DETECTING OSTEOPOROSIS FROM DENTAL RADIOGRAPHS USING ACTIVE SHAPE MODELS

P. D. Allen¹, J. Graham¹, D. J. J. Farnell¹, E. J. Marjanovic¹, J. Adams¹, R. Jacobs²
K. Karayianni³, C. Lindh⁴, P. F. van der Stelt⁵, K. Horner⁶, H. Devlin⁶.

¹ISBE, University of Manchester, UK. ²Oral Imaging Centre, Katholieke Universiteit Leuven, Belgium. ³Dental School, University of Athens, Greece. ⁴Faculty of Odontology, Malmö University, Sweden. ⁵Academic Centre for Dentistry, Amsterdam, The Netherlands. ⁶School of Dentistry, University of Manchester, UK.

ABSTRACT

We describe a novel method of estimating reduced bone mineral density (BMD) from dental panoramic tomograms (DPTs) which show the entire mandible. Careful expert width measurement of the inferior mandibular cortex has been shown to be predictive of BMD in hip and spine osteopenia and osteoporosis. We have implemented a method of automatic measurement of the width by active shape model search, using as training data 132 DPTs of female subjects and tested on a test data set of 606 DPTs of female subjects whose BMD has been established by dual-energy X-ray absorptiometry. We demonstrate that the sensitivity of the automatic method to osteoporosis is equivalent to that of expert manual measurement.

Index Terms—Osteoporosis, Dental Radiographs, Bone Mineral Density, Active Shape Model, Image Segmentation.

1. INTRODUCTION

Osteoporosis is a general loss of bone mineral density and can lead to an increased risk of fracture. Based on factors such as previous fracture, family history, and height loss, patients deemed to be at risk are referred for bone mineral density (BMD) assessment using dual-energy X-ray absorptiometry (DXA). However, there has recently been great interest among dental researchers in the possibility of identifying those at risk of reduced BMD from dental radiographs since mandibular BMD is related to systemic BMD [1].

Figure 1 shows an example of a dental panoramic tomogram (DPT) of a normal (non-osteoporotic) patient. The positions of anatomical points key to manual annotation are shown. The mental foramen is a hole in the mandible through which the mental nerve passes, and which is faintly discernible on the DPT. The Gonion and Ante-Gonion are recognisable points where the curvature of the mandible changes.

Fig. 1. An example of a panoramic dental tomogram of a normal (non-osteoporotic) patient. The positions of anatomical points key to manual annotation are shown. The mental foramen is a hole in the mandible through which the mental nerve passes, and which is faintly discernible on the DPT. The Gonion and Ante-Gonion are recognisable points where the curvature of the mandible changes.

Fig. 2. Schematic diagram of the right mandible, showing the point at which the inferior mandibular cortex thickness is measured by dentists (Mental Index MI).

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).
thickness of the IMC at a point closest to the mental foramen, referred to as the Mental Index (MI) (figure 2), has been found to be the best indicator of low BMD compared with the equivalent indices at the gonion (GI) and the ante-gonion (AI) (figure 1) [2].

There is considerable room for subjectivity in the precise placement of the MI measurement, the mental foramen is an indistinct feature and the endosteal border (see figure 2) can become particularly faint in cases of osteoporosis. These factors do not pose significant problems for an expert radiologist. However, for general dental practitioners (GDP), they lead to considerable variability in MI measurement, even with individual training, and so routine assessment of low BMD risk from dental radiographs by GDPS is not practical [3].

Here we describe an automatic method of measuring radiographic indices of BMD using computer image analysis.

2. DATA

Two sets of data were used in this study: a training set on which the ASM was trained and segmentation accuracy was assessed, and a separate test data set on which sensitivity to BMD was assessed.

The training data set, collected for a previous study [4], was obtained from 132 female patients aged 45-55 who attended the University of Manchester Dental Hospital for routine dental treatment. The DPT films were digitised at a resolution of 25.64 pixels/mm. A full manual annotation of these digitised images was made as described below in section 3.1.

The test data was collected specifically for this study from four European centres and consisted of 606 ambulant female patients recruited from the area surrounding these centres of which 133 were osteoporotic. Digital and conventionally processed dental panoramic radiography machines were used. The conventional films were digitised as above, but their resolution reduced digitally to the same resolution (7.69 pixels/mm) as the digital radiographs so that they could be grouped into a consistent data set. In each image, a 3mm diameter ball bearing set in a plastic block and gripped between the teeth served as a calibration object.

