Journal of Cerebral Blood Flow and Metabolism
11:A26–A30 © 1991 The International Society of Cerebral Blood Flow and Metabolism

The Precision of Positron Emission Tomography: Theory and Measurement

Nathaniel M. Alpert, W. Craig Barker, *Andrew Gelman, Stephen Weise, Michio Senda, and John A. Correia

*Department of Radiology, Massachusetts General Hospital, Boston, and *Department of Statistics, Harvard University, Cambridge, Massachusetts, U.S.A.*

Summary: The limits of quantitation with positron emission tomography (PET) are examined with respect to the noise propagation resulting from radioactive decay and other sources of random error. Theoretical methods for evaluating the statistical error have been devised but seldom applied to experimental data obtained on human subjects. This paper extends the analysis in several ways: (1) A Monte Carlo method is described for tracking the propagation of statistical error through the analysis of in vivo measurements; (2) Experimental data, obtained in phantoms, validating the Monte Carlo method and other methods are presented; (3) A difference in activation paradigm, performed on regional CBF (rCBF) data from five human subjects, was analyzed on 1.6-cm diameter regions of interest to determine the mean fractional statistical error in PET tissue concentration and in rCBF before and after stereotactic transformation; and (4) A linear statistical model and calculations of the various statistical errors were used to estimate the magnitude of the subject-specific fluctuations under various conditions. In this specific example, the root mean squared (RMS) noise in flow measurements was about three times higher than the RMS noise in the concentration measurements. In addition, the total random error was almost equally partitioned between statistical error and random fluctuations due to all other sources. **Key Words:** Positron emission tomography—Random error—Statistical error.

At an elementary level of description, positron emission tomography (PET) estimates the radioactivity concentration in one or more slices through a three-dimensional object. Many factors influence the accuracy and precision of the local concentration measurements with PET. Systematic errors (deterministic inaccuracies) arise from the finite resolution of the PET scanner, both in-plane and axial, and from the coarse sampling of slice-by-slice measurements. The limitations imposed by incomplete three-dimensional sampling are not widely appreciated. Another class of systematic errors arises from scanner calibration and corrections due to factors such as photon attenuation, scattered coincidences, and scanner deadtime. The precision of the PET concentration measurements per se are determined by the propagation of random statistical error through the image reconstruction procedure, and statistical error is inherent in the nuclear decay and measurement process.

The measurement of radioactivity concentration is seldom an end in itself. Rather, the local concentration data and other measurements, such as arterial blood concentration histories, are used to calculate regional estimates of physiological variables. Typical examples include measurements of blood flow, glucose metabolism, and oxygen metabolism. Converting the concentration data to a measurement of a physiological variable involves a mathematical model and results in a quantity with a different accuracy and precision. Understanding the propagation of the systematic and random errors is important when devising new measurement strategies.

A description of the accuracy and precision of PET measurements applied to studies of human or animal subjects results in additional complexity.

Address correspondence and reprint requests to Dr. N. M. Alpert, Division of Nuclear Medicine, Massachusetts General Hospital, Fruit Street, Boston, MA 02114, U.S.A.

Abbreviations used: BGO, bismuth germanate; FWHM, full width half maximum; PET, positron emission tomography; rCBF, regional CBF; RMS, root mean squared; XCT, x-ray enhanced computed tomography.

One reason for this additional complexity is the need to compare measurements between different subjects or the mean of measurements among groups of subjects. Random variations in physiological state, as well as differences between subjects or groups, are superimposed on the statistical errors associated with radioactive counting measurements.

It is well known that when the resolution length of the PET scanner (in all dimensions) is less than one half the size of the object to be measured, the concentration is approximately independent of the resolution; otherwise the result depends on object size (finite resolution) and axial position (partial volume). If all subjects had the same anatomy and could be imaged in identical geometric positions, PET would measure uniform systematic errors specific to the object measured; however, in the real world normal anatomic variation among subjects is seen as a variation superimposed on all the other variations affecting PET measurements.

The goals of this paper are (1) to review the available methods for assessing the precision of PET measurements, (2) to assess the accuracy of these methods, and (3) to demonstrate their application to a realistic measurement situation.

