Simplified Quantification and Reproducibility Studies of Dopamine D2-Receptor Binding with Iodine-123-IBF SPECT in Healthy Subjects

Masanori Ichise, James R. Ballinger, Douglass Vines, Scott Tsai and Hank F. Kung

Department of Nuclear Medicine, Mount Sinai Hospital; Princess Margaret Hospital, University of Toronto, Toronto, Ontario, Canada; and Division of Nuclear Medicine, Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania

The purpose of this study was to assess the feasibility of a simplified SPECT scan protocol to quantify D2-receptor binding using [^{123}I]iodobenzofuran (IBF) and to evaluate reproducibility of quantitative IBF-SPECT imaging without blood data. **Methods:** Twenty healthy volunteers participated in the study, six had test/retest studies separated by 1 wk. Scans were acquired every 5 min for 180 min using a triple-headed SPECT camera after a bolus injection of IBF (292 MBq). The receptor parameter was determined by using our previously proposed variation of graphical analysis that derives the distribution volume ratio (Rv = V_3/V_2) from multiple scan data without blood data. Rv' was determined from three 20-min scan data obtained at 0–20, 50–70 and 160–180 min postinjection and compared with Rv as determined from scans obtained at 0–180 min. **Results:** The mean Rv' (2.93 ± 0.59) underestimated the mean R_V (3.10 ± 0.50) by 5%. The mean variability (mean percent absolute difference) between Rv' and Rv was low (10%) with excellent reliability (intraclass correlation coefficient, $\rho = 0.90$). The relationship between Rv' and Rv was linear ($r = 0.95$, $p < 10^{-5}$). The mean test/retest Rv and mean test/retest Rv' were (3.19 ± 0.70/3.18 ± 0.80) and (3.16 ± 0.81/3.01 ± 0.94), respectively, and these measures were not significantly different between test/retest studies. The mean test/retest variability of Rv was low (5%) with excellent reliability ($\rho = 0.98$). In addition, the mean test/retest variability of Rv' was low (10%) with excellent reliability ($\rho = 0.94$). **Conclusion:** Three short (20 min) IBF-SPECT scans allowing for rest periods between scans permit reliable measurements of the dopamine D2-receptor parameter V_3/V_2. Quantitative IBF-SPECT imaging without blood data is reliable and reproducible.

Key Words: iodine-123-iodobenzofuran; dopamine D2-receptors; brain SPECT; simplified quantification; reproducibility

J Nucl Med 1997; 38:31–37

SPECT imaging of dopamine D2-receptors using ^{123}I-labeled benzamide analogs such as [^{123}I]iodobenzamide (IBZM) and [^{123}I]iodobenzofuran (IBF) is a potentially useful tool to evaluate patients with several neuropsychiatric disorders (*1–12*). Given the widespread availability and lower costs of SPECT compared with PET, D2-receptor SPECT imaging promises to play an important clinical role. Critical to the application of this technique to clinical studies, however, is the availability of simple, readily accessible, yet valid methods to obtain quantitative information about the receptor. In addition, information on the reliability and reproducibility of quantitative SPECT measurements is essential in order to perform power calculation for prospective SPECT investigations or to evaluate the significance of serial changes in the dopamine receptor status in patients, for example, with progressive extrapyramidal disorders.

Received Jan. 29, 1996; revision accepted Apr. 10, 1996.

For correspondence or reprints contact: Masanori Ichise, MD, FRCP (C), Rm. 635, Nuclear Medicine, Dept. of Medical Imaging, Mount Sinai Hospital, 600 University Ave., Toronto, Ontario, Canada M5G 1X5.

With IBZM or IBF-SPECT, the tissue ratio of the basal ganglia to the frontal cortex or the cerebellum measured at a fixed time after a bolus injection has been commonly used as an outcome measure of D2-receptor binding (*10*). This is a simple and practical approach compared to a more complex tracer kinetic modeling, which requires both invasive arterial blood sampling and rapid serial SPECT acquisition (*13*), limiting its usefulness as a routine clinical tool. With both IBZM and IBF, which bind reversibly to D2 receptors, this tissue ratio at a fixed time postinjection may not accurately reflect the receptor density. This ratio reflects the dynamic interplay of the receptor density/affinity and nonreceptor factors such as ligand delivery or regional cerebral blood flow (rCBF) and the clearance of ligand from plasma.

We previously showed that the receptor parameter k_3/k_4, ratio of the transfer constants between the intracerebral nondisplaceable and specifically bound receptor compartments, can be measured by using a variation of the graphical analysis method that derives the ratio of ligand distribution volumes ($R_V = V_3/V_2$) from serial SPECT scan data (*14,15*). Advantages of this approach are: elimination of invasive arterial blood sampling, outcome measure Rv is independent of rCBF and plasma clearance (*14,16*), distribution volume ratio is a more stable measure than individual kinetic constants (*13,17,18*) and Rv, being obtained without blood data, is additionally free of errors of plasma measurements including errors of metabolite corrections as well as errors of cross-calibration between plasma and tissue measurements. However, a disadvantage of this method is the need for continuous SPECT scanning for at least 2 hr to obtain stable values of Rv (*14*). This may be difficult to implement, particularly for those patients who are elderly, agitated or have hyperkinetic movement disorders.

Therefore, in the present study, we assessed the feasibility of a simplified procedure to obtain this outcome measure in which only three separate 20-min scans are required, allowing for rest periods between scans. In addition, we evaluated the reproducibility of SPECT measurements of IBF binding without blood data for both the original and simplified procedures.

MATERIALS AND METHODS

Subjects

Twenty healthy volunteers (10 women, 10 men; mean age 31.3 ± 9.3 yr) with no current or past history of neuropsychiatric disorders or family history of movement disorders were included. All patients were free of drugs. Of these, 14 were included in our previous study (*14*) and the remaining six (3 women, 3 men; mean age 30.7 ± 7.9 yr) were newly recruited for test/retest studies. Before and after the IBF study, subjects were orally given 400 mg of potassium perchlorate. Every patient gave written informed consent. The project was approved by the Human Subjects Review Committee of the University of Toronto.

