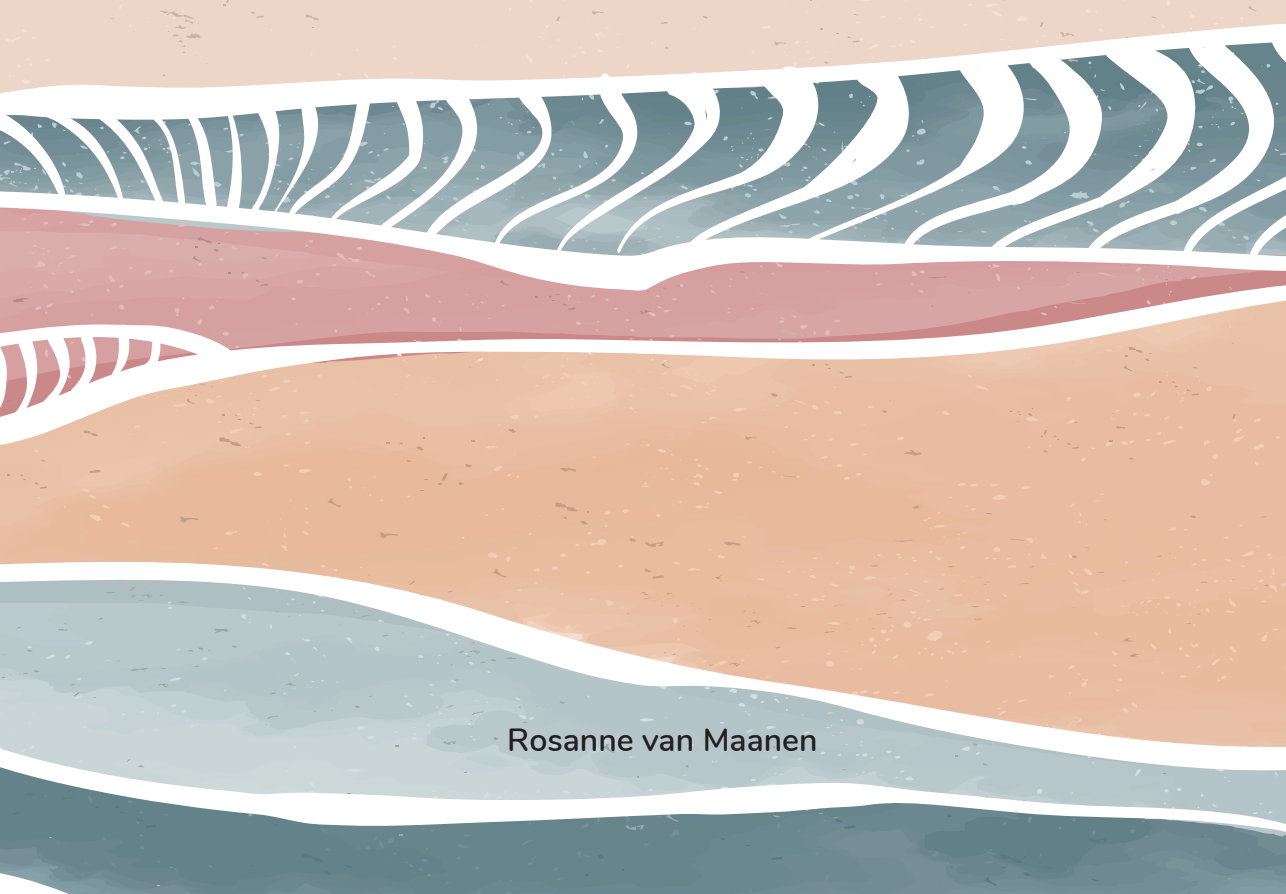
Diagnostic challenges of venous thromboembolism in primary care



Rosanne van Maanen

Diagnostic challenges of venous thromboembolism in primary care

Rosanne van Maanen

Diagnostic challenges of venous thromboembolism in primary care
Author: Rosanne van Maanen
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht
PhD thesis, Utrecht University, the Netherlands
ISBN: 978-90-393-7533-4

Printed by Ipskamp Printing, proefschriften.net
Cover design and layout by Rowen Aker, persoonlijkproefschrift.nl

The printing of this thesis was financially supported by the SBOH, employer of GP trainees. Part of the research reported in this thesis was funded by ZonMw.

Diagnostic challenges of venous thromboembolism in primary care

Diagnostische dilemma's bij veneuze trombo-embolie in de eerste lijn

(met een samenvatting in het Nederlands)

Proefschrift

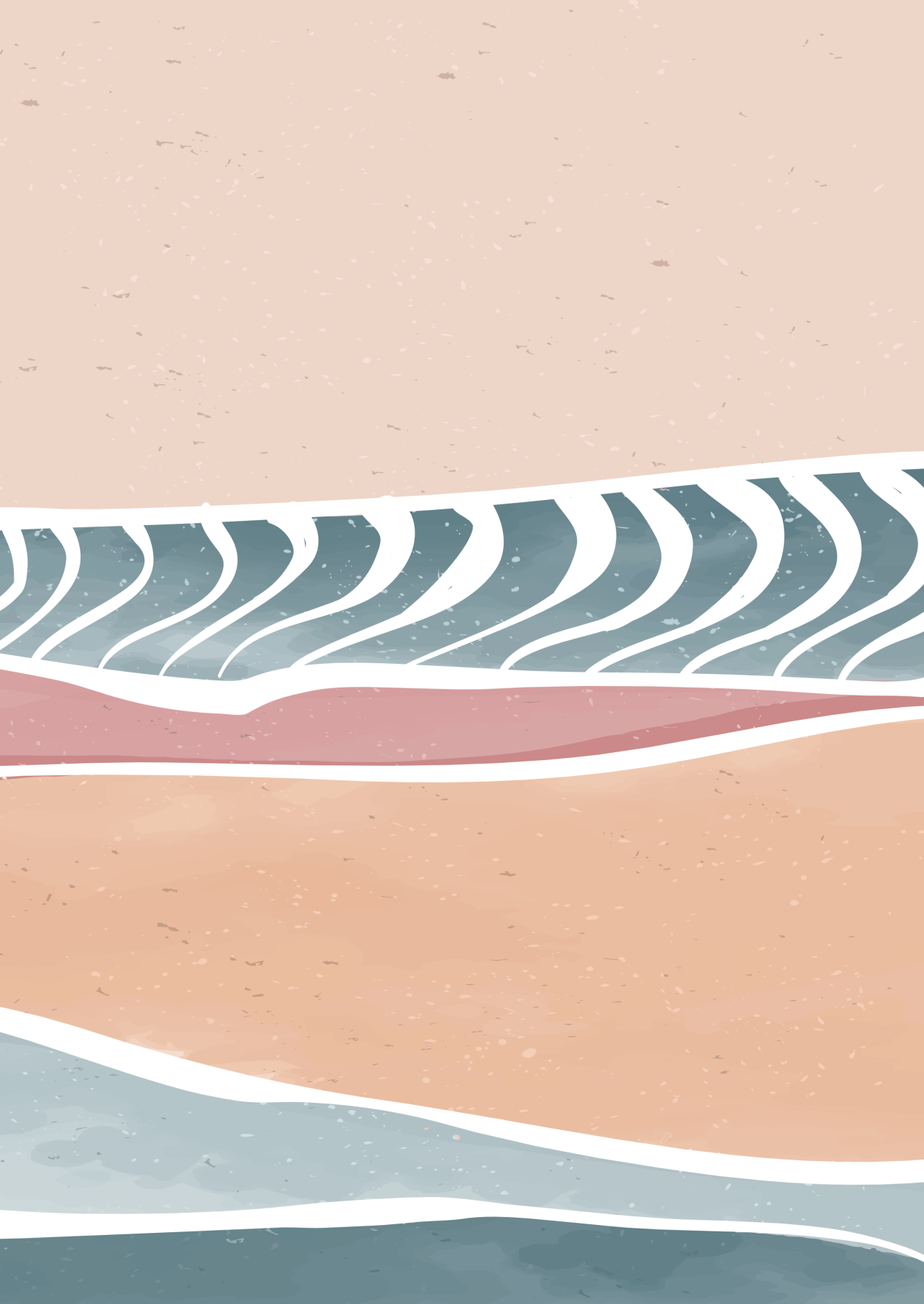
ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
donderdag 16 februari 2023 des middags te 4.15 uur

door

Rosanne van Maanen
geboren op 5 juli 1989
te Utrecht

Promotoren:	Prof. dr. F.H. Rutten
	Prof. dr. K.G.M. Moons
Copromotoren:	Dr. G.J. Geersing
	Dr. J.W. Blom
Beoordelingscommissie:	Prof. dr. R.E.G. Schutgens (voorzitter)
	Prof. dr. ir. Y.T. van der Schouw
	Dr. H.L. Koek
	Prof. dr. R.A.M.J. Damoiseaux
	Prof. dr. S. Middeldorp

