

The Precision and Sensitivity of ^{18}F -Fluoride PET for Measuring Regional Bone Metabolism: A Comparison of Quantification Methods

Musib Siddique¹, Michelle L. Frost¹, Glen M. Blake¹, Amelia E.B. Moore¹, Yosra Al-Beyatti¹, Paul K. Marsden², Paul J. Schleyer², and Ignac Fogelman¹

¹King's College London, Osteoporosis Screening and Research Unit, Guy's Hospital, London, United Kingdom; and ²King's College London, PET Imaging Centre, St. Thomas' Hospital, London, United Kingdom

The planning of research studies requires an understanding of the minimum number of subjects required. The aim of this study was to evaluate different methods of analyzing ^{18}F -fluoride PET (^{18}F -PET) dynamic spine scans to find the approach that requires the smallest sample size to detect a statistically significant response to treatment. **Methods:** Eight different approaches to ^{18}F -PET analysis (3 variants of the Hawkins 3-tissue compartmental model, 3 variants of spectral analysis, deconvolution, and Patlak analysis) were used to evaluate the fluoride plasma clearance to bone mineral (K_i). Standardized uptake values (SUVs) were also studied. Data for 20 women who had ^{18}F -PET spine scans at 0, 6, and 12 mo after stopping long-term bisphosphonate treatment were used to compare precision errors. Data for 18 women who had scans at baseline and 6 mo after starting teriparatide treatment were used to compare response to treatment. **Results:** The 4 approaches that fitted the rate constant k_4 describing the reverse flow of ^{18}F from bone as a free variable showed close agreement in K_i values, with correlation coefficients greater than 0.97. Their %CVs were 14.4%–14.8%, and treatment response to teriparatide was 23.2%–23.8%. The 3 methods that assumed $k_4 = 0$ gave K_i values 20%–25% lower than the other methods, with correlation coefficients of 0.83–0.94, percentage coefficients of variation (%CVs) of 12.9%–13.3%, and treatment response of 25.2%–28.3%. A Hawkins model with $k_4 = 0.01 \text{ min}^{-1}$ did not perform any better (%CV, 14.2%; treatment response, 26.1%). Correlation coefficients between SUV and the different K_i methods varied between 0.60 and 0.65. Although SUV gave the best precision (%CV, 10.1%), the treatment response (3.1%) was not statistically significant. **Conclusion:** Methods that calculated K_i assuming $k_4 = 0$ required fewer subjects to demonstrate a statistically significant response to treatment than methods that fitted k_4 as a free variable. Although SUV gave the smallest precision error, the absence of any significant changes make it unsuitable for examining response to treatment in this study.

Key Words: ^{18}F -fluoride; bone metabolism; positron emission tomography; precision errors; sensitivity

J Nucl Med 2011; 52:1748–1755

DOI: 10.2967/jnumed.111.093195

Quantitative radionuclide imaging of the skeleton using ^{18}F -fluoride PET (^{18}F -PET) (1,2) is a valuable tool for research studies examining the pathophysiology of metabolic bone diseases and the response of patients to treatment (3–7). Treatments for osteoporosis and Paget disease generally have a profound effect on bone remodeling (8–11), and studies of bone metabolism have an important role in the evaluation of the effect of treatment on bone tissue (12). Bone biopsy with double-tetracycline labeling is considered the gold standard for the direct assessment of bone turnover activity but is invasive, costly, and restricted to a single site, the iliac crest (13,14). The most practical method for the assessment of bone turnover is the measurement of biochemical markers in serum and urine (15,16). However, bone turnover markers provide information on the integrated response across the whole skeleton and cannot give insight into the changes occurring at specific sites such as the spine and hip, or differences between cortical and trabecular bone. Radionuclide imaging using ^{18}F -PET provides a unique way of studying regional bone metabolism that reflects bone blood flow and osteoblastic activity and can complement these other methods (17).

Quantitative ^{18}F -PET is often performed using the dynamic scan method described by Hawkins (1). The bone time–activity curve and the arterial input function are analyzed to find the fluoride plasma clearance to bone mineral (K_i) ($\text{mL} \cdot \text{min}^{-1} \cdot \text{mL}^{-1}$) (1,3–5). A simpler method of quantifying PET studies that avoids having to find the input function is to measure standardized uptake values (SUVs) by normalizing the mean ^{18}F concentration in the bone region of interest for injected activity and body weight (mean SUV = mean kBq/mL \times body weight [kg]/injected activity [MBq]) (6).

Received May 13, 2011; revision accepted Jul. 14, 2011.

For correspondence or reprints contact: Musib Siddique, Osteoporosis Research Unit, 1st Floor, Tower Wing, Guy's Hospital, London, SE1 9RT, United Kingdom.

E-mail: muhammed.siddique@kcl.ac.uk

Published online Oct. 11, 2011.

COPYRIGHT © 2011 by the Society of Nuclear Medicine, Inc.