MATERIALS AND METHODS

Review of methods

Many investigators have addressed the problem of reconstruction of data from projections for PET and x-ray enhanced computed tomography (XCT). The propagation of statistical error in the reconstruction has also received considerable attention. We divide reconstruction techniques into two classes, iterative and noniterative. We will restrict our discussion to the noniterative methods, as they are used in most laboratories. We will assume that there are only minor differences in noise propagation among the noniterative methods; most of our discussion applies directly to the filtered backprojection technique.

Approximation methods. Approximation formulas have been developed by several groups for estimating the noise-to-signal ratio in PET. Theoretical analyses have established the relationship among the RMS noise, the number of detected events, and the spatial resolution for the case of a disk with uniform activity distribution and no photon attenuation (Budinger et al., 1977; Brownell et al., 1979). These studies were extended to somewhat more complicated distributions of radioactivity by Budinger et al. (1978). Tanaka and Murayama (1982) have derived more general approximation formulas for the estimation of the noise variance at an arbitrary point in the reconstruction.

These approximate methods have established the general relationships among the variables. The simple approximation formula, popularized by Budinger et al. (1977), is often used to estimate the RMS noise level. However, in the experimental situation the simple approximation formulas serve best as a convenient way to compare the RMS noise in similar experiments; in absolute terms they serve only as a rough guide. In RMS noise measurements performed in our laboratory, the approximate formulas and experimental data are not in good agreement, often differing by a factor of two or more. This disagreement is not surprising considering the many factors not included in these formulas.

Analytic methods. Several analytic methods have been described for the calculation of the local variance for arbitrary distributions of radioactivity. All of these depend on the linearity of the filtered backprojection method. Huesman (1977) and Tanaka and Murayama (1982) derived formulas for computation of the local statistical noise in PET scans. Alpert et al. (1982) and Palmer et al. (1985, 1986) derived similar formulas explicitly including the effects of noise propagated by corrections to the projection data for photon attenuation and random coincidences. Huesman (1984) further extended these calculations to include the noise propagation for arbitrarily shaped regions of interest (ROIs).

Monte Carlo methods. The initial work of Budinger et al. (1977, 1978) summarized the results of Monte Carlo simulations for simple geometric distributions of radioactivity in the now well known analytic formulas alluded to above. However, the use of simulation techniques for following the propagation of random error has not, to our knowledge, been reported. The Monte Carlo method, when applied to the filtered backprojection algorithm, is straightforward, requiring only the simulation of Poisson distributed random variables with means equal to the expected values of the elements of the projection arrays. Estimates can be obtained using real data by approximating the expected values by the measured data. Calculations using the Monte Carlo approach are the most flexible of the available methods. We illustrate the use of this approach below to study the propagation of random error in experimental data for a group of subjects whose PET data are used to compute flow maps that are transformed to a stereotactic coordinate system. The other techniques described above cannot be used as conveniently for this purpose because the PET data do not adequately sample the three-dimensional space.

Noise propagation and measurements in phantom studies

In this section we present data comparing calculated and measured estimates of noise propagated by the PET image reconstruction procedure. Ideally the estimation of random errors in the reconstructed activity concentrations could be derived by computing the sample variance of repeated measurements on the same test object. In practice, such measurements are complicated by radioactive decay, making it impossible to obtain replicas that differ only by statistical noise. In fact, the relative noise level of the measurements increases as the radioactive counts diminish in number. In our studies, we fit the measured concentrations to a decreasing exponential and used the residual errors to estimate the noise as a function of time. We assumed that corrections for random coincidences, coincidence deadtime, and other rate-dependent factors were known exactly.

Two test objects were used in these studies, a cylindrical phantom with a diameter of 19.3 cm and a height of 17.5 cm, and a Capintec brain phantom (Capintec, Inc). The "20"-cm phantom was filled with ^{18}F in water to a concentration of 0.4 μCi/cc. The brain phantom, with

compartments representing grey and white matter, was filled with ^{18}F in a concentration ratio of 4:1 (4.8:1.2 μCi/cc). The phantoms were imaged with a Scanditronix PC-384 positron tomograph with three rings of bismuth germanate (BGO) crystals (Litton et al., 1984). A sequence of images was measured during the decay of the radioactivity and the data were reconstructed at a resolution of 8 mm full width half maximum (FWHM), using a standard filtered backprojection technique. An analytic correction, with contour fitting, was used to compensate for photon attenuation. The projection data were also corrected for random coincidences, scatter coincidences (Bergstrom et al., 1983), variation in wobble speed, and detector nonuniformity. After reconstruction, a set of ROIs were marked on the images and the decay curve for each was extracted. The decay curves were fit to a decreasing exponential and the residual errors of the fit were used to estimate the variance of the concentration data.