CONTENTS



Introduction

VENOUS THROMBOEMBOLISM

Venous thromboembolism (VTE) is a disease caused by inappropriate fibrin formation which leads to the development of thrombi or emboli and causes vein obstruction. Most often such thrombi are located in the lower extremities, called deep vein thrombosis (DVT). However, a thrombus located in the legs may also dislodge and occlude (one of the branches of) the pulmonary arteries, called pulmonary embolism (PE). DVT and PE are both parts of the spectrum of VTE with an annual incidence rate of 1-2 per 1000 person-years.[1,2] VTE is associated with frequent hospitalizations and can have long-term consequences such as chronic thromboembolic pulmonary hypertension or a post-thrombotic syndrome of the leg.[3,4] Moreover, especially PE may be life-threatening. Historically, the mortality rate of PE was reported to be 25% if left untreated [5]. Once diagnosed, VTE is treated with anticoagulant medication such as vitamin K antagonists, low-molecular-weight heparin, or one of the newer direct oral anticoagulants, significantly reducing the mortality risk.

Nevertheless, the diagnosis of VTE could be challenging. Signs and symptoms of VTE are non-specific and mimic other conditions such as cellulitis in patients suspected of DVT, or respiratory tract infections or acute coronary syndrome in patients suspected of PE.[6] To illustrate this dilemma further, the classical diagnostic triad for PE of pleuritic chest pain, shortness of breath, and haemoptysis, occurs in less than 10% of all patients with PE.[7] Performing diagnostic imaging in all patients *suspected* of VTE would obviously decrease the risk of missing a VTE diagnosis. However, this seems not to be the most desirable solution given the low prevalence of VTE in suspected patients, the high healthcare costs of such approach, increased patient burden, and finally the increased risk of contrast nephropathy caused by such a diagnostic strategy.[8]

DIAGNOSTIC CHALLENGES

All of the above-mentioned aspects lead to a diagnostic challenge with on the one hand the fear of missing a VTE diagnosis, and on the other hand knowing that it is not desirable to refer every suspected patient for compression ultrasound or CT pulmonary angiography, the reference standards for DVT and PE respectively. Nevertheless, a rather low failure rate of 3% missed VTE diagnosis is internationally accepted as a safety margin in the diagnostic management of VTE.[9] Increasing this safety to miss fewer diagnoses of VTE will of course lead to more hospital referrals

and diagnostic testing. On the contrary, when the efficiency is increased by only referring patients with a high(er) suspicion of VTE more diagnoses will be missed. This trade-off in the diagnostic management of VTE is challenging and contributes to the diagnostic dilemma for physicians. General practitioners (GPs) have the difficult task to decide which patients are at high enough risk of VTE for referral to the hospital for further testing, and which patients are at such a low risk of VTE that they may be managed at home without referral nor anticoagulation, likely being treated for an alternative mimicking disorder. Currently, most patients referred because of the suspicion of VTE do not have a confirmative diagnosis.[10,11] On the other hand, VTE is listed among the diagnoses most frequently missed and is often delayed diagnosed.[12–14] An important cause for a missed or delayed diagnosis of VTE is that (initially) the physician does not consider VTE. The absence of the suspicion of VTE as a differential diagnosis for symptoms of a patient is still a major problem and is certainly related to mimicking disorders causing similar symptomatology.

So, both over-diagnosing (resulting in unnecessary diagnostic testing and referrals) and underdiagnosing (resulting in a missed or delayed VTE diagnosis) are common problems in VTE. Underdiagnosing is possibly the most important clinical problem from a patients' perspective. Patients will less likely blame their doctor for over-diagnosing although this obviously depends on the burden of the extra though unnecessary testing and their results, whereas a missed VTE diagnosis may lead to accusations. In line with the patients, also physicians tend to minimize the risk of missing a serious condition such as VTE]. However, overdiagnosing could lead to (iatrogenic) harm to the patient, such as the risk of contrast nephropathy and the exposition to ionising radiation which may increase the cancer risk later in life. [8,15] Hence, it can still be debated what is the optimal trade-off in this diagnostic challenge of VTE.

DIAGNOSTIC PREDICTION MODELS AND CLINICAL DECISION RULES

When physicians only rely on their gut feeling, or 'gestalt', it was shown that they overestimate the risk of VTE in suspected patients as compared to the observed risks.[16–18] This results in a high referral rate and thus a lower efficiency. To help both primary and secondary care physicians with this diagnostic challenge, several diagnostic prediction models and clinical decision rules (CDRs) have been developed and validated in the past decades.[10,11,19–21] A diagnostic prediction model

calculates the probability of the presence or absence of the disease for individual patients in whom that disease is suspected. Such models are often transformed to CDRs to use in clinical practice, in which round points are assigned to each diagnostic item of the model. This leads to a sum score and classifies the patients in, most often, a high or low risk of having the disease with corresponding recommendations to physicians for their subsequent management. For example, in CDRs for VTE the physician scores a list of patient characteristics and predictors for VTE with subsequent D-dimer testing, leading to a risk estimation of VTE presence in suspected patients. Patients with a low-risk estimate of VTE are left untreated, while patients with a high-risk estimate of VTE are referred for diagnostic imaging. Applying CDRs lead to fewer referrals as compared to using gestalt only, and is safe to exclude VTE in patients with a low estimated risk and a normal D-dimer level.[10,11] However, despite the use of CDRs, only 25-30% of referred patients from primary care are eventually diagnosed with respectively DVT or PE.[10,11] This percentage is even lower in elderly patients, due to an often falsely elevated D-dimer level.[22]

CASE

Mrs. Pecan is 74 years old and lives with her dog in a village in the vicinity of Utrecht. She visits her GP infrequently and is mostly seen by the practice nurse for regular control visits because of her longstanding hypertension. However, now she has made an appointment with her GP because she experiences shortness of breath when she walks her dog. This started rather abruptly three weeks ago, followed by another two episodes and resulting eventually in problems walking her dog. She also feels some chest discomfort, albeit Mrs. Pecan finds it difficult to describe this. This morning while making breakfast, she experienced some dizziness, although her GP knows this is not uncommon for her, especially during periods of stress. During the physical examination, the GP detects a slightly elevated heart rate of 92 beats per minute, an oxygen saturation of 97% on pulsoximetry, and she can provoke localised chest pain. Auscultation of heart of lungs is without abnormalities. She scrutinizes the patient file and uncovers that Mrs. Pecan had a DVT in the puerperium 40 years ago. She considers: "Could Mrs. Pecan now have pulmonary embolism? Or is it new slow-onset heart failure or anaemia? Or just non-specific chest pain causing anxiety?". The GP decides to use the Wells rule as recommended in the Dutch GP guidelines on VTE, and calculates a risk score of 1.5 points (for a history of VTE). According

to the Wells rule, the estimated probability of PE is low and therefore D-dimer blood testing is recommended. Mrs. Pecan goes home awaiting the D-dimer result that will follow later that day. Around 5 p.m. she is called; the D-dimer is 820 ng/mL, which is above the (age-independent) threshold of 500 ng/ml as mentioned in the GP guidelines. After some persuasion and discussions who will take care of her dog while she is in the hospital, the GP refers Mrs. Pecan for further imaging testing. At the emergency department, a CT pulmonary angiography is ordered without findings related to PE. Blood testing also comes back normal, as is resting ECG. Having spent the full evening in the hospital, Mrs. Pecan was sent home at 11 p.m. with a diagnosis of intercostal neuralgia, and her symptoms were released after knowing that she did not have a PE. The GP initially felt relieved that she had not missed a PE diagnosis, yet swiftly followed by feelings of having send Mrs. Pecan to the hospital for no good reason.