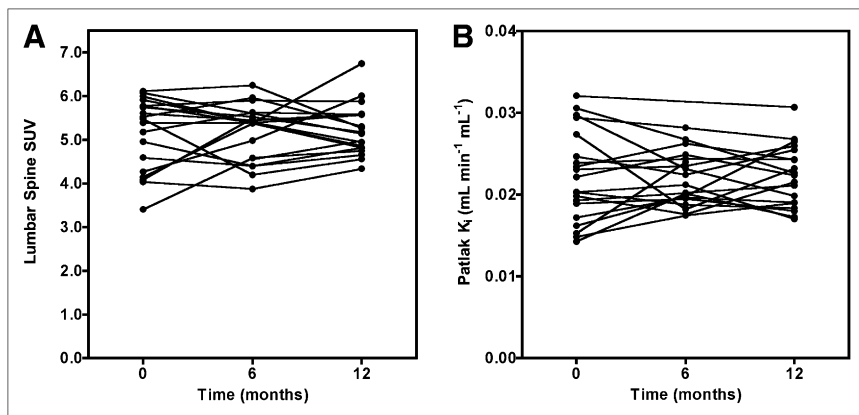


FIGURE 1. Plots of SUVs (A) and fluoride plasma clearance to bone mineral (Patlak K_i) (B) at baseline, 6 mo, and 12 mo in 20 women used to evaluate precision errors of $^{18}\text{F}^-$ lumbar spine PET tracer kinetics.

When planning research studies using $^{18}\text{F}^-$ PET, it is important to estimate reliably the number of subjects required for a statistically significant result. Among other factors, this depends on the precision error of the technique and, for longitudinal studies that quantify the effects of pharmacologic treatments, the change in the measurement variable during the study. If the study design involves subjects who serve as their own controls with a baseline scan and a single follow-up scan at the end of the treatment period, the number of subjects N required to achieve a specified level of statistical significance is given by the following equation (18):

$$N = (Z_{\alpha/2} + Z_{\beta})^2 \times \frac{(\sigma_b^2 + 2\sigma_p^2)}{\Delta_{treat}^2} \quad \text{Eq. 1}$$

In Equation 1, $Z_{\alpha/2}$ and Z_{β} are z scores corresponding to type 1 and type 2 errors, respectively, Δ_{treat} is the treatment response expressing the average change in the measurements between baseline and the end of treatment, σ_p is the random scan-to-scan precision error in the PET variable, and σ_b is the inherent biologic variability in treatment response between subjects.

Several different approaches have been used for the quantitative analysis of $^{18}\text{F}^-$ PET scans, including the measurement of SUV and the derivation of K_i using either Patlak graphical analysis or nonlinear regression

analysis based on the Hawkins 3-tissue compartmental model (19). Alternative approaches to scan analysis include deconvolution and spectral analysis (20,21). Although these latter techniques are widely used in functional imaging, we are not aware of whether they have previously been applied to $^{18}\text{F}^-$ studies of bone metabolism.

The purpose of the present study was to evaluate several different approaches to the quantitative analysis of $^{18}\text{F}^-$ PET. The different approaches were compared in terms of the consistency of the numeric results obtained between the methods and the differences in precision error and treatment response. The aim was to determine the optimum approach to scan analysis in terms of the minimum numbers of subjects required for a statistically significant result. Nine different methods were compared: 3 variants of the Hawkins model, 3 variants of spectral analysis, deconvolution, Patlak analysis, and SUV.

MATERIALS AND METHODS

Subjects

The precision of lumbar spine dynamic $^{18}\text{F}^-$ PET was studied by analyzing data for 20 postmenopausal women with osteoporosis who had been treated with a bisphosphonate for an average of 5 y (range, 3–8 y) at the time of their baseline scan. The women stopped their bisphosphonate treatment the day after their first PET scan and had follow-up examinations 6 and 12 mo later. Analysis of the data showed no evidence for any statistically sig-

TABLE 1
Demographics and Baseline Characteristics of Precision and Treatment Response Study Groups

Variable	Precision group ($n = 20$)	Treatment response group ($n = 20$)
Age (y)	66.2 \pm 6.8	65.3 \pm 8.2
Years postmenopausal	16.5 \pm 7.9	17.7 \pm 9.1
Body mass index (kg/m ²)	24.0 \pm 5.0	24.6 \pm 3.13
Lumbar spine T-score	−2.0 \pm 0.7	−3.0 \pm 0.7
Femoral neck T-score	−1.8 \pm 0.6	−2.0 \pm 0.6
Total hip T-score	−1.6 \pm 0.8	−1.7 \pm 0.6

Data are mean \pm SD.

nificant change in SUV or K_i during the 12-mo period after the bisphosphonate had been stopped (Fig. 1). The mean change in SUV was 1.7% (SD, 13.9%; range, -25% to +27%) ($P = 0.61$) at 6 mo and 0.0% (SD, 16.5%; range, -19% to +34%) ($P = 0.99$) at 12 mo. For Patlak K_i , the changes were 1.3% (SD, 18.7%; range, -41% to +39%) ($P = 0.77$) at 6 mo and -1.0% (SD, 20.1%; range, -45% to +47%) ($P = 0.83$) at 12 mo. Because of the lack of any significant changes, the data were deemed suitable for an analysis of scan precision.

To study treatment response, we analyzed PET data for 18 postmenopausal women with osteoporosis who participated in a clinical trial of the bone anabolic agent teriparatide (7). Dynamic ^{18}F -PET scans of the lumbar spine were performed at baseline and after 6 mo of treatment with teriparatide, 20 $\mu\text{g}/\text{d}$.

Informed written consent was obtained from all participants, and the local ethics committee and the Administration of Radioactive Substances Advisory Committee approved the studies. Baseline characteristics of the women who participated in the precision and treatment studies are compared in Table 1.

