# Acute Pulmonary Embolism: Sensitivity and Specificity of Ventilation-Perfusion Scintigraphy in PIOPED II Study

**H. Dirk Sostman, MD**  
Paul D. Stein, MD  
Alexander Gottschalk, MD  
Fadi Matta, MD  
Russell Hull, MBBS, MSc  
Larry Goodman, MD

**Purpose:** To use Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED) II data to retrospectively determine sensitivity and specificity of ventilation-perfusion (V/Q) scintigraphic studies categorized as pulmonary embolism (PE) present or PE absent and the proportion of patients for whom these categories applied.

**Materials and Methods:** The PIOPED II study had institutional review board approval at all participating centers. Patient informed consent was obtained; the study was HIPAA compliant. Approval and consent included those for future retrospective research. Patients in the PIOPED II database of clinical and imaging results were included if they had diagnosis at computed tomographic (CT) angiography, Wells score, and diagnosis at V/Q scanning. V/Q scan central readings were recategorized as PE present (PIOPED II reading = high probability of PE), PE absent (PIOPED II reading = very low probability of PE or normal), or nondiagnostic (PIOPED II reading = low or intermediate probability of PE). A composite reference standard was used: the PIOPED II digital subtraction angiographic (DSA) result, or if there was no definitive DSA result, CT angiographic results that were concordant with the Wells score (i.e., positive CT angiographic result and Wells score > 2 or negative CT angiographic result and Wells score < 6). Sensitivity and specificity of recategorized central readings were computed.

**Results:** With the exclusion of patients with intermediate or low probability, the sensitivity of a high probability (PE present) scan finding was 77.4% (95% confidence interval [CI]: 69.7%, 85.0%), while the specificity of very low probability or normal (PE absent) scan finding was 97.7% (95% CI: 96.4%, 98.9%). The percentage of patients with a PE present or PE absent scan finding was 73.5% (95% CI: 70.7%, 76.4%).

**Conclusion:** In a population similar to that in PIOPED II, results of V/Q scintigraphy can be diagnostically definitive in a majority of patients; thus, it can be considered an appropriate pulmonary imaging procedure in patients for whom CT angiography may be disadvantageous.

© RSNA, 2008

---

1 From the Office of the Dean, Weill Cornell Medical College and the Methodist Hospital, 6565 Fannin St, Houston, TX 77030 (H.D.S.); Department of Research, St Joseph Mercy Oakland Hospital, Pontiac, Mich, and Department of Medicine, Wayne State University, Detroit, Mich (P.D.S., F.M.); Department of Radiology, Michigan State University, East Lansing, Mich (A.G.); Department of Medicine, University of Calgary, Calgary, Alberta, Canada (R.H.); and Department of Radiology, Medical College of Wisconsin, Milwaukee, Wis (L.G.). Received February 8, 2007; revision requested April 12; revision received April 26; accepted May 29; final version accepted July 19. Supported by UO1 Grants HL63981, HL63940, and HL067453 from the U.S. Department of Health and Human Services, Public Health Services, National Heart, Lung, and Blood Institute, Bethesda, Md. Address correspondence to H.D.S. (e-mail: dsostman@tmhs.org).

© RSNA, 2008
The Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED) II investigators reported that, among patients with images of adequate quality, multidetector computed tomographic (CT) angiography in the pulmonary arteries had a sensitivity of 83% and a specificity of 96% for the diagnosis of acute pulmonary embolism (PE) in patients who were suspected of having the disorder, and patients were assigned a final reference diagnosis by using a composite diagnostic reference standard (1). The addition of venous-phase CT venography improved sensitivity to 90%, with a specificity of 95%. Finally, CT angiograms of adequate quality and clinical probability assessment by using the Wells score (2) helped yield positive or negative predictive values greater than 90% in 89% of the patients; this was in marked contrast to the results of the original PIOPED study, or PIOPED I (3), in which the combination of ventilation-perfusion (V/Q) scintigrams and clinical assessment resulted in predictive values higher than 90% in only 22% of patients.

By 2001, CT angiography replaced V/Q scanning as the predominant imaging modality for PE diagnosis in the United States (4). However, CT angiography has limitations, which include cost, relatively high radiation doses, and inapplicability in patients who have contraindications (eg, reduced renal function or iodine allergy) to iodinated contrast material. In PIOPED II, of an initial contact population of 7284, 1350 patients had an elevated creatinine level or were receiving dialysis and 272 patients were allergic to intravenous contrast material. In addition, recently proposed diagnostic pathways (5) based on PIOPED II data and results of other recent studies are not definitive with regard to further imaging options if CT results are inconclusive or discordant with the clinical probability assessment.

