvement
U – unacceptable	does not provide information sufficient to assess tooth sites for detection of vertical bone defect or the bone between the roots for detection of furcation involvement

Remarks

- If it is not possible to make an assessment because **image quality = U** (unacceptable), mark the **detection for vertical bone defect** or **furcation involvement** with / in the protocol
- If it is possible to assess *i.e.* **image quality = A** (acceptable) mark the **detection of a vertical bone defect** or a **furcation involvement** with **X** in the protocol.

<u>Panoramic radiography</u>

Radiographs should be assessed using a light box and a magnifying viewer under subdued lighting. The observer must take a <u>pause after every hour of work</u>.

ALVEOLAR BONE LEVEL

The **image quality for the assessment of the alveolar bone level** is assessed and the **alveolar bone level** is scored with the ruler at the mesial and distal tooth sites noted in the protocol *"Assessment of radiographs"*. The distal bone level of distal roots and the mesial bone level of mesial roots are assessed in multi-rooted teeth.

Image quality is assessed at each site as:

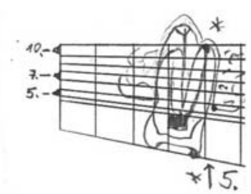
E – excellent	provides necessary information (clear appearance of the cemento-enamel junction (CEJ) with no overlapping, no filling or artificial crown obscuring the CEJ, clear appearance of the alveolar crest and the periodontal ligament space (PDL); good – appropriate density, contrast resolution; correct projection – no image distortion, no overlapping;)
A – acceptable	some defects but still acceptable for the purpose of scoring the alveolar bone level
U – unacceptable	does not provide necessary information for the purpose of scoring the alveolar bone level

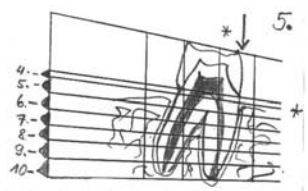
The reference points are:

- the cusp tip or incisial edge of each teeth
- the CEJ of the mesial and distal side of the tooth
- the apex of root in one-rooted tooth; the apex of mesial or distal root of lower jaw in mesial or distal side respectively; the apex of mesiobuccal or distobuccal root of upper jaw in mesial or distal side respectively
- the alveolar bone at that level where the PDL is considered to have radiological normal width

How to use the ruler

6. <u>The vertical lines</u> of the ruler are placed parallel to the longitudinal axis of the tooth.
7. The ruler is placed with <u>the coronal reference line</u> over the cusp tip or incisial edge of the crown and the second <u>reference line over the CEJ</u> and the <u>apical reference line</u> over the apex of root. Examples:





8. <u>The reference line over the alveolar bone</u>, where the PDL is considered a have a normal width, represents the alveolar bone level.

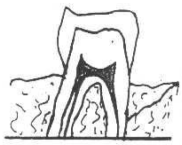
9. If the alveolar bone crest has two edges on the radiograph, the PDL is considered to have a normal width from the lowest.
10. If the CEJ is not clearly visible, the alveolar bone loss is scored from an imaginary point, where the CEJ would have been on the tooth.

Remarks

- If it is not possible to make an assessment because **image quality = U** (unacceptable), mark the score **alveolar bone level** with / in the protocol
- If it is possible to score *i.e.* **image quality = E (**excellent) or **= A** (acceptable) the **alveolar bone level** is scored by **4, 5, 6, 7, 8, 9 or 10.**

VERTICAL BONE DEFECT

The **image quality for detection of vertical bone defect** and **detection of vertical bone defect** (vertical bone resorption, three-; two-; one-walled bone defects, see examples) is assessed at the mesial and distal tooth sites noted in the protocol. Examples:



FURCATION INVOLVEMENT

The **image quality for detection of furcation involvement** and **detection of furcation involvement** (when a radiolucency is observed between the roots, see example) is assessed for the teeth noted in the protocol. Example:

Image quality for **detection of vertical bone defect** and **furcation involvement** is assessed as:

A – acceptable	provides information sufficient to assess tooth sites for detection of vertical bone defect and the bone between the roots for detection of furcation involvement
U – unacceptable	does not provide information sufficient to assess tooth sites for detection of vertical bone defect and the bone between the roots for detection of furcation involvement

Remarks

- If it is not possible to make an assessment because **image quality = U** (unacceptable), mark the **detection for vertical bone defect** and **furcation involvement** mark with / in the protocol

If it is possible to assess *i.e.* **image quality = A** (acceptable) mark the **detection of a vertical bone defect** or a **furcation involvement** with **X** in the protocol.

Appendix 2 Protocol *"Assessment of radiographs"*

Patient group: ….. Patient no: ….. Observer: ….. Observation no: ….. Radiograph(s): …..