Labeling of Iodine-123-IBF

Labeling of [^{123}I]IBF was performed as described previously (*14*). The radiochemical yield was 86.3% ± 6.3% and the radiochemical purity was 97.2% ± 2.4%. Retrospective sterility testing was negative.

SPECT Imaging

SPECT imaging was performed using a triple-headed system equipped with ultra-high-resolution, fan-beam collimators and interfaced to a computer. Each subject received an intravenous IBF injection 292 ± 37 MBq (7.9 ± 1.0 mCi) over 30 sec, while six subjects had a second IBF injection 292 ± 26 MBq (7.9 ± 0.7 mCi) for retest 1 wk later, at the same time of the day, under the same experimental conditions. Scans were acquired using a continuous scan mode every 5 min for 180 min. For each scan, 100 twenty 7.5-sec projection images were obtained using 3° angle intervals, 128 × 128 matrix, 360° rotating each head 120°. The radius of rotation was fixed at 13.5 cm. During the retest scan, efforts were made to position the subject's head in the headholder dedicated for brain imaging in a reproducible fashion as well as to place the fiducial markers to determine the canthomeatal line (CML) in the same fashion by identifying the previously used anatomical landmarks in each subject. FWHM of the system was 9.1 mm in water at the center of the field of view. The mean sensitivity of the system was 612 ± 12 cpm/μCi and varied less than 2% and 5% within and between the experiments, respectively. Dead time count losses over this activity range were negligible (*14*).

SPECT images were reconstructed every 10 min on a 64 × 64 matrix to ensure adequate count-per-pixel statistics. One-pixel thick (4.33 mm) transaxial slices from the vertex of the brain to the level of the CML, as identified by the fiducial markers, were reconstructed parallel to the CML using a three-dimensional Butterworth postreconstruction filter (order = 10, cutoff frequency = 0.25) after applying a ramp backprojection filter. Attenuation correction was performed by assuming uniform attenuation equal to that of water (μ = 0.15 cm^{-1}) (*19*) within an ellipse drawn around the skull as identified by the fiducial markers.

Data Analysis

From each set of transaxial images, four consecutive slices (1.73 cm thick) corresponding to the highest signal in the basal ganglia were summed. Regions of interest (ROIs) were placed using templates over the right and left basal ganglia (BG, volume = 4.87 cm^3 each) and frontal cortex (FC, volume = 6.82 cm^3) as described previously (*14*). Average counts per pixel from each region were decay-corrected to the time of injection. Counts from right and left basal ganglia were averaged. No attempts were made to correct for partial volume or scatter effects. Time-to-peak uptake in the basal ganglia, the frontal cortex and specific binding (basal ganglia-frontal cortex), was quantified after fitting a sum of three exponentials to the regional time-activity data as described previously (*14*). Time-to-peak uptake in the basal ganglia and specific binding and distribution volume ratios as described below were used to evaluate the reproducibility of IBF binding.

Variation of Graphical Analysis Method

The variation of the graphical analysis method of Logan et al. (*20*) as we proposed previously (*14*) allows calculation of the receptor parameter, $R_V = k_3/k_4$, for radioligands that reversibly bind to receptors without requiring the knowledge of blood ligand concentrations. This method is similar to the method described by Eckernas et al. (*21,22*) in that a reference region instead of an arterial input function is used, linear regression is also used. The operational equation in the variation of the graphical analysis is given by:

$$\frac{\int_0^t C_{BG}(t)dt}{C_{BG}(t)} = \left(\frac{a}{a'}\right)\frac{\int_0^t C_{FC}(t)dt}{C_{BG}(t)} + \left(-\frac{ab'}{a'}\right)\frac{C_{FC}(t)}{C_{BG}(t)} + b, \qquad \text{Eq. 1}$$

for times in which secular equilibrium can be assumed between the local precursor pool and the receptor pool, where $C_{BG}(t)$ and $C_{FC}(t)$ represent time-activity measurements in the D2-receptor-rich basal ganglia and receptor-devoid frontal cortex, respectively, and a, a′, b and b′ are constants. Equation 1 is a multilinear equation with partial regression coefficients, a/a′, −ab′/a′ and b. Of interest to us is a/a′ which is given by $(V_2 + V_3 + V_P)/(V_2^* + V_P)$ where V_2 and V_3 are the equilibrium distribution volumes of the ligand in the nondisplaceable and receptor compartments of the basal ganglia, respectively, whereas V_2^* is the equilibrium distribution volume of the nondisplaceable compartment of the frontal cortex; and V_P is the plasma volume within the tissue. Although this linear relationship cannot be graphically analyzed by using conventional graph plotting techniques, these coefficients can be obtained by multilinear regression analysis. If we assume $V_2 = V_2^*$ and V_P is negligible as described previously (*14*), then the ratio of V_3 to V_2 (distribution volume ratio, R_V):

$$R_V = \frac{V_3}{V_2} = \frac{a}{a'} - 1. \qquad \text{Eq. 2}$$

In experiments using tracer doses of ligand, the binding potential (BP) or B_{max}/K_d is identical to V_3 and R_V is identical to k_3/k_4 (*14*). Hence,

$$R_V = \frac{BP}{V_2} = \frac{k_3}{k_4} = \frac{a}{a'} - 1. \qquad \text{Eq. 3}$$

The b in Equation 1 is not constant but reaches constant after some time t* when the transport of ligand from plasma to tissue becomes unidirectional. This point in time is identified in the original graphical analysis method using blood data by observing when the plot becomes linear (*20*). In the present method, this time point was estimated by examining the residual values after fitting a multilinear regression equation using all data points. The residual value is the deviation of a particular point from the value predicted by the regression equation. Residual analysis identifies outliers that significantly deviate from the regression (*23,24*). Thus, the residuals from those early points before some time t* are expected to be significantly larger than the rest. Outliers were defined as those points with residuals outside the limit of ± 2.5 s.d. of residuals. To illustrate the significance of residual analysis, the observed values of the dependent variable, in Equation 1, were plotted against the predicted values of the same variable by the regression in the present study.

