

## Automated osteoporosis risk assessment by dentists: A new pathway to diagnosis

H. Devlin<sup>a,\*</sup>, P.D. Allen<sup>b</sup>, J. Graham<sup>b</sup>, R. Jacobs<sup>d</sup>, K. Karayianni<sup>c</sup>, C. Lindh<sup>e</sup>,  
P.F. van der Stelt<sup>f</sup>, E. Harrison<sup>b</sup>, J.E. Adams<sup>b</sup>, S. Pavitt<sup>a</sup>, K. Horner<sup>a</sup>

<sup>a</sup> School of Dentistry, University Dental Hospital, Higher Cambridge Street, Manchester, M15 6FH, UK

<sup>b</sup> Imaging Science and Biomedical Engineering, University of Manchester, UK

<sup>c</sup> Dental School, University of Athens, Greece

<sup>d</sup> Oral Imaging Centre, School of Dentistry, Oral Pathology and Maxillofacial Surgery, Katholieke Universiteit Leuven, Belgium

<sup>e</sup> Faculty of Odontology, Malmö University, Sweden

<sup>f</sup> Academic Centre for Dentistry, Amsterdam, The Netherlands

Received 28 May 2006; revised 29 September 2006; accepted 29 October 2006

Available online 22 December 2006

### Abstract

General dental practitioners use a vast amount of panoramic radiography in their routine clinical work, but valuable information about patients' osteoporotic status is not collected. There are many reasons for this, but one of the prime reasons must be the disruption involved in clinical routine with lengthy manual radiographic assessment. We have developed computer software, based on active shape modeling that will automatically detect the mandibular cortex on panoramic radiographs, and then measure its width. Automatic or semi-automatic measurement of the cortical width will indicate the osteoporotic risk of the patient. The aim of our work was to assess the computer search technique's ability to measure the mandibular cortical width and to assess its potential for detection of osteoporosis of the hip, spine and femoral neck.

Mandibular cortical width was measured using the manually initialized (semi-automatic) method and, when assessed for diagnosing osteoporosis at one of the three measurement sites, gave an area under the ROC curve ( $A_z$ )=0.816 (95% CI=0.784 to 0.845) and for the automatically initialized searches,  $A_z$ =0.759 (95% CI=0.724 to 0.791). The difference between areas=0.057 (95% Confidence interval=0.025 to 0.089),  $p<0.0001$ . For diagnosing osteoporosis at the femoral neck, mandibular cortical width derived from the manually initialized fit gave an area under the ROC curve ( $A_z$ )=0.835 (95% CI=0.805 to 0.863) and for the automatically initialized searches  $A_z$ =0.805 (95% CI=0.773 to 0.835). The difference in  $A_z$  values between active shape modeling search methods=0.030 (95% CI=-0.010 to 0.070), and this was not significant,  $p=0.138$ .

We concluded that measurement of mandibular cortical width using active shape modeling is capable of diagnosing skeletal osteoporosis with good diagnostic ability and repeatability.

© 2006 Elsevier Inc. All rights reserved.

**Keywords:** Osteoporosis; Active shape modeling; Risk assessment; Mandible; Radiography

### Introduction

Mandibular cortical width on dental panoramic radiographs is significantly correlated with bone mineral density at the hip [1], lumbar spine [2] and forearm [3], the most common sites of fracture related to osteoporosis in post-menopausal women. Measuring mandibular cortical width could be used for diag-

nosis as a screening tool for osteoporosis. Taguchi et al. [2] found that mandibular cortical thickness was related to the bone mineral density of the third lumbar vertebra. Devlin and Horner [3] found that mandibular cortical width had moderate accuracy when used to diagnose skeletal osteopenia. Subsequent work [4] advised that a cortical thickness of less than 3 mm in the mental foramen region should be a trigger for referral for dual energy X-ray absorptiometry (DXA).

While DXA facilities are often limited, millions of dental panoramic radiographs are taken every year across Europe. A recent study based in the United Kingdom, showed that 61% of

\* Corresponding author. Fax: +44 161 275 6480.

E-mail address: [Hugh.Devlin@manchester.ac.uk](mailto:Hugh.Devlin@manchester.ac.uk) (H. Devlin).

general dental practitioners used panoramic radiography equipment [5]. Measurements of mandibular cortical width from them may prove to be a cost-effective, efficient triage method of selecting those patients at high-risk of osteoporosis [4].

One important barrier to using cortical width measurements in primary dental care is the significant observer variability in measurement that is not improved by individualized instruction [6]. Furthermore, manual measurement of cortical width may be seen as a time-consuming interruption by the dentist in their busy schedule.

We have developed computer software, based on active shape modeling [7], that will automatically detect the mandibular cortex on panoramic radiographs and then measure its width. Active shape modeling is a technique widely used in computer vision to detect shapes and analyze them and has been used successfully to replicate the shape of vertebrae [8] and accurately detect the edge shape of bone in digitized radiographs [9]. Once the mandibular cortex has been detected using the active shape model, multiple measurements of width and further analysis of the endosteal border become possible with minimal user interaction.

In 2003, the 3-year OSTEODENT project was commenced, consisting of collaboration by five European centers to investigate the role of dental radiographs in the diagnosis of osteoporosis. The overriding aim of this project was to identify the most valid and effective radiographic index, or combination of radiographic and clinical indices, for the diagnosis of osteoporosis applicable for use by dentists. The aim of the work reported here was to assess the computer search technique's ability to measure the mandibular cortical width and to assess its potential for detection of osteoporosis of the hip, spine and femoral neck.

