Reviewer's report

Title: A Method for Determining Skeletal Sizes from DXA Images

Version: 1 Date: 8 May 2007

Reviewer: Roger I Price

Reviewer's report:

General

This study has comparatively assessed skeletal dimensions in 90 Caucasian females (actually only 88), at two time points separated by 4 years, from whole body fan beam DXA images, using two techniques – one based on an electronic ruler applied directly to the DXA image data, and the other based on manual analysis of hardcopy printouts. Data from a subset of the subjects were also compared with direct anthropometry, and data from another subset with morphometric analysis of radiographs.

This is a potentially useful report in that there is significant interest in the possibility of bone dimensions being obtainable from whole body DXA scans, partly because of the existence of large and diverse retrospective data sets. With the advent of fan beam DXA systems however, there is concern about geometric distortion, at least along one axis of the scanning plane. (For the Hologic 4500, a circular object only produces a circular image at about 12cm above scanning cushion). In respect of radiographs being used for the same purpose, there is concern about magnification effects. Finally, though estimates of bone dimensions can obviously be made from both DXA images and radiographs the precision error (both comparative and absolute) of such measurements is clearly of interest.

The questions that arise about the manuscript relate mainly to the data presentation, and poor delineation of the subject groups. This is discussed below.

-----------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Page 4, para 2. Osteoporosis is not a disease of poor mineralisation, (unlike osteomalacia) – the bone is of normal chemical composition, there is just too little of it, when normalised for gender and age. Also, I am not sure if Paget's disease has much bearing on a discussion of bone dimensions, despite the possible long term effect of the disease on long bones. Diagnosis is based on symptoms, radiology (radio-opacity, lytic lesions) and biochemistry (bone turnover). Likewise, treatment efficacy is assessed by a combination of these.

Page 4, Para 3. The “gold standard” for osteoporosis diagnosis is surely evidence of fragility fractures. DXA is a “silver standard”, when fractures are not present.

Pages 6 & 7. The subject groups studied could be clarified better, under a specific "subjects" sub-heading of "Methods". The exact composition and purpose of each subject group would be better declared all in one place in a single paragraph - so the the origins of the different comparisons could be seen.

It is not clear how the numbers of subjects actually "add up"; (i) 50 had two DXA whole body measurements four years apart. These data were used for LPC & RET-based measurements of the same set of variables (short & long term precision); (ii) 20 were measured for these variables by direct anthropometry, and must also have had DXA measurements, because the "direct" method was compared with DXA derived methods ("Results" para of Abstract, plus last para before "Discussion").(iii) Another 20 had leg X-rays, and were assessed for "leg" variables. These also must have had DXA for the same reason as in "(ii)".

Furthermore, under Methods, "Fifty patients with whole body DXA scans....". These were presumably not the normal subjects referred to in the Abstract, unless this was a misuse of the word.

Also, of the 20 subjects who had direct anthropometry, 2 were male - though in the Abstract all study subjects were said to be female. All this needs clarification.
Under methods/Reproducibility it is stated that the subjects were scanned between 1999 and 2003. However, in the first para under "Results" it is stated that "Collectively, subjects were scanned between 1999 and 2005".

----------------------------------------------------------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Abstract & top page 7. Use of the term "inter-observer error". This usually means two separate observers, where the variability of human judgement is being tested. In this case, the same observer appears to have made measurements on images obtained 4 years apart. This is correctly designated as "long-term precision", but is not "inter-observer".

Page 6, para 3. The words “width” and “vertical” need clarifying. For non-densitometry readers, “vertical” might be interpreted (with confusion) as vertically above table, instead of parallel to the long axis of the scanning table.

Figure 1 "B". The number of horizontal scan lines seems to be the measure of the length of femur. Is this robust when the femur is angled?

Figure 2. It is not possible to read the legend for all 4 vertical bars for each bone.

Figure 3. Using a single vertical axis to represent both “mm” and “number of lines” is economical, but it invites a direct visual comparison of the two methods, for each bone, which is not the intention. For RET, 1 mm = 2.06cm actual length, and for LPC, 1 line = 1.3cm actual length. What is important is the RMS SD in “real” units in relation to bone length. Though this is made clearer in the next figure, at least a calibration of LPC and RET units would be a useful addition to the legend of this figure.

In discussion of limitations, the possible effect of BMI on average height of patient above scanning table, and therefore variation in image dimension across the table in fan beam systems should be mentioned, briefly.

----------------------------------------------------------------------------------------------------------------------------------

Discretionary Revisions (which the author can choose to ignore)

Page 7; top para. It would be interesting to determine any statistically significant systematic differences in the measured variables between the 1999 and the 2003 data sets. Since the women were of different (presumably reasonably overlapping) baseline ages, it might be possible to piece together a hybrid cross-sectional & longitudinal age-dependent rate of change for each morphometry variable, though underpowering (due to sample size) might yield significant error. This issue has bearing as well on the assessment of long term precision, since a "true" systematic change in a variable over a particular observation period is not a valid contributor to long term precision error over the same period - since it is not an "error".

The predictive equations (or at least a sample of them) for (say) LPC in determining RET would be of interest, in addition to a comparison of their errors.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests