V/Q SPECT Interpretation for Pulmonary Embolism Diagnosis: Which Criteria to Use?

Pierre-Yves Le Roux1–3, Philippe Robin1–3, Aurélien Delluc1,2,4, Ronan Abgral1–3, Alexandra Le Duc-Pennec1–3, Emmanuel Nowak5, Francis Couturaud1,2,4, Grégoire Le Gal1,4,6, and Pierre-Yves Salaun1–3

1Université Européenne de Bretagne, Brest, France; 2Université de Brest, EA3878 (GETBO) IFR 148, Brest, France; 3CHRU de la Cavale Blanche, Service de Médecine Nucléaire, Brest, France; 4CHRU de la Cavale Blanche, Département de Médecine Interne et de Pneumologie, Brest, France; 5INSERM CIC 05-02 IFR148, Brest, France; and 6Université de Brest, INSERM CIC 05-02 IFR148, Brest, France

Ventilation-perfusion (V/Q) SPECT has been reported to improve the diagnostic performance of V/Q imaging for the diagnosis of pulmonary embolism (PE). However, only sparse data based on an objective reference test are available, and the criteria used for interpretation have varied widely. Therefore, the aim of our study was to assess the performance of V/Q SPECT using various criteria for interpretation, in comparison with a validated independent diagnostic strategy. Methods: The SPECT study included patients for whom V/Q SPECT data were compared with the results of an independent and validated diagnostic algorithm for PE. V/Q SPECT scans were performed after intravenous injection of 99mTc-macro-aggregated albumin and simultaneous ventilation with 81mKr gas. Interpretation was performed independently by 2 nuclear medicine physicians who were not aware of the clinical history, diagnostic strategy conclusion, or patient’s outcome. Sensitivity, specificity, and likelihood ratios were evaluated for various combinations of mismatched defect numbers and sizes (segmental or subsegmental). Generation of receiver-operating-characteristic curves was based on the number of mismatch defects and the number of subsegmental mismatch defects or equivalent. Results: Of the 249 patients who were analyzed, the diagnosis of PE was confirmed in 49 and ruled out in 200 according to the previously validated independent strategy. Of all the tested criteria, the best performance was achieved using a diagnostic cutoff of at least 1 segmental or 2 subsegmental mismatches, with sensitivity and specificity of 0.92 (95% confidence interval, 0.84–1) and 0.91 (95% confidence interval, 0.87–0.95), respectively. With a negative V/Q SPECT result, the posttest probability of PE was 0.010, 0.037, and 0.119 for a low, intermediate, and high clinical probability. With a positive V/Q SPECT result, the posttest probability of PE was 0.531, 0.814, and 0.939 for a low, intermediate, and high probability. Conclusion: For V/Q SPECT interpretation, a diagnostic cutoff of 1 segmental or 2 subsegmental mismatches seems best for confirming or excluding acute PE.

Key Words: V/Q SPECT; pulmonary embolism; criteria

DOI: 10.2967/jnumed.112.113639

Planar ventilation-perfusion (V/Q) lung scintigraphy was the first validated noninvasive procedure for the diagnosis of pulmonary embolism (PE). In the late 1980s, chest CT overtook V/Q scanning in most institutions as the initial imaging modality for suspected PE, mainly because of the high rate of inconclusive results from planar V/Q scintigraphy using the probabilistic interpretation proposed in the PIOPED study (prospective investigation of pulmonary embolism diagnosis) (1).

More recently, there has been a renewed interest in V/Q imaging, primarily in response to concerns about high radiation exposure from CT angiography (2). Indeed, the radiation dose is lower with V/Q scintigraphy than with chest multidetector CT (MDCT), particularly to the female breast (3–5). Another drawback of MDCT is that it cannot be performed in many patients because of contraindications such as renal failure or an allergy to the contrast agent. Insufficient quality is also responsible for inconclusive results in 5%–10% of cases. Moreover, V/Q imaging might be the best-adapted modality for follow-up, especially for the diagnosis of recurrence (6) or for therapeutic assessment. In view of these considerations, V/Q imaging remains a powerful tool for the diagnosis of PE. In addition, the use of SPECT has been reported to improve the performance of V/Q imaging thanks to better spatial resolution and limited overlapping of perfusion defects by other structures. Studies have suggested that the performance of V/Q SPECT could at least equal that of MDCT (7).

