Ventilation-Perfusion Scintigraphy in the PIOPED Study. Part II. Evaluation of the Scintigraphic Criteria and Interpretations

Alexander Gottschalk, H. Dirk Sostman, R. Edward Coleman, Jack E. Juni, James Thrall, Kenneth A. McKusick, Jerry W. Froelich and Abass Alavi

Department of Radiology, Michigan State University, East Lansing, Michigan; Department of Radiology, Duke University Medical Center, Durham, North Carolina; Department of Radiology, William Beaumont Hospital, Royal Oak, Michigan; Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts; Radiology Imaging Associates, Swedish Medical Center, Englewood, Colorado; and Department of Radiology, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania

This article presents an evaluation of the criteria used for categorical interpretation of the ventilation-perfusion (V/Q) scans performed in the PIOPED study. In addition, the correlation of percent probability estimates with the actual frequency of pulmonary embolism (PE) is presented. Cases which met the PIOPED criteria for various diagnostic categories were selected by computerized search of the detailed scan descriptions that had been done as part of the study. The process by which the scans were described was detailed in Part I of this report. Most of the criteria appropriately categorized V/Q scans which satisfied them. However, we recommend that three criteria should be reconsidered:

1. A single moderate perfusion defect is appropriately categorized as intermediate, rather than as low probability.
2. Extensive matched V/Q abnormalities are appropriate for low probability, provided that the chest radiograph is clear. On the other hand, single-matched defects may be better categorized as intermediate probability. Although due to the small number of cases with this finding, no definite, statistically founded recommendation can be made.
3. Two segmental mismatches may not be the optimum threshold for high probability, and in some cases should be considered for intermediate probability. However, due to the small number of cases with this finding, no definite, statistically founded recommendation can be made.

We suggest that the revised criteria resulting from these adjustments should now be used for the interpretation of V/Q scans.

J Nucl Med 1993; 34:1119–1126

This study uses the computerized data from the consensus ventilation-perfusion (V/Q) scan description obtained in the PIOPED study to evaluate retrospectively: (1) the PIOPED criteria for categorical interpretation of V/Q scans and (2) the estimates of the percent probability of pulmonary embolism (PE) by the investigators.

METHODS AND RESULTS

In the analyses that follow, we discuss each criterion used in the PIOPED study for categorical probability assessment. The data reported here were obtained by isolating various components of the detailed consensus description made from the V/Q scans by members of the Nuclear Medicine Working Group* and comparing these to descriptors of the angiographic findings derived from the angiographic findings determined by the PIOPED Angiography Working Group†.

This is possible since the Angiography Working Group also described their findings on a form intended for computerization. This form included a detailed description of the location of any embolus found. Pulmonary emboli could be described in the main pulmonary artery (PA), left or right PA and regional lobar, segmental or peripheral PA depending on the size and location of the embolus. Therefore, it was possible to describe regional findings on the V/Q scan and determine whether the patient had a pulmonary embolus and if its location corresponded to the location of those particular findings on the V/Q scan. In those patients who had PE diagnosed upon angiography of the first lung studied and therefore underwent angiography of only one lung (*1*), regional correlations were performed in the present study only for those lung regions in which angiograms were obtained.

A small number of patients (*1*) were unable to complete ventilation scans. These patients are not included in this study.

Data from the trial (*1*) were initially reported on the basis of findings from the PIOPED Clinical Outcome Committee, which reviewed not only the pulmonary angiogram results, but also the

Received December 14, 1992; revision accepted March 30, 1993.

For correspondence or reprints contact: H. Dirk Sostman, MD, Dept. of Radiology, Box 3808, Duke University Medical Center, Durham, NC 27710.

*The PIOPED Nuclear Medicine Working Group included the following individuals: Abass Alavi, MD, Chairman (University of Pennsylvania); R. Edward Coleman, MD (Duke University); Jerry W. Froelich, MD (Henry Ford Hospital); Alexander Gottschalk, MD (Yale University); Jack E. Juni, MD (University of Michigan); Kenneth A. McKusick, MD (Massachusetts General Hospital); H. Dirk Sostman, MD (Yale University); James Thrall, MD (Henry Ford Hospital).

†The PIOPED Angiography Working Group included the following individuals: Richard H. Greenspan, MD, Chairman (Yale University); N. Reed Dunnick, MD (Duke University); P.C. Shetty, MD (Henry Ford Hospital); Christos Athanasoulis, MD (Massachusetts General Hospital); Kyung J. Cho, MD (University of Michigan); Gordon K. McLean, MD (University of Pennsylvania); Donald F. Denny Jr., MD (Yale University).

patient's clinical status. For example, patients who had a V/Q scan but no angiogram, but who nevertheless were followed for 1 yr, were not anticoagulated and had no relevant clinical events were considered by the outcome committee not to have PE. For the purposes of evaluating the scan criteria, however, we have in almost all instances analyzed only those patients who had angiograms which were read definitely as PE present or PE absent and thus could be correlated with the V/Q scans. The only exception to this is consideration of the "normal" scan interpretations, which includes patients classified both by angiography and by outcome.