In addition, the adequacy of the regression model was tested by examining the correlation between the residuals, after adjusting for the two independent variables on the right-hand side of Equation 1. These partial correlations, r_A for the first term and r_B for the second term, represent the unique contribution of the respective independent variable to the prediction of the dependent variable and vary from 0 (no contribution) to 1 (total contribution) (*23,24*).

Identifiability of partial regression coefficients, a/a′, −ab′/a′ and b, were assessed by examining the standard errors of the estimate of the respective coefficient expressed in percentage of estimate. Like the standard errors given by the diagonal of the covariance matrix in the case of parameter estimations using nonlinear regression (*25*), these standard errors are a measure of the identifiability of the parameters by the multilinear regression process. It should be noted that they are not the same as the s.d. of the parameter in the sample.

Simplified Measurement of Distribution Volume Ratio (Rv')

To measure the distribution volume ratio from multiple IBF-SPECT scan data in such a way that the scanning may be more tolerable than the continuous 3-hr scan, we devised and evaluated an alternative scan protocol in which subjects will undergo three separate 20-min scans, the first scan immediately after injection, second beginning at 50 min and third at 160 min postinjection, respectively. This protocol was chosen from several protocols after preliminary evaluation to see if they satisfied predefined, arbitrary criteria. The rationale and criteria for devising such protocols were as follows: the multilinear regression analysis using Equation 1 requires at least four data points. Each data point corresponds to 10 min in scan acquisition time. Thus, a 20-min scan provides two data points. The average variability between the parameter obtained using the simplified protocol (Rv') and that obtained from the original 3-hr protocol (Rv) must be within 10%. Because Equation 1 requires areas under $C_{BG}(t)$ and $C_{FC}(t)$ time-activity curves (TAC), which are numerically calculated using the trapezoid rule, datasets should include the peak times of $C_{BG}(t)$ and $C_{FC}(t)$, respectively. Each scanning session must be $\leq$ 20 min in acquisition duration and the number of such sessions must be as low as possible.

Statistical Analysis

All statistical analyses including multilinear regression analysis were implemented in STATISTICA (StatSoft, Inc., Tulsa, OK). Two-tailed Student's t-tests for paired samples were used for a comparison between Rv' and Rv and between test and retest measures. Linear regression analysis was used to determine the relationship between R_V and R_V'. The test/retest variability of the regional peak time, Rv, and Rv' as well as the variability between Rv' and Rv were calculated as the absolute value of the difference between the two measurements, expressed as a percentage of the mean value of both measurements. Analysis of variance (ANOVA) with a repeated measure design was used to compare differences of the s.e. of estimate between partial regression coefficients as well as differences of the test/retest variability among regional peak times and distribution volume ratios. When an F-test was significant, individual means were compared by a posthoc Sheff test to correct for multiple comparisons (*26*). The reliability of the two measurements between simplified and original protocols as well as the reliability of the two measurements between test and retest was assessed by calculating the intraclass correlation coefficient, according to the following equation (*27*):

$$\rho = \frac{MSBS - MSWS}{MSBS + (k-1)MSWS}, \qquad \text{Eq. 4}$$

where MSBS and MSWS are the mean sum of squares between and within subjects, respectively, and k is the number of within-subject measurements, being 2 in the present study. This coefficient is an estimate of the reliability of the two sets of measurements and varies from 0 (no reliability) to 1 (total reliability). Statistical significance was defined as $p < 0.05$. Summaries of study variables were expressed as mean $\pm$ s.d.

RESULTS

Regional TAC

In the 20 subjects, basal ganglia activity showed a peak uptake at 59 $\pm$ 16 min (range 33–92 min) postinjection and a slower washout than that for the frontal cortex, whereas frontal cortex activity showed an early peak uptake (<20 min) and rapid washout. Specific binding activity showed a peak occurring later (95 $\pm$ 16 min, range 67–122 min) than that for the basal ganglia. In the subgroup of six subjects who had test/retest studies, the mean variability of the basal ganglia peak uptake

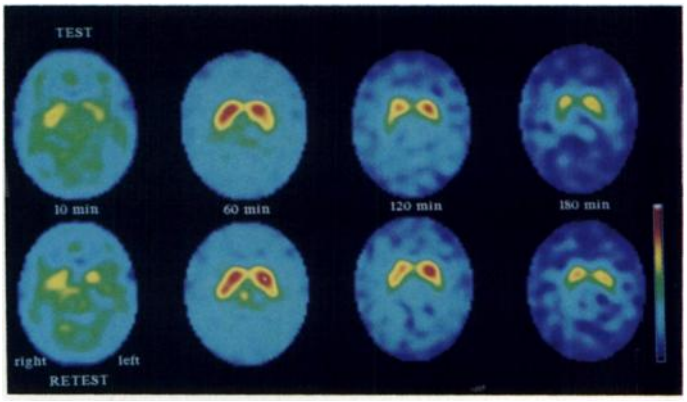

FIGURE 1. Selected serial transaxial IBF-SPECT images at the level of the striatum after injection of 7.59 mCi for test (top row) and 7.77 mCi for retest 1 wk later, respectively, in a 25-yr-old male (Subject 13). Images have been corrected for decay. Test/retest Rv and test/retest Rv' were (4.18/4.42) and (4.46/4.76), respectively.

time between test and retest was low (9.6% $\pm$ 8.3%) with relatively good reliability (intraclass correlation, $\rho = 0.70$). The mean variability of the specific binding peak time was likewise low (11.6% $\pm$ 6.9%) but with somewhat lower reliability (ρ = 0.60) compared with that for the basal ganglia. Selected IBF images of a 25-yr-old man (Patient 13) who had test/retest studies are shown in Figure 1 and the corresponding regional IBF TAC in Figure 2.

Distribution Volume Ratio

Examination of the residuals after multilinear regression analysis including all 18 data points for the original protocol and all six data points for the simplified protocol, respectively, showed no outliers. In particular, the residuals of the initial four points fell within $\pm$ 2 s.d. of residuals in each subject. The linear relationship between the observed values of the dependent variable in Equation 1 and the predicted values by the regression for Subject 13 is illustrated in Figure 3. Thus, the b in Equation 1 was assumed to become constant very early for IBF and all early data points were included in the regression analysis.

