
Very Low Probability Interpretation of V/Q Lung Scans in Combination with Low Probability Objective Clinical Assessment Reliably Excludes Pulmonary Embolism: Data from PIOPED II

Alexander Gottschalk¹, Paul D. Stein^{2,3}, H. Dirk Sostman⁴, Fadi Matta², and Afzal Beemath²

¹Department of Radiology, Michigan State University, East Lansing, Michigan; ²St. Joseph Mercy Oakland Hospital, Pontiac, Michigan; ³Wayne State University, Detroit, Michigan; and ⁴Office of the Dean, Weill Cornell Medical College and Methodist Hospital, Houston, Texas

Use of a very low probability interpretation of ventilation/perfusion (V/Q) lung scans, if verified by prospective evaluation to have a low positive predictive value (PPV), will reduce the number of nondiagnostic interpretations of V/Q scans and may be particularly useful in patients with a relative contraindication to CT. The purpose of this investigation was to test the hypothesis that a very low probability interpretation of the V/Q scan has a PPV of <10%. **Methods:** Data are from PIOPED II (Prospective Investigation of Pulmonary Embolism Diagnosis II). Very low probability criteria are (a) nonsegmental perfusion abnormalities, (b) perfusion defect smaller than corresponding radiographic lesion, (c) ≥ 2 matched V/Q defects with regionally normal chest radiograph, (d) 1–3 small segmental perfusion defects (<25% of a segment), (e) solitary triple matched defect in middle or upper lung zones, (f) stripe sign around the perfusion defect(s), and (g) perfusion defect from pleural effusion equal to one third or more of the pleural cavity with no other perfusion defect. **Results:** A very low probability consensus interpretation of the V/Q scan was made in 56% of patients. The PPV of a very low probability interpretation of the V/Q scans was 36 of 440 patients (8.2%). Among patients with suspected pulmonary embolism who had a low clinical probability objective clinical assessment and a very low probability V/Q scan, the PPV was 8 of 259 patients (3.1%). Among women ≤ 40 y, the PPV of the very low probability V/Q with a low objective clinical assessment was 1 of 50 (2%). **Conclusion:** The very low probability V/Q scan together with a low probability clinical assessment reliably excludes pulmonary embolism.

Key Words: V/Q scintigraphy; venous thromboembolism; very low probability criteria

J Nucl Med 2007; 48:1411–1415
DOI: 10.2967/jnumed.107.040998

A “low probability” interpretation of a ventilation/perfusion (V/Q) lung scan has been stated to be misleading and even dangerous because of the unacceptably high frequency of pulmonary embolism (PE) among patients with such an interpretation (1). Since the conclusion of the first Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED) (2), retrospective analyses of the PIOPED database identified several abnormalities on the V/Q scan that are associated with <10% probability of PE (3). These criteria are defined as a “very low probability” interpretation. Use of the very low probability interpretation—if verified by the prospective case series from PIOPED II to have such a low positive predictive value (PPV)—would reduce the number of nondiagnostic interpretations of V/Q scans. Pulmonary scintigraphy is still the recommended imaging test of choice in patients with renal failure or serious allergy to iodinated contrast material and in patients in whom the dose of radiation is an important concern (4). Therefore, we evaluated data from PIOPED II to test the hypothesis that a very low probability interpretation of V/Q scans has a PPV of <10%.

MATERIALS AND METHODS

Data are from the PIOPED II investigation, which was a prospective multicenter study to determine the diagnostic validity of multislice contrast-enhanced spiral CT pulmonary angiography (CTA) and the combination of CT angiography with venous phase imaging of the pelvic and proximal leg veins for the diagnosis of acute PE (5). In the present investigation of the PPV of very low probability V/Q scans, we used CTA and digital subtraction angiography (DSA) as the diagnostic reference standards for PE. Excluding inconclusive studies, in PIOPED II, the sensitivity of CTA for acute PE was 150 of 181 (83%) and the specificity was 567 of 592 (96%) (5).

