Sparseness of the trabecular pattern on dental radiographs: visual assessment compared with semi-automated measurements

¹W G M GERAETS, MSc, PhD, ²C LINDH, DDS, ODONT Dr and ¹H VERHEIJ, MSc, PhD

¹Department of Oral and Maxillofacial Radiology, Academic Centre for Dentistry Amsterdam (ACTA), Amsterdam, Netherlands, and ²Department of Oral Radiology, Faculty of Odontology, Malmö University, Malmö, Sweden

Objective: In diagnostic imaging; human perception is the most prominent, yet least studied, source of error. A better understanding of image perception will help to improve diagnostic performance. This study focuses on the perception of coarseness of trabecular patterns on dental radiographs. Comparison of human vision with machine vision should yield knowledge on human perception.

Method: In a study on identifying osteoporotic patients, dental radiographs were made from 505 post-menopausal women aged 45–70 years. Intra-oral radiographs of the lower and upper jaws were made. Five observers graded the trabecular pattern as dense, sparse or mixed. The five gradings were combined into a single averaged observer score per jaw. The radiographs were scanned and a region of interest (ROI) was indicated on each. The ROIs were processed with image analysis software measuring 25 image features. Pearson correlation and multiple linear regression were used to compare the averaged observer score with the image features.

Results: 14 image features correlated significantly with the observer judgement for both jaws. The strongest correlation was found for the average grey value in the ROI. Other features, describing that osteoporotic patients have fewer but bigger marrow spaces than controls, correlated less with the sparseness of the trabecular pattern than a rather crude measure for structure such as the average grey value.

Conclusion: Human perception of the sparseness of trabecular patterns is based more on average grey values of the ROI than on geometric details within the ROI.

Received 2 February 2011
Revised 8 July 2011
Accepted 20 July 2011

DOI: 10.1259/bjr/32962542

© 2012 The British Institute of
Radiology

Image perception is an important aspect of diagnostic imaging [1, 2]. According to the UNSCEAR 2000 report (Volume 1: Annex D), the average number of diagnostic radiological examinations in countries with Level I healthcare is about 1000 per year per 1000 population. Therefore, it can be estimated that each European has about one radiological examination per year.

The interpretation of radiographs is complicated by the variations in human anatomy and the spatial information that is lost while projecting the patient body on a two-dimensional plane [3]. Visual clues are overlooked or misinterpreted [4–6]. The diagnostic process of radiologists can be improved by the use of computers [7–11]. Pattern-recognition techniques have been designed to draw the attention of the radiologist to regions in mammograms that need careful scrutiny and interpretation [12]. Fully automated methods can screen chest radiographs for features of tuberculosis [13]. Although the results compete with human performance,

the automated methods do not outperform the radiologists. It is expected that computers may eventually replace human observers in the analysis of data. However, complete replacement of the human observer is still a remote possibility [11]. For the foreseeable future, human interpretation will continue to be an inseparable element of medical imaging [14]. We need to understand the images and the technologies used to acquire and display them, but as patient treatment and care depend to a large extent on radiologists interpreting images, we also need to understand human perception and cognition. Many parameters are involved in the process of image acquisition, image processing and image display, and it is largely unknown how they should be optimised for human interpretation. Understanding the perceptual and cognitive processes involved in reading medical images will help to enhance the most useful properties of the images to improve diagnostic performance and reduce error rates [2, 3, 6, 14–17].

In dental radiography, many radiographs show bone with a radiographic trabecular pattern, an irregular meshwork of vague bright lines with fuzzy dark meshes (Figure 1). Visual assessment of the trabecular pattern in intra-oral radiographs is a method to identify women at risk of having osteoporosis. Dense trabeculation is a strong indicator of healthy bone, whereas sparse trabeculation is a sign of osteoporosis [18–20].

Address correspondence to: Dr Wil Geraets, Department of Oral and Maxillofacial Radiology, Academic Centre for Dentistry Amsterdam (ACTA), Gustav Mahlerlaan 3004, 1081 LA, Amsterdam, Netherlands. E-mail: W.Geraets@acta.nl

This work was supported by a research and technological development project grant from the European Commission Fifth Framework Programme “Quality of Life and Management of Living Resources” (QLK6-2002-02243; “Osteodent”).