## Method

Ethical approval was given for the recruitment of female subjects (aged 45–70 years) following their informed consent. The study was open to all female patients in this age-group, except those who suspected that they might be pregnant. No one was excluded from recruitment based on race or pre-existing medical condition such as secondary osteoporosis. With Ethics Committee approval, those who have previously had a bone density scan performed and were identified as having a below average bone density were recruited into the study. Recruitment of osteoporotic individuals was encouraged to provide a sufficient sample size with narrow confidence intervals around both sensitivity and specificity values of the diagnostic tests.

We compared the diagnostic ability of clinical risk indices with that of the computer radiographic measurements. Two well established indices were chosen, that of the National Osteoporosis Foundation (NOF) index [10] and the Osteoporosis Risk Assessment Index (ORAI) [11]. The NOF index scores 1 point for each of the following: patient is >65 years, weight <57.6 kg, maternal/paternal history of fracture, current cigarette smoking, and a personal history of fracture. ORAI used the following subject scoring system: age >75 years (+15), age 65–74 years (+9), age 55–64 (+5), body weight <60 kg (+9), body weight 60–70 kg (+3), estrogen therapy (+2). The total score for each patient was calculated for the two indices.

### *Central dual energy X-ray absorptiometry (DXA)*

Dual energy X-ray absorptiometry (DXA) scans were performed on the Hologic QDR 4500, Hologic Discovery (Hologic Inc., Bedford, Massachusetts,

USA) and the GE Lunar Prodigy (GE Lunar Corporation, Madison, Wisconsin) at four centers throughout Europe. The four centers used were located in Leuven (Belgium), Athens (Greece), Manchester (UK), and Malmö (Sweden) and ambulant female patients were recruited from the area surrounding these centers. Stewart's rules were used to monitor quality assurance throughout the study period [12].

The European spine phantom was used to standardize measurements between different manufacturers using the method described by Pearson and colleagues [13]. *T* and *Z* scores were calculated using Hologic reference data for the lumbar spine and NHANES reference data for the proximal femur [14].

Patients were diagnosed as osteoporotic according to the World Health Organization (WHO) criteria, i.e. those with a bone mineral density *T*-score value 2.5 S.D. or more below the mean value of the young sex matched reference population.

## Radiography

The subjects received a dental panoramic radiograph examination while biting on a plastic block in the left premolar region. The plastic block contained a spherical, steel ball bearing (3.175 mm diameter), which was used to compensate for the image magnification.

Digital and conventionally processed dental panoramic radiography machines were used. Leuven (Belgium) and Malmö (Sweden) used a Cranex III (Soredex, IL, USA) whereas Athens (Greece) and Manchester (UK) used a Planmeca (Planmeca USA, Roselle, IL, USA). The imaging parameters also varied but typically were 70 kV at 8 mA for 15 s. In Leuven, ADC Solo (Afga, Mortsel, Belgium) was used as the photostimulable phosphor plate system for image capture and digital read out, but other centers used a conventional film/cassette.

### *Point distribution model*

The radiographs were digitized using a Kodak LS85 digitizer (Eastman Kodak, Rochester, NY) at a resolution of 25.64 pixels/mm. Using a previous training set of 132 DPR images, a point distribution model (PDM) [7] of the inferior mandibular cortex was created by manual annotation of the endosteal and periosteal borders. Two experts performed this task independently using a graphical user interface, outlining the inferior mandibular cortex by placing equally spaced points on the computer images between the mental foramen and antegonial region. The mean point position of both experts was used to define the shape of the cortex, and the PDM built using 200 points interpolating between the manually placed points. The point distribution model was used to search for and identify the inferior mandibular cortex. Principal component analysis applied to the covariance matrix of the point position allows the main “modes of variation” of the target shape to be determined. The point distribution model captured the principal modes of variation of the shape of the inferior mandibular cortex. These modes of variation were manipulated in an image search program to find the region of the image whose shape was contained within the observed shape distribution of the training data, and which provided the best fit to the expected image appearance. The point distribution model was then used to locate the inferior mandibular cortex in the OSTEODENT

sample, a set of dental panoramic radiographs which had not been included in the training set. A search strategy was used in which, from a given starting shape and pose, the point distribution model iteratively deformed in an attempt to align the points with the strongest image gradient (edge) found within a predefined search area around each point. This combination of point distribution model and search mechanism is known as an active shape model.

The initial stage of the active shape model search was to locate just the lower border of the inferior mandibular cortex using a point distribution model built only from points corresponding to the lower border. This is because the periosteal border of the inferior mandibular cortex is a much more clearly defined feature than the upper endosteal border. The final stage of the active shape model search used a point distribution model of the endosteal and periosteal borders of the inferior mandibular cortex.

The points defining the inferior mandibular cortex were positioned on a border which was well defined in a direction perpendicular to the edge, but not along it. In other words whilst we can be reasonably confident that the endosteal and periosteal borders of the inferior mandibular cortex were correctly defined in the active shape model search, the correct anatomical positioning of the points with respect to the mental foramen and antegonial regions could not be guaranteed.