DEVELOPMENTS IN DIAGNOSING PULMONARY EMBOLISM

Ideally, Mrs. Pecan would not have been referred to the hospital. However, another patient with exact these signs, symptoms, and this D-dimer value might have had a PE. So, when is a referral in retrospect considered as “over-diagnosis” and when should we speak of “correctly referred for further testing, however, no PE”?

To objectify the suspicion of VTE and guide in their subsequent actions, physicians can use CDRs in the diagnostic management of VTE. New CDRs have been developed and existing CDRs have been modified to optimize the safety, efficiency, and usability of the strategies. For patients suspected of PE, the most commonly used CDRs are the revised (simplified) Geneva score and the Wells rule, of which the last one is prospectively validated in primary care and nowadays implemented in practice guidelines on VTE for GPs in the Netherlands.[10,11,23] In the Wells rule, the physician scores seven items with present or absent: clinical signs and symptoms of DVT, heart rate above 100 beats/minute, immobilization or surgery in the past four weeks, previous VTE, haemoptysis, active malignancy, and (a subjective item) whether the physician thinks PE is the most likely diagnosis. The main difference between the Wells rule and the Geneva score is that in the latter only eight *objective* items have to be scored. Each item of both CDRs corresponds with 1-3 points,

resulting in a risk estimation of PE. Patients with a high-risk estimate should directly be referred for diagnostic imaging, while patients with a low-risk estimate could first undergo a D-dimer test. If the D-dimer is below the threshold, PE is considered to be ruled out. The standard threshold is 500ng/mL, but an age-adjusted D-dimer cut-off (defined as age × 10 in patients 50 years or older) could also be applied and it is known to result in a higher efficiency.[24,25] Another recently developed and validated CDR in a hospital setting is the YEARS algorithm, which applies flexible D-dimer thresholds.[19,26,27] This strategy starts with the assessment of only three items, namely whether there are (i) clinical signs of DVT, (ii) haemoptysis, and (iii) is PE the most likely diagnosis according to the physician. In addition, D-dimer testing is performed in all suspected patients. In patients without any of the 'YEARS items', the D-dimer threshold to be used is 1000 ng/mL, while in patients with one or more 'YEARS items' the threshold remains the standard cut-off of 500 ng/mL. As compared to applying a fixed D-dimer threshold of 500 ng/mL, the YEARS algorithm increased the proportion of patients in whom diagnostic imaging was not required and importantly, the failure rate remained low (0.61%). Nowadays, this strategy is implemented as standard care at emergency departments in the Netherlands. It seems tempting to use this new algorithm in primary care as well given the higher efficiency. However, the YEARS strategy was developed and validated in the hospital setting and cannot directly be translated to primary care because of a lower pre-test probability, a different case mix of patients with relatively less severe cases of PE, and fewer experience with PE of the GP as compared to the hospital specialist. Thus, without first performing a prospective study evaluating the YEARS algorithm in primary care, it is currently not recommended that the GP would consider watchful waiting instead of referral of Mrs. Pecan.

THESIS OBJECTIVE

The diagnostic challenge of VTE translates into both over-diagnosing resulting in unnecessary diagnostic testing and referrals, as well as under-diagnosing resulting in a missed or delayed VTE diagnosis with possible life-threatening consequences. Therefore, the objective of this thesis is to contribute to improving the diagnostic management of patients suspected of VTE.

THESIS OUTLINE

The first two chapters describe the challenges of diagnosing VTE. Diagnostic delay in patients with PE is common, but not yet systematically analysed. **Chapter 1** describes a systematic review and meta-analysis of diagnostic delay in PE, to assess the prevalence, extent, and determinants of such delay. Although CDRs for VTE are widely used in primary care, the failure rate and safety of these CDRs, as well as determinants for, and consequences of incorrect application in real-life primary care are currently unknown. **Chapter 2** shows the real-life impact of the use of CDRs for both PE and DVT in primary care, based on a cross-sectional cohort study. In most CDRs for PE, one important item is subjective, namely whether PE is considered to be the most likely diagnosis. In **Chapter 3** we describe the diagnostic value of the subjective 'gestalt' item of CDRs for PE across patient subgroups, healthcare settings and countries, using an individual patient data meta-analysis. The following two chapters are about the PECAN study: a diagnostic management study to evaluate the YEARS algorithm for diagnosing PE in primary care. **Chapter 4** describes the rationale and design of the PECAN study. **Chapter 5** shows the results of the PECAN study and describes the safety and efficiency of the YEARS algorithm for patients suspected of PE in primary care. Finally, as became apparent during the conduct of the studies described in this thesis, performing prospective diagnostic VTE studies in primary care is challenging given the high workload of GPs, time-consuming study procedures, and low prevalence of VTE, leading to poor patient accrual. **Chapter 6** presents a new way of performing diagnostic studies embedded in daily routine primary care which bypasses these barriers. This newly described route of patient accrual into diagnostic studies was also used in this thesis (i.e. for the PECAN study, Chapter 4).

REFERENCES

1 Heit JA, Spencer FA, White RH. The epidemiology of venous thromboembolism. *J Thromb Thrombolysis* 2016;41:3–14. doi:10.1007/s11239-015-1311-6

2 Cohen AT, Agnelli G, Anderson FA, *et al*. Venous thromboembolism (VTE) in Europe - The number of VTE events and associated morbidity and mortality. *Thromb Haemost* 2007;98:756–64. doi:10.1160/TH07-03-0212

3 Klok FA, van der Hulle T, den Exter PL, *et al*. The post-PE syndrome: A new concept for chronic complications of pulmonary embolism. *Blood Rev* 2014;28:221–6. doi:10.1016/j.blre.2014.07.003

4 Huisman M V, Barco S, Cannegieter SC, *et al*. Pulmonary embolism. *Nat Rev Dis Prim* 2018;4:18028.https://doi.org/10.1038/nrdp.2018.28

5 Barrit DW, Jordan SC. Anticoagulant drugs in the treatment of pulmonary embolism. A controlled trial. *Lancet (London, England)* 1960;1:1309–12. doi:10.1016/s0140-6736(60)92299-6

6 Erkens PMG, Lucassen WAM, Geersing GJ, *et al*. Alternative diagnoses in patients in whom the GP considered the diagnosis of pulmonary embolism. *Fam Pract* 2014;31:670–7. doi:10.1093/fampra/cmu055

7 Meyer G, Roy PM, Gilberg S, *et al*. Pulmonary embolism. *BMJ* 2010;340:974–6. doi:10.1136/bmj.c1421

8 Mitchell AM, Jones AE, Tumlin JA, *et al*. Prospective study of the incidence of contrast-induced nephropathy among patients evaluated for pulmonary embolism by contrast-enhanced computed tomography. *Acad Emerg Med* 2012;19:618–25. doi:10.1111/j.1553-2712.2012.01374.x

9 Dronkers CEA, van der Hulle T, Le Gal G, *et al*. Towards a tailored diagnostic standard for future diagnostic studies in pulmonary embolism: communication from the SSC of the ISTH. *J Thromb Haemost* 2017;15:1040–3. doi:10.1111/jth.13654

10 Geersing GJ, Erkens PMG, Lucassen WAM, *et al*. Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in Primary care: Prospective cohort study. *BMJ* 2012;345:1–10. doi:10.1136/bmj.e6564

11 Buller HR, Cate-hoek AJ, Hoes AW, *et al*. Safely Ruling Out Deep Venous Thrombosis in Primary Care. *Ann Intern Med* 2009;150:229–36. doi:10.7326/0003-4819-150-4-200902170-00003