Figure 1. Radiograph of the right side of the lower jaw with region of interest 3.7×5.8 mm between the first and second premolar. This is used to measure mean and standard deviation of the grey value.

At the oral radiology department of the Academic Centre for Dentistry Amsterdam, methods were developed for semi-automatic analysis of the trabecular pattern of radiographs. Measurements on the trabecular pattern of intra-oral radiographs were found to predict bone mineral density and osteoporosis [21, 22].

When both the visual assessment and the semi-automatic analysis had been applied to the same set of radiographs, there arose an opportunity to compare the two and to gain more insight into the human perception of the coarseness of the radiographic trabecular pattern.

Methods and materials

In 2003, the European Union granted a research project, named OSTEODENT, to five European universities at Manchester, Amsterdam, Athens, Leuven and Malmö.

Subjects and radiographs

In the project, women from Manchester, Athens, Leuven and Malmö were recruited [20, 23]. Local ethical approval for the study was obtained in each recruiting centre and informed consent was obtained from all subjects. From each subject, intra-oral radiographs were made from the upper right and lower right premolar region using one of three Planmeca Prostyle Intra devices (60–63 kV; Planmeca Oy, Helsinki, Finland) or with a Siemens Heliodent MD (60 kV; Sirona, Bensheim, Germany).

The radiographic trabecular pattern was graded by three experienced radiologists and two general practitioners [20]. They were given three reference images from the upper and the lower jaw, and they were asked to classify the trabecular pattern between the roots of the premolars as dense, alternating dense and sparse, or sparse. Additional instructions to observers described dense trabeculation as having many trabeculae connected to each other and small or few marrow spaces. Sparse trabeculation was described as having fewer trabeculae, larger marrow spaces and being darker. Any trabecular pattern that was ambiguous had to be assigned to the intermediate category. Lamina dura, mandibular cortex and maxillary sinus, as well as diseased areas, were excluded from the assessment.

Subjects who had not been graded by all the observers were excluded from the study [20]. Bone mineral density (BMD) values were measured by dual energy X-ray absorptiometry scans of the left hip and the lumbar spine. Complete sets of data, including BMD of hip and spine, two intra-oral radiographs and five observer gradings, were obtained from 505 subjects, of whom 21% were diagnosed as osteoporotic in accordance with the criteria of the World Health Organization if the T-score of the hip or spine was -2.5 or less.

Image processing

The intra-oral radiographs were scanned with a flatbed scanner (Agfa Duoscan T₁200; Agfa Gevaert, Mortsel, Belgium; fixed sensitivity settings) at a resolution of 118 pixels per centimetre (300 dpi). Most radiographs displayed three interdental regions, of which the widest was used by an observer to select a region of interest (ROI) containing a trabecular pattern only (Figure 1). The ROI was subjected to automated image analysis procedures measuring various image features that have proven their relevance for bone structure and osteoporosis extensively [21–27]. Firstly, the mean (MEAN) and standard deviation of the grey values were determined on the raw unfiltered ROI (Figure 1).

Isolated pixels with deviating grey values were adjusted with a 3×3 median filter. Large-scale variations in grey value caused by varying thickness of cortex and soft tissues were removed with an unsharp self-masking filter. Then the ROI was segmented using the mode of the histogram as the threshold value. This resulted in a version of the ROI consisting of black and white segments (Figure 2). The segments were used to measure the fractal dimension according to the caliper method (FRACT), the number of black segments (N_{black}), the number, area and the perimeter of the white segments (N_{white} , BV/TV, BS/TV, respectively) and an index of orientation in horizontal direction (from 0° to 165° in steps of 15° : LFD 0, LFD 15, ..., LFD 165).