Multilinear regression analysis was highly significant ($F_{2,15} = 4700 \pm 2500$ (range, 1400–8800), $r = 0.999 \pm 0.001$, $p < 10^{-6}$ for 3-hr protocol and $F_{2,3} = 5200 \pm 4500$ (range,

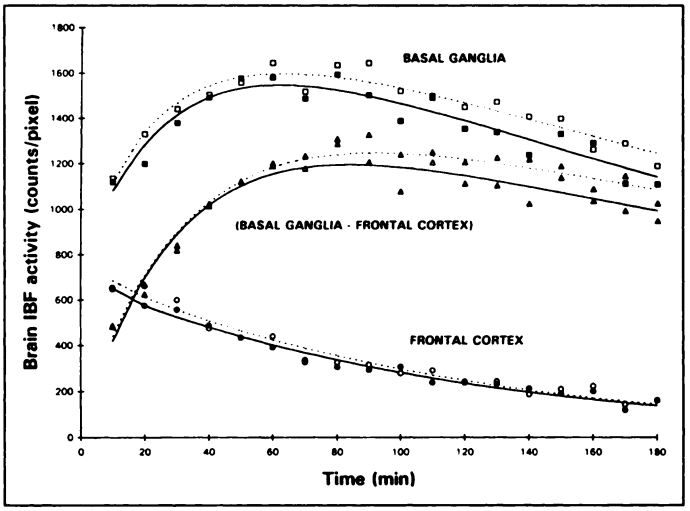

FIGURE 2. Regional IBF activity curves of the same subject in Figure 1 including basal ganglia activity (■ and □), frontal cortex activity (● and ○) and specific binding (indicated by ▲ and △). Solid fitted lines and solid symbols belong to test whereas the fitted dotted lines and hollow symbols to retest.

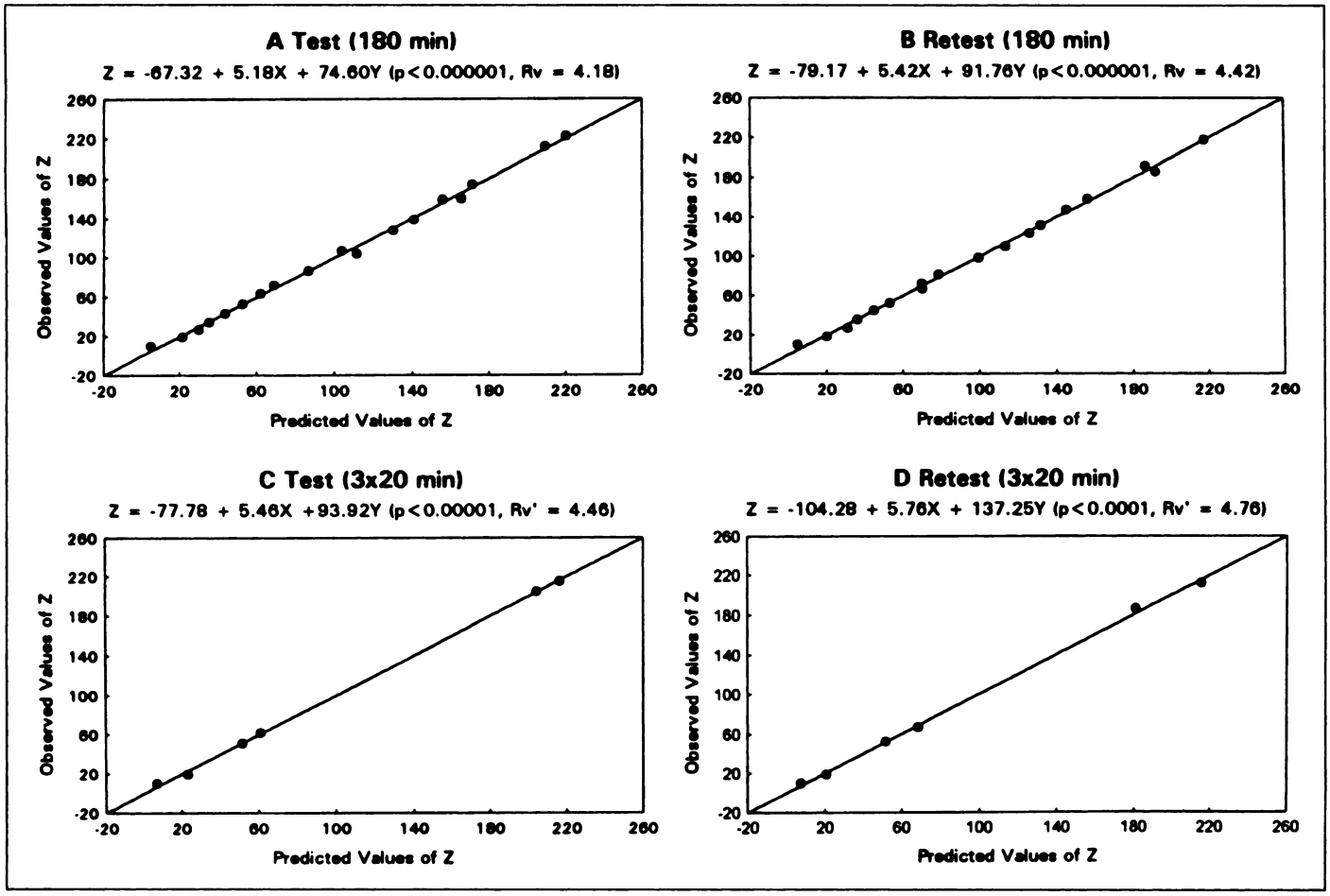

FIGURE 3. Relationship between the observed values of the dependent variable (●) and the predicted values by the regression for the same subject in Figures 1 and 2. Solid lines represent the line of identity with the predicted value. A (test) and B (retest) represent the original protocol whereas C (test) and D (retest) represent the simplified protocol.

500–16000), r = 0.999 ± 0.002, $p < 0.0001$ for simplified protocol, respectively).

