Diagnosing osteoporosis: A new perspective on estimating bone density

R. Cassia-Moura\textsuperscript{a,b,c,*}, A.D. Ramos\textsuperscript{d}, C.S. Sousa\textsuperscript{d}, T.A.S. Nascimento\textsuperscript{c}, M.M. Valença\textsuperscript{e}, L.C.B.B. Coelho\textsuperscript{f}, S.B. Melo\textsuperscript{c}

\textsuperscript{a}International Centre for Theoretical Physics, Trieste 34100, Italy
\textsuperscript{b}Instituto de Ciências Biológicas, DCF - Biofísica, Universidade de Pernambuco, Caixa Postal 7817, Recife 50670 000, Brazil
\textsuperscript{c}Centro de Informática, Universidade Federal de Pernambuco, Recife 50740-540, Brazil
\textsuperscript{d}Departamento de Estatística, Universidade Federal de Pernambuco, Recife 50740-540, Brazil
\textsuperscript{e}Centro de Ciências da Saúde, Universidade Federal de Pernambuco, Recife 50740-540, Brazil
\textsuperscript{f}Departamento de Bioquímica, Universidade Federal de Pernambuco, Recife 50740-540, Brazil

Received 2 June 2006; received in revised form 19 September 2006
Available online 19 March 2007

Abstract

Osteoporosis may be characterized by low bone density and its significance is expected to grow as the population of the world both increases and ages. Our purpose here is to model human bone mineral density estimated through dual-energy x-ray absorptiometry, using local volumetric distance spline interpolants. Interpolating the values means the construction of a function \( F(x, y, z) \) that mimics the relationship implied by the data \((x_i, y_i, z_i; f_i)\), in such a way that \( F(x_i, y_i, z_i) = f_i \), \( i = 1, 2, \ldots, n \), where \( x, y \) and \( z \) represent, respectively, age, weight and height. This strategy greatly enhances the ability to accurately express the patient's bone density measurements, with the potential to become a framework for bone densitometry in clinical practice. The usefulness of our model is demonstrated in 424 patients and the relevance of our results for diagnosing osteoporosis is discussed.

\( \odot \) 2007 Elsevier B.V. All rights reserved.

Keywords: Densitometry; DEXA; Interpolation methods; Human patient; Scattered data; Mathematical modeling

1. Introduction

Over the past 10 years, osteoporosis has emerged as a major clinical challenge for physicians and patients, with regard both to its prevalence and to the morbidity and mortality of associated fractures. Moreover, resultant fractures in the hip are associated with a greater number of deaths and disabilities, increasing socioeconomic costs that are expected to rise in the future as the population of the world both increases and ages. In their article, Delmas and Fraser \cite{1} present a compelling argument describing the potential health crisis the world will face if osteoporosis is not made a high priority by the world health community.
They describe several health-related consequences of this disease, especially in terms of increases in human pain and suffering and the continual increase in global health care costs [1]. Despite the fact that osteoporosis may affect the entire skeleton, fractures constitute the only important adverse health consequence of this disorder [2], often from minor falls or spills that would not normally result in bone fracture. Osteoporosis is defined as a progressive systemic skeletal disease characterized by low bone mineral density (BMD), microarchitectural deterioration of bone tissue or both, with a consequent increase in bone fragility and susceptibility to fracture [3,4]. It has been estimated that 13–18% of women aged 50 and over have osteoporosis, and for those over the age of 80, the proportion rises to 70% [5] and [6]. Osteoporosis is one of the most debilitating diseases particularly in peri- or post-menopausal women, but there are many other diseases (not related to gender, age or hereditary) that can cause this condition, e.g. hyperthyroidism, diabetes, hypogonadism, to mention just a few.

Bone density is the best single predictor of a future bone fracture and measurement of BMD with dual-energy X-ray absorptiometry (Dexa) is the current “gold standard” physical method for diagnosing osteoporosis. Dexa measures the BMD and compares this measurement with a reference population based on age, weight, height, gender and ethnic background. Bone densitometry is a simple, non-invasive and painless procedure that examines the hip, lumbar spine (L₂–L₄) and occasionally the wrist, because these are the sites where osteoporosis first appears. By using local volumetric distance spline interpolants, a major component of this study is to model the BMD generated with Dexa measurements. Here we present one method for solving variants of the following problem: given a finite set of n scattered BMD data points in the three-dimensional physical space, how can one find a surface that interpolates a given set of points? A manifold interpolation is proposed for comparing the patient’s BMD with a reference healthy population based on age, weight, height, gender and ethnic background. This is a mathematical paper on solving an experimental medical problem using an algebraic approach.

