Journal of Zhejiang University SCIENCE B
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Activity optimization method in SPECT: A comparison with ROC analysis*

DÍAZ Marlén Pérez†[1], RIZO Oscar Díaz[2], DÍAZ Adlin López[3],
APARICIO Eric Estévez[4], DÍAZ Reinaldo Roque[4]

([1]*Center for Studies on Electronics & Info Tech. Images Group, Central University of Las Villas, Santa Clara 54830, Cuba*)
([2]*Department of Nuclear Physics, Institute for Sciences and Advanced Technology, Havana 10400, Cuba*)
([3]*Department of Nuclear Medicine, Cardiovascular Surgery Institute, Havana 10400, Cuba*)
([4]*Department of Nuclear Medicine, University Hospital Celestino Hernández, Santa Clara 50200, Cuba*)
†E-mail: mperez@uclv.edu.cu

Received Aug. 11, 2006; revision accepted Sept. 14, 2006

Abstract: A discriminant method for optimizing activity in nuclear medicine studies is validated by comparison with ROC (received operating characteristic)-curves. The method is tested in 21 single photon emission computerized tomography (SPECT), performed with a cardiac phantom. Three different lesions (L_1, L_2 and L_3) were placed in the myocardium-wall by pairs for each SPECT. Three activities (84, 37 or 18.5 MBq) of 99mTc were used as background. Linear discriminant analysis was used to select the parameters that characterize image quality among the measured variables in the images [(Background-to-Lesion (B/L_i) and Signal-to-Noise (S_i/N) ratios)]. Two clusters with different image quality (P=0.021) were obtained. The ratios B/L_1, B/L_2 and B/L_3 are the parameters used to construct the function with 100% of cases correctly classified into the clusters. The value of 37 MBq was the lowest tested activity for which good results for the B/L_i ratios were obtained. The result coincides with the applied ROC-analysis (r=0.89).

Key words: Activity optimization, Image quality, SPECT, Discriminant analysis, ROC curves
doi:10.1631/jzus.2006.B0947 **Document code:** A **CLC number:** R8

INTRODUCTION

Image quality in nuclear medicine tomography is critically dependent on the activity undergone by the patient. The optimal activity is the smallest amount of activity which preserves diagnostic accuracy. Nevertheless, the optimum depends on the gamma camera used for imaging, the size of the patient and the imaging application (Mattsson *et al*., 1998).

To optimize, one has to study how diagnostic accuracy depends on the activity undergone for each particular study, taking into account the above factors. An optimization method can be obtained by using a set of images with shorter and longer registration times than the usual, but this method can only be used for static images. Other optimization method for static images is that used by Fjalling *et al*.(1997)[1]. They graded the images by experienced nuclear medicine physicians using received operating characteristic (ROC) analysis with successful results in the reduction of the activity to be undergone. In the case of dynamic studies (Moonen and Jacobsson, 1997) used patient studies and simulated varying activity levels by adding different amounts of randomized statistical noise. ROC analysis was also used to optimize procedures between filtered and unfiltered studies (Stark and Carlsson, 1997). In other cases, the current situa-

* Project supported by the Third World Academy of Sciences (TWAS), Cuba

[1] Fjalling, M., Forssell Aronsson, E., Grétarsdóttir, J., Jacobsson, L., 1997. Personal Communication. Sahlgrenska University Hospital, SE-413 45 Goteborg, Sweden

tion is described in terms of the statistical distribution of activities undergone in a region, having similar equipment and processing methods. The last one is a defensive and descriptive approach, which may prolong the use of high activity levels (Mattsson *et al*., 1998). This group of authors also arrived at the conclusion that the final decision on the used activity undergone should be based on ROC analysis. Vestergren *et al*.(1996; 1998; 1999) developed some mathematical models to optimize activities in pediatric nuclear medicine, keeping the same image quality for patients of all ages.

The use of the scores and probabilities derived from statistical methods such as discriminant analysis and clustering techniques has increased their popularity since 1975 (Venables and Ripley, 1994). It was used first in marketing and after that, in some medical applications (Juez and Díaz, 1996). They offer a more continuous scale than the rank techniques based on the ROC curve analysis (Metz, 1978).

Perez *et al*.(2002a; 2002b; 2002c; 2003) developed a mathematical method for activity optimization in nuclear medicine studies, based on clustering and discriminant analysis of image quality. It constitutes a new approach to the same problem: to reduce the patient dose while image quality keeps good (Mattsson *et al*., 1998; ICRP 52, 1987; ICRP 53, 1987; ICRP 60, 1991). This work presents a validation of this new method by comparison with the well-known ROC analysis.

METHODS

Image acquisition

A cardiac phantom presented in Fig.1 (model ECT/CAR/I from GAMMASONICS http://www.gammasonics.com/) (Greer and Scarfone, 1999) was used. Three solid pathological structures (Fig.1) were used following the combinations that appear in Table 1.

