Role of ventilation scintigraphy in diagnosis of acute pulmonary embolism: an evaluation using artificial neural networks

Eva Evander[1], Holger Holst[2], Andreas Järund[2], Mattias Ohlsson[3], Per Wollmer[2], Karl Åström[4], Lars Edenbrandt[2]

[1] Department of Clinical Physiology, University Hospital, Lund University, Lund, Sweden
[2] Department of Clinical Physiology, Malmö University Hospital, Malmö, Sweden
[3] Department of Theoretical Physics, Lund University, Lund, Sweden
[4] Department of Mathematics, Lund Institute of Technology, Lund, Sweden

Received: 1 November 2002 / Accepted: 2 March 2003 / Published online: 14 May 2003

Abstract. The purpose of this study was to assess the value of the ventilation study in the diagnosis of acute pulmonary embolism using a new automated method. Either perfusion scintigrams alone or two different combinations of ventilation/perfusion scintigrams were used as the only source of information regarding pulmonary embolism. A completely automated method based on computerised image processing and artificial neural networks was used for the interpretation. Three artificial neural networks were trained for the diagnosis of pulmonary embolism. Each network was trained with 18 automatically obtained features. Three different sets of features originating from three sets of scintigrams were used. One network was trained using features obtained from each set of perfusion scintigrams, including six projections. The second network was trained using features from each set of (joint) ventilation and perfusion studies in six projections. A third network was trained using features from the perfusion study in six projections combined with a single ventilation image from the posterior view. A total of 1,087 scintigrams from patients with suspected pulmonary embolism were used for network training. The test group consisted of 102 patients who had undergone both scintigraphy and pulmonary angiography. Performances in the test group were measured as area under the receiver operation characteristic curve. The performance of the neural network in interpreting perfusion scintigrams alone was 0.79 (95% confidence limits 0.71–0.86). When one ventilation image (posterior view) was added to the perfusion study, the performance was 0.84 (0.77–0.90). This increase was statistically significant (P=0.022). The performance increased to 0.87 (0.81–0.93) when all perfusion and ventilation images were used, and the increase in performance from 0.79 to 0.87 was also statistically significant (P=0.016). The automated method presented here for the interpretation of lung scintigrams shows a significant increase in performance when one or all ventilation images are added to the six perfusion images. Thus, the ventilation study has a significant role in the diagnosis of acute lung embolism.

Keywords: Image processing – Artificial neural networks – Pulmonary embolism

Eur J Nucl Med Mol Imaging (2003) 30:961–965
DOI 10.1007/s00259-003-1182-5

Eva Evander (✉)
Department of Clinical Physiology, University Hospital, Lund University, 221 85 Lund, Sweden
e-mail: eva.evander@skane.se
Tel.: +46-46-173304, Fax: +46-46-151769

Introduction

Lung scintigraphy has been used to diagnose pulmonary embolism for almost 40 years. Initially, perfusion scintigraphy only was performed, and the evaluation of the study was based on an analysis of several characteristics of the perfusion defects (e.g. their number, size, shape and demarcation). The first technique for ventilation scintigraphy to become widely used was the single breath/wash-out of xenon-133. With this technique, information on ventilation-perfusion mismatch is generally obtained in one posterior view only. Subsequently, krypton-81m and radioactive aerosols were introduced for ventilation scintigraphy. With these techniques, multiple images of ventilation are obtained in the same projections as perfusion images, allowing a more direct comparison of the distribution of ventilation and perfusion. Introduction of the ventilation scan caused the interpretation of lung scintigrams to focus much more on the ventilation-perfusion mismatch present in pulmonary embolism. Several studies showed improved diagnostic accuracy with the addition of a ventilation study [1, 2], and combined ventilation-perfusion scintigraphy became the method of choice. This practice was challenged some years ago in the PISA-PED study [3], which showed excellent diagnostic accuracy of perfusion scintigraphy only. The value of the ventilation scan is thus still uncertain.

An objective means of evaluating the value added by the ventilation scan is offered by automatic image interpretation using computerised image processing and artificial neural networks [4]. The interpretation of lung scintigrams based on artificial neural networks uses automatically selected features from the images in order to achieve complete objectivity. The purpose of this study was to evaluate the value of the ventilation study in the diagnosis of pulmonary embolism using a new automated method employing either perfusion scintigrams alone or the combination of ventilation and perfusion scintigrams as the only source of information regarding pulmonary embolism. We also wanted to assess the influence on performance when the network received information from the perfusion study combined with a single posterior ventilation image, as this mode of investigation is used in many studies.

Materials and methods

Training group. During the period 1 January to 30 June 1997 and 1 January to 30 June 1998 a total of 1,096 patients underwent lung scintigraphy due to suspected pulmonary embolism at the University Hospital, Lund, Sweden. The scintigram consisted of a ventilation and a perfusion study with six projections in both studies. Patients who had been examined more than once during the period contributed only with their first lung scintigram. Nine patients in the training group also fulfilled the inclusion criteria for the test group and were excluded, leaving 1,087 scintigrams in the training group (Table 1). In order to present the network with a binary classification, all studies were re-examined independently by two experienced physicians, who provided a consensus assessment for either "pulmonary embolism" or "no pulmonary embolism". Also, intermediate scintigrams were classified as "pulmonary embolism" or "no pulmonary embolism" depending on, for example, the presence of typical wedge-shaped perfusion defects without matching ventilation abnormalities. The prevalence of embolism in the training material was 15%. The training group has been described in detail previously [4, 5].

Table 1. Study population

	Training group	Test group
No. of patients	1,087	102
Female (%)	60	58
Age, mean (range) years	62 (7–98)	58 (17–93)
Prevalence of pulmonary embolus (%)	15	55

Test group. All patients who during the period 1 January 1993 to 31 March 1998 had undergone both lung scintigraphy and pulmonary angiography at the University Hospital in Lund, with no more than 2 days between the examinations, were included in the test group. The test group consisted of 102 patients and none of these patients were included in the training group. The result of pulmonary angiography was used as the gold standard. Angiography was performed according to standardised requirements and interpreted by experienced radiologists. The prevalence of embolism was 55%. The test group has been described in detail previously [4].

Imaging protocol. The ventilation study was always performed before the perfusion study. The patient inhaled an aerosol of 15–25 MBq of technetium-99m diethylene triamine penta-acetate (Solco/Sorin DTPA, Solco Nuclear, Birsfelden-Basel, Switzerland) in the supine position. Immediately after the ventilation study, perfusion imaging was performed following intravenous administration of 100 MBq of ^{99m}Tc-macroaggregated albumin (TechneScan LyoMAA, Mallinckrodt Medical, Petten, Holland) with the patient in the supine position. The scintigrams were obtained with the patient sitting in front of a large-field-of-view gamma camera with a low-energy general-purpose collimator (Toshiba GCA 901A/ECT, Toshiba Corporation, Tokyo, Japan), with images acquired in six projections (anterior, posterior, left posterior oblique, right posterior oblique, left lateral and right lateral). Images were stored digitally in 128×128 matrix size. This standard protocol was unchanged during the study period [6].

Image processing. Image processing was performed to reduce the abundance of information in the lung images to a few features relevant to the diagnosis of pulmonary embolism. The image processing was performed in several consecutive steps, previously described in detail [4]. In a first step, any hot spots in the ventilation images, due to local deposition of radioactive particles in central airways, were removed and replaced with an interpolation of the pixels surrounding each hot spot. Next, in order to compensate for variations in the size and shape of the lungs and differences in the position of the patient in the ventilation study and the corresponding perfusion study, all scintigrams were aligned to templates representing lungs of normal size and shape. The alignment included translation, rotation, scaling and skew of the images and also scaling of the intensity.

Three types of alignment were used depending on the images in the study. When only perfusion scintigrams were used, the six images were aligned to the corresponding templates. When joint ventilation-perfusion scintigrams in six projections were used, the ventilation images were first aligned to the corresponding templates. Thereafter, the six perfusion images were aligned to the corresponding ventilation images that had previously been aligned. In the study using perfusion scintigrams in six views and only the ventilation scintigram from the posterior view, the single ventilation image was first aligned to a template and thereafter the perfusion image in the posterior view was aligned to the ventilation image that had previously been aligned. The five remaining single perfusion images were aligned to the corresponding templates.

In order to describe mismatches in the lung images, a pixel to pixel perfusion/template quotient image or a perfusion/ventilation quotient image was calculated. This resulted in a new image, a quotient image based on the perfusion and the template images or a quotient image based on the perfusion and the ventilation images. Quotient images were calculated in all six projections for both types of alignment. Quantitative measures of mismatch were calculated. The lungs of each patient were divided into 18 segments based on an anatomical atlas [7]. For each segment, the sum of pixels in all projections with a significant mismatch (defined as a perfusion/template or perfusion/ventilation quotient lower than 0.75) was calculated. These sums were divided by the total number of pixels for the corresponding segments. The resulting 18 values were used as input to the neural network.

Artificial neural network. A general discussion of artificial neural networks and their workings may be found in the work of Cross et al. [8]. A more detailed description has been presented by Bishop [9]. For the present study, a multilayer perceptron architecture was used with one input layer, one hidden layer and one output layer [10].

Three different neural networks were trained, one for the interpretation of perfusion scintigrams alone and the other two for the interpretation of joint perfusion and ventilation scintigrams using six and one ventilation projections, respectively. The structure of the three networks was exactly the same; only the training data differed. The input layer contained one unit for each of the 18 input variables while the hidden layer possessed five units. The output layer consisted of a single unit that encoded, with a value between 0 and 1, the probability for pulmonary embolism. During training, the desired output was 1 for "pulmonary embolism" and 0 for "no pulmonary embolism". The first neural network was trained using the 18 features derived solely from each patient's perfusion study, the 1,087 perfusion studies of the training group being used for this purpose. The second network was trained using the 18 features derived from both the ventilation and perfusion studies of the training group. The third network was trained using the 18 features derived from each patient's perfusion study and the sole posterior ventilation study. The networks were trained using the back-propagation algorithm. To avoid "overtraining" a weight elimination technique was used [11]. The network weights were frozen after the training process had been completed. Each of the three neural networks was thereafter evaluated using the corresponding projections in the test group.

Each network presented an average output value between 0 and 1 for each test case. A threshold in this interval was used above which all values were regarded as consistent with pulmonary embolism. By varying this threshold, a receiver operating characteristic (ROC) curve was obtained. The performance of each network was measured as the area under the ROC curve, calculated according to the standard method [12]. The 95% confidence limits of the area were estimated by a bootstrap technique [13].

The difference in performance between two networks was measured as the difference in area under the ROC curves. The statistical significance of such an observed area difference was assessed by means of a permutation test (Monte Carlo version) as follows [14]:

A new classification list was created by randomly selecting for each of the 102 test cases either the classification made with the first network or the classification made with the other network. A second list was created from the classification not included in the first one. The two lists were used to construct two ROC curves, and the areas under the curves were calculated, as was the area difference (test statistic). The procedure was repeated 100,000 times. The relative frequency of area differences that had an absolute value greater than the actual difference was taken as the probability of obtaining at least the actual area difference if no true difference existed. The obtained *P* values are used to find statistical significant differences between the performances of the three classification methods used in this paper. Training and testing of the artificial neural network was undertaken using the JETNET 3.0 package [15].

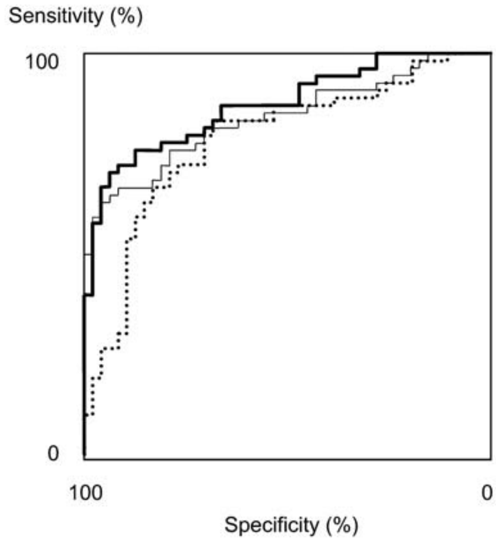

Fig. 1. Three ROC curves presenting the performance of the neural networks interpreting perfusion scintigrams alone (*dotted line*), perfusion scintigrams together with one ventilation image (posterior view) (*thin line*), and perfusion scintigrams together with all ventilation studies (*thick line*)

Results

The performances of the three neural networks trained to detect pulmonary embolism using different sets of lung scintigrams are presented in Fig. 1. The performance of the neural network interpreting perfusion scintigrams alone was 0.79 (95% confidence limits 0.71–0.86). When one ventilation image (posterior view) was added to the perfusion studies, the performance increased to 0.84 (0.77–0.90). This increase was statistically significant (P=0.022). The performance increased to 0.87 (0.81–0.93) when all ventilation studies were used together with the corresponding perfusion studies. The increase in performance from 0.79 to 0.87 was statistically significant (P=0.016), but the increase from 0.84 to 0.87 was not significant.

Discussion

The main finding of this study is that the automated method presented here for the interpretation of lung scintigrams shows a significant increase in performance when one or all ventilation images are added to the six perfusion images. Thus, the ventilation study contributes significantly to the diagnosis of acute lung embolism.

The effect of ventilation images on observer interpretation of lung perfusion examinations was investigated by McLaughlin and co-workers in 1977 [16]. They found a statistically significant reduction in the diagnosis of pulmonary embolism with the addition of the ventilation image. The value of ventilation/perfusion lung scans in the diagnosis of acute pulmonary embolism was also described in the results of the collaborative study of the Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED) [17]. The diagnosis of pulmonary embolism was made on the basis of the number of mismatched segmental equivalent perfusion defects. Ninety-eight patients were randomly selected from the study population of 1,389 patients in order to assess the value of ventilation/perfusion lung scans compared with that of perfusion scans alone in the diagnosis of acute pulmonary embolism [18]. The perfusion scans were read independently of, as well as in combination with, the ventilation scans. The sensitivity of high-probability ventilation/perfusion scans did not differ from the high-probability perfusion scans and the specificity was identical.

The frequency of pulmonary embolism did not differ among patients with a low-probability ventilation/perfusion scan or a low-probability perfusion scan alone. There was a tendency for patients with perfusion scans alone to have a higher percentage of indeterminate readings. The authors concluded that reliable information can be obtained if the interpretation of the perfusion is high or low probability, or near normal/normal. If the perfusion scan is interpreted as intermediate probability for pulmonary embolism according to the PIOPED criteria [17], the addition of a ventilation study may change the interpretation to a more definitive probability.

In a study performed by Miniati and co-workers [3], the value of perfusion lung scan in the diagnosis of pulmonary embolism was evaluated using different scintigraphic interpretation criteria than in the PIOPED study [17]. Only perfusion scans were used and the diagnosis of lung embolism was made if single or multiple wedge-shaped perfusion defects existed, usually together with wedge-shaped areas of overperfusion. The sensitivity and specificity were 92% and 87%, respectively. The authors concluded that an accurate diagnosis of pulmonary embolism is possible by perfusion scanning alone, without ventilation imaging. As no ventilation scintigrams were performed, the additional value of such a scan could not be established.