The structure of this paper is as follows: in Section 2, the current scores are described, which are used to identify high-risk fracture patients; Section 3 starts by introducing the interpolation methods, after which the volumetric distance spline method is described, which will lead to the presentation of our theoretical model; the experimental procedure is described in Section 4; Section 5 focuses on the results and discussion, in which the importance of proper monitoring of the osteoporosis, and the need for a smooth interpolation for BMD as a function of physical parameters of interest, are discussed; and in Section 6 we present our conclusions.

2. The raw report

From the need to make an individualized approach to osteoporosis screening more practical and simplified for physicians, several clinical prediction rules (CPRs) have been developed, several of which attempt to predict BMD outcomes and others that attempt to predict fracture outcomes [7]. The quality of most of these studies has been rated fair or poor, most often because of methodological limitations [7]. A CPR that was developed and tested in studies with good methodological ratings was the simple calculated osteoporosis risk assessment estimation (score) [8]. In most reports, the number of standard deviation (sd) away from the mean is represented as two scores, the T-score and the Z-score, which are used to compare the subject’s BMD value with the mean values in others of the same weight, height, gender and ethnic background. T-score is compared with young normal subjects, which relates more closely to fracture risk, and Z-score is expressed relative to the average of persons of the same age as the person being tested.

Despite the fact that the T-score is an industry standard to interpret Dexa results, in 1994, using the young adult women as the reference group, the World Health Organization (WHO) Osteoporosis Study Group [9] has classified patients with a T-score above −1 as normal; those with a T-score of −1 to −2.5 are classified as low BMD (i.e., osteopenic, and a risk for developing osteoporosis); and subjects with a T-score below −2.5 are considered to have an abnormally low BMD (i.e., diagnostic of osteoporosis). Although there is no true “fracture threshold” (i.e., a value of BMD above which fractures do not occur, since sufficient trauma will break any bone), the fracture risk increases considerably at bone densities that are less than 2 sd below the young normal mean. A Z-score below −1.5 raises concern with factors other than aging as contributing to osteoporosis, which may include tobacco use, malnutrition, to mention just a few.
3. Theoretical model

For many scientific visualization systems, it is desirable to have the input data defined over a regular grid, but this is usually not the case. This is where the mathematical modeling of data comes in. A mathematical model can be evaluated over a uniform grid and these sampled data are used as input to a visualization tool. The visualization system can then use interpolation schemes to generate uniform grid data in cases where scattered data need to be displayed. Scattered data is a term used to convey the idea of data that are at irregular positions. In medical imaging, scattered data interpolation is essential to construct a closed surface of human organs from CT scans, MRI, SPECT and some ultrasound techniques. It requires scattered data interpolation to determine values at arbitrary positions, not just those at which the data are available. The reader is referred to Nielson [10] for various tests and comparison of some methods for scattered data interpolation. In Fig. 1 we can observe an example of interpolation of a data set defined in $\mathbb{R}$:

The goal of interpolation is to construct an underlying function that may be evaluated at any desired set of positions [11]. Interpolation means finding a curve or surface that satisfies some imposed constraints exactly [12]. This is a paper on an interpolation scheme and its use for the study of osteoporosis.

Normally the interpolant is a linear combination of some basis functions. This is due to some basic assumptions, such as continuity and smoothness, a scaling requirement and the preservation of linear combinations. For instance, if we scale up the values $F_i$, $i = 1, 2, \ldots, n$ by a factor of $k$, then we expect the produced interpolant to be equivalent to the previous interpolant scaled up by $k$. Additionally, if we decompose the values $F_i$ as $G_i + H_i$, then we expect to be able to decompose the interpolant $F$ as a sum of $G$ and $H$, the interpolants corresponding to $G_i$ and $H_i$, $i = 1, 2, \ldots, n$, respectively. Thus, we can say that the mathematical modeling process is a search for adequate basis functions.

A modeling function defined over the entire domain is determined so that it may interpolate or it may approximate the given scattered data. It is clear from Section 2 that it is very important to rely more on the precision of the function values, pointing towards interpolant models as opposed to the approximating ones (nowadays, the usual procedure by using the number of $sd$ away from the mean value of the BMD). This is the reason why we propose here one interpolating function $F(x, y, z)$, which in some sense fits the scattered BMD data. Interpolating the values means the construction of a function $F(x, y, z)$ that mimics the relationship implied by the data $(x_i, y_i, z_i; f_i)$, in such a way that $F(x_i, y_i, z_i) = f_i$, $i = 1, 2, \ldots, n$. By using the BMD data set generated by Dexa, $x, y$ and $z$ represent, respectively, age, weight, and height. Our aim is to obtain $F$ applied to an arbitrary point, not necessarily coincident with any one of the given data points. The data points represent the clinical information obtained from a population during a given period of time, restricted to a site of the bone (hip, lumbar spine or wrist). The arbitrary point represents what should be expected from a patient who is being examined by the physician at a particular time. The actual measured BMD is then compared with the model’s result for diagnosing osteoporosis.

