# A mathematical model for diagnosing osteoporosis

**R. Cassia-Moura** [a,b] **S. B. Melo** [a],

[a] Centro de Informática, Universidade Federal de Pernambuco, Recife 50740-540 Brazil

[b] International Centre for Theoretical Physics, Trieste 34100 Italy

E-mail:rcassiamoura@yahoo.com.br, sbm@cin.ufpe.br

**Abstract** − 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. 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.

**Keywords:** *densitometry, interpolation methods, scattered data, mathematical modeling.*

## 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 [1]. 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 [1]. In their article, Delmas and Fraser [2] 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 [2]. 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 [6] and [9].

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 [1]. 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 $(L_2 - L_4)$ and occasionally the wrist, because these are the sites where osteoporosis first appears [1]. By using local volumetric distance spline interpolants, a major component of this study [1] is to model the BMD generated with Dexa measurements. The goal of interpolation is to construct an underlying function that may be evaluated at any desired set of positions [7]. Interpolation means finding a curve or surface that satisfies some imposed constraints exactly [3]. Here we present [1] 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 [1] for comparing the patient's BMD with a reference healthy population based on age, weight, height, gender and ethnic background.

## Theoretical Model

(for more on this, see Cassia-Moura et al [1])

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 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 [1] 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 $F(x_i, y_i, z_i) = f_i$, $i = 1, 2, ..., n$. 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.

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 $\mathcal{C}^2$ piecewise cubic splines. Let $F'(x) = \frac{\partial}{\partial x}F(x)$ and $F''(x) = \frac{\partial^2}{\partial x^2}F(x)$. 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 Greville [5]). 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 $F''(x_1) = F''(x_n) = 0$. 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 linear system of equations $AX = B$, where $A$ is the matrix:

$$\begin{pmatrix} 1 & x_1 & 0 & |x_1 - x_2|^3 & ... & |x_1 - x_n|^3 \\ 1 & x_2 & |x_2 - x_1|^3 & 0 & ... & |x_2 - x_n|^3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & |x_n - x_1|^3 & |x_n - x_2|^3 & ... & 0 \\ 0 & 0 & 1 & 1 & ... & 1 \\ 0 & 0 & x_1 & x_2 & ... & x_n \end{pmatrix} \quad (1)$$

$$X = \begin{pmatrix} c_0 \\ c_x \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \quad (2)$$

and B is the matrix:

$$\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \\ 0 \\ 0 \end{pmatrix} \quad (3)$$

To generalize this idea to volumetric BMD data points, we start by letting $F : \mathbb{R}^3 \to \mathbb{R}$ 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 $v_{ij} = (x_i - x_j, y_i - y_j, z_i - z_j)$, we get a system $AX = B$, where $A$ is:

$$\begin{pmatrix} 1 & x_1 & y_1 & z_1 & 0 & ||v_{12}||^3 & ... & ||v_{1n}||^3 \\ 1 & x_2 & y_2 & z_2 & ||v_{21}||^3 & 0 & ... & ||v_{2n}||^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & y_n & z_n & ||v_{n1}||^3 & ||v_{n2}||^3 & ... & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & ... & 1 \\ 0 & 0 & 0 & 0 & x_1 & x_2 & ... & x_n \\ 0 & 0 & 0 & 0 & y_1 & y_2 & ... & y_n \\ 0 & 0 & 0 & 0 & z_1 & z_2 & ... & z_n \end{pmatrix}$$

$X$ is the matrix:

$$\begin{pmatrix} c_0 \\ c_x \\ c_y \\ c_z \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

and B is:

$$\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

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 $w_k : \mathbb{R}^3 \to [0, 1]$, whose support is the $k^{th}$ region. In addition, they satisfy the unity partition property: $\sum w_k(x, y, z) = 1$ for any $(x, y, z) \in \mathbb{R}^3$. 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 $F(x_i,y_i,z_i) = f_i$, $i = 1,2,...,n$, provided that $w_k's$ regions of support cover $F's$ entire domain. As for the $w_k$ 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.

## 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, U.S.A.), from 1991 to 1999. From this database we have selected the entire set of white women $L_2 - L_4$-BMD readings, composed of 5,761 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 [8], 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 1,070 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 to 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 [3], 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. [8] in a reference group correspond to its central point; for instance, in the reference group 40 to 49 years-old and 50 to 59 Kg, the center point is 45 years-old and weight 55 Kg; in the neighboring reference group 50 to 59 years-old and 60 to 69 Kg, the center point is 55 years-old and weight 65 Kg. From the weight of 55 Kg to 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 interpolation we obtained 2,856 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. [8]. 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. [8]. Through this procedure we obtained our 424 individuals that formed the reference healthy population used to build (as shown in Section 2) our local volumetric distance spline function [1].

## Results and Discussion

(for more on this, see Cassia-Moura et al [1])

In this study we present [1] 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.