It is important to note that these analyses focus upon individual PIOPED criteria. In many instances, combined patterns were excluded. In addition, we do not report the data pertaining to other findings which were not enumerated in the PIOPED criteria.

Combined patterns involving mismatched defects were considered in the analysis of criteria for "high probability," since mismatched defects and high probability diagnoses took priority over other patterns in the PIOPED criteria. For example, a patient with three segmental mismatches (which meets criteria for high probability) and matched V/Q defects (which meets the criteria for low probability) should have been assigned to the high probability category. Thus, all such patients (including those with combined patterns) were considered in the analysis of mismatched defects for high probability.

However, when evaluating low probability criteria, since these did not take precedence in scan categorization, it was necessary to isolate analysis to those patients who did not have scan findings which would place them in a higher category. For example, a patient with a matched perfusion/ventilation defect and an area of matched perfusion/ventilation/radiographic abnormality *would be excluded* from the analysis of matched perfusion defects. A total of 137 such patients were excluded. Patients without V/Q mismatch who had pleural effusions combined with other abnormalities were also excluded. A total of 102 patients fell into this group. Therefore, many patients in the study database with mixed patterns that were not enumerated specifically in the PIOPED criteria do not contribute to the analysis of low probability criteria and are referred to future analyses.

The analyses which follow correlate scintigraphic patterns that fulfilled individual PIOPED criteria with angiograms read as definitely positive or negative for PE. There were a total of 731 patients in the randomized, mandatory angiography group who had definitive angiographic results. We report here the analysis of scan-angiogram correlations in 393 of those patients (53.8%) that were pertinent to *individual* PIOPED criteria.

For the purposes of this analysis, we define "low probability" as a 0%–19% likelihood of PE, "high probability" as an 80%–100% likelihood of PE and "intermediate" as 20%–79%. We recognize that these ranges are somewhat arbitrary, but we have used them since there are no universally accepted thresholds and these are the ranges used to group the PIOPED Clinical Science estimates of percent probability into categories (*1*).

Based on the review described here, we conclude that some of the PIOPED criteria were not appropriate for the category to which they were originally designated. The results of this analysis are therefore assembled into a revised set of scintigraphic criteria for the diagnosis of PE.

Criterion Analysis for Less Than Low Probability

Normal Perfusion Images. A truly normal scan was one in which both readers considered the perfusion images to be normal. There were 21 studies with "normal" interpretations. None had PE as determined by pulmonary angiogram (3 patients) or by 1 yr of careful clinical follow-up (18 patients).

Near Normal/Normal. In the initial PIOPED publication (*1*), a scan category of "near normal/normal" was used. This category was developed because few patients in the "normal" and "very low probability" categories had pulmonary angiograms. It was defined precisely as "readings of very low probability by one reader and low probability by the other, very low probability by both, or very low probability by one and normal by the other." Of the patients in this category who had angiography, 9% had PE, although based upon the Outcome Committee classification (which includes patients who had only follow-up), only 4% had PE. Unfortunately, the use of the term "near normal/normal" has caused much confusion since the "near normal/normal" category has been considered by many to be equivalent to truly normal (*2*). To avoid this confusion and to conform with common current usage, we called this "near normal" group "very low probability." In addition, since the criteria for very low probability involved only small perfusion defects, these were combined with "low probability" and a single analysis of small lesions was made.

Criterion Analysis for Low Probability

Nonsegmental Perfusion Defects. Twenty-nine cases were identified in which the perfusion scan demonstrated either cardiomegaly, enlargement of the aorta or hila, an elevated diaphragm on one or both sides or any combination thereof with no other perfusion defects present. None (0%) of these 29 patients had PE identified on angiography.

When pleural effusion caused an isolated perfusion defect on the V/Q scan, with the size of the perfusion defect congruent with the size of the pleural effusion, and when the pleural effusion was limited to the costophrenic angle, two of eight (25%) patients had PE; only two of these patients had unilateral effusions and neither (0%) had PE. The number of patients with this finding was probably too small to be meaningful. When all sizes of effusion were considered, 4 of 27 (15%) such patients had PE.

In addition, we investigated as a nonsegmental perfusion defect the finding of either linear opacity or subsegmental atelectasis on chest radiographs with associated small perfusion defects. To do this, all patients with either atelectasis or linear atelectasis identified in less than 25% of one lung zone with a corresponding abnormality on the perfusion scan were identified. There were 20 such patients. Two of the patients (10%) had PE. However, none (0%) of the 20 lung zones involved had pulmonary emboli in them.

