How to Use a Gestalt Interpretation for Ventilation–Perfusion Lung Scintigraphy

Petronella J. Hagen, MD[1]; Ieneke J.C. Hartmann, MD[2,3]; Otto S. Hoekstra, MD, PhD[4,5];
Marcel P.M. Stokkel, MD, PhD[3]; Gerrit J.J. Teule, MD, PhD[4]; and Martin H. Prins, MD, PhD[6],
for the ANTELOPE Study Group

[1]Department of Pulmonary Medicine, Vrije Universiteit Medical Center, Amsterdam, The Netherlands; [2]Department of Radiology, University Medical Center, Utrecht, The Netherlands; [3]Department of Nuclear Medicine, University Medical Center, Utrecht, The Netherlands; [4]Department of Nuclear Medicine, Vrije Universiteit Medical Center, Amsterdam, The Netherlands; [5]Department of Clinical Epidemiology, Vrije Universiteit Medical Center, Amsterdam, The Netherlands; and [6]Department of Clinical Epidemiology and Medical Technology Assessment, Academic Hospital Maastricht, Maastricht, The Netherlands

The use of a so-called gestalt interpretation, an integration of different sets of criteria and the physician's own experience, has been advocated in the interpretation of lung scintigraphs of patients with clinically suspected pulmonary embolism. However, data on the reliability of this approach are limited. The aim of this study was to investigate the observer variability and accuracy of the gestalt interpretation of perfusion scintigraphy (combined with chest radiography) as well as the impact of adding ventilation scintigraphy and clinical pretest information. **Methods:** Three experienced observers independently reviewed the chest radiograph and ventilation–perfusion scans of 101 consecutive patients with clinically suspected pulmonary embolism. All datasets were reviewed twice by each observer, using a visual analog scale to indicate the estimated probability of pulmonary embolism. The results of the gestalt interpretations were analyzed against the presence or absence of pulmonary embolism. **Results:** All 3 gestalt interpretations had a good-to-excellent interobserver variability (intraclass correlation coefficient [ICC], 0.73–0.89), with similar intraobserver agreement (ICC, 0.76–0.95). The performance of all 3 readers was comparable. The areas under the curve (AUCs) of all 3 observers were high and similar (for observer 1, the AUCs were 0.96 [95% confidence interval (CI)], 0.93–1.00), 0.96 (95% CI, 0.93–1.00), and 0.95 (95% CI, 0.90–1.00), respectively, for the 3 gestalt interpretations). **Conclusion:** A gestalt interpretation is a useful classification scheme with good-to-excellent intra- and interobserver variability. However, the interpretation and the consequences of this result are dependent on the observer. Unexpectedly, the addition of information on ventilation scintigraphy and clinical information did not affect the overall assessment.

Key Words: ventilation–perfusion lung scintigraphy; pulmonary embolism; gestalt interpretation

J Nucl Med 2002; 43:1317–1323

Lung scintigraphy is widely applied as a noninvasive and readily available technique in patients with clinically suspected pulmonary embolism. However, a diagnosis of pulmonary embolism with ventilation–perfusion (V/Q) scintigraphy is based on the presence of perfusion defects relative to the ventilation rather than on a direct visualization of the embolus, as is the case with pulmonary angiography or spiral CT angiography. Therefore, the result of V/Q lung scintigraphy essentially can be regarded as a probability estimate of the presence or absence of pulmonary embolism that is based on the number and size of defects. Several well-defined interpretation schemes are used that divide V/Q scan probabilities in 3 or 4 categories from normal to high probability (1–3). Indeed, several large clinical studies have shown that a normal scan virtually excludes clinically significant pulmonary embolism (remaining incidence, 0%–2%), whereas 80%–90% of the patients with a high-probability lung scan do have pulmonary embolism (4–6). However, a considerable category will be in the midrange probability, in which the diagnosis cannot be rejected.

Many physicians do not adhere rigidly to a single diagnostic scheme but appear to use integrated knowledge of published algorithms, ancillary findings, clinical data, and complex interrelationships in a so-called gestalt interpretation (5). The term "gestalt" stems from psychology and describes perceptions as a system of phenomena so integrated as to constitute a functional unit with properties not derivable from its parts. With the gestalt interpretation, the experienced nuclear medicine physician may be able to provide a more accurate interpretation of the lung scan than is provided by criteria alone. It has been suggested that experienced nuclear medicine physicians should incorporate this interpretation into their final report—for example, as percentage of probability (3,5). The gestalt interpretation might be influenced by information about the patient's history and risk factors. In a large database we found that

the 3 most commonly used sets of criteria (the revised Prospective Investigation of Pulmonary Embolism Diagnosis [PIOPED] criteria, the Hull criteria, and a gestalt interpretation) had a similar accuracy (7). Because gestalt reading requires integration of several sources of knowledge, including experience, data on observer variability are essential. However, data on the observer agreement and accuracy of gestalt reading are limited.