Thus, two search strategies were investigated: a constrained search using manually placed initialization points, and an unconstrained search with no manual initialization. In the constrained search, only four points were planted by the user on the periosteal border of the inferior mandibular cortex at the left and right mental foramina and antegonial regions. The point distribution model of the mean shape was then warped so that its mental foramina and antegonial points matched those placed by the user, and this was used as the search start point with no further reference to the initialization points being used during the search. Accurate placement of the initialization points by the user with respect to the mandibular cortex was not required since the border was located by the active shape model search. The unconstrained, active shape model search was completely automatic and started from the mean shape and pose found in the training set—in this case a multi-resolution coarse-to-fine search strategy was required for robust search results. The endosteal border was not defined by the user in either the constrained or unconstrained searches, but was located by the active shape model search.

After convergence of the search, the mandibular cortical width was measured at a series of contiguous locations along the lower border of the mandible between the antegonion and the mental foramina for all subjects. The measured width at each location was averaged over an interval of approximately 10% of the length of the lateral cortical border. For each of these locations the correlation was calculated between the measured cortical width and the skeletal bone mineral density ground-truth. Fig. 1(a) shows the correlation values as a function of location along the mandible. Statistically significant correlations were found in the lateral region of the inferior mandibular cortex, the highest correlation occurring at 10–20% of the

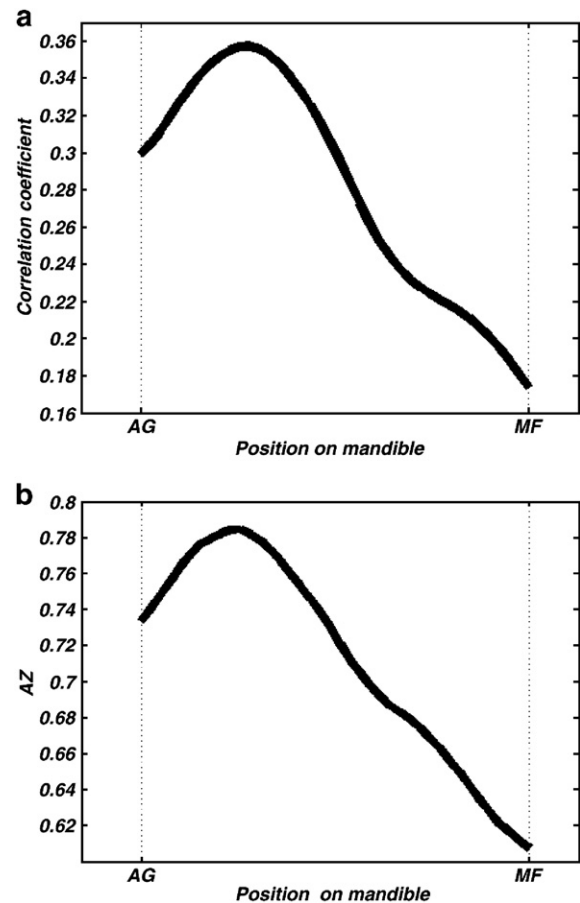


Fig. 1. From the manually initialized ASM fit, the mandibular cortical thickness was calculated between the antegonial and mental regions. The cortical width measurements from the lateral region of the mandible were (a) most highly correlated with the minimum  $T$ -score derived from all three BMD measurement sites or (b) provided the highest  $A_z$  values in ROC analysis. In (a), the correlation coefficient significance levels are 0.098 and 0.069 at the  $p=0.01$  and  $p=0.05$  levels, respectively; with all our data points comfortably exceeding the 1% threshold. AG=antegonial position, MF=mental foramen position.

distance from the antegonion to the mental foramen. Fig. 1(b) shows a similar curve plotting the  $A_z$  values arising from ROC analysis. Maximum sensitivity is obtained for measurements made in the same region. The “optimal measurement region” defined by this measurement is indicated in Fig. 2. Results reported from the manually-initialized and fully automatic searches all refer to measurements made in this region.

The width of the ball bearing image was used to scale the linear cortical width measurements in the constrained and unconstrained fits. On each of the digital images the position of the ball bearing was marked manually. The image was then cropped around this position to a size larger than the expected size of the ball bearing. A Canny edge detector [15] was then used to detect the edges in this cropped region and a Hough transform [16] used to isolate those edges belonging to an ellipsoidal object. From the detected elliptical image, the dimensions of the ball bearing were then calculated. The measured dimensions of the ball-bearing were used to scale the width measurements made at different centers.

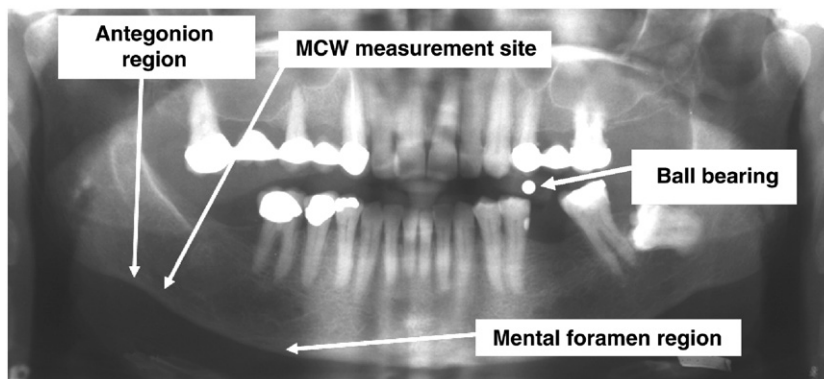


Fig. 2. Dental panoramic radiograph (DPR) showing antegonial and mental regions, between which the measurements of mandibular cortical width were made. The ball-bearing used for inter-center calibration is visible on the left side of the mandible.

