

Detection of Serial Changes in Absolute Myocardial Perfusion with ^{82}Rb PET

Robert A. deKemp, Terrence D. Ruddy, Tanya Hewitt, Mary M. Dalipaj, and Rob S.B. Beanlands

Cardiac PET Centre, Division of Cardiology, University of Ottawa Heart Institute, Ottawa, Ontario, Canada

Serial changes in myocardial perfusion may represent an important marker of disease progression or regression or the effects of therapy for patients with coronary artery disease (CAD). Quantitative methods have not been developed for the assessment of serial changes in perfusion. The objective of this study was to use receiver operator characteristic (ROC) analysis to determine the sensitivity and specificity of direct paired comparisons (DPCs) to detect changes in absolute myocardial perfusion measured with ^{82}Rb PET. **Methods:** Repeated dynamic ^{82}Rb PET scans were obtained on 8 dogs at rest and during hyperemia induced with dobutamine ($n = 4$) or atrial pacing ($n = 4$). Radiolabeled microspheres were used to verify perfusion changes. Polar maps of absolute ^{82}Rb retention and associated SD were estimated from the dynamic images. Paired comparisons were then performed using a t test on each of the 532 polar map sectors. Rest–rest and stress–stress differences were used to assess specificity and reproducibility, and stress–rest differences were used to assess sensitivity. **Results:** ^{82}Rb retention differences of 20% over baseline were detected with 85%–90% sensitivity and specificity, using the optimal DPC probability value and image smoothness. The average ^{82}Rb retention differences correlated well with microspheres ($r = 0.74$; $P = 0.001$). Reproducibility of the mean retention values was $4.7\% \pm 2.1\%$. As reproducibility varies, the DPC probability value can be adjusted to maintain specificity. These ROC results are directly applicable to other image modalities that produce measurements with similar SEs ($3.7\% \pm 0.9\%$). **Conclusion:** The developed method of DPCs is sensitive and specific for the detection of changes in absolute myocardial perfusion measured with ^{82}Rb PET.

Key Words: perfusion; myocardial blood flow; ^{82}Rb ; PET; serial imaging

J Nucl Med 2000; 41:1426–1435

Myocardial perfusion is an important prognostic marker for patients with coronary artery disease (CAD). Serial changes in perfusion are used to assess the progression of CAD or manage the therapy for individual patients. PET and SPECT have been used to evaluate the effect of therapy on perfusion (1–6). In these research studies significant changes are shown between groups of patients but not in individual patients. In the clinical setting it would be useful to determine whether a given therapy is improving perfusion in

the individual patient. A sensitive method to detect individual changes over time would be useful to evaluate new and existing treatments of heart disease such as diet, exercise, and surgical and drug therapies.

Analysis of myocardial perfusion studies is often performed using a predefined template of myocardial regions corresponding to territories of the coronary arteries or their major branches. However, perfusion changes rarely occupy these template regions exactly, thereby limiting the sensitivity of region-based measurements to detect small changes over time. Manual drawing of regions can also be used, but this is a tedious and subjective process. Image-based approaches have been developed for the diagnosis of CAD by comparing relative perfusion measurements against normal database values (1,2,7–9). This diagnostic approach is used widely to quantify the extent and severity of CAD but may not be optimal for monitoring individual changes over time. Image-based methods for change detection in individual subjects has found widespread use in brain activation studies with PET and MRI (10–13) but has not been reported for the analysis of serial myocardial perfusion studies. We have developed an image-based method to detect and localize individual changes in myocardial perfusion by performing a direct paired comparison (DPC) of serial dynamic PET perfusion studies with ^{82}Rb (14). The purpose of this study was to validate the method and to assess the sensitivity and specificity of DPC to detect perfusion changes of varying magnitudes.

MATERIALS AND METHODS

Animal Surgical Preparation

Eight mongrel dogs weighing 26 ± 4 kg were studied. Animals were fasted overnight and premedicated with fentanyl. Anesthesia was induced with chloralose and supplemented with a constant infusion of fentanyl and inhalation of isoflurane. The dogs were then intubated and mechanically ventilated. In the left lateral recumbent position, a left fifth thoracotomy was performed, the lungs were retracted, and the pericardium was opened. A catheter was placed in the left atrial appendage for injection of radiolabeled microspheres and left atrial pressure monitoring. Arterial blood samples were withdrawn from the femoral artery during injection of microspheres and periodically throughout the experiment to monitor pH, PCO_2 , and PO_2 . Respiration was adjusted to maintain arterial blood gases within the normal physiologic range. Systolic arterial blood pressure was monitored using a catheter in the carotid artery. Intravenous accesses for the administration of ^{82}Rb , medica-

Received Jul. 14, 1999; revision accepted Feb. 16, 2000.

For correspondence or reprints contact: Robert A. deKemp, PhD, University of Ottawa Heart Institute, 40 Ruskin St., Ottawa, Ontario K1Y 4W7, Canada.

tions, and fluids were placed in the femoral, jugular, and cephalic veins. The study protocol was approved by the Animal Care Committee at the University of Ottawa.

Study Protocol

Repeated dynamic ^{82}Rb PET scans were obtained in 8 dog studies at rest and during stress. Dobutamine infusion or atrial pacing was used to increase myocardial perfusion for the stress condition. Increases in the heart rate times systolic blood pressure were used to estimate and adjust the corresponding perfusion increases. The first 4 dogs underwent 4 repeated scans at rest and 4 scans after dobutamine stress for a total of 8 scans each. The last 4 dogs underwent 4 repeated scans at rest and 4 scans at each of 3 different paced heart rates for a total of 16 scans each.

Dynamic ^{82}Rb PET Imaging

PET imaging was performed using a Siemens/CTI ART scanner (Siemens Medical Systems, Hoffman Estates, IL) (15,16). A 4-min transmission scan containing at least 100 million counts was acquired to perform attenuation correction. This was followed by 4 rapid rest or stress 10-min dynamic ^{82}Rb scans (15×6 s, 5×12 s, 30 s, 1 min, 2 min, 4 min) as described above. A custom infusion system was used to control the administered dose rate and to flush all ^{82}Rb activity out of the patient intravenous line at the end of the infusion. For each dynamic scan, 185 MBq (5 mCi) ^{82}Rb were infused over a 1-min period, and the acquisitions were started when the activity arrived in the scanner field of view. This infused dose rate was selected to limit the detector dead time to <50% so that the peak blood activity was measured accurately.