The aim of this study was to investigate the inter- and intraobserver variability and accuracy of the gestalt interpretation in 3 observers with a different level of experience. Furthermore, the effect of adding ventilation scintigraphy and clinical information on the accuracy and observer agreement of the gestalt interpretation was investigated.

MATERIALS AND METHODS

This study was a reevaluation of the V/Q lung scintigraphs from a prospective study on diagnostic methods in patients with clinically suspected pulmonary embolism between May 1997 and October 1998 (8–12). The chest radiographs and V/Q scintigraphs of 101 consecutive patients of 1 participating center were reviewed.

Study Protocol

Before any further testing, the attending physician gave a probability estimate for pulmonary embolism based on evaluation of the clinical history, physical examination, chest radiograph, and electrocardiogram. This probability estimate was performed with a visual analog scale of 0% to 100%. Within 24 h of referral, venous duplex sonography, a D-dimer blood test, and perfusion lung scintigraphy were performed. Patients were stratified according to the lung scan result. In case of a normal perfusion scan, the investigations were stopped and no further tests were performed. Ventilation scintigraphy was indicated for all patients with at least 1 segmental perfusion defect. Spiral CT angiography was performed on all patients with perfusion defects, irrespective of the size of these defects. Pulmonary angiography was performed on all patients with a "nonhigh" probability lung scan result and on patients with discordance between V/Q scintigraphy and spiral CT angiography. The study protocol was violated in patients who had a contraindication for spiral CT angiography or pulmonary angiography. These patients were not excluded from the study. We aimed at performing the complete study protocol within 48 h after the first perfusion scan, with a maximum of 24 h between the examinations under study. The diagnosis of pulmonary embolism was made on the basis of pulmonary angiography, as the strongest source of evidence, or a high-probability lung scan result. A normal perfusion scan or pulmonary angiogram ruled out pulmonary embolism. In all cases, the final diagnosis was established by independent, blind reading of the diagnostic images.
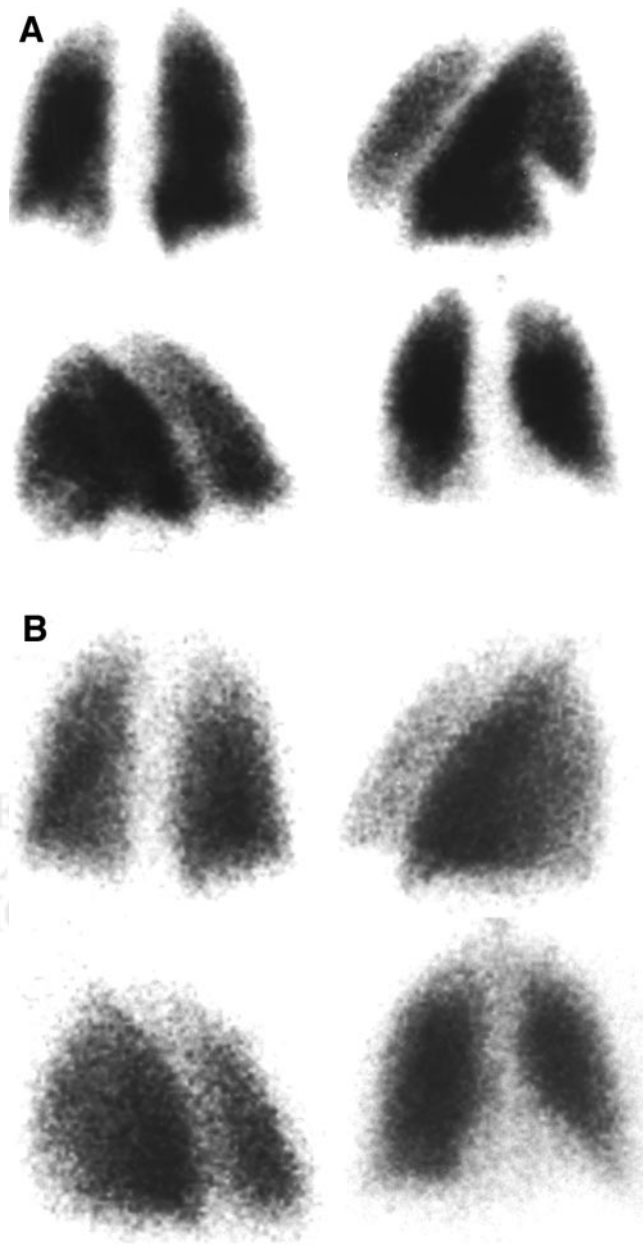
Perfusion Scintigraphy

