Critical Review Form
Clinical Practice Guidelines

Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Suspected Pulmonary Embolism, *Ann Emerg Med* 2011; 57: 628-652

Objectives: **To help clinicians answer the following critical questions:**

1) **Do objective criteria provide improved risk stratification over gestalt clinical assessment in the evaluation of patients with possible PE?**
2) **What is the utility of the Pulmonary Embolism Rule-Out Criteria (PERC) in the evaluation of patients with suspected PE?**
3) **What is the role of quantitative D-dimer testing in the exclusion of PE?**
4) **What is the role of CT pulmonary angiogram of the chest as the sole diagnostic test in the exclusion of PE?**
5) **What is the role of venous imaging in the evaluation of patients with suspected PE?**
6) **What are the indications for thrombolytic therapy in patients with PE?**

This critical appraisal will not evaluate Question #6 – instead refer to our July 2010 Journal Club where this topic was reviewed in detail.

Methods: **This updates the 2003 ACEP clinical policy (Ref). This policy was reviewed by representative of the American College of Chest Physicians, ACR, ACEP's Emergency Ultrasound section, and ACEP's Quality and Performance Committee, although their endorsement is not implied.**

Evidence quality was graded by at least 2 subcommittee members.

Design/ Class	Therapy[†]	Diagnosis[‡]	Prognosis[§]
1	Randomized, controlled trial or meta-analyses of randomized trials	Prospective cohort using a criterion standard	Population prospective cohort
2	Nonrandomized trial	Retrospective observational	Retrospective cohort Case control
3	Case series Case report Other (eg, consensus, review)	Case series Case report Other (eg, consensus, review)	Case series Case report Other (eg, consensus, review)

[*]Some designs (eg, surveys) will not fit this schema and should be assessed individually.

[†]Objective is to measure therapeutic efficacy comparing ≥2 interventions.
[‡]Objective is to determine the sensitivity and specificity of diagnostic tests.
[§]Objective is to predict outcome including mortality and morbidity.

Next, subcommittee members graded the evidence on 6 dimensions felt to be most important for guideline development: blinding of outcome assessment, blinded/randomized allocation, direct or indirect outcome measures, bias (selection, detection, transfer), external validity and sample size. The recommendations generated were then classified as Level A, B, or C.

Approach to downgrading strength of evidence.

	Design/Class		
Downgrading	1	2	3
None	I	II	III
1 level	II	III	X
2 levels	III	X	X
Fatally flawed	X	X	X

Level A – high degree of clinical certainty based on Class 1 strength of evidence or preponderance of Class II studies.
Level B – moderate clinical certainty based on Class II studies or strong consensus of Class III studies.
Level C – other strategies for patient management that are based on Class III studies, or in the absence of any adequate published literature, based on panel consensus.

Searches were limited to English language studies and human studies. The key phrases for each question are listed.

Question 1: Risk stratification, PE, ED, emergency services, risk assessment diagnostic strategies, Well's criteria, Wicki criteria, Kline criteria, Geneva score, revised Geneva score, PISA model, and variations and combinations of the key words/phrases; years 2000 through Dec. 2009.

Question 2: PERC, pulmonary embolism rule out criteria, block rule, PE, and various combinations of the key words/phrases; year's 2000-Dec. 2009.

Question 3: PE, Fibrin Fragment D, sensitivity, specificity, D-dimer, differential diagnosis, and variations and combinations of the key words/phrases, years 2001-Dec. 2009.

Question 4: X-ray computed tomography, CT, spiral computed tomography, PE, angiography, thromboembolic, outcome, follow-up, recurrent, morbidity, mortality, false negative, false positive, prognosis, treatment outcome, and variation and combination of the key words/phrases, years 2001-Dec. 2009.
Question 5: PE, venous ultrasonography, sensitivity, specificity, probability, likelihood, and variations and combinations of the key words/phrases, years 2002 through Dec. 2009.