Next, the white segments were eroded to a wire frame (Figure 3) that was used to measure the length of the frame, the number of terminal points and the number of nodes (TSL_{white} , N_{white} , $N_{\text{Nd}_{\text{white}}}$, respectively). Similarly, the black segments were eroded to a wire frame that was used to measure the length, the number of terminal points and the number of nodes (TSL_{black} , N_{black} , $N_{\text{Nd}_{\text{black}}}$, respectively). Various methods for

Figure 2. The region of interest in Figure 1 has been filtered and segmented into black and white segments. This is used to measure fractal dimension, numbers of black and white segments, area and perimeter of the white segments and orientation.

filtering and measuring image features have been described before [21, 23–32]. Similar image features are current in studies on osteoporosis and bone structure [33–39].

Statistics

To reduce the variations between the individual gradings and to simplify the analysis, the gradings of the five observers were combined by equating the gradings dense, alternating and sparse with numbers 1, 2 and 3, respectively. For each subject and each jaw the five numbers were combined into a single averaged observer score, resulting in 1010 averaged observer scores pertaining to 505 subjects.

For the interobserver agreement, values of the kappa index ranged from 0.32 (fair) to 0.55 (moderate) [20]. As kappa is the agreement, it can be seen that the disagreement, or noise, in the individual observer is 0.45–0.68. To estimate the noise in the averaged observer score, the factor $1/\sqrt{5}$ is used, leading to a noise level of 0.20–0.30. Obviously 30% of the variation in the averaged observer judgement must be considered noise and 70% of the

Figure 3. The white segments in Figure 2 have been eroded. The eroded parts are displayed in grey. The remaining wire structure is displayed in white. Each white pixel contributes to the length of the white frame (TSL_{white}). Each white pixel with one (or no) white neighbours is an end point and contributes to the number of end points ($N.Tm_{white}$). Each white pixel with three (or four) white neighbours is a node and contributes to the number of nodes ($N.Nd_{white}$).

variation in the averaged observer score is the maximum that can be accounted for by any set of features.

With respect to the image features, it can be said that the semi-automatic measurements were very reproducible as most of the associated values of Cronbach's α exceeded 0.8, and even 0.9 [27].

To test the relation between the averaged observer score and the image features, the Pearson correlation was calculated. In addition, stepwise multiple linear regression was applied to calculate the multiple correlation between the averaged observer score and the image features. Single and multiple correlations were computed with the SPSS package (version 18; SPSS Inc., Chicago, IL). To define significance, $\alpha=0.05$ was used. Additional computations of confidence intervals were done in accordance with Hays [40].

Results

Table 1 summarises the correlations of the image features on the lower and upper jaws with the averaged

observer score. Of the 25 features that were investigated, 14 correlated significantly for both jaws and 6 correlated significantly for only 1 jaw. The difference between the correlation coefficients for the upper and lower jaws was less than the critical value of 0.124 for all features except BS/TV. Allowing 1 in 20 features to differ, this implies that the correlations for the upper and lower jaws correspond.

The highest correspondence is found for the average grey value (MEAN) in the lower jaw (-0.39), as well as in the upper jaw (-0.37). This implies that MEAN accounts for 15% of the variance in the averaged observer score in the lower jaw and 14% in the upper jaw.

Using stepwise multiple linear regression, the percentage variation accounted for n increased to 19% ($r=0.43$) for the upper jaw and 27% ($r=0.52$) for the lower jaw. The analysis started with predictor MEAN and then predictors were added one by one until the prediction improved insignificantly. Table 2 shows the predictors and the order in which they entered the regression equation. Six predictors were entered for the lower jaw and four for the upper jaw. The three most important features in lower and upper jaw corresponded; they were MEAN, BS/TV and LFD 75.

Discussion

The assessments of the trabecular pattern by the observers were similar for the upper and lower jaw [20]. In addition, the semi-automatic measurements of upper and lower jaws correspond to a large extent [29].

This may explain the correspondence of upper and lower jaws in Table 1.

In a previous study, several features discriminated significantly between osteoporotic patients and non-osteoporotic controls, of which the strongest was the $N.Tm_{black}$ [29]. To a large extent this is confirmed by the correlations provided in Table 1. It can be seen that the trabecular patterns of osteoporotic patients have less geometrical detail than the patterns of the non-osteoporotic controls. Obviously, osteoporotic patients have fewer but bigger marrow spaces than non-osteoporotic controls. This is consistent with the finding that sparse trabeculation is a sign of osteoporosis.