Comparison with analytic methods. Figure 1 depicts the comparison between the measured and calculated noise-to-signal ratio for a set of regions applied to the "20"-cm phantom and the brain phantom. The regions were circular, with diameters ranging from 10 to 163 mm. In the "20"-cm phantom the regions were placed both concentrically and eccentrically. Linear regression with no intercept yielded a regression slope of 1.03, a value not significantly different than 1.

Comparison with Monte Carlo studies. Figure 2 depicts the comparison between measurement and the Monte Carlo estimates. These studies, using a set of regions including those described above, yielded a regression coefficient of 0.89, which was not statistically different than 1.

Noise propagation in human studies

The patient studies consisted of blood flow measurements on five subjects, each was scanned during two different visual tasks. The underlying hypotheses of these experiments is not relevant to our purpose, the examination of the precision of a typical experimental PET paradigm. A more detailed report of this ongoing study will be presented elsewhere. Blood flow was measured using a variant of the equilibrium method (Senda et al., 1988) on 9 slices with a center-to-center spacing of 9.3 mm. We defined a dependent variable that we call the relative difference in regional activation as the difference in flow between the two conditions divided by the mean flow. We also estimated mean concentration and flow for each region analyzed; the standard deviation divided by the mean was taken as measure of the precision of these measurements.

Noise analyses were made at two stages in the analysis, in the "laboratory coordinate system" and in a stereotactic coordinate system. Measurement of flow and tissue concentration in the laboratory were made on manually selected ROIs, 1.6 cm in diameter. For measurements in the stereotactic coordinate system, we used the following procedure: An individually cast plastic-foam-molded headholder was fitted to each subject. The headholder was mounted on a frame that was rigidly attached to the computed tomography and PET scanner. The size and position (shift and rotation) of the brain were determined by (a) the midsagittal plane, located on the stack of PET images, (b) the line connecting limbus sphenoidale and sulcus transversus, identified on the scout view of XCT image and assumed to be parallel to, and a known distance from the line connecting anterior and posterior commissures (reference line of the atlas), and (c) brain contour. Assuming proportionality of the brain structures between subjects, the laboratory coordinate system was transformed to the coordinate system of the stereotactic atlas of the human brain. To reduce the effect of local morphometric variation, the brain was divided into three segments to which different scaling factors were applied. The stereotactic atlas of Talairach and Sziklag (1967) was digitized three-dimensionally, which allowed for the generation and overlaying of anatomical maps upon PET images on the image display system.

The general scheme in analyzing the data was to describe the regional difference in activation with a linear statistical model that includes classification variables for subject and region effects. Assuming the statistical model to be adequate, the total mean square error is an estimate of the overall experimental precision, which includes the effect of both the statistical error of PET scans and other

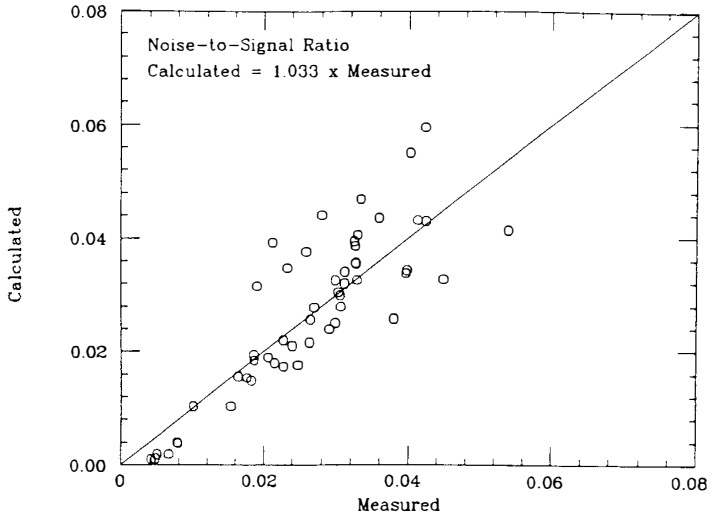

FIG. 1. Comparison of analytic calculation of RMS noise versus measured data. The calculations were performed using Huesman's (1977, 1984) method and compared to measurements on circular ROIs with varying radius and location. The solid line is the linear regression, with zero intercept. The regression equation is indicated in the graph. See text for details.