PET Image Analysis

Dynamic PET images were reconstructed using a Hann filter (0.6 cycle/cm) producing a reconstructed image resolution of ~12 mm. These dynamic images were reoriented automatically into short-axis sections (17). From the mean uptake images (2.5–10 min), maximum activity profiles were used to define the 3-dimensional shape of the left ventricular (LV) myocardium with combined cylindric and hemispheric (bottlebrush) sampling (18). The resulting set of 532 midmyocardial coordinates was used to resample the short-axis images into dynamic polar maps containing myocardial time-activity curves ($\text{Cm}(t)$ cps/g) for each of 532 sectors. The sector values correspond to midmyocardial voxels of ~4 mm³ depending on heart size and can be displayed as a conventional 2-dimensional polar plot or as a 3-dimensional surface. An arterial blood time-activity curve ($\text{Ca}(t)$ cps/mL) was obtained from the average of 4 image regions (1.4 cm² each) placed

automatically in the center of the LV cavity on 4 basal short-axis images (17). Polar maps of absolute ^{82}Rb retention were computed sector by sector using Equation 1 (19), where $\text{Cm}(T)$ is the average myocardial uptake from the last $n = 4$ dynamic frames (2.5–10 min), R_c is the partial volume recovery coefficient for the myocardial sectors estimated from phantom studies (0.7), and $\int \text{Ca}(t)dt$ is the integral of the arterial time-activity curve from 0 to 2 min. In addition to the retention value, the variability is also estimated for each sector using the sample SD of the last 4 dynamic frames. Therefore, for each dynamic ^{82}Rb PET scan, 2 polar maps are produced: the retention value and the associated SD ($\text{Ret} \pm \text{SD}$).

$$\text{Retention} = E \times F = \frac{\text{Cm}(T)/R_c}{\int \text{Ca}(t) dt} (\text{mL/min/g}). \quad \text{Eq. 1}$$

^{82}Rb retention is equal to the product of absolute myocardial perfusion (F) times the extraction fraction (E) of ^{82}Rb . Although absolute perfusion can be obtained by performing a nonlinear correction for extraction (20), this calculation was not performed because the increase in variability of the estimates would reduce the statistical significance of serial changes, particularly at high perfusion values.

Paired Comparisons

DPC identifies statistically significant differences between 2 serial perfusion scans (baseline and follow-up). Polar maps of the ^{82}Rb Ret and SD are created for the baseline ($\text{Ret1} \pm \text{SD1}$) and follow-up ($\text{Ret2} \pm \text{SD2}$) scans as illustrated in Figure 1. The statistical significance of the differences between the 2 scans is assessed by performing a 2-tailed, 2-sample t test (Eq. 2) on each polar map sector.

$$t\text{-statistics} = \frac{\text{Ret2} - \text{Ret1}}{\sqrt{(\text{SD1}^2 + \text{SD2}^2)/n}} \\ \sim t_{2n-2} \text{ (under the null hypothesis)}. \quad \text{Eq. 2}$$

The t -statistics polar map is displayed, and the frequency distribution of the 532 t -statistics is calculated. To compare 2 scans, the t -statistics threshold is set at a critical DPC probability value, α (e.g., $P < 0.05$). The number of significantly changed sectors is then calculated (N_{sig} increased or decreased) and expressed as a percentage (%LV) of the total number of sectors.

$$\%LV = N_{\text{sig}}/532 \times 100\% (P < \alpha). \quad \text{Eq. 3}$$

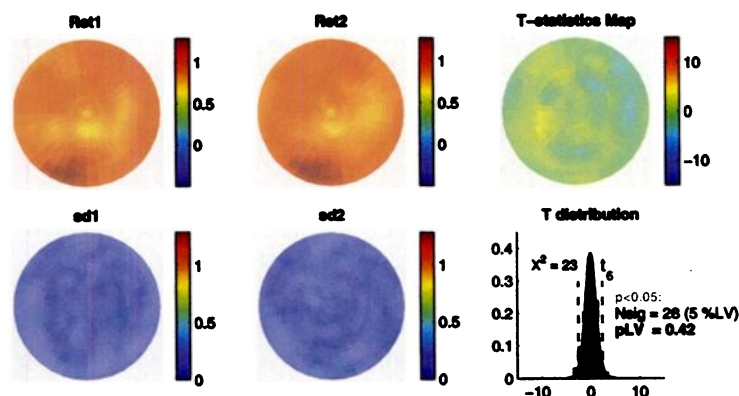


FIGURE 1. DPC of serial perfusion images (rest-rest). ^{82}Rb retention mean and SD polar maps are created at baseline ($\text{Ret1} \pm \text{SD1}$) and follow-up ($\text{Ret2} \pm \text{SD2}$). Differences are assessed using t -statistics map, with threshold set at desired DPC probability value to detect significant changes between baseline and follow-up scans. Frequency distribution of t -statistics from all 532 polar map sectors follows expected t distribution under null hypothesis ($\text{Ret1} - \text{Ret2} = 0$), which verifies that SD estimates are accurate. There is no significant global change in t -statistics map as a whole (probability of obtaining %LV changed sectors by chance alone [p_{LV}] = 0.42). N_{sig} = number of significantly changed sectors.

The probability $P(LV)$ of observing N_{sig} out of 532 sectors by chance is also estimated (10) and represents the significance of global changes in the t -statistics polar map as a whole.

For each stress–rest comparison, the underlying ^{82}Rb retention increase was computed as the mean percent difference (D) between the 2 polar maps:

$$D = \sum_{532} \frac{(\text{stress} - \text{rest}) / \text{rest}}{532} \times 100\%. \quad \text{Eq. 4}$$

For the rest–rest and stress–stress comparisons, the absolute value of the mean difference (Eq. 4) was also computed (21,22) and is referred to as reproducibility (R). Small values of R indicate good reproducibility.

Verifying SD Estimates

If the estimated sector SDs are accurate, then under the null hypothesis when there is no mean difference between 2 polar maps ($D = 0$), the calculated t -statistics should follow the t distribution with $2n - 2$ degrees of freedom (Eq. 2). If the SD estimates are not accurate, then the expected t distribution will not be obtained under the null hypothesis, violating the basic assumptions underlying the subsequent use of the t test. Therefore, it is important to first verify that the expected t distribution is obtained under the null hypothesis. Verification was performed by comparing the frequency distribution of the 532 calculated t -statistics against the expected t distribution with $2n - 2 = 6$ degrees of freedom. For each dog study the observed and expected distributions were compared using a χ^2 test for 1 pair of scans with $R \leq 1\%$ (where the null hypothesis was assumed to be true).

ROC Analysis

Receiver operator characteristic (ROC) analysis was used to determine the sensitivity and specificity of DPC to detect changes of a given magnitude (D).

Specificity ($1 - \text{false-positive rate}$) was assessed using paired comparisons of the rest–rest and stress–stress pairs. Ideally there should be no significant changes between these pairs of scans ($D = 0$, $N_{sig} = 0$). Therefore, any sectors that were detected as significantly changed were assumed to be false-positives (i.e., %LV from the rest–rest and stress–stress comparisons gives the false-positive rate).

Sensitivity (true-positive rate) was assessed using paired comparisons of the stress–rest pairs. Ideally all polar map sectors should be different between the rest and stress perfusion scans ($D > 0$, $N_{sig} = 532$). Therefore, all sectors that were detected as significantly changed were assumed to be true-positives (i.e., %LV from the stress–rest comparisons gives the true-positive rate).

