Published in final edited form as:

J Clin Densitom. 2008 ; 11(2): 211–220. doi:10.1016/j.jocd.2007.10.004.

REPRODUCIBILITY AND SOURCES OF VARIABILITY IN RADIOGRAPHIC TEXTURE ANALYSIS (RTA) OF DENSITOMETRIC CALCANEAL IMAGES

Tamara J. Vokes, M.D.¹, Ann Pham, B.S.¹, Joel Wilkie, B.S.², Masha Kocherginsky, Ph.D.³, Siu-Ling Ma, M.D.¹, Michael Chinander, Ph.D.², Theodore Karrison, Ph.D.³, Octavia Bris, and Maryellen L. Giger, Ph.D.²

¹Department of Medicine, University of Chicago

²Department of Radiology, University of Chicago

³Department of Health Studies, University of Chicago

Abstract

Radiographic texture analysis (RTA) is a computerized analysis of the spatial pattern of radiographic images used as a way of evaluating bone structure. We have shown that RTA performed on high resolution heel images obtained using a portable densitometer differentiates subjects with and without osteoporotic fractures. In the present study, short-term precision of RTA was examined on densitometric heel images obtained from 33 subjects scanned 8 times each, with 3 observers placing a region of interest (ROI) 3 times on each image. The long-term precision was examined on images obtained from 10 subjects 3 times on each of 3 days separated by one week, with 2 observers placing an ROI on each image. The RTA features examined included the RMS (root mean square variation, a measure of the contrast between the light and dark areas of the image), the FMP (first moment of the power spectrum, a measure of the spatial frequency of the trabecular pattern), and Minkowski fractal (MINK, a measure of roughness/smoothness of the trabecular pattern). The precision of the RTA features expressed as coefficient of variation ranged between the lowest of 0.5–0.7% for MINK and the highest of 14–16% for RMS. The short- and long-term precision was similar, and was not significantly influenced by repositioning and rescanning, or by ROI placement by the same or different observers. Significant sources of variability of RTA were the between-subject differences and differences between regions of the heel, but not differences due to repositioning, rescanning in the same position, or ROI placement by the same or different observers. We conclude that technical aspects of image acquisition and processing are adequate to allow further development of RTA of the densitometric images for clinical application as a method for non-invasive assessment of bone structure.

Corresponding author: Tamara Vokes, Section of Endocrinology, University of Chicago, 5841 S. Maryland, MC1027, Chicago. IL 60637, Tel: 773-702-5718, tvokes@uchicago.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Presented in part at the 2006 Annual Meeting of the American Society for Bone and Mineral Research

INTRODUCTION

Radiographic texture analysis (RTA) is a method for assessing the spatial pattern or texture of radiographs. It has been applied to various medical images including mammograms, chest CT, and also to bone radiographs, which like the bone that they image, have naturally structured trabecular pattern.

Osteoporotic fractures are common and have significant impact on morbidity and loss of quality of life in the aging population. It is estimated that the number of patients with fractures will continue to increase due to the aging of the population coupled with increase in the secular trend in fracture incidence (1). Currently, the main challenge in osteoporosis research and clinical practice is to improve the assessment of fracture risk in order to help clinicians to better select patients for costly and potentially toxic therapy. Developing novel methods for assessment of fragility and predicting fracture risk is of utmost importance at this time.

Measurement of bone mineral density (BMD) using DXA is the predominant approach for estimating bone mass. Although BMD predicts fracture risk reasonably well in population studies (2–4), it performs less well in individuals as there is considerable overlap in BMD values of subjects who do and those who do not sustain fragility fractures (2,3,5,6). This is because fragility is determined not only by bone quantity (measured as bone density), but also by its “quality” (7), which cannot be easily assessed using current diagnostic methods. Although it is not completely clear what constitutes bone quality (8), most agree that bone architecture or trabecular structure represent its significant component. The most direct way of assessing bone structure is through bone biopsy which is invasive, costly, associated with long turnaround time, and available at only a few select institutions. Bone microarchitecture can be studied using high-resolution 3-D imaging techniques such as microCT, MRI, and XtremeCT (9–14), which have provided important information about bone structure underlying fragility and its changes with aging (12,15). Although these methods have the advantage of three-dimensional assessment they are not practical for widespread clinical use due to high cost and/or radiation exposure, limited availability or competing demand for equipment use for other clinical purposes. An alternative approach is assessment of radiographic pattern or texture of radiographs. Several groups have shown that computerized radiographic texture analysis performed on digitized radiographs can be used to characterize fragility: RTA of the spine radiographs differentiated patients with and without vertebral fractures better than spine BMD (16,17); RTA of femoral neck specimens predicted bone strength *in vitro* (18,19); RTA of the proximal femur radiographs differentiated patients with hip fractures from age matched controls (20); fractal analysis of the calcaneus radiographs differentiated subjects with and without osteoporotic fractures (21,22) when used alone, and even better when combined with anisotropy indices derived from fast Fourier transform (23,24). The same analysis, performed on cadaveric calcaneal specimens *ex vivo*, correlated with histomorphometric findings (25) and with biomechanical properties (26).

We have applied the RTA methods previously used for radiographs to calcaneal images obtained using a portable densitometer specially equipped with high resolution camera (27–29), and found that RTA of densitometric calcaneal images performed well in differentiating patients with and without prevalent vertebral fractures (29), with the magnitude of difference between fracture and non-fracture patients that is quite similar to what has been reported for calcaneal radiographs (21,22). In addition, we found that RTA performed as well as hip BMD in the task of differentiating patients with and without fractures. Furthermore, when we combined BMD and RTA in a logistic regression model with presence of vertebral fractures as a binary outcome, we found that both BMD and RTA had significant effect (29), suggesting that they captured different aspects of bone fragility, presumably BMD capturing bone mass or quantity and RTA some aspects of bone structure or quality. However, RTA of densitometric

bone images is a new method which has not been extensively evaluated. Specifically, it is not known which factors impact the precision and/or contribute to the variability of the measurement. In this study we examined the reproducibility and sources of variability of the following four RTA features: Root Mean Square variation (RMS), standard deviation of RMS (sdRMS), First Moment of the Power spectrum (FMP), and Minkowski fractal (MINK).