The negative value of the correlation between MEAN and the observer score implies that low values of MEAN are associated with sparse trabecular patterns, which is consistent with the loss of bone mineral and increased sparseness of the radiographic trabecular pattern of osteoporotic patients [20, 41]. Considering that MEAN represents the average grey value of the pixels in the ROI, one might infer that MEAN is a crude measure for structure. Therefore, it is striking that in this study MEAN has a stronger correlation with the observer grading than the other features (such as $N.Nd_{black}$ and TSL_{white}) that reflect relevant structural aspects of trabecular microarchitecture. MEAN even surpasses the features N_{black} and BV/TBV , which are closely related with the concept of sparse and dense trabeculation. Part of a possible explanation can be found in the instructions to the observers stating that sparse trabeculation associates with darker images.

Our main conclusion is that the human perception of the sparseness of trabecular patterns is based more on

Table 1. Correlations between observer grading and image features. Osteoporotic patients tended to have high observer gradings, low grey values and few details. Healthy controls tended to have low observer gradings, high grey values and many details

Feature	Code	Upper jaw	Lower jaw
Average grey value	MEAN	-0.37	-0.39
Standard deviation of grey value	SD	-0.10	-0.13
Fractal dimension	FRACT	$+0.16$	$+0.09$
Number of white segments	N_{white}	$+0.15$	$+0.11$
Number of black segments	N_{black}	-0.18	-0.18
Area of white segments	BV/TV	-0.19	-0.26
Perimeter white segments	BS/TV	-0.16	-0.32
Orientation horizontal	LFD 0	ns	$+0.12$
Orientation along 15°	LFD 15	ns	ns
Orientation along 30°	LFD 30	ns	ns
Orientation along 45°	LFD 45	ns	-0.13
Orientation along 60°	LFD 60	-0.14	-0.18
Orientation along 75°	LFD 75	-0.15	-0.19
Orientation vertical	LFD 90	-0.13	-0.16
Orientation along 105°	LFD 105	-0.10	-0.15
Orientation along 120°	LFD 120	-0.11	ns
Orientation along 135°	LFD 135	ns	ns
Orientation along 150°	LFD 150	ns	$+0.11$
Orientation along 165°	LFD 165	ns	$+0.11$
Length white wire	TSL_{white}	-0.21	-0.32
End points white wire	$N.Tm_{white}$	ns	ns
Nodes white wire	$N.Nd_{white}$	-0.14	-0.19
Length black wire	TSL_{black}	ns	ns
End points black wire	$N.Tm_{black}$	-0.19	-0.22
Nodes black wire	$N.Nd_{black}$	-0.09	ns

ns, not significant.

Table 2. Results of stepwise multiple linear regression

Upper jaw			
Order in regression	Feature in upper jaw	Code	Variance accounted for
1	Average grey value	MEAN	14%
2	Perimeter white segments	BS/TV	16%
3	Orientation along 75°	LFD 75	18%
4	Standard deviation of grey value	SD	19%
Lower jaw			
Order in regression	Feature in lower jaw	Code	Variance accounted for
1	Average grey value	MEAN	15%
2	Perimeter white segments	BS/TV	21%
3	Orientation along 75°	LFD 75	23%
4	Length of white wire	TSL _{white}	25%
5	Orientation horizontal	LFD 0	26%
6	Nodes white wire	N.Nd _{white}	27%

Stepwise multiple linear regression first selects the single feature best describing the observer grading. The second feature is selected to increase the descriptive power the most. This is continued as long as significant improvement can be achieved.

the average grey value of the ROI than on other structural aspects of the ROI.