The BMD of the test data patients was measured using Dual energy X-ray absorptiometry (DXA) scans at the four recruitment centers, calibrated using the European spine phantom. Patients were diagnosed osteoporotic according to the World Health Organization (WHO) criteria, i.e. those with a bone mineral density T-score value 2.5 SD or more below the mean value of the young sex matched reference population. Full details of the collection of these data are given in [1].

3. THE ACTIVE SHAPE MODEL METHOD

The ASM method has been extensively documented already elsewhere [5, 6] and only a brief description will be given here. At its core is a Point Distribution Model (PDM) that describes the principal modes of variation of a set of landmark points used to describe the object of interest. The model is ‘trained’ using points placed on a training set of example shapes, usually manually (see section 3.1), at anatomically consistent locations around the border of the object. Point Distribution Models may be used in image search as Active Shape Models [5]. Here, the best fit of the model to the data is found through a process of iterative local refinement in which the position of each landmark point is updated to the nearest best grey-level match within the constraints of the PDM. In this study the best match is based on edge strength and edge direction.

Since the lower border of the IMC is far more clearly defined than the upper border (figure 2), the search in this study is divided into two phases. The first phase uses a model built from the points on the lower border only to locate that edge, defining the overall shape and pose of the mandible. This result is used as the initialisation of phase 2, which is a search using the complete model of the IMC to obtain the positions of both lower and upper borders.

3.1. Manual Annotation of the Training Images

The points required to build a PDM of the IMC from the set of training images were placed manually. This was done by two experts using a custom-written graphical user interface. The positions of the upper and lower border of the IMC were marked at two key anatomical landmarks: the points closest to the mental foramen, and the points at the ante-gonion (see figure 1), henceforth referred to as the MF and AG points. Between these, 50 equally spaced points were placed to define the upper and lower borders of the IMC on each side (200 points in all). The PDM was defined using the mean of the two sets of manually positioned points.

Between the left and right MF points the shadow of the spine is unavoidably superimposed on the center of the image of the mandible resulting in poor definition of the cortex. Lateral to the AG points there can also be superimposition of shadows of the opposite side of the mandible due to the tomographic image acquisition, making the endosteal border indistinguishable and the lower mandibular border difficult to discern from other structures (1). Thus the best region from which to measure cortical thickness is between the AG and MF points, and it is from these that the PDM was built.

3.2. Experimental Procedure

To evaluate accuracy of segmentation using the ground-truth available in the training set, leave-one-out cross-validation was used. For the test data set, a leave one out method was unnecessary since none of these were used in training the ASM. No annotation was available for this set, which was used to evaluate BMD estimates against DXA measurement. Two versions of the ASM search described in 3 were tested experimentally on both data sets:
The first was a free search without any manual initialisation points. The first phase ASM search for the lower IMC border was initialised from the mean position and pose of the training data. For some images, the correct shape and pose are some distance from this starting point. ASM search uses a multi-resolution coarse-to-fine search strategy in such circumstances [5] and that was employed in this case. We refer to fits determined this way as ‘unconstrained fits’ or ‘UFits’.

The second version used four manually defined reference points on the lower mandible edge at the left and right AG and MF as starting points for the first phase search. To start the search the mean example of the lower mandible border PDM was stretched and positioned such that its AG and MF points matched the manually placed start points. An edge based ASM search was then initiated, making no further reference to the manual points during the search. The use of this straightforward interaction allowed us to decouple the effects of location and shape in ASM search. Starting the search so close to the true position guarantees that the search will finish up with the correct pose. The quality of ASM fit is determined solely by the ability of the PDM to represent the variation in shape that occurs among the images. We refer to fits determined this way as ‘constrained fits’ or ‘4PFits’.

In both cases the results of the first phase search were then used to initialise the second phase full endosteal border ASM search.

### 4. RESULTS

To estimate the accuracy of model fits we compare them with the manual annotation using the mean point-to-point, and the mean point-to-curve differences. Ultimately we wish to test the sensitivity of the derived measurements to osteoporosis and this is done by calculating the correlation coefficient between the parameter in question and BMD, and by plotting a Receiver Operator Characteristic (ROC) curve [7]. The area under the curve (AUC) can be used to quantify the overall diagnostic efficacy of the parameter in question - ranging from 0.5 (no better than chance) to 1.0 (perfect discrimination).