Research on scintigraphy since PIOPED I suggested that the accuracy of interpretation could be improved by refining diagnostic criteria (6). In addition, retrospective analyses of the PIOPED I database identified findings on the V/Q scan that are associated with a less than 10% probability of PE (7), which is defined as a “very low probability” interpretation; thus, this reduces the number of nondiagnostic interpretations of V/Q scans. It also has been proposed that scintigrams should be interpreted like other imaging tests—that is, positive for PE, negative for PE, and uncertain (8). However, the overall sensitivity and specificity of the use of such diagnostic categories must be determined. Thus, the purpose of our study was to use PIOPED II data to retrospectively determine the sensitivity and specificity of V/Q scintigraphic studies categorized as PE present or PE absent and the proportion of patients for whom these categories apply.

### Materials and Methods

Data were abstracted by one author (F.M.) from the database of clinical findings and imaging results of the PIOPED II investigation, which was a prospective multicenter study to determine the diagnostic validity of multidetector contrast material–enhanced spiral CT angiography and the combination of CT angiography with CT venography for the diagnosis of acute PE (1). The PIOPED II study had institutional review board approval at all participating centers, and patient informed consent was obtained. Approval and consent included those for future retrospective research. The study was Health Insurance Portability and Accountability Act compliant.

In PIOPED II, all patients suspected of having acute PE at clinical evaluation who were seen at the eight participating clinical centers between September 2001 and July 2003 were potentially eligible for recruitment. Exclusion criteria were age of less than 18 years, inability to complete tests within 36 hours, critical illness or hemodynamic instability, ventilatory support, shock or hypotension, myocardial infarction within 1 month, ventricular fibrillation or sustained ventricular tachycardia within 24 hours, abnormal serum creatinine level, chronic renal dialysis (first 14 months of recruitment only), allergy to contrast material, pregnancy, chronic pulmonary hypertension, treatment with long-term use of anticoagulants, thrombolytic therapy planned in the next 24 hours, inferior vena cava filter, deep venous thrombosis of the upper extremity, patients who were prisoners, and patients who were previously enrolled.

Image interpretations for CT angiography, pulmonary digital subtraction angiography (DSA), and V/Q scanning were on the basis of agreement of two blinded PIOPED II central readers from centers other than the one from...
which the image was obtained. The Wells scores were delineated by experienced pulmonary physicians at each site. The methods for obtaining Wells scores, V/Q scans, DSA images, and CT angiograms, as well as the methods of central reading, have been described (1). The PIOPED II central readers (1) were the members of the PIOPED II Ventilation-Perfusion Scan Working Group and are listed in the Appendix.

**Patients**

Patients in the PIOPED II database were included in our Health Insurance Portability and Accountability Act–compliant retrospective study if they had the following: (a) a diagnosis recorded at V/Q scanning and either (b) a Wells score (1,2) recorded prospectively and a diagnosis of PE present or PE absent at CT angiography or (c) a diagnosis of PE present or PE absent at DSA. All patients had given informed consent in PIOPED II for further retrospective evaluation of their study data, and no patient was identified personally in this study.

**Data Analysis**

We (H.D.S.) recorded the V/Q scan central readings and compared them with our study’s reference diagnoses and also compared the distribution of scan readings with those published for PIOPED I (3). We also recorded the computed sensitivity and specificity at cutoffpoints representing the probability category readings for both the PIOPED II data in our study and the published PIOPED I data (3).

We (H.D.S., P.D.S., F.M.) reclassified the PIOPED II V/Q scan central readings (Table 1) as PE present (PIOPED II reading = high probability of PE), PE absent (PIOPED II reading = very low probability of PE or normal), or nondiagnostic (PIOPED II reading = low probability or intermediate probability of PE).

**Statistical Analysis**

We (H.D.S., P.D.S., F.M.) computed sensitivity and specificity by using two reference standards. First, we compared the above recategorization of the PIOPED II V/Q scan central readings with the PIOPED II DSA results. If DSA was not performed or if there was no definitive DSA result, we compared the recategorizations of the V/Q scan central readings with PIOPED II CT angiographic results that were concordant with the Wells score (ie, positive CT angiographic result and Wells score ≥ 2 or negative CT angiographic results and Wells score < 6), which in the PIOPED II study had both positive and negative predictive values of 93% (1).