The ROC curves were constructed by varying the DPC probability value from 1 to 0.000001 to vary the false-positive and true-positive rates described above. Each ROC curve represents the ability to detect global changes of a given magnitude equal to the average stress–rest difference (\bar{D}), as shown in Table 1 for the dog study 2. For each of the first 4 dog studies (4 rest scans and 4 dobutamine stress scans), 1 ROC curve was generated by plotting the average false-positive rate against the average true-positive rate. For each of the last 4 dogs, 3 ROC curves were generated corresponding to the average changes induced at the 3 paced heart rates. The area under each ROC curve ($n = 16$ total) was tabulated as a measure of overall performance.

Sensitivity and Specificity

Sensitivity is dependent primarily on the magnitude of change to be detected. To show the increasing ability to detect changes of increasing magnitude, the ROC curves were used to construct a plot of sensitivity against the average stress–rest difference (\bar{D}), where the measured specificity was equal to 90%.

Specificity is affected directly by the reproducibility of the measurements. As reproducibility is degraded, the distribution of t -statistics (Fig. 1) is shifted away from zero, causing the false-positive rate to increase above the theoretic value of $P \times 100\%$. To investigate the effect of varying reproducibility, the DPC probability value was plotted against the average reproducibility (\bar{R}) of the rest–rest and stress–stress comparisons, where the measured specificity was equal to 90% and 99%, respectively. To relate the effects of reproducibility back to sensitivity, the DPC probability value was also plotted against the average stress–rest difference (\bar{D}), where sensitivity was equal to 50%, 90%, and 99%.

Microspheres Analysis

The method of radiolabeled microspheres was used to verify the average stress–rest perfusion differences. Four microspheres were used in order of decreasing photon energy: ^{95}Nb , ^{46}Sc , ^{113}Sn , and ^{141}Ce . The first 4 dogs had microspheres injected before and after rest and then before and after dobutamine stress. The last 4 dogs had microspheres injected before the scans at rest and before the scans at each of the 3 paced heart rates. At the end of each experiment, dogs were euthanized by inducing ventricular fibrillation, and the hearts were removed. The LV myocardium and the arterial blood reference samples were assayed for the different microspheres to determine the average LV perfusion (F) at the corresponding times of injection (23). Briefly, myocardial perfusion was calculated as $F = Nm \times Qa/Na$ (mL/min/g), where Nm is the assayed myocardium counting rate (cpm/g), Qa is the blood sample withdrawal rate (mL/min), and Na is the total arterial blood sample counting rate (cpm).

RESULTS

Verifying SD Estimates

Comparisons of the observed and expected t distributions under the null hypothesis are shown in Table 2. These comparisons were performed using a DPC probability value of 0.05; therefore, if reproducibility was perfect (0%), then %LV would theoretically equal 5%. In this subgroup of cases in which the mean reproducibility was good ($R = 0.6\%$), the mean %LV was $4.4\% \pm 2.6\%$, which is close to the theoretic value of 5% ($P = 0.52$).

The observed and expected t distributions were not significantly different in 6 of the 8 cases ($P > 0.05$ after correction for multiple comparisons), indicating that the sector SD estimates are accurate in general. In 1 case, the observed distribution was slightly skewed (study 1), but the %LV was very close to the theoretic value of 5% ($P(LV) = 0.49$), indicating that there was no global change in perfusion. In the other case with a significant χ^2 value (study 3), the observed distribution was wider than expected and had a higher false-positive rate (%LV = 9.6%; $P(LV) > 0.01$) than expected. On close examination, it was found that there had been marked fluid loading between the 2 scans, producing regional changes in ^{82}Rb retention likely caused by inaccur-

TABLE 1
ROC Analysis for Dog Study 2

| Specificity comparison | R (%) | Nsig | | | | | |
|-------------------------|-------|----------------|----------------|-----------------|------------------|-------------------|--------------------|
| | | <i>P</i> < 1.0 | <i>P</i> < 0.1 | <i>P</i> < 0.01 | <i>P</i> < 0.001 | <i>P</i> < 0.0001 | <i>P</i> < 0.00001 |
| Rest1–rest2 | 5.7 | 532 | 169 | 44 | 9 | 0 | 0 |
| Rest1–rest3 | 6.3 | 532 | 236 | 60 | 10 | 0 | 0 |
| Rest1–rest4 | 8.3 | 532 | 265 | 79 | 14 | 2 | 0 |
| Rest2–rest3 | 0.7 | 532 | 36 | 4 | 1 | 0 | 0 |
| Rest2–rest4 | 2.8 | 532 | 55 | 6 | 1 | 0 | 0 |
| Rest3–rest4 | 2.1 | 532 | 41 | 3 | 0 | 0 | 0 |
| Stress1–stress2 | 1.1 | 532 | 33 | 0 | 0 | 0 | 0 |
| Stress1–stress3 | 1.2 | 532 | 85 | 16 | 3 | 0 | 0 |
| Stress1–stress4 | 8.8 | 532 | 165 | 34 | 6 | 0 | 0 |
| Stress2–stress3 | 0.1 | 532 | 67 | 6 | 0 | 0 | 0 |
| Stress2–stress4 | 7.6 | 532 | 166 | 43 | 10 | 2 | 0 |
| Stress3–stress4 | 7.5 | 532 | 199 | 36 | 2 | 0 | 0 |
| Average (\bar{R}) | 4.3 | 532 | 126.4 | 27.6 | 4.7 | 0.3 | 0 |
| False-positive rate (%) | | 100 | 23.8 | 5.2 | 0.9 | 0.1 | 0 |

| Sensitivity comparison | D (%) | Nsig | | | | | |
|------------------------|-------|----------------|----------------|-----------------|------------------|-------------------|--------------------|
| | | <i>P</i> < 1.0 | <i>P</i> < 0.1 | <i>P</i> < 0.01 | <i>P</i> < 0.001 | <i>P</i> < 0.0001 | <i>P</i> < 0.00001 |
| Stress1–rest1 | 31 | 532 | 509 | 398 | 245 | 71 | 13 |
| Stress2–rest1 | 32 | 532 | 519 | 423 | 248 | 92 | 15 |
| Stress3–rest1 | 32 | 532 | 503 | 407 | 247 | 95 | 22 |
| Stress4–rest1 | 42 | 532 | 526 | 468 | 348 | 148 | 39 |
| Stress1–rest2 | 39 | 532 | 507 | 435 | 286 | 132 | 30 |
| Stress2–rest2 | 40 | 532 | 514 | 454 | 303 | 138 | 34 |
| Stress3–rest2 | 40 | 532 | 504 | 425 | 304 | 151 | 38 |
| Stress4–rest2 | 51 | 532 | 526 | 476 | 361 | 203 | 67 |
| Stress1–rest3 | 40 | 532 | 525 | 460 | 321 | 160 | 64 |
| Stress2–rest3 | 41 | 532 | 527 | 487 | 341 | 173 | 66 |
| Stress3–rest3 | 41 | 532 | 512 | 475 | 343 | 167 | 55 |
| Stress4–rest3 | 52 | 532 | 531 | 497 | 404 | 239 | 89 |
| Stress1–rest4 | 43 | 532 | 527 | 458 | 311 | 148 | 53 |
| Stress2–rest4 | 44 | 532 | 530 | 465 | 326 | 162 | 67 |
| Stress3–rest4 | 44 | 532 | 512 | 462 | 331 | 167 | 62 |
| Stress4–rest4 | 55 | 532 | 532 | 497 | 395 | 230 | 92 |
| Average (\bar{D}) | 42 | 532 | 519 | 455 | 320 | 155 | 50 |
| True-positive rate (%) | | 100 | 97.5 | 85.5 | 60.1 | 29.1 | 9.4 |

Corresponding ROC curve is plotted in Figure 2 (\bar{D} = 42%).

rate attenuation correction. The observed wider distribution can be attributed to these regional differences. This illustrates the importance of accurate registration between the transmission and emission scans and also shows that the paired comparison technique is sensitive to this type of data acquisition error.