A pair of cold lesions were positioned in two different locations over the wall of the myocardium in each of 21 studies included in this work. The myocardium was constituted as a hot background, filled with one of the three, a priori selected activity levels of 99mTc (activity 1=84 MBq, activity 2=37 MBq or activity 3=18.5 MBq), diluted into the water volume of this cavity for each SPECT. Activities were measured using a dose calibrator PTW CURIEMENTER 3 (1998, Germany). The phantom was placed on a magnetic stirrer for 5 min to achieve homogeneity of the myocardium activity, prior to each imaging session.

The phantom was positioned over the table

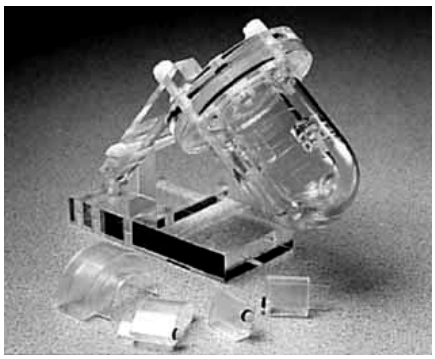

Fig.1 Cardiac phantom

Table 1 Combinations of lesions and activities used to acquire the SPECT

Case	Lesion	Position	Background activity (MBq)
1	L_1	1	18.5
2	L_1	1	37.0
3	L_1	1	84.0
4	L_2	2	18.5
5	L_2	2	37.0
6	L_2	2	84.0
7	L_3	1	18.5
8	L_3	1	37.0
9	L_3	1	84.0
10	L_1 and L_2	1 and 2	18.5
11	L_1 and L_2	1 and 2	37.0
12	L_1 and L_2	1 and 2	84.0
13	L_3 and L_2	1 and 2	18.5
14	L_3 and L_2	1 and 2	37.0
15	L_3 and L_2	1 and 2	84.0
16	L_1 and L_3	1 and 2	18.5
17	L_1 and L_3	1 and 2	37.0
18	L_1 and L_3	1 and 2	84.0
19	−	−	18.5
20	−	−	37.0
21	−	−	84.0

L_1: 45° (angular extension), 1.5 cm (height), 1.0 cm (width); L_2: 60° (angular extension), 2.0 cm (height), 0.5 cm (width); L_3: 60° (angular extension), 2.0 cm (height), 1.0 cm (width)

(plane *XY* of the gamma camera). The axis *Y* of the phantom coincided with axis *Y* of the camera for all the SPECT, to guarantee the reproducibility in the location of the phantom under the detector center. Although the position of the phantom over the *XY* plane is not exactly the same as that in a human heart (in our case, the phantom was not rotated with respect to *Y* axis in the *XY* plane), the angular position of the phantom apex with respect to the *XY* plane of the camera resembled the human apex position.

A digital Sopha camera (1000, circular DCX, France) equipped with a general purpose, mean resolution parallel-hole collimator (HRBE8-140) was used for imaging. The gamma camera was tested with NEMA Protocol (NU-1-1994 address SPECT camera performance) (NEMA, 1994).

Thirty-two views were acquired over 180° in each SPECT, using a circular orbit starting at −45° right anterior oblique (RAO) and ending at 135° left posterior oblique (LPO) (30 s per projection). Images were acquired in 64×64 pixel matrix size. The rotation radius was 16 cm for all the studies, as this is the nearest position between the phantom and the camera detector. The energy window used was 15% around the 140 keV photopeak.

Image reconstruction

The reconstruction was developed following filtered back projection algorithm. The projections were prefiltered with a Ramp [cutoff frequency f_c=1.0 Nyquist (Ny)] and then with a Butterworth filter (order 4, f_c=0.18 Ny) (interslices=1 pixel). Attenuation conditions were not taken into account in this preliminary experiment.

ROC analysis

Forty-one slices of the 21 SPECT were selected to perform the lesion detection analysis. An example is shown in Fig.2. The slices were selected to cover the lesion sites and a background region among images obtained with different activities. As the position of the myocardium and the detector orbit remained constant for all the acquisitions, identical reconstruction slices and angles were used to generate oblique long- and short-axis slices and coronal slices of the phantom. The 41 selected reconstructed images were shown to three observers blinded to the used activity, experienced in myocardium SPECT (medical doctors, with specialty in cardiology and nuclear medicine: 15, 10 and 12 years, respectively; and more than 8 years at the Hospital Nuclear Medicine Service). Each image contains 10 regions to evaluate (Fig.2). The observers rated 82 abnormal regions (corresponding to lesion sites) and 54 normal regions, with each activity, on a 4-point scale in order to obtain a score. The evaluation was as follows: 1=certainly normal, 2= probably normal, 3=probably pathological and 4= certainly pathological. Normal and abnormal regions were asked at random. Furthermore, all the observers graded each SPECT as good image quality (5 points), fair image quality (4 points) or poor image quality (3 points), as is conventional. The order for analyzing the studies was from low to high activity to avoid a negative predisposition of the observers to grade first images with supposed better image quality.