Overall, in the groups investigated, there were 76 patients, of whom six (8%; 95% confidence interval, 2%–14%) had pulmonary emboli. Because nonsegmental perfusion defects were associated with a less than 19% probability of PE, we conclude that they generally are appropriate for low probability. However, since there was considerable variation in the frequency with which individual patterns were associated with PE, we recommend that nonsegmental defects be considered in context of the pattern seen in each patient.

Single Moderate Mismatched Segmental Perfusion Defect with Normal Chest Radiograph. Twenty-eight patients were identified with a single moderate mismatch. Ten of these patients (36%; 95% confidence interval, 18%–54%) had PE. It is clear that this was *not* a valid criterion for low probability. Scans with this finding should be considered intermediate probability for PE.

Any Perfusion Defect with a Substantially Larger Chest Radiographic Abnormality. For this criterion, the ventilation scan find-

TABLE 1
Frequency of PE in Patients with Only Matched V/Q Defects

	No.	% PE	95% C.I.
Patients with V/Q matches meeting original PIOPED criteria for low probability			
Single defect <75% zone	21	24	
Multiple defects <75% zone and <50% lung	39	15	
Overall	60	18	8%–28%
Patients with V/Q matches not meeting original PIOPED criteria for low probability			
Single defect >75% zone	2	50	
Multiple defects, in total <50% either lung	13	15	
Multiple defects, in total >50% one lung	7	14	
Multiple defects, in total >50% two lungs	7	0	
Overall	29	14	1%–27%
Summary (all matched defects)			
Single defect (any size)	23	26	8%–44%
Multiple defects (any size)	66	14	6%–22%
All matched defects	89	17	9%–25%

The observed frequency of pulmonary embolism (% PE) is shown for patterns that did or did not meet the original PIOPED criteria for low probability. The 95% confidence interval (95% C.I.) is shown for summary groups.

ings were irrelevant. Twelve zones were found with this pattern in eleven patients. Two of these patients (18%) had PE. However, only one of the twelve (8%) zones involved had pulmonary emboli in the zone. Furthermore, in the six instances in which radiographic opacity was smaller than 25% of the lung zone, none (0%) of the zones involved had pulmonary emboli. Therefore, this criterion seems valid for low probability, particularly if radiographic opacity is small.

Small Segmental Perfusion Defects with a Normal Chest Radiograph. For this criterion, the findings on the ventilation scan were irrelevant. There were 29 patients in which only small perfusion defects were present. Two of these patients (7%; 95% confidence interval, 0%–16%) had PE. This criterion is therefore appropriate for the low probability category.

Large or Moderate Segmental Perfusion Defects Involving No More Than Four Segments in One Lung and No More Than Three Segments in One Lung Region, with MATCHING Ventilation Defect, Either Equal to or Larger in Size, and Chest Radiograph Either Normal or with Abnormalities Substantially Smaller than Perfusion Defects. In essence, this rather complicated criterion states that if the matched perfusion defect does not involve more than 50% of one lung, or does not involve more than 75% of one lung zone (region), then the scan should be categorized as low probability.

The results in the patient population with matched V/Q defects are summarized in Table 1. We included in this group patients who had hilar or mediastinal abnormalities on radiographs and small defects on scintigrams, since these had been shown to be acceptable for low probability. We excluded from the analysis those with pleural effusions or significant parenchymal abnormalities on the radiograph and those with mismatched defects on the V/Q scintigram.

There were 21 patients who had a single matched V/Q defect, less than 75% of the affected lung zone. Of these, five (24%) had PE. In two (40%) of the five, the embolus was located in the same lung zone as the matched defect. Since this was a quite surprising result, the final categorical readings and consensus probabilities for these patients were reviewed to ensure that no errors in the search algorithm were present. The consensus probability for these patients ranged from 3% to 15% (average, 10%). The categorical reading was low/very low in 3, low probability in 16 and intermediate in 2 (one with a consensus probability of 15% had PE and one with consensus probability of 10% did not).

There were 39 patients who had multiple matched defects, each less than 75% of the affected zone and in some less than 50% of either lung. Six of these patients (15%) had PE. Only 3 of the 147 involved lung zones, however, had emboli in them. In total, there were 60 patients identified who met the above criteria. Of these, 11 (18%) had PE, but in only 5 (3%) of the 168 zones. In general, matched V/Q defects as described in the PIOPED criteria are appropriate for low probability, but patients with single matched defects appear to have a higher likelihood of PE and should be considered for intermediate probability.