What does RTA measure?

RTA is a computerized analysis of the trabecular pattern of a region-of interest (ROI) within a bone image. Currently our analyses can be broadly divided into two categories. One involves the Fourier transformation of the image data and the other represents a form of fractal analysis (16–19,30). The Fourier analysis (17) yields two descriptors: RMS (root mean square variation) and FMP (first moment of the power spectrum), as well as related directional measures. RMS is a measure of the variability in the radiographic texture pattern, the relative difference in the contrast between light and dark areas, such that a bone image with a washed-out appearance due to loss of trabecular structure (as that seen in osteoporosis), will have a low value for RMS because there will be relatively little contrast between lighter and darker areas of the image (17,29). During the RTA, the Fourier spectrum is divided into 24 angular sectors (of 15 degrees each) to yield another descriptor, sdRMS, (standard deviation of the RMS), which is the average difference in the RMS from the 24 angular sectors and represents a measure of anisotropy. In previous study, we found that sdRMS was the best RMS feature for differentiating patients with and without fractures (29). FMP is the calculation of the first moment of the power spectrum, which describes the spatial frequency of the radiographic pattern and the underlying trabecular structure. Thus, the FMP corresponds to the coarseness or fineness of the radiographic texture pattern. High level of FMP indicates a structure with thin and closely spaced trabeculae, such as those in osteoporotic bone. In contrast, low FMP indicates widely spaced dark areas usually corresponding to strong thick trabecular structure.

For the fractal analysis we used Minkowski fractal, MINK (16,18,19,29), which in our preliminary studies performed better on densitometric images than other forms of fractal analysis (27,28). Mathematical fractal is a measure of self-similarity such that a perfectly fractal object has the same pattern at different levels of magnification with fractal number representing the ratio of the log of magnification to the log of size. Fractal analysis is often applied to natural objects which have partially fractal nature, including bone images. In the simple form of fractal analysis, the gray level of each pixel is regarded as height and the pixel size as the surface area. Adjacent pixels are combined to yield progressively larger pixel sizes with a new gray level averaged from the combined pixels. The fractal dimension corresponds to the slope of the least square-fitted line relating log surface area to log pixel size at different magnifications. MINK employs morphological filtering to generate “volume” between dilated and eroded versions of each ROI at multiple (usually 8–10) scales. Minkowski fractal dimension is calculated from the slope of the least-square fitted line relating log volume (difference between dilated and eroded versions of each ROI) and log scale (magnification). Similar to FMP, MINK is also a measure of the coarseness or fineness of the radiographic pattern with the higher value representing a more washed-out or smooth pattern and the lower value representing a coarser, stronger textural pattern.

METHODS

1. Short term precision and variability

Subjects—The 33 participants (2 males) included in this analysis were representative of our densitometry populations. There were 13 African Americans, 1 Asian, 1 Hispanic and 18 Caucasians. They had a wide range of ages (21 to 87 years, mean±standard deviation of 61.5 ± 16.5 years), heel sizes and BMD values (0.189 to 0.873gm/cm², mean of 0.439±0.148; T-

scores of -3.7 to $+4.8$, mean of -0.8 ± 1.8), height (57 to 72 inches, mean of 63.5 ± 3.7) and weight (94 to 256 lbs, mean 152.5 ± 5.7).

Subject imaging procedures—The procedures used to obtain the data are schematically shown in Figure 1A. Thirty three volunteers had 8 scans of the left calcaneus on the same day. The scans were obtained using PIXI (Peripheral Instantaneous X-ray Imager, GE Medical Systems, Madison, Wisconsin). The subjects were instructed to place the foot in the cradle of the densitometer with the radiology technologist adjusting the position so that the medial maleolus was snug next to the wall of the cradle and the tip of the heel positioned as deeply as possible in the round groove in the center of the bottom of the positioner (Figure 2A). The subjects were instructed not to move their foot while 2 scans were obtained with the usual time between the scans of 1–2 minutes. After 2 scans, the subjects were instructed to take their foot out of the scanner and then place it back, with the process repeated 3 more times. (It is not practical to attempt to obtain more than 2 images without repositioning because the subjects will often become uncomfortable and un-intentionally move the foot.) Overall the protocol yielded 8 images per subject resulting from 4 positions with 2 images in each position (Figure 1A). For each scan an electronic digital image file was generated and used for RTA.

Image acquisition and selection of ROI—The PIXI was especially equipped with a high spatial-resolution camera that obtains a dual energy image of the calcaneus for the calculation of BMD measurement. The PIXI unit provides a 512 by 512 pixel high- and low-energy (80 and 55 kVp) 12-bit images with 0.2 mm pixels. It uses Gd_2O_2S phosphorus detector coupled with a CCD camera via the lens system. Low energy images were used for RTA (Figure 2B).

On each image each of 3 observers blinded to the identity of the subjects, placed an ROI 3 times (Figure 2C). The ROI selection was performed using a display program written in IDL (ITT Visual Information Solution, Boulder, Colorado), which allows visualization of the radiographic texture (Figure 2D) for placement of the ROI in the region with most prominent trabecular structure. An ROI of 128×128 pixels (Figure 2C) was placed on all images regardless of the bone size. All of the observers were advised to place the ROI as close to the talus as possible while avoiding the talus and bone edges (Figure 2C). This placement was based on our earlier findings as well as that of others that bone trabecular structure is richer in the proximal than in the distal parts of the calcaneus (29,31–34).

For the RTA, each 128×128 pixels ROI was subsequently divided into 5 quadrants of 64×64 pixels (ROI1 through ROI5) as shown in Figure 1 and Figure 2.