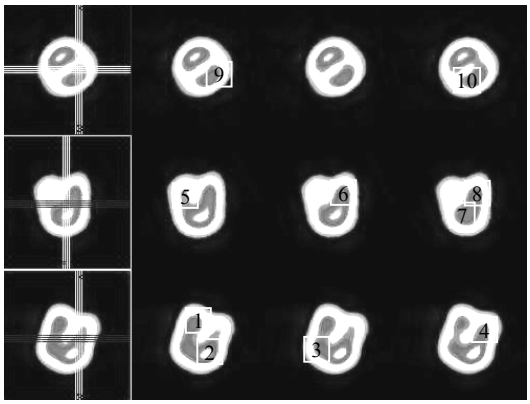

Fig.2 Example of reconstructed image graded by the observers

Conventional receiver operating characteristic (ROC) curve analysis (Metz, 1978) was applied to analyze the data from the tests. The area under the ROC curve was calculated using Wilcoxon statistics (Hanley and McNeil, 1982).

Measured variables

The maximum number of counts per pixel in the reconstructed images was determined in regions of interest (ROIs of 3×3 pixels). These regions contain the lesion sites on the images (L_1, L_2 or L_3), the apex (A) (which it was always considered as a background region in our study), and two more additional background regions (P_1 and P_2) taken at the same position for all the studies. These variables were used to con-

struct a group of Background-to-Lesion ratios (B/L_1, B/L_2 and B/L_3). The mean background for each study is calculated as the mean value among all the considered background sites (A, P_1 and P_2) in each analyzed image. The ratios were calculated as:

$$B/L_i=\text{Background}/(\text{Background}-\text{Lesion}),\quad i=1, 2, 3. \tag{1}$$

These Background-to-Lesion ratios and the Signal-to-Noise ratios were variables evaluated for the quantitative criterion of image quality. They were introduced at the software Statistical Package for Social Science (SPSS 9.0) to develop the discriminant analysis.

The Signal-to-Noise ratios were calculated following Hoffman and Phelps equation (Phelps *et al*., 1975).

$$S_i/N=(N_{ri})^{1/2}/R^{1/4},\quad i=1, 2, 3, \tag{2}$$

where N_{ri} is the reconstructed counts, R is the number of pixels containing the activity.

Image contrast was calculated as the reciprocal of B/L_i ratios per 100%.

Mathematical procedure

1. Clustering technique

Clustering techniques can be used to determine if, from a group of measured variables, different levels of image quality can be detected with significant differences among them (Perez *et al*., 2002a; 2003). This technique may be necessary in those cases in which observers cannot distinguish different levels of image quality by simple observation. Initial data obtained from the N_c cases or studies (21 in our experiment) were analyzed to reconstruct images (41 images in our study corresponding to 21 SPECT or cases), can be used to calculate the distance between pairs of images corresponding to cases in which common variables have been measured, including all the possible combinations among the cases (36 by subgroups). It is calculated by the traditional Euclidean distance:

$$d_{A,B}=[(X_{A1}-X_{B1})^2+(X_{A2}-X_{B2})^2+...+(X_{Ai}-X_{Bi})^2]^{1/2}, \tag{3}$$

where $X_{A1}\ldots$ to X_{Ai} are the values of each of the X_i measured variables of one image (Group of B/L_i and S_i/N ratios in this particular study) and $X_{B1}\ldots$ to X_{Bi} are the values of each of the X_i measured variables (the same) in the other image. As in all images the same B/L_i and S_i/N ratios are not measured, then 0 is attributed in the corresponding difference X_A-X_B for each particular not measured variable.

The images compared belonged to the following cases: subgroup 1: cases 1, 2, 3, 10, 11, 12, 16, 17 and 18; subgroup 2: cases 4, 5, 6, 10, 11, 12, 13, 14 and 15; subgroup 3: cases 7, 8, 9, 13, 14, 15, 16, 17 and 18.

The k-means clustering method chooses a specified number of cluster centroids. The two images with the highest distance were selected by subgroups. The variable's values for these extreme cases were considered the initial centroids of the initial clusters of each subgroup. The distance from each centroid was calculated for each image belonging to each subgroup. Each case was then placed in the cluster whose distance to the centroid is the lowest by subgroups. This procedure confirms 6 clusters (2 by subgroups). Nevertheless, the cases can be grouped in only two final clusters by transitivity, taking into account that each case belongs to two clusters at the same time. The centroids are recalculated for each final cluster. Under these conditions each cluster is ensured to have a different level of image quality according to the X_i measured variables.

We have been working with a model of two clusters. Nevertheless, a model of three clusters is also possible to construct, following basically, the same procedure. In this model the centroids of the initial clusters would be the variables corresponding to the three images with highest distance among them. The rest of the procedure is the same.

2. Discriminant analysis

Discriminant analysis was then used for the obtained cluster's centroids. The objective is the construction of linear image quality discriminant functions to select from all the measured variables, which determine image quality. One function is necessary in a model of two clusters, but if we have 3 clusters, then, 2 functions will be necessary. In our particular case we constructed two clusters. The method also determines the relative relevance of each selected variable in the above discrimination procedure by its correlation coefficient with the function (Venables and Ripley, 1994).