12 Hendriksen J, Lucassen W, Erkens P, *et al*. Klinische blik versus beslisregel bij longembolie. *Huisarts Wet* 2017;60:152–4. doi:10.1007/s12445-017-0098-4

13 Schiff GD, Hasan O, Kim S, *et al*. Diagnostic Error in Medicine: Analysis of 583 Physician-Reported Errors. *Arch Intern Med* 2009;169:1881–7. doi:10.1007/s10459-009-9187-x

14 Walen S, Damoiseaux RAMJ, Uil SM, *et al*. Diagnostic delay of pulmonary embolism in primary and secondary care: A retrospective cohort study. *Br J Gen Pract* 2016;66:e444–50. doi:10.3399/bjgp16X685201

15 Zondervan RL, Hahn PF, Sadow CA, *et al*. Body CT Scanning in Young Adults: Examination Indications, Patient Outcomes, and Risk of Radiation-induced Cancer. *Radiology* 2013;267:460–9. doi:10.1148/radiol.12121324

16 Barais M, Morio N, Cuzon Breton A, *et al*. 'I can't find anything wrong: It must be a pulmonary embolism': Diagnosing suspected pulmonary embolism in primary care, a qualitative study. *PLoS One* 2014;9:1–8. doi:10.1371/journal.pone.0098112

17 Kline JA, Stubblefield WB. Clinician Gestalt Estimate of Pretest Probability for Acute Coronary Syndrome and Pulmonary Embolism in Patients With Chest Pain and Dyspnea. *Ann Emerg Med* 2014;63:275–80. doi:10.1016/J.ANNEMERGMED.2013.08.023

18 Hendriksen JMT, Lucassen WAM, Erkens PMG, *et al*. Ruling Out Pulmonary Embolism in Primary Care: Comparison of the Diagnostic Performance of 'Gestalt' and the Wells Rule. *Ann Fam Med* 2016;14:227–34. doi:10.1370/afm.1930

19 van der Hulle T, Cheung WY, Kooij S, *et al*. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet* 2017;390:289–97. doi:10.1016/S0140-6736(17)30885-1

20 Le Gal G, Righini M, Roy P-M, *et al*. Prediction of Pulmonary Embolism in the Emergency Department: The Revised Geneva Score. *Ann Intern Med* 2006;144:165–71.

21 Wicki J, Perneger T V, Junod AF, *et al*. Assessing clinical probability of pulmonary embolism in the emergency ward: a simple score. *Arch Intern Med* 2001;161:92–7. doi:10.1001/archinte.161.1.92

22 Schouten HJ, Geersing GJ, Oudega R, *et al*. Accuracy of the Wells Clinical Prediction Rule for Pulmonary Embolism in Older Ambulatory Adults. *J Am Geriatr Soc* 2014;62:2136–41. doi:10.1111/jgs.13080

23 NHG-werkgroep Diepe veneuze trombose en longembolie. NHG-Standaard Diepe veneuze trombose en longembolie (tweede partiële herziening). *Huisarts Wet* 2017;60:460.

24 Righini M, Van Es J, Den Exter PL, *et al*. Age-adjusted D-dimer cutoff levels to rule out pulmonary embolism: The ADJUST-PE study. *JAMA - J Am Med Assoc* 2014;311:1117–24. doi:10.1001/jama.2014.2135

25 van Es N, Kraaijpoel N, Klok FA, *et al*. The original and simplified Wells rules and age-adjusted D-dimer testing to rule out pulmonary embolism: an individual patient data meta-analysis. *J Thromb Haemost* 2017;15. doi:10.1111/jth.13630

26 Van Es J, Beenen LFM, Douma RA, *et al*. A simple decision rule including D-dimer to reduce the need for computed tomography scanning in patients with suspected pulmonary embolism. *J Thromb Haemost* 2015;13:1428–35. doi:10.1111/jth.13011

27 van der Pol LM, Tromeur C, Bistervels IM, *et al*. Pregnancy-Adapted YEARS Algorithm for Diagnosis of Suspected Pulmonary Embolism. *N Engl J Med* 2019;380:1139–49. doi:10.1056/NEJMoa1813865

Authorship statement

The idea and set-up of the general introduction were mine; I conducted the literature search and wrote the general introduction. During the whole process, I asked for and implemented input and feedback from my supervisory team.

A systematic review and meta-analysis of diagnostic delay in pulmonary embolism

Rosanne van Maanen*
Emmy M. Trinks-Roerdink*
Frans H. Rutten
Geert-Jan Geersing

*both authors contributed equally

European Journal of General Practice 2022; 28:1,165-172

ABSTRACT

Introduction Diagnostic delay in patients with pulmonary embolism (PE) is common, yet the proportion of patients with PE that experienced delay and for how many days is less well described, nor are determinants for such delay. The purpose of this study was to assess the prevalence and extent of delay in diagnosing PE.

Methods A systematic literature search was performed to identify articles reporting on delay in diagnosing PE. The primary outcome was mean delay (in days) or a percentage of patients with diagnostic delay (defined as PE diagnosis > 7 days after symptom onset). The secondary outcome was determinants of delay. Random effect meta-analyses were applied to calculate a pooled estimate for mean delay and to explore heterogeneity in subgroups.

Results The literature search yielded a total of 10,933 studies, of which 24 were included in the final analysis. The pooled estimate of the mean diagnostic delay based on 12 studies was 6.3 days (95% prediction interval 2.5 to 15.8). The percentage of patients having > 7 days of delay varied between 18% and 38%. All studies assessing the determinants coughing (n=3), chronic lung disease (n=6) and heart failure (n=8) found a positive association with diagnostic delay. Similarly, all studies assessing recent surgery (n=7) and hypotension (n=6), as well as most studies assessing chest pain (n=8), found a negative association with diagnostic delay of PE.

Conclusion Patients may have symptoms for almost one week before PE is diagnosed and in about a quarter of patients the diagnostic delay is even longer.

KEY MESSAGES

- In this systematic review and meta-analysis with an extensive scope of all existing relevant studies on delay in diagnosing pulmonary embolism (PE), the mean diagnostic delay was almost one week and in a quarter of patients the delay was even longer.
- This emphasizes the importance of increasing awareness on PE and educating patients and physicians on how to recognize PE.

INTRODUCTION

Pulmonary embolism (PE) is the most serious condition within the spectrum of venous thromboembolic (VTE) conditions, given its associated high mortality rate, as well as its related morbidity and frequent hospitalization.[1, 2] Prompt and early recognition of PE is thus paramount. Clinical prediction rules – such as the Wells criteria, Geneva rule or YEARS algorithm – can assist physicians in diagnosing PE in suspected patients.[3-5] However, these rules are useful only when the physician actually has a clinical suspicion of PE. It can be extremely challenging to diagnose PE in a timely manner because symptoms of PE can differ widely in severity, and are often non-specific.[6,7] In some patients ultimately diagnosed with PE, the suspicion either never arose, or occurred only after multiple consultations. As an example, the so-called 'classical' PE-triad of chest pain, dyspnoea, and haemoptysis occurs in less than 10% of patients.[8]

Insight into the proportion of patients with PE that experienced delay and determinants that are associated with delay may help to increase awareness among physicians and patients, and thereby help to reduce diagnostic delay. This is especially meaningful for general practitioners (GPs) since patients with symptoms of PE often seek medical advice from their GP first. No previous study has systematically assessed the prevalence and extent of delay in diagnosing PE. Therefore, the purpose of this study was to systematically review the literature on studies reporting on delay in diagnosing PE. The primary objective was to assess the proportion of patients with PE that experienced diagnostic delay and the extent of this delay. A secondary objective was to identify determinants associated with a delayed diagnosis of PE.