### Statistical analysis

ROC curve analysis was used to measure the diagnostic abilities of the computed measurements of cortical width in diagnosis of osteoporosis. In this respect, separate analyses were performed for a diagnosis of osteoporosis at any measured site (lumbar spine, femoral neck or total hip) and at the femoral neck alone. The areas under ROC curves ( $A_z$ ) were calculated using the Medcalc® software program (MedCalc Software, Mariakerke, Belgium).

### Repeatability of automated methods

It was not possible in our study to measure repeatability of the cortical width measurements on radiographs obtained from multiple exposures of the same patient for ethical reasons. The robustness of the technique was tested by repeated measurement of the same radiographic sample set.

#### Automatically initialized search

The same model applied to the same digital image produced identical searches each time, so the reproducibility error was zero in this case.

#### Manually initialized search

There is a source of variability in manually initialized search, as four initialization points need to be specified interactively. The variability arising from doing this was simulated by perturbing each of the four initialization points by a set distance in a random direction and repeated for 10 searches. The size of the perturbation was calculated from the two sets of manual mark-ups and corresponded to the mean distance between the two sets of four points placed manually at the mental foramen and antegonial positions. The mean within subject variance was calculated for the 10 searches using one-way ANOVA. The repeatability is the difference between two measurements for the same subject and is expected to be less than 2.77 times the within-subject standard deviation for 95% of pairs of observations.

### Results

671 subjects were recruited in total. 10 subjects were eliminated from the study because 8 were outside the age range of the inclusion criteria (45–70 years) and bone mineral density data was incomplete on 2. The radiographs of a further 9 subjects were unsuitable for further analysis due to unacceptable quality, accidental loss, data corruption or the absence of a ball bearing image and were eliminated from the study. Of the remaining 652 subjects that formed the study population, 140 had osteoporosis at one of the three measurement sites and 65 had osteoporosis at the hip. The mean age of the subjects was 54.9 years (S.D. = 6.10).

For the manually and automatically initialized searches, there were significant differences between the mandibular cortical widths of normal subjects and those with osteoporosis at one of the three measurements sites (Table 1; using Mann–Whitney  $U$  test,  $p < 0.0001$  for both searches).

There were significant correlations between the mandibular cortical widths derived from both computer image searches and bone mineral density at the total hip, spine and femoral neck (Table 2).

ROC curves were plotted of cortical width derived from manually initialized searches for diagnosing osteoporosis at one of the three sites (Fig. 3) or at the femoral neck (Fig. 4). The manually initialized searches gave higher  $A_z$  values for the ROC curves than the automatically initialized searches (Figs. 5 and 6).

Table 1  
Mean mandibular cortical width (MCW) and standard deviation (S.D.) derived from automatically and manually initialized searches for osteoporotic and normal individuals

|                               | <i>N</i>                           | Mean MCW | S.D.  |
|-------------------------------|------------------------------------|----------|-------|
| Manually initialized fit      | Normal                             | 3.747    | 0.596 |
|                               | osteoporosis at any of three sites | 3.031    | 0.552 |
| Automatically initialized fit | Normal                             | 3.778    | 0.681 |
|                               | osteoporosis at any of three sites | 3.117    | 0.682 |

Table 2

Mandibular cortical width derived from automatically and manually initialized search correlated with bone mineral density measured at the total hip, spine and femoral neck using dual X-ray energy absorptiometry

|                                  | Spearman's rho          | Manually initialized search | Automatically initialized search |
|----------------------------------|-------------------------|-----------------------------|----------------------------------|
| Automatically initialized search | Correlation coefficient | 0.722                       | 1.000                            |
|                                  | Sig. (2-tailed)         | $p < 0.001$                 |                                  |
| Total hip                        | Correlation coefficient | 0.399                       | 0.328                            |
|                                  | Sig. (2-tailed)         | $p < 0.001$                 | $p < 0.001$                      |
| Femoral neck                     | Correlation coefficient | 0.460                       | 0.376                            |
|                                  | Sig. (2-tailed)         | $p < 0.001$                 | $p < 0.001$                      |
| Lumbar spine                     | Correlation coefficient | 0.433                       | 0.359                            |
|                                  | Sig. (2-tailed)         | $p < 0.001$                 | $p < 0.001$                      |

### Comparison of methods

The difference in area ( $A_z$ ) under the ROC curves was used to identify the most effective method for diagnosis of osteoporosis in any site. Mandibular cortical width derived from the manually initialized fits gave an area under the ROC curve ( $A_z$ ) of 0.816 (95% CI=0.784 to 0.845) and for the automatically initialized searches,  $A_z$  was 0.759 (95% CI=0.724 to 0.791). The manually initialized search model had a significantly greater  $A_z$  than the automatically initialized search model ( $A_z$  difference=0.057; 95% CI 0.025 to 0.089,  $p=0.0001$ ).