Guide		Comments
I.	**Are the Recommendations Valid?**	Answer questions IA-D below
A.	**Did the recommendations consider all relevant patient groups, management options, and possible outcomes?**	No. "This guideline is **not intended to address the care of patients with PE in the presence of cardiac arrest or pregnancy, patients with absence of symptoms suggestive of PE, or pediatric patients.**" (p. 630)
B.	**If necessary, was an explicit, systematic, and reliable process used to tap expert opinion?** *You should look for a clear description of how the panel was assembled along with the members' specialties and any organizations they are representing.*	No. The process of ACEP Guideline committee or subcommittee members is not transparent. The expertise of sub-committee members for the diagnostic and therapeutic management of PE is not stated. At the end of the guideline the following CDI disclosure is provided: "There are no relevant industry relationships disclosed by the subcommittee or committee members." (p. 646)
C.	**Is there an explicit, systematic specification of values or preferences?** *Panelists' ratings presumably reflect the risk-benefit trade-offs of specific interventions, but whether other physicians or patients themselves would make the same decisions remains uncertain. Whether given options are value or preference related should be clearly stated in the guideline.*	Maybe. Although the authors do not explicitly describe the reference point for values or preferences (patient, family, nurse, physician, third party payer or society), they do provide outcomes of interest for all perspectives.
D.	**If the quality of the evidence used in originally framing the criteria was weak, have the criteria themselves been correlated with patient outcomes?** *When the studies utilized to produce guidelines are less than randomized-controlled trials, conclusions can be strengthened by noting how outcomes can be correlated with adherence to the guidelines.*	No, the results of weaker evidence were not consistently correlated with patient outcomes.

II.	**Were the Criteria Applied Appropriately?**	Answer questions II A-B below.
A.	**Was the process of applying the criteria reliable, unbiased, and likely to yield robust conclusions?**	The ACEP guidelines have not been field tested. Whether each of the recommendations can be accurately and reliably applied across heterogeneous emergency settings (rural/urban, academic/community boarded EP/non boarded EP) and whether patients/society/malpractice attorney will accept and respect these evidence-based/consensus guidelines remains to be determined.
B.	**What is the impact of uncertainty associated with evidence and values on the criteria based ratings of process of care?**	There is no discussion about uncertainty around the point estimates of diagnostic accuracy or individual trials' exclusion criteria limiting the external validity of diagnostic strategies in dissimilar populations.
III.	**How Can I Apply the Criteria to Patient Care?**	

A.	**Are the criteria relevant to your practice setting?** *Medical practice is shaped by an amalgam of evidence, values, and circumstances; clinicians should consider their local medical culture and practice circumstances before importing a particular set of audit criteria.* **Geneva Criteria**	Yes, despite the above critiques, this clinical policy represents the most up-to-date critical appraisal of an overwhelming body of literature from the perspective of EP's. Q1: • **Level B recommendation** "either objective criteria or gestalt clinical assessment can be used to risk stratify patients with suspected PE. There is insufficient evidence to support the preferential use of one method over another. • Geneva Criteria – the simplified revised Geneva score is based on one Class II study (Klok 2008) of 330 patients (65% of cohort) with "PE unlikely" score of 0-2 and a negative D-dimer no patient was diagnosed with VTE on presentation or 3-month follow-up. (p. 631) • Wells Criteria – simple scoring system prospectively validated in a Class II investigation from 4 ED's at tertiary care Canadian

Variable	Points: Revised Geneva	Points: Simplified Revised Geneva
Risk factors		
Age ≥65 y	1	1
Previous DVT/PE	3	1
Recent surgery/fracture (4 wk)	2	1
Active malignancy	2	1
Symptoms		
Unilateral lower-limb pain	3	1
Hemoptysis	2	1
Clinical signs		
Heart rate		
75–94 beats/min	3	1
≥95 beats/min	5	2*
Pain on lower-limb deep venous palpation and unilateral edema	4	1