![Fig. 1. Interpolation of a scattered data set using the distance spline method, where the interpolant $F$ is defined as $F : \mathbb{R} \rightarrow \mathbb{R}$.](image)
In order to make the above method usable and apply it to very large data sets, we adopt the idea of localizing the method. This requires localizing functions, which are smooth and have a small region of support. In general any method may be used to obtain the local interpolants. A local method that produces smoother interpolants is the local volumetric distance spline method. This is the generalization of the local cubic spline interpolation method, which in itself is a modification of the C^2 piecewise cubic splines. Let F'(x) = (∂/∂x)F(x) and F''(x) = (∂^2/∂x^2)F(x). This univariate interpolant is the solution to the following problem: given the BMD data (x_i, f_i), i = 1, 2, ..., n, a < x_1 < x_2 < ... < x_n < b, find, among all piecewise functions defined over [a, b], a function that minimizes \( \int_a^b (F''(x))^2 \, dx \), subject to the interpolation conditions F(x_i) = f_i, i = 1, 2, ..., n. This minimization condition is chosen in order to prevent the graph of the function from excessively wiggling, so typical of high degree interpolating polynomials, and additionally it allows continuity in the second derivative at the junctions (for more on this, see Ref. [13]). The piecewise polynomials are the simplest functions, and the cubic ones are those with the lowest degree that have enough freedom to allow continuity in the second derivative at the junctions x_i's. The resulting function is called the natural cubic spline, which can be characterized by the following conditions: (i) F, F', F'' continuous over [a, b]; (ii) F(x_i) = f_i, i = 1, 2, ..., n; (iii) F is piecewise cubic, i.e., F is a cubic polynomial on each interval [x_i, x_{i+1}], i = 1, 2, ..., n - 1; (iv) F is linear on [a, x_1] and [x_n, b], which means F''(x_i) = F''(x_n) = 0. The interpolant can then be expressed as

\[
F(x) = c_0 + c_x x + \sum_{i=1}^n c_i |x - x_i|^3.
\]

The basis is \{1, x, |x - x_1|^3, |x - x_2|^3, ..., |x - x_n|^3\}, and all of its components satisfy (i). It is straightforward to verify that this is so, even for the junctions x_i, i = 1, ..., n. By adding conditions (ii) and (iv), and doing some algebraic manipulations (see the Appendix), we end up obtaining the following linear system of equations:

\[
\begin{pmatrix}
1 & x_1 & 0 & |x_1 - x_2|^3 & \cdots & |x_1 - x_n|^3 \\
1 & x_2 & |x_2 - x_1|^3 & 0 & \cdots & |x_2 - x_n|^3 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_n & |x_n - x_1|^3 & |x_n - x_2|^3 & \cdots & 0 \\
0 & 0 & 1 & 1 & \cdots & 1 \\
0 & 0 & x_1 & x_2 & \cdots & x_n
\end{pmatrix}
\begin{pmatrix}
c_0 \\
c_x \\
c_1 \\
c_2 \\
0 \\
0
\end{pmatrix}
= \begin{pmatrix}
f_1 \\
f_2 \\
f_1 \\
f_2 \\
f_n \\
f_n
\end{pmatrix}.
\]

To generalize this idea to volumetric BMD data points, we start by letting F : \( \mathbb{R}^3 \rightarrow \mathbb{R} \) be defined as follows:

\[
F(x, y, z) = c_0 + c_x x + c_y y + c_z z + \sum_{i=1}^n c_i ||(x - x_i, y - y_i, z - z_i)||^3,
\]

where \( ||(x - x_i, y - y_i, z - z_i)|| = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} \) and, by imposing the corresponding conditions for the trivariate case, and if we let \( v_i = (x_i - x_j, y_i - y_j, z_i - z_j) \), we get