The form of the function is:

$$Z = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_{n^*} x_{n^*}, \quad (4)$$

where Z is the discriminant punctuation of each case [in our case Z is the cluster value which represent image quality (IQ)].

The λ_s values are calculated from n^* linear equations.

$$\begin{aligned} &\lambda_1\sigma_{11} + \lambda_2\sigma_{12} + ... + \lambda_{n^*}\sigma_{1n^*} = \mu_{11} - \mu_{12}, \\ &\lambda_1\sigma_{21} + \lambda_2\sigma_{22} + ... + \lambda_{n^*}\sigma_{2n^*} = \mu_{12} - \mu_{22}, \\ &... \\ &\lambda_1\sigma_{n^*1} + \lambda_2\sigma_{n^*2} + ... + \lambda_{n^*}\sigma_{n^*n^*} = \mu_{1n^*} - \mu_{n^*2}, \end{aligned} \quad (5)$$

where σ_{kp} (k, p=1, …, n^*) are the covariance matrices corresponding to the X_{n^*} variables. They are calculated as:

$$\sigma_{kp} = \sum_{i=1}^{N_c} (X_{ki} - \overline{X}_k)(X_{pi} - \overline{X}_p). \quad (6)$$

μ_{ij} are the mean values of each of the selected variables, where i takes the values 1 or 2 (groups) and (j=1, ..., n^*).

Depending on the Z value ($Z<C$ or $Z>C$), each case was classified into a cluster, according to the new selected variables X_{n^*} (in our case, they are those selected values of B/L_i and/or S_i/N from the initially introduced which are not correlated among them).

C is considered the threshold between clusters and may not be determined by visual observation, but can be calculated as:

$$C = (\overline{Z}_N + \overline{Z}_A)/2, \quad (7)$$

where $\overline{Z}_N$ and $\overline{Z}_A$ are the discriminant punctuations for both final groups according to the X_{n^*} discriminant variables.

To reduce the number of variables relevant to image quality from the initial measured X_n to the final X_{n^*} ($n \le n^*$), the correlation factors among all the variables are calculated, following the relevance order of the variables for image quality. Only those variables with the highest relevance for image quality, among the totally entire measured are included in the function formed with variables not correlated or with poor correlation (r<0.2) (Venables and Ripley, 1994). Other variables correlated with the selected variables are neglected. The parameters that finally determine image quality in a particular study will be those that cannot be extracted from the function without affecting the percentage of cases correctly classified (it means, correctly placed in a cluster using discriminant analysis). Then, the activity optimization criterion consisting of selecting among a group of tested values, the minimum administered activity which permits good results for the reduced number of parameters selected by the function (Perez *et al*., 2002a; 2002b; 2002c; 2003).

RESULTS

ROC analysis

Fig.3 shows the ROC curves for each observer and the three tested activities. Table 2 summarizes the area under the ROC curves for the observers. The

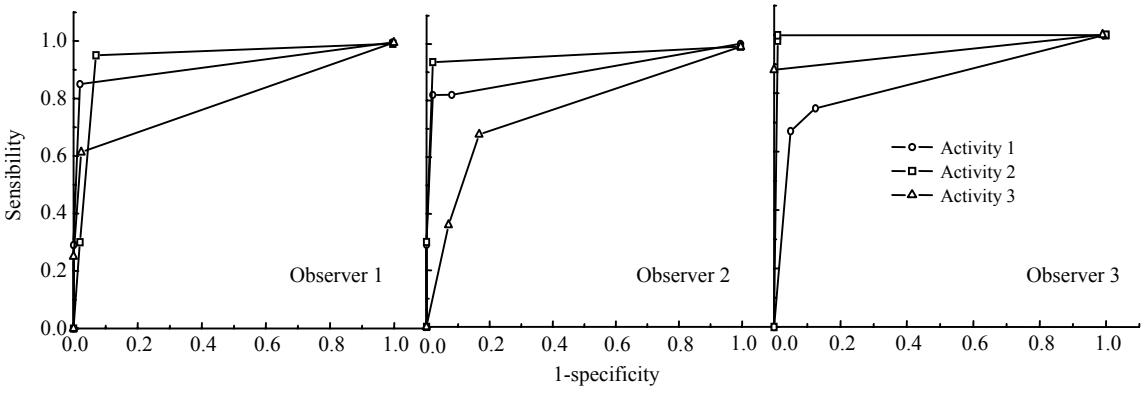

Fig.3 ROC curves from each observer for the three activity values tested

Table 2 Mean area under the ROC curves for the three observers

Activity (MBq)	Area (±*SD*) under ROC curves
84.0	0.92±0.08
37.0	0.92±0.05
18.5	0.79±0.07

difference between the results for 37 MBq and 84 MBq was not significant (P=0.32), at 95% confidence level, in the case of the three observers. Nevertheless, these differences were significant for 18.5 MBq, mainly for the observer 1 (P=0.008).

It is also important to explain the behavior of the false-positives and the false-negatives rates for each observer. The curves for the observer 1 show that when the activity is increased, the false-positive rate is increased from 0.024 with 18.5 MBq to 0.073 with 84 MBq. No significant differences were seen between 18.5 MBq and 37 MBq (P>0.05), at 95% confidence level. In the case of the calculated false-negative rate for the observer 1, the reduction in the administered activity produces an increase of this rate from 0.036 for 84 MBq to 0.38 for 18.5 MBq. For the last value, 21 false-negatives were reported, 8 for 37 MBq and just 2 for 84 MBq. For the observer 2, the false-positive rate was 0.02 for 37 and 84 MBq and it was increasing up to 0.097 for 18.5 MBq. The calculated false-negative rate was 0.05 for 84 MBq, 0.18 for 37 MBq and 0.31 for 18.5 MBq. For the observer 3, the false-negative rates were: 0.12 for 84 MBq, 0.02 for 37 MBq and 0.25 for 18.5 MBq. The false-positives rates were: 0 for 84 MBq, 0.011 for 37 MBq and 0.14 for 18.5 MBq. For the last value, 6 false-positives and 29 false-negatives were reported. From this point of view, it appears to not be possible to reduce the administered activity below 37 MBq in this experiment without affecting the precision in the observer evaluation, mainly for the false-negatives rate. The differences in the false-negative rate between 37 and 84 MBq were not statistically significant (P=0.07) for 95% confidence level and the false-positive rate was a minor problem.

Collateral result to the ROC analysis

Fig.4 shows the values of the contrast for each lesion as a mean value of the three slices for each activity. Significant differences in image contrast were obtained with 18.5 MBq (P<0.05 for all the cases) from the 37 and 84 MBq cases for the three lesions. This variation was not significantly different between 37 and 84 MBq (P>0.05).

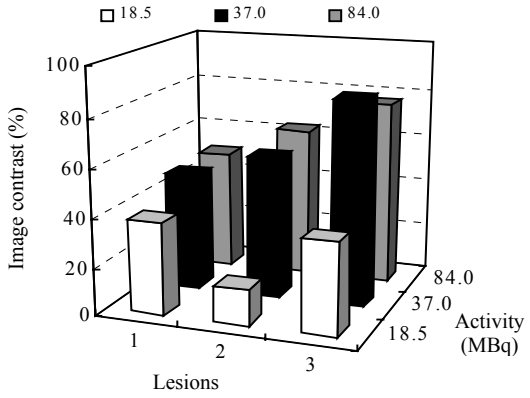

Fig.4 Mean contrast for each lesion

Subjective analysis

Fig.5 shows the behavior of the mean score given by the observers, when they graded each lesion with each activity.

The value of 37 MBq was adequate for an accurate evaluation of the lesions. The mean score given by the observers was not significantly different regarding the score for 84 MBq (P=0.08). The value of 18.5 MBq, however, had the worst result in this sense. The scores given for each lesion were lower than that of the other two activities in this case.

Discriminant analysis

The three observers graded all images observed as good image quality (5 points), independent of the activity level in each case. Nevertheless, two clusters with differentiated image quality (P=0.021) were obtained according to the Background-to-Lesion ratios. The first one involved the studies performed with 37 and 84 MBq, and the second one included the studies performed with 18.5 MBq. The centroids of the variables for the first cluster were: B/L_1=1.93, B/L_2=1.71 and B/L_3=1.27. The centroids of the second cluster were B/L_1=2.68, B/L_2=6.64 and B/L_3=2.61. We thus can assume that the first cluster contains the studies with better image quality from the quantitative point of view. In this case, as the lesions are cold regions, the increase of the activity implies a decrease in Background-to-Lesion ratios, when the dispersive conditions are constant and partial volume effects are negligible.

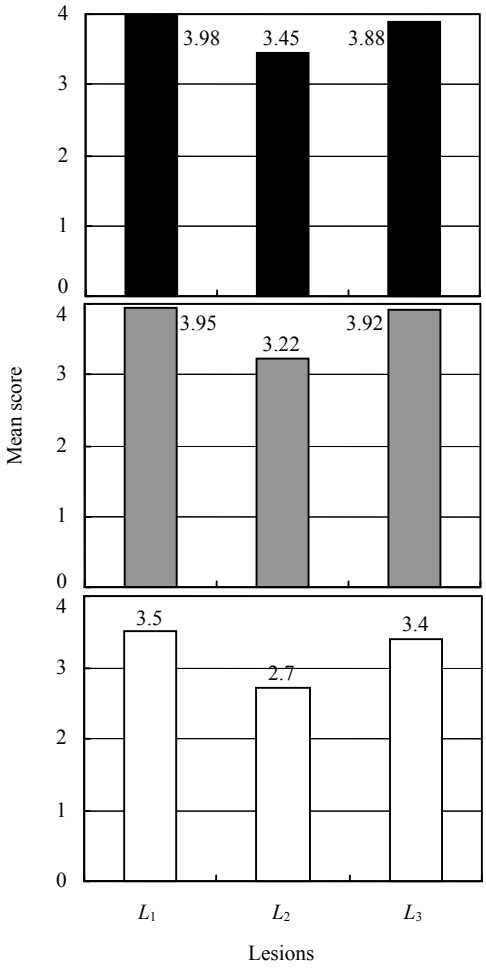

Fig.5 Mean score for each lesion size with each activity [37 MBq (top), 84 MBq (middle); 18.5 MBq (bottom)]

The results of the cluster classification were strongly correlated with the areas under ROC curves (r=0.89).

Following the cluster criterion a discriminant analysis was applied. The obtained linear discriminant function was:

$$IQ=0.987B/L_1-0.383B/L_2+0.813B/L_3. \quad (8)$$

IQ is cluster discrimination between the centroids (Z value).

The function was significant. The Wilks' λ was equal to 0.006 (Venables and Ripley, 1994). All of the functions studied were correctly classified into the clusters. The correlation coefficient of each variable with the function were B/L_1 (r=0.895), B/L_3 (r=0.461) and B/L_2 (r=0.198). These coefficients represent the relative weight of each variable in the discrimination procedure.

Table 3 shows the mean values of B/L_i for the three activities. The variables were in general better for 84 MBq but without significant differences with respect to 37 MBq (P=0.078).

Table 3 Mean Background-to-Lesion ratios for each activity

Activity (MBq)	Background-to-Lesion ratios		
	B/L_1	B/L_2	B/L_3
18.5	2.32±0.21	6.57±0.27	2.61±0.18
37.0	2.07±0.15	1.75±0.20	1.26±0.19
84.0	1.95±0.14	1.67±0.31	1.28±0.16

Table 4 shows the results of the mean values of the tomographic Signal-to-Noise ratios (S_i/N) in the analyzed images, for the three lesions and the three activities. S_i/N showed correlation with the Lesion-to-Background ratios (r=0.922). For this reason they were extracted from the discriminant function.

Table 4 Mean tomographic Signal-to-Noise ratios in the three slices

Activity (MBq)	Signal-to-Noise ratios		
	S_1/N	S_2/N	S_3/N
18.5	7.57±0.04	6.56±0.04	7.87±0.05
37.0	10.70±0.03	10.41±0.02	11.22±0.03
84.0	12.04±0.03	13.37±0.01	13.07±0.02

DISCUSSION

A cardiac phantom was used to obtain the images with different activities. This phantom has the disadvantage that the lesion sites always are the same and that the images have few differences among them. On the other hand, although this phantom seems fairly unrealistic (heart phantom in air, without attenuation, there was radioactivity in the region corresponding to heart muscle, but non elsewhere, and the position of the phantom over the table for the experiment was not exactly the same as that of a human heart), those aspects do not have serious implications for this paper's purpose (to validate a mathematical procedure for activity optimization in SPECT images, taking into account physical parameters by comparison with ROC).

For optimizing real activities for cardiac SPECT, this work is still a preliminary work, which does not yield meaningful results for realistic situations like the case of 99mTc-MIBI SPECT in patients. More investigations to apply the method in this particular study are needed in order to consider common values of background and noise in routine practice with patients, due to realistic conditions of scatter and attenuation, but this is not the aim of the present work.

The a priori selected activities were chosen only for the purpose of this experiment with a phantom, in order to guarantee enough image counts which represent different levels of image quality. Nevertheless, neither of the activities used in the experiment are typical radioactivity levels for cardiac SPECT nor are the results of the measured S_i/N ratios typical. If researchers want to reproduce this procedure in patients, the activities need to be carefully chosen and tested in routine practice. The ways to adequately apply the discriminant procedure to other real patient studies are deeply discussed in (Perez *et al*., 2002a; 2002b; 2002c; 2003).

Another way to improve image quality in our experiment to maintain low levels of administered activity and guarantee the observation of all lesion sizes, could be to change the projection acquisition time from 30 s to, for example, 60 s or more. This option may not be adequate in practice for some patients, due to the overall increase in image acquisition time.

ROC analysis is recommended for finding optimal activity values to detect the lesions containing 37 MBq (the area under the ROC curve was not statistically different from the 84 MBq case). Furthermore, from the point of view of the false-positive and false-negative rates, activity levels lower than 37 MBq appear to adversely affect the precision of lesion observation in the experiment involving all observers. The differences in the false-positive rate between 37 and 84 MBq were not significant at a 95% confidence level and the false-negative rate was a minor problem for both activity levels. Nevertheless, when the activity was reduced to 18.5 MBq the observers had problems with threshold detection and the false-negative rate increased. Image contrast, for the specific case of L_2, diminished almost to the detection limit of the human eye (13%) (Goodenough, 1981). This lesion (the thinnest) was the one most often related to the high rate of false-negatives. The image contrast grew with increase of the activity.

All lesions were larger than (1/1.75) times the full width at half maximun (FWHM) of the imaging system (The spatial resolution of the system used with collimator was 5.25 mm). For this reason, the possible influence of partial volume effects is considered to be negligible over the obtained false-negative rates (Tsui and Zhao, 1994). In fact, this rate was particularly low for 37 and 84 MBq.

