Image Processing Assessment of Femoral Osteopenia

R.L. Lee, J.E. Dacre, and M.F. James

Visual assessment of femoral osteopenia (the radiographic presentation of osteoporosis) is unreliable. Many of the short-comings of observer grading can be overcome by digital image analysis. Our group has developed algorithms to make automatic assessment of osteopenia from clinical radiographs. Texture Analysis Models (TA) commonly used in image analysis were investigated as measures of osteopenia. Unlike densitometric methods, TA characterizes properties of the *structure* of the image (ie, trabecular patterns). A group of women were analyzed whose subjects ranged from those at risk of osteoporosis ($n = 24$) to normal ($n = 40$). Using an IBM PC, frame-grabber, camera, and light-box, we appraised five statistical TA algorithms for assessment of the femoral neck in standard pelvic radiographs: (1) *Fractal Signature* (FS) describes the image's fractal nature. (2) *Auto-Correlation* of unaltered and Sobel Edge Transformed images (ACSE) measures image spatial self-similarity. (3) *Co-occurrence Matrices* (CM) gives the joint probability of greylevels with distance/direction and describes statistical relationships of image variation. (4) *Textural Spectrum* (TS) neighborhood pixel relationships measure regional directional and pixel-inversion properties. (5) *Eular Numbers* (EN) describe texture by properties (such as connectivity) of binary images. Good reproducibility from repeated analysis of radiographs was shown using both paired t-tests and Altman-Bland's methods. We have shown a correlation between femoral neck bone mineral density (BMD—the "gold standard" of osteoporosis assessment) and textural measures for all five algorithms. Significant measures of osteopenia were: ACSE ($r = 0.6$, $P < .001$), CM ($r = -0.69$, $P < .001$), FS ($r = 0.35$, $P < .01$), TS ($r = 0.52$, $P < .001$) and EN ($r = -0.39$, $P < .01$). Relationships were also found between textural characteristics and age/weight. TA techniques characterize the radiographic changes of bone in osteoporosis. Technology based on these ideas may have a place alongside BMD measurements in the assessment of this condition.

Copyright © 1997 by W.B. Saunders Company

KEY WORDS: osteoporosis, femoral radiographs, digital image processing, texture analysis.

From CHIME, UCL Medical School, Whittington Hospital, London (RLL, JED); and SmithKline Beecham Pharmaceuticals, Essex, UK (MFJ).

Supported by SmithKline Beecham Pharmaceuticals, New Frontiers Science Park, Harlow, Essex, UK, PhD Studentship.

Address reprint requests to R.L. Lee, Centre for Health Informatics & Multiprofessional Education, UCL Medical School, 4th Floor, Archway Wing, Whittington Hospital, London, N19 5NE, UK.

Copyright © 1997 by W.B. Saunders Company
0897-1889/97/1003-1071\$3.00/0

CONVENTIONAL ASSESSMENT of osteoporosis by densitometry has limitations¹ and subjective radiographic scoring² has never proven satisfactory. We are developing an alternative technology to quantify bone matrix from clinical radiographs by a methodology not based on densitometry or subjective scoring.

We applied Texture Analysis (TA), a digital image processing paradigm, to quantifying the trabecular pattern. TA algorithms statistically characterize properties of visual texture ("a global pattern arising from the repetition, either deterministically or randomly, of local sub-patterns"³). Our research compared TA measures and risk of osteoporosis for the upper femur.

MATERIALS AND METHODS

Patients. Hip radiographs from 66 female patients were selected primarily on the basis of left femoral neck Bone Mineral Density (BMD) data estimated by QDR-2000 DEXA (Hologic Inc, Waltham, MA). Patients were divided into four groups by theoretical osteoporosis risk (Table 1).

Texture Analysis Methods

Radiographs were aligned with the left femoral neck horizontal and captured using a Pulnix TM-765 CCD camera (Pulnix Europe Ltd, UK) and ITI OFG digitizing card (Imaging Technology Inc, Bedford, UK) (gamma = 1, automatic gain). Central femoral neck regions (128 by 200 pixels; 16.28 mm by 26.8 mm) were analyzed by custom software running within an Optimas 5.1 image processing system (Optimas Corp, Seattle, WA).