PET Image Acquisition

The protocol for PET image acquisition and blood sampling to estimate the input function was the same for both studies (7). Briefly, 60-min dynamic scans of the lumbar spine (time frames, 24×5 s, 4×30 s, and 14×240 s) were acquired on a Discovery PET/CT scanner (GE Healthcare) with a 15.4-cm axial field of view. Image acquisition was commenced simultaneously with the bolus injection of 90 MBq of ^{18}F -sodium fluoride, resulting in 47×3.27 -mm slices for each frame, with a pixel size of 1.66 mm in the transaxial plane. All activity measurements were corrected for radioactive decay back to the time of injection. Regions of interest were defined by summing the frames from 12 to 60 min and using a sagittal projection image to identify a set of transaxial slices through the middle of each vertebral body L_1 to L_4 , avoiding the end plates and the disk space. These axial slices were summed to produce a transaxial image of each vertebra in which an elliptic region of interest was placed within the vertebral body. The final lumbar spine time-activity curve for each subject was produced by averaging results for the 4 individual vertebral bodies. Measurements of lumbar spine K_i were obtained by estimation of the arterial plasma input function using a semipopulation method (7). Venous blood samples were taken at 30, 40, 50, and 60 min after injection, and a single exponential was fitted to define the terminal exponential for the 0- to 60-min dynamic scan. To reconstruct the

entire arterial input function, a population residual curve representing the bolus peak and sum of the early fast exponentials obtained by direct arterial sampling from 10 postmenopausal women studied by Cook et al. (22) was scaled for injected activity and added to each individual's terminal exponential curve.

^{18}F -PET Scan Analysis

The data from the PET dynamic spine scan were analyzed using 8 different methods to estimate K_i in the lumbar vertebral bodies, together with a measurement of SUV. For any one dynamic scan, the same input function was used for each of the K_i calculations. The 9 different methods are described below and are summarized in Table 2.

Method 1: Hawkins Model with Freely Fitted k_4 . K_i was found using the Hawkins model as previously described (1,7) and calculated from the following equation:

$$K_i = K_1 \times k_3 / (k_2 + k_3) \text{ mL} \cdot \text{min}^{-1} \cdot \text{mL}^{-1}. \quad \text{Eq. 2}$$

Method 2: Hawkins Model with $k_4 = 0$. Method 2 is a variant of the Hawkins model with the rate constant k_4 set to zero on the assumption that no release of tracer from bone mineral occurs during the 1-h scan (23). It has the theoretic advantage that eliminating the random scan-to-scan errors in k_4 should improve the precision of the K_i measurements.

Method 3: Hawkins Model with $k_4 = 0.01 \text{ Min}^{-1}$. Method 3 is similar to method 2 but with k_4 set to its population average value. As well as possibly improving precision, it has the additional advantage of giving a better fit to the bone time-activity curves.

Methods Based on Spectral Analysis. Spectral analysis is a technique that assumes that the impulse response function (IRF) for tissue tracer kinetics measured by PET can be expressed as a sum of exponentials (20). The technique produces a spectrum of the kinetic components that relates the tissue response to the plasma activity curve with minimal modeling assumptions. From this summary of the kinetic components, the IRF can be derived as the weighted sum of exponentials:

$$\text{IRF}_s = \sum_{i=1}^n \alpha_i e^{-\beta_i t}, \quad \text{Eq. 3}$$

where α_i is the weighting factor and β_i is the frequency of the i^{th} exponential, n is the number of exponentials used, and t is time after injection. We used an iterative approach using a least-squares curve fit method to minimize the following objective function:

TABLE 2
List of Approaches to PET Scan Quantification Used in Current Study

Method	Technique	Abbreviation	Value of k_4	Group
Method 1	Hawkins model	Hawkins 1	Free variable	Group 1
Method 2	Hawkins model	Hawkins 2	0	Group 2
Method 3	Hawkins model	Hawkins 3	0.01 min^{-1}	—
Method 4	Spectral analysis ($n = 500$)	Spectral 1	Free variable	Group 1
Method 5	Spectral analysis ($n = 2$)	Spectral 2	Free variable	Group 1
Method 6	Spectral analysis ($n = 2, \beta_1 = 0$)	Spectral 3	0	Group 2
Method 7	Deconvolution	Deconvolution	Free variable	Group 1
Method 8	Patlak analysis	Patlak	0	Group 2
Method 9	Mean SUV	SUV	Not applicable	—

n = number of discrete frequencies used in spectral analysis; β_1 = value of first frequency.

$$\min_{\alpha_i, \beta_i} \left[C_t - C_p \times \sum_{i=1}^n \alpha_i e^{-\beta_i t} \right]^2. \quad \text{Eq. 4}$$

The value of K_i was found from the intercept of the terminal exponential fitted to the 10- to 60-min points of the IRF.