Pulmonary embolism can be diagnosed automatically by a neural network which has been trained on either the combination of ventilation and perfusion scans or perfusion scans alone [4, 5]. As the method is not dependent upon manual feeding of subjective scintigraphic, radiographic or clinical information into the computer, the results are completely objective. At the same time, information from, for example, chest X-ray or medical history that potentially could have contributed to the diagnosis was not taken into account. In this study we have shown that the addition of a complete ventilation study to the perfusion study increases the performance of the network significantly. It can be assumed that the ventilation study does not contribute much to the diagnosis if the perfusion study is normal. On the other hand, it can be assumed that in patients with primarily ventilation disorders, such as obstructive lung disease, when ventilation and perfusion typically show the same (matching) pattern, the ventilation study should contribute to the diagnosis. This has, however, not been evaluated in the present study.

The neural networks used in this study were trained on a limited number of features from the scintigrams representing differences in perfusion and ventilation, or, when ventilation was missing, differences in perfusion compared with an ideal perfusion image. All features were gained in a completely automated fashion and could be modified or extended. The features in the current setting were in several respects close to those used in conventional evaluation of ventilation/perfusion lung scans in the diagnosis of acute pulmonary embolism, such as the PIOPED study [17]. Therefore, it is not surprising that the addition of the ventilation study resulted in a significantly better performance of the neural network, reflecting the findings of Stein et al. in the selected patient material from the PIOPED study [18].

Different features from the images can be extracted automatically for training of neural networks. Recently, Tourassi et al. [19] presented an automated texture analysis of regions of interest in the perfusion study. The complexity or "roughness" of each image was calculated. The method was then tested for the diagnosis of pulmonary embolism using a semi-automatic computer aid, by interpreting the textural features from perfusion scintigrams from 45 patients [20]. No information from the ventilation images was used. Performance of the computer-assisted diagnostic tool was shown to be substantially better than that of the average clinician and comparable to that of an experienced nuclear medicine physician. In the present study, we did not make any comparison with interpretations by physicians. This was done in an earlier study [5].

Scott and co-workers have also recently published two studies of automated artificial neural network image interpretion in which texture analysis, including both ventilation and perfusion features, was used [21, 22]. In contrast to Tourassi et al., Scott and co-workers extracted features representing the average size and degree of mismatch as well as measures of lung size and the vertical centroid. The authors concluded that computers can perform comparably to experienced observers in patients with normal findings on chest radiographs.

Thus, pulmonary embolism can be diagnosed quickly and without variability by the use of image processing and artificial neural networks. We have previously presented the results of networks trained on either joint ventilation and perfusion scintigrams or perfusion scintigrams alone [4, 5]. To the authors' knowledge, the value of the ventilation study in automated interpretation has not been studied before.

An advantage of using artificial networks in the interpretation of lung scintigrams is that they are good at pattern recognition. The use of the network requires a training set of scintigrams with known classification, e.g. pulmonary embolism or not. In this study, the classification was made by experts in nuclear medicine. There may have been cases in the training group with incorrect classification, which may have resulted in a less than optimal performance of the network when tested on the test material. In addition, we did not investigate any subgroups of the cases taking into account, for example, the location or extent of any lung emboli. However, this could not have interfered with the results when perfusion scintigrams and the combination of ventilation and perfusion scintigrams were used.

In conclusion, we have shown that pulmonary embolism can be diagnosed automatically by a neural network. The network shows a significant increase in performance when one or all ventilation images are added to the six perfusion images. Consequently, the ventilation study has a significant role in the diagnosis of acute lung embolism.

Acknowledgements. This study was supported by grants from the Swedish Medical Research Council (09893 and 10841).