Binomial 95% confidence limits of sensitivity and specificity were computed (9). A software program was not used to perform the calculations. (Calculations were performed by H.D.S. and F.M.)

**Results**

**Patients and Readings**

A total of 951 patient records met the selection criteria (Figure). Of these, 41 were eliminated because there was no DSA result and the CT angiographic result and Wells score were discordant (ie, the CT angiographic result was negative for PE and the Wells score was greater than 6, indicating a high probability at clinical examination, or the CT angiographic result was positive for PE and the Wells score was less than 2, indicating a low probability at clinical examination). This resulted in a sample of 910 patients. Among these patients, a definitive DSA result was present in 216 and a concordant CT angiographic result and Wells score were present in 694.

The V/Q scan readings were categorized as follows: 102 scans were categorized as high probability; 152 scans, as intermediate probability; 89 scans, as low probability; 415 scans, as very low probability; and 152 scans, as normal (Table 2). The positive predictive value of a high probability scan was 87.2% (89 of 102); that of an intermediate probability scan, 30.9% (47 of 152); that of a low probability scan, 6.7% (six of 89); that of a very low probability scan, 5.8% (24 of 415); and that of a normal scan was 1.3% (two of 152). The very low probability category was not used in PIOPED I, while PIOPED I also included few truly normal scans (3).

**Categorization of V/Q Scan Central Readings**

The sensitivity and specificity for different scan category cutoff points are
shown for the PIOPED II and PIOPED I data in Table 3. In our study, with the exclusion of patients with intermediate or low probability, the sensitivity of a high probability (PE present) scan was 77.4% (89 of 115), while the specificity of a very low probability or normal (PE absent) scan was 97.7% (541 of 554). A total of 241 scans (26.5%) were categorized as nondiagnostic (ie, intermediate or low probability) at our predetermined analysis. The percentage of patients with a high or a very low probability or normal scan was 73.5% (669 of 910). Therefore, the PE present (high probability) and PE absent (very low probability or normal) categories showed a sensitivity for acute PE of 77.4% (95% confidence interval: 69.7%, 85.0%) and a specificity of 97.7% (95% confidence interval: 96.4%, 98.9%) among patients remaining after exclusion of those with nondiagnostic (intermediate or low probability) interpretations.

### Discussion

The goal of our study was to evaluate the sensitivity and specificity of V/Q scans categorized as PE present or PE absent and to determine in what proportion of patients these categories applied. The study conformed to the Standards for Reporting of Diagnostic Accuracy criteria for reporting of studies of diagnostic accuracy (10). Our results are that the V/Q scintigram was categorized as PE present or PE absent in a large percentage (73.5%) of patients and the sensitivity and specificity of those scans were defined according to PIOPED II (1) criteria: low clinical probability (Wells score, 2) with CT angiographic result positive for PE, or high clinical probability (Wells score, >6) with CT angiographic result negative for PE.

### Table 2

<table>
<thead>
<tr>
<th>Category</th>
<th>PIOPED I</th>
<th>PIOPED II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PE Present (n = 251)</td>
<td>PE Present (n = 168)</td>
</tr>
<tr>
<td>High probability</td>
<td>102</td>
<td>89</td>
</tr>
<tr>
<td>Intermediate probability</td>
<td>105</td>
<td>47</td>
</tr>
<tr>
<td>Low probability</td>
<td>39</td>
<td>6</td>
</tr>
<tr>
<td>Very low probability</td>
<td>NA</td>
<td>24</td>
</tr>
<tr>
<td>Normal</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

Note.—Data are number of readings. NA = not available.
tant with the Wells score. 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 at DSA, but the PIOPED II DSA readings were independent of other imaging or clinical information. The CT angiographic readings and Wells score determinations likewise were independent data observations. We chose to use only those CT angiographic results that were concordant with the Wells scores because the PIOPED II study results (1) showed that CT angiography result and Wells score concordance was associated with high positive and negative predictive values. The sensitivity and specificity of the V/Q scan computed without reference to DSA results, with only a concordant CT angiographic result and Wells score as the reference standard, were 85.1% and 98.2%, respectively, after patients with intermediate or low probability V/Q scans were eliminated. This was not substantially different from the sensitivity of 77.4% and specificity of 97.7% obtained by using DSA as a component of the reference standard.