RTA (Radiographic Texture Analysis)—RTA was performed as previously described (16,17,27–29). Prior to Fourier analysis, background trend correction using a second-order polynomial least square fit is performed on each ROI. The magnitude and spatial frequency of the trabecular bone pattern of each image is characterized using the Fourier transform of each ROI. To characterize the magnitude of the trabecular texture pattern the root-mean-square (RMS) variation is calculated using the following formula:

$$RMS = \sqrt{\sum_m \sum_n |F_{m,n}|^2}$$

To characterize the frequency or coarseness of the trabecular pattern, the first moment of the power spectrum (FMP) is calculated using the following formula:

$$FMP = \frac{\sum_m \sum_n \sqrt{m^2 + n^2} |F_{m,n}|^2}{\sum_m \sum_n |F_{m,n}|^2}$$

For both RMS and FMP formulae, $F_{m,n}$ corresponds to the Fourier transform of the background corrected ROI of the trabecular pattern, and m and n are the indices of the ROI array. RMS represents the variation in terms of relative image intensity and is expressed in grayscale level. FMP is expressed in terms of spatial frequency (in cycles/mm).

Minkowski fractal dimension $D_M[f]$, was calculated using the following formula:

$$D_M[f] = \lim_{\varepsilon \rightarrow 0} \frac{\log [V_g(\varepsilon) / \varepsilon^3]}{\log (1/\varepsilon)},$$

where f corresponds to the ROI image data. For a structuring element g at scale ε , $V_g(\varepsilon)$ is the "volume" between two processed versions of f obtained using morphological operators. The volume $V_g(\varepsilon)$ is computed by

$$V_g(\varepsilon) = \sum_{m=0}^{64} \sum_{n=0}^{64} \{(f \oplus \varepsilon g) - (f \otimes \varepsilon g)\},$$

where $(f \oplus \varepsilon g)$ and $(f \otimes \varepsilon g)$ are the dilated version and the eroded version, respectively, of the image obtained using a structuring element g at scale ε . Finding the slope of the least-squares fitted line between $\log[V_g(\varepsilon)/\varepsilon^3]$ and $\log(1/\varepsilon)$ gives the estimated fractal dimension.

Analysis of the Data—Precision (coefficients of variation, CV) for BMD and texture features were calculated using a modified version of the precision calculator developed by Dr. Michael Lewiecki for the ISCD (35), where CV is calculated for each of n subjects, and the root mean square average is found using

$$\text{averageCV}(\%) = \sqrt{\frac{1}{n} \sum_{j=1}^n \frac{SD_j^2}{\bar{x}_j^2}} * 100,$$

where SD_j is the standard deviation of the subject j and $\bar{x}_j$ is the average of the measurements for the subject j (36). Intra - and inter-observer CVs, as well as the CVs for variability due to positioning were calculated separately for each of the 5 ROI sub-regions (shown on Figure 1 and Figure 2). Intra-observer CV reflects the variability of the repeated ROI placements by the same observer, and was estimated using the measurements from 3 ROI placements by the same observer on the first of the 8 images for each patient. Inter-observer CV reflects the degree of between-observer variability and was estimated based on the first ROI placed by each of the 3 observers on the first of the 8 images from each subject.

We further examined the degree to which each of the following components contributed to overall variability:

Subject - the between-subject difference in RTA and BMD between the 33 participants.

Position - 4 foot positions for each subject

Re-imaging - 2nd image in the same position

ROI - placed by each observer 3 times in each image.

ROI sub-region - for each ROI placement, 5 ROI sub-regions were defined as in Fig. 1 and 2.

Observer - 3 observers placed an ROI 3 times on each image.

Variance components for the nested random effects model were estimated using PROC VARCOMP in the SAS software. Separate models were fitted for each of the four RTA features (response variables). The nested effects were subject, position within subject, and re-imaging

within position and subject. The observer random effect was included as a crossed random effect. A model including the ROI sub-region as a crossed random effect was fitted (Table 2), and additional models with identical random effects structure were fitted for each of the 5 ROI sub-regions (Table 3). Due to the differences in scales in the four RTA features, the model-based proportion of variability attributable to each factor, estimated by dividing each variance component by the total variance, is reported in Table 2 and Table 3. Another set of models using only the first image obtained in each position and the ratings of only one observer (thus mimicking the data obtained for BMD) was used to estimate the variability in the RTA features attributable to subject and position, and to compare them to those of BMD (Table 4). These models were fitted separately for each ROI, and only the result for ROI 5 are shown in Table 4. In order to compare the three observers and the 5 ROI's, a mixed effects analysis of variance model was fitted with the nested random effects of subject, position, image, but treating the observer and ROI as a fixed effects. This model was estimated using PROC MIXED in the SAS software and the pairwise comparisons between the 5 ROIs were performed using least square means with the Tukey-Kramer multiple comparison adjustment.

2. Long-term Precision

For this study 10 volunteers were recruited among the clinic employees. The subjects included 9 women and 1 man; 1 African American, 5 Asians, and 4 Caucasians. The age range was 15 to 53 years (mean±standard deviation of 36.1 ± 8.5 years), the BMD 0.352 to 0.712 gm/cm² (0.554 ± 0.129); T-scores of -1.9 to +2.5 (mean of 0.5 ± 1.5), height 59 to 72 inches (64 ± 4), and weight 110 to 230 lbs (150 ± 42). One day a week for 3 weeks each subject had 3 scans with repositioning after each scan resulting in total of 9 images per subject (Figure 1B). The image acquisition and ROI placement were as described for short term precision except that 2 observers placed ROI once on each image. Long-term precision was estimated using precision calculator as described in the data analysis for short-term precision study. The inter-observer variability was estimated using measurements from the ROI placement by two observers on the first image from each subject. Long term precision was calculated for each observer using measurements from the first scan obtained from each subject once a week for 3 weeks, and by using all 9 scans from each subject (3 scans every week for 3 weeks).