METHODS

Search strategy

On the 31st of August 2021, we performed a literature search in Medline and Embase databases without date limit or language restrictions. The key terms in the search consisted of "pulmonary embolism" and synonyms, combined with "diagnostic delay", "time to diagnosis", "misdiagnosis" and alternative terms. See Appendix 1 for the full search syntax. Two reviewers (RvM and EMTR) screened the abstracts independently and selected original studies, describing any form of delay in the diagnostic management of PE. Subsequently, both reviewers independently selected full-text articles. In case of no consensus between these two researchers about the selection of a full-text article, a third researcher (GJG) was asked to screen the article in question, and a consensus was reached by discussion. We performed a cross-reference check for all included articles.

Definitions and study selection

For this study 'diagnostic delay' was defined as the time between the onset of symptoms (as reported by patients and described in the original publication) until confirmation of the diagnosis of PE. The primary objective was to quantify the presence of 'diagnostic delay', expressed as either a mean or median delay, or as a percentage of patients with diagnostic delay >7 days. The secondary objective was to quantify determinants for such delay. Studies conducted in general practices, emergency departments and hospital wards were considered for this review. We excluded systematic reviews, case reports, and articles describing the outcome in a highly specific population, e.g. paediatric populations, only post-operative patients, or pregnant women. Also, articles that only considered "logistic delay", for example, the time between admission and confirmation of the diagnosis with imaging, were excluded from our review since our primary aim was to obtain a pooled point estimate of the total diagnostic delay. Last, if there was no definition of delay mentioned or if we could not derive the definition of delay, the article was excluded.

Risk of bias and applicability assessment

No validated risk of bias tool was available for observational cross-sectional studies at the time we performed this review. Therefore, two reviewers independently assessed the risk of bias with modified criteria based on the QUADAS-2 tool.[9] We scored the risk of bias as high, low or unclear, within the following three domains: selection of study population (to assess generalizability and selection bias), validity of diagnostic testing (to assess information bias), and assessment of delay (to assess recall and

information bias). Moreover we scored the applicability of studies to primary care. Studies performed in general practice or studies in which patients were referred by the GP are considered very applicable to primary care. Studies in which a part of the included patients were referred by their GP are considered likely applicable to primary care. Studies in which patients were included from emergency departments are considered as possibly applicable. Studies in which patients were included from hospital wards are considered not applicable to primary care. If it was not clear from which setting patients were included, we considered the applicability to primary care as unclear. See Appendix 2 for the modified risk of bias and applicability tool that was used, including further clarification of these domains.

Data extraction and data analysis

The data were extracted using a standardized data extraction form. In addition to the primary objective to assess diagnostic delay of PE, we also collected data concerning our secondary objective, i.e. determinants for delay. Both determinants tested in univariable analysis and determinants tested in multivariable analysis were considered. We created an overview of clinically relevant determinants that were studied more than once and described whether a (significant) positive or negative association was found in the individual studies.

We performed a meta-analysis with studies that reported a mean delay since most studies reported a mean delay and not a median delay. Studies only reporting a median delay were excluded from this meta-analysis. We have sought contact with authors of studies only reporting a mean delay to obtain the median delay as well, but unfortunately we received no response. We log-transformed the data because we assumed that the mean delay of the individual studies was not normally distributed. Random effects meta-analysis was applied to calculate a pooled estimate with a 95% confidence interval and prediction interval for the mean diagnostic delay (defined in days). The prediction interval represents the range of estimates for the mean delay that can be found in future studies with a similar study design and thus can be considered as a measure for heterogeneity across studies.[10] Next, we performed meta-analyses to explain the heterogeneity in the following subgroups: studies that included only patients in the emergency department, studies with a low risk of bias due to misclassification, studies with the same definition of delay (time from onset of symptoms to diagnosis) and studies with prospective and retrospective data collection. Statistical analyses were performed in R version 3.4.1.

RESULTS

The literature search yielded a total of 10,933 studies. After screening on title and abstract we identified 50 articles, which we assessed for eligibility. Twenty-four articles met our in-and exclusion criteria.[11-34] For an overview of the literature search and article selection, see Figure 1. The 24 studies were published between 1998 and 2021. Data were collected retrospectively in 13 studies and collected prospectively in 11 studies. The included studies were performed in different settings, namely: primary care practices (n=1), emergency departments (n=7), hospital wards (n=9), or combinations (n=7). The characteristics of the included studies are presented in Table 1. The risk of bias regarding the domains of patient selection and valid diagnosis was assessed as 'low' in most studies. The risk of bias due to misclassification (assessment of delay) was assessed as 'high' in ten studies, mostly because of retrospective data collection. Two studies were assessed as very applicable to primary care, five studies as likely applicable, five studies as possibly applicable, six studies as not applicable and for six studies the applicability to primary care was unclear. See Appendix 3 for the risk of bias and applicability assessment.

Diagnostic delay

In total, 12 studies presented a mean delay with standard deviation. Figure 2 shows the forest plot of all 12 studies reporting a mean delay in diagnosing PE. The reported mean delay ranged from 2.5 to 11.9 days. The pooled point estimate of the mean delay was 6.3 days (95% CI 4.8 to 8.2) with a wide prediction interval (95% PI 2.5 to 15.8 days). The mean delay in studies performed in emergency departments was 7.7 days (95% PI 4.6 to 12.8). In our further pre-defined subgroup analyses (i.e. analyses of only studies with a low risk of bias, with a uniform definition of delay, or only using either prospective or retrospective data collection) the prediction intervals remained wide, indicating residual and unexplained heterogeneity. Sixteen studies reported a percentage of patients with diagnostic delay. Thirteen of these fifteen studies categorized delay beyond seven days. More than seven days of delay varied between 18% and 38%. The main outcomes are presented in Table 1.

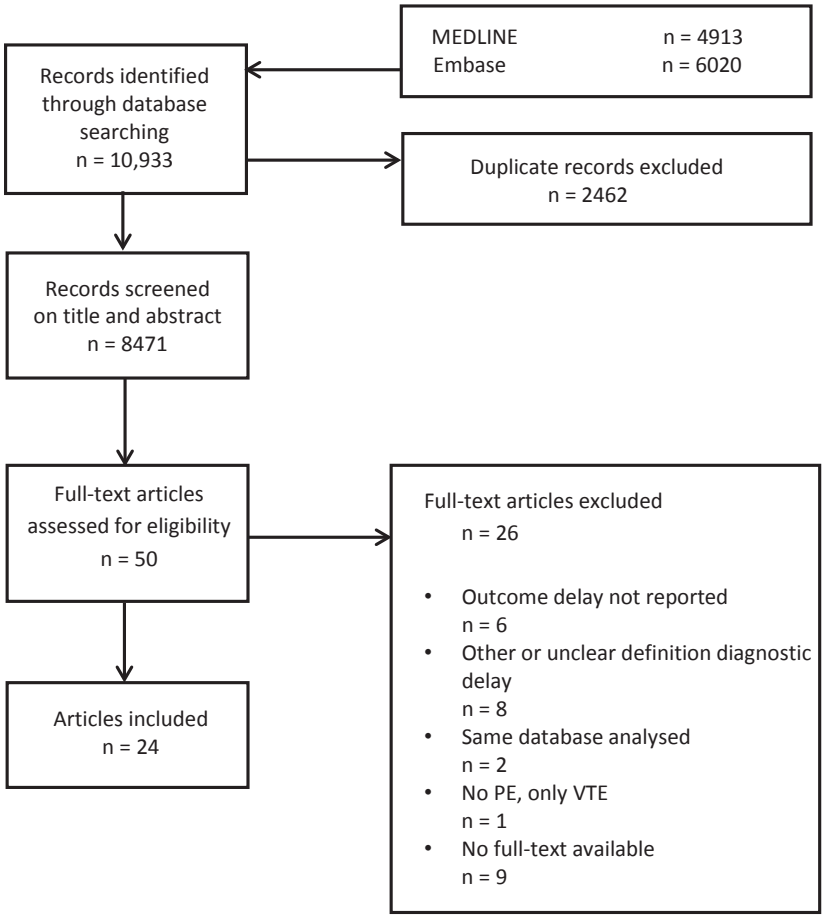


Figure 1. Flow-chart article selection

Table 1. Studies that assessed diagnostic delay in patients with pulmonary embolism