Method 4: Spectral Variant 1. Method 4 makes no a priori assumptions regarding the number of components required to describe the IRF. For this study, we chose 500 discrete values of β_i by equally sampling $\log(\beta_i)$ in a predefined range and performed spectral analysis to obtain the optimal values of α_i . After convergence, only a handful of β_i values were nonzero, demonstrating that only a few frequencies were present in the spectrum.

$$IRF_s = \sum_{i=1}^{500} \alpha_i e^{-\beta_i t}, \text{ where } \beta_i \text{ s are fixed.} \quad \text{Eq. 5}$$

Method 5: Spectral Variant 2. In method 5, we take $n = 2$ to mimic the Hawkins model and perform spectral analysis to obtain the optimal pair of values of α and β . In this respect, the method is similar to the Hawkins model with k_4 freely fitted (method 1). The objective function takes the form:

$$\min_{\alpha_i, \beta_i} \left[C_t - C_p \times \sum_{i=1}^2 \alpha_i e^{-\beta_i t} \right]^2. \quad \text{Eq. 6}$$

Method 6: Spectral Variant 3. In method 6, we assume $n = 2$ as above but fix 1 value of β to be zero and perform spectral analysis to obtain the optimal values of the other 3 parameters. This method is similar to method 2 in assuming that the tracer is irreversibly bound to the bone mineral compartment. The objective function takes the form:

$$\min_{\alpha_1, \alpha_2, \beta_2} \left[C_t - C_p \times (\alpha_1 + \alpha_2 e^{-\beta_2 t}) \right]^2. \quad \text{Eq. 7}$$

Method 7: Deconvolution. In recent years, the deconvolution method has been applied in many areas, including signal processing, geophysics, and communications (21). In the present study, the convolution of the plasma input function with the IRF gives the best-fit curve to the observed tissue data. The analysis makes no a priori assumptions about the IRF and hence is model-independent and nonparametric. All the elements of the IRF vector are variables and need to be determined:

$$C_t(t) = IRF(t) \times C_p(t). \quad \text{Eq. 8}$$

We used a spectral guided deconvolution algorithm to estimate the IRF. The deconvolution is performed iteratively, whereby each step improves the estimation of the IRF. A good estimate of the IRF is helpful for quicker convergence and is obtained by running a few iterations in combination with the spectral algorithm. We use a least-squares curve fit method to minimize the following objective function:

$$\min_{IRF_d} \left[C_t - IRF_d \times C_p \right]^2. \quad \text{Eq. 9}$$

Method 8: Patlak Analysis. Patlak analysis (1,23) is a graphical technique for estimating K_i that, like method 2, assumes that $^{18}\text{F}^-$ is irreversibly bound to bone mineral. To allow for equilibration between tracer in plasma and the bone extracellular fluid compartment, the dynamic scan data were fitted from 10 min to the end of

data acquisition at 60 min. The advantage of the method is that the derivation of K_i does not involve any sophisticated computer programming.

Method 9: Mean SUV. Measurement of the mean SUV in the tissue region of interest provides a particularly simple method of evaluating $^{18}\text{F}^-$ skeletal kinetics because no information about the arterial input function is required (6,23). SUVs represent tissue activity within a region of interest corrected for injected activity and body weight. For the present study, we derived SUV by averaging the last 2 frames of the dynamic study (52–60 min).

These 9 methods are summarized in Table 2, where they are categorized into those methods that allow for the rate constant k_4 describing the reverse flow of tracer out of bone mineral to be freely fitted to the data, those methods that assume a fixed value of k_4 , and SUV.

Statistical Analysis

The baseline characteristics of the 2 study populations were expressed by their mean and SD. Scatterplots were drawn to show the correlation between the different scan analysis methods. For each scatterplot, the Pearson correlation coefficient was calculated to assess the strength of the correlation. When appropriate, the fits of the different methods to the bone time–activity curves were compared using the Akaike information criterion (24). When 2 methods were compared, the binomial distribution was used to evaluate whether the number of times one method gave a better fit than another was statistically significantly different from 50%. For the precision study, the precision error was described as the percentage coefficient of variation (%CV), defined as the root mean square SD of the measurements expressed as a percentage of their overall mean along with the 95% confidence interval estimated using the χ^2 distribution (25). The statistical significance of the differences in precision error between the methods was assessed using the F test. For the treatment response study, the changes in the PET parameters were used to calculate the percentage change from baseline, expressed as the mean and SD and evaluated using the paired Student *t* test. For all statistical tests, a 2-tailed *P* value of 0.05 or less was considered statistically significant.

RESULTS

Figure 2 shows the scatterplots for $^{18}\text{F}^-$ lumbar spine PET scan quantification by methods 2–9 against the results from method 1. The plots show the combined data from all scans performed on both groups ($n = 94$). These comprise 36 scans for the 18 subjects in the teriparatide study and 58 scans for the 20 subjects in the precision study. Two subjects in the latter were each missing 1 follow-up scan.

The 4 methods that fitted k_4 as a free variable (methods 1, 4, 5, and 7, hereafter referred to as group 1) showed close agreement in K_i values, with correlation coefficients greater than 0.97 (Figs. 2A–2C). When assessed by binomial testing for the lowest Akaike information criterion values, these methods gave a statistically significantly better fit to the bone time–activity curves than the methods that assume $k_4 = 0$, with method 1 giving the best fit. The 3 methods with $k_4 = 0$ (methods 2, 6, and 8, hereafter