Spatial Distribution Descriptions of Texture—ACSE, CM and TS

TA algorithms generally describe spatial arrangements (distribution) of image intensity. We implemented three spatial distribution algorithms: auto-correlation⁴ (ACSE), co-occurrence matrix (CM),⁴ and texture spectrum (TS).⁵

Auto-correlation. ACSE is a process similar to sliding a radiograph over a copy of itself and measuring total transmitted light. Preprocessing with a Sobel edge algorithm⁶ removes background lighting variations and enhances trabecular lines. We calculated the ACSE perpendicular to the femoral neck (across the trabecular lines), over a range of displacements.

Co-occurrence matrix. CM describes the statistical distribution of pixel intensities found at a specified displacement d and angle θ apart.⁴ The image is high-pass filtered using a FFT Butterworth filter⁷ to remove background lighting variations, and "histogram equalized" from 256 to 64 greylevels⁸ (standardizing the intensity histogram). Entries in a 64 by 64 CM matrix are the summation of row intensity values found (d, θ) from column intensity values. We calculate a "feature" of the matrix

Table 1. Patient Data

Data	Risk Grouping			
	1 (very low)	2 (low)	3 (nonspecific)	4 (high)
No.	5	28	9	24
Mean height (SD)	162.2 mm (5.1)	164.7 mm (6.8)	163.8 mm (8.8)	157.8 mm (4.7)
Mean weight (SD)	70.9 kgs (14.1)	77.1 kgs (7.4)	64.3 kgs (10.1)	57.5 kgs (1.9)
Mean age (SD)	54 (5)	47 (1)	59 (3)	63 (5)
Femoral neck BMD (by DEXA)	>0.95 gms/cm ²	<0.95 gms/cm ² >0.89 gms/cm ²	<0.89 gms/cm ² >0.6 gms/cm ²	<0.6 gms/cm ²

Patients were devided into osteoporosis risk groups on the basis of BMD data using population statistics. Older, frailer, patients with low femoral neck BMD have highest risk.

called the Inverse Difference Moment, with θ perpendicular to the femoral neck.

Texture spectrum. TS requires no preprocessing to give a localized image description derived by comparing a central pixel intensity with its surrounding eight neighbors. The unique configuration of higher and lower neighborhood intensities at each central pixel throughout the image is used to generate a distribution of localized pixel intensity relationships. This distribution is known as the “texture spectrum.” By manipulating the spectrum, we compute a measure related to localized directional properties: degree of direction (DD).

Nonspatial Distribution TA Measures–FS and EN

Textural properties can also be described by methods not directly based on the spatial distributions of intensity. We have applied two such methods: the fractal signature (FS)⁹ and Euler numbers (EN).¹⁰

Fractal signature. FS relates the “fractal” dimension with scale by an extension of the fractal dimension philosophy. Texturally, the FS is a measure of information at different image scales, and thus the strength and spatial size(s) of texture. Our FS algorithm was implemented by a morphological method,⁹ which calculates the “Hausdorff-Besicovich Dimension” H (iteratively) from the log of the change in “morphological gradient”¹¹ over a specified scale range. Fractal dimension D is: $D = 3 - H$.

EN. EN is a measure of structural properties of the texture quite separate from any spatial properties of size, distribution, strength, or orientation. We used a simple index derived by high-pass filtering with an averaging filter and conversion from 256 greyscale to binary. A network-like appearance is generated by this process, from which we calculated the number of “dark,” unconnected image regions, called “MESH.”¹⁰

Statistics

Reproducibility was assessed for all algorithms. On two occasions, films were captured and analyzed. Paired Students t -test and plots of mean versus difference (Altman-Bland plots¹²) were used to test reproducibility, with Coefficient of Variation (COV)¹³ and Signal-to-Noise ratio (S/N: proportion of residual standard deviation (SD) to explained SD) calculations of intrasubject measurement error.

MS Excel 5.0 (Microsoft Corp) and SPSS 6.1 (SPSS Inc, Chicago, IL) were used for data analysis.