Both of the two independent variables significantly contributed to the prediction of the dependent variable in Equation 1 (r_A = 0.996 ± 0.002, $p < 10^{-6}$, r_B = 0.866 ± 0.062, $p < 0.0001$ for 3-hr protocol and r_A = 0.998 ± 0.001, $p < 0.001$, r_B = 0.950 ± 0.050, $p < 0.01$ for simplified protocol, respectively). However, the mean value of the partial correlation r_A was significantly higher than that of r_B for both the original and simplified protocols (paired t-test, $p < 0.001$).

The mean s.e. of estimate of the partial regression coefficients were 2.1% ± 0.7%, 14.6% ± 4.3% and 9.0% ± 3.0% for a/a′, −ab′/a′ and b, respectively, in the original protocol; and 2.6% ± 1.5%, 17.4% ± 9.3% and 12.1% ± 6.8%, respectively, in the simplified protocol. The mean s.e. of estimate was significantly lower for a/a′ compared to −ab′/b′ or b (ANOVA: $p < 10^{-5}$ and post hoc comparison: $p < 10^{-5}$) in both the original and simplified protocols.

Table 1 summarizes the individual values of distribution volume ratios for the group of 20 subjects. The mean value of Rv′ (2.93 ± 0.59) underestimated that of R_V (3.10 ± 0.50) by 5% (paired t-test, p = 0.001). The mean variability between Rv′ and Rv was low (9.7% ± 6.9%, range 0.6%-18.9%) with excellent reliability (ρ = 0.90). The relationship between Rv′ and Rv was linear (slope = 1.13, r = 0.95, $p < 0.00001$) (Fig. 4).

Table 2 summarizes the individual values of distribution volume ratios for the subgroup of six subjects who had test/retest studies. The mean values of test/retest Rv and

TABLE 1
Distribution Volume Ratios, Variability and Reliability between Original (Rv) and Simplified (Rv′) Protocols in 20 Healthy Subjects

Subject no.	Age (yr)	Sex	Rv	Rv′	Variability*
1	19	F	3.31	2.97	10.8
2	21	F	3.50	3.52	0.6
3	22	F	3.66	3.42	6.8
4	23	F	3.01	2.49	18.9
5	25	F	2.99	2.78	7.3
6	25	F	2.85	2.82	1.1
7	33	F	3.04	2.85	6.5
8	33	F	3.23	3.02	6.7
9	38	F	2.76	2.52	9.1
10	38	F	2.29	2.20	4.0
11	22	M	2.95	2.93	0.7
12	23	M	3.78	3.89	2.9
13	25	M	4.18	4.46	6.5
14	26	M	2.48	2.46	0.8
15	34	M	3.41	3.24	5.1
16	41	M	2.61	2.21	16.6
17	41	M	3.46	3.29	5.0
18	44	M	3.24	2.81	14.2
19	45	M	3.01	2.85	5.5
20	47	M	2.16	1.96	9.7
Mean ± s.d.	31.3 ± 9.2		3.10 ± 0.50	2.93 ± 0.59	6.9 ± 5.2
Reliability†			0.90		

*Absolute values of the difference between Rv and Rv′ expressed as percentage of the mean of the Rv and Rv′.
†Intraclass correlation coefficient.

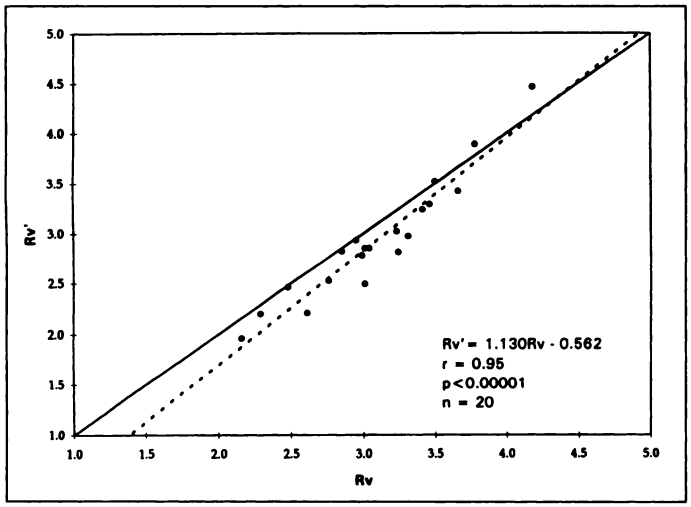

FIGURE 4. Relationship between Rv' (●) and R_V. The solid and dotted lines represent the line of identity with R_V and the regression line, respectively.

test/retest Rv' were (3.19 ± 0.70/3.18 ± 0.80) and (3.16 ± 0.81/3.01 ± 0.94), respectively, and these measures were not significantly different between test/retest studies. The mean test/retest variability of Rv was low (4.4% ± 3.1%, range 0.6%-5.6%) with excellent reliability (ρ = 0.98). The mean test/retest variability of Rv' was also low (9.4% ± 4.9%, range 4.0%-15.3%) with equally excellent reliability (ρ = 0.94). Although Rv had a lower mean test/retest variability than did Rv' or peak uptake times, there were no statistically significant variability differences between these measures due to the relatively large intersubject variation in variability (F = 1.5, p = 0.25).

DISCUSSION

Simplified Quantification

The variation of the graphical analysis method to derive the distribution volume ratio without blood data used in the present study is insensitive to changes in rCBF and plasma clearance (*14*). Rv, which is the same measure as the V_3'' as defined by Laruelle et al. (*13*), should be a more stable measure than BP itself because Rv is free of the errors of cross calibration and metabolite corrections. However, the major limitation of Rv is that Rv can reflect BP only if V_2 is uniform across subjects, but this assumption may not be valid (*13*). The clinical significance of this measure, which incorporates V_2 in addition to BP, needs to be evaluated in various disease conditions.

In the present study, two additional potential problems associated with our variation of the graphical analysis method were addressed. The first problem is that of identifying when Equation 1 becomes linear. Residual analysis used for this purpose may be an indirect method, but the regression can be improved iteratively by deleting early outliers each time. This iteration was not necessary in the present study because the time point t* appears to be reached very early for IBF. However, other ligands such as IBZM may have a different value of t*.