For diagnosis of osteoporosis at the femoral neck, mandibular cortical width derived from the manually initialized fit gave an area under the ROC curve ( $A_z$ )=0.835 (95% CI=0.805 to 0.863) and for the automatically initialized searches  $A_z$ =0.805 (95% CI=0.773 to 0.835). There was no significant difference

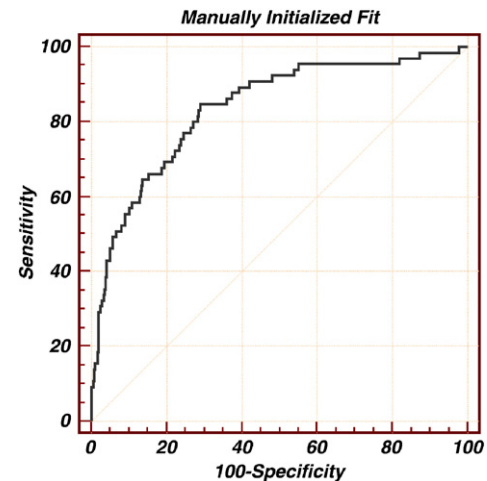


Fig. 4. Receiver Operating Characteristic curve for the measurements of cortical width obtained by the manually initialized method in the diagnosis of osteoporosis at the femoral neck.  $A_z$ =0.835 (95% CI=0.805 to 0.863).

in  $A_z$  between the two methods ( $A_z$  difference=0.03, 95% CI −0.009 to 0.070,  $p=0.135$ ).

There was a significant correlation between manually initialized and automatically initialized search models (Spearman's  $\rho$ =0.722,  $p < 0.0001$ , 95% CI=0.683 to 0.757). A Passing & Bablok plot [17] was used to compare the manually initialized and automatic search results (Fig. 7). The 95% confidence intervals of the slope and intercept were used to determine significant differences from 1 and 0, respectively.

Manual Search ( $Y$ )=0.30+0.90 Automatic Search ( $X$ ).

Intercept=0.30 (95% CI: 0.13 to 0.47) and Slope=0.90 (95% CI: 0.85 to 0.95).

There was no significant deviation in linearity between the automatic or manually initialized search methods ( $p > 0.10$ , using the Cusum test).

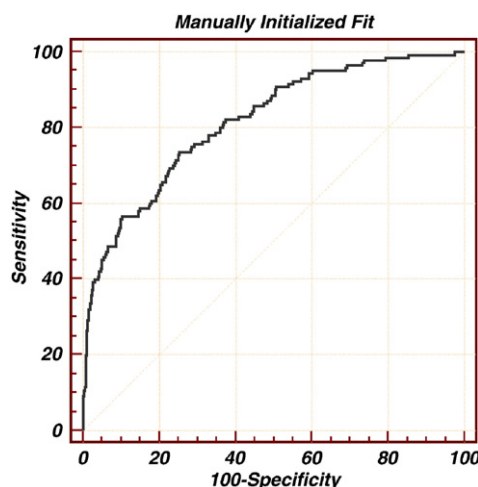


Fig. 3. Receiver Operating Characteristic curve for the measurements of cortical width obtained by the manually initialized method in the diagnosis of osteoporosis at any site (lumbar spine, femoral neck or total hip).  $A_z$ =0.816 (95% CI=0.784 to 0.845).

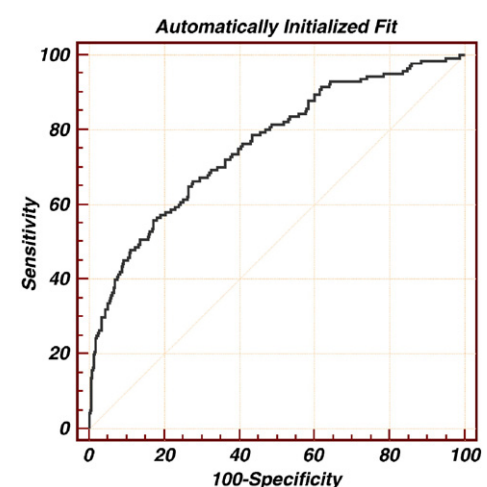


Fig. 5. Receiver Operating Characteristic curve for the measurements of cortical width obtained by the automatically initialized method in the diagnosis of osteoporosis at any site (lumbar spine, femoral neck or total hip).  $A_z$ =0.759 (95% CI=0.724 to 0.791).

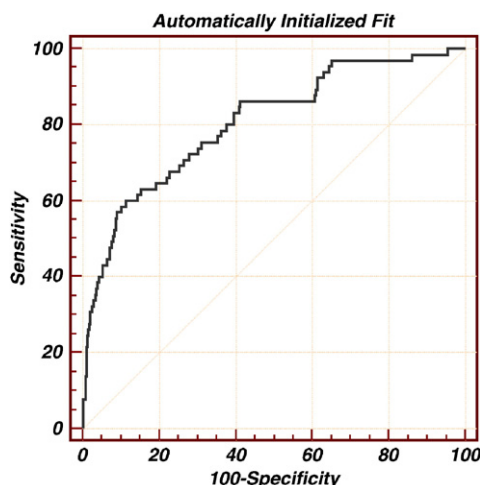


Fig. 6. Receiver Operating Characteristic curve for the measurements of cortical width obtained by the automatically initialized method in the diagnosis of osteoporosis at the femoral neck.  $A_z=0.805$  (95% CI=0.773 to 0.835).