Score Range	Probability of PE, % (95% CI)	Patients With This Score, %	Interpretation of Risk
Revised Geneva Score[20]			
0–3	7.9 (5.0–12.1)	37.0	Low
4–10	28.5 (24.6–32.8)	57.4	Moderate
11–25	73.7 (61.0–83.4)	5.5	High
Simplified Revised Geneva Score[21]			
Traditional interpretation			
0–1	7.7 (5.2–10.8)	36.0	Low
2–4	29.4 (25.9–33.1)	60	Moderate
5–7	64.3 (48.0–78.5)	4.0	High
Alternative interpretation			
0–2	12.9 (10.5–15.7)	64.9	PE unlikely
3–7	41.6 (36.5–46.8)	35.1	PE likely

*The original table from Klok et al[21] lists 1 point for heart rate ≥95 beats/min, but the assessment of score states, "[b]ecause of the weight of heart rate in the original score, we attributed 1 point to a heart rate between 75 and 94 beats/min and an additional point for a heart rate of 95 beats/min or more." Thus, a patient with a heart rate of 100 beats/min would receive a total of 2 points (personal communication, F. A. Klok, MD, PhD, Department of General Internal Medicine, Leiden University Medical Center/Bronovo Hospital Den Haag, May 2010).

Well's Criteria (below)

Criteria	Points
Suspected DVT	3.0
An alternative diagnosis is less likely than PE	3.0
Heart rate >100 beats/minute	1.5
Immobilization or surgery in the previous 4 weeks	1.5
Previous DVT/PE	1.5
Hemoptysis	1.0
Malignancy (on treatment, treated in the last 6 months, or palliative)	1.0

Score Range, Points	Probability of PE (%)	% With This Score	Interpretation of Risk
Traditional interpretation			
0–1	3.6 (2.0–5.9)	40.3	Low
2–6	20.5 (17.0–24.1)	52.6	Moderate
>6	66.7 (54.3–77.6)	7.1	High
Alternate interpretation			
0–4	7.8 (5.9–10.1)	71.5	PE unlikely
>4	40.7 (34.9–46.5)	28.5	PE likely

Kline Rule (below)

Hospitals (Wells 2001) with 1.3% PE incidence in low pretest prob, (95% CI 0.5-2.7%) 16.2% (95% CI 12.5-20.6%) in moderate pre-test prob, and 40.6% (95% CI 28.7-53.7%) in the high pre-test prob group of 437 patients with low pre-test prob and negative D-dimer only 1 (0.2%) developed a PE during follow-up. Another Class II investigation (van Belle 2006) dichotomized patients into "PE unlikely" or "PE likely" in a multicenter evaluation of 3306 patients finding 66.7% PE unlikely and 1057 of those also had a negative D-dimer with 5/1057 (0.5%) developing VTE at 3-months with no deaths. In the PE likely group, CT confirmed PE in 20.4%. Although multiple Class II (Anderson 2005, Kabrhel 2009) and Class III (Chagnon 2002, Moores 2004, Miniati 2005, Ollenberger 2006, Goekoop 2007, Hogg 2006, Kruip 2002, Sanson 2000, Wolf 2004, Yap 2007) have validated Wells score for risk stratification, the major critique is that the subjective variable represents physician judgment overriding the objective elements of the rule and moving patients to intermediate probability (Wicki 2001, Klok 2008, Iles 2003, Testuz 2006)

- Kline Rule – derived in a Class II study from prospective (before imaging)

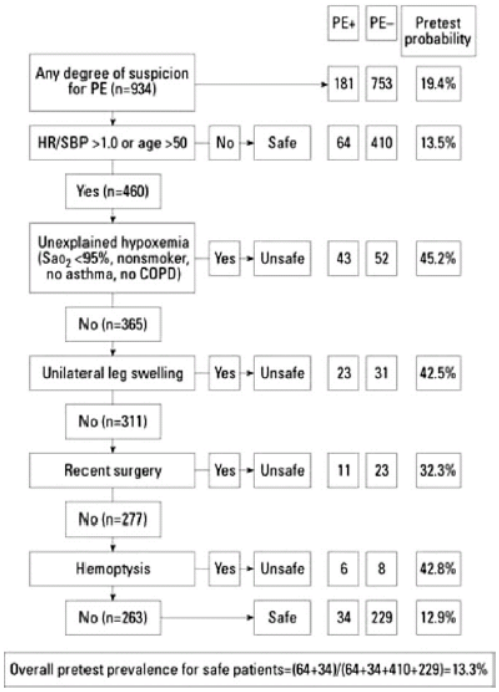


Figure. Kline decision rule for excluding PE.[23] Reprinted from *Annals of Emergency Medicine*, 39, Kline JA, Nelson RD, Jackson RE, et al. Criteria for the safe use of D-dimer testing in emergency department patients with suspected pulmonary embolism: a multicenter US study, 144-152, 2002, Copyright from the American College of Emergency Physicians, [2002].

Flow diagram demonstrating the Kline decision rule in selecting patients in whom D-dimer assay less than 500 ng/ml can reliably rule out PE. This decision rule splits the patients into 2 groups, four fifths of whom are eligible for D-dimer testing ("safe" patients with pretest probability of PE of 13.3%) and one fifth of whom are ineligible for D-dimer testing ("unsafe" patients with pretest probability of 42.1%).

data collection in 7 U.S. EDs to categorize under 80% of patients as safe to exclude PE with D-dimer. 197/934 (21%) met high risk criteria with 42.1% (95% CI 35.5%-49.6%) having a PE. In the non-high risk patients the incidence of PE was 13.3% (95% CI 10.9%-15.9%) and authors opined that a D-dimer with LR-<0.07 would reduce PE probability to under 1%. No prospective studies had validated these criteria.

- Pisa model – Class II study to stratify patients as low (≤ 10% PE probability), intermediate (> 10% to 50% probability), moderately high (>50% to 90%) or high (>90% probability) from 1100 consecutive patients with suspected PE at a single institution. The results were low probability 39% with 4% PE, intermediate 26% with 22% PE, moderately high 7% with 74% PE, and high 28% with 98% PE. Miniati has validated the PISA model after removing the chest x-ray from the model.
- Multiple comparison studies (Klok 2008, Runyon 2005, Chagnon 2002, Miniati 2005, Ollenberger 2006) have not suggested that one CDR is superior over another.
- Multiple studies (PIOPED I, Barghouth 2000, Musset 2002, Nilsson 2001, Perrier 1996, Perrier 2000, Kabrhel 2009, Sanson 2000, Runyon 2005) have suggested that CDR's are neither more

		accurate nor more reliable than clinical gestalt although non-significant trends favor the more experienced clinician since the more experienced clinicians "recognize the difficulties in ruling out PE and are reluctant to exclude it on clinical grounds." (p. 635) Kabrhel Well's / Gestalt / Actual Incidence PE Low (69%) / Low (68%) / 3% Int (28%) / Int (26%) / 13% High (3%) / High (6%) / 36% Q2: • **Level B recommendation** "In patients with a low pretest probability for suspected PE, consider using the PERC to exclude the diagnosis based on historical and physical examination data alone." • PERC rule was derived in a Class II prospective study of 3148 ED patients with suspected PE (Kline 2004) and validated in a separate Class II multicenter prospective trial (Kline 2008). In the first validation trial 20.4% of patients were very low risk by PERC and the negative LR was 0.12 (Carpenter 2009). The second validation was a Class III trial of 120 consecutive ED patients and yielded LR^+ 1.3, LR^- 0. The rule was designed for use in patients considered low risk by clinical gestalt to avoid ordering D-dimers with high false positive rates.

Table 4. The Pulmonary Embolism Rule Out Criteria (PERC).