The second potential problem with the multilinear equation (Eq.1) is related to the considerable differences of the two dependent variables in terms of their magnitudes and their variation with time. The values of partial correlations, however, indicated that both independent variables significantly contributed to the prediction of the dependent variable in the present study although the correlation was significantly higher for the first term. Thus, the coefficient a/a', which is the parameter of interest, was very well identified by the regression. However, this may not always be the case depending on the ligand and experimental conditions.

Our method is an alternative to constant tracer infusion paradigms, which allow single scan measurements of V_3/V_2 during a prolonged equilibrium phase (*28,29*), although it takes 7 hr for IBF to reach equilibrium after a bolus injection followed by a constant infusion (*29*). However, the three 20-min SPECT scan protocol presented in this study permits measurements of the distribution volume ratio with low variability and excellent reliability as compared with the original 3-hr protocol.

In order to accurately estimate the time integrals in the multilinear equation, data points at the time to peak uptake in the basal ganglia are required. In the simplified protocol, the data points at the average time to peak uptake were used. This simplification introduces errors in estimating the volume ratio. This vulnerability of Rv' but not Rv to the intersubject variation in time to peak uptake, however, is indirect in that this variation affects only a part of the multilinear equation to derive the volume ratio. The specific binding-to-frontal cortex ratio measured at the average time of peak specific binding (empirical ratio) is more directly vulnerable to the intersubject variation in time to peak uptake. Rv' correlated better with Rv than did this empirical ratio with Rv in the present study (data not shown). Although the simplified method was reasonable in estimating the volume ratio in our young healthy subjects, there may be significantly larger variations in the peak uptake time with age and disease states. In that case, the simplified method without a priori knowledge of the peak uptake may introduce larger errors in estimating the volume ratio.

Crucial to implementing this simplified protocol might be the

TABLE 2
Distribution Volume Ratios, Variability and Reliability between Test and Retest in Six Healthy Subjects

	Rv (3-hr protocol)			Rv' (simplified protocol)		
Subject no.	Test	Retest	Variability*	Test	Retest	Variability*
2	3.50	3.39	3.2	3.52	3.02	15.3
8	3.23	3.25	0.6	3.02	2.88	4.8
10	2.29	2.08	9.6	2.20	1.93	13.1
13	4.18	4.42	5.6	4.46	4.76	6.5
14	2.48	2.59	4.3	2.46	2.56	4.0
17	3.46	3.36	2.9	3.29	2.90	12.6
Mean ± s.d.	3.19 ± 0.70	3.18 ± 0.80	4.4 ± 3.1	3.16 ± 0.81	3.01 ± 0.94	9.4 ± 4.9
Reliability†	0.98			0.94		

*Absolute values of the test/retest difference expressed as percentage of the mean of the test and retest.
†Intraclass correlation coefficient.

reproducibility of repositioning of subject's head in the head holder. However, the three-dimensional volume imaging capability of a gamma-camera type SPECT should be advantageous over a ring detector type SPECT or PET because it allows accurate reorientation of tomographic slices according to the fiducial markers without sacrificing z-axis resolution, provided that these markers have been securely held in the same place throughout the experiment.

Reproducibility Studies

In the present study, the respective reproducibility of peak uptake times and distribution volume ratios was all high. However, the reliability of the time to peak uptake was considerably lower than that of the volume ratio. This lower reliability is due to the smaller relative intersubject variance in Equation 4 compared to the volume ratio, presumably because time to peak uptake is determined by multiple factors such as rCBF, plasma clearance and receptor affinity, which together may have reduced the intersubject variance.

The test/retest reproducibility of dopamine D2-receptor binding recently has been reported by Volkow et al. using ^{11}C-raclopride and PET (*30*). The mean test/retest variability in the distribution volume ratio was 5.3%. Despite the significant methodology differences between their study and ours, our reproducibility study results are in excellent agreement with theirs. The limitation of a lower sensitivity of SPECT compared with PET may have been offset by the longer time for IBF to reach peak uptake in the basal ganglia (60 min) compared with ^{11}C-raclopride (14 min), allowing adequate sampling time for SPECT to characterize regional time activity required for the derivation of the distribution volume ratio.

Abi-Dargham et al. (*31*) and Seibyl et al. (*32*) reported the reproducibility of SPECT measurements of benzodiazepine receptor binding and dopamine transporters, respectively. The test/retest variability of ^{123}I-iomazenil distribution volume ratios was 5% to 8% whereas that of V_T' was 10%. The test/retest variability of V_3'' using ^{123}I-β-CIT was 6.8%. A significant factor for this highly reproducible SPECT measurement in the latter study may have been adequate sampling (45 min), which was feasible because of the prolonged equilibrium and a long physical half-life of ^{123}I. Although there are major differences in imaging and data analysis techniques as well as receptor systems studied, the reproducibility results of these SPECT studies, including the current study, are in good agreement, supporting the notion that quantitative SPECT imaging of neuroreceptors is a potentially viable methodology.

In addition, although care was taken in the present study to ensure accurate repositioning of the subject's head as well as accurate replacement of the fiducial markers to identify the CML, the reproducibility of this procedure needs to be formally evaluated. Other potential sources of test/retest variability include subject movement during scan and the method of ROI placement. Furthermore, the ROI size may affect the reliability of measurements (*33*). The former needs to be addressed because subject movement is likely to be of major significance in patients with movement disorders. We are currently addressing the latter issue by developing an automated computer program in which SPECT and MRI data in the same subject are coregistered three-dimensionally. This method should improve the reproducibility of ROI placement and allow evaluation of effects of ROI size in a reproducible fashion.

CONCLUSION

The simplified scan protocol consisting of three short IBF-SPECT scans allowing for rest periods presented in this study permits measurements of the D2-receptor parameter V_3/V_2 with low variability and excellent reliability compared with the original 3-hr protocol. In addition, quantitative IBF-SPECT imaging without blood data is highly reliable and reproducible. Further studies are warranted to evaluate the feasibility of implementing the simplified protocol clinically, validity of this technique in both aged and diseased subjects and reproducibility studies in the aged and diseased subjects.