ROC Analysis

Sixteen ROC curves were generated from the experiments in the 8 dogs. Table 3 shows the average stress–rest differences in ^{82}Rb retention (\bar{D}) for each of the 16 comparisons and the total ROC area as a measure of overall performance. Differences of 100% or greater were detected with nearly perfect performance (>99% ROC area). The performance for changes below 5% was poor (i.e., similar to that obtained by chance alone [50%]). The retention differ-

ences (\bar{D}) were well correlated with the corresponding perfusion differences (ΔF) determined by microspheres ($r = 0.74$; $P = 0.001$). The fitted linear relation ($\bar{D} = 0.71 \times \Delta F + 4\%$) did not differ significantly from the line of identity (slope $P = 0.1$; intercept $P = 0.76$), although a slope of <1 is consistent with decreasing extraction of ^{82}Rb at higher flows (20).

Four representative ROC curves are shown in Figure 2 for average differences of 12%, 20%, 42%, and 96%. The ROC curves for smaller changes are closer to the diagonal line (chance performance), indicating a reduced ability to detect changes of smaller magnitude.

Sensitivity

Sensitivity is shown in Figure 3 as a function of the average stress–rest difference (\bar{D}) for 90% specificity. Larger

TABLE 2
Paired Comparison of Scans with Best Reproducibility
($R \leq 1\%$)

| Dog study | R (%) | %LV ($P < 0.05$) | $P(LV)$ | χ^2_{31} |
|---------------|-------|--------------------|---------|---------------|
| 1 | 0.9 | 5.1 | 0.49 | 146* |
| 2 | 0.7 | 2.8 | 0.99 | 50 |
| 3 | 0.4 | 9.6 | <0.01 | 114* |
| 4 | 0.4 | 1.9 | 1.00 | 24 |
| 5 | 0.7 | 4.1 | 0.84 | 33 |
| 6 | 0.4 | 2.1 | 1.00 | 42 |
| 7 | 1.0 | 6.2 | 0.13 | 23 |
| 8 | 0.5 | 3.0 | 0.99 | 42 |
| 0.6 ± 0.2 | | 4.4 ± 2.6 | | |

* $P < 0.05$.

R = mean rest–rest or stress–stress difference (see Eq. 4);
 $P(LV)$ = probability of obtaining %LV changed sectors by chance alone; χ^2 statistic compares observed and expected distribution of t -statistics.

differences are detected with higher sensitivity as expected, and small differences of 15%–20% are still detected with ~50% sensitivity. A sensitivity value of 50% means that in a region of myocardium with a uniform increase in perfusion, half of the sectors in the region would be detected as significantly changed. In practice, these sectors tend to cluster at the center of the region, so sensitivity is related to a reduction in size of the detected regions.

The sensitivity to detect significant changes is also determined by the SE of the measurements (Eq. 2). In this study, the average SE of the sector values of ^{82}Rb retention was $3.7\% \pm 0.9\%$.

Specificity

Specificity was found to be a function of reproducibility. The results in Table 2 show that the theoretic specificity is obtained when comparing scans measured with nearly perfect reproducibility ($R \leq 1\%$). However, in practice,

TABLE 3
ROC Performance of DPCs

| Dog study | Comparison | \bar{D} (%) | ROC area (%) |
|-----------|-------------|---------------|--------------|
| 1 | Stress–rest | 147 | 99 |
| 2 | Stress–rest | 42 | 95 |
| 3 | Stress–rest | 74 | 97 |
| 4 | Stress–rest | 17 | 78 |
| 5 | Pace1–rest | 8 | 67 |
| 5 | Pace2–rest | 12 | 68 |
| 5 | Pace3–rest | 18 | 78 |
| 6 | Pace1–rest | 1 | 50 |
| 6 | Pace2–rest | 2 | 51 |
| 6 | Pace3–rest | 5 | 53 |
| 7 | Pace1–rest | 2 | 50 |
| 7 | Pace2–rest | 19 | 80 |
| 7 | Pace3–rest | 20 | 83 |
| 8 | Pace1–rest | 32 | 85 |
| 8 | Pace2–rest | 96 | 99 |
| 8 | Pace3–rest | 131 | 99 |

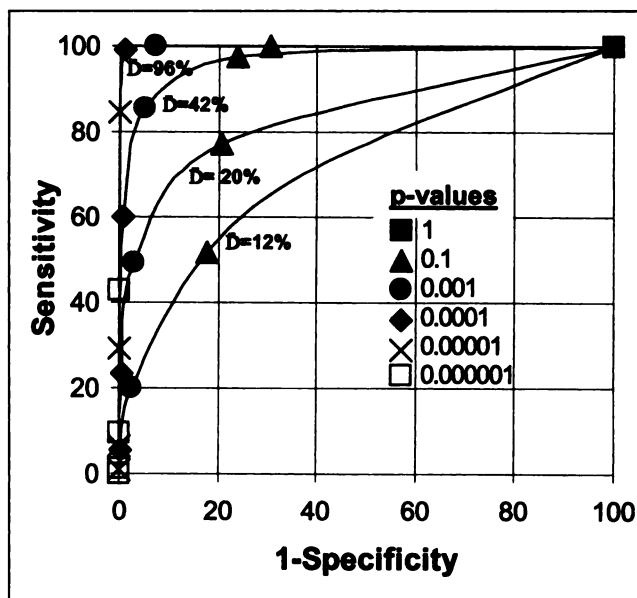


FIGURE 2. ROC curves from 4 dog studies with average stress–rest retention differences (\bar{D}) = 12%, 20%, 42%, and 96%. DPC probability values used to construct curves are shown.

reproducibility is not perfect, and this lowers the actual specificity observed with DPC. For example, in Figure 2, the points corresponding to $P = 0.1$ have actual false-positive rates of 20%–30%, which is greater than the theoretic value of 10%. This effect is summarized in Figure 4, which shows that specificity can be maintained by lowering the DPC probability value to exclude the false-positive values associated with poor reproducibility. For comparison, the mean