Study Name first author+ year of publication	***n*** Patients with PE	**Patient characteristics** Mean age +SD (years)	**Female** (%)	**Setting inclusion** Emergency department (ED), during hospital admission (HA), general practice (GP)	**Design of data collection**	**Definition delay** Moment of start counting delay – moment of stop counting delay	**Mean delay** Mean (days) + standard deviation	**Delay > 7 days** Percentage of patients in this category	**Other categories of delay**
Ageno 2008	542	59.8	57.4	HA+ED	Prospective	Symptoms - diagnosis			<5 days: 64% 5-10 days: 20% >10 days: 16%
Alonso-Martinez 2004	106	72±11	46.2	HA	Prospective	Symptoms – hospital admission	10±12		
Alonso-Martinez 2010	375	Median 75 IQR 15	49.6	HA	Prospective	Symptoms - diagnosis	Median 6 IQR 12		>6 days: 50% >14 days: 25% >21 days: 10%
Aranda 2021	150	61.2 ±18	51.3	HA	Prospective	Symptoms - diagnosis		26%	
Aydogdu 2013	53	65±17	54.7	ED	Prospective	Symptoms - diagnosis	6.8±7.7	38%	>1 day: 93%
Berghaus 2011	248	64.2±16.4	60.5	HA	Retrospective	Symptoms - diagnosis	2.5±1.9		
Bulbul 2009	178	60.4±16.8	53.9	HA+ED	Retrospective	Symptoms - diagnosis	9.3±11.6		
Bulbul 2011	156	64.1±15.9	62.2	ED	Prospective	Symptoms - diagnosis	7.93±10.05		
Chan 2020	302	†	§	HA	Retrospective	Symptoms - diagnosis		24%	
Den Exter 2013	849	52±18/56±18 *	58.2‡	HA+ED	Prospective	Symptoms - diagnosis		19%	
Elliott 2005	344	61.3±16.4	57.3	HA	Retrospective	Symptoms - diagnosis	4.8±20.2	17%	>25 days: 5%
Goyard 2018	514	Median 65 IQR28	51.2	HA	Prospective	Symptoms - diagnosis	Median 3 IQR 8	27%	>3 days: 47%
Hendriksen 2017	128	56±15/62±18 **	53.1	GP	Retrospective	First GP contact - diagnosis		26%	
Ilvan 2015	100	58.31±15.13	46	ED	Retrospective	Symptoms - diagnosis	11.9±22.6	28%	

Table 1. (Continued)

Study Name first author+ year of publication	**n** Patients with PE	**Patient characteristics** Mean age +SD (years)	**Female** (%)	**Setting inclusion** Emergency department (ED), during hospital admission (HA), general practice (GP)	**Design of data collection**	**Definition delay** Moment of start counting delay – moment of stop counting delay	**Mean delay** Mean (days) + standard deviation	**Delay > 7 days** Percentage of patients in this category	**Other categories of delay**
Jenab 2014	195	59.2±17.1	42.1	ED	Prospective	Symptoms – presentation hospital	5.6±7.9		<1 day: 31% <3 days: 57% >1 month: 1%
Jimenez 2007	397	69	55.4	ED	Prospective	Symptoms - diagnosis	Median 7	18%	>25 days: 6%
Kayhan 2012	189	57.95±16.36	55.0	HA	Retrospective	Symptoms - diagnosis		37%	
Menéndez 1998	102	64 range 21-88	54.9	HA+ED	Retrospective	Symptoms - diagnosis	Median 4 range 3-11		
Ozlem 2016	11	71.5±7.9	72.7	ED	Retrospective	Symptoms – ED admission	10.6 range 3-30		
Ozsu 2011	408	62.12±16.2	57.4	HA+ED	Retrospective	Symptoms - diagnosis	6.9±8.5	28%	
Pasha 2014	113	56±17	46.9	HA+ED	Retrospective	Symptoms - presentation hospital	5.7±9.2	18%	>1 month: 4%
Rahimi-Rad 2013	88	54.46±17.27 £	43.6 ₸	HA+ED	Prospective	Symptoms – treatment	3.05±6.42		
Walen 2016	261	60.6±16.9	47.9	ED	Retrospective	Symptoms - diagnosis	8.6±25.5	24%	>1 month: 6%
Zycinska 2013	53			HA	Retrospective	Symptoms - diagnosis	5		

** complaints<7 days: 52±18, complaints >7 days: 56±18 (suspected PE patients), ** 56±15 (no diagnostic delay) 62±18 (diagnostic delay), † 115 patients (38.1%) <65 years, 152 patients (50.3%) 65–84 years and 35 patients (11.6%) ≥85 years, § 77 female patients (67.0%) <65 years, 100 female patients (65.8%) 65-84 years and 25 female patients (71.4%) ≥ 85 years, £ Baseline characteristics of 353 patients with PE, PE/DVT or DVT. ‡ Suspected female PE patients. ₸ 43.6 % female patients in a group of 353 patients with DVT, PE and DVT+PE patients*

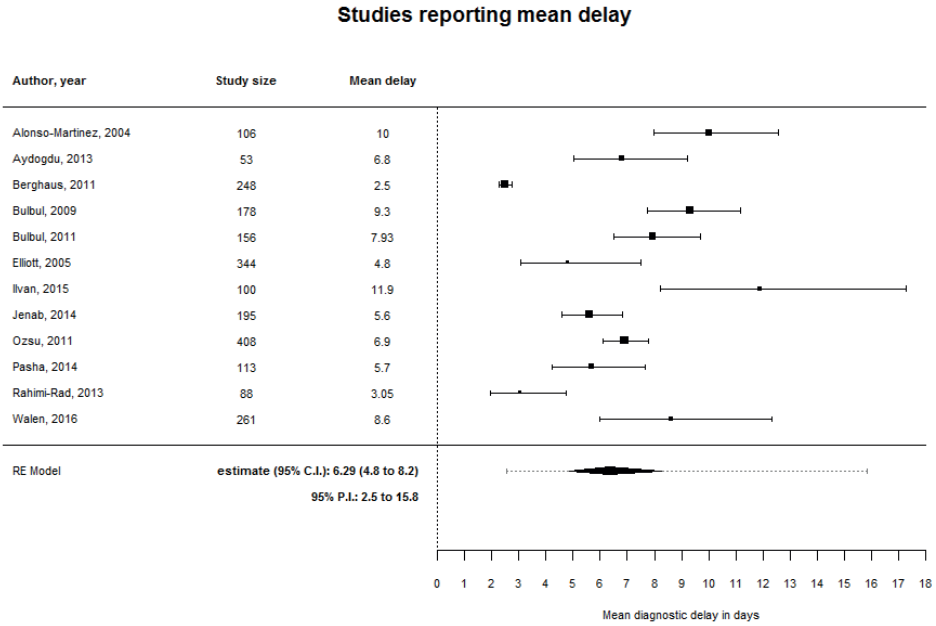


Figure 2. Meta-analysis of studies reporting mean delay

Determinants associated with delay

Fourteen studies assessed determinants potentially associated with diagnostic delay. Figure 3 summarizes these determinants and the positive or negative association with diagnostic delay that was found in the individual studies. See Appendix 4 for the full overview. For many of the explored determinants, findings were inconclusive and sometimes conflicting across different studies. Nevertheless, from a narrative synthesis we identified several determinants positively and negatively associated with diagnostic delay based on univariable and/or multivariable analyses, albeit not all statistically significant (see Figure 3 and Appendix 4). First, all of the three studies analysing coughing symptoms, all of the six studies analysing chronic lung disease and all of the eight studies analysing heart failure found a positive association of these determinants with diagnostic delay. Second, all of the seven studies analysing recent surgery and all of the six studies analysing hypotension found a negative association of these determinant with diagnostic delay. Last, seven out of nine studies analysing chest pain and six out of seven studies analysing tachycardia found a negative association with diagnostic delay.