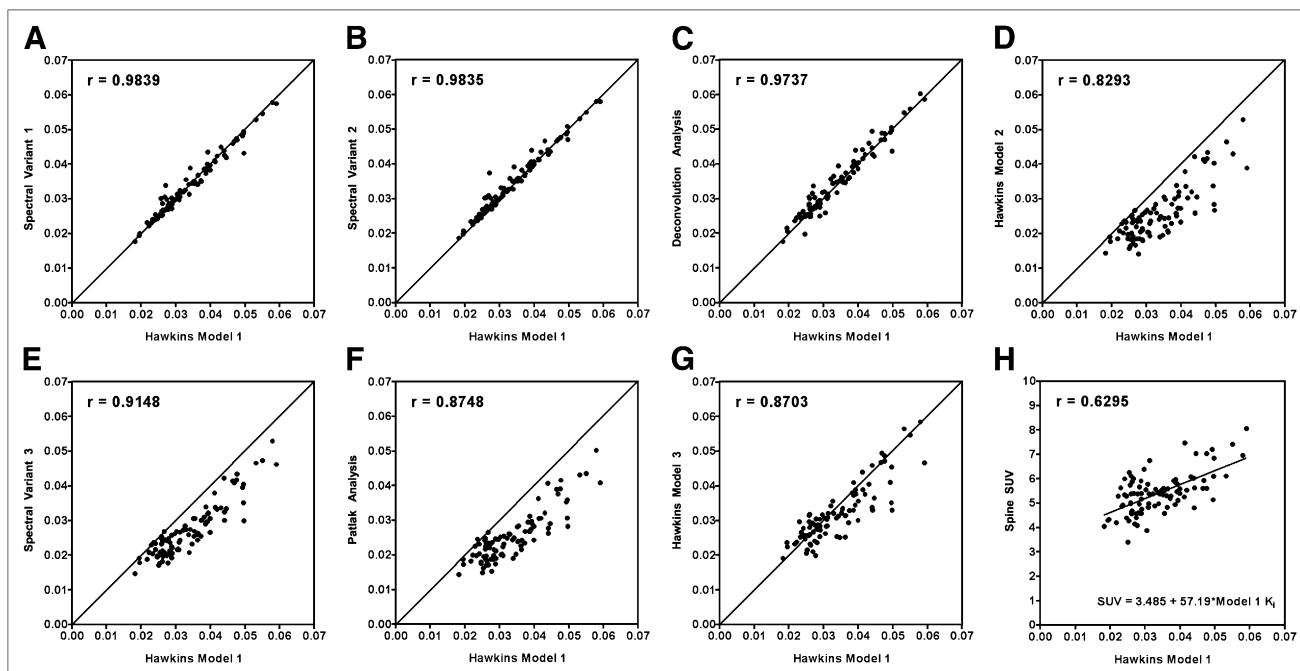


FIGURE 2. Scatterplots for ^{18}F -lumbar spine PET scan quantification by methods 2–9 listed in Table 2 against results from method 1.

referred to as group 2) gave K_i values that were 20%–25% lower than the methods in group 1 and had correlation coefficients with method 1 that varied between 0.83 and 0.91 (Figs. 2D–2F). When the group 2 methods were plotted against each other, the correlation coefficients between them were greater than 0.97. Method 3 with k_4 fixed at 0.01 min^{-1} had correlation coefficients of 0.87–0.90 with the group 1 methods and 0.97–0.98 with the group 2 methods (Fig. 2G). Finally, correlation coefficients between SUV and the 8 methods of evaluating K_i varied between 0.60 and 0.65 (Fig. 2H).

The results for the precision errors of the 9 approaches to ^{18}F -PET scan analysis are plotted in Figure 3A with their 95% confidence intervals. SUV had the smallest %CV (10.1%), and the deconvolution method had the largest (14.8%). The %CV results for K_i were larger for the methods in group 1 (range, 14.4%–14.8%) than for the methods in group 2 (range, 12.9%–13.3%). When k_4 took a fixed value of 0.01 min^{-1} , the precision was 14.2%. When assessed by the F test, the precision error for SUV was significantly smaller than for any of the methods in group 1 ($P = 0.011$ to 0.019). When SUV was compared with the 3 methods in group 2, the differences in the precision errors were not statistically significant ($P = 0.051$ – 0.075). None of the other differences were significant.

Table 3 shows the mean values of K_i estimated by 8 different methods at baseline and after 6 mo of treatment with teriparatide along with the SUV results. When treatment response was expressed as the percentage change in K_i , for the methods in group 1 the increase varied from 23.2% to 23.8%, and for the methods in group 2 the increase varied from 25.2% to 28.3%. For all 8 K_i methods,