Figure 1 was produced by using a technique called *ray tracing* (see [4]), which allows the user to estab-

Table 1: Distribution of the number of normal patients by age and weight in the reference groups.

| Age×Weight | 40-49 | 50-59 | 60-69 | 70-79 |
|---|---|---|---|---|
| 20-29 | 5 | 4 | 0 | 1 |
| 30-39 | 17 | 42 | 35 | 16 |
| 40-49 | 39 | 264 | 268 | 136 |
| 50-59 | 83 | 356 | 465 | 257 |
| 60-69 | 73 | 250 | 338 | 207 |

lish the parallelepiped bounds, viewing configuration and color attributes. Figure 1a shows a partial volumetric graph of our local volumetric distance spline function. A point in the rectangular parallelepiped is a triplet of numbers representing age in the range 40 to 65 years, weight in the range 50 to 60 Kg and height in the range 140 to 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. Figures 1b, 1c and 1d show the rendering of the function by considering one of the dimensions varying one unit.

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 study.

**Acknowledgements**

We are very grateful to the 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.

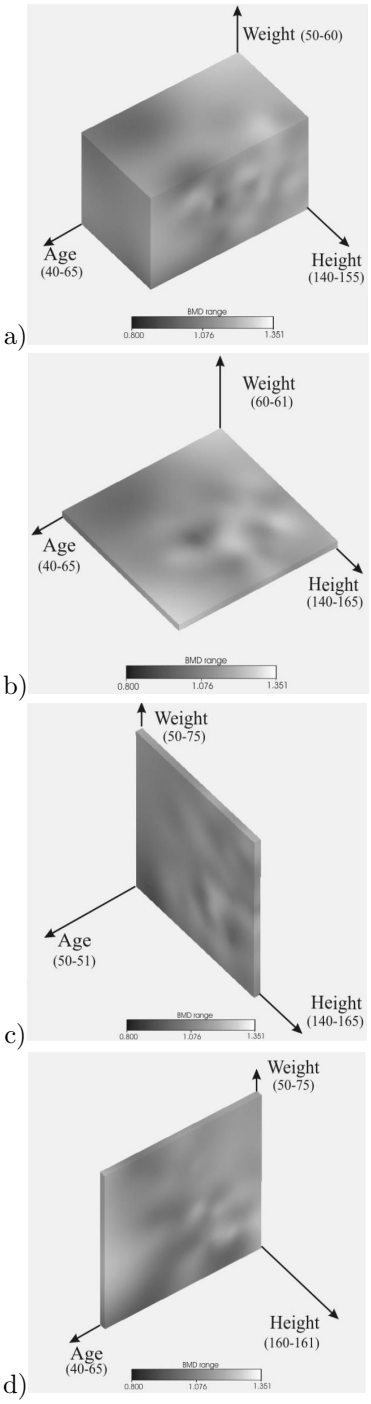


Figure 1: *Distance spline interpolated $L_2 - L_4$-BMD shown on a graph by age (year), weight (Kg) and height (cm). a) Partial volumetric graph of the interpolating function. b) The distance spline interpolation for patients with weight between 60 and 61 Kg. c) The distance spline interpolation for 50 to 51-year-old patients. d) The distance spline interpolation for patients with height between 160 and 161 cm.*

## References

[1] R. Cassia-Moura, A.D. Ramos, C.S. Sousa, T.A.S. Nascimento, M.M. Valença, L.C.B.B. Coelho, S.B. Melo, Diagnosing osteoporosis: A new perspective on estimating bone density, Physica A 381 (2007) 273-284.

[2] P.D. Delmas, M. Fraser, Strong bones in later life: luxury or necessity?, Bull. World Health Organ. 77 (1999) 416-422.

[3] G. Farin, Curves and Surfaces for Computer Aided Geometric Design - a Practical Guide, Academic Press, San Diego, 1997.

[4] J.D. Foley, A. van Dam, S. Feiner, J.F. Hughes, Computer Graphics: Principles and Practice, second ed., Adison Wesley, Reading MA, 1995.

[5] T.N.E. Greville, Theory and Applications of Spline Functions, Academic Press, New York, 1969.

[6] I. Kulenovic, S. Rasic, E. Kulenovic, Osteoporosis current trends in diagnosis and management, Bosn. J. Basic Med. Sci. 6 (2006) 24-28.

[7] S. Lee, G. Wolberg, S.Y. Shin, Scattered data interpolation with multilevel B-splines, I.E.E.E. Trans. Vis. Comput. Graphics 3 (1997) 228-244.

[8] S. Lewin, C.H.A. Gouveia, M.M.S. Marone, S. Wehra, L.F. Malvestiti, A.C. Bianco, Densidade mineral ossea vertebral e femoral de 724 mulheres brancas brasileiras: influencia da idade e do peso corporal, Rev. Ass. Med. Brasil. 43 (1997) 127-136.

[9] P. Werner, Knowledge about osteoporosis: assessment, correlates and outcomes, Osteoporos. Int. 16 (2005) 115-127.