RESULTS

Reproducibility of texture features

Short-term precision—Table 1 shows the coefficients of variation (CV) for four RTA features examined in the short term precision study. The inter-observer CVs were similar to, and in the range of intra-observer CVs. The highest CVs were observed for sdrms, and the smallest for MINK. There were no significant differences in the variability of the RTA measurements between the 5 ROI sub-regions (Table 1), although they tended to be slightly higher for ROI1.

The effect of repositioning and repeated (duplicate) scanning without repositioning was examined by comparing the first ROI selection by each of the 3 observers on all 8 images and on 1st of the 2 images obtained in each position (images 1, 3, 5, and 7) in each subject. The CVs derived from 8 images and those derived from 4 images were very similar. Consequently, only the CVs obtained from all 8 images are included in Table 1. The CVs from 8 images were generally in the range of values for intra- and inter-observer CVs. It should be noted that there was no variability due to RTA per se – RTA performed 3 times on one image with one ROI selection produced identical results (data not shown).

The CV for the heel BMD measurement obtained from 8 scans was 2.75%, while that obtained from the 1st scan in each of 4 positions was 2.7%. This unusually high variability was due to

several images (9 out of 256) with obviously poor positioning. When these were eliminated from precision calculation, the CV was 2.0%. For the first 2 images obtained from each subject without repositioning, the precision was 1.5%, while for 2 images obtained with repositioning, the precision was 2.0%. The long term in vitro precision using phantom scanned daily over 3 years was 1.05%.

Sources of variability in RTA measurements

As shown in Table 2, for all features, the between-subject variability was much greater than the variability due to positioning, or ROI placement by the same or by different observers. Another significant source of variability was due to the difference between the 5 ROI sub-regions. For Fourier based features, RMS, sdRMS and FMP, 61–74% of variability was attributable to differences between the subjects, and 2–16% was due to differences between the 5 sub-regions of the ROI. Variability due to positioning, repeated scanning, and ROI placement by the same or different observers each accounted for less than 1% of the variability in the measurement. For these Fourier features, there was residual variability of 21–30%. Variability due to the 3 ROI placements by the same observer is contained in the residual (error) term.

These models suggest that there is still residual variability in RTA that could not be accounted for by any of the factors examined. Residual variance is highest for MINK (49%), with the between-subject variability of only 20%, and the variability due to differences between the 5 ROI sub-regions of 28%. The residual variability was lower and the variability due to subject was higher when it was examined for each ROI sub-region separately (Table 3).

We also compared the sources of variability of BMD and RTA measurements (Tables 4) by examining only factors that could contribute to the BMD variability, namely the variability due to subject, and repeated scanning with and without repositioning. For this purpose we used RTA measurements resulting from the first (out of 3) ROI placement by one observer on each of 8 images from 33 subjects. Again, because the calculation of variability is limited by the degrees of freedom, variability due to rescanning without repositioning can not be derived separately and is contained in the residual (error) term. (It is safe to assume however, that the variability of RTA due to repositioning and due to re-scanning in the same position are similar based on findings presented in Table 3.) The variability of BMD measurements was largely due to differences between the subjects, with rescanning with and without repositioning contributing only about 0.5% of the total variability. For Fourier based features the majority of variability was also due to differences between the subjects. It should be noted that although we used only the first ROI placement by one of the observers, there is still variability due to ROI placement which is included in and accounts for relatively higher residual variance (error). This is particularly true for MINK, which has the highest residual variance because this feature is most influenced by even small differences in ROI placement.

Differences in RTA results between the observers and between 5 ROI sub-regions

As expected based on the low variability due to ROI placement by the same or different observers, the RTA measurements did not differ significantly between the 3 observers, except for iRMS which was slightly higher in ROIs placed by one of the observers. In contrast, there were clear differences between the 5 sub-regions in RTA values such that MINK was significantly lower in the ROI sub-region 1 compared to all other sub-regions ($p < 0.0001$), while RMS based measurements were lower ($p < 0.0001$) in ROI 4.

Long-term precision

Table 5 shows the coefficients of variations resulting from scanning 10 subjects over 3 weeks. The long—term precision (Table 5) was similar to the short-term precision described earlier

(Table 1). The precision using results from the first scan of each week and using all 9 scans (3 scans from each of 3 weeks) from each subject is given for each observer separately. The results show that the precision is similar between observers and also similar whether just the first or all 3 scans from each week are used.

DISCUSSION

In order to develop RTA performed on densitometric heel images for eventual clinical application we examined the reproducibility and the sources of variability of RTA measurements. We found that both short- and long-term precision of Fourier based RTA features RMS and FMP, were in the same range as the precision of calcaneal BMD measurements (Table 1 and Table 5). The variability of sdRMS was considerably higher, presumably because this feature is highly dependent on the directionality of the underlying bone structure which is likely to vary even within small regions of the image and thus be more influenced by small differences in the position of the ROI or the region of the heel examined. However, the variability due to technical factors such as positioning, or ROI selection by the same or different observers, relative to the variability due to differences between subjects, was similar for sdRMS and for other Fourier based features. This indicates that sdRMS would be at least as useful as the other features in assessing structural parameters. Consistent with this, we previously reported that sdRMS separated subjects with and without fractures quite well (29), suggesting that sdRMS captures relevant properties of bone structure. On the other end of the precision spectrum is fractal measurement MINK which had a very low coefficient of variation (Table 1 and Table 5), reflecting narrow range of values resulting from the nature of the calculation for this feature (log-log ratio). Our observations that the inter-observer CVs were similar to, and in the range of intra-observer CVs (Table 1 and Table 5), suggest that ROI placement by different observers is not a significant source of analytical error. Similarly, we found that CVs from 8 images from the short-term precision study (Table 1) and 9 images from the long-term precision study (Table 5) were generally in the range of values for intra- and inter-observer CVs, suggesting that repositioning and repeated scanning is not a significant source of analytical error in radiographic texture analysis. In contrast, the heel region used for RTA may be important source of error since we find slightly higher variability in areas closest to the talus which are likely to have more variability in the texture depending on whether a denser portion of the calcaneus located just below the talus shadow was included in the ROI or not. We also found that the absolute values of RTA features differed between ROI sub-regions. Thus, MINK values were significantly lower in the ROI1 compared to all other sub-regions ($p < 0.0001$), presumably due to denser trabecular structure with less variability in texture (Figure 2). In contrast, RMS based measurements were lower ($p < 0.0001$) in ROI 4, probably because this sub-region is most “washed-out” being furthest from the more proximal area with rich trabecular structure. This is consistent with our observation that the best separation between patients with and without fractures is achieved using RTA performed on ROI5 sub-region (29), which is not affected by talus shadow yet has sufficiently rich trabecular structure to characterize it radiographically. Similarly, other investigators using MRI have found that osteoporotic subjects with and without fractures are better differentiated on the proximal rather than distal portion of the calcaneus (33,34).