There are several differences between our results and the PIOPED I study results with regard to V/Q scanning. One of the most striking is the proportion of patients who received a “definite” diagnosis; there were many fewer intermediate probability readings in PIOPED II and many more low, very low, and normal readings (Table 2). We believe that this is due to two independent factors. First, criteria for interpreting V/Q scans have improved considerably since the PIOPED I study (and largely thanks to the data it provided) (11–16). Second, the frequency of different readings from V/Q scans likely will vary depending on the population studied (eg, patients in an intensive care unit would be expected to have a higher proportion of uncertain results). The patient samples in PIOPED I and PIOPED II were quite different, as shown in Table 4. Many studies of PE diagnosis (17–19) have a composition similar to that of PIOPED II (eg, heavily weighted toward outpatients), while some (20–22) more closely resemble PIOPED I (eg, with more representation of inpatients and the critically ill). There is little literature on the performance of V/Q scanning in different populations (23), but most nuclear medicine physicians would probably agree that scans in inpatients and critically ill patients are more difficult to interpret than scans in outpatients.

The positive predictive values for low probability and very low probability were quite similar in our study. This suggests the possibility of combining these categories. If one were to do this, it would have the same effect as adding low probability to the PE absent category. The resulting sensitivity would be 73.6%, with a specificity of 98.0%. The nondiagnostic category would consist of 16.7% of the patient population. However, this combination was not our pre-specified hypothesis and would require an independent validation study. Similarly, if one combines the Wells score with the V/Q categories of PE present, PE absent, and nondiagnostic, a total of 617 (68.9%) of 896 patients who had both V/Q scanning results and Wells scores recorded would have findings with a positive or negative predictive value greater than 90%. This is also a post hoc analysis and would require independent confirmation, particularly because the prevalence of PE in the PIOPED I low probability category was 14%; it is not clear without further data whether the difference reflects evolution of interpretive criteria or differences in the two patient populations.

In comparison with the results of CT angiography in PIOPED II, the results reported here for V/Q scanning have both similarities and differences. There were many more nondiagnostic results with V/Q scanning (26.5% of patients) than with CT angiography (6.2% of patients). Furthermore, the nondiagnostic CT angiographic results were virtually all because of technically inadequate scans (24), while the nondiagnostic V/Q scanning results were all because of inconclusive interpretations. Once nondiagnostic studies were removed from the study sample, the sensitivity and specificity of V/Q scanning (77% and 98%, respectively) were similar to those of CT angiography (83% and 96%, respectively).
Our study had limitations. The fact that DSA was not performed in all patients was a limitation, albeit one that all future studies of PE will almost certainly share because of the vanishing use of invasive pulmonary angiography in clinical practice. The use of concordant CT angiographic results and Wells scores as an alternative diagnostic reference standard is supported by the PIOPED II study results (1) but has not been prospectively validated in an independent study. Therefore, we can state with assurance that V/Q scans have the reported test-operating characteristics as a substitute for CT angiography when CT angiography is not desirable, but not necessarily as an independent alternative to CT angiography. Another potential limitation relates to technique. Planar imaging was used for V/Q studies in PIOPED II, which is the historical standard. Tomographic imaging has theoretical advantages and has been used in some centers (25), but it has not been the subject of a large prospective clinical trial for PE diagnosis.

Our study results have implications with regard to the role of V/Q imaging in patients suspected of having acute PE in the current era of CT angiography. The well-recognized advantages of CT angiography—rapid and widely available examinations, the ability to concurrently perform lower extremity venous imaging, the higher proportion (94% at CT angiography vs 74% at V/Q scanning in PIOPED II) of definitive diagnostic results, and the ability of CT to help make nonvascular diagnoses (26,27) that may be of clinical importance—secure a place for CT angiography as the first-line diagnostic imaging method. However, the advantages of V/Q scintigraphy—lower radiation dose, lack of iodinated contrast material—can now be considered together with a contemporary characterization of the test’s diagnostic performance in a predominantly outpatient population. We conclude that V/Q scintigraphy often yields diagnostically definitive results and can be considered an appropriate pulmonary imaging procedure in patients for whom CT angiography may be disadvantageous.

### References