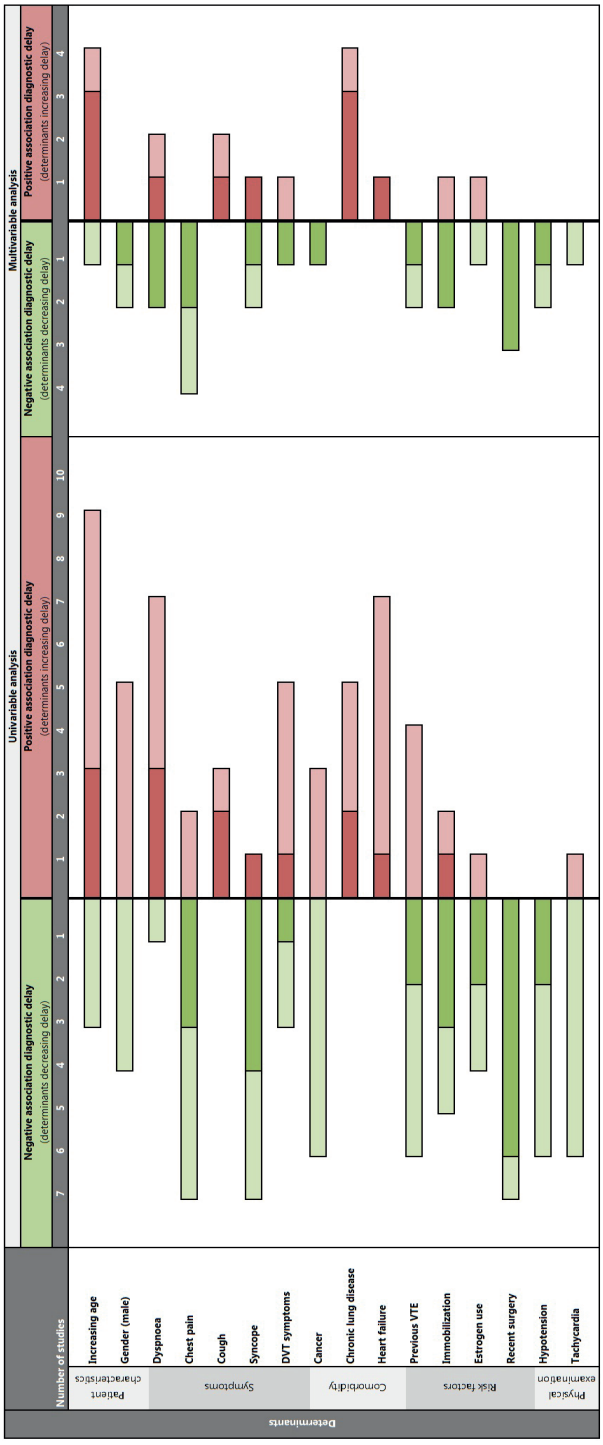


Figure 3. Determinants associated with diagnostic delay

DISCUSSION

This systematic review shows that delay in diagnosing PE is common with a pooled point estimate of a mean diagnostic delay of almost one week, albeit with a wide prediction interval indicating considerable heterogeneity between studies. About a quarter of patients had more than seven days of delay. Existing data suggest that patients with chronic cardiopulmonary co-morbidity or symptoms of coughing are at greater risk for delay. Yet, these observations were made only out of narrative synthesis from the included studies as formal meta-regression on determinants for delay was considered inappropriate due to differences in determinant definition and analytical techniques used.

Strengths and limitations

To the best of our knowledge, this is the first study to systematically describe the full scope and extent of delay in diagnosing PE. We performed a complete literature search without date or language restrictions and therefore we were able to provide an extensive scope of all existing relevant studies. Thereby, we were able to summarize the existing body of evidence on this important topic, hoping to provide some 'base evidence' for future studies embarking on this topic, allowing to compare findings from these new studies with the inferences found in our review. Furthermore, we pooled the mean delay using random-effect meta-analyses and explored heterogeneity. Some limitations however need to be taken into account. First, the mean diagnostic delay in days is probably not normally distributed, and therefore providing a pooled estimate of the median delay would have been preferable. However, most studies only reported a mean delay with a standard deviation and therefore we had to use the mean delay to calculate a pooled estimate. Second, in some of the included studies, delay was not clearly defined, necessitating us to use a proxy instead. The definition of delay also differed between the studies. Most of the included studies analysed the time from the onset of symptoms until the definitive confirmative diagnosis of PE. However, some studies reported the time from onset of symptoms until hospital admission, emergency department admission, or the start of treatment. For future diagnostic studies on PE, we would recommend reporting on diagnostic delay uniformly. We would suggest reporting the time between symptom onset (patient-reported) and confirmation of the PE diagnosis, and preferably also the time between symptom onset and the moment that the patient actually seeks medical attention, in order to distinguish between patients and physicians delay. Third, the methodology of the included studies differed, for example, in determining the duration of diagnostic delay. In some studies, patients were interviewed after a confirmative diagnosis, and

this could introduce recall bias, for which it is difficult (or even impossible) to adjust for. Finally, probably as an overall consequence of these above-described limitations, the between-study heterogeneity was considerable. An important cause of heterogeneity was that patients were included from different settings (hospital wards, emergency departments and primary care). In our review both studies categorized as very applicable to primary care, found a similar percentage of patients delay of more than seven days (24% and 26%). However, since both patients delay and physicians delay as well as the clinical implications of delay will be largely dependent on the setting of inclusion, this should be taken into account when interpreting our results.

Clinical implications

In our review, we focused primarily on the prevalence and extent of diagnostic delay of PE. Although not the purpose of our study, we could hypothesize on possible explanations for the diagnostic delay of approximately a week. First and foremost, it might be that PE-symptoms are often not timely recognized by the physician and/or the patient. As mentioned before, symptoms of PE are often non-specific and can vary in severity. As a consequence, it can be challenging to differentiate PE from alternative diagnoses, leading to a delay in the diagnostic process. This is supported by the fact that we found that delay seemed to occur more frequently in patients with comorbidities. Moreover, the decreasing prevalence of proven PE in suspected patients in diagnostic studies might suggest that physicians do think of PE quite often but still are struggling to correctly and timely identify PE in the right patients. [35,36] This emphasizes the importance of increasing awareness of PE and educating physicians and patients on how to recognize PE, e.g. during (albeit not exclusively) events like World Thrombosis Day.[37]

Second, another explanation for the diagnostic delay we found might be that PE is not an acute disease per se in *all* PE patients. With an average duration of symptoms of almost a week before diagnosis, PE might rather be a subacute condition with a slower onset of unfolding symptoms in a subset of patients, leading to a 'delayed', or perhaps better framed as a protracted and evolving, presentation. Should this be true, it could be that in the patients with such a milder clinical trajectory, the delay in diagnosis *might* be associated with less negative clinical consequences. In that respect, it could well be that delay happens more often in patients with sub-segmental PE compared to patients with lobular or more central PE's. Both possible explanations could also be true simultaneously. Yet given the fact that PE can also have serious (long-term) implications, more research is urgently needed to gain insight into the outcomes of patients with and without a delayed diagnosis.

We were unable to study the clinical consequences of diagnostic delay since only few of the included studies reported on clinical outcomes, such as recurrent PE or mortality. For instance, none of the included studies reported on clinical outcomes such as chronic thromboembolic pulmonary hypertension (CTEPH) or post-embolic syndrome. However, we know from the sparsely, existing literature on post-embolic syndromes that a delayed diagnosis might be a risk factor for developing CTEPH.[38]

Conclusion

Delay in diagnosing PE is common. Patients may have symptoms for almost one week before PE is diagnosed and in about a quarter of patients the diagnostic delay is even longer.

REFERENCES

1 Cohen AT, Agnelli G, Anderson FA, *et al.* Venous thromboembolism (VTE) in Europe - The number of VTE events and associated morbidity and mortality. *Thromb Haemost* 2007;98:756–64. doi:10.1160/TH07-03-0212

2 S Goldhaber, L Visani MDR. Acute pulmonary embolism clinical outcomes in the International Cooperative Pulmonary Embolism Registry (ICOPER). *Lancet* 1999;353:1386–9.