Our precision results are similar to what has been reported by investigators who performed the RTA on calcaneal radiographs (37,38). Thus, using fractal analysis measuring the Maximum Likelihood Estimator based on Fractional Brownian Motion model, Benhamou et al. reported intra-observer, inter-observer, and transfer of region of interest coefficients of variation (CV) of 0.61, 0.68 and 0.53%, similar to our findings of 0.5 to 0.9% for CV for MINK, with lower CVs (0.5–0.7%) for ROI4 and ROI5 sub-regions, which correspond to the location of the ROI used in the study or Benhamou et al (37). Similarly, our CVs for Fourier based texture features RMS and FMP were of the order of magnitude of CV of 2.8 reported for an anisotropy indexed

based on Fast Fourier transform of the calcaneal radiographs (23). It should be noted, however, that our precision study included considerably larger number of subjects than the studies using radiographs. In addition, we included subjects with wide variability in heel anatomy with at least some subject's having heel size and shape for which it is difficult to obtain an optimal image using a densitometer which has a fixed size of the positioner.

We found that the CV of 2–3% for the BMD measurements of our 33 subjects were higher than previously reported values between 1.4 and 1.7% (39,40). Despite that, the BMD variability due to these technical factors (repositioning and re-imaging in the same position) was a very small component of total variability, with differences between subjects still contributing 98.9% to the total variance (Table 4). In contrast, the variability due to differences between subjects is somewhat lower for the Fourier features with residual variance (error term) being higher (Table 4). This is most likely due to the fact that BMD measurement is integrated over a relatively large area of the heel, while RTA quantifies texture pattern in a very small area with the RTA features reflecting a much more variable structure. Consequently, even small changes in ROI placement will produce larger differences in RTA. While this is associated with poorer precision, the ability of RTA to capture subtle differences in bone structure may result in greater sensitivity of RTA to detect structural differences underlying bone fragility. Further studies are needed to examine this possibility.

Among the RTA features examined, MINK had the highest residual variability. This is largely explained by extreme sensitivity of MINK to the region analyzed since the residual variability was lower and the variability due to subject was higher when it was examined for each ROI sub-region separately (Table 3). This variability is not due to RTA *per se*, since repeating analysis on the same image with the same ROI placement produces identical results (data not shown). Thus, it seems that MINK is most sensitive to subtle differences in texture pattern between even closely positioned areas of the radiographic image. Interestingly, MINK which has the highest residual variance (error term in Table 2 to Table 4), i.e. the percent of total variance that is not attributable to any of the factors examined, performed better than the Fourier based features in the task of differentiating subjects with and without vertebral fractures in our study of postmenopausal women (29). This suggests that some of the residual variability in MINK is due to differences in bone structure related to bone strength. This possibility, however, can not be evaluated in the present study due to small number of subjects.

Sources of error in bone structural parameters have also been examined for micro-MRI based techniques in 6 healthy volunteers each having 3 scans of distal radius and distal tibia (41). The CVs for different parameters varied between 4 and 7%, again in the range of CVs that we observed for our Fourier based RTA features. The main sources of variability in the MRI study were involuntary patient motion, and failure to accurately match the analysis volume in the repeated studies. While patient motion is not a factor in our and other radiography-based methods, matching of the analysis volume is analogous to the ROI placement in our study. We also observe that because of the variability in bone structural pattern within relatively small anatomic area, the alterations in ROI position, like the selection of analysis volume for MRI based measurements, have a significant effect on the results.

A potential limitation of our study is that the densitometer used in our study is no longer produced and that it required special equipment to generate high resolution images. However, our study is a “proof of principle” that a portable densitometer can be used to assess both bone mass and bone structure. Encouraging results from studies like ours may lead to development of appropriate equipment for wide spread clinical use and population screening.

In summary, our study of RTA performed on densitometric heel images shows reasonably good reproducibility, comparable to other techniques that assess bone structure. Further, we find that

variability due to technical aspects of image acquisition and ROI selection is relatively low relative to the variability resulting from the difference between subjects, which indicates potential utility of this method for examining differences in radiographic texture and presumably in underlying bone structure in physiologic and pathologic conditions affecting bone.