In a recent study, we showed satisfactory accuracy for V/Q SPECT in the diagnosis of PE integrated into a validated diagnostic strategy, despite the use of the revised PIOPED criteria for interpretation, which were initially developed only for V/Q planar scans and are probably inadequate for V/Q SPECT (8). Before V/Q SPECT is integrated into routine practice, it is of the utmost importance that the interpretation criteria for V/Q SPECT be validated. Indeed, the medical literature shows an important variability in the criteria used for the diagnosis of PE. For example, some studies suggested that 1 mismatch defect, regardless of size, was sufficient for the diagnosis of PE (9–11), whereas in others, 1 segmental mismatch (12) or 2 mismatches (7) were necessary. In 2009, the European Association of Nuclear Medicine published guidelines for V/Q scintigraphy in which criteria for reading V/Q SPECT images were proposed. However, these criteria were not validated according to a clear, independent gold standard (13).
The aim of our study was to assess the accuracy of various criteria for V/Q SPECT interpretation, using as a reference standard a validated independent diagnostic strategy.

**MATERIALS AND METHODS**

The eligible study population consisted of patients with suspected PE who had been consecutively included in a prospective diagnostic accuracy study that assessed the performance of V/Q SPECT (8,14). The methods of the study have been extensively described elsewhere (8,14). Briefly, the study consisted of patients aged 18 y or older who were seen as inpatients or outpatients at Brest University Hospital because of symptoms suggestive of PE. The clinical probability of PE was assessed by the physicians in charge according to the clinical model described by Wells et al. and was rated as low, intermediate, or high (15). To be considered for inclusion, patients had to have either a high clinical probability of PE or an abnormal concentration of plasma enzyme-linked immunosorbent assay dimer (500 μg/mL). Exclusion criteria were pregnancy, breastfeeding, life expectancy of less than 3 mo, impossibility of follow-up, anticoagulant treatment long-term or beginning more than 48 h before the time of screening, and confirmed or massive PE. The protocol was approved by the Ethics Committee of our institution and was registered with institutional review board authorization number 04.036 on the clinical trials.gov Web site (NCT01183026). Written informed consent was obtained from all patients. All included patients underwent a standardized diagnostic strategy. First, the patients underwent lower-limb vein compression sonography. PE was considered present in patients with proximal deep venous thrombosis (16). Patients with no deep venous thrombosis on sonography underwent a V/Q planar lung scan. Dual-isotope (99mTc-macroaggregated albumin/81mKr gas) V/Q planar lung scans were acquired with 6 standard views (anterior, posterior, right and left lateral, and right and left posterior oblique), and the nuclear medicine physician on duty classified them on the basis of the revised PIOPED criteria (normal, low, intermediate, or high) (17). The results of chest radiography did not affect the choice of imaging procedure in the diagnostic algorithm but were available for interpretation. PE was diagnosed in patients with an intermediate or high clinical probability of PE and a high-probability V/Q scan. PE was ruled out in patients with normal V/Q findings, a low or intermediate clinical probability of PE and a low-probability V/Q scan, or a low clinical probability of PE and an intermediate-probability V/Q scan. All other patients underwent chest MDCT, which was interpreted by a vascular radiologist. Detailed results have previously been reported (8,14).

The safety of the diagnostic strategy was assessed by monitoring the risk of thromboembolic events during the 3-mo follow-up period in patients deemed not to have PE according to the initial diagnostic work-up.

**V/Q SPECT Acquisition and Interpretation**

All patients who required a planar V/Q scan also underwent pulmonary V/Q SPECT at the same time. Pulmonary SPECT consisted of perfusion SPECT and ventilation SPECT performed simultaneously (i.e., 128 projections of 12 s each, 360°) immediately after the planar images were acquired, with the patient remaining supine. Planar and SPECT scans were obtained after a single intravenous injection of 200 MBq of 99mTc-macroaggregated albumin and simultaneous ventilation with 81mKr gas obtained using an 85Rb/81mKr generator. Images were acquired using a dual-head camera (DST Xli; SMV) equipped with low-energy, high-resolution collimators in a 128 × 128 matrix. Images were collected with 2 energy windows centered on a photopeak of 140 keV ± 10% (99mTc) for perfusion and 190 keV ± 10% (81mKr) for ventilation. V/Q SPECT images were numerically stored but not interpreted until the completion of the patient’s follow-up.