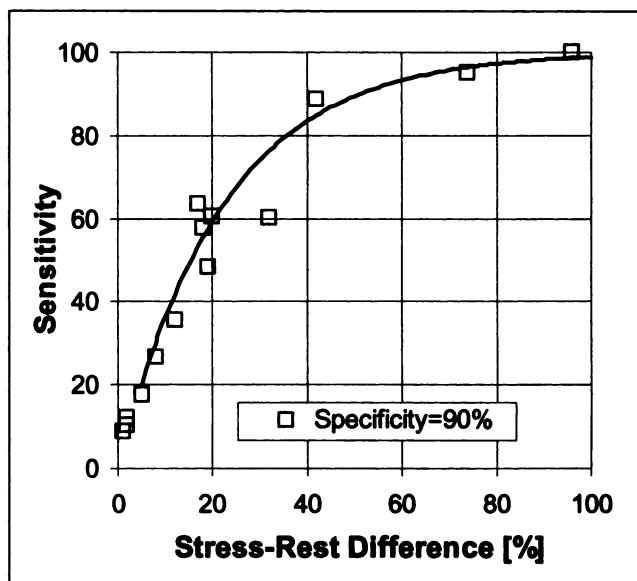


FIGURE 3. Sensitivity values derived from all ROC curves where specificity = 90%. Sensitivity improves as magnitude of change (\bar{D}) increases, as expected. Sensitivity is 100% for stress–rest differences of 100% or more. Point at 32% appears as outlier because of poor reproducibility.

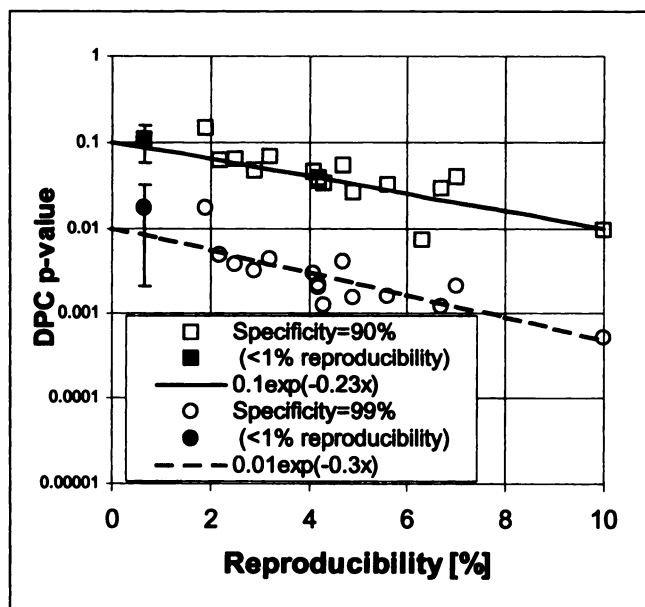


FIGURE 4. DPC probability values derived from all ROC curves where specificities = 90% and 99%. Probability values are a function of study reproducibility (average rest–rest and stress–stress differences). To maintain desired specificity, DPC probability values can be lowered on basis of known or assumed reproducibility.

DPC probability value for the cases with $R \leq 1\%$ is also plotted. In this subgroup, the measured values are not significantly different ($P > 0.1$) from the theoretic values of 0.1 and 0.01 for 90% and 99% specificity, respectively.

Figure 4 can be used to determine what DPC probability value should be used to maintain the desired specificity as a function of reproducibility. For example, the mean reproducibility for all studies was $4.7\% \pm 2.1\%$, and the DPC probability value needed to achieve 90% specificity in this case would be $P = 0.034$ instead of the theoretic value of $P = 0.1$.

To evaluate serial perfusion changes in individual patients, the minimum detectable changes can then be determined from Figure 5, which plots the DPC probability value against the stress–rest differences (\bar{D}) for sensitivity values of 99%, 90%, and 50%. For the example above, with a DPC probability value set at $P = 0.034$ (for 90% specificity), the technique would be 50% sensitive to changes of 18% and 90% sensitive to changes of 45% or larger.

The best possible results (with perfect reproducibility) can be obtained directly from Figure 5. For example, if 90% sensitivity and specificity are desired, then the detectable difference is 36%, corresponding to the point where the 90% sensitivity line crosses $P = 0.1$.

The effect of reproducibility is also illustrated in study 8 ($\bar{D} = 32\%$), which has the largest reproducibility value of 10%. From Figure 4, the DPC probability value needed for 90% specificity in this case is $P = 0.01$, which is $>50\%$ sensitive to changes of 32% (from Fig. 5). In Figure 3, the corresponding point appeared as an outlier because it had

poorer reproducibility than did the other studies. For the analysis or design of new studies, Figures 4 and 5 can be used to determine the sensitivity and specificity to detect changes as a function of the expected reproducibility.

Improving ROC Performance

Most of the false-positive changes found in the specificity comparisons had a magnitude (D) that was small (5%–10%). Specificity could be improved by applying a constraint on the magnitude of the change in addition to the statistical significance. On a plot of the t -statistic against D (Fig. 6), the optimal discriminator of the specificity comparisons (rest–rest) from the sensitivity comparisons (stress–rest) should also include a component of D . In fact, an added magnitude constraint of $D > t_\alpha \times 10$ improves specificity markedly in the range of $P = 0.001$ –1.0 as shown in Figure 7. The ROC curve for $\bar{D} = 20\%$ (study 7) is reproduced from Figure 2. With this D -constraint added, specificity is typically improved from 90% to 95% (with constant or improved sensitivity) (i.e., the 90% specificity curve in Fig. 4 increases to $\sim 95\%$, and the DPC probability values can be set 8 times higher). Alternatively, the effect can be viewed on Figure 7 as improving sensitivity by 10%–20% in the 90%–99% specificity range, which is true even for small changes of 5%–10%.

Increased image smoothing also improves the ROC performance as shown in Figure 7. If the images from the same study ($\bar{D} = 20\%$) are reconstructed with 20-mm instead of 12-mm resolution, the average SE decreases from 3.7% to 1.8%, and sensitivity improves an additional 10%. With the typical image resolution of 20 mm used clinically with ^{82}Rb , 85%–90% sensitivity and specificity are achieved for the detection of 20% changes.

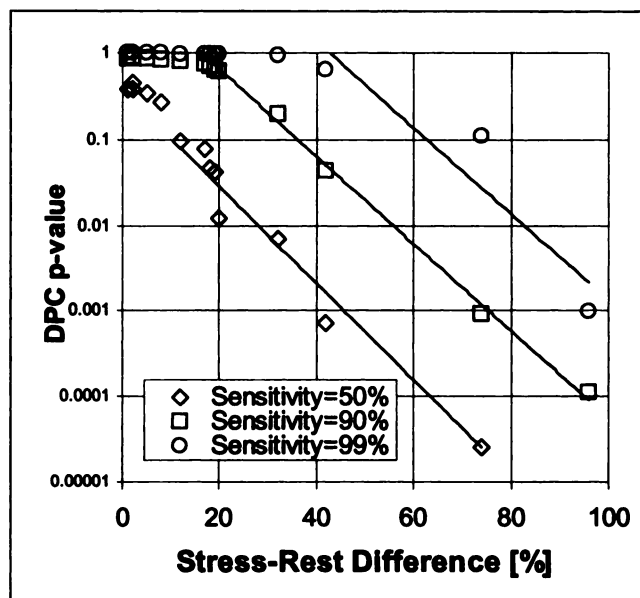


FIGURE 5. DPC probability values derived from all ROC curves where sensitivities = 50%, 90%, and 99%. Given DPC probability value has improved sensitivity for changes of increasing magnitude (stress–rest difference). This relation is independent of effects of reproducibility shown in Figure 4.