Acknowledgements

Funding support: supported by grants K23 AR048205-01A1, AR42739-S1, and AR42739 from the National Institute of Health

References

1. Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. *Lancet* 2002;359(9319):1761–1767. [PubMed: 12049882]
2. Cummings SR, Black DM, Nevitt MC, Browner W, Cauley J, Ensrud K, Genant HK, Palermo L, Scott J, Vogt TM. Bone density at various sites for prediction of hip fractures. The Study of Osteoporotic Fractures Research Group. *Lancet* 1993;341(8837):72–75. [PubMed: 8093403]
3. Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ* 1996;312(7041):1254–1259. [PubMed: 8634613]
4. Seeley DG, Browner WS, Nevitt MC, Genant HK, Scott JC, Cummings SR. Which fractures are associated with low appendicular bone mass in elderly women? The Study of Osteoporotic Fractures Research Group. *Ann Intern Med* 1991;115(11):837–842. [PubMed: 1952469]
5. Melton LJ 3rd, Kan SH, Frye MA, Wahner HW, O'Fallon WM, Riggs BL. Epidemiology of vertebral fractures in women. *Am J Epidemiol* 1989;129(5):1000–1011. [PubMed: 2784934]
6. Ross PD, Davis JW, Vogel JM, Wasnich RD. A critical review of bone mass and the risk of fractures in osteoporosis. *Calcif Tissue Int* 1990;46(3):149–161. [PubMed: 2137720]
7. Ammann P, Rizzoli R. Bone strength and its determinants. *Osteoporos Int* 2003;14:S13–S18. [PubMed: 12730800]
8. Recker RR, Barger-Lux MJ. The elusive concept of bone quality. *Curr Osteoporos Rep* 2004;2(3):97–100. [PubMed: 16036089]
9. Majumdar S, Kothari M, Augat P, Newitt DC, Link TM, Lin JC, Lang T, Lu Y, Genant HK. High-resolution magnetic resonance imaging: three-dimensional trabecular bone architecture and biomechanical properties. *Bone* 1998;22(5):445–454. [PubMed: 9600777]
10. Link TM, Bauer JS. Imaging of trabecular bone structure. *Semin Musculoskelet Radiol* 2002;6(3):253–262. [PubMed: 12541203]
11. Link TM, Vieth V, Stehling C, Lotter A, Beer A, Newitt D, Majumdar S. High-resolution MRI vs multislice spiral CT: which technique depicts the trabecular bone structure best? *Eur Radiol* 2003;13(4):663–671. [PubMed: 12664101]
12. Wehrli FW, Saha PK, Gomberg BR, Song HK, Snyder PJ, Benito M, Wright A, Weening R. Role of Magnetic Resonance for Assessing Structure and Function of Trabecular Bone. *Top Magn Reson Imaging* 2002;13(5):335–355. [PubMed: 12464746]
13. Gasser JA, Ingold P, Grosios K, Laib A, Hammerle S, Koller B. Noninvasive monitoring of changes in structural cancellous bone parameters with a novel prototype micro-CT. *J Bone Miner Metab* 2005;23:90–96. [PubMed: 15984422]
14. MacNeil JA, Boyd SK. Load distribution and the predictive power of morphological indices in the distal radius and tibia by high resolution peripheral quantitative computed tomography. *Bone* 2007;41(1):129–137. [PubMed: 17442649]
15. Khosla S, Riggs BL, Atkinson EJ, Oberg AL, McDaniel LJ, Holets M, Peterson JM, Melton LJ 3rd. Effects of sex and age on bone microstructure at the ultradistal radius: a population-based noninvasive in vivo assessment. *J Bone Miner Res* 2006;21(1):124–131. [PubMed: 16355281]
16. Caligiuri P, Giger ML, Favus M. Multifractal radiographic analysis of osteoporosis. *Med Phys* 1994;21(4):503–508. [PubMed: 8058015]

17. Caligiuri P, Giger ML, Favus MJ, Jia H, Doi K, Dixon LB. Computerized radiographic analysis of osteoporosis: preliminary evaluation. *Radiology* 1993;186(2):471–474. [PubMed: 8421753]
18. Chinander MR, Giger ML, Martell JM, Jiang C, Favus MJ. Computerized radiographic texture measures for characterizing bone strength: a simulated clinical setup using femoral neck specimens. *Med Phys* 1999;26(11):2295–2300. [PubMed: 10587210]
19. Jiang C, Giger ML, Chinander MR, Martell JM, Kwak S, Favus MJ. Characterization of bone quality using computer-extracted radiographic features. *Med Phys* 1999;26(6):872–879. [PubMed: 10436887]
20. Gregory JS, Stewart A, Undrill PE, Reid DM, Aspden RM. Identification of hip fracture patients from radiographs using Fourier analysis of the trabecular structure: a cross-sectional study. *BMC Med Imaging* 2004;4(1):4. [PubMed: 15469614]
21. Benhamou CL, Poupon S, Lespessailles E, Loiseau S, Jennane R, Siroux V, Ohley W, Pothuau L. Fractal analysis of radiographic trabecular bone texture and bone mineral density: two complementary parameters related to osteoporotic fractures. *J Bone Miner Res* 2001;16(4):697–704. [PubMed: 11315997]
22. Pothuau L, Lespessailles E, Harba R, Jennane R, Royant V, Eynard E, Benhamou CL. Fractal analysis of trabecular bone texture on radiographs: discriminant value in postmenopausal osteoporosis. *Osteoporos Int* 1998;8(6):618–625. [PubMed: 10326070]
23. Chappard C, Brunet-Imbault B, Lemineur G, Giraudeau B, Basillais A, Harba R, Benhamou CL. Anisotropy changes in post-menopausal osteoporosis: characterization by a new index applied to trabecular bone radiographic images. *Osteoporos Int* 2005;16(10):1193–1202. [PubMed: 15685395]
24. Brunet-Imbault B, Lemineur G, Chappard C, Harba R, Benhamou CL. A new anisotropy index on trabecular bone radiographic images using the fast Fourier transform. *BMC Med Imaging* 2005;5:4. [PubMed: 15927072]
25. Lespessailles E, Roux JP, Benhamou CL, Arlot ME, Eynard E, Harba R, Padonou C, Meunier PJ. Fractal analysis of bone texture on os calcis radiographs compared with trabecular microarchitecture analyzed by histomorphometry. *Calcif Tissue Int* 1998;63(2):121–125. [PubMed: 9685516]
26. Lespessailles E, Jullien A, Eynard E, Harba R, Jacquet G, Ildefonse JP, Ohley W, Benhamou CL. Biomechanical properties of human os calcanei: relationships with bone density and fractal evaluation of bone microarchitecture. *J Biomech* 1998;31(9):817–824. [PubMed: 9802782]
27. Wilkie JR, Giger ML, Chinander MR, Vokes TJ, Li H, Dixon L, Jaros V. Comparison of radiographic texture analysis from computed radiography and bone densitometry systems. *Med Phys* 2004;31(4):882–891. [PubMed: 15125006]
28. Wilkie JR, Giger ML, Chinander MR, Vokes TJ, Nishikawa RM, Carlin MD. Investigation of physical image quality indices of a bone densitometry system. *Med Phys* 2004;31(4):873–881. [PubMed: 15125005]
29. Vokes TJ, Giger ML, Chinander MR, Karrison TG, Favus MJ, Dixon LB. Radiographic texture analysis of densitometer-generated calcaneus images differentiates postmenopausal women with and without fractures. *Osteoporos Int* 2006;17(10):1472–1482. [PubMed: 16838099]
30. Chinander MR, Giger ML, Martell JM, Favus MJ. Computerized analysis of radiographic bone patterns: effect of imaging conditions on performance. *Med Phys* 2000;27(1):75–85. [PubMed: 10659740]
31. Rupprecht M, Pogoda P, Mumme M, Rueger JM, Puschel K, Amling M. Bone microarchitecture of the calcaneus and its changes in aging: a histomorphometric analysis of 60 human specimens. *J Orthop Res* 2006;24(4):664–674. [PubMed: 16514636]
32. Chappard D, Guggenbuhl P, Legrand E, Basle MF, Audran M. Texture analysis of X-ray radiographs is correlated with bone histomorphometry. *J Bone Miner Metab* 2005;23(1):24–29. [PubMed: 15616890]
33. Lin JC, Amling M, Newitt DC, Selby K, Srivastav SK, Delling G, Genant HK, Majumdar S. Heterogeneity of trabecular bone structure in the calcaneus using magnetic resonance imaging. *Osteoporos Int* 1998;8(1):16–24. [PubMed: 9692073]
34. Wehrli FW, Hilaire L, Fernandez-Seara M, Gomberg BR, Song HK, Zemel B, Loh L, Snyder PJ. Quantitative magnetic resonance imaging in the calcaneus and femur of women with varying degrees