The PERC criteria negative (PERC–) require the clinician to answer no to the 8 questions below.[60] If a patient is low risk by gestalt impression and PERC–, the posttest probability of venous thromboembolism is <2%. Reprinted with permission. Copyright © John Wiley & Sons Ltd, Publisher. Kline JA, Courtney DM, Kabrhel C, et al. Prospective multicenter evaluation of the pulmonary embolism rule-out criteria. *J Thromb Haemost*. 2008;6:772-780.

1. Is the patient older than 49 years of age?
2. Is the pulse rate greater than 99 beats/min^{-1}?
3. Is the pulse oximetry reading <95% while the patient breathes room air?
4. Is there a present history of hemoptysis?
5. Is the patient receiving exogenous estrogen?
6. Does the patient have a prior diagnosis of venous thromboembolism (VTE)?
7. Has the patient had recent surgery or trauma (requiring endotracheal intubation or hospitalization in the previous 4 weeks)?
8. Does the patient have unilateral leg swelling (visual observation of asymmetry of the calves)?

- No prospective outcome studies have evaluated PERC impact on clinical decision making, test ordering or patient outcomes.

Q3:

- **Level A recommendation** "in patients with a low pretest probability for PE a negative quantitative D-dimer assay result can be used to exclude PE.
- Level C recommendation "In patients with an intermediate pretest probability for PE, a negative quantitative D-dimer assay result may be used to exclude PE.
- One RCT (Kearon 2006) using a qualitative D-dimer assay demonstrated that in low risk Wells those randomized to home after negative D-dimer vs. V/Q after negative D-dimer had no difference in 6-month outcomes.
- Class I systematic reviews (Di Nisio 2007, Brown 2002, Brown 2003) identify quantitative D-dimer **pooled sensitivity 0.93-0.96** and **pooled specificity 0.39-0.51**. translating to LR^- under 0.01 which would reduce low risk probability from 10% to 1% supported by multiple Class I (van Belle 2006), II (Courtney 2010, Perrier 2005, Steeghs 2005, Than 2009) and Class III (Goekoop 2007, Hogg 2006, Goergen 2005, Carrier 2009, Legnani 2010, Runyon 2008) studies.
- The use of D-dimer testing alone to exclude PE in

		intermediate probability patients is controversial as Righini noted upper limit of 95% CI 5% for 3-month VTE in that subgroup. • False-positive rates of D-dimer are higher in older and pregnant patients (Brown 2002, Sohne 2005) so different thresholds for abnormal may exist in this subset. The impact of raising the D-dimer threshold on sensitivity needs to be examined (Le Gal 2006). • One decision analysis concluded that using D-dimer is not cost effective if CT is available but his was based on the assumption that the patient had no other completing diagnosis. (Duriseti 2006) Q4: • **Level B recommendation** "for patients with a low or PE unlikely (Wells score ≤4) pretest probability for PE who require additional diagnostic testing (e.g. positive D-dimer result or highly sensitive D-dimer test not yet available), a negative multi detector CT pulmonary angiogram alone can be used to exclude PE." • Level C recommendation 1) for patients with an intermediate pretest probability for PE and a negative CT pulmonary angiogram result in whom a clinical concern for PE still exists and CT venogram has not already been performed, consider additional diagnostic testing (e.g., D-dimer, lower extremity imaging, VQ

		scanning, traditional pulmonary arteriography) prior to exclusion of VTE disease. 2) For patients with a high pretest probability for PE and a negative CT angiogram result, and CT venogram has not already been performed, perform additional diagnostic testing (e.g., D-dimer, * lower extremity imaging, VQ scanning traditional pulmonary arteriography) prior to exclusion of VTE disease. *A negative, highly sensitive, quantitative D-dimer result in combination with a negative multidetector CT pulmonary angiogram result theoretically provides a posttest probability of VTE less than 1%. • This critical appraisal will only focus on multi detector CT's which have been evaluated by one Class II (Stein 2006) study and two Class III studies (Winer-Muram 2004, Coche 2003) and one Class III meta-analysis (Stein PD 2007) with sensitivities 83-90% and specificity 94-100% (see July 2011 Journal Club for PIOPED II Synopsis). • Clinician should use caution when evaluating high pretest probability patients with a negative CTPA. Of 13 prospective trials three Class I (van Belle 2006, Musset 2002, Righini 2008) , four Class II (Anderson 2007, Perrier 2004, Stein 2006, van Strijen 2003), and six Class III (Anderson