\[
\begin{pmatrix}
1 & x_1 & y_1 & z_1 & 0 & ||v_{11}||^3 & ||v_{1n}||^3 \\
1 & x_2 & y_2 & z_2 & ||v_{21}||^3 & 0 & ||v_{2n}||^3 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & x_n & y_n & z_n & ||v_{n1}||^3 & ||v_{n2}||^3 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & \cdots & 1 \\
0 & 0 & 0 & 0 & x_1 & x_2 & \cdots & x_n \\
0 & 0 & 0 & 0 & y_1 & y_2 & \cdots & y_n \\
0 & 0 & 0 & 0 & z_1 & z_2 & \cdots & z_n
\end{pmatrix}
\begin{pmatrix}
c_0 \\
c_x \\
c_y \\
c_z \\
c_1 \\
c_2 \\
c_n
\end{pmatrix}
= \begin{pmatrix}
f_1 \\
f_2 \\
\vdots \\
f_n \\
n \\
0 \\
0
\end{pmatrix}.
\]
In order to localize this interpolant, we first subdivide the domain into regions, with non-empty intersections, each region with roughly the same number of data points. We then define smooth functions $w_k : \mathbb{R}^3 \rightarrow [0, 1]$, whose support is the $k^{th}$ region. In addition, they satisfy the unity partition property: $\sum w_k(x, y, z) = 1$ for all $(x, y, z) \in \mathbb{R}^3$. We compute the localized interpolants $F_k$ such that $F_k(x_i, y_i, z_i) = f_i$ for all data points in the support of $w_k$, by solving linear systems of equations like the one above, one for each region. We then take $F(x, y, z) = \sum w_k(x, y, z)F_k(x, y, z)$ as our interpolant, since $F(x_i, y_i, z_i) = f_i$, $i = 1, 2, \ldots, n$, provided that $w_k$ functions, it is not hard to build them through piecewise tricubics.

The local volumetric distance spline method ranks favorably against others because it is relatively easy to implement, is local and very smooth, and can be applied to very large databases. The detail that requires attention is the choice of the covering regions of support. The application needs to pick regions in such a way that their number of BMD data points is approximately constant. Alternatively, it may let the user make this choice, but it should be robust enough to handle regions with too few data points. Sometimes the local convex hull of the BMD data points is degenerate, and does not present the form of a solid, which is necessary for building a volumetric interpolation.

4. Materials and methods

We tested our theoretical model on a data set from Unidade de Densitometria Ossea do Recife/Brazil, composed of lumbar spine BMD readings ($L_2$–$L_4$-BMD) of a Dexa (DPX-L, Lunar Radiation Corp., Madison-WI, USA), from 1991 to 1999. From this database we have selected the entire set of white women $L_2$–$L_4$-BMD readings, composed of 5761 individual entries. We generated a table from the patient’s information needed for our modeling, as follows: age, weight, height and $L_2$–$L_4$-BMD. In order to build our reference healthy population that is used to produce the local volumetric distance spline function, we based our selection on the paper [14], which established the normality standards of $L_2$–$L_4$-BMD readings, for a Brazilian white female population, using the same equipment (DPX-L). In order to remain within that paper’s limits, we excluded 1070 individuals who did not satisfy the following conditions: age between 20 and 70 years, and weight between 40 and 80 kg. The corresponding normal patient’s readings in that paper are presented in a table in which each entry is referred to as a reference group and corresponds to the mean $L_2$–$L_4$-BMD and the sd ranges of 10 years by 10 kg. For instance, a patient whose age is in the range of 40–49 years with weight between 50 and 59 kg should present a BMD reading of 1.117, with a sd of 0.142.

We decided to select the reference population by using a piecewise bilinear interpolation [12], which means that in order to find an individual reference BMD, we consider that the BMD values in a reference group change linearly as the age moves into the neighboring reference group, and the same is done as the weight moves from one reference group into the next one. We consider that the values presented by Lewin et al. [14] in a reference group correspond to its central point; for instance, in the reference group 40–49 years-old and 50–59 kg, the center point is 45 years-old and weight 55 kg; in the neighboring reference group 50–59 years-old and 60–69 kg, the center point is 55 years-old and weight 65 kg. From the weight of 55–65 kg the BMD values and the sd will vary linearly. From the age of 45 years to 55 years, the BMD values and the sd will also vary linearly. The same applies to the other ranges. In this way there will be no appreciable difference between the reference BMD used for the age of 49.9 and that used for 50.0 years. By applying the piecewise bilinear

<table>
<thead>
<tr>
<th>Age x Weight</th>
<th>40–49</th>
<th>50–59</th>
<th>60–69</th>
<th>70–79</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–29</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>30–39</td>
<td>17</td>
<td>42</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>40–49</td>
<td>39</td>
<td>264</td>
<td>268</td>
<td>136</td>
</tr>
<tr>
<td>50–59</td>
<td>83</td>
<td>356</td>
<td>465</td>
<td>257</td>
</tr>
<tr>
<td>60–69</td>
<td>73</td>
<td>250</td>
<td>338</td>
<td>207</td>
</tr>
</tbody>
</table>
interpolation we obtained 2856 normal $L_2$–$L_4$-BMD readings (see Table 1). Our reference population is a further selection in this normal population.