Retrospectively, V/Q SPECT results were interpreted independently by 2 nuclear medicine physicians who did not know the clinical history, reference strategy imaging results, diagnostic conclusion, or patient’s outcome. Any difference of interpretation was resolved by consensus.

The location, type (matched or mismatched), and extent (segmental or subsegmental) of each defect were recorded. A mismatched defect was defined as a pulmonary region affected by a severe reduction or complete loss of perfusion along with distinctly reduced ventilation. The extent of each defect was assessed visually. A defect was defined as segmental if it involved more than 75% of a segment and subsegmental if it involved less than 75%. The tested criteria were as follows: 1 mismatch, regardless of size; at least 1 segmental mismatch; at least 2 mismatches, regardless of size; at least 1 segmental or 2 subsegmental mismatches; more than 1 segmental mismatch; at least 2 segmental mismatches; and at least 3 mismatches, regardless of size.

**Data Analysis**

Standard statistical measures of diagnostic accuracy were used for each tested criterion, including sensitivity, specificity, and likelihood ratios. Generation of receiver-operating-characteristic curves was based on the number of mismatch defects and the number of subsegmental mismatch defects or equivalent. Areas under the curve were calculated from the receiver-operating-characteristic curves.

**RESULTS**

We included 321 patients over the 28-mo inclusion period, with a median age of 72 y (range, 18–95 y).

**Diagnostic Strategy Including Conventional V/Q Scans**

Sonography showed a proximal deep venous thrombosis, establishing venous thromboembolism, in 43 patients (13%). A V/Q scan was therefore required in 278 patients (87%). Among the remaining 278 patients, the diagnosis of PE was confirmed in 55 (20%) and ruled out in 220 (79%). Three patients (1%) did not undergo chest MDCT as per study protocol and hence could not be classified regarding the diagnosis of PE. All 3 received anticoagulant therapy on the basis of a distal-only deep venous thrombosis and were excluded from further analysis. Follow-up was completed successfully for all patients. The 3-mo thromboembolic risk in patients not given anticoagulants, based on the results of the diagnostic protocol, was 0.53% (95% confidence interval [CI], 0.09–2.94).

**V/Q SPECT Interpretation**

Among 278 patients who underwent conventional planar V/Q scans, V/Q SPECT data could not be analyzed in 26 patients (9%) mainly because of handling and operative constraints (n = 5), technical pitfalls (n = 3), or loss of numerically stored data (n = 18). Therefore, 249 patients with V/Q SPECT and clinical outcome were analyzed. Figure 1 summarizes the results of the diagnostic strategy for these 249 patients.

Sensitivity and specificity were calculated for each of the criteria tested (Table 1). The highest sensitivity (0.92; 95% CI, 0.84–1) was achieved using a diagnostic cutoff either of 1 segmental or subsegmental mismatch or of 1 segmental or 2 subsegmental mismatches. Regarding these 2 criteria, specificity was 0.91 (95% CI, 0.87–0.95) for 1 segmental or subsegmental mismatch versus 0.84 (95% CI, 0.79–0.98) for 1 segmental or 2 subsegmental mismatches. Figure 2 illustrates the challenge of V/Q SPECT interpretation criteria.
Receiver-operating-characteristic curves were generated according to the number of mismatches, regardless of their size (Fig. 3) and the number of subsegmental mismatches or equivalent (considering 1 segmental mismatch equal to 2 subsegmental mismatches) (Fig. 3B). The areas under the curve were, respectively, 0.9374 and 0.9376. Considering these 2 curves, the best compromise sensitivity–specificity was achieved for 2 subsegmental mismatches or equivalent, corresponding to 1 segmental or 2 subsegmental mismatches.

Likelihood ratios with V/Q SPECT are presented in Table 1. The positive likelihood ratio using a diagnostic cutoff of 1 segmental or 2 subsegmental mismatches was 10.2 (95% CI, 6.5–16). The negative likelihood ratio was 0.09 (95% CI, 0.04–0.23).