All patients ≥ 18 y old with clinically suspected acute PE, whether inpatients or outpatients—seen at the 8 participating clinical centers between September 2001 and July 2003—were

Received Feb. 19, 2007; revision accepted Apr. 17, 2007.
For correspondence or reprints contact: Alexander Gottschalk, MD, Department of Radiology, Michigan State University, East Lansing, MI 48824-1303.
E-mail: Alex.Gottschalk@radiology.msu.edu
COPYRIGHT © 2007 by the Society of Nuclear Medicine, Inc.

potentially eligible for recruitment. Patients sent for a diagnostic imaging test for PE were identified for recruitment as well as patients for whom the study nurse was aware of a consultation request for suspected PE. Exclusion criteria were an inability to complete tests within 36 h, critical illness or hemodynamic instability, ventilatory support, shock or hypotension, myocardial infarction within 1 mo, ventricular fibrillation or sustained ventricular tachycardia within 24 h, abnormal serum creatinine, chronic renal dialysis (first 14 mo of recruitment only), allergy to contrast material, pregnancy, chronic pulmonary hypertension, treatment with long-term anticoagulants, thrombolytic therapy planned in the next 24 h, inferior vena cava filter, deep vein thrombosis (DVT) of the upper extremity, prisoners, and patients previously enrolled.

Image interpretations of CTA, DSA, and V/Q scans were based on agreement of two PIOPED II readers from centers other than that at which the image was obtained. They were given only the imaging studies with no other clinical or imaging information. The methods for obtaining V/Q scans, pulmonary DSAs, and CTA were described (6). The CTA was obtained with 4-, 8-, or 16-head scanners. Scanners with 4-detector arrays were used in 691 patients. Eight-detector scanners were used in 37 patients, and 16-detector scanners were used in 45 patients (5). The V/Q findings that constitute a very low probability interpretation were defined previously on the basis of retrospective analysis of the data from PIOPED I (3,7–11). This served as the basis for very low probability criteria used in PIOPED II, although some of the criteria were defined somewhat differently.

Criteria for a very low probability of PE in PIOPED II are as follows:

- (a) Nonsegmental perfusion abnormalities. These were enlargement of the heart or hilum, elevated hemidiaphragm, linear atelectasis, or costophrenic angle effusion with no other perfusion defect in either lung.
- (b) Perfusion defect smaller than corresponding radiographic lesion.
- (c) ≥ 2 matched V/Q defects with regionally normal chest radiograph and some areas of normal perfusion elsewhere in the lungs.
- (d) 1–3 small segmental perfusion defects ($< 25\%$ of a segment).
- (e) Solitary triple matched defect (defined as a matched V/Q defect with associated matching chest radiographic opac-

ity) in the middle or upper lung zone confined to a single segment.

- (f) Stripe sign, which consists of a stripe of perfused lung tissue between a perfusion defect and the adjacent pleural surface (best seen on a tangential view).
- (g) Pleural effusion equal to one third or more of the pleural cavity with no other perfusion defect in either lung.

Objective clinical probability was assessed by the Wells test (12). The criteria used for this test are shown in Table 1.

Among patients with a consensus V/Q interpretation of very low probability for PE, we considered PE to be present if DSA showed PE and, in the absence of DSA, the diagnosis was based on CTA.

In addition, an analysis using CTA only as the reference test was performed.

The PIOPED II composite reference standard was not used because the V/Q scan was often a key component of it and, thus, incorporation bias would have been present. We chose to use the DSA result where available, as DSA is the most widely accepted imaging reference standard. If no DSA was available, we used the CTA result. The V/Q scan had some influence on selection of patients to undergo DSA and, therefore, affected the spectrum of patients who had a diagnosis by DSA, but the PIOPED II DSA readings were independent of other imaging or clinical information. The CTA readings and the Wells score determinations likewise were independent data observations. PE was considered absent if DSA was negative or, in the absence of DSA, if CTA was negative. PPV of a very low probability interpretation was calculated as the number of patients with PE who had a very low probability interpretation divided by the total number of patients with a very low probability interpretation. Patients were stratified according to objective clinical probability based on the Wells test (12).