In order to obtain a satisfactory distribution of patients, we sorted the list of entries by increasing age, and accepted an entry if the BMD value was less than 0.07 under the previous entry’s BMD, and was less than 0.05 above the previous entry’s BMD. This difference in threshold produces a bias toward a smaller BMD as age increases; this is an expected behavior according to Lewin et al. [14]. When this was done, our reference population decreased to 904 white females. We then sorted by decreasing weight, and accepted an entry if the BMD value was less than 0.09 under the previous entry’s BMD, and was less than 0.11 above the previous entry’s BMD, creating another bias toward a larger BMD as weight increases; this is also an expected behavior according to Lewin et al. [14]. Through this procedure we obtained our 424 individuals that formed the reference healthy population used to build (as shown in Section 3) our local volumetric distance spline function.

5. Results and discussion

5.1. Experiments with a human population

Fig. 2 was produced by using a technique called ray tracing (see Ref. [15]), which allows the user to establish the parallelepiped bounds, viewing configuration and color attributes. Fig. 2a shows a partial volumetric graph of our local volumetric distance spline function. A point in the rectangular parallelepiped is a triplet of coordinates.
numbers representing age in the range 40–65 years, weight in the range 50–60 kg and height in the range 140–155 cm. The parallelepiped corresponding to these constraints is entirely contained in the interior of the volume corresponding to our reference healthy population (i.e., 424 patients, with ages ranging from 20 to 70 years, weights ranging from 40 to 80 kg and heights ranging from 140 to 175 cm). The BMD value of a point is represented in shades of gray, where a white point is associated with a BMD of 1.351 (maximum value) and a black point is associated with a BMD of 0.8. Fig. 2b–d show the rendering of the function by considering one of the dimensions varying one unit.

In Fig. 2b, when the weight is kept constant at the value of 61 kg, with ages in the range of 40–65 years, the BMD value will be positively influenced by the increase in height in the range of 140–165 cm. In the Fig. 2c it may be seen that, in 51-year-old patients with a low BMD value, the BMD value was not significantly influenced by change in weight from 50 to 75 kg and in height from 140 to 165 cm. Fig. 2d reveals that, in patients aged 40–65 years, a weight of 50–75 kg and height of 161 cm, the individuals with a lower weight have a higher risk of developing osteoporosis.

The oscillations in gray tonality observed in Fig. 2 reflect a behavior influenced by a noise naturally present in biological systems, suggesting the existence of factors external to our model, which are compatible with those pointed out by Lewin et al. [14]. The analysis of some patient findings takes into consideration a range of tolerance, from which such oscillations may be expected to be absorbed. The methodology proposed in this study may have advantages when performing analyses employing regression rather than interpolation, which probably would minimize the appearance of oscillations.

With the aim of validating our model, we carried out a statistical analysis of the set of white women L2–L4-BMD readings, composed of 5761 individual entries. We found that the mean of the BMD values, μ, is 1.0328108. Since we are interested in knowing how far, on average, our data are close to or distant from μ, we calculated their average deviation (i.e., root mean square, rms), expressed by

\[ \text{rms}(x - \mu) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2}, \]

where \( x = (x_1, \ldots, x_{5761}) \) represent the BMDs of our set of data. For our \( x \), the value of \( \text{rms}(x - \mu) \) is 0.1830667, which implies that our data are on average at a distance of 0.1830667 from the value of μ. In order to assess the variance in our values of this rms, that is, to what extent this measurement may not reflect the overall behavior of our 5761 patients, we calculated

\[ \text{Var}(x) = (\text{rms}(x - \mu))^2, \]

which in our case was 0.0335134. In view of the above, we may conclude that on average our data are really manifested at a distance of 0.1830667 from μ. In our set of 5761 items of data the minimum value found was 0.415, while the maximum value was 1.882. This interval [0.415;1.882] enables us to conclude that the \( \text{rms}(x - \mu) = 0.1830667 \) is small. Therefore, a larger quantity of our data is manifested close to the value of μ.

