**Critical Review Form**

 **Clinical Prediction or Decision Rule**

[Eddy M, Robert-Ebadi H, Richardson L, et al. External validation of the YEARS diagnostic algorithm for suspected pulmonary embolism. J Thromb Haemost. 2020 Dec;18(12):3289-3295.](http://pmid.us/32869501)

PGY-3

**Objectives: "In the YEARS study, a simplified diagnostic strategy was proposed in which three clinical elements of the Wells score were used along with differential D-dimer cutoff values...before recommending the use of this simplified management strategy in a clinical setting, further external validation with use of an independent cohort was lacking. We aimed to fill this knowledge gap by conducting a post hoc analysis of previous PE diagnostic studies." (p. 3290)**

**Methods: This retrospective analysis included prospectively collected data from 3 outcomes studies conducted at multiple sites in Belgium, France, and Switzerland. The studies were conducted between October 2000 and June 2002, August 2002 and November 2003, and January 2005 and August 2006. In these studies, consecutive patients presenting to the ED were eligible for inclusion if there was a clinical suspicion for PE due to acute onset of new or worsening shortness of breath or chest pain without another obvious etiology. Clinical probability was assessed using the [Geneva score](https://www.wikidoc.org/index.php/Geneva_score_calculator%22%20%5Ct%20%22_blank%22%20%5Co%20%22https%3A//www.wikidoc.org/index.php/Geneva_score_calculator) in two of the studies and the [modified Geneva score](https://www.mdcalc.com/geneva-score-revised-pulmonary-embolism%22%20%5Ct%20%22_blank%22%20%5Co%20%22https%3A//www.mdcalc.com/geneva-score-revised-pulmonary-embolism) in the third study. In addition to data from these clinical scores, all three of the [YEARS criteria](https://www.mdcalc.com/years-algorithm-pulmonary-embolism-pe%22%20%5Ct%20%22_blank%22%20%5Co%20%22https%3A//www.mdcalc.com/years-algorithm-pulmonary-embolism-pe) (including "PE most likely diagnosis") and D-dimer levels were part of prospectively collected data in all three of the studies.**

**In patients with no YEARS criteria, the authors evaluated the diagnostic accuracy and potential impact of both a D-dimer < 1000 ng/mL and an [age-adjusted D-dimer](https://www.mdcalc.com/age-adjusted-d-dimer-venous-thromboembolism-vte%22%20%5Ct%20%22_blank%22%20%5Co%20%22https%3A//www.mdcalc.com/age-adjusted-d-dimer-venous-thromboembolism-vte) to rule out PE. For patients with one or more YEARS criteria, the authors evaluated the diagnostic accuracy and potential impact of a D-dimer < 500 ng/mL. For patients in whom PE would have been ruled out by the YEARS algorithm, the proportion of patients who were diagnosed with a PE at initial testing or within 3 months constituted the overall failure rate.**

**Out of 3414 patients enrolled in the three studies, 100 (2.9%) were excluded for missing data, leaving 3314 patients in the final analysis. Of these, 1783 (53.8%) had no YEARS criteria and 1531 (46.2%) were positive for one or more criteria. PE was diagnosed in 731 (22.1%) of the cohort.**