RESULTS

A very low probability consensus interpretation of the V/Q scan was made in 460 of 824 patients with suspected PE (56%). Among these patients, the CTA or DSA obtained was of diagnostic quality in 440 patients, and 430 had an objective clinical assessment according to the Wells scoring system (12). The PPV of a very low probability interpretation

TABLE 1
Clinical Probability of PE*

Clinical features	Score (points)
Clinical signs and symptoms of DVT (objectively measured leg swelling and pain with palpation in deep vein system)	3.0
Heart rate > 100 beats/min	1.5
Immobilization ≥ 3 consecutive days (bedrest except to access bathroom) or surgery in previous 4 wk	1.5
Previous objectively diagnosed PE or DVT	1.5
Hemoptysis	1.0
Malignancy (cancer patients receiving treatment within 6 mo or receiving palliative treatment)	1.0
PE as likely as or more likely than alternative diagnosis (based on history, physical examination, chest radiograph, electrocardiogram, and blood tests)	3.0

*According to Wells et al. (12).

Score: < 2.0 , low probability; 2.0–6.0, moderate probability; > 6.0 , high probability.

of the V/Q scans in patients who had an objective clinical assessment was 32 of 430 patients (7.4%) (Table 1). Among patients with suspected PE who had a low probability clinical assessment (Wells score < 2) and a very low probability V/Q scan interpretation, the PPV was 8 of 262 (3.1%) (Table 2). Among patients with an intermediate-probability clinical assessment (Wells score = 2–6), the PPV of a very low probability V/Q scan was 19 of 151 (12.6%). Among patients with a high probability clinical assessment (Wells score > 6) and a very low probability V/Q interpretation, the PPV was 5 of 17 (29.4%).

Because V/Q scans have been suggested by some to be the imaging test of choice in women of reproductive age, we analyzed the PPV of a very low probability interpretation of the V/Q scan in women ≤ 40 y old. There were 79 women no older than 40 y in PIOPED II with a very low probability V/Q scan interpretation, and 77 also had an objective clinical assessment. The PPV of a very low probability interpretation of the V/Q scans in young women who also had an objective clinical evaluation was 7 of 77 (9.1%) (Table 2). Among women ≤ 40 y old with suspected PE who had a low probability clinical assessment (Wells score < 2) and a very low probability V/Q scan interpretation, the PPV was 1 of 50 (2.0%) (Table 2). Among young women with an intermediate probability clinical assessment (Wells score = 2–6), the PPV of a very low probability V/Q scan was 5 of 21 (23.8%). In young women with a high probability clinical assessment, the PPV for PE was 1 of 6 (16.7%).

Analysis using CTA only as the reference test gave comparable, if somewhat better, results (Table 3).

DISCUSSION

The strengths of this investigation are the meticulous methods of data collection in PIOPED II and the availability of prospectively acquired patient data.

A potential weakness is the possibility that the local V/Q scan reading influenced the selection of DSA according to the study protocol and, therefore, altered the spectrum of disease in the group of patients who underwent DSA. We do not believe this had a significant impact on our results. Although the protocol indicated consideration of DSA in

certain circumstances, the use of DSA was always at the clinical discretion of the attending physician. It is likely that DSA was used in difficult cases, but this would, if anything, strengthen our conclusions, as the best-accepted diagnostic reference standard (DSA) would have been used in the most difficult cases.

In PIOPED II, a very low probability V/Q interpretation was made in 56% of the patients. The importance of this finding is that it increases the number of patients whose V/Q scan result enables enough diagnostic certainty for definitive clinical management. The problem with the PIOPED results was not that the V/Q scan could not be sensitive or specific for the diagnosis or exclusion of acute PE. Rather, it was that only a minority of patients (approximately 25% of the whole patient sample) had a reading that enabled clinical management without further testing. As we have shown, in PIOPED II, a very low probability reading is made frequently, and this category appears to have a negative predictive value comparable to a negative CTA (4).

These data also indicate that the combination of a very low probability interpretation of the V/Q scan in a patient with a low probability clinical assessment by the Wells scoring system (12) reliably excludes PE. In PIOPED II, among 824 patients with a reference test diagnosis and completed CT study, 262 patients (32%) had both a very low probability interpretation and a low probability clinical assessment. Importantly, among women ≤ 40 y of age, in whom radiation from CTA in the opinion of some is a diagnostic concern (13,14), the combination of a very low probability interpretation and a low probability clinical assessment reliably excluded PE (PPV = 2.0%).