We evaluated whether, by means of the \( F \) interpolation function (obtained from the 424 items of data extracted from the set of 5761 patients, as described in Section 4), the new BMDs obtained from the ages, weights and heights of these 5761 patients, present a behavior qualitatively similar to that of the measurements of the 5761 original BMDs (i.e., mean, rms, variance, etc). We found that the mean of the new BMD values, \( \mu_F \), obtained by means of the \( F \) is 0.9817268. In this case the value of the \( \text{rms} \) is 0.1556418, which implies that this is the mean distance of our new BMDs as far the value of \( \mu_F \). We also calculated the variance of these values in order to discover to what extent this average behavior may not reflect the overall behavior of our 5761 new BMDs, thus obtaining a value of 0.0242244. In view of the above, we may conclude that, on average, the new BMD values are really manifested at a distance of 0.1556418 from the value of \( \mu_F \). In our set of 5761 new BMDs the minimum value found was 0.0381, while the maximum value was 1.6375. This interval [0.0381;1.6375] enables us to conclude that the value of \( \text{rms} \) is small. Therefore, a larger quantity of our data is manifested close to the value of \( \mu_F \). We may thus consider that our model, despite having used less than 10% of the total set of data (i.e., 424 patients), very accurately reflected the measurements used from the total set of data (i.e., 5761 patients).
5.2. Modeling issues

By means of computer simulation, parameters in the model were adjusted, enabling us to obtain a figure of approximately 10% of the total data for the purpose of adapting our theoretical model to the experimental data obtained in 5761 human patients. From the article of Lewin et al. [14] we obtained the criteria for constructing our model, which generates an interpolant function representing the adjustment of data from a population composed of 424 normal individuals, a population that serves as a baseline for the detection of osteoporosis. We adopted the following criteria proposed by Lewin et al. [14]: (i) data on white women aged from 20 to 70 years and weighing between 40 and 80 kilos; (ii) data on patients whose BMD tends to diminish with increasing age; (iii) data on patients whose BMD tends to increase with increasing weight. In addition, we decided that for each reference group GR, (i.e., weight versus age) a BMD value associated with its respective SD was adopted, referred to here as \(B_i\) and \(sd_i\), respectively, from the center of each \(GR_i\), so that the \([B_i + sd_i, B_i - sd_i]\) interval is linearly interpolated with the \([B_j + sd_j, B_j - sd_j]\) interval of its neighbor. This being so, what we actually did was select the data in such a way as to decrease the noise. Our model thus comes to act as an approximation in relation to the 5761 patients and is in fact an interpolation in relation to the 424 patients that comprise the model. And in accordance with our model, the 424 points represent the best adjustment possible when we consider the expected behavior of a population of normal patients. In this sense, the main finding of this study is that the manifold we propose here can greatly enhance the ability to accurately inform the patient’s bone density measurements.

Fitting a model to the BMD data can help visualize otherwise hidden structure in the data. Understanding the overall behavior of the functional models is important in deciding between finding an approximating or an interpolating model. The statistical approach, which nowadays is the usual procedure (described in Section 2), is equivalent to find an approximating model, whose graphic form is an hyperplane orthogonal to the density axis. There then it follows a \(sd\)-normalization of the distance between the individual measured BMD and the corresponding linear value obtained from the approximating model. Thus, only global information is used to estimate the BMD at a particular point, ignoring relevant neighboring BMD data. This approach is more suitable for noisy data points, which is not the case with BMD data from a particular densitometer brand. Furthermore, we do not expect sudden changes in bone density, for we measure it continuously over a period of time, taking into consideration variables such as age, height, weight, from populations classified by ethnic background and gender, with the data separated by selected sites of the bone. In the particular case of bone densitometry, the BMD at a given point is already a mean of values acquired over a site, which by itself filters out most of the undesirable high frequency effects. Thus, instead of averaging BMD from patients of the same age, in number of years, to confront with the individual measured data (nowadays, the usual procedure, as described in Section 2), here we construct a pool of 424 BMD data from a reference healthy population (as described in Section 4), of different ages, crossed with different weights and heights, for a particular site of bone, from populations classified by ethnic background and gender. In this way we are able to build a database with the evolution of a normal population’s BMD, as the population ages. From these data, a smooth interpolating function is constructed (as shown in Section 3), and the individual BMD is compared with the corresponding evaluated density from the function. Once we utilize a function, it becomes implicit that we cannot have a single triplet \((x, y, z)\) with different BMD values. One advantage of our method is that the proposed model has a smooth and flexible behavior within the range adopted for the data analysis, unlike the method routinely used in clinical practice.