![Study flow chart](image)

**FIGURE 1.** Study flow chart. Interm. = intermediate; OAT = oral anticoagulant therapy; VTE = venous thromboembolism.

**FIGURE 3.** Receiver-operating characteristic (ROC) curve analysis. V/Q scan = ventilation–perfusion scintigraphic study; C138 = confidence interval; PIOPED = Prospective Investigation of Pulmonary Embolism Diagnosis; V/Q = ventilation–perfusion; VTE = venous thromboembolism. ROC curves were generated according to the number of mismatches, regardless of their size (Fig. 3A) and the number of subsegmental mismatches or equivalent (considering 1 segmental mismatch equal to 2 subsegmental mismatches) (Fig. 3B). The areas under the curve were, respectively, 0.9374 and 0.9376. Considering these 2 curves, the best compromise sensitivity–specificity was achieved for 2 subsegmental mismatches or equivalent, corresponding to 1 segmental or 2 subsegmental mismatches.

**DISCUSSION**

In this study, we assessed different criteria for V/Q SPECT interpretation in the face of suspected PE using as a reference standard a validated independent diagnostic strategy. The best performance was achieved using a diagnostic cutoff of 1 segmental or 2 subsegmental mismatches, with sensitivity and specificity of 0.92 (95% CI, 0.84–1) and 0.91 (95% CI, 0.87–0.95), respectively.

Indeed, many studies have suggested that the transition from planar imaging to SPECT may improve the diagnostic performance of V/Q imaging for the diagnosis of PE. However, only sparse data based on an objective reference test are available, and the criteria used for interpretation varied widely (9).

Before integrating V/Q SPECT into routine practice, 3 validating steps are needed: the first step was the validation of V/Q SPECT accuracy compared with a validated diagnostic strategy, which was the aim of our previous study (8); the second step, addressed in this study, is the validation of interpretation criteria dedicated to V/Q SPECT; the last step will be a large management outcome study including V/Q SPECT in the diagnostic algorithm, in which patients with a negative diagnostic work-up would be left untreated and followed up over time.

Historical PIOPED studies used probabilistic interpretation, with results proposed on a scale from high to low probability (17). This terminology led to confusion for physicians and was probably partly responsible for the declining popularity of the V/Q scan. There is currently a consensus that V/Q SPECT readings should be binary: positive or negative. However, the diagnostic cutoff varies widely (19). Indeed, PE has been diagnosed according to different criteria in different studies, such as “1 mismatched defect, regardless of size” (9–11), “1 segmental mismatched defect” (12), or “2 mismatched segmental or subsegmental mismatched defects” (7). The European Association of Nuclear medicine also proposed criteria for reading V/Q SPECT scans in which PE is reported for a mismatch of at least “1 segment or 2 subsegments that conform to the pulmonary vascular anatomy” (13). However, these criteria were not validated on the basis of an objective reference standard.

In our study, the best diagnostic cutoff for the diagnosis of PE was 1 segmental or 2 subsegmental mismatches, with sensi-

**TABLE 1**

<table>
<thead>
<tr>
<th>SPECT criteria</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive likelihood ratio</th>
<th>Negative likelihood ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 segmental or subsegmental mismatch</td>
<td>0.92 (0.84–1)</td>
<td>0.84 (0.79–0.89)</td>
<td>5.7 (4.1–8)</td>
<td>0.10 (0.04–0.25)</td>
</tr>
<tr>
<td>≥1 segmental mismatch</td>
<td>0.88 (0.76–0.96)</td>
<td>0.93 (0.89–0.96)</td>
<td>11.4 (6.9–18.8)</td>
<td>0.15 (0.08–0.31)</td>
</tr>
<tr>
<td>≥2 segmental or subsegmental mismatches</td>
<td>0.88 (0.79–0.97)</td>
<td>0.95 (0.91–0.98)</td>
<td>16.0 (8.9–28.6)</td>
<td>0.13 (0.06–0.27)</td>
</tr>
<tr>
<td>≥1 segmental or 2 subsegmental mismatches</td>
<td>0.92 (0.84–1)</td>
<td>0.91 (0.87–0.95)</td>
<td>10.2 (6.5–16)</td>
<td>0.09 (0.04–0.23)</td>
</tr>
<tr>
<td>≥2 segmental mismatches</td>
<td>0.82 (0.71–0.93)</td>
<td>0.96 (0.93–0.99)</td>
<td>20.4 (10.2–40.7)</td>
<td>0.19 (0.11–0.35)</td>
</tr>
<tr>
<td>≥3 segmental or subsegmental mismatches</td>
<td>0.67 (0.55–0.82)</td>
<td>0.99 (0.97–1)</td>
<td>44.9 (14.4–140.4)</td>
<td>0.33 (0.22–0.50)</td>
</tr>
</tbody>
</table>