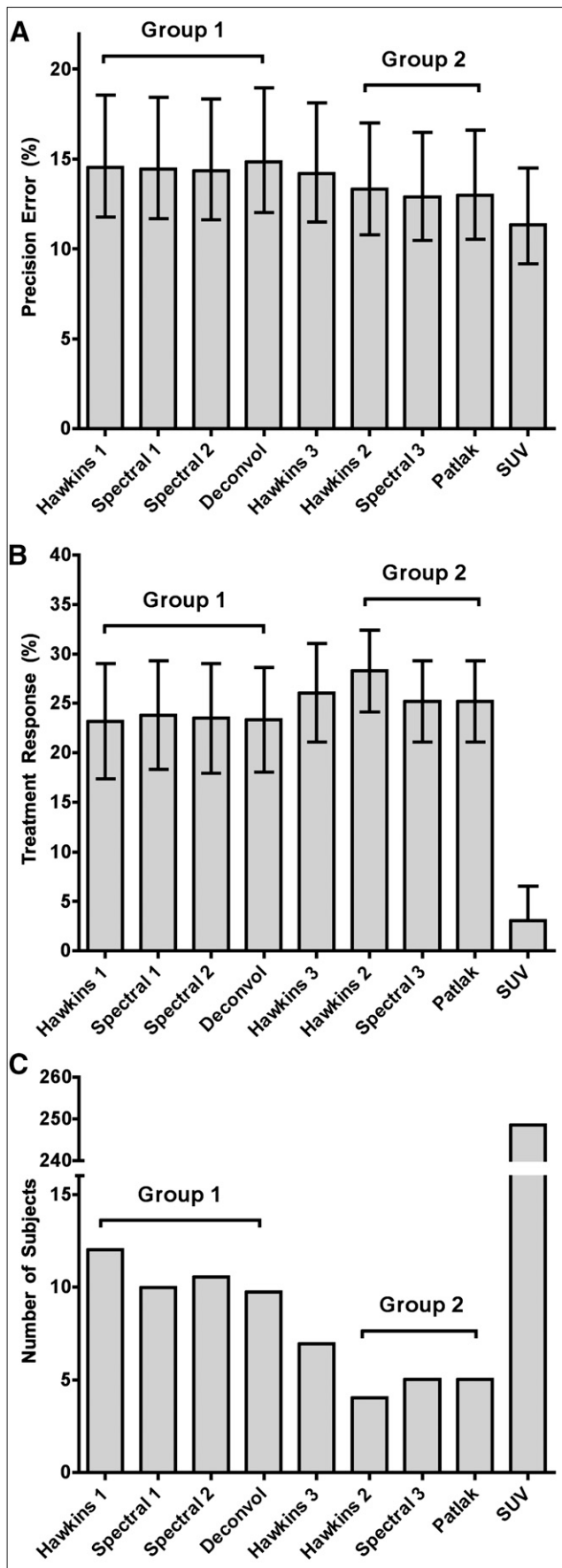
the response to treatment was highly statistically significant ($P < 0.003$). For SUV, the change was 3.1% and was not significant ($P = 0.71$). The results for treatment response are plotted in Figure 3B with their 95% confidence intervals. The treatment response for SUV was statistically significantly smaller than for the 8 K_i methods ($P = 3.8 \times 10^{-4}$ – 2.5×10^{-5}). None of the differences between the K_i methods were statistically significant.

The data on treatment response and SD in Table 3 were substituted in Equation 1 to estimate the number of subjects required for a type 1 error of $\alpha = 0.05$ and a statistical power of 90% (Fig. 3C). Methods in group 2 required half the number of subjects compared with those in group 1.

DISCUSSION

The different methods of evaluating ^{18}F -bone tracer kinetics are conveniently divided into, first, methods for estimating K_i that allow the reverse flow of ^{18}F from bone mineral to be fitted to the data as a free variable (group 1); second, methods for estimating K_i that assume that tracer is irreversibly bound to bone (group 2); and third, SUV. Methods in group 1 gave almost identical numeric values of K_i , and the results were highly correlated. Methods in group 2 also gave results that were highly correlated with each other. Given these high correlations ($r = 0.974$ – 0.997), it is not surprising that the methods within each group also gave similar results for precision, treatment response, and the minimum number of subjects required for a study to show a statistically significant response to treatment (Fig. 3).

When the K_i measurements in group 2 were compared with those in group 1, they gave results that were lower by



20%–25% on average. The lower values reflect the fact that when K_i is estimated, these methods do not take into account the tracer taken up into bone mineral that is released back into plasma before the end of the 60-min scan. Correlation coefficients between the group 1 and group 2 methods were smaller, at 0.83–0.94, and this reflected the range of k_4 values (0–0.024 min⁻¹) found in individual scans.

When the quality of the fits to the bone time–activity curves were assessed using the Akaike information criterion (24), the analysis showed that the methods in group 1 performed better than the methods in group 2. When compared with method 2, method 1 gave a better curve fit in 61 of 94 scans (65%), compared with 50% for the null hypothesis that the 2 methods perform equally well ($P = 0.005$). Because methods with a nonzero value of k_4 gave a better fit to the bone time–activity curves, we also evaluated the Hawkins model with a fixed nonzero value of k_4 (method 3). When assessed by the Akaike information criterion, method 3 lay between methods 1 and 2 in the quality of the curve fits but was not statistically significantly different from either of them (method 3 vs. method 1: $P = 0.256$; method 3 vs. method 2: $P = 0.353$).

When the precision errors of the ¹⁸F– PET scan results were assessed, SUV measurements had the best performance (%CV, 10.2%), the methods in group 2 had the second-best performance (%CV, 12.9%–13.3%), and the methods in group 1 had the largest error (%CV, 14.4%–14.8%). SUV is expected to have the smallest precision error because the measurement involves evaluating only bone uptake and avoids the additional sources of error from measuring the arterial input function. The methods in group 1 are expected to have the worst precision because of the additional errors entailed in measuring k_4 . In the context of the 60-min dynamic scan, the numeric value of k_4 is relatively small ($1/k_4$, ~100 min), making a reliable measurement difficult. As a consequence, the precision error for k_4 was relatively poor (%CV, 32%), and the resulting scan-to-scan variations in k_4 explain the poorer precision of the group 1 K_i measurements. Interestingly, when precision errors for measurements in single lumbar vertebrae were compared with those for L₁–L₄, they were almost unchanged despite the 4 times smaller volume of bone. This finding suggests that the size of the precision error is set by the calibration of the PET scanner rather than by counting statistics (26).

In a previous study of the precision of lumbar spine ¹⁸F– PET measurements, Frost et al. reported %CV values of 14.4% for SUV, 13.8% for Patlak analysis, 12.2% for the Hawkins model with $k_4 = 0$, and 26.6% for the Hawkins

FIGURE 3. (A) Results for precision errors of 8 estimates of fluoride plasma clearance to bone mineral (K_i) and SUV expressed as % CV with 95% confidence intervals. (B) Results for treatment response to teriparatide at lumbar spine for methods in A. Error bars are SE. (C) Minimum number of subjects required for verifying statistically significant treatment response to teriparatide at lumbar spine, as calculated from data in Table 3. Group 1 = scan analysis methods that fit k_4 as free variable; group 2 = scan analysis methods that assume $k_4 = 0$.