#### Clinical indices

Table 3 summarizes the area under the ROC curve for the radiographic measurements and the clinical indices in diagnosing osteoporosis. ORAI performed significantly better than the NOF index at detecting osteoporosis at the femoral neck ( $p=0.001$ ) and at any one of the three measurement sites ( $p<0.001$ ). For detecting osteoporosis at the femoral neck there was no significant difference in  $A_z$  values between the manually initialized fit and ORAI ( $p=0.431$ ), and between the automatically initialized fit and ORAI ( $p=0.109$ ). For detecting osteoporosis at one of the three sites, there was no difference in  $A_z$  values between the manually initialized fit and ORAI ( $p=0.641$ ) and between the automatically initialized fit and ORAI ( $p=0.135$ ).

#### Numbers of patients detected

If a  $<3$  mm threshold was applied to the manually initialized search data, then 119 patients would have “failed” the radio-

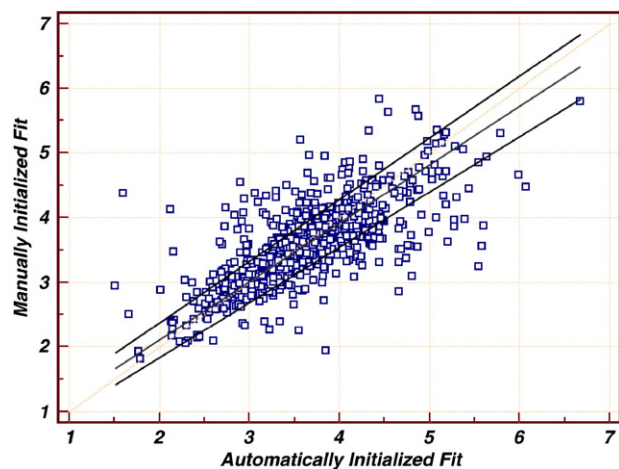


Fig. 7. A Passing & Bablok plot to compare the manually initialized and automatic search results. Linear regression line (with 95% CI).

Table 3

Summary table comparing radiographic assessment by computer with the clinical indices National Osteoporosis Foundation (NOF) index and the Osteoporosis Risk Assessment Index (ORAI)

|                               | AUC for osteoporosis at the femoral neck $A_z$ | AUC for osteoporosis at any of three sites $A_z$ |
|-------------------------------|------------------------------------------------|--------------------------------------------------|
| Manually initialized fit      | 0.835                                          | 0.816                                            |
| Automatically initialized fit | 0.805                                          | 0.759                                            |
| ORAI                          | 0.861                                          | 0.803                                            |
| NOF                           | 0.732                                          | 0.671                                            |

AUC=area under the ROC curve.

graphic test and been referred out of a total population of 652. Of these 119 referred patients, 72 were found to have osteoporosis at one site as measured using DXA. The probability that the referred patients had osteoporosis, given that they had failed the radiography test, was therefore 60.5%. This contrasts with the prior probability of osteoporosis in our study population of 21.5%. Using the automatically initialized search data with a cut-off threshold of  $<3$  mm, 126 patients would have been referred with osteoporosis, of whom 67 were found to have osteoporosis using DXA. The probability that any of these referred patients had osteoporosis, given that they had failed the radiographic test was 53.2%, over double the prior probability of osteoporosis. A high threshold of  $<4$  mm with the manually initialized search gave an excellent sensitivity of 96.4% (but poor specificity of 29.5%).

#### Repeatability of manually initialized search

For all 652 subjects, the mean mandibular cortical width was 3.59 mm (range 1.81 to 5.83 mm). The mean within-subject variance for the subsequent 10 manually initialized searches (corrected for image magnification) was 0.062 mm, giving a measurement error of 0.25 mm and repeatability of 0.69 mm. This error range indicates that the difference between two measurements for the same subject is expected to be less than 0.69 mm for 95% of pairs of observations.

#### Discussion

Mass screening for the detection of osteoporosis is not recommended as cost-effective. Instead a cheaper method of detecting those at high risk of osteoporotic fracture is desirable. Such a method should require minimal input from the clinician to be cost-effective. Active shape modeling (ASM) search has been used for robust location of anatomical features in a number of medical imaging studies [18]. Of particular interest to the current study is the location and measurement of the shape of the femur in radiographs [19], and detecting the edges of bone in digitized radiographs [9]. In the present study, the automated image analysis software performed in an equivalent manner to that reported previously with manual measurements by experts [20]. Manually initialized ASM measurement of cortical width produced an  $A_z=0.816$  in ROC curve analysis that was high for detecting osteoporosis at either the total hip, femoral neck and lumbar spine and for the automatically initialized searches,

$A_z=0.759$ . Taguchi et al. [20] found that the  $A_z$  for their expert manual radiographic measurement was 0.802 (95% CI, 0.705 to 0.899) in detecting osteoporosis. With our image analysis software, we anticipate that the automated computer measurement would alert the examining dentist to the patient's thin mandibular cortex. The dentist could follow up by an examination of the patient's risk factors (low body mass index, age, steroid use etc) and then consider the advisability of further referral for central dual energy X-ray absorptiometry. Our analytical software has a comparable diagnostic ability in detecting osteoporosis to clinical indices such as ORAI, as given by the area under the ROC curve.