Acknowledgments

The reported work was carried out in Amsterdam and Malmö. Thanks to Professor Paul Van der Stelt (Academic Centre for Dentistry Amsterdam, The Netherlands), Professor Keith Horner, Dr Hugh Devlin, Professor Judith Adams and Dr Elisabeth Marjanovic (University of Manchester, United Kingdom), Professor Reinhilde Jacobs (KU Leuven, Belgium) and Professor Kety Nicopoulou-Karayianni (University of Athens, Greece) for organising the project, recruiting patients and making the radiographs.

References

- Bankman IN. Improvement of visual perception. In: Bankman IN, ed. *Handbook of medical imaging processing and analysis*. Burlington, MA: Academic Press; 2000. p. 788.
- Krupinski EA, Kundel HL. Update on long-term goals for medical image perception research. *Acad Radiol* 1998;5: 629–33.
- Krupinski EA. The future of image perception in radiology: synergy between humans and computers. *Acad Radiol* 2003;10:1–3.
- Gale A, Krupinsky EA, Manning DJ. Medical image perception society conference XI. *Br J Radiol* 2006;79: S109–10.
- Van Dijk JAM, Verbeek LM, Hendriks JHCL, Holland R. The current detectability of breast cancer in a mammographic screening program. *Cancer* 1993;72:1933–8.
- Krupinski EA. Medical image perception issues for PACS deployment. *Semin Roentgenol* 2003;38:231–43.
- Kundel HL, Nodine CF, Krupinski EA. Computer-displayed eye position as a visual aid to pulmonary nodule recognition. *Invest Radiol* 1990;25:890–6.
- Nakamura K, Yoshida H, Engelmann R, MacMahon H, Katsuragawa S, Ishida T, et al. Computerized analysis of the likelihood of malignancy in solitary pulmonary nodules with use of artificial neural networks. *Radiology* 2000;214: 823–30.
- Floyd CE, Lo JY, Yun AJ, Sullivan DC, Kornguth PJ. Prediction of breast cancer malignancy using an artificial neural network. *Cancer* 1994;74:2944–8.
- Baker JA, Kornguth PJ, LO JY, Floyd CE. Artificial neural network: improving the quality of breast biopsy recommendations. *Radiology* 1996;198:131–5.
- Robinson PJA. Radiology's Achilles' heel: error and variation in the interpretation of the Röntgen image. *Br J Radiol* 1997;70:1085–98.
- Te Brake GM, Karssemaier JHCL. Automated detection of breast cancer carcinomas not detected in a screening program. *Radiology* 1998;207:465–71.
- Van Ginneken B, Katsuragawa S, Ter Haar Romeny BM, Doi K, Viergever MA. Automatic detection of abnormalities in chest radiographs using local texture analysis. *IEEE Trans Med Imaging* 2002;21:139–48.
- Samei E. Why medical image perception? *J Am Coll Radiol* 2006;3:400–1.
- Hendee WR, Wells PNT. Visual perception as an opportunity for radiologic research. *Invest Radiol* 1989;24:575–6.
- Manning DJ, Gale A, Krupinsky EA. Perception research in medical imaging. *Br J Radiol* 2005;78:683–5.
- Madsen MT, Berbaum KS, Ellingson AN, Thompson BH, Mullan BF, Caldwell RT. A new software tool for removing, storing and adding abnormalities to medical images for perception research studies. *Acad Radiol* 2006;13:305–12.
- Jonasson G, Bankvall G, Kiliaridis S. Estimation of skeletal bone mineral density by means of the trabecular pattern of the alveolar bone, its interdental thickness, and the bone mass of the mandible. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2001;92:346–52.
- Lindh C, Petersson A, Rohlin M. Assessment of the trabecular pattern before endosseous implant treatment. Diagnostic outcome of periapical radiography in the mandible. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1996;82:335–43.
- Lindh C, Horner K, Jonasson G, Olsson P, Rohlin M, Jacobs R, et al. The use of visual assessment of dental radiographs for identifying women at risk of having osteoporosis: the OSTEODENT project. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2008;106:285–93.
- Geraets WGM, Verheij JGC, Van der Stelt PF, Horner K, Lindh C, Nicopolou-Karayianni K, et al. Prediction of bone mineral density with dental radiographs. *Bone* 2007;40: 1217–21.
- Verheij JGC, Geraets WGM, Van der Stelt PF, Horner K, Lindh C, Nicopoulou-Karayianni K, et al. Prediction of