Perfusion lung scintigraphy was performed using 50 MBq ^{99m}Tc-labeled macroaggregated albumin. The tracer was injected intravenously in the supine position, whereas imaging was performed in the sitting position. Acquisition was performed in at least 4 standard positions (anterior, posterior, left and right posterior oblique) with at least 150 kilocounts per view (low-energy, high-resolution collimator, 128 × 128 matrix). In most cases, lateral views were also obtained.

^{81m}Kr Ventilation Scintigraphy

Inhalation imaging with ^{81m}Kr was performed either immediately after perfusion scintigraphy or using dual-isotope scanning. In case ^{81m}Kr was not available, inhalation imaging was executed the next day, but at least within 24 h. Each image was made with at least 22 kilocounts per view. Ventilation scans were obtained in the same projections as the perfusion scans. An example of V/Q scintigraphy of a 35-y-old patient with dyspnea and pleuritic chest pain is given in Figure 1.

V/Q Scintigraphy Assessment

After the clinical part of the study was completed, 3 experienced observers reinterpreted all scans, independently and unaware of the

FIGURE 1. A 35-y-old patient with acute dyspnea and pleuritic chest pain. V/Q scintigraphy shows 1 segmental perfusion defect with mismatch. (A) Perfusion scintigraphy. (B) Ventilation scintigraphy.

results of the consensus reading of the V/Q scans and the results of pulmonary angiography or spiral CT. The observers had 2-, 9-, and 15-y' experience in practice after their training. They were all experienced in application of the Hull and (revised) PIOPED criteria. In all sessions, a lung segment reference chart was available (*13*).

In each session, scans were interpreted according to only 1 of the 3 gestalt interpretations. For this gestalt interpretation the observers were asked to make a probability estimate of the presence of pulmonary embolism on a visual analog scale of 0% to 100%. In the first reading, the observer was asked to give the percentage of probability based on the perfusion lung scan and the chest radiograph. In the second gestalt reading, the clinical information (malignancy, prior surgery, chronic obstructive pulmonary disease, history of venous thromboembolism, productive cough, chest pain, dyspnea, or hemoptysis) was added to the perfusion scan and chest radiograph. Finally, in the third reading, the observer gave a percentage of probability based on the V/Q scintigraphy, chest radiography, and clinical information. The interval between the reading sessions was at least 4 wk and the order of lung scans was randomized. To measure the intraobserver agreement the procedure was repeated after a 6-mo interval. The gestalt probability estimate was based solely on the personal experience and opinion of the observers.

Statistics

For all 3 gestalt interpretations, the inter- and the intraobserver agreement was assessed using intraclass correlation coefficients (ICCs).

The probability estimates obtained for the different gestalt interpretations were compared with the presence or absence of pulmonary embolism according to the final diagnosis. For this purpose, receiver operating characteristic (ROC) analysis and areas under the ROC curve (AUCs) were used as objective measures to evaluate the overall accuracy of the 3 different gestalt interpretations (*14*). In addition, sensitivities and specificities of the gestalt interpretations were calculated using conventional cutoffs of 20%, 50%, and 80%.

All statistical analyses were performed with SPSS software (SPSS, Inc., Chicago, IL) (*15*).

RESULTS

Patient Population

For this study the V/Q scintigraphs of 101 patients evaluated prospectively for clinically suspected pulmonary embolism were used. The demographic and clinical characteristics of these patients are given in Table 1.

For all 101 patients, the 3 gestalt interpretations of the 3 observers were available. For 81 of the 101 patients, a final diagnosis with regard to the presence or absence of pulmonary embolism was available. Of these 81 patients, 25 (31%) had a diagnosis of pulmonary embolism.

Inter- and Intraobserver Agreement of 3 Gestalt Interpretations

All 3 gestalt interpretations (*n* = 101) showed good-to-excellent interobserver variability. For the gestalt interpretation based only on perfusion scintigraphy and the chest radiograph, the ICC varied between 0.73 and 0.80, whereas

TABLE 1
Demographic and Clinical Characteristics of 101 Study Patients with Clinically Suspected Pulmonary Embolism

Characteristic	Study patients (*n* = 101)
Female	54 (53%)
Mean age ± SD (y)	57.0 ± 17.5
Inpatient	21 (21%)
Median duration of symptoms (d) (interquartile range)	2 (1–9)
History of PE/DVT*	10/100 (10%)
COPD*	16/101 (16%)
Malignancy*	14/99 (14%)
Surgery*	19/101 (19%)

*Numerator varies because data of some patients were missing. PE/DVT = pulmonary embolism/deep vein thrombosis; COPD = chronic obstructive pulmonary disease.

this statistic was 0.79–0.84 when clinical information was added and 0.79–0.89 when ventilation scintigraphy was made available. The scores of observers 1 and 2 for the 3 gestalt interpretations are depicted in Figures 2A–2C. Other interobserver graphs gave similar results (data not shown).

The intraobserver variability was 0.76–0.92 for the gestalt reading based only on perfusion scan and chest radiograph. For the gestalt reading in which clinical information was added, the intraobserver variability was 0.88–0.92; adding ventilation scintigraphy resulted in an intraobserver variability of the gestalt interpretation of 0.93–0.95. The performance of all 3 readers was comparable. The scores of the first and second readings of observer 1 for the 3 gestalt interpretations are depicted in Figures 2A–2C. Other intraobserver graphs gave similar results (data not shown).

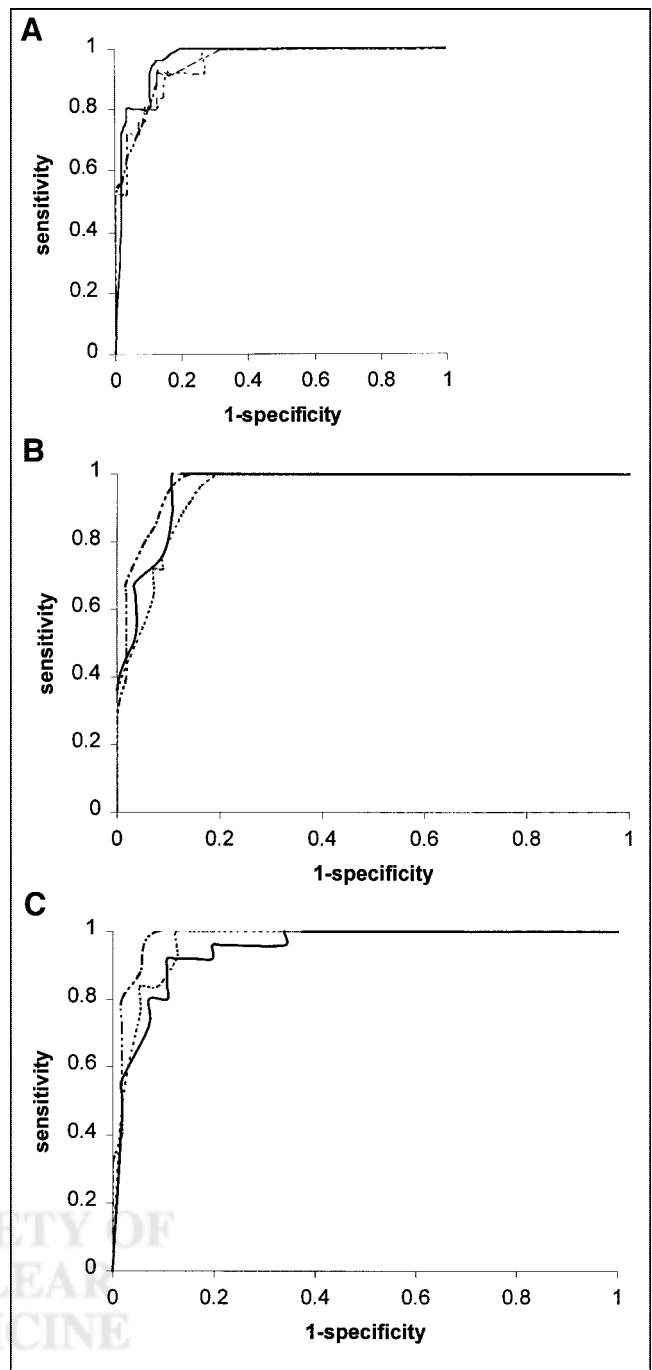### Accuracy of 3 Gestalt Interpretations