ACKNOWLEDGMENTS

This work was supported by the SPECT Research Fund from the Department of Medical Imaging, Mount Sinai Hospital, partly contributed to by Picker International Canada Inc.

REFERENCES

1. Kung HF, Pan S, Kung MP, et al. In vitro and in vivo evaluation of [^{123}I]IBZM: a potential CNS D2 dopamine receptor imaging agent. *J Nucl Med* 1989;30:88–92.
2. Kung MP, Kung HF, Billings J, Yang Y, Murphy RA, Alavi A. The characterization of IBF as a new selective dopamine D2 receptor imaging agent. *J Nucl Med* 1990;31:648–654.
3. Brucke T, Podreka I, Angelberger P, et al. Dopamine D2 receptor imaging with SPECT: studies in different neuropsychiatric disorders. *J Cereb Blood Flow Metab* 1991;11:220–228.
4. Ichise M, Toyama H, Fornazzari L, Ballinger JR, Kirsh JC. Iodine-123-IBZM dopamine D2 receptor and technetium-99m-HMPAO brain perfusion SPECT in the evaluation of patients with and subjects at risk for Huntington's disease. *J Nucl Med* 1993;34:1274–1281.
5. Oertel WH, Schwarz J, Tatsch K, Arnold G, Gasser T, Kirsch CM. IBZM-SPECT as predictor for dopamimetic responsiveness of patients with de novo parkinsonian syndrome. *Adv Neurol* 1993;60:519–524.
6. Tatsch K, Schwarz J, Oertel WH, Kirsch CM. SPECT imaging of dopamine D2 receptors with ^{123}I-IBZM: initial experience in controls and patients with Parkinson's syndrome and Wilson's disease. *Nucl Med Commun* 1991;12:699–707.
7. Schwarz J, Tatsch K, Arnold G, et al. Iodine-123-iodobenzamide SPECT predicts dopaminergic responsiveness in patients with de novo parkinsonism. *Neurology* 1992;42:556–561.
8. Laulumaa V, Kuikka JT, Soininen H, Bergstrom K, Lansimies E, Riekkinen P. Imaging of D2 dopamine receptors of patients with Parkinson's disease using SPECT and iodobenzamide ^{123}I. *Arch Neurol* 1993;50:509–512.
9. van Royen E, Verhoeff NF, Speelman JD, Wolters EC, Kuiper MA, Janssen AG. Multiple system atrophy and progressive supranuclear palsy. Diminished striatal D2 dopamine receptor activity demonstrated by ^{123}I-IBZM SPECT. *Arch Neurol* 1993; 50:513–516.
10. Schlosser R, Schlegel S. D_2-receptor imaging with [^{123}I]IBZM and SPECT in psychiatry: a survey of current status. *J Neural Transm* 1995;99:173–185.
11. Knable MB, Jones DW, Coppola R, et al. Lateralized differences in iodine-123-IBZM uptake in the basal ganglia in asymmetric Parkinson's disease. *J Nucl Med* 1995;36: 1216–1225.
12. Buck A, Westera G, Sutter M, Albani C, Kung HF, vonSchulthess GK. Iodine-123-IBF SPECT evaluation of extrapyramidal diseases. *J Nucl Med* 1995;36:1196–1200.
13. Laruelle M, van Dyck C, Abi-Dargham A, et al. Compartmental modeling of iodine-123-iodobenzofuran binding to dopamine D2-receptors in healthy subjects. *J Nucl Med* 1994;35:743–754.
14. Ichise M, Ballinger JR, Golan H, et al. Noninvasive quantification of dopamine D2-receptors with iodine-123-IBF SPECT. *J Nucl Med* 1996;37:513–520.
15. Ichise M, Ballinger JR. From graphical analysis to multilinear regression analysis of reversible radioligand binding [Letter]. *J Cereb Blood Flow Metab* 1996;16:750–751.
16. Holthoff VA, Koeppe RA, Frey KA, Paradise AH, Kuhl DE. Differentiation of radioligand delivery and binding in the brain: validation of a two-compartment model for [^{11}C]flumazenil. *J Cereb Blood Flow Metab* 1991;11:745–752.
17. Volkow ND, Ding YS, Fowler JS, et al. A new PET ligand for the dopamine transporter: studies in the human brain. *J Nucl Med* 1995;36:2162–2168.
18. Sawada Y, Hiraga S, Francis B, et al. Kinetic analysis of 3-quinuclidinyl 4-[^{125}I]iodobenzilate transport and specific binding to muscarinic acetylcholine receptor in rat brain in vivo: implications for human studies. *J Cereb Blood Flow Metab* 1990;10: 781–807.
19. Chang L. A method for attenuation correction in radionuclide computed tomography. *IEEE Trans Nucl Sci* 1978;NS-25:638–643.
20. Logan J, Fowler JS, Volkow ND, et al. Graphical analysis of reversible radioligand binding from time-activity measurements applied to [N-^{11}C-methyl]-(-)-cocaine PET studies in human subjects. *J Cereb Blood Flow Metab* 1990;10:740–747.
21. Eckernas SA, Aquilonius SM, Hartvig P, et al. PET in the study of dopamine receptors in the primate brain: evaluation of a kinetic model using ^{11}C-N-methyl-spiperone. *Acta Neurol Scand* 1987;75:168–178.
22. Hartvig P, Eckernas SA, Ekblom B, et al. Receptor binding and selectivity of three ^{11}C-labeled dopamine receptor antagonists in the brain of rhesus monkeys studied with PET. *Acta Neurol Scand* 1988;77:314–321.
23. Armitage P, Berry G. Multiple measurements. In: *Statistical methods in medical research*, 2nd ed. Oxford: Blackwell Scientific Publications; 1987:296–357.
24. Multiple regression. In: *STATISTICA for Windows*. Tulsa OK: StatSoft, Inc.; 1994: 1485–1530.
25. Carson RE. Parameter estimation in positron emission tomography. In: Phelps ME, Mazziota JC, Schelbert HR, eds. *Positron emission tomography. Principles and applications for the brain and the heart*. New York: Raven Press; 1986:347–390.