We also investigated patients whose matched defects exceeded the limit established for low probability. In 29 patients, 57 lung zones were identified that had matching abnormalities larger than 75% of the zone and correlative angiograms in the zone. There was only one (2%) pulmonary emboli found in the involved lung zones (in a single matched defect), and 4 of the 29 patients (14%) had PE. Of seven patients with defects involving more than 50% of one lung, 14% had PE, whereas none (0%) of seven patients with defects involving more than 50% of the combined lung fields had PE. Of a total of 21 lungs with more than 50% involvement by matched defects, none had emboli in the same lung.

Therefore, patients with extensive chronic obstructive pulmonary disease (COPD) did not have a higher frequency of PE. On

TABLE 2
Frequency of PE in Patients with Various Patterns of Mismatched Perfusion Defects

Type of mismatch	No. of patients	% PE observed (95% C.I.)
1 segmental equivalent	33	52 (35–69)
One large defect	24	46
Two moderate defects	9	67
1.5 segmental equivalents	18	72 (51–93)
One large + one moderate defect	11	73
Three moderate defects	7	71
2 segmental equivalents	7	71 (37–100)
Two large defects	5	80
One large + two moderate defects	1	0
Four moderate defects	1	100 (69–100)
2.5 segmental equivalents	10	100
Two large + one moderate defect	8	100
One large + three moderate defects	1	100
Five moderate defects	1	100

the other hand, single V/Q matches could be associated with a relatively higher frequency of PE, although the difference was not statistically significant. This finding suggests that further analyses which include combined patterns may define other subgroups of patients who have a V/Q match and a higher frequency of PE. The size of the single matched defect, however, did not seem to yield any diagnostic information. Of the six patients who had PE and a single matched defect, the size of the defect was <25% of the zone in three, 25%–50% in one, 51%–75% in one and >75% in one. Of the 17 patients with a single matched defect who did not have PE, the size of the defect was <25% of the zone in six, 25%–50% in eight, 51%–75% in two and >75% in one.

We conclude that multiple matched V/Q abnormalities, even when relatively extensive, are properly categorized as low probability. Single matched defects are borderline between low probability and intermediate, and by the PIOPED interpretive guidelines should be considered intermediate.

Criterion Analysis for Intermediate Probability

The intermediate probability category was not defined explicitly. Accordingly, we do not consider it warranted to conduct an extensive and detailed exploration of hypothetical patterns.

First, as described above, extensive matched V/Q abnormalities were considered intermediate probability in PIOPED, since they exceeded the criterion used for low probability. The data show that this was incorrect and that this criterion for low probability was too stringent. On the other hand, single matched defects should be categorized as intermediate.

Second, as described below, the criteria for high probability may not have been sufficiently stringent, and mismatched perfusion defects involving less than 2.5 segments, although still officially classified as high probability, should be considered as intermediate.

Criterion Analysis for High Probability

Analysis of Mismatched Perfusion Defects. In the PIOPED trial, criteria for "high probability" were two or more large segmental mismatched perfusion defects, or at least one large defect plus two moderate defects, or at least four moderate defects. Small defects were excluded from consideration. These criteria employed the concept of segmental equivalents (*4*), with the high probability cutoff set at two mismatched large perfusion defects or the arithmetic equivalent in moderate defects. The concept of "segmental equivalents" was that moderate perfusion defects can be added together to achieve the same diagnostic significance as the equivalent number of large defects. A *moderate* perfusion defect (one that is 25%–75% of a segment) is considered equivalent to one-half of a *large* perfusion defect (one that is greater than 75% of a segment).

It is thus possible to assemble combinations of moderate and large segments and evaluate their efficacy in detecting pulmonary emboli. Data from the scan description were correlated to the angiographic findings to test the performance of various thresholds for high probability based on this concept.

Mismatches Not Satisfying the Original PIOPED Criteria for High Probability. As noted in the discussion of the criteria for low probability, a single moderate mismatched perfusion defect is properly considered intermediate, with a 36% probability for PE. Table 2 illustrates the data accumulated for other subsets of the PIOPED V/Q scan population who had mismatched perfusion defects and pulmonary angiograms interpreted as presence or absence of PE. These data indicate that a single moderate defect, a single large defect (or equivalent) and a large plus a moderate mismatched perfusion defect (or equivalent) are patterns which should be considered "intermediate probability" for PE.

Mismatches Satisfying the Original PIOPED Criteria for High Probability. Criteria for "high probability" were two or more large defects, one large and two moderate defects or four moderate defects. We analyzed 101 cases that the central readers called high probability for PE based upon mismatched segmental defects and who had angiograms interpreted as PE present. Characteristics of mismatched defects in these patients are detailed in Table 3. Note that the numbers of patients for each degree of mismatch do not equal those in Table 2, since Table 3 refers to only those patients who were called "high probability" by independent central readers, whereas Table 2 includes all patients who were read by the consensus reading teams. Since the consensus description was not necessarily performed by independent central readers, disagreements were possible and did occur between the central readings of scan category and the category corresponding by the PIOPED criteria to the consensus descriptions.