- of osteopenia and vertebral deformity status. *J Bone Miner Res* 2002;17(12):2265–2273. [PubMed: 12469921]
35. ISCD. Precision calculator. [Accessed September 11, 2007].
<http://www.iscd.org/visitors/resources/calc.cfm>
36. Gluer CC, Blake G, Lu Y, Blunt BA, Jergas M, Genant HK. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. *Osteoporos Int* 1995;5(4):262–270. [PubMed: 7492865]
37. Benhamou CL, Lespessailles E, Jacquet G, Harba R, Jennane R, Loussot T, Tourliere D, Ohley W. Fractal organization of trabecular bone images on calcaneus radiographs. *J Bone Miner Res* 1994;9(12):1909–1918. [PubMed: 7872056]
38. Fidouh F, Harba R, Jacquet G, Loussot T, Lespessailles E. Long-term reproducibility optimization of an x-ray process for bone architectural evaluation during osteoporosis. *Phys Med Biol* 1999;44(1):N1–N8. [PubMed: 10071892]
39. Hakulinen MA, Saarakkala S, Toyras J, Kroger H, Jurvelin JS. Dual energy x-ray laser measurement of calcaneal bone mineral density. *Phys Med Biol* 2003;48(12):1741–1752. [PubMed: 12870580]
40. Pacheco EM, Harrison EJ, Ward KA, Lunt M, Adams JE. Detection of osteoporosis by dual energy X-ray absorptiometry (DXA) of the calcaneus: is the WHO criterion applicable? *Calcif Tissue Int* 2002;70(6):475–482. [PubMed: 11960208]
41. Gomberg BR, Wehrli FW, Vasilic B, Weening RH, Saha PK, Song HK, Wright AC. Reproducibility and error sources of micro-MRI-based trabecular bone structural parameters of the distal radius and tibia. *Bone* 2004;35(1):266–276. [PubMed: 15207767]

Figure 1.

The schematic representation of study design. A. Short-term precision study; B. Long-term precision study.

Figure 2.

Acquisition and processing of the heel image used for RTA: A. Positioning of the heel for acquisition of PIXI image; B. An example the heel PIXI image; C. The usual placement of the 128×128 ROI with 5 smaller, 64×64 sub-ROIs; D. Appearance of the image after image processing which allows the texture pattern to be visualized more clearly.

Table 1

Short-term precision: Coefficients of variation (%) obtained by 3 ROI placements by the same or by 3 different observers on the 8 images obtained on the same day from each of 33 subjects.

ROI 1	Intra-observer*	Inter-observer	8 images*
RMS	1.5 – 3.6	3.5	2.4 – 3.0
sdRMS	10.1 – 19.1	21.8	16.0 – 18.7
FMP	1.1 – 2.8	3.4	2.1 – 2.6
MINK	0.6 – 0.9	0.9	0.6 – 0.9
ROI 2			
RMS	0.9 – 2.1	2.9	1.9 – 2.2
sdRMS	7.5 – 14.5	15.0	14.0 – 15.9
FMP	1.1 – 2.4	3.0	2.2 – 2.7
MINK	0.5 – 0.7	0.8	0.7 – 0.7
ROI 3			
RMS	2.1 – 4.7	4.4	3.5 – 4.8
sdRMS	7.3 – 14.6	12.2	13.3 – 15.1
FMP	1.3 – 2.4	2.9	2.4 – 2.9
MINK	0.6 – 0.7	0.8	0.7 – 0.8
ROI 4			
RMS	2.3 – 3.6	4.4	2.6 – 4.0
sdRMS	6.8 – 13.0	18.8	14.0 – 15.2
FMP	1.4 – 2.7	3.2	2.3 – 3.1
MINK	0.5 – 0.7	0.7	0.6 – 0.7
ROI 5			
RMS	1.0 – 3.6	2.3	2.5 – 4.2
sdRMS	7.0 – 13.4	13.1	14.0 – 16.1
FMP	1.0 – 2.4	2.4	2.1 – 2.5
MINK	0.5 – 0.7	0.6	0.6 – 0.8
BMD	Not applicable	Not applicable	2.8

ROI 1 through 5 designate ROI sub-regions 1–5.