In this study we present an interpolant function for experimental data obtained in a healthy population. With a view to its use in the clinical detection of osteoporosis, as we are dealing with a smooth (i.e., continuous) function, we may be able to make it a density function \(f_d\), thereby enabling us to obtain a \(sd\) associated with it. Therefore, one suggested use in diagnosis is to analyze the BMD value of the patient being studied, comparing it with the value of the \(F\) interpolation function considered for this patient. In this case the patient is situated in a given range obtained from the \(sd\) in relation to the \(F\) interpolation function through which the patient may be regarded as normal, osteopenic or osteoporotic, according to the range in which his or her BMD is situated. For the construction of the range of normality, osteopenia and osteoporosis it will be necessary to carry out a subsequent, more thorough study based on clinical data, with a view to using the \(sd\) for obtaining the ranges in question.
By using our model, scattered data sets that contain noise or unwanted detail can be smoothed at low cost by the appropriate choice of the basis function. Implicit modeling of surfaces with basis function offers the ability to interpolate across large, irregular holes in incomplete surface data without constraining the topology of a structure or relying on a priori knowledge of its shape. A single function from surface data can be considered as a volume function. This spatial function may represent a signed “distance” from the bone’s surface and is an explicit function of position. The bone’s surface may be defined implicitly as the zero set of this function, and points inside the bone have a negative “distance”, while points outside are positive. Thus, the patient’s BMD data are evaluated on a regular three-dimensional grid, which can be useful for multi-planar re-slicing or for extracting isosurfaces that can be combined for visualization systems. Multiple surfaces can be visualized at once with appropriate choices of color and transparency for the different surfaces. Multiple isosurface thresholds may be displayed in consecutive thick slices through the data. The isosurface may be shaded by the magnitude of the gradient at each point, which may show how fast the BMD is changing on the isosurface.

For clinical usage of our model, a rendering of BMD function over bone models can be provided that allows the user to navigate through the model. The standard coloring can use shades of red when the observed patient BMD value falls below the corresponding interpolated function value, and shades of green when the opposite occurs. Additionally, an index can be formulated using absolute differences between the actual patient BMD at a certain bone site and the estimated one from the interpolating function, in order to evaluate the patient’s possible abnormality. A table can be established, with the reference values in each case, sorted by age, weight, height, gender and ethnic background, which might be a valuable tool for diagnosing osteoporosis. A calibration factor will be necessary to minimize differences between brands of densitometers. The generalizability of these tools is currently limited and will remain so until their applicability in the clinical setting can be tested in a prospective manner. We will address these issues in an upcoming paper.

5.3. BMD estimates: magnitude of the clinical problem

As soon as bone densitometry entered into clinical practice during the 1980s, debates began about its appropriate clinical use [16]. By the early 1990s, evidence from prospective studies suggested, but did not prove, that screening BMD reduces fracture rates [16]. Recently Kern et al. [17] suggested that their results provide controlled trial evidence that densitometry screening reduces the risk of hip fracture, but population-based screening for osteoporosis, or low bone density, is [17] controversial—the US Preventive Services Task Force has joined the National Osteoporosis Foundation in recommending routine osteoporosis screening with BMD testing for all women 65 years and older and for those women aged 60–64 years who are at increased risk for osteoporotic fracture [18,19]; on the other hand, the National Institutes of Health argue that evidence is insufficient to recommend population-based screening for women [20]. There is consensus on the recommendation that BMD testing should be individualized, but disagreement about how this individualized approach for screening should be achieved [7]. Opinions differ as to which site should be used to aid in the initial clinical assessment [21]. Gnudi and Malavolta [22] showed that the T-score leads to diagnostic inconsistencies among different skeletal sites. It is widely agreed, however, that the WHO definitions for normal, osteopenic and osteoporotic results may be applied to Dexam measurements of the lumbar spine and hip. Two guidelines acknowledge the problem of osteoporosis in men but do not make screening recommendations [18,20]. The optimum use of densitometry in men and non-white women remains unclear [16].

Though prevalent in white postmenopausal women, osteoporosis occurs in all populations and at all ages and has significant physical, psychosocial, and financial consequences [20]. Over 1.5 million fractures per year are attributable to osteoporosis in the United States; these fractures result in 500,000 hospitalizations, 800,000 emergency room visits, 2.6 million physician visits, 180,000 nursing home placements, and 12 billion dollars to 18 billion dollars in direct healthcare costs each year [23]. Population-based studies of the impact of bone densitometry and osteoporosis remain challenging because of the incomplete and fragmented test data found in most regions of the world [24]. Almost all the evidence and estimates of cost-effectiveness come from the United States and Western Europe, so the current guidelines cannot be directly applied to other regions of the world where rates of fracture are lower and densitometry and pharmacologic treatment is less affordable [25].
As a result of the lack of accepted practice guidelines that recommend measuring BMD for the detection of osteoporosis, physicians either ignore the potential problem of osteoporotic fracture, or are obliged to resort to the use of non-validated clinical factors to determine which subjects are at increased risk of osteoporosis [26]. At present, no clear consensus exists as to which factors should be used to guide this decision [26]. Clinicians need to continue for now to rely on surrogate markers of bone fragility, including BMD, prevalent fracture, and other important risk factors for fracture [27]. Considerable efforts are currently being made worldwide to provide a means of comparing all bone densitometry results, from all physical methods and all commercial systems, for all measurement sites, in both men and women subjects of all ethnic backgrounds. Its use as a clinical tool in the diagnosis of osteoporosis is now established, but the choice of technique, interpretation of results, decision to institute therapy and time to repeat the examination, are all matters of debate. Here we present a new strategy for enhancing BMD estimates, which is useful for diagnosing osteoporosis. As software and hardware innovations become more available and less expensive, manufacturers may soon adapt the model that was presented in Section 3 to provide a better visualization of the BMD results than what is currently commercially available and economically feasible. The advantages include earlier detection of the osteoporosis, more precise determination of the disease stage and better planning of the therapy.