A mismatch of two segments, the cutoff used in the PIOPED trial, does not seem to be a particularly good threshold for the high probability category, since only 71% of such patients had PE on

TABLE 3
Magnitude of Segmental Mismatch in 101 Cases of Angiographically Proven PE Categorized as "High Probability"

No. of segments	Cases	Cumulative (%)
9 or more	47	48
4.5–8.5	31	77
4	3	80
3.5	2	82
3	6	88
2.5	7	95
2	2	97
1.5	2	99
1	1	100

angiography. In contrast, however, all ten of the patients with 2.5 segmental mismatches had PE. In addition, this was a relatively uncommon pattern. Finally, it should be noted from inspection of Tables 2 and 3 that the central readers had more difficulty with lesser degrees of mismatch. Only two of seven patients with two segments of mismatch were called high probability, whereas seven of ten patients with 2.5 mismatched segments were categorized correctly. Cases with larger numbers of mismatched defects were categorized accurately as high probability.

Absent or decreased perfusion in combination throughout an entire lung was seen in 23%, whereas bilateral mismatched defects were present in 80% of the 101 patients in this series. Thus, a solitary whole lung perfusion defect ("the white-out") is a relatively uncommon appearance for PE. To use the scan description form to investigate whether a solitary *lobar* defect was commonly present in PE, it was necessary to assume that a lobar perfusion defect consisted of three segmental lesions confined to one upper lobe, four segmental lesions confined to one lower lobe and two segmental lesions in either the lingula or the middle lobe. By using these assumptions, only two patients in this series had solitary lobar defects. Thus, a solitary lobar perfusion defect is an unusual presentation for PE.

Use of Percent Probability Estimates

The relationship between the frequency of angiographically proven PE and the "experiential" or "gestalt" percent probability (obtained by averaging the percent probability estimate listed for the two final readers) is shown in Table 4 and Figure 1.

DISCUSSION

Perfusion lung scintigraphy is sensitive but not specific for PE. Years of experience with perfusion imaging has demonstrated that nearly all pulmonary diseases, including neoplasms, infections and COPD, can produce decreased pulmonary arterial blood flow to affected lung zones (*5*).

More than two decades ago Wagner (*6*) and DeNardo (*7*) suggested combined V/Q lung imaging as a means to improve the specificity of radionuclide methods for diagnosing pulmonary emboli. This recommendation came nearly 15 yr after Knipping and colleagues (*8*) pioneered the use of ^{133}Xe to study pulmonary ventilation.

McNeil et al. (*9*) highlighted the findings of numerous investigators by pointing out that abnormalities in the perfusion scan matched by zones of abnormal ventilation rarely represent pulmonary emboli, whereas mismatched abnormalities, given a normal chest radiograph, have a high correspondence with the angiographic diagnosis of PE. Alderson and coworkers (*10*) later demonstrated that the overall diagnostic accuracy for scintigraphic detection of pulmonary emboli was significantly improved when ^{133}Xe ventilation studies were added to perfusion scans and chest radiographs.

TABLE 4
Scan Reader Estimated % Probability for PE Correlated to Angiogram

Scan % probability	Angiogram PE+	Angiogram PE−	% with PE
0– 10.9	19	176	10
11– 20.9	32	140	19
21– 35.9	39	99	28
36– 64.9	53	49	52
65– 79.9	12	5	71
80– 89.9	30	5	86
90–100	66	6	92

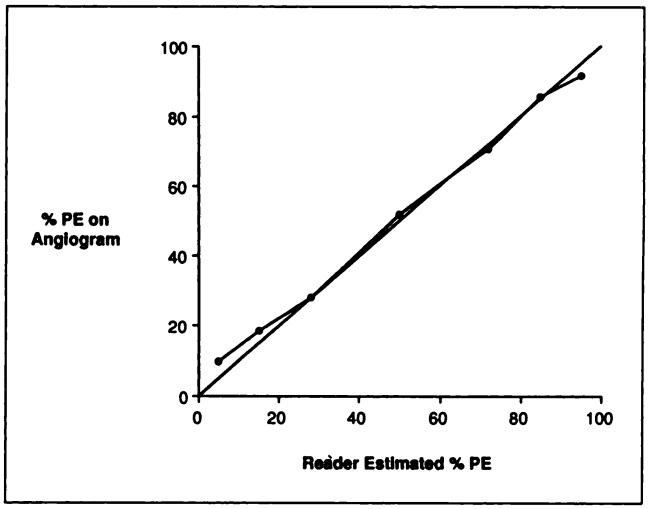

FIGURE 1. Readers' "gestalt" estimate of likelihood of PE versus incidence of emboli found on angiography. Solid line is line of identity.