RESULTS

Measurement reproducibility. Paired t -test showed no significant difference between visits. However, four subjects (6.2%) were found to lie consistently outside 2 SD in Altman-Bland plots for all TA measures. On further investigation, these four patients were found to have poor quality radiographs and were difficult to capture adequately with the CCD camera. After exclusion of these subjects’ reproducibility improved.

Measurement error (Table 2) is different by COV and S/N ratio. We found that the COV data is misleading due to a dependence on the measurement mean: the higher the mean in regard to the SD, the lower the apparent error. For example, ACSE has a low COV error (1%) because it has a high mean, but has a S/N ratio that is much higher (49%). This effect is demonstrable with FS, which could be reported as H , not D , and COV error for H is 24%, not 4%, however S/N remains unchanged by the use of H .

Correlations. TA measures correlated with BMD in all cases (Table 3). Controlling for height and weight, these correlations were reduced; FS was not significantly correlated. Further controlling for age reduced the correlations, with only CM significantly correlated with BMD.

High BMD was associated with the younger, heavier patients (Table 4), with our selection of patients by risk heightening this relationship.

Table 2. Measurement Error

	FS	ACSE	CM	TS	EN
COV (%)	4	1	13	3	9
S/N ratio (%)	45	49	58	59	54

Intrasubject measurement error as determined by COV (Residual SD/Mean $\times$ 100%) and Signal-to-Noise Ratio (Residual SD/Explained SD $\times$ 100%), from analysis of variance of two visits on the same films.

Table 3. TA Correlations With BMD

Statistic	TA Measure				
	FS	ACSE	CM	TS	EN
Pearson's <i>r</i> correlation	0.35 (<i>P</i> < .01)	0.6 (<i>P</i> < .001)	−0.69 (<i>P</i> < .001)	0.52 (<i>P</i> < .001)	−0.39 (<i>P</i> < .01)
Corrected for height and weight	nonsignificant	0.25 (<i>P</i> < .05)	−0.55 (<i>P</i> < .001)	0.39 (<i>P</i> < .01)	−0.32 (<i>P</i> < .01)
Corrected for height, weight, and age	nonsignificant	nonsignificant	−0.37 (<i>P</i> < .05)	nonsignificant	nonsignificant

Correlations of five TA measures with DEXA BMD (*n* = 62). Pearson's *r* correlations and partial correlations controlling for height, weight, and age.

Three examples of digitized upper femoral radiographs chosen at extremes of the CM measure range are shown in Fig 1. In correspondence with the grading system of Singh et al,² the examples show a change in the trabecular pattern from fine and strong, becoming coarser, and finally disappearing entirely with increasing osteoporosis.

DISCUSSION

We have investigated TA as a method of determining trabecular bone architecture visible in hip radiographs. Our TA algorithms (which are representative of available techniques) have been applied to the femoral neck because of its high morbidity in osteoporosis.

The data show significant correlations of textural measures with femoral neck bone mineral density but, although highly significant, these correlation *r* values are moderate, which may be due to the disparate properties measured (mineralization compared with mineralized matrix). Height, weight, and age also correlated with BMD. Yet, even adjusting for height and weight, four of five TA measures were still significantly related to BMD. Further adjusting for all factors, only one (CM) textural measure appears to have an independent relationship to BMD.

This residual correlation of CM and BMD, after controlling for height, weight, and age is almost certainly due to the known relationship of BMD and matrix loss in osteoporosis. We conclude, therefore, that TA is capable of quantifying trabecular *structure* and is a potentially useful method of augmenting densitometry.

Measurement reproducibility on good quality radiographs was shown and we have estimated the intrasubject measurement error. COV and S/N ratio error have been shown to disagree because COV is dependent on a mean value, making the S/N a more appropriate method of comparing algorithms. Although S/N ratio is still dependent on the chosen patient range, the patient range is consistent between algorithms, whereas TA means are not.

On the basis of *r* alone, the CM algorithm is the best measure. The choice is more complex, however, when one considers measurement error. With broadly similar S/N for all the algorithms, spatial distribution TA as implemented with ACSE and CM, characterizes the visual change in bone matrix more adeptly than TS, FS, or EN. It is worth noting, perhaps, that ACSE and CM are very closely related, and that the spatial displacements for these two algorithms was similar (data not shown).