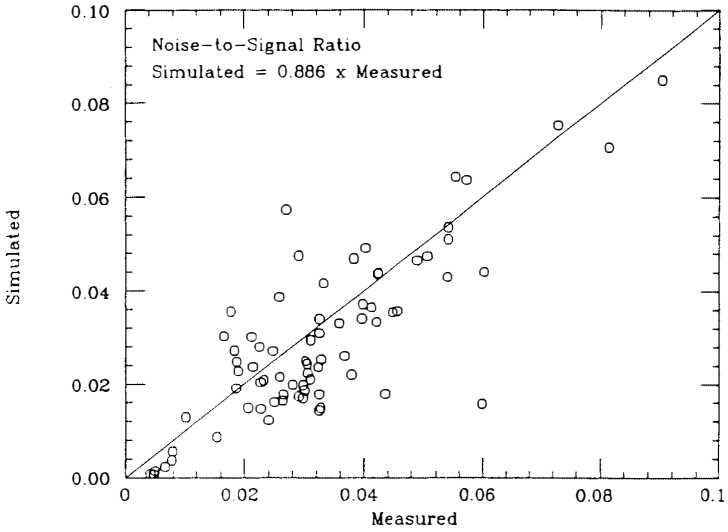

FIG. 2. Comparison of Monte Carlo estimate of RMS noise versus measured data. The Monte Carlo method was applied to measured data (as described in the text) and compared to measurements on circular ROIs with varying radius and location. The solid line is the linear regression, with zero intercept. The regression equation is indicated in the graph.

fluctuations. By subtracting an estimate of the statistical error, we estimated the magnitude of the other fluctuations and thus partitioned the mean square error as shown in Table 1. The latter error term includes experimental noise not due to fluctuations in radioactive counts, as well as true patient-specific changes in blood flow between the two measurements, perhaps due to the difference between the two tasks. Additional details are given below.

In order to follow the noise propagation through the various processing steps, we used Monte Carlo simulation techniques to produce 10 replicas per image set. FORTRAN subroutines "GGNML" and "GGPOS" from the IMSL library were used to generate independent Poisson random variables for each in-plane pair of detectors, with mean equal to the coincidence count for the detector pair. We modified the standard Scanditronix reconstruction program by replacing the raw projection data with Poisson simulations and then reconstructed to produce a set of replicas that represent samples from the distribution of possible outcomes of repeated measurements of the concentration maps. From these data we computed regional sample means and variances at each stage in the data processing.

The statistical model used in this work decomposes the difference in blood flow Δ_{ij} between the two visual tasks, for subject i and region j as described by the equation,

$$\Delta_{ij} = \delta + p_i + ROI_j + \epsilon_{ij} \quad (1)$$

where δ represents the grand mean difference in blood flow between tasks over all subjects and ROIs, p_i is the difference from the grand mean due to the i^{th} subject, ROI_j is the difference from the grand mean due to the j^{th} region and ϵ_{ij} is the residual random error associated with the ij^{th} observation. The variance of ϵ_{ij} was determined with the general linear model procedure from the SAS statistical software package.

TABLE 1. *Decomposition of mean square error*

Coordinate mapping	Statistical error	All other fluctuations	Total MSE error
Laboratory	0.0069	0.0054	0.0124
Stereoatlas	0.0020	0.0026	0.0046

RESULTS

The mean RMS noise level for the regional concentration data was 0.024 with a range of 0.016–0.040. When the concentration data were used in the ^{15}O equilibrium model to compute blood flow and the mean regional RMS noise, the mean noise-to-signal level rose to 0.059, with a range of 0.024–0.176. After transformation to stereotactic coordinates, the mean regional noise level in CBF was 0.030, with a range of 0.016–0.118. Table 1 presents the decomposition of mean square error for two methods of analysis, one based on circular ROIs defined in laboratory coordinates, and the other using stereotactic mapping to a proportional brain atlas, with ROIs of the same size chosen according to the positions for the anatomic structures specified by the brain atlas. Inspection of this table shows that the stereotactic mapping reduced the apparent statistical error, a finding we had anticipated because of the averaging done in three-dimensional interpolation and because the axial resolution of our tomograph is substantially worse than the in-plane resolution. We also note that the noise due to other sources of error was reduced, probably due to the averaging of the stereotactic transformation and perhaps as a result of the more systematic anatomic location of the ROI definitions.