#### 4.1. Fit Accuracy (Training Data)

The point-to-point and point-to-curve differences are presented in table 1 as mean (standard deviation). Four comparisons are made:

- **Manual 1-2** - The manual annotations of the two observers.
- **Manual-UFit** - The mean of the manual points with those of the UFit.
- **Manual-4PFit** - The mean of the two sets of manual points with those of the 4PFit.
- **Fit1-Fit2** - The 4PFit involves user interaction, and so there is a certain degree of subjectivity involved in the exact placement of the four initialisation points. To estimate the magnitude of this effect we perform two 4PFits, each initialised by a different observer.

The fully automatic method (UFit) results in much larger point-to-point error than the difference between two manual observers. This is because for any given point on the edge of the mandible, its orthogonal distance from the edge is clearly defined by the local grey level gradient. However there are no features such as edges to define its position along the edge of the mandible, and so without the 4 point initialisation, although the border of the IMC will be correctly identified, the correct anatomical position of the resulting points with respect to the AG and MF points cannot be guaranteed. This is borne out by the relatively low corresponding point-to-curve error.

The Fit1-Fit2 results exhibit a much lower point-to-point error than the difference between two manual observers. However, this is not surprising since the 4PFit is initialised using points that are the mean of points from the two observers. The corresponding point-to-curve errors are very similar to those between the two manual observers. Further details on model evaluation can be found in [8].

#### 4.2. Sensitivity to Reduced BMD (Test Data)

The correlation between measured IMC thickness and BMD varies along the mandible. Figure 3 shows this variation for the 4PFit measurement and DXA measured at the hip. A similar curve is obtained for the other DXA sites, and the curve

<table>
<thead>
<tr>
<th>Comparison</th>
<th>point-to-point</th>
<th>point-to-curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual 1-2</td>
<td>2.45 (2.45)</td>
<td>0.31 (0.33)</td>
</tr>
<tr>
<td>Manual-UFit</td>
<td>5.73 (4.57)</td>
<td>0.49 (1.58)</td>
</tr>
<tr>
<td>Manual-4PFit</td>
<td>0.59 (0.54)</td>
<td>0.31 (0.40)</td>
</tr>
<tr>
<td>Fit1-Fit2</td>
<td>2.31 (2.44)</td>
<td>0.14 (0.24)</td>
</tr>
</tbody>
</table>

*Fig. 3. Correlation coefficient between cortical thickness and Hip BMD plotted as a function of position along mandible.*
derived by taking the minimum BMD measurement from the three sites is also shown. At all positions along the mandible the correlations exceed the 1% significance threshold of 0.098. This indicates that the highest sensitivity to BMD is found by measuring not at the MF, but in the lateral half of the mandible.

Figure 4 shows the ROC curves obtained using the cortical thickness averaged over this optimum region of the mandible for both 4PFit and UFit points. These are also compared with the results of measuring the cortical thickness manually from the radiographs - the average of 5 expert observers is shown [1]. The difference in AUC between the UFit and manual data is not statistically significant, but the difference between the 4PFit and the manual data is (p=0.007).

5. DISCUSSION

We can conclude from the above that it is possible to accurately measure the width of the inferior mandibular cortex in panoramic dental tomograms using an edge-based ASM method. For these measurements to have an exact anatomical correspondence, four manually placed initialisation points are required. This initialisation is a reasonable level of interaction since only the lower mandible edge need be identified - a clearly visible feature in all patients - and the points only need to be placed close to the border, not exactly on it, since the ASM search will locate the exact position of the local edge anyway.

Correlation of the cortical thickness with the BMD measured from the spine, hip or femoral neck, was highest for the lateral portion of the AG-MF region of the mandible. Using this optimal region to test for Osteoporosis, the resulting ROC curve suggests improved performance over traditional manual measurements for the 4PFit, and equivalent performance from the UFit.

In [9] we discuss the clinical applicability of estimation of BMD by ASM search on DPTs. Osteoporosis diagnosis is not part of everyday dental practice and so any involvement in this task should be facilitated for the dentist. Measurement of cortical width will only be practical if it is fully automatic or very nearly so. The limited and straightforward interaction described here may be sufficiently unobtrusive to be practical. However the results of automatic search indicate that useful measurement can be made without dentist involvement at all. The improved specificity and sensitivity arising from being able to make measurements at anatomically precise locations holds out the possibility of improved diagnostic performance.

6. REFERENCES