ROC curves for the 3 different gestalt interpretations (*n* = 81) were calculated for all observers and are shown in Figures 3A–3C. The AUCs for the 3 gestalt interpretations of all 3 observers were high and similar. The AUC of the gestalt interpretation based on perfusion scintigraphy and chest radiography of observer 1 was 0.96 (95% confidence interval [CI], 0.93–1.00); adding clinical information resulted in an AUC of 0.96 (95% CI, 0.93–1.00), whereas the AUC after addition of the ventilation scan was 0.95 (95% CI, 0.90–1.00). For observer 2, these values were 0.95, 0.95, and 0.96, respectively, whereas the AUCs of observer 3 were 0.95, 0.97, and 0.98, respectively.

For 2 observers, we found a trend toward higher intraobserver agreement when additional information (clinical data or ventilation scintigraphy) was added. A similar trend was seen for interobserver variability (Table 2).

In Table 3, the sensitivities and specificities of the complete gestalt interpretation to which these ROC curves translate at specific cutoff probabilities for each of the 3 observers are given.
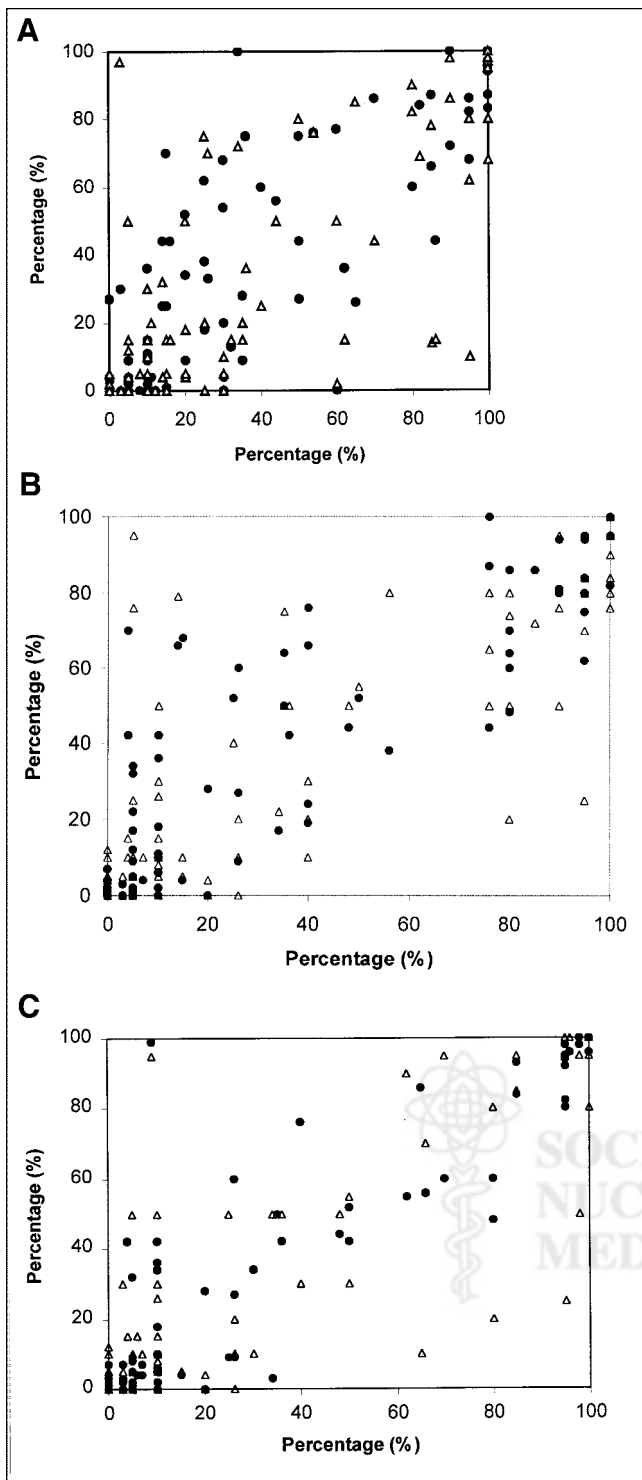
FIGURE 2. Intraobserver variability of observer 1 and interobserver variability of observers 1 and 2 of 3 different gestalt interpretations. (A) Gestalt interpretation of perfusion scintigraphy and chest radiography. (B) Gestalt interpretation of perfusion scintigraphy, chest radiography, and clinical information. (C) Gestalt interpretation of V/Q scintigraphy, chest radiography, and clinical information. ●, Intraobserver variability of observer 1; △, interobserver variability of observers 1 and 2.