TABLE 3
Mean $^{18}\text{F}^-$ Kinetic Parameters at Baseline and After 6 Months of Treatment with Teriparatide

Method	Baseline mean	6-mo mean	Mean change*	P^\dagger
Hawkins 1	0.035 (0.011)	0.041 (0.011)	+23.2 (24.8)	0.00261
Hawkins 2	0.028 (0.006)	0.035 (0.009)	+28.3 (17.5)	1.59E-6
Hawkins 3	0.033 (0.008)	0.041 (0.009)	+26.1 (21.2)	3.65E-5
Spectral 1	0.034 (0.010)	0.041 (0.010)	+23.8 (23.2)	0.00057
Spectral 2	0.034 (0.011)	0.041 (0.010)	+23.5 (23.5)	0.00099
Spectral 3	0.029 (0.008)	0.036 (0.009)	+25.2 (17.5)	6.25E-6
Deconvolution	0.035 (0.010)	0.042 (0.010)	+23.3 (22.5)	0.00048
Patlak	0.027 (0.006)	0.033 (0.008)	+25.2 (17.4)	8.48E-6
SUV	5.75 (1.02)	5.83 (0.84)	+3.1 (14.9)	0.71

*Percentage change from baseline for individual subjects.

†Calculated using paired t test.

Data in parentheses are SD.

model with k_4 fitted as a free variable (23). With the exception of the last result, these results are in good agreement with the present study. The study by Frost et al. was based on data for 16 women who had two $^{18}\text{F}^-$ PET scans 6 mo apart. The study therefore had 16 degrees of freedom for finding the precision error (25). The present precision study was based on 58 scans performed on 20 women over 12 mo, giving 38 degrees of freedom, and so is substantially larger. In general, studies with at least 30 degrees of freedom are recommended for the reliable measurement of precision errors (25). The poorer precision Frost et al. found for method 1 may be explained by the larger precision error of 75% reported for k_4 .

For all the methods evaluating K_i , the treatment response was highly statistically significant. To compare the overall performance of the different approaches with PET scan analysis, the smallest number of subjects required to verify a statistically significant treatment response was estimated using Equation 1 (Fig. 3C). On this measure, the methods in group 2 performed better because they combined a smaller precision error with a larger treatment response.

In contrast to the large treatment response of the K_i measurements shown in Figure 3B, the spine SUV at the end of the 60-min dynamic scan increased by only 3%. At first sight, the finding of a large and statistically significant change in K_i in association with almost no change in SUV is paradoxical. The difference was explained by the effect of teriparatide treatment on the area under the ^{18}F plasma concentration curve, which decreased by 20% when subjects were treated (27). Intuitively, a 23% increase in K_i (Table 3, method 1) would be expected to lead to a 23% increase in SUV, and a 20% decrease in area under the curve would lead to a 20% decrease in SUV. The small change in SUV of 3% is explained by the near cancellation of these 2 effects (27).

This study had some important limitations. The precision study was performed on $^{18}\text{F}^-$ PET scan data originally obtained in a 12-mo trial of the effects of patients stopping bisphosphonate treatment for osteoporosis. Therefore, the subjects' bone turnover was not strictly in a steady state.

However, given that there was no evidence for any significant changes in SUV or K_i in the lumbar spine, and that the average changes seen were much smaller than the precision errors previously reported (23), the data were judged suitable for an analysis of scan precision. Ideally, precision and treatment response should be studied in the same group of subjects. However, both groups studied here were postmenopausal women with osteoporosis whose demographics were well matched (Table 1). Another important limitation of the present study was that the information about treatment response was restricted to the effect of teriparatide on $^{18}\text{F}^-$ PET measurements at the lumbar spine. The unexpectedly small treatment response of the SUV measurements was a consequence of this choice because of the fortuitous cancellation of the change in K_i by the change in the 0- to 60-min area under the curve. Studies of other osteoporosis treatments and studies of teriparatide at other skeletal sites might lead to different conclusions regarding the role of SUV measurements. Nevertheless, the present study emphasizes that some care is necessary in the choice of measurement variables in future studies, since measurements of SUV and K_i may not lead to the same conclusions (27).

CONCLUSION

This study showed that lumbar spine $^{18}\text{F}^-$ PET trials analyzed using methods that assume no reverse flow of tracer from the bone mineral compartment required the fewest subjects. Because Patlak analysis is computationally simpler than nonlinear regression or spectral analysis, it is suggested that this is the best approach to measuring K_i . This recommendation is in agreement with Brenner et al. (19). The methods that fitted k_4 as a free variable were found to give more accurate fits to the bone time-activity curves, although their reliability was affected by the relatively poor precision of the k_4 figures.

DISCLOSURE STATEMENT

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely

to indicate this fact, this article is hereby marked “adversetisement” in accordance with 18 USC section 1734.

ACKNOWLEDGMENT

No potential conflict of interest relevant to this article was reported.