		2007, Perrier 2004, van Belle 2006, Righini 2008, Stein 2006) including 5 using multi detectors CT) the following subsequent PE rates (between 3-6 months) were found in those with negative CTP who did not receive anticoagulation: Study* — PE at follow-up Christopher 2006 — 1.3% (0.7-2%) Anderson — 1.7% Van Strijen — 2% Musset 2002 — 1.8% PIOPED II 2003 — 17% Study* — PE at follow-up Vigo D-dimer ⊕ — 19.7% Vigo D-dimer ⊖ — 1.17% *Christopher, Anderson, Van Strigen, and Musset used various combinations of post-CT ultrasound on venography where as PIOPED II is reporting all patients with a negative CT. Q5: • **Level B recommendations**. When a decision is made to perform venous ultrasound as the initial imaging modality, *a positive finding in a patient with symptoms consistent with PE can be considered evidence for diagnosis of VTE disease and may preclude the need for additional diagnostic imaging in the ED.

		*Examples of situations in which a venous ultrasound may be considered as initial imaging may include patients with obvious signs of DVT for whom venous ultrasound is readily available, patients with relative contraindications for CT scan (e.g., borderline renal insufficiency, CT contrast agent allergy), and pregnant patients. • Level C recommendations. 1) For patients with an intermediate pretest probability for PE and a negative CT angiogram result, for whom a clinical concern for PE still exists and CT venogram has not already been performed, consider lower extremity venous ultrasound as an additional test to exclude VTE disease (see question 4). 2) In patients with a high pretest probability for PE and a negative CT angiogram result, and CT venogram has not already been performed, perform additional testing to exclude VTE disease (see question 4). As one of these additional tests, consider lower extremity venous ultrasound to exclude VTE disease (see question 4). • The role of venous imaging for PE has been evaluated in three Class I (Anderson 2005, Musset 2002, Righini 2008) and three Class II (Stein 2006, Elias 2004, Le Gal 2006) and five Class III studies (Au 2001, Coche 2001, Begemann 2003, Johnson 2006, Loud 2001).

		• Venous US of bilateral LE does not evaluate abdominal and pelvic venous systems and often increases ED LOS, but should be considered as first line tests in some patients (renal disease, dye allergy, pregnancy). • One Class I study (Anderson 2005) demonstrated that obtaining US before CT or proceeding directly to CT were equally effective but 11 patients would need to have US for one to avoid CT. • Anderson et al performed US on everybody with low risk and elevated D-dimer with negative CTPA and identified an additional 3.1% with DVT. Others using similar protocols have replicated these findings: Musset 6.0%, LeGal 0.9%, Au 2.6%, Coche 2.3%. • Venous ultrasound and CT venous imaging are equally useful.
B.	**Have the criteria been field-tested for feasibility of use in diverse settings, include settings similar to yours?**	The ACEP guidelines have not been field tested. Whether each of the recommendations can be accurately and reliably applied across heterogeneous emergency settings (rural/urban, academic/community boarded EP/non boarded EP) and whether patients/society/malpractice attorney will accept and respect these evidence-based/consensus guidelines remains to be determined.