References

1. Alderson PO, Rujanavech, N, Secker-Walker RH, McKnight RC. The role of ^{133}Xe ventilation studies in the scintigraphic detection of pulmonary embolism. *Radiology* 1976; 120:633–640.
2. Goris ML, Daspit SG. Krypton-81m ventilation scintigraphy for the diagnosis of pulmonary embolism. *Clin Nucl Med* 1981; 6:207–212.
3. Miniati M, Pistolesi M, Marini C, et al. Value of perfusion lung scan in the diagnosis of pulmonary embolism: results of the prospective investigative study of acute pulmonary embolism diagnosis (PISA-PED). *Am J Respir Crit Care Med* 1996; 154:1387–1393.
4. Holst H, Åström K, Järund A, et al. Automated interpretation of ventilation-perfusion lung scintigrams for the diagnosis of pulmonary embolism using artificial neural networks. *Eur J Nucl Med* 2000; 27:400–406.
5. Holst H, Måre K, Järund A, et al. An independent evaluation of a new method for automated interpretation of lung scintigrams using artificial neural networks. *Eur J Nucl Med* 2001; 28:33–38.
6. Tägil K, Evander E, Wollmer P, Palmer J, Jonson B. Efficient lung scintigraphy. *Clin Physiol* 2000; 20:95–100.
7. Mandell CH. *Scintillation camera lung imaging: an anatomic atlas and guid*e. New York: Grune and Stratton, 1976.
8. Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. *Lancet* 1995; 346:1075–1079.
9. Bishop CM. *Neural networks for pattern recognition*. Oxford: Oxford University Press, 1995.
10. Rumelhart DE, Hinton GE, Williams RJ. Learning internal representation by error propagation. In: Rumelhart DE, McCleland JL, eds. *Parallel distributed processing: explorations in the microstructure of cognition.* Cambridge, Mass: MIT Press/Bradford Books; 1986:318–363.
11. Hanson SJ, Pratt LY. Comparing biases for minimal network construction with backpropagation. In: Touretzky DS, ed. *Advances in neural information processing systems.* San Mateo, Ca.: Morgan Kaufmann; 1989:177–185.
12. Woods K, Bowyer KW. Generating ROC curves for artificial neural networks. *IEEE Trans Med Imaging* 1997; 16:329–337.
13. Efron B. Computers and the theory of statistics: thinking the unthinkable. *Siam Rev* 1979; 21:460–480.
14. Wehrens R, Putter H, Buydens LMC. The bootstrap: a tutorial. *Chemom Intell Lab Syst* 2000; 54:35–52.
15. Peterson C, Rögnvaldsson T, Lönnblad L. JETNET 3.0: a versatile artificial neural network package. *Comp Phys Comm* 1994; 81:185–220.
16. McLaughlin GL, Burt RW, DePalma B, Gubler R. Effect of ventilation images on observer interpretation of lung perfusion examinations. *Am J Roentgenol* 1977; 128:1037–1038.
17. The PIOPED Investigators. Value of the ventilation/perfusion scan in acute pulmonary embolism: results of the Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED*). JAMA* 1990; 263:2753–2759.
18. Stein PD, Terrin ML, Gottschalk A, Alavi A, Henry J. Value of ventilation/perfusion scans versus perfusion scans alone in acute pulmonary embolism. *Am J Cardiol* 1992; 69:1239–1241.
19. Tourassi GD, Frederick ED, Vittitoe NF, Coleman RE. Fractal texture analysis of perfusion lung scans. *Comput Biomed Res* 2000; 33:161–171.
20. Tourassi GD, Frederick ED, Floyd CE Jr, Coleman RE. Multifractal texture analysis of perfusion lung scans as a potential diagnostic tool for acute pulmonary embolism. *Comput Biol Med* 2001; 31:15–25.
21. Scott JA. Using artificial neural network analysis of global ventilation-perfusion scan morphometry as a diagnostic tool. *Am J Roentgenol* 1999; 173:943–948.
22. Scott JA, Palmer EL, Fischman AJ. How well can radiologists using neural network software diagnose pulmonary embolism? *Am J Roentgenol* 2000; 175:399–405.