REFERENCES

- Hawkins RA, Choi Y, Huang S-C, et al. Evaluation of the skeletal kinetics of fluorine-18-fluoride ion with PET. *J Nucl Med*. 1992;33:633–642.
- Schiepers C, Nuyts J, Bormans G, et al. Fluoride kinetics of the axial skeleton measured in vivo with fluorine-18-fluoride PET. *J Nucl Med*. 1997;38:1970–1976.
- Cook GJR, Blake GM, Marsden PK, et al. Quantification of skeletal kinetic indices in Paget’s disease using dynamic ^{18}F -fluoride positron emission tomography. *J Bone Miner Res*. 2002;17:854–859.
- Frost ML, Cook GJR, Blake GM, et al. A prospective study of risedronate on regional bone metabolism and blood flow at the lumbar spine measured by ^{18}F -fluoride positron emission tomography. *J Bone Miner Res*. 2003;18:2215–2222.
- Installé J, Nzeusseu A, Bol A, et al. ^{18}F -fluoride PET for monitoring therapeutic response in Paget’s disease of bone. *J Nucl Med*. 2005;46:1650–1658.
- Uchida K, Nakajima H, Miyazaki T, et al. Effects of alendronate on bone metabolism in glucocorticoid-induced osteoporosis measured by ^{18}F -fluoride PET: a prospective study. *J Nucl Med*. 2009;50:1808–1814.
- Frost ML, Siddique M, Blake GM, et al. Differential effects of teriparatide on regional bone formation using ^{18}F -fluoride positron emission tomography. *J Bone Miner Res*. 2011;26:1002–1011.
- Neer RM, Arnaud CD, Zanchetta JR, et al. Effect of recombinant human parathyroid hormone (1-34) fragment on spine and non-spine fractures and bone mineral density in postmenopausal osteoporosis. *N Engl J Med*. 2001;344:1434–1441.
- Chesnut CH, Skag A, Christiansen C, et al. Effects of oral ibandronate administered daily or intermittently on fracture risk in postmenopausal osteoporosis. *J Bone Miner Res*. 2004;19:1241–1249.
- Black DM, Delmas PD, Eastell R, et al. Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. *N Engl J Med*. 2007;356:1809–1822.
- Cummings SR, San Martin J, McClung MR, et al. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. *N Engl J Med*. 2009;361:756–765.
- McClung MR, San Martin J, Miller PD, et al. Opposite bone remodeling effects of teriparatide and alendronate in increasing bone mass. *Arch Intern Med*. 2005;165:1762–1768.
- Arlot M, Meunier PJ, Boivin G, et al. Differential effects of teriparatide and alendronate on bone remodeling in postmenopausal women assessed by histomorphometric parameters. *J Bone Miner Res*. 2005;20:1244–1253.
- Lindsay R, Zhou H, Cosman F, et al. Effects of a one-month treatment with PTH (1-34) on bone formation on cancellous, endocortical, and periosteal surfaces of the human ilium. *J Bone Miner Res*. 2007;22:495–502.
- Garnero P, Weichung JS, Gineyts E, et al. Comparison of new biochemical markers of bone turnover in late postmenopausal osteoporotic women in response to alendronate treatment. *J Clin Endocrinol Metab*. 1994;79:1693–1700.
- Glover SJ, Eastell R, McCloskey EV, et al. Rapid and robust response of biochemical markers of bone formation to teriparatide therapy. *Bone*. 2009;45:1053–1058.
- Blake GM, Park-Holohan S, Cook GJR, Fogelman I. Quantitative studies of bone with the use of ^{18}F -fluoride and $^{99\text{m}}\text{Tc}$ -methylene diphosphonate. *Semin Nucl Med*. 2001;31:28–49.
- Sowers MF, Karvonen-Gutierrez CA. Epidemiological methods in studies of osteoporosis. In: Marcus R, Feldman D, Nelson DA, Rosen CJ, eds. *Osteoporosis*. 3rd ed. Burlington, MA: Elsevier Academic Press; 2007:645–665.
- Brenner W, Vernon C, Muzi M, et al. Comparison of different quantitative approaches to ^{18}F -fluoride PET scans. *J Nucl Med*. 2004;45:1493–1500.
- Cunningham VJ, Jones T. Spectral analysis of dynamic PET studies. *J Cereb Blood Flow Metab*. 1993;13:15–23.
- Lau CY, Lun DP-K, Feng D. Non-invasive quantification of physiological processes with dynamic PET using blind deconvolution. *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, May 12–15, 1998, Seattle, Washington*. Vol 3. Piscataway, NJ: IEEE;1998:1805–1808.
- Cook GJ, Lodge MA, Marsden PK, et al. Non-invasive assessment of skeletal kinetics using fluorine-18 fluoride positron emission tomography: evaluation of image and population-derived arterial input functions. *Eur J Nucl Med*. 1999;26:1424–1429.
- Frost ML, Blake GM, Park-Holohan SJ, et al. Long-term precision of ^{18}F -fluoride PET skeletal kinetic studies in the assessment of bone metabolism. *J Nucl Med*. 2008;49:700–707.
- Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974;19:716–723.
- Glüer C-C, 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;5:262–270.
- Lockhart CM, MacDonald LR, Alessio AM, McDougald WA, Doot RK, Kinahan PE. Quantifying and reducing the effect of calibration error on variability of PET/CT standardized uptake value measurements. *J Nucl Med*. 2011;52:218–224.
- Blake GM, Siddique M, Frost ML, Moore AE, Fogelman I. Radionuclide studies of bone metabolism: do bone uptake and bone plasma clearance provide equivalent measurements of bone turnover? *Bone*. 2011;49:537–542.