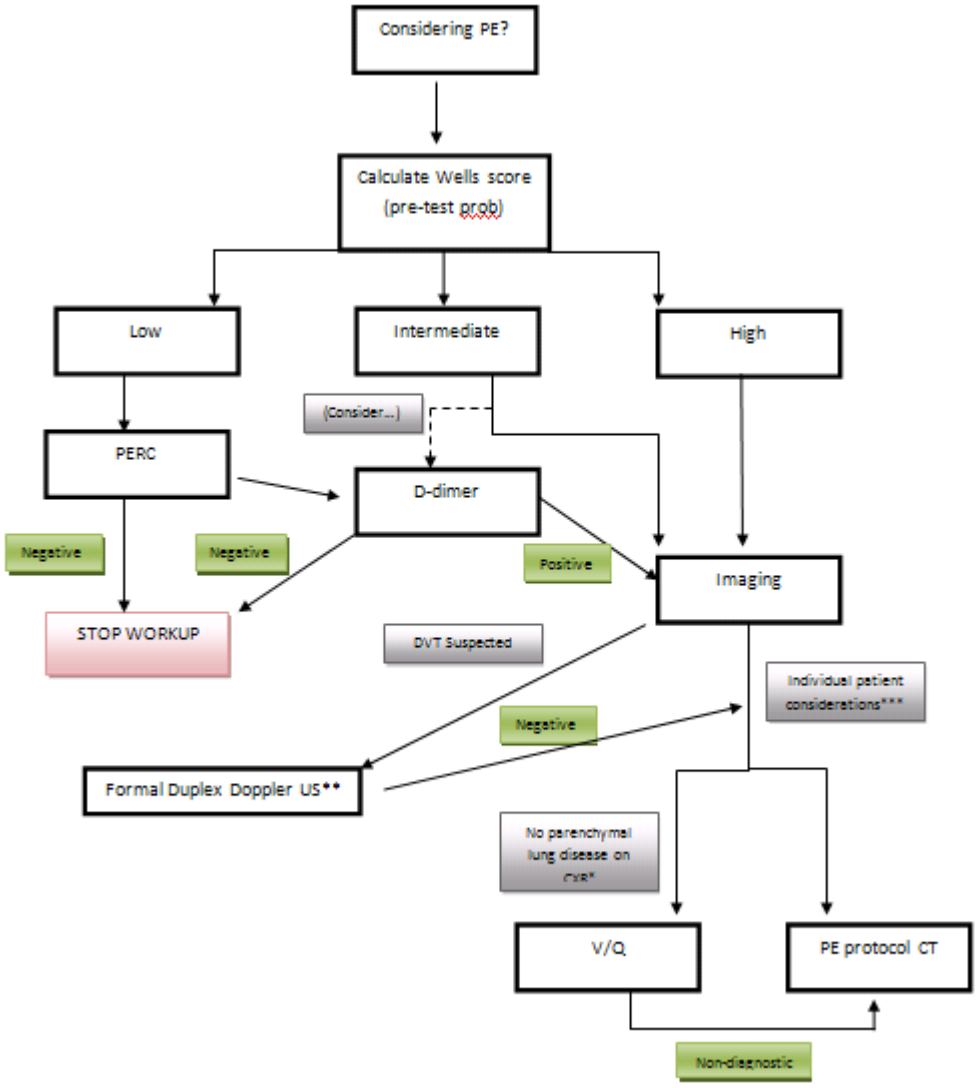
Limitations

1. **Exclusion of pregnant and pediatric patients.**
2. **Failure to use GRADE criteria and provide unequivocal actionable statements.**

3. **Failure to consider barriers to use of this clinical policy (malpractice, business of medicine, etc.)**
4. **Failure to incorporate patient perspective or patient communication tools.**
5. **No analysis of emergency physician performed lower extremity ultrasound.**

Bottom Line

Based upon these guidelines, local opinion, resource availability and acceptable level of risk aversion the following diagnostic algorithm was accepted by our working group that included four dozen emergency physicians from 6 hospitals, four radiologists, risk management and three malpractice defense attorneys.



* Chronic changes, small bilateral pleural effusions or mild pulmonary edema are unlikely to interfere with Nuclear Medicine interpretation of a V/Q study.

** Emergency physician performed ultrasound may be appropriate in credentialed centers, though no high quality evidence is available.

*** Examples young female (breast/thyroid cancer risk), advanced age, or renal diseases favor V/Q, whereas thoracic diagnostic considerations other than PE (aortic dissection) favor CT.