FIGURE 3. ROC curves of observers 1–3 of 3 different gestalt interpretations. (A) Gestalt interpretation of perfusion scintigraphy and chest radiography. (B) Gestalt interpretation of perfusion scintigraphy, chest radiography, and clinical information. (C) Gestalt interpretation of V/Q scintigraphy, chest radiography, and clinical information. Solid line, observer 1; dotted line, observer 2; dashed-dotted line, observer 3.

DISCUSSION

In this study 3 different gestalt interpretations available in V/Q lung scintigraphy were compared in consecutive patients with suspected pulmonary embolism. All 3 interpretations had a good-to-excellent inter- and intraobserver vari-

TABLE 2
Intraobserver and Interobserver Variability of 3 Different
Gestalt Interpretations (ICC) with 95% CIs

Observer	Gestalt XQ (95% CI)	Gestalt XQ + info (95% CI)	Gestalt XQ + V + info (95% CI)
	Intraobserver variability		
1	0.87 (0.81–0.91)	0.90 (0.85–0.93)	0.94 (0.91–0.96)
2	0.92 (0.88–0.94)	0.92 (0.88–0.95)	0.93 (0.90–0.95)
3	0.76 (0.67–0.83)	0.88 (0.83–0.92)	0.95 (0.93–0.97)
	Interobserver variability		
1 and 2	0.80 (0.71–0.86)	0.84 (0.77–0.89)	0.89 (0.85–0.93)
2 and 3	0.79 (0.71–0.85)	0.79 (0.70–0.85)	0.84 (0.77–0.89)
3 and 1	0.73 (0.62–0.81)	0.70 (0.71–0.86)	0.79 (0.71–0.86)

XQ = chest radiography and perfusion scintigraphy; info = clinical information; V = ventilation scintigraphy.

ation (ICC, 0.73–0.95). The accuracy of the gestalt interpretation based on perfusion scintigraphy and chest radiography was good (AUC, 0.95–0.96) and did not improve when clinical information or ventilation scintigraphy (or both) was added (0.95–0.97 and 0.95–0.98, respectively).

The gestalt theory is based on the principle that the entirety is more than the sum of the different parts. Thus, the combination of, for example, radiographic and scintigraphic abnormalities may have quite different implications than the same findings when present alone. Also, the presence of scintigraphic abnormalities may have different implications if the clinical presentation varies. The PIOPED investigators reported that the combination of a clinical assessment with the lung scan interpretation improved the overall chance of reaching a definitive diagnosis (6). Not surprisingly, experienced nuclear medicine physicians often override reference criteria with their own subjective gestalt, based on extensive experience with reading and interpreting. Conceptually, this approach might provide better results than the standard algorithms because the latter are "distillations of decision making into finite linear steps" (16) that cannot account for complex interrelationships. Published evidence on ancillary scintigraphic findings is clearly teachable, and from time to time criteria are revised

accordingly. Personal experience may be more difficult to transfer to others. However, we found no direct association between experience and accuracy.

After the PIOPED study, Sostman et al. (3) described the gestalt interpretation in the evaluation of V/Q scintigraphy. In a group of 104 patients with clinically suspected pulmonary embolism, the gestalt interpretation was compared with the (revised) PIOPED criteria. Although the gestalt estimate of probability was not statistically significantly better, it yielded the best accuracy for assessing the presence of pulmonary embolism both in the individual and in the consensus readings (area under the ROC curve, 0.78–0.84). The authors suggest that experienced readers should incorporate a likelihood estimate into the final scan report, based on their experience and personal judgment.

In another study, the gestalt interpretation was compared with the McNeil and Biello schemes in 98 patients with suspected pulmonary embolism (17). Again, no statistically significant differences were found in overall accuracy.

Christiansen et al. (18) performed a study in 170 patients with suspected pulmonary embolism. The PIOPED criteria were combined with a probability estimate on a visual linear scale. There was no statistically significant difference in the area under the ROC curve between the PIOPED categorization and the estimate probability. Therefore, they concluded that adding a visual linear scale probability to the PIOPED criteria was not useful.