Data in parentheses are 95% confidence intervals.
tivity and specificity of 0.92 (95% CI, 0.84–1) and 0.91 (95% CI, 0.87–0.95), respectively. These data are concordant with those previously published, reporting sensitivity and specificity of at least 90% in almost all studies (19). The main strength of our study is that we used as a reference standard an independent validated diagnostic strategy whose conclusions were not based on V/Q SPECT results.

Moreover, in our study, calculations of posttest probability using pretest probability and likelihood ratios according to the Bayes theorem showed that the diagnostic conclusions were satisfactory when clinical assessment was not discordant with the V/Q SPECT interpretation. However, the posttest probability of PE with a negative V/Q SPECT result but a high clinical probability was 12%, and the posttest probability with a positive V/Q SPECT result but a low clinical probability was 53%. These data are concordant with previously published studies, especially with CT pulmonary angiography. In the PIOPED II trial, CT pulmonary angiography readings were false-positive 42% of the time among patients with a low clinical probability of PE and false-negative 40% of the time among patients with a high clinical probability (20). Thus, when the clinical probability and the results of imaging procedures are discordant, the posttest probability of PE is probably neither sufficiently low nor sufficiently high to permit therapeutic decisions. Under these circumstances, additional objective testing is necessary.

Our study had some potential limitations. First, the subset in whom SPECT was analyzed did not include “all comers” but only patients with either a high clinical probability or positive d-dimer result and negative sonography results, thus reducing the number of patients with PE. On the other hand, given the diagnostic algorithm, this limitation is also an advantage because the subset represents the group of patients who would have V/Q SPECT using a diagnostic algorithm that has sonography as a first examination. One other limitation of this study could be the potential correlation between the planar V/Q scan used in the reference standard and V/Q SPECT. Indeed, because these 2 modalities assess the same physiopathologic process with different technical approaches, the accuracy of V/Q SPECT could have been artificially increased. However, our gold standard was not only V/Q planar scanning but also a sequential diagnostic algorithm including clinical probability, d-dimer testing, sonography, and MDCT, along with a formal 3-mo clinical follow-up. Thus, we compared V/Q SPECT results to a reference diagnostic strategy, limiting the impact of a possible correlation between planar and SPECT acquisition. Finally, anatomic imaging was not used for V/Q SPECT interpretation. Indeed, the aim of the study was to assess the performance of V/Q SPECT as the cornerstone of the PE diagnosis. Moreover, the impact of anatomic imaging on the interpretation of V/Q scans is not clearly established. The development of dual-modality imaging with SPECT/CT, which provides simultaneous tomographic anatomic imaging, may improve the performance of V/Q SPECT and might be evaluated in future studies.

Besides its diagnostic performance, V/Q scanning offers many advantages over MDCT. First, the radiation exposure is lower with V/Q scanning, particularly to the female breast. Second, V/Q scanning can be proposed for almost any patient. In the PIOPED 2 study, over 40% of patients did not undergo MDCT, mainly because of renal impairment or contrast allergy. Third, V/Q SPECT might be the best-adapted modality for follow-up, especially for the diagnosis of recurrence or for therapeutic assessment.
CONCLUSION

From these considerations and using a diagnostic cutoff of 1 segmental or 2 subsegmental mismatches, V/Q SPECT seems able to position itself as the first imaging test in patients with suspected PE. The last step now required is to perform a large management outcome study with venous thromboembolism recurrence and mortality at 3 mo as outcome measures to definitely confirm these data.

DISCLOSURE

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734. This study was partially funded by the Projet Hospitalier de Recherche Clinique 2004 (French Ministry of Health). No other potential conflict of interest relevant to this article was reported.

REFERENCES