- osteoporosis with dental radiographs and age. *Dentomaxillofac Radiol* 2009;38:431–7.
23. Geraets WGM, Verheij JGC, Van der Stelt PF, Horner K, Lindh C, Nicopoulou-Karayianni K, et al. Selecting regions of interest on intraoral radiographs for the prediction of bone mineral density. *Dentomaxillofac Radiol* 2008;37:375–9.
 24. Geraets WGM, Van der Stelt PF, Elders PJM. The radiographic trabecular bone pattern during menopause. *Bone* 1993;14:859–64.
 25. Geraets WGM, Van der Stelt PF, Lips P, Van Ginkel FC. The radiographic trabecular pattern of hips in patients with hip fractures and in elderly control subjects. *Bone* 1998;22:165–73.
 26. Geraets WGM. Comparison of two methods for measuring orientation. *Bone* 1998;23:383–8.
 27. Geraets WGM, Verheij JGC, Van der Stelt PF, Horner K, Lindh C, Karayianni K, et al. Osteoporosis and the general dental practitioner: reliability of some digital dental radiological measures. *Community Dent Oral Epidemiol* 2007;35:465–71.
 28. Lee BD, White SC. Age and trabecular features of alveolar bone associated with osteoporosis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2005;100:92–8.
 29. White SC, Rudolph DJ. Alterations of the trabecular pattern of the jaws in patients with osteoporosis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1999;88:628–35.
 30. White SC, Cohen JM, Mourshed FA. Digital analysis of trabecular pattern in jaws of patients with sickle cell anemia. *Dentomaxillofac Radiol* 2000;29:119–24.
 31. White SC. Oral radiographic predictors of osteoporosis. *Dentomaxillofac Radiol* 2002;31:84–92.
 32. White SC, Atchison KA, Gornbein JA, Nattiv A, Paganini-Hill A, Service SK, et al. Change in mandibular trabecular pattern and hip fracture rate in elderly women. *Dentomaxillofac Radiol* 2005;34:168–74.
 33. Cendre E, Kaftandjian V, Peix G, Jourlin M, Mitton D, Babot D. An investigation of segmentation methods and texture analysis applied to tomographic images of human vertebral cancellous bone. *J Microsc* 2000;197:305–16.
 34. Chen H, Zhou X, Shoumura S, Emura S, Bunai Y. Age- and gender-dependent changes in three-dimensional microstructure of cortical and trabecular bone at the human femoral neck. *Osteoporos Int* 2010;21:627–36.
 35. Homminga J, McCreddie BR, Weinans H, Huiskes R. The dependence on the elastic properties of osteoporotic cancellous bone on volume fraction and fabric. *J Biomech* 2003;36:1461–7.
 36. Licks R, Licks V, Ourique F, Radke Bittencourt H, Fontanella V. Development of a prediction tool for low bone mass based on clinical data and periapical radiography. *Dentomaxillofac Radiol* 2010;39:224–30.
 37. Rizzoli R. Microarchitecture in focus. *Osteoporos Int* 2010;21:S403–6.
 38. Thomsen JS, Laib A, Koller B, Prohaska S, Mosekilde L, Gowin W. Stereological measures of trabecular bone structure: comparison of 3D micro computed tomography with 2D histological sections in human proximal tibial bone biopsies. *J Microsc* 2005;218:171–9.
 39. Zhang ZM, Li ZC, Jiang SD, Dai LY. Micro-CT and mechanical evaluation of subchondral trabecular bone structure between postmenopausal women with osteoarthritis and osteoporosis. *Osteoporos Int* 2010;21:1383–90.
 40. Hays WL. *Statistics*. 4th edn. Austin, TX: Holt, Rinehart and Winston Inc; 1988.
 41. Nackaerts O, Jacobs R, Devlin H, Pavitt S, Bleyen E, Yan B, et al. Osteoporosis detection using intraoral densitometry. *Dentomaxillofac Radiol* 2008;37:282–7.