26. Armitage P, Berry G. Comparison of several groups: multiple comparisons. In: *Statistical methods in medical research*, 2nd ed. London: Blackwell Scientific Publications; 1996:200–205.
27. Snedecor GW, Cochran WG. Specific indices of interater reliability. In: Snedecor GW, Cochran WG, eds. *Statistical Methods*, 6th ed. Ames, IO: Iowa State University Press; 1989:147–146.
28. Carson RE, Channing MA, Blasberg RG, et al. Comparison of bolus and infusion methods for receptor quantitation: application to [^{18}F]cyclofoxy and positron emission tomography. *J Cereb Blood Flow Metab* 1993;13:24–42.
29. Laruelle M, Abi-Dargham A, Van Dyck CH, et al. D-amphetamine induced dopamine release reduces the binding potential of [^{123}I]IBF for dopamine D2/D3 receptors in humans [Abstract]. *J Nucl Med* 1994;35:84P.
30. Volkow ND, Fowler JS, Wang GJ, et al. Reproducibility of repeated measures of carbon-11-raclopride binding in the human brain. *J Nucl Med* 1993;34:609–613.
31. Abi-Dargham A, Gandelman M, Zoghbi SS, et al. Reproducibility of SPECT measurement of benzodiazepine receptors in human brain with iodine-123-iomazenil. *J Nucl Med* 1995;36:167–175.
32. Seibyl JP, Laurelle M, Van Dyck CH, et al. Reproducibility of iodine-b-CIT SPECT brain measurement of dopamine transporters. *J Nucl Med* 1996;37:222–228.
33. Vingerhoets FJ, Snow BJ, Schulzer M, et al. Reproducibility of fluorine-18-6-fluorodopa positron emission tomography in normal human subjects. *J Nucl Med* 1994;35:18–24.

EDITORIAL

"Yet there is method in't" (Shakespeare)

meth-od* (n.): *a way, technique, or process of or for doing something (Webster's Third International Dictionary) (1)

There is the method school of acting, and as The Bard has noted, there may be method in one's madness. What I would like to consider at this time is method in science: more specifically, the Methods Section of scientific articles. This unsung and often neglected portion of our scientific literature deserves, it seems to me, much more respect than it frequently gets.

Nuclear medicine is a technical specialty. It is reasonable to expect technical excellence in its practitioners and detail in the presentation of technically complex procedures. The Methods section of our literature thus takes on special significance. It is only through the Methods Section that we can get a feeling for the technical competence of a group of investigators. Sadly, a careful reading of much of our literature suggests, all too commonly, a lack of basic understanding of fundamental elements of technique. Even worse perhaps is the complete omission of information needed to assess that technical competence.

Recently this was put into sharp focus when I came across an article comparing SPECT imaging with planar technique for a specific application. The authors had reached the surprising (to me) conclusion that planar imaging was better. On reviewing the images that accompanied the article, the probable reason for this seemingly backward result was obvious. The SPECT images were some of the worst I had ever seen. It was bad enough that these authors could not recognize that their images were dreadful, but when I turned to the Methods Section, it was not possible to determine what had been done wrong. Did they choose the wrong collimator? Was their equipment old and out-of-date? Did they use the wrong processing parameters? It was impossible to tell because all such information was lacking from the Methods Section.

Often a neglected stepchild in the preparation, review and reading of scientific articles, the Methods Section is in fact of vital importance when a reader or reviewer tries to critically assess the value and validity of the results presented. No amount of statistical manipulation can salvage useful meaning if the data has been badly acquired. Furthermore, if a reader wishes to introduce a promising technique into her/his laboratory, the Methods Section must describe the technique in question in sufficient detail to permit its duplication.

Unfortunately, this is often not the case. Although omission of almost all technical detail as described above is unusual, the absence of small but significant details is very common. For example, as I have noted in an earlier essay on the SUV (*2*), it is frequently impossible to tell whether the SUV was calculated on the basis of the average counts within an ROI or if the maximum value was used. The time from injection of tracer to imaging is often omitted. Yet these small variations in technique can have *large* effects on the measured value.

For another example, which seems trivial until it is closely examined, let us look at the manner in which the reconstruction filter is described in SPECT and PET articles. A common description is, "A Shepp-Logan (or Hamming, or Butterworth, etc.) filter was used with a cut-off of 0.3." 0.3 what? This number is *not* dimensionless. It has units. Because different manufacturers use different units, when specifying filters, the exact units used must be given, if this number is to be of any use to the reader.

How common are such problems? A survey of three consecutive recent issues of *The Journal of Nuclear Medicine* gave the following results: Of 30 articles on either PET or SPECT, 16 gave no information about the filter used at all, 8 described the filter but gave no units, and only 6 articles completely and properly specified the filter used.

One might argue that omitting such information is trivial. Certainly articles on CT and MRI never discuss reconstruction filters. For most commercial CT systems, the filter functions are closely held proprietary information and not subject to user scrutiny or modification. This is not, however, the case for nuclear medicine systems. Due to the varying resolution and noise characteristics found with different collimators, tracers and doses administered, patients and system configurations, it is common to use different filters for different types of studies. This is further complicated by widely varying user preferences in terms of final image appearance. It is thus common practice to have the filter type and cut-off value as user specifiable parameters.

Investigators who publish an article lacking this information are doing both themselves and their readers a disservice. Certainly reviewers who pass an article lacking such information are not doing their job properly. How can it be said that one has critically reviewed an article when the reviewer does not know for sure what was done?

Largely ignored, the Methods Section is, perhaps, the single most important part of a scientific article. Only from a careful reading of the methods can one decide whether to believe the results of a study. Only with a complete explication of the

Received June 17, 1996; accepted June 17, 1996.
For correspondence or reprints contact: John W. Keyes, Jr., MD, Bowman Gray School of Medicine, Division of Radiologic Sciences, PET Center, Medical Center Blvd., Winston-Salem, NC 27157.