DISCUSSION AND CONCLUSIONS

We have reviewed several methods for estimating the statistical error in the reconstruction of PET

scans. Our discussion was limited to noniterative techniques, but we note that other approaches, such as the EM method (Carson, 1986), may be used more widely in the future. Comparison of measured fluctuations with both Monte Carlo simulation studies and projection-based calculations on the RMS noise in ROIs show that the statistical errors can be computed with an accuracy of 10–20%. The source(s) of the remaining differences among the measurements, calculations, and simulations are relatively small and more refined experiments would be needed to understand them. The calculations using Huesman's (1977, 1984) method on the ROIs do not take into account the correlations induced in the projections when they are interpolated to give equal linear sampling. Another difficulty in interpreting the apparent differences between theory and measurement is the potential for systematic errors in the measured noise.

We conclude that both the direct calculation and Monte Carlo simulation are accurate enough for most practical applications. The example given in the text illustrates the decomposition of the random error at several stages of processing in a relatively complex visual activation study. It shows that the fluctuations due to noise propagation in the reconstructions of our regional averages contribute about half the variance. Extrapolating the results of this example should be done with caution as they depend on many details that may be unique to our laboratory and to properties of the tomograph used in our studies. Similar studies may be useful in designing experiments and/or optimizing experimental paradigms.

Acknowledgment: This research was supported in part by NINCDS Stroke Center grant number NS10828, USPH Training Grant CA09362, Schizophrenia Program Project Grant MH31154, and a National Science Foundation graduate fellowship for Andrew Gelman. The authors also wish to thank Donald Rubin for his helpful comments.

REFERENCES

Alpert N, Chesler D, Correia J, Ackerman R, Chang J, Finklestein S, Davis S, Brownell G, Taveras J (1982) Estimation of the local statistical noise in emission computed tomography. *IEEE Trans Med Imag* 1:142–146

Bergstrom M, Eriksson L, Bohm C, Blomqvist G, Litton J (1983) Corrections for scattered radiation in a ring detector positron camera by integral transformation of the projections. *J Comput Assist Tomogr* 7:42–50

Brownell G, Correia J, Zamenhof R (1979) Positron instrumentation. In: *Recent Advances in Nuclear Medicine, Vol. 5,* New York: Grune & Stratton, Inc., pp 13–20

Budinger T, Derenzo S, Gullberg G, Greenberg W, Huesman R (1977) Emission computer assisted tomography with single-photon and positron annihilation photon emitters. *J Comput Assist Tomogr* 1:131–145

Budinger T, Derenzo S, Greenberg W, Gullberg G, Huesman R (1978) Quantitative potentials of dynamic emission computed tomography. *J Nucl Med* 19:309–315

Carson R (1986) A maximum likelihood method for region-of-interest evaluation in emission tomography. *J Comput Assist Tomogr* 10:654–663

Huesman R (1977) The effects of a finite number of projection angles and finite lateral sampling of projections on the propagation of statistical errors in transverse section reconstruction. *Phys Med Biol* 22:511–521

Huesman R (1984) A new fast algorithm for the evaluation of regions of interest and statistical uncertainty in computed tomography. *Phys Med Biol* 29:543–552

Litton J, Bergstrom M, Eriksson L, Bohm C, Blomquist G (1984) Performance study of the PC-384 PET camera for the brain. *J Comput Assist Tomogr* 8:74–87

Palmer M, Bergstrom M, Beddoes M, Pate B (1985) Effects of detector wobble motion on image noise in positron emission tomography. *IEEE Trans Med Imag* 4:58–62

Palmer M, Bergstrom M, Pate B, Beddoes M (1986) Noise distribution due to emission and transmission statistics in positron emission tomography. *IEEE Trans Nucl Sci* 33:439–442

Senda M, Buxton RB, Alpert NM, Correia JA, Mackay BC, Weise SB, Ackerman RH (1988) The ^{15}O steady state method: correction for variation in arterial concentration. *J Cereb Blood Flow Metab* 8:681–690

Talairach J, Sziklag L (1967) *Atlas of Stereotactic Anatomy of the Telencephalon*. Paris, Masson and Cie

Tanaka E, Murayama H (1982) Properties of statistical noise in positron emission tomography. In: *Proceedings International Workshop on Physics and Engineering in Medical Imaging,* New York, IEEE, pp 158–164