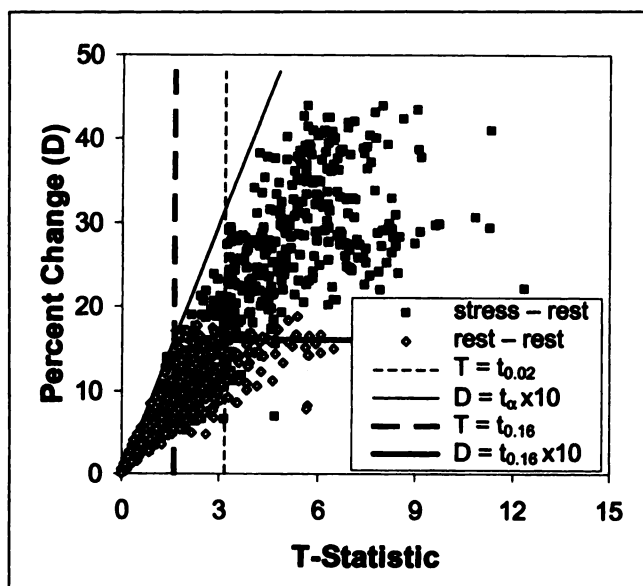


FIGURE 6. Plot of sector t -statistics versus magnitude of change (D) from 1 rest-rest and 1 stress-rest comparison of study 7 ($\bar{D} = 20\%$). Added constraint on magnitude of change improves specificity by reducing number of false-positive rest-rest changes. For example, specificity is 89% for $|t| > t_{0.02}$, which increases to 95% when D-constraint is added ($D > t_{0.16} \times 10$ and $|t| > t_{0.16}$). Sensitivity remains constant at 69%.

DISCUSSION

The performance of DPCs has been characterized for the detection of serial myocardial perfusion changes measured with absolute ^{82}Rb retention and PET. DPC requires accurate SD estimates, which can be verified against the theoretic t distribution when there is no mean difference between 2 images. It assumes that the temporal changes in normal heterogeneity of perfusion are small relative to the measurement variability (24). The sample SD is a simple and robust estimator of the measurement variability and was shown to be accurate using this methodology.

A novel method has been developed to correct the DPC probability values for reproducibility to maintain the desired specificity. To detect perfusion changes between serial clinical scans in patients, a mean reproducibility value should be determined to select the appropriate correction from Figure 4. For example, reproducibility may be worse if the serial scans are obtained on separate days (21,22) compared with this study, in which reproducibility was determined using rapid serial scans on the same day. The sensitivity results (Fig. 5) are directly applicable to the results of other studies with SEs in the same range as in this study. Clinical ^{82}Rb studies performed in our institution have average SEs of $\sim 5\% \pm 1\%$; therefore, the sensitivity to detect significant changes should be similar to that shown in Figure 5.

Previous studies evaluating medical therapies for CAD have reported changes in the extent or severity of relative perfusion defects across groups of subjects (1,2). These methods are based on normal population databases, which

do not assess the statistical significance of perfusion changes in individuals nor do they assess increases and decreases independently as performed in this study. Other studies have reported changes in absolute perfusion after medical (3–5) or interventional (6) therapy indicating that these treatments improve perfusion across a group of subjects but do not determine the effects for a given individual as developed in this study.

DPC detects significant differences in serial myocardial perfusion studies by creating polar map images of the t -statistic using individual subjects as their own controls. This is analogous to the method of statistical parametric mapping (SPM), which has found widespread application for the analysis of brain activation studies with PET and MRI (10–12). SPM uses an analysis of covariance to detect adjusted differences relative to the whole-brain mean perfusion, whereas the 2-sample t test used by DPC is equivalent to a 1-way ANOVA of absolute perfusion measurements. In the single-subject SPM paradigm (13), repeated scans are used to assess the measurement variability, which therefore includes the effect of reproducibility. DPC uses sequential frames within 1 dynamic scan to assess the measurement variability, but if repeated scans were performed as part of the experimental design, reproducibility could then be included as an additional factor in the ANOVA. DPC can also be performed on adjusted differences analogous to SPM, simply by subtracting the LV mean from each sector value before t testing. Adjusting for the mean does not change the sector SDs, but it does improve the reproducibility (hence, specificity) to essentially perfect values by setting the global difference equal to zero as in Figure 1.

The DPC probability values used to construct the ROC

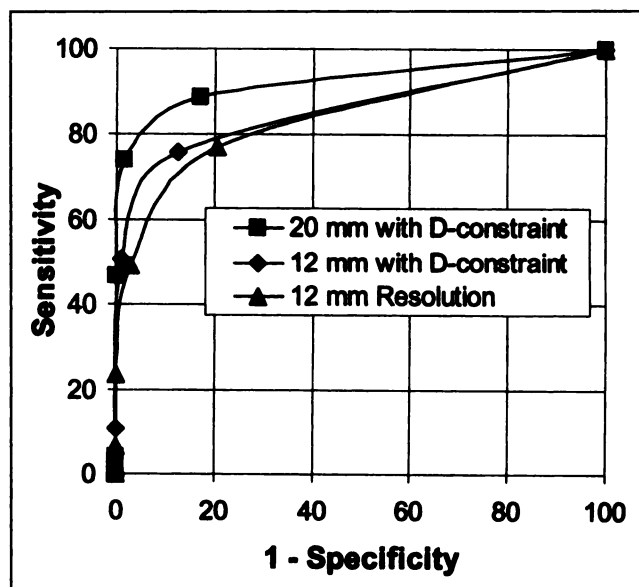


FIGURE 7. ROC curves from study 7 ($\bar{D} = 20\%$). ROC performance is improved by adding D-constraint on magnitude of change and also by altering reconstructed image resolution from 12 to 20 mm.

curves were not corrected for multiple comparisons. In this case, only the total number of detected differences (%LV) are used to report the significance $P(LV)$ of global changes in the polar map as a whole (10) as shown in Figure 1. To report individual regions of significant change, the probability value must be corrected for the number of independent comparisons, which is a function of image smoothness (11). The number of independent regions in an image can be estimated by the number of resolution elements, which is a function of the reconstructed image resolution and the pixel size (25). The polar map images used in this study had a reconstructed resolution of 12 mm and a sector size of 4 mm, which corresponds to ~60 independent regions in a polar map with 532 sectors. Therefore, for a desired specificity, the DPC probability value must be divided by 60 if individual regions of change are reported. This Bonferroni correction was verified simply by measuring the false-positive rate of regions detected (9 connected sectors) with the corrected DPC probability value. For a relatively poor reproducibility of 12%, the DPC probability value needed for 90% specificity is $P = 0.0063$ (Fig. 4); dividing by 60 gives a corrected value of $P = 0.0001$. Using all possible comparisons (from all studies) with a difference near 12%, only 2 of 16 comparisons (13%) had a significant region detected, which is close to the expected value of 10% for 90% specificity. Similarly, for a relatively good reproducibility of 2%, the corrected DPC probability value is $P = 0.001$, and the observed false-positive rate was 1 of 34 (3% versus 10% expected).