|  |  |
| --- | --- |
| **Guide** | **Comments** |
| **I.** | ***Is this a newly derived instrument* (Level IV)?** |  |
| A. | Was validation restricted to the retrospective use of statistical techniques on the original database? (If so, this is a Level IV rule & is not ready for clinical application). | No. Validation was performed retrospectively using data from 3 prospectively collected datasets separate from the original YEARS study. |
| **II.** | **Has the instrument been validated? (Level II or III). If so, consider the following:** |  |
| 1a | Were all important predictors included in the derivation process? | N/A |
| 1b | Were all important predictors present in significant proportion of the study population? | N/A |
| 1c | Does the rule make clinical sense? | Yes. The YEARS algorithm is a simplified clinical decision tool that uses 3 of the questions from the Wells' PE criteria. Using a [Bayesian statistical](https://www.analyticsvidhya.com/blog/2016/06/bayesian-statistics-beginners-simple-english/) approach, patients with no YEARS criteria present who hence have a low pre-test probability for PE were considered ruled out with a higher D-dimer threshold (1000 mg/dL), while patients with ay YEARS criteria were felt to be higher risk and hence were ruled out with a lower D-dimer threshold (500 mg/dL). |
| 2 | Did validation include prospective studies on several different populations from that used to derive it (II) or was it restricted to a single population (III)? | Somewhat. While this was a retrospective analysis, it utilized data that were collected prospectively. These data were obtained from multiple centers in Belgium, France, and Switzerland and included all aspects of the YEARS criteria and a D-dimer value, allowing proper evaluation of this algorithm. |
| 3 | *How well did the validation study meet the following criteria?* |  |
| 3a | Did the patients represent a wide spectrum of severity of disease? | Yes. The authors make no specific mention of excluding patients felt to be at high-risk for PE (based on clinical gestalt or Geneva risk scores). As a result, this study appears to include a wide spectrum of potential disease severity and should be at low risk of [spectrum bias](http://www.cmaj.ca/content/173/4/385.full.pdf). There were 1783 patients (53.8%) with 0 YEARS criteria, 1235 (37.3%) with 1 YEARS criteria, 288 (8.7%) with 2 YEARS criteria, and 8 (0.2%) with 3 YEARS criteria. |
| 3b  | Was there a blinded assessment of the gold standard? | Uncertain. The authors do not mention whether radiologists or follow-up outcome assessors reading CTPAs or venous compression ultrasonography were blinded to risk stratification scores (included components of the YEARS score), D-dimer results, or level of clinician suspicion for PE. |
| 3c | Was there an explicit and accurate interpretation of the predictor variables & the actual rule without knowledge of the outcome? | Likely yes. The authors mention that, "a sequential assessment of clinical probability, plasma D-dimer measurement, lower limb venous compression ultrasonography (CUS), and CTPA with some variations" was utilized in the evaluation for PE. This suggests that clinical probability assessments (including YEARS criteria) were made prior to D-dimer results, which were obtained prior to imaging. |
| 3d | Did the results of the assessment of the variables or of the rule influence the decision to perform the gold standard? | Yes. CTPA was performed in 1891 patients, representing 57% of the overall cohort. The remaining patients were followed up at 3 months, though details of the follow-up process in each of the studies were not provided ([partial verification bias](https://www.youtube.com/watch?v=qIFbU84IfmI)). |
| 4 | How powerful is the rule (in terms of sensitivity & specificity; likelihood ratios; proportions with alternative outcomes; or relative risks or absolute outcome rates)? | * Based on the YEARS algorithm, 1423 (42.9%) patients would have had PE excluded without CTPA. Of these, 17 were diagnosed with PE at initial testing, for a failure rate of 1.2% (95% CI 0.8-1.9). No additional PEs were diagnosed at follow-up.
	+ All 17 missed cases had 0 YEARS criteria and a D-dimer < 1000 ng/mL.
* Among 870 patients with no YEARS criteria and a negative age-adjusted D-dimer, there were no PEs diagnosed: 0.0%, 95% CI 0.0-0.4%.
 |
| **III.** | **Has an impact analysis demonstrated change in clinical behavior or patient outcomes as a result of using the instrument? (Level I). If so, consider the following:** |  |
| 1 | How well did the study guard against bias in terms of differences at the start (concealed randomization, adjustment in analysis) or as the study proceeded (blinding, co-intervention, loss to follow-up)? | While this was a retrospective analysis, it utilized prospectively collected data. There is no mention of blinding of radiologists reading the images and just over half of patients underwent the gold standard test for PE. While there is no specific mention of blinding practitioners to downstream test results, the sequential performance of testing (risk stratification followed by D-dimer testing followed by CTPA) suggests they were not aware of downstream results. |
| 2 | What was the impact on clinician behavior and patient-important outcomes? | N/A. The authors did not perform any analysis on whether use of the YEARS algorithm would impact CT scan ordering or outcomes. |

**Limitations:**

1. **Performed early to mid-2000s with less advanced CT scanners. There is a risk of missing smaller segmental PEs.**
2. **The gold standard test (CTPA) was only performed in 57% of the cohort (**[**partial verification bias**](https://www.youtube.com/watch?v=qIFbU84IfmI)**).**
3. **There is no mention of blinding of radiologists to clinical criteria or D-dimer results, raising the possibility of** [**incorporation bias**](https://first10em.com/ebm/incorporation-bias/)**.**
4. **The authors mention that patients were followed-up at 3 months in the included studies, but do not describe how this follow-up was obtained. They also do not mention what percent of patients were lost to follow-up and these missing data were handled.**

**Bottom Line:**

**In this retrospective analysis of prospectively collected data, the use of the standard YEARS algorithm (D-dimer < 1000 ng/mL in patients with 0 YEARS criteria and < 500 ng/mL in patients 1 or more YEARS criteria) resulted in a miss rate of 1.2% (95% CI 0.8-1.9). Use a modified YEARS algorithm with an age-adjusted D-dimer cutoff for patients with no YEARS criteria resulted in a miss rate of 0.0%, 95% CI 0.0-0.4%.**