Fonseca et al. (19) used a simplified gestalt interpretation consisting of a 5-point scale. V/Q scintigraphs of 204 patients were reviewed retrospectively according to this 5-point-scale, gestalt interpretation (using terms as definitely, probable, possible, and uncertain) and the modified Biello criteria. A very low agreement (47.5%) compared with the Biello criteria (agreement, 77%) for this gestalt interpretation was found. One limitation of this study was that pulmonary angiography was available in only 8% of the patients. The difference in interobserver variability was explained by the addition of 1 diagnostic category.

In comparison with these literature observations, our results were similar or even better. Although the ROC curves are approximately comparable, the interpretation of these results differs per observer and is influenced by different cutoffs, as shown in Table 3. Gray et al. (20) showed a

TABLE 3
Sensitivities and Specificities of Complete Gestalt Interpretation (V/Q Scintigraphy, Chest Radiography, and Clinical Information) at Different Cutoffs of 3 Observers

Observer	Cutoff probability					
	20%		50%		80%	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
1	96	79	92	89	80	89
2	100	84	92	88	84	91
3	88	95	72	98	64	98

comparable result in their analysis of the understanding of a verbal probability language used in lung scan reports. A wide variation in the interpretation of this probability language was found. This variation in subjective probabilities may be influenced by individual and institutional variations in factors such as prior training, inherent abilities, local experience, and personal biases. However, as in our results, the individual physicians were very reliable in their interpretation on different occasions.

Also, in comparison with results reported in the literature, the gestalt interpretation performed well, and the lack of a better performance of the gestalt interpretation in comparison with the PIOPED and Hull criteria (7) is, therefore, unlikely to be due to inexperience. Alternatively, the gestalt reading may have been dominated by an almost instantaneous cognitive referral to known diagnostic algorithms.

Our data suggest a limited impact of adding ventilation scintigraphy or clinical data on gestalt reading performance. However, there was a trend toward a higher observer agreement when adding information (Table 2).

Although the results of this study provide a good view on the use of gestalt interpretation, several methodologic aspects deserve comment. For 81 of the 101 patients, a final diagnosis was available, mainly because of differences in the local and central reading of the different investigations, and is unlikely to have influenced the results.

To avoid recall bias, the scans were assessed in random order with a time interval of at least 4 wk. The given data, therefore, are unlikely to be influenced by memory and reflect the variation based on true differences between and within observers. Furthermore, the observers received summarized clinical information about the patient and his or her history, which is certainly different from the direct view of a patient, as is the case in daily practice. Therefore, it remains to be shown whether a similar lack of a clear effect of having clinical information on interobserver agreement will be found in a prospective study. Obviously, the acute presentation of the disease does not allow evaluation of intraobserver variation.

Finally, with only 3 experienced observers it was not possible to confirm that the gestalt interpretation is influenced by experience, as is suggested in the literature.

CONCLUSION

A gestalt interpretation for perfusion scintigraphy in patients with suspected pulmonary embolism is a useful classification scheme with good-to-excellent intra- and interobserver variability. However, the interpretation and the consequences of this result do seem to be dependent on the observer. Furthermore, addition of clinical information or a ventilation study did not improve our results.

The differences between observers in sensitivities and specificities, based on the absolute percentages as reported, suggest that knowledge of their own operating characteristics in gestalt interpretation of lung scintigraphy could be valuable to individual observers. Because most observers will not have this knowledge, in daily practice, the use of the predefined Hull or PIOPED criteria will be less dependent on the observer.

APPENDIX

The results of this study are reported on behalf of the ANTELOPE (Advances in New Technologies Evaluating the Localization of Pulmonary Embolism) Study Group of the Dutch prospective multicenter trial on the diagnosis of pulmonary embolism.