As mentioned above, the Bonferroni correction changes with the degree of image smoothing, and it limits the portions of the ROC curve that can be achieved in practice. For example, images with a reconstructed resolution of 20 mm would use a correction factor of only 20, as opposed to a factor of 60 required for images with 12-mm resolution. Increased image smoothing improves the ROC performance to detect global changes (Fig. 6) and lowers the required Bonferroni correction but degrades the spatial resolution to detect small regional changes. Conversely, images reconstructed with higher resolution (less smoothing) can detect smaller sized regions of change but with poorer ROC performance.

For clinical application, the desired specificity is set at 95% ($P < 0.05$), a default reproducibility of 5% is assumed, and the additional magnitude D-constraint is used to perform the paired comparisons. These settings yield a DPC probability value of $P = 0.13$ used to set the threshold of the t -statistics and assess the overall significance of the global changes $P(LV)$. To detect individual regional changes, the DPC probability value is corrected to $P = 0.0065$ for images with a reconstructed resolution of 20 mm as discussed above. An example is illustrated in Figure 8, which shows serial studies in a patient with a coronary stent inserted as primary therapy for occlusion of the left anterior descending (LAD) artery. There is significantly increased perfusion to the apex in the distribution of the distal LAD artery at 6 wk versus 5 d after revascularization. The perfusion defect in the lateral wall is unchanged, consistent with known disease in the left circumflex artery.

PET perfusion imaging has been used to quantify regional perfusion in patients with CAD or global perfusion in patients with hypertrophy, syndrome X, and cardiac transplantation (3–6,26–30). DPC provides a sensitive method for detecting serial changes in regional and global perfusion in these conditions. This approach may be useful for monitoring the progression of disease and provide a guide for therapy modification. For example, changes in the degree of coronary stenosis on angiography with cholesterol-lowering therapy may be small but represent important modification in plaque stability that reduces cardiac events. The change in functional significance of the stenosis could be monitored with serial perfusion imaging and DPC analysis. Therapy could be adjusted to more aggressive drug or interventional therapy should the desired improvement not occur. The DPC method also provides an important research tool to evaluate the effects of new therapies on perfusion (and other parameters measured by PET). The relevant clinical difference would be preset, and the hypothesis that the therapy results in this degree of change would be tested.

The methods developed in this study can also be used to generalize the application of DPC to compare serial images acquired with other PET tracers or with other imaging modalities. PET ^{13}N -ammonia retention images could be analyzed using the identical methodology. The extraction of

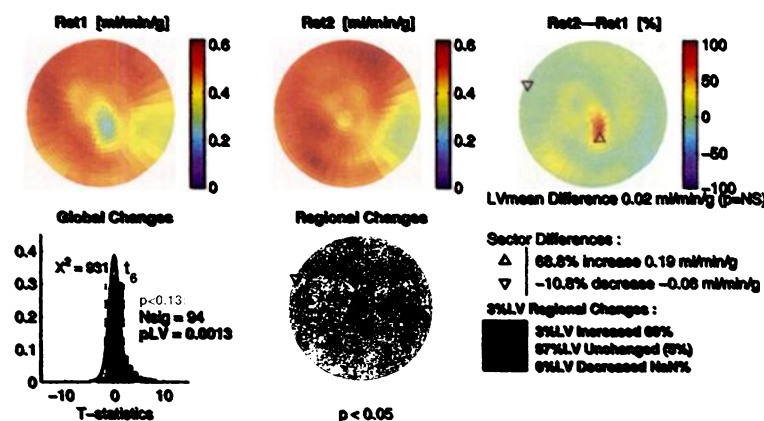


FIGURE 8. DPC of serial ^{82}Rb perfusion studies in 91-y-old man after coronary stent revascularization of left anterior descending (LAD) artery. Polar map images show perfusion to LV myocardium at 5 d (Ret1) and 6 wk (Ret2) after revascularization as well as percent changes (Ret2 - Ret1). Frequency distribution of t -statistics from all 532 polar map sectors (Global Changes) indicates significant change in LV as a whole (probability of obtaining %LV changed sectors by chance alone [p_{LV}] = 0.0013). t -statistics map set at threshold (Regional Changes) shows significantly increased perfusion (red) to apex in distribution of distal LAD artery. Perfusion defect in lateral wall is unchanged (green), consistent with known disease in left circumflex artery. NS = not significant.

ammonia is substantially higher than that of rubidium over the entire physiologic flow range (20), which would produce larger differences in retention for a given change in absolute perfusion. This would be advantageous, particularly for detecting small changes under stress (hyperemic) conditions. This study assessed the ability of DPC to detect changes relative to resting values only.

DPC can also be used to detect serial changes in kinetic rate constants estimated by compartmental or graphical modeling. In this case it is essential to obtain (and verify) accurate SD estimates of the model parameters (31–33). Compartmental modeling can also have the advantage of including resolution recovery (34). In this study a constant recovery coefficient was used with the retention model, but this does not affect the ability to detect individual serial changes if the myocardial wall thickness also remains constant.

Finally, the DPC approach has the potential to be applied to other imaging modalities, including SPECT, to improve sensitivity for detecting serial changes with these modalities. Further studies are necessary to evaluate this approach for gated imaging and relative perfusion SPECT.

CONCLUSION

DPC detects small global and regional changes in serial myocardial perfusion images. Significant differences are displayed as polar maps of the *t*-statistic, which have thresholds set at the appropriate DPC probability value. Specificity is determined by the probability value, which is corrected for multiple comparisons and for reproducibility to produce the desired false-positive rate of detected image regions. Sensitivity is a function of the SE of the measurements and the magnitude of the changes in perfusion.

The DPC method developed in this study can assess the significance of global changes in perfusion associated with diffuse myocardial disease and can assess regional changes associated with therapy or progression of CAD in individual patients.

ACKNOWLEDGMENTS

The authors thank May Aung for technical assistance in obtaining the PET scans and Richard Seymour for technical assistance from the Animal Care and Veterinary Service. This work was supported by grant MA-14319 and personnel awards from the Medical Research Council of Canada and by the Heart and Stroke Foundation of Canada.