Gottschalk and colleagues (*4*) introduced the concept of "segmental equivalents" (i.e., that two subsegmental perfusion defects may be added to produce the same diagnostic significance as a single segmental defect). A subsequent retrospective study by Kotlyarov and Reba supported the usefulness of this approach (*11*).

Extensive work by Biello and collaborators (*12,13*) further categorized perfusion defects matched by ventilatory or radiographic abnormalities and provided grounds for reducing the number of "indeterminate" diagnoses. Further evaluation of this work (*14*) indicated that this diagnostic scheme provides improved interobserver consistency and a 30% reduction in "indeterminate" readings than results from an older scheme.

Before PIOPED, the probability of PE for a number of specific image findings, such as the combination of a segmental and a subsegmental mismatch or the finding of two subsegmental mismatches, had not been ascertained. The original Biello criteria (*12*) categorized the former pattern as high probability for PE, since the high probability category was defined as findings of "one or more segmental mismatches" in zones with a normal radiographic appearance. Data (*10,11,15*) published subsequent to that original paper, however, indicated the need to employ two zones of mismatch for a high probability categorization. Thus, a study showing a segmental mismatch and two subsegmental mismatches would be considered high probability for PE if it showed the equivalent involved volume of two segmental mismatches (*15*). Little data existed, however, regarding other specific image pattern subcategories.

In assessing the utility of various criteria for low proba-

TABLE 5
Revised PIOPED V/Q Scan Criteria

High Probability (≥80%)
≥2 Large mismatched segmental perfusion defects or the arithmetic equivalent in moderate or large + moderate defects*.
Intermediate Probability (20%–79%)
One moderate to two large mismatched segmental perfusion defects or the arithmetic equivalent in moderate or large + moderate defects*.
Single matched ventilation-perfusion defect with clear chest radiograph†.
Difficult to categorize as low or high, or not described as low or high.
Low Probability (≤19%)
Nonsegmental perfusion defects (e.g., cardiomegaly, enlarged aorta, enlarged hila, elevated diaphragm).
Any perfusion defect with a substantially larger chest radiographic abnormality.
Perfusion defects matched by ventilation abnormality† provided that there are: (1) clear chest radiograph and (2) some areas of normal perfusion in the lungs.
Any number of small perfusion defects with a normal chest radiograph.
Normal
No perfusion defects or perfusion outlines exactly the shape of the lungs seen on the chest radiograph (note that hilar and aortic impressions may be seen and the chest radiograph and/or ventilation study may be abnormal).

*Two large mismatched perfusion defects are borderline for "high probability." Individual readers may correctly interpret individual scans with this pattern as "high probability." In general, it is recommended that more than this degree of mismatch be present for the "high probability" category.

†Very extensive matched defects can be categorized as "low probability." Single V/Q matches are borderline for "low probability" and thus should be categorized as "intermediate" in most circumstances by most readers, although individual readers may correctly interpret individual scans with this pattern as "low probability."

bility of PE, one must remember that in the PIOPED study, when only patients who had pulmonary angiograms were considered, the frequency of PE in patients in the low and very low probability categories combined was 15%. This rather high frequency stems from two factors. First, a large number of patients with low probability and very low probability scans (and presumably, low clinical suspicion) were withdrawn from the study prior to pulmonary angiography. If the outcome committee classification is used to categorize these patients as PE present or absent, the frequency of PE in the combined low and very low probability categories is 11%. Second, the original PIOPED criteria for low probability included a single moderate segmental mismatch. If this finding places the scan in the intermediate category, the frequency of PE in the low/very low probability patients with angiograms would be reduced to 13% and in the whole cohort of low/very low probability patients would be reduced to 9%. It is interesting in this regard that the retrospective literature (*16*) indicates that small segmental defects are essentially never associated with PE, whereas in PIOPED, 7% of the cases with small segmental defects identified had PE. It is important to note that in PIOPED pulmonary angiography was performed usually within 12 hr of the V/Q scan and always within 24 hr. This is in sharp distinction to the retrospective data, where pulmonary angiography often was performed up to 3 days later. If it is postulated that small emboli can lyse and disappear rapidly, this could suggest that there may be patients with emboli in the PIOPED group who do not need to be treated. This interpretation, however, would be difficult to test.

As mentioned above, the analyses reported here focus upon individual criteria. Most types of combined patterns were excluded in analyses of low probability. For example, a patient with a matched perfusion/ventilation defect plus a moderate segmental mismatch would be excluded from the analysis of the significance of matched perfusion defects to the "low probability" category. This implies ineluctably that many patients in the study database do not contribute to this analysis. It implies equally that other scintigraphic patterns exist which may be of great importance to increasing the accuracy of V/Q scan interpretation. It remains for future analyses to address this issue, which is of obvious importance and may help to explain the good performance of the "gestalt" percent probability readings.