Tooth	Site	Alveolar bone level/rezorbtion		Vertical bone defect		Site	Tooth	Furcation involvement	
		Image quality E A U	Score 4 5 6 7 8 9 10	Image quality A U	Detection + or -			Image quality A U	Detection + or -
17	d					d	17		
17	m					m	17		
16	d					d	16		
16	m					m	16		
15	d					d	15		
15	m					m	15		
14	d					d	14		
14	m					m	14		
13	d					d	13		
23	d					d	23		
24	m					m	24		
24	d					d	24		
25	m					m	25		
25	d					d	25		
26	m					m	26		
26	d					d	26		
27	m					m	27		
27	d					d	27		
37	d					d	37		
37	m					m	37		
36	d					d	36		
36	m					m	36		
35	d					d	35		
35	m					m	35		
34	d					d	34		
34	m					m	34		
33	d					d	33		
43	d					d	43		
44	m					m	44		
44	d					d	44		
45	m					m	45		
45	d					d	45		
46	m					m	46		
46	d					d	46		
47	m					m	47		
47	d					d	47		

Appendix 3 Protocol *"Inclusion or exclusion of publications in the systematic review"*

1st author:…………………………………………..Publication number:…………………

Journal:……………………………………..Year: . …Volume:………Pages:……………

Publication type ☐ Primary study ☐ Review ☐ Other …………..………………

Relevance for this review ☐ Yes ☐ No

If No reason for exclusion…………………………………………………………………..

Is there a well-defined aim/purpose? ☐ Yes ☐ No ☐ Cannot tell

My interpretation is:…………………………….………………………………………..

Technique for panoramic radiography:

☐ Cranex ☐ Orthopantomograph ☐ Ortho Ceph 10 ☐ Orthophos ☐ Orthoralix

☐ Panelipse ☐ Planmeca ☐ Scanora, dental ☐ Scanora, jaw

☐ Other ……...

COMPARISON WITH INTRAORAL RADIOGRAPHIC METHOD

☐ ***Bitewing radiography*** ☐ ***Periapical radiography*** ***Other:*** ______________________

ASSESSMENT OF:

☐ Bone loss ☐ Vertical bone defects ☐ Furcation involvements ☐ Image quality

---

OVERALL RESULTS

☐ INCLUDE ☐ EXCLUDE

Data extraction made by:………………………. Date: ………………..

Appendix 4 Protocol based on the QUADAS-tool *"Interpretation of publications in the systematic review"*

First author ……………………………………………………… **Article nr.** …………………………...
Journal **Year** **Volume** **Page**

Are the results of the study valid?	Yes	No	Unclear
1. Was the sample appropriate concerning: Number of subjects ………Number of teeth: ……Number of sites: ……….. Description of selection criteria. Subjects: ……………Sites: ………………. Is the sample representative of those receiving the test in practice?			
2. Was criterion standard used? Yes □ No □ Criterion standard(s): ……………………………………………............ Comparative method(s): ……………………………………………... **Is the criterion standard likely to correctly classify the target condition** (alveolar bone level, bone defects, furcation involvement)?			
3. Did patients receive the same criterion standard/comparative examination regardless of what the panoramic radiograph showed?			
4. Was the panoramic radiograph independent of the criterion standard (i.e. not form part of the criterion standard)?			
5. Was the panoramic radiography described in sufficient detail to permit replication of the test? Panoramic machine………………………… kVp □ mA □ Collimator □ Receptor □ Processing □ Enlargement □			
6. Was the criterion standard/comparative method described in sufficient detail to permit replication of the test? Surgery: Reference points □ Splint □ Probe □ Other…………………….. Execution of measurements and used units □ ……………………………….. Radiography: Method □ X-ray unit □ kVp □ mA □ s □ Collimator □ Focus-skin-distance □ Receptor □ Holder □ Processing □ Execution of measurements and used units □ ……………………………….. Consensus radiographic standard: ……………………………………..			
7. Was the interpretation of panoramic and intraoral images described in sufficient detail concerning: Instruction/calibration of observers concerning diagnostic criteria □ Execution of measurements and used units □……………………………... Number of observers :…………. Settings for observations ………………… Alveolar bone: Alveolar bone level: Reference points: Diagnostic criteria for: □ bone defects □ furcation involvement			
8. Were the results from the panoramic radiography interpreted without knowledge of the result from the criterion standard or comparative method and vice versa?			
9. Were withdrawals (of patients) and uninterpretable test results (images/sites) reported?			
10. Were appropriate results presented and were these calculated appropriately? Percentage of correct diagnosis □ Sensitivity □ Specificity □ Predictive values □ Measures of ROC □ Likelihood ratios □ Other relevant measures……………………………………………………			
Include □ Exclude □ Reason for exclusion…………………………………………………….			
Signature **Date**			

ACKNOWLEDGEMENTS

I would like to express my gratitude to the Swedish Institute for the grant Ref nr 2419/1998 (380/123), that made my, Deimante Ivanauskaite's studies possible, during 1999-2000 in the Department of Oral and Maxillofacial Radiology, Faculty of Odontology, Malmö University, Malmö, Sweden.