Using the manually initialized fit for predicting osteoporosis at the hip,  $A_z$  was 0.835, and for the automatically initialized fit was 0.805 indicating these provided good diagnostic tests in predicting osteoporosis at this site.

The manually initialized fit used defined points on the dental panoramic radiograph in the medio-lateral direction placed by the observer, which limited the search along the mandible. The automatically initialized fit was unable to use any edge features on the image with which to establish correct anatomical placement of the points with respect to the antegonion and mental foramen regions. In other words, there were no features along the edge of the mandible which could be used to define the position of the points in the fully automated search. The mean values for cortical width measured using the automatically initialized search were greater than for the manually initialized search (Fig. 7 and Table 1), resulting in reduced sensitivity. This resulted in consistently greater cortical width values for the automatically initialized fit than the manually initialized fit in Fig. 6, with a slope significantly different from unity. The mean values for cortical width measured using the automatically initialized search were greater than for the manually initialized search (Table 1). The manually initialized search strategy provided better  $A_z$  values than the automatic search. At the appropriate operating point, both sensitivity (true positive fraction of those with osteoporosis) and specificity (true negative rate) were improved, with corresponding reductions in false positives and false negatives. For maximum diagnostic accuracy, some minimal observer interaction is therefore necessary.

Given the less than perfect diagnostic accuracy of the cortical width measurements, the dental panoramic radiograph would not be taken for osteoporosis screening *per se*, but some unrelated dental investigation. In addition, the patient's case history and medical data must be considered before undertaking further referral and investigation. Radiographic measurements cannot be used as the sole basis for referral. With these limitations, our computer methods of osteoporosis triage could be cost-effective as the assessment is performed automatically on digital films, with minimal intervention from the dentist. Setting the threshold for further investigation to achieve a low percentage of false positive diagnoses would be possible by considering the high specificity end of the ROC curve. Our future investigations will consider the precise threshold required to minimize false positive diagnosis but yet still provide a good diagnostic test. Manual measurements of mandibular cortical

width by general dental practitioners have poor repeatability [6], but this problem would be eliminated using the computer-based technology we have described.

The repeatability of the manually initialized search was better than that previously described for the manual measurements of different experts. Devlin and Horner [3] found that the limits of agreement, which indicate the interval in which 95% of measurement differences lie, were +1.32 mm for manual measurements of cortical width. Larger limits of agreement were found when the manual cortical width measurements were made by general dental practitioners [6].

We have previously recommended that patients with a mandibular cortical width of <3 mm should be referred for further DXA investigation. That recommendation was on the basis of our manual measurements of cortical width [4]. The higher sensitivity and specificity and better reproducibility that arises from the digital analysis would, of course, result in an improved analysis at this threshold. Of more importance is to consider the proportion of women labeled "osteoporotic" by this test who are not truly osteoporotic (as measured by DXA). Taguchi et al. [20] found that 60% of their patients who had a cortical width of less than 3 mm were osteoporotic. Threshold values for manual measurements on dental radiographs should be chosen which balance sensitivity and specificity for the prevailing health care systems. Using a high threshold for mandibular cortical width, such as 4.785 mm below which patients are classed as osteoporotic [21], will result in excellent sensitivity, but poor specificity. In the environment seen in many European countries of inadequate availability of DXA [22,23], a low sensitivity/high specificity strategy may be more appropriate, at least where DXA availability is less than the minimum recommended 8 units per million population.

Digital analysis increases the diagnostic yield of radiographs [24]. Gregory et al. [19] developed an active shape model of the femur. They found that the gross morphology of the femur could be used to identify patients who may develop a hip fracture in the future. Despite differences in the positioning of patients and femur in their study, as well as variable magnification of the images, they found that their active shape model was more robust than other methods. We have compensated for magnification errors by asking the patient to bite on a plastic block incorporating a ball bearing of 3.175 mm diameter during the radiographic exposure. Magnification in the DPR is about 20–36% for most machines. Another computer-aided diagnostic technique [25] also used panoramic radiographs to provide osteoporotic pre-screening, but required considerable operator input e.g. to correctly identify the mental foramen. This is a potential weakness in view of the non-visibility of the mental foramen in a minority of patients [26] and the low intra-examiner agreement in localizing the mental foramen on DPRs [27]. Arifin et al. [25] described a semi-interactive method based on image processing in which the cortex was distinguished by thresholding, high-pass filtering and morphological operations to enhance the image. They showed that this method delivered measurements of similar diagnostic value to manual measurement. Their measurements were restricted to the region of the mental foramen, and required

input from an expert user in defining the position of the mental foramen.

In conclusion, active shape modeling is capable of automatic mandibular morphometry of dental panoramic radiographs, producing a diagnostic test of skeletal osteoporosis that is comparable to that achieved by manual measurement. Our technique does have some shortcomings and sources of error, but it requires minimal interaction by the clinician and provides automated warning if the patient is at high risk of osteoporosis. Future work will use further automated image analysis of the morphological features of the cortex to improve the diagnostic accuracy of our methodology. In particular, we believe our methodology has further potential for development where automated detection of low bone density may be beneficial such as implantology, and assessment of the effect of osteoporosis on fracture healing, tooth loss and periodontitis [28].