Our retrospective analyses indicate that three major adjustments to the PIOPED criteria should be considered. These adjustments to the PIOPED criteria are included in the revised criteria shown in Table 5. It is recognized that these revised criteria should themselves be subjected to prospective testing.

First, it is incorrect to consider a single moderate segmental mismatch a criterion for low probability. Scans demonstrating this pattern should be classified as intermediate.

Second, the PIOPED criteria for interpreting matched perfusion defects should be modified. The data indicate that if matched V/Q defects are identified within a clear region of the chest radiograph, the matched lesions do not usually hide a pulmonary embolus. Thus, even an extensive V/Q match is an acceptable criterion for low probability. Although this appears inconsistent with prior data, it

may actually not be. Previous authors (*12*) have presented data which suggest a higher frequency of PE (19% of patients; 95% confidence intervals, 0%–38%) in patients with very extensive matched V/Q defects due to COPD when compared to those patients with lesser degrees of matched defects (3% of patients; 95% confidence intervals, 0%–8%). However, due to the small number of patients in both groups, there is not a statistically significant difference between the two groups in that study (*12*) or between that study and the present report. If any trend exists in our data, it is that more extensive matched defects have a lower association with PE. Note, however, that our data analysis assumes that there are some areas of the lung in which normal ventilation and perfusion occur. Note also that *single* areas of a V/Q match with a clear radiograph have a somewhat higher likelihood of PE than multiple defects of similar size, and we have thus categorized them as intermediate in Table 5. Although the difference between single and multiple defects is not statistically significant, the general guideline in PIOPED was to assign cases which were not clear-cut to the intermediate category. It is in this general area that we consider further analyses will most likely reveal useful additional data. For example, the present analysis does not address the significance of combined patterns which include matched V/Q defects, since such combined patterns were not enumerated in the PIOPED criteria.

Third, two large mismatched defects did not provide a reliable interpretation of "high probability," whereas 2.5 mismatched "segmental equivalents" provided more accurate categorization. We cannot definitely recommend changing the threshold for high probability, however, since the number of patients with this pattern is very small. However, the data are suggestive, and also it was striking that studies with more mismatch were more likely to be read accurately. Patients with high probability V/Q scans and angiographically documented PE typically had a large number of mismatched segments (85% of patients had >3 mismatched large defects and 48% had >9 mismatched large defects). Thus, an accurate high probability categorization usually is very easy to make. Any difficulty in assigning a scan to the high probability category should lead the reader to consider the intermediate category, as suggested by both the original and revised PIOPED criteria.

Certain adjunctive signs were tested in PIOPED. In 1982, Sostman and Gottschalk (*17*) described the "stripe sign" as an indication that a region with a perfusion defect showing the sign did not contain a pulmonary embolus. In the initial data these authors presented, 92% of zones with a stripe sign had no pulmonary emboli. This criterion was tested in PIOPED, where we found that 79 of 85 (93%) lung zones with the stripe sign had no PE present on angiography (*18*). This sign remains a useful finding. In 1985, Bedont and Datz described pleural effusions that caused an isolated perfusion defect on V/Q scans (*19*). When the size of this perfusion defect was equal to the size of the pleural effusion, these authors found a low (4%) incidence of PE. In PIOPED, we found that 4 of 27 (15%) patients had PE. Consequently, this sign did not perform as well in PIOPED as it had previously, although it still falls within the low probability range.

The "gestalt" or "experiential" percent probability estimates correlated well with the frequency of PE on angiography. Similar data have been published previously (*20*). This result may call into question the use of standardized criteria for categorical interpretation of V/Q scans. Indeed, it is clear that even the specific criteria validated by the PIOPED experience must be considered as representing a range of probabilities of PE. Certainly, combinations of radiographic and scintigraphic findings may have quite different implications from the same findings when present in isolation. Therefore, all readers of lung scans must still exercise appropriate judgment when interpreting individual cases. However, because the members of the Nuclear Medicine Working Group had particularly extensive experience in the interpretation of V/Q scans, we have no way of knowing how the data regarding "gestalt" probabilities should be utilized most effectively by others with less experience. We suggest, however, that experienced readers should consider incorporating percent probability estimates into the scan interpretation report along with the scan category interpretation. For example, it is likely that a reading of "low probability" with the likelihood of PE estimated to be approximately 15% has a different patient management connotation to the referring physician than a reading indicating a "low probability" for PE with the estimated probability being approximately 5%. Further research into combined scintigraphic patterns may yield further insight into lung scan interpretation.