REFERENCES

- Mahmari J, Fenimore NL, Marks JF, et al. Transdermal nitroglycerin patch therapy reduces the extent of exercise-induced myocardial ischemia: results of a double-blind, placebo-controlled trial using quantitative thallium-201 tomography. *J Am Coll Cardiol*. 1994;24:25–32.
- Gould KL, Martucci JP, Goldberg DI, et al. Short-term cholesterol lowering decreases size and severity of perfusion abnormalities by positron emission tomography after dipyridamole in patients with coronary artery disease: a potential noninvasive marker of healing coronary endothelium. *J Nucl Med*. 1994;35:1530–1538.
- Fallen EL, Nahmias C, Scheffel A, et al. Redistribution of myocardial blood flow with topical nitroglycerin in patients with coronary artery disease. *Circulation*. 1995;91:1381–1388.
- Guethlin M, Kasel AM, Coppenrath K, et al. Delayed response of myocardial flow reserve to lipid-lowering therapy with fluvastatin. *Circulation*. 1999;99:475–481.
- Henes CG, Bergmann SR, Perez JE, et al. The time course of restoration of nutritive perfusion, myocardial oxygen consumption, and regional function after coronary thrombolysis. *Coron Artery Dis*. 1990;1:687–696.
- Walsh M, Geltman E, Steele R, et al. Augmented myocardial perfusion reserve after coronary angioplasty quantified by positron emission tomography with $H_2^{15}O$. *J Am Coll Cardiol*. 1990;15:119–127.
- Hicks K, Ganti G, Mullani N, Gould KL. Automated quantitation of three-dimensional cardiac positron emission tomography for routine clinical use. *J Nucl Med*. 1989;30:1787–1797.
- van Train KF, Maddahi J, Berman DS, et al. Quantitative analysis of tomographic stress thallium-201 myocardial scintigrams: a multicenter trial. *J Nucl Med*. 1990;31:1168–1179.
- van Train KF, Garcia EV, Maddahi J, et al. Multicenter trial validation for quantitative analysis of same-day rest-stress technetium-99m-sestamibi myocardial tomograms. *J Nucl Med*. 1994;35:609–618.
- Friston KJ, Frith CD, Liddle PF, et al. The relationship between global and local changes in PET scans. *J Cereb Blood Flow Metab*. 1990;10:458–466.
- Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ. Comparing functional PET images: the assessment of significant change. *J Cereb Blood Flow Metab*. 1991;11:690–699.
- Watson JDG, Meyers R, Frackowiak RSJ, et al. Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb Cortex*. 1993;3:79–94.
- Silbersweig DA, Stern E, Frith CD, et al. Detection of thirty-second cognitive activations in single subjects with positron emission tomography: a new low-dose $H_2^{15}O$ regional cerebral blood flow three-dimensional imaging technique. *J Cereb Blood Flow Metab*. 1993;13:617–629.
- deKemp RA, Ruddy TD, Hewitt T, et al. Monitoring the response to therapy using direct paired comparison of serial myocardial perfusion images [abstract]. *J Nucl Med*. 1999;40:161P.
- Bailey D, Young H, Bloomfield PM, et al. ECAT ART: a continuously rotating PET camera—performance characteristics, initial clinical studies, and installation considerations in a nuclear medicine department. *Eur J Nucl Med*. 1997;24:6–15.
- Watson CC, Newport D, Casey ME, deKemp RA, Beanlands RS. Evaluation of simulation-based scatter correction for 3-D PET cardiac imaging. *IEEE Trans Nucl Sci*. 1997;44:90–97.
- deKemp RA, Nahmias C. Automated determination of the left ventricular long axis in cardiac positron tomography. *Physiol Meas*. 1996;17:95–108.
- Laubenbacher C, Rothley J, Sitomer J, et al. An automated analysis program for the evaluation of cardiac PET studies: initial results in the detection and localization of coronary artery disease using nitrogen-13-ammonia. *J Nucl Med*. 1993;34:968–978.
- Bellina CR, Parodi O, Camici P, et al. Simultaneous in vitro and in vivo validation of nitrogen-13-ammonia for the assessment of regional myocardial blood flow. *J Nucl Med*. 1990;31:1335–1343.
- Yoshida K, Mullani N, Gould KL. Coronary flow and flow reserve by PET simplified for clinical applications using rubidium-82 or nitrogen-13-ammonia. *J Nucl Med*. 1996;37:1701–1712.
- Nagamachi W, Czernin J, Kim AS, et al. Reproducibility of measurements of regional resting and hyperemic myocardial blood flow assessed with PET. *J Nucl Med*. 1996;37:1626–1631.
- Beauchesne LM, Ruddy TD, deKemp RA, et al. Reproducibility of myocardial perfusion measurements with rubidium-82 PET imaging [abstract]. *Circulation*. 1998;98:1-222.
- Heymann MA, Payne BD, Hoffman JIE, Rudolf AM. Blood flow measurements with radionuclide-labeled particles. *Progr Cardiovasc Dis*. 1977;20:55–79.
- King RB, Bassingthwaite JB, Hales JRS, Rowell LB. Stability of heterogeneity of myocardial blood flow in normal awake baboons. *Circ Res*. 1985;57:285–295.
- Budinger TF, Derenzo SE, Greenberg WL, Gullberg GT, Huesman RH. Quantitative potentials of dynamic emission computed tomography. *J Nucl Med*. 1978;19:309–315.
- Demer LL, Gould KL, Goldstein RA, et al. Assessment of coronary artery disease severity by positron emission tomography: comparison with quantitative arteriography in 193 patients. *Circulation*. 1989;79:825–835.
- Beanlands R, Muzik O, Melon P, et al. Noninvasive quantification of regional myocardial blood flow reserve in stenosed and angiographically normal vessels of patients with coronary atherosclerosis. *J Am Coll Cardiol*. 1995;26:1465–1475.

28. Krivokapich J, Stevenson LW, Kobashigawa J, Huang SC, Schelbert HR. Quantification of absolute myocardial perfusion at rest and during exercise with positron emission tomography after human cardiac transplantation. *J Am Coll Cardiol.* 1991;18:512-517.
29. Camici P, Chiriac G, Lorezoni R, et al. Coronary vasodilation is impaired in both hypertrophied and non-hypertrophied myocardium of patients with hypertrophic cardiomyopathy. *J Am Coll Cardiol.* 1991;17:879-886.
30. Geltman EM, Hennes GC, Senneff MJ, Sobel BE, Bergmann SR. Increased myocardial perfusion at rest and diminished perfusion reserve in patients with angina and angiographically normal coronary arteries. *J Am Coll Cardiol.* 1990;16:586-595.
31. Landaw EM, DiStefano JJ. Multiexponential, multicompartmental, and noncompartmental modeling. II. Data analysis and statistical considerations. *Am J Physiol.* 1984;246:R665-R677.
32. Coxson PG, Brennan KM, Huesman RH, Lim S, Budinger TF. Variability and reproducibility of rubidium-82 kinetic parameters in the myocardium of the anesthetized canine. *J Nucl Med.* 1995;36:287-296.
33. Huesman RH, Klein GJ, Reutter BW, et al. Strategies for extraction of quantitative data from volumetric dynamic cardiac positron emission tomography data. *Cardiology.* 1997;88:54-61.
34. Hutchins GD, Caraher JM, Raylman RR. A region of interest strategy for minimizing resolution distortions in quantitative myocardial PET studies. *J Nucl Med.* 1992;33:1243-1250.