*

The range represents the range of CVs for the 3 observers

Variability due to subject factors and image acquisition and processing steps expressed as percent of total variability in the data collected from all sub-regions (ROI 1 to ROI 5).

% variability due to	RMS	sdRMS	FMP	MINK	BMD
Subject	61.4	65.6	74.3	20.7	98.9
Repositioning	0.6	0.6	0.7	0.9	
2 nd scan same position	0.3	0.6	0.3	0.2	0.6
Observer	0.4	0.1	0.1	0.8	N/A
ROI sub-region (1-5)	16.1	3.9	1.8	29.0	N/A
Residual (error)	21.2	29.3	22.8	48.5	0.5

Table 3

Variability due to the same factors as in table 2 but calculated separately for the 5 ROI sub-regions (ROI 1 through ROI 5).

% variability due to	RTA features			
ROI 1	RMS	sdRMS	FMP	MINK
Subject	76.9	69.4	83.8	34.9
Repositioning	2.1	3.6	1.9	2.4
2 nd scan same position	2.0	0.9	1.8	0.5
Observer	4.2	5.3	1.6	8.9
Residual (error)	14.8	20.8	10.8	53.3
ROI 2	RMS	sdRMS	FMP	MINK
Subject	88.4	79.1	86.6	43.0
Repositioning	0.9	3.0	2.4	4.7
2 nd scan same position	0.8	1.2	0.9	1.9
Observer	2.3	2.1	1.8	4.4
Residual (error)	7.5	14.6	8.3	46.0
ROI 3	RMS	sdRMS	FMP	MINK
Subject	81.5	87.8	88.8	42.7
Repositioning	2.2	1.3	1.9	3.5
2 nd scan same position	2.4	2.5	2.2	6.7
Observer	0.8	0.5	0.1	0.1
Residual (error)	13.1	8.0	7.0	47.0
ROI 4	RMS	sdRMS	FMP	MINK
Subject	82.2	81.7	86.0	62.0
Repositioning	1.8	2.2	2.0	3.2
2 nd scan same position	1.5	3.0	2.4	1.7
Observer	0.4	0.3	0.1	1.7
Residual (error)	14.2	12.8	9.5	31.4
ROI 5	RMS	sdRMS	FMP	MINK
Subject	85.2	83.5	88.4	50.2
Repositioning	2.0	1.9	1.95	2.6
2 nd scan same position	0.9	2.0	1.33	3.8
Observer	0.5	0.4	0.12	1.1
Residual (error)	11.4	12.2	8.2	42.4

Table 4
Variability or RTA and BMD due to subject factors and image acquisition as percent of total variability

% variability due to	RMS	sdRMS	FMP	MINK	BMD
Subject	92.6	90.6	93.0	61.0	98.9
Repositioning	2.3	3.6	3.2	3.0	0.6
Residual (error)	5.1	5.8	3.8	36.0	0.5

The data for RTA is shown for ROI 5 only. The analysis was performed using the first ROI selection by one of the observers on 8 images from each of 33 patients. The residual variability (error) includes the variability due to rescanning without repositioning as well as the true "residual" variability. In the case of RTA features, it also includes the variability due to ROI selection.

Table 5 Long-term precision: Coefficients of variation (%) obtained by ROI placements by the same or by 2 different observers on the 9 images (3 images on each of 3 days separated by one week) from 10 subjects.

	1 st image day 1 Inter-Observer	3 images same day		3 images 1 week apart		All 9 images	
		Observer 1	Observer 2	Observer 1	Observer 2	Observer 1	Observer 2
ROI 1							
RMS	3.3	1.8	1.4	2.4	2.4	2.6	2.2
sdRMS	17.3	9.9	8.9	12.7	14.3	14.1	12.6
FMP	3.0	1.9	1.7	2.4	2.4	2.3	2.4
MINK	0.7	0.5	0.4	0.3	0.5	0.5	0.5
ROI 2							
RMS	1.6	1.6	1.5	2.3	2.5	2.3	2.5
sdRMS	8.4	13.2	11.1	14.2	12.4	14.7	15.4
FMP	2.0	1.9	1.7	2.7	2.5	2.6	2.7
MINK	0.5	0.4	0.6	0.5	0.5	0.5	0.6
ROI 3							
RMS	1.8	1.8	1.5	1.3	1.6	1.7	1.9
sdRMS	12.1	10.5	10.0	9.7	12.3	11.7	12.9
FMP	1.6	2.3	1.7	1.6	1.9	2.0	2.2
MINK	0.5	0.5	0.5	0.4	0.5	0.5	0.5
ROI 4							
RMS	1.5	2.0	1.9	2.1	2.1	2.3	1.9
sdRMS	14.0	9.1	9.3	14.6	13.2	14.7	12.5
FMP	1.3	2.3	1.5	2.3	2.5	2.7	2.1
MINK	0.4	0.5	0.6	0.6	0.5	0.6	0.6
ROI 5							
RMS	1.6	1.7	1.4	1.8	2.1	2.1	2.1
sdRMS	8.7	9.4	11.0	9.8	10.6	12.0	12.4
FMP	1.6	2.1	2.2	2.0	1.7	2.4	2.1
MINK	0.6	0.5	0.4	0.5	0.6	0.6	0.6
BMD	N/A	1.4		2.2			2.0

ROI 1 through 5 designate ROI sub-regions 1–5.