SUMMARY AND CONCLUSIONS

In summary, a review of the criteria used in the PIOPED study documents that most of the criteria appropriately categorized V/Q scans. However, we recommend reconsideration of three criteria:

1. A single moderate perfusion defect be appropriately categorized as intermediate rather than as low probability.
2. Multiple and relatively extensive matched V/Q abnormalities are appropriate for low probability, provided that the chest radiograph is clear. On the other hand, single matched defects may be better categorized as intermediate probability, although this cannot be definitely validated statistically.
3. Two segmental mismatches may not be the optimum threshold for high probability, and in some cases should be considered for intermediate probability. However, due to the small number of cases with this finding, no definite, statistically founded recommendation can be made.

We suggest that the revised criteria resulting from these adjustments should now be used for the interpretation of

V/Q scans. Further studies of the PIOPED database are possible and desirable to evaluate, for example, the diagnostic import of mixed scan patterns not specifically enumerated in the PIOPED criteria.

Finally, physicians who are experienced in the interpretation of V/Q scans can predict the likelihood of PE, based upon their personal experience with an accuracy comparable to that obtained by use of categorical criteria.

ACKNOWLEDGMENTS

The members of the Nuclear Medicine Working Group gratefully acknowledge the hard work and intellectual contributions of their co-investigators at the Maryland Medical Research Institute and the National Heart Lung and Blood Institute to the PIOPED study. Supported by contracts N01-HR-34007 through N01-HR-34013 of the National Heart, Lung and Blood Institute.

REFERENCES

1. 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.
2. Bone RC. Ventilation/perfusion scans in pulmonary embolism: the emperor is incompletely attired. *JAMA* 1990;263:2794–2795.
3. Biello DR, Mattar AG, Osei-Wusu A, et al. Interpretation of indeterminate lung scintigrams. *Radiology* 1979;133:189–194.
4. Neumann RD, Sostman HD, Gottschalk A. Current status of ventilation-perfusion imaging. *Semin Nucl Med* 1980;10:198–217.
5. Secker-Walker RH, Siegel BA. The use of nuclear medicine in the diagnosis of lung disease. *Radiol Clin North Am* 1973;11:215–241.
6. Wagner HN Jr, Lopez-Majano V, Langan JK, et al. Radioactive xenon in the differential diagnosis of pulmonary embolism. *Radiology* 1968;91:1168–1174.
7. DeNardo GL, Goodwin DA, Ravasini R, et al. The ventilatory lung scan in the diagnosis of pulmonary embolism. *N Engl J Med* 1970;282:1334–1336.
8. Knipping HW, Bolt W, Venrath H, et al. Eine neue methode zur prüfung der herz-und lungenfunktion die regionale funktionsanalyse in der lungen und herzklinik mit hilfe des radioaktiven endelgases xenon-133 (isotopen thorakographie). *Deutsch Med Wochenschr* 1955;80:1146–1147.
9. McNeil BJ, Holman L, Adelstein J. The scintigraphic definition of pulmonary embolism. *JAMA* 1974;227:753–756.
10. Alderson PO, Rujanavech N, Secker-Walker RH, et al. The role of ^{133}Xe ventilation studies in the scintigraphic detection of pulmonary embolism. *Radiology* 1976;120:633–640.
11. Kotlyarov EV, Reba RC. The concept of using abnormal V/Q segment equivalents to refine the diagnosis of pulmonary embolism [Abstract]. *Invest Radiol* 1981;16:383.
12. Alderson PO, Biello DR, Sachariah KG, et al. Scintigraphic detection of pulmonary embolism in patients with obstructive pulmonary disease. *Radiology* 1981;138:661–666.
13. Biello DR, Mattar AG, McKnight RC, et al. Ventilation-perfusion studies in suspected pulmonary embolism. *AJR* 1979;133:1033–1037.
14. Carter WD, Brady TM, Keyes JW, et al. Relative accuracy of two diagnostic schemes for detection of pulmonary embolism by ventilation-perfusion scintigraphy. *Radiology* 1982;145:447–451.
15. Rosen JM, Palestro CJ, Markowitz D, Alderson PO. Significance of single V-P mismatches found in Kr-81m/Tc-99m lung scans. *J Nucl Med* 1986;27:361–365.
16. Alderson PO, Martin EC. Pulmonary embolism: diagnosis with multiple imaging modalities. *Radiology* 1987;164:297–312.
17. Sostman HD, Gottschalk A. The stripe sign: a new sign for diagnosis of nonembolic defects on pulmonary perfusion scintigraphy. *Radiology* 1982;142:737–741.
18. Sostman HD, Gottschalk A. Prospective validation of the stripe sign in ventilation-perfusion scintigraphy. *Radiology* 1992;184:455–459.
19. Bedont RA, Datz FL. Lung scan perfusion defects limited to matching pleural effusions: low probability of pulmonary embolism. *AJR* 1985;145:1155–1160.
20. 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.