How might this result affect clinical care? This question leads naturally to a reappraisal of the role of V/Q scintigraphy in suspected acute PE. A high probability interpretation (2 or more mismatched segmental equivalent perfusion defects in a patient with no prior PE) (2,15) allows the V/Q scan to be interpreted as “PE present” with high specificity and PPV, but low sensitivity (2). We propose that normal perfusion and very low probability interpretations, when combined, allow a “PE absent” interpretation of the V/Q scan. “PE uncertain” is indicated by all other interpretations. We believe this is

TABLE 2
PPV of PE Based on DSA If Done and CT Angiography If DSA Was Not Done

Subjects	Low probability clinical assessment*	Moderate probability clinical assessment†	High probability clinical assessment‡	Total PE+/no. of patients (%)
All patients: very low probability V/Q	8/262 (3.1)	19/151 (12.6)	5/17 (29.4)	32/430 (7.4)
Women ≤ 40 y old: very low probability V/Q	1/50 (2.0)	5/21 (23.8)	1/6 (16.7)	7/77 (9.1)

*Wells score < 2.
†Wells score 2–6.
‡Wells score > 6.
Wells score data are expressed as PE+/no. of patients (%).

TABLE 3
PPV of PE Based on CT Angiography Only

Subjects	Low probability clinical assessment (%) [*]	Moderate probability clinical assessment (%) [†]	High probability clinical assessment (%) [‡]	Total PE+/no. of patients (%)
All patients: very low probability V/Q	10/259 (3.9)	13/140 (9.3)	5/14 (35.7)	28/413 (6.8)
Women ≤ 40 y old: very low probability V/Q	1/50 (2.0)	1/23 (4.3)	1/2 (50.0)	3/75 (4.0)

*Wells score < 2.

†Wells score 2–6.

‡Wells score > 6.

Wells score data are expressed as PE+/no. of patients (%).

comparable to the diagnostic choices usually used when interpreting the CT angiogram as well as to their diagnostic accuracy.

How relevant is this concept? It was shown previously that 89% of the patients with a normal or near-normal chest radiograph could have a definitive scintigraphic diagnosis (PE present or PE absent). (6). Therefore, it seems likely that the normal or near-normal chest radiograph can be used to preselect patients for scintigraphy that can be reliably diagnosed as PE present or PE absent.

There are several differences between our results and the original PIOPED study results with regard to the V/Q scan. There were fewer intermediate probability readings in PIOPED II and many more low, very low, and normal readings. We believe that this is due to 2 independent factors. First, there were no formal criteria for interpreting very low probability in PIOPED. Analysis of the PIOPED database provided the basis for the development of the current criteria (7–11). Second, the frequency of different interpretations of V/Q scans likely will vary depending on the population studied—for example, patients in an intensive care unit would be expected to have a higher proportion of uncertain results. The patient samples in PIOPED and PIOPED II were quite different. Many recent studies of PE diagnosis (16–18) have a composition similar to PIOPED II (e.g., heavily weighted toward outpatients), whereas some studies (19–21) more closely resemble PIOPED (e.g., with more representation of inpatients and the critically ill). Literature on the performance of V/Q scans in different populations is sparse (22), but most nuclear medicine physicians would probably agree that inpatients and critically ill patients are more likely to have nondiagnostic V/Q scans than outpatients.

CONCLUSION

This study has implications with regard to the role of V/Q imaging in patients with suspected acute PE in the current era of CTA. Recognized advantages of V/Q scintigraphy are lower radiation dose than CTA and lack of need for iodinated contrast material. Here, we show that in a large proportion of patients with suspected acute PE who have a very low

probability V/Q scan interpretation combined with a low probability clinical assessment, PE can be reliably excluded in a contemporary, predominantly outpatient population. This gives further support for the selection of V/Q imaging as the pulmonary imaging procedure of choice for patients in whom CTA may be disadvantageous.

ACKNOWLEDGMENTS

We thank the PIOPED II nuclear medicine central readers for their effort: Manuel Brown, M. Edward Coleman, Stanley J. Goldsmith, Christine P. Molnar, Edwin L. Palmer, Henry D. Royal, and James A. Scott. This study was supported by grant HL67453 from the U.S. Department of Health and Human Services, Public Health Services, National Heart, Lung, and Blood Institute, Bethesda, Maryland.