Good image quality implies maintenance of the same noise level and the same spatial resolution in tomography. Spatial resolution is not dependent on the level of administered activity, but depends only on parameters such as the distance between the collimator and the studied object and the filter used in the reconstruction. These parameters were kept constant in all the studies. Those are possible reasons why the observers judged all analyzed studies as good tomographic image quality.

There was good threshold discrimination between useful signals (lesions and structures) and noise, with differences of more than two standards deviations for the observer's grades. This suggests that good discrimination was obtained, in general, between normal and abnormal sites as the observers and ROC curves moved upwards in areas near to 1 (Goodenough, 1981; Evans, 1981).

The ROC methodology requires a very large number of observations to arrive at conclusions (410 observations were done in this study by 3 observers). This kind of analysis is also very determined by the experience degree of the observers and is always affected by negative and positive false rates (Metz, 1978; Goodenough, 1981; Evans, 1981). Nevertheless, in this particular study we did not observe significant inter-observer variability ($P>0.05$), mainly due to the similar expertise degree of the three observers.

Nevertheless, the proposed method gives an easier discrimination than the subjective observer's opinions and ROC analysis. The method application found slight differences in tomographic image quality, considering the values of the Background-to-Lesion ratios and also the Signal-to-Noise ratios, which were strongly correlated with the Background-to-Lesion ratios, the Signal-to-Noise ratios were extracted from the function, but in our study they are good for re-

marking the same results and can be used in future work as determinant variables for discriminant procedures in nuclear medicine activity optimization.

There were real differences in the image quality obtained for activity levels of 37 and 84 MBq compared to 18.5 MBq (the studies belong to different clusters with significant differences between them). These differences were not evident for typical observers.

The applied method classified more than 75% of cases correctly into the clusters. This is the minimum percentage admissible to consider the method application as successful (Venables and Ripley, 1994). It ensures the possibility to predict what image quality would be a completely new case with each tested activity. Then, we only need to verify the above measuring B/S_i ratios or S_i/N ratios in the images.

The relevant order for image quality in our experiment was as follows: L_1, L_3 and L_2. The small lesion was the most relevant in the mathematical discrimination. The ratios of L_2 regarding background had the lowest correlation with the obtained classification criterion. Following this order we establish the optimization criterion. The value of 37 MBq was considered optimal for this experiment after method application. This value was the minimum for which image quality was always good with the quantitative criterion of image quality (all the cases belong to the cluster with better image quality). For this value the Background-to-Lesion ratios did not present significant differences with respect to 84 MBq. Signal-to-Noise ratios were also reinforced in this result.

Another result possibility, in the case of the model of two clusters, would be that the cluster with better image quality would contain only the studies performed with the highest activity. In this case this activity would be the optimal, but we also should apply discriminant analysis for obtaining the variables which determine this result. We also develop multivariate analysis among the results of the clusters with the lowest image quality, trying to see if the other two activities produce image quality with significant differences between them. If the three activities produce significant differences in image quality, then the activity cannot be reduced for optimization purpose. This type of model also permits analysis of more than three activities (Perez *et al*., 2002a).

If we select a model of three clusters and find significant differences among the three activities regarding image quality, determined by the selected variables with discriminant analysis; then, the optimum is the activity which produces the best image quality contained in one cluster. In this model, if we obtain a result similar to the one obtained in the present paper, then, one cluster will be empty, another contains the cases belonging to the two highest activities, which produce similar and good image quality and the other one will contain the studies performed with the lowest activity, which contains the studies with the lowest image quality. The optimum will be the lowest activity which permits keeping the cases in the cluster with better images quality. Generalizing this methodology, we can construct a model of n clusters to test m activities ($n \leq m$).

Another advantage of the proposed method is that it permits focusing the analysis of which activity is selected as optimal just on a reduced group of variables. This fact can be particularly useful in patient studies like (Perez *et al*., 2002a; 2002b; 2002c; 2003), where the number and diversity of variables with influence over image quality can be very high. In these cases, if a robust statistical technique is not applied, it is not obvious which variables determine the criterion of the observer who performs the diagnosis from the medical images.

This method can be applied for any kind of the current gamma cameras and is particularly useful for patient radiological protection (ICRP 52, 1987), according to the ALARA (as low as reasonably achievable) principle (ICRP 60, 1991), because by reducing the administered activity we can reduce the absorbed dose in many tissues and organs of the patients, which imply a reduction in the radiological risk (ICRP 53, 1987).

The discriminant method has disadvantage in the necessity to a priori establish the activity values that the researcher needs to test for each type of nuclear medicine study. It is not possible to obtain an interpolated value as the optimal activity.

The optimal obtained activity with the proposed procedure coincides with the activity for which the area under the ROC curve is the highest, the sensitivity is the best, with false-positive and false-negative rates being a minor problem in our work. Nevertheless, in our opinion, discriminant method provides an easier and faster procedure than the ROC method for ana-

lyzing the image quality for an activity optimization procedure. The discriminant method is, besides, completely independent of the observer subjectivity.

CONCLUSION

Discriminant analysis to select which variables can describe image quality appears to be a fast and easy method to choose parameters to take into account the activity optimization procedure.