## Acknowledgments

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

## References

- [1] White SC, Taguchi A, Kao D, Wu S, Service SK, Yoon D, et al. Clinical and panoramic predictors of femur bone mineral density. *Osteoporos Int* 2005;16:339–46.
- [2] Taguchi A, Tanimoto K, Suei Y, Ohama K, Wada T. Relationship between the mandibular and lumbar vertebral bone mineral density at different postmenopausal stages. *Dentomaxillofac Radiol* 1996;25:130–5.
- [3] Devlin H, Horner K. Mandibular radiomorphometric indices in the diagnosis of reduced skeletal bone mineral density. *Osteoporos Int* 2002;13:373–8.
- [4] Horner K, Devlin H, Harvey L. Detecting patients with low skeletal bone mass. *J Dent* 2002;30:171–5.
- [5] Tugnait ACV, Hirschmann PN. Radiographic equipment and techniques used in general dental practice: a survey of general dental practitioners in England and Wales. *J Dent* 2003;31:197–203.
- [6] Devlin CV, Horner K, Devlin H. Variability in measurement of radiomorphometric indices by general dental practitioners. *Dentomaxillofac Radiol* 2001;30:120–5.
- [7] Cootes TF, Cooper DH, Graham J. Active shape models—Their training and application. *Comput Vis Image Underst* 1995;61:38–59.
- [8] Smyth PP, Adams JE. Vertebral shape: automatic measurement with active shape models. *Radiology* 1999;211:571–8.
- [9] Behiels G, Vandermeulen D, Suetens P. Evaluation of image features and search strategies for segmentation of bone structures in radiographs using active shape models. *Med Image Anal* 2002;6:47–62.
- [10] Cadarette SM, Jaglal SB, Murray TM, McIsaac WJ, Joseph L, Brown JP. Evaluation of decision rules for referring women for bone densitometry by dual-energy X-ray absorptiometry. *JAMA* 2001;286:57–63.
- [11] Richy F, Gourlay M, Ross PD, Sen SS, Radican L, De Ceulaer F, et al. Validation and comparative evaluation of the osteoporosis self-assessment tool (OST) in a Caucasian population from Belgium. *Q J Med* 2004;97:39–46.
- [12] Orwoll ES, Oviatt SK. Longitudinal precision of dual-energy X-ray absorptiometry in a multicenter study. The Nafarelin/Bone Study Group. *J Bone Miner Res* 1991;6:191–7.
- [13] Pearson J, Dequeker J, Henley M, Bright J, Reeve J, Kalender W, et al. European semi-Anthropomorphic spine phantom for the calibration of bone densitometers—Assessment of precision, stability and accuracy—The European Quantitation of Osteoporosis Study Group. *Osteoporos Int* 1995;5:174–84.
- [14] Looker A, Wahner H, Dunn W, Calvo M, Harris T, Heyse S, et al. Updated data on proximal femur bone mineral levels of US adults. *Osteoporos Int* 1998;8:468–90.
- [15] Canny J. A computational approach to edge-detection. *IEEE Trans PAMI* 1986;8:679–98.
- [16] Illingworth J, Kittler J. A survey of the Hough transform. *Comput Vis Graph Image Process* 1988;44:87–116.
- [17] Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry: I. *J Clin Chem Clin Biochem* 1983;21:709–20.
- [18] Rogers M. Robust active shape model search. In: Heyden ASG, Nielsen M, Johansen P, editors. *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision, Copenhagen, Denmark, vol. 2353; 2002. p. 517–30.*
- [19] Gregory JS, Stewart A, Undrill PE, Reid DM, Aspden RM. A method for assessment of the shape of the proximal femur and its relationship to osteoporotic hip fracture. *Osteoporos Int* 2004;15:5–11.
- [20] Taguchi A, Ohtsuka M, Nakamoto T, Tanimoto K. Screening for osteoporosis by dental panoramic radiographs. *Clin Calcium* 2006;16:67–73.
- [21] 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.
- [22] Compston J, Papapoulos SE, Blanchard F. Report on osteoporosis in the European Community: current status and recommendations for the future. Working Party from European Union Member States. *Osteoporos Int* 1998;8:531–4.
- [23] Compston J. Action plan for the prevention of osteoporotic fractures in the European Community. *Osteoporos Int* 2004;15:259–62.
- [24] Geraets WG, 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* 1988;22:165–73.
- [25] Arifin AZ, Taguchi A, Nakamoto T, Ohtsuka M, Tanimoto K. Computer-aided system for measuring the mandibular cortical width on panoramic radiographs in osteoporosis diagnosis. *Proc. SPIE Med Imaging, vol. 5747. Image Processing Conference; 2005. p. 813–21.*
- [26] Yosue T, Brooks SL. The appearance of mental foramina on panoramic radiographs. I. Evaluation of patients. *Oral Surg, Oral Med Oral Pathol* 1989;68:360–4.
- [27] Sakakura CE, Monteiro Loffredo L de C, Scaf G. Diagnostic agreement of conventional and inverted scanning panoramic radiographs in the detection of the mandibular canal and the mental foramen. *J Oral Implant* 2004;30:2–6.
- [28] Hohlweg-Majert B, Schmelzeisen R, Pfeiffer BM, Schneider E. Significance of osteoporosis in craniomaxillofacial surgery: a review of the literature. *Osteoporos Int* 2006;17:167–79.