Given the high S/N ratios of 45% to 60%, TA would not appear adequate as a tool of diagnosis without further investigation of the sources of measurement error. However, used to augment densitometry and/or as a measure in population studies, these errors are less important.

Finally, the visual appearance of films (Fig 1) with a range of CM measures corresponds to the grading of Singh et al,² giving further, intuitive support to the hypothesis that TA can measure bone matrix properties from radiographs.

This study has demonstrated a definite relationship between BMD and TA values. Further studies are underway to consolidate the role of TA in the imaging of osteoporosis.

Table 4. BMD Correlations With Patient Data

	Height	Weight	Age
Correlation with BMD	0.44 (<i>P</i> < .001)	0.75 (<i>P</i> < .001)	−0.86 (<i>P</i> < .001)

Correlations of height, weight, and age with DEXA BMD (*n* = 66).

ACKNOWLEDGMENT

SmithKline Beecham Pharmaceuticals, Harlow, UK (CASE Studentship, equipment and software). Dr T. Spector, The Twin Research and Osteoporosis Unit, St Thomas' Hospital, UK (radiographs and patient data). Dr AJ Lynch, UMDS, Guy's Hospital, UK (Fractal Signature Algorithm).

Fig 1. Visual appearance of TA grades. Three examples of femoral neck images in the CM textural measure range. (A) Low CM (high BMD): fine, strong trabecular pattern. (B) Median CM (median BMD): course trabecular pattern. (C) High CM (low BMD): few, weak trabecular pattern and general low contrast between soft-tissue and bone.

REFERENCES

1. Wallach S, Feinblatt JD, Avioli LV: The Bone "Quality" Problem. *Calcif Tissue Int* 51:169-172, 1992
2. Singh M, Nagrath AR, Maini PS: Changes in the trabecular pattern of the upper end of the femur as an index of osteoporosis. *J Bone Joint Surg* 52-A:457-467, 1970
3. Zucker SW, Terzopoulos D: Finding structure in co-occurrence matrices for texture analysis. *Computer Graphics and Image Processing* 12:286-308, 1980
4. Sonka M, Hlavac V, Boyle R: Texture, in Sonka M, Hlavac V, Boyle R (eds): *Image Processing, Analysis and Machine Vision*. London, UK, Chapman and Hall, 1993, pp 480-485
5. Dong-Chen H, Li W: Texture features based on texture spectrum. *Pattern Recognition* 24:391-399, 1991
6. Gonzalez RC, Woods RE: Image segmentation, in Gonzalez RC, Woods RE (eds): *Digital Image Processing*. New York, NY, Addison Wesley, 1992, pp 416-429
7. Gonzalez RC, Woods RE: Image Enhancement, in Gonzalez RC, Woods RE (eds): *Digital Image Processing*. New York, NY, Addison Wesley, 1992, pp 209-213
8. Rosenfeld A, Kak AC: Enhancement, in Rosenfeld A, Kak AC (eds): *Digital Picture Processing*. New York, NY, Academic, 1976, pp 173-175
9. Lynch JA, Hawkes DJ, Buckland-Wright JC: A robust and accurate method for calculating the fractal signature of texture in macroradiographs of osteoarthritic knees. *Med Inform* 16:241-251, 1991
10. Geraets WGM, Van Der Stelt PF, Netelenbos CJ, et al: A new method for automatic recognition of the radiographic trabecular pattern. *J Bone Miner Res* 5:227-233, 1990
11. Gonzalez RC, Woods RE: Representation and Description, in Gonzalez RC, Woods RE (eds): *Digital Image Processing*. New York, NY, Addison Wesley, 1992, pp 548-560
12. Altman DG, Bland JM: Measurements in medicine: The analysis of method comparison studies. *The Statistician* 32:307-317, 1983
13. Bland M: Clinical measurement, in Bland M (ed): *An Introduction to Medical Statistics* (ed 2). Oxford, UK, Oxford Medical Publications, 1995, pp 266-269