The results of this procedure were similar to the results obtained from the well-known ROC analysis. The value of 37 MBq was selected as the optimal for obtaining good image quality with the technical conditions applied for this particular study using ROC and discriminant method.

ACKNOWLEDGEMENT

The authors thank Prof. Michael Stabin, DSc. for helping with manuscript preparation.

References

Evans, A., 1981. The Evaluation of Medical Images. *In*: Adams Hilger Ltd. (Ed.), Image Evaluation by Signal Detection Theory. Medical Physics Handbook. Techno House, UK, p.80-113.

Goodenough, J., 1981. Technical Aspects of CT. *In*: Newton 8 Potts (Ed.), Psychophysical Perception of Computed Tomography Images. Washington DC, p.3993-4021.

Greer, L., Scarfone, C., 1999. Data Spectrum's SPECT User's Manual. Data Spectrum Corporation 437 Dimmock Mill Road Hillsboroug, NC 27278.

Hanley, J., McNeil, B., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **143**(1):29-36.

ICRP (International Commission on Radiological Protection) 52, 1987. Protection of the Patient in Nuclear Medicine. Publication 52 Ann., ICRP **17**(4). Pergamon Press, Oxford.

ICRP (International Commission on Radiological Protection) 53, 1987. Radiation Dose to Patients from radiopharmaceutical. Publication 53 Ann., ICRP **18**(1-4). Pergamon Press, Oxford.

ICRP (International Commission on Radiological Protection) 60, 1991. Recommendations of the International Commission on Radiological Protection. Publication 60 Ann., ICRP **21**(1-3). Pergamon Press, Oxford.

Juez, P., Díaz, F.,1996. Probabilities and Statistic in Medicine Applications in the Clinic Practice and Sanitary. Madrid, p.189-222.

Mattsson, S., Jacobsson, L., Vestergren, E., 1998. The basic principles in assessment and selection of reference doses: considerations in nuclear medicine. *Rad. Prot. Dosim.*, **80**(1):23-27.

Metz, C., 1978. Basic principles of ROC-analysis. *Sem. Nucl. Med.*, **8**(1):283-298.

Moonen, M., Jacobsson, L., 1997. Effect of administered activity on precision in the assessment of renal function using gamma camera renography. *Nucl. Med. Commun.*, **18**(4):346-351.

NEMA (National Electrical Manufacturers Association), 1994. Performance Measurements of Scintillation Cameras. Washington DC.

Perez, M., Quevedo, J., Ponce, F., Diaz-Rizo, O., 2002a. Administered activity optimisation in patients studied by equilibrium gated radionuclide ventriculography using pyrophosphate and Tc-99m. *Nucl. Med. Commun.*, **23**(3):347-353.

Perez, M., Quevedo, J., Diaz-Rizo, O., Dopico, R., Estévez, A., Viamonte, A., 2002b. Administered activity optimization in skeletal scanning using MDP labelled 99m-Tc. *ALASBIMN J.*, **16**(4):AJ16-5.

Perez, M., Diaz-Rizo, O., Dopico, R., Estéves, E., 2002c. Administered activity optimization in renal scintigraphy with DMSA-99mTc. *ALASBIMN J.*, **28**(5):AJ20-5.

Perez, M., Estévez, E., Diaz-Rizo, O., Roque, R., Hernández, C., 2003. Administered activity optimization in ^{99m}Tc-MAG3 renography for adults. *J. Nucl. Med. Technol.*, **31**(1):216-221.

Phelps, M., Hoffman, E., Mullani, N., Ter-Pogossian, M., 1975. Application of annihilation coincidence detection to transaxial reconstruction tomography. *J. Nucl. Med.*, **16**(2):210-224.

Stark, S., Carlsson, S., 1997. Digital filtering of bone scans: a ROC study. *Nucl. Med. Commun.*, **18**(2):98-104.

Tsui, B., Zhao, X., 1994. Quantitative single-photon computed tomography: basics and clinical considerations. *Sem. Nucl. Med.*, **24**(1):38-65. [doi:10.1016/S0001-2998(05)80248-X]

Venables, W., Ripley, B., 1994. Statistics and Computing. Modern Applied Statistics with S-Plus. Springer-Verlag, New York, p.311-318.

Vestergren, E., Jacobsson, L., Bjure, J., Mattsson, S., 1996. Radiopharmaceutical dosage of 99mTc-HMPAO for cerebral blood flow SPECT studies in children. *Nucl. Med. Commun.*, **17**(4):423-429.

Vestergren, E., Jacobsson, L., Lind, A., Sixt, R., Mattson, S., 1998. Administered activity of Tc-99m-DMSA for kidney scintigraphy in children. *Nucl. Med. Commun.*, **19**(6):695-701.

Vestergren, E., Jacobsson, L., Moonen, M., Eklind, I., Sixt, R., Mattsson, S., 1999. Administered activity of 99mTc-MAG-3 for gamma camera renography in children. *Nucl. Med. Commun.*, **20**(8):799-806.