Academic Medical Center, Amsterdam, The Netherlands

B.J. Sanson, MD; M.H. Prins, MD; H.R. Büller, MD

Leiden University Medical Center, Leiden, The Netherlands

W. de Monyé, MD; M.V. Huisman, MD; P.M.T. Pattynama, MD

Leyenburgh Hospital, The Hague, The Netherlands

M.J.L. van Strijen, MD; G.J. Kieft, MD

Slotervaart Hospital, Amsterdam, The Netherlands

M.R. Mac Gillavry, MD; D.P.M. Brandjes, MD

Vrije Universiteit Medical Center, Amsterdam, The Netherlands

P.J. Hagen, MD; O.S. Hoekstra, MD; P.E. Postmus, MD

University Medical Center, Utrecht, The Netherlands

I.J.C. Hartmann, MD; J.D. Banga, MD; P.F.G.M. van Waes, MD

ACKNOWLEDGMENT

REFERENCES

1. Biello DR, Mattar AG, McKnight RC, Siegel BA. Ventilation-perfusion studies in suspected pulmonary embolism. *AJR*. 1979;133:1033–1037.
2. Hull RD, Hirsh J, Carter CJ, et al. Diagnostic value of ventilation-perfusion lung scanning in patients with suspected pulmonary embolism. *Chest*. 1985;88:819–828.
3. Sostman HD, Coleman RE, DeLong DM, Newman GE, Paine S. Evaluation of revised criteria for ventilation-perfusion scintigraphy in patients with suspected pulmonary embolism. *Radiology*. 1994;193:103–107.
4. Alderson PO. Scintigraphic diagnosis of pulmonary embolism: where do we go from here? *Radiology*. 1994;193:22–23.
5. Gottschalk A, Sostman HD, Coleman RE, et al. Ventilation-perfusion scintigraphy in the PIOPED study. Part II. Evaluation of the scintigraphic criteria and interpretations. *J Nucl Med*. 1993;34:1119–1126.
6. 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.
7. Hagen PJ, Hartmann IJ, Hoekstra OS, Stokkel MP, Postmus PE, Prins MH. Different criteria for interpretation of ventilation perfusion lung scintigraphy [abstract]. 11th Annual Meeting of the European Respiratory Society; September 22–26, 2001; Berlin, Germany. Abstract P1565.
8. de Monye W, van Strijen MJ, Huisman MV, Kieft GJ, Pattynama PM. Suspected pulmonary embolism: prevalence and anatomic distribution in 487 consecutive patients: Advances in New Technologies Evaluating the Localisation of Pulmonary Embolism (ANTELOPE) Group. *Radiology*. 2000;215:184–188.

9. Hartmann IJ, Hagen PJ, Melissant CF, Postmus PE, Prins MH. Diagnosing acute pulmonary embolism: effect of chronic obstructive pulmonary disease on the performance of d-dimer testing, ventilation/perfusion scintigraphy, spiral computed tomographic angiography, and conventional angiography. *Am J Respir Crit Care Med.* 2000;162:2232–2237.

10. Mac Gillavry MR, Sanson BJ, Buller HR, Brandjes DP. Compression ultrasonography of the leg veins in patients with clinically suspected pulmonary embolism: is a more extensive assessment of compressibility useful? ANTELOPE-Study Group. *Thromb Haemost.* 2000;84:973–976.

11. Mac Gillavry MR, de Monye W, Lijmer JG, et al. Clinical evaluation of a monoclonal antibody-based enzyme immunoassay for fibrin degradation products in patients with clinically suspected pulmonary embolism. ANTELOPE-Study Group. *Thromb Haemost.* 2000;83:892–895.

12. Sanson BJ, Lijmer JG, Mac Gillavry MR, Turkstra F, Prins MH, Buller HR. Comparison of a clinical probability estimate and two clinical models in patients with suspected pulmonary embolism. ANTELOPE-Study Group. *Thromb Haemost.* 2000;83:199–203.

13. Lensing AW, van Beek EJ, Demers C, et al. Ventilation-perfusion lung scanning and the diagnosis of pulmonary embolism: improvement of observer agreement by the use of a lung segment reference chart. *Thromb Haemost.* 1992;68:245–249.

14. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143:29–36.

15. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* 1983;148:839–843.

16. Freeman LM, Krynyckyi B, Zuckler LS. Enhanced lung scan diagnosis of pulmonary embolism with the use of ancillary scintigraphic findings and clinical correlation. *Semin Nucl Med.* 2001;31:143–157.

17. Sullivan DC, Coleman RE, Mills SR, Ravin CE, Hedlund LW. Lung scan interpretation: effect of different observers and different criteria. *Radiology.* 1983;149:803–807.

18. Christiansen F, Nilsson T, Mare K, Carlsson A. Adding a visual linear scale probability to the PIOPED probability of pulmonary embolism. *Acta Radiol.* 1997;38:458–463.

19. Fonseca RB, Spies WG, Spies SM. A simplified method for "Gestalt" interpretation of ventilation/perfusion (V/Q) scintigrams in suspected pulmonary embolism (PE): comparison with modified Biello criteria and assessment of interobserver variability [abstract]. *J Nucl Med.* 1997;38(suppl):299P.

20. Gray HW, McKillop JH, Bessent RG. Lung scan reporting language: what does it mean? *Nucl Med Commun.* 1993;14:1084–1087.