6. Concluding remarks

We have proposed a model that is able to take advantage of the precision and dependability of BMD readings, producing an interpolating function of the BMD scattered data points. The interpolant may be rendered in such a way as to facilitate clinical interpretation and standardization. Since our methodology of trivariate splines is innovative, the originality of this theoretical work lies in finding an application of splines to experimental BMD data analysis. The present paper shows an interesting mathematical model to interpret BMD data, which is of great importance in the detection of osteoporosis, with the potential to become a framework for bone densitometry in clinical practice.

Acknowledgments

We gratefully acknowledge the contributions of the anonymous reviewers. We are very grateful to the Unit of Bone Densitometry from Recife (Unidade de Densitometria Ossea do Recife—Grupo Albuquerque do O, Rua Prof. Joao Medeiros, 1088/Boa Viagem, Recife, Brazil), and special thanks are due to Dr. Lenarte Maia, who provided us with the human database and for fruitful discussion on osteoporosis. CS Sousa and AD Ramos acknowledge financial support from Fapesb/Brazil.

Appendix

In this appendix we show how the linear system of equations for the univariate interpolating function was obtained. The first \( n \) rows are simply

\[
F(x_j) = c_0 + c_j x_j + \sum_{i=1}^{n} c_i |x_j - x_i|^3 = f_j, \quad j \in \{1, 2, \ldots, n\}.
\]

The last two equations \( (\sum_{i=1}^{n} c_i = 0 \) and \( \sum_{i=1}^{n} c_i x_i = 0) \) of linear system (1) in the p. 6 are not so obvious. The equations are obtained from \( F''(x_1) = 0 \) and \( F''(x_n) = 0 \). We first notice that, by derivating twice the function \( F \), we get

\[
F''(x) = 6 \sum_{i=1}^{n} c_i |x - x_i|.
\]

By applying \( x_1 \) to this expression, equaling it to zero, and observing that \( x_1 < x_i, \ i \in \{2, 3, \ldots, n\} \) implies that \( |x_1 - x_i| = (x_i - x_1), \ i \in \{2, 3, \ldots, n\} \), we can then rewrite (2) with

\[
F''(x_1) = 6 \sum_{i=2}^{n} c_i (x_i - x_1) = 0,
\]
which is the same as
\[
\sum_{i=2}^{n} c_i x_i - x_1 \sum_{i=2}^{n} c_i = 0; \tag{3}
\]
by doing the same in the equation \(F''(x_n) = 0\), but remarking that \(x_n > x_i\), for \(i \in \{1, 2, \ldots, n-1\}\), so \(|x_n - x_i| = (x_i - x_n)\), we get
\[
x_n \sum_{i=1}^{n-1} c_i - \sum_{i=1}^{n-1} c_i x_i = 0. \tag{4}
\]
By adding Eqs. (3) and (4) we obtain
\[
-c_1 x_1 + c_n x_n - x_1 c_n + x_n c_1 + (x_n - x_1) \sum_{i=2}^{n-1} c_i = 0
\]
\[
\Rightarrow c_1 (x_n - x_1) + c_n (x_n - x_1) + (x_n - x_1) \sum_{i=2}^{n-1} c_i = 0 \Rightarrow (x_n - x_1) \sum_{i=1}^{n} c_i = 0.
\]

Since \(x_n \neq x_1\) it follows that \(\sum_{i=1}^{n} c_i = 0\). The equation \(\sum_{i=1}^{n} c_i x_i = 0\) is obtained by adding and subtracting the term \(c_1 x_1\) in Eq. (3), which results in \(\sum_{i=1}^{n} c_i x_i - x_1 \sum_{i=1}^{n} c_i = 0\); since \(\sum_{i=1}^{n} c_i = 0\) we get \(\sum_{i=1}^{n} c_i x_i = 0\), as we wished.

References