REFERENCES

- Hull RD, Raskob GE. Low-probability lung scan findings: a need for change. *Ann Intern Med.* 1991;114:142–143.
- 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.
- Stein PD, Gottschalk A. Review of criteria appropriate for a very low probability of pulmonary embolism on ventilation-perfusion lung scans: a position paper. *Radiographics.* 2000;20:99–105.
- Stein PD, Woodard PK, Weg JG, et al. Diagnostic pathways in acute pulmonary embolism: recommendations of the PIOPED II investigators. *Am J Med.* 2006;119:1048–1055.
- Stein PD, Fowler SE, Goodman LR, et al. Multidetector computed tomography for acute pulmonary embolism. *N Eng J Med.* 2006;354:2317–2327.
- Gottschalk A, Stein PD, Goodman LR, Sostman HD. Overview of Prospective Investigation of Pulmonary Embolism Diagnosis II. *Semin Nucl Med.* 2002;32:173–182.
- Stein PD, Relyea B, Gottschalk A. Evaluation of the positive predictive value of specific criteria used for the assessment of low probability ventilation/perfusion lung scans. *J Nucl Med.* 1996;37:577–581.
- Stein PD, Henry JW, Gottschalk A. Small perfusion defects in suspected pulmonary embolism. *J Nucl Med.* 1996;37:1313–1316.
- Gottschalk A, Stein PD, Henry JW, Relyea B. Matched ventilation/perfusion defects and chest radiographic abnormalities: re-evaluation of the triple matched defect in the assessment of acute pulmonary embolism. *J Nucl Med.* 1996;37:1636–1638.
- Sostman HD, Gottschalk A. The stripe sign: a new sign for diagnosis of nonembolic defects on pulmonary perfusion scintigraphy. *Radiology.* 1982;142:737–741.
- Sostman HD, Gottschalk A. Prospective validation of the stripe sign in ventilation-perfusion scintigraphy. *Radiology.* 1992;184:455–459.

12. Wells PS, Anderson DR, Roger M, et al. Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and D-dimer. *Ann Intern Med.* 2001;135:98–107.
13. Parker MS, Hui FK, Camacho MA, Chung JK, Broga DW, Sethi N. Female breast radiation exposure during CT pulmonary angiography. *AJR.* 2005;185:1228–1233.
14. Milne ENC. Female breast radiation exposure [letter]. *AJR.* 2006;186:E24.
15. Gottschalk A, Sostman HD, Coleman RE, et al. Ventilation-perfusion scintigraphy in the PLOPED study. Part II. Evaluation of scintigraphic criteria and interpretations. *J Nucl Med.* 1993;34:1119–1126.
16. Perrier A, Howarth N, Didier D, et al. Performance of helical computed tomography in unselected outpatients with suspected pulmonary embolism. *Ann Intern Med.* 2001;135:88–97.
17. Perrier A, Roy P-M, Sanchez O, et al. Multidetector-row computed tomography in suspected pulmonary embolism. *N Engl J Med.* 2005;352:1760–1768.
18. The Christopher Study Investigators. Effectiveness of managing suspected pulmonary embolism using an algorithm combining clinical probability, D-dimer testing, and computed tomography. *JAMA.* 2006;295:172–179.
19. van Strijen MJL, de Monye W, Schiereck J, et al. Single-detector helical computed tomography as the primary diagnostic test in suspected pulmonary embolism: a multicenter clinical management study of 510 patients. *Ann Int Med.* 2003;138:307–314.
20. Revel MP, Petrover D, Hernigou A, Lefort C, Meyer G, Fria G. Diagnosing pulmonary embolism with four-detector row helical CT: prospective evaluation of 216 outpatients and inpatients. *Radiology.* 2005;234:265–273.
21. Kelly AM, Patel S, Carlos RC, Cronin P, Kazerooni EA. Multidetector row CT pulmonary angiography and indirect venography for the diagnosis of venous thromboembolic disease in intensive care unit patients. *Acad Radiol.* 2006;13:486–495.
22. Worsley DF, Alavi A, Palevsky HI, Kundel HL. Comparison of diagnostic performance with ventilation-perfusion lung imaging in different patient populations. *Radiology.* 1996;199:481–483.