3 Wells PS, Anderson DR, Rodger M, *et al.* Derivation of a Simple Clinical Model to Categorize Patients Probability of Pulmonary Embolism: Increasing the Models Utility with the SimpliRED D-dimer. *Thromb Haemost* 2000;83:416–20.

4 Le Gal G, Righini M, Roy P-M, *et al.* Prediction of Pulmonary Embolism in the Emergency Department: The Revised Geneva Score. *Ann Intern Med* 2006;144:165–71.

5 van der Hulle T, Cheung WY, Kooij S, *et al.* Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet* 2017;390:289–97. doi:10.1016/S0140-6736(17)30885-1

6 Pineda LA, Hathwar VS, Grant BJB. Clinical Suspicion of Fatal Pulmonary Embolism. *Chest* 2001;120:791–5. doi:10.1378/CHEST.120.3.791

7 Barais M, Morio N, Cuzon Breton A, *et al.* 'I can't find anything wrong: It must be a pulmonary embolism': Diagnosing suspected pulmonary embolism in primary care, a qualitative study. *PLoS One* 2014;9:e98112. doi:10.1371/journal.pone.0098112

8 Meyer G, Roy P-M, Gilberg S, *et al.* Pulmonary embolism. *BMJ* 2010;340:c1421. doi:10.1136/bmj.c1421

9 Whiting PF, Rutjes AWS, Westwood ME, *et al.* QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011;155:529–36. doi:10.7326/0003-4819-155-8-201110180-00009

10 IntHout J, Ioannidis JPA, Rovers MM, *et al.* Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 2016;6:e010247. doi:10.1136/bmjopen-2015-010247

11 Ageno W, Agnelli G, Imberti D, *et al.* Factors associated with the timing of diagnosis of venous thromboembolism: results from the MASTER registry. *Thromb Res* 2008;121:751–6. doi:10.1016/j.thromres.2007.08.009

12 Alonso Martínez JL, Echegaray Agara M, Urbieta Echezarreta MA, *et al.* Embolismo pulmonar agudo. Un registro de 10 años: de mayo de 1992 a mayo de 2002. *Rev Clin Esp* 2004;204:521–7. doi:10.1157/13066175

13 Ilvan A, Celikdem M. The causes of misdiagnosis and pulmonary embolism. 2015;63:13–21.

14 Jenab Y, Alemzadeh-Ansari MJ, Fehri SA, *et al.* Effect of delay in hospital presentation on clinical and imaging findings in acute pulmonary thromboembolism. *J Emerg Med* 2014;46:465–71. doi:10.1016/j.jemermed.2013.09.014

15 Jiménez Castro D, Sueiro A, Díaz G, *et al.* Prognostic significance of delays in diagnosis of pulmonary embolism. *Thromb Res* 2007;121:153–8. doi:10.1016/j.thromres.2007.03.028

16 Kayhan S, Ünsal M, İnce Ö, *et al.* Delays in Diagnosis of Acute Pulmonary Thromboembolism: Clinical Outcomes and Risk Factors. *Eur J Gen Med* 2012;9:124–9. doi:10.29333/ejgm/82476

17 Menéndez R, Nauffal D, Cremades MJ. Prognostic factors in restoration of pulmonary flow after submassive pulmonary embolism: A multiple regression analysis. *Eur Respir J* 1998;11:560–4. doi:10.1183/09031936.98.11030560

18 Ozlem B, Gokhan E, Baris G, *et al*. THE DIAGNOSIS OF PULMONARY EMBOLISM IN PATIENTS WITH NORMAL D-DIMER LEVELS. *Acta Medica Mediterr* 2016;32:171–7. doi:10.19193/0393-6384

19 Ozsu S, Oztuna F, Bulbul Y, *et al*. The role of risk factors in delayed diagnosis of pulmonary embolism. *Am J Emerg Med* 2011;29:26–32. doi:10.1016/j.ajem.2009.07.005

20 Pasha SM, Klok FA, van der Bijl N, *et al*. Right ventricular function and thrombus load in patients with pulmonary embolism and diagnostic delay. *J Thromb Haemost* 2014;12:172–6. doi:10.1111/jth.12465

21 RAHIMI-RAD MH, RAHIMI-RAD S, ZARRIN S. Delays in diagnosis and treatment of venous thromboembolism in a developing country setting. *Tuberk Toraks* 2013;61:96–102. doi:10.5578/tt.5348

22 Walen S, Damoiseaux RAMJ, Uil SM, *et al*. Diagnostic delay of pulmonary embolism in primary and secondary care: A retrospective cohort study. *Br J Gen Pract* 2016;66:e444–50. doi:10.3399/bjgp16X685201

23 Alonso-Martínez JL, Sánchez FJA, Echezarreta MAU. Delay and misdiagnosis in sub-massive and non-massive acute pulmonary embolism. *Eur J Intern Med* 2010;21:278–82. doi:10.1016/j.ejim.2010.04.005

24 Zycińska K, Wiktorowicz M, Tomasik D, *et al*. Clinical presentation of pulmonary embolism in general practice. *Fam Med Prim Care Rev* 2013;15:430–3.

25 Goyard C, Côté B, Looten V, *et al*. Determinants and prognostic implication of diagnostic delay in patients with a first episode of pulmonary embolism. *Thromb Res* 2018;171:190–8. doi:10.1016/j.thromres.2018.08.015

26 Chan TF, Ngian VJJ, Hsu K, *et al*. Pulmonary embolism: clinical presentation and diagnosis in the oldest old. *Intern Med J* 2020;50:627–31. doi:10.1111/imj.14824

27 Aranda C, Peralta L, Gagliardi L, *et al*. A significant decrease in D-dimer concentration within one month of anticoagulation therapy as a predictor of both complete recanalization and risk of recurrence after initial pulmonary embolism. *Thromb Res* 2021;202:31–5. doi:10.1016/j.thromres.2021.02.033

28 Aydoğdu M, Doğan NÖ, Sinanoğlu NT, *et al*. Delay in diagnosis of pulmonary thromboembolism in emergency department: Is it still a problem? *Clin Appl Thromb* 2013;19:402–9. doi:10.1177/1076029612440164

29 Berghaus TM, Von Scheidt W, Schwaiblmair M. Time between first symptoms and diagnosis in patients with acute pulmonary embolism: Are patients with recurrent episodes diagnosed earlier? *Clin Res Cardiol* 2010;100:117–9. doi:10.1007/s00392-010-0217-8

30 Bulbul Y, Ozsu S, Kosucu P, *et al*. Time delay between onset of symptoms and diagnosis in pulmonary thromboembolism. *Respiration* 2009;78:36–41. doi:10.1159/000167409

31 Bulbul Y, Ayik S, Oztuna F, *et al*. The relationship between socio-demographic characteristics of patients and diagnostic delay in acute pulmonary thromboembolism. *Ups J Med Sci* 2011;116:72–6. doi:10.3109/03009734.2010.530701

32 den Exter PL, van Es J, Erkens PMG, *et al*. Impact of delay in clinical presentation on the diagnostic management and prognosis of patients with suspected pulmonary embolism. *Am J Respir Crit Care Med* 2013;187:1369–73. doi:10.1164/rccm.201212-2219OC

33 Elliott CG, Goldhaber SZ, Jensen RL. Delays in diagnosis of deep vein thrombosis and pulmonary embolism. *Chest* 2005;128:3372–6. doi:10.1378/chest.128.5.3372

34 Hendriksen JMT, Koster-van Ree M, Morgenstern MJ, *et al*. Clinical characteristics associated with diagnostic delay of pulmonary embolism in primary care: a retrospective observational study. *BMJ Open* 2017;7:e012789. doi:10.1136/bmjopen-2016-012789

35 Klok FA, van der Hulle T, den Exter PL, *et al*. The post-PE syndrome: A new concept for chronic complications of pulmonary embolism. *Blood Rev* 2014;28:221–6. doi:10.1016/j.blre.2014.07.003

36 Kearon C, de Wit K, Parpia S, *et al*. Diagnosis of Pulmonary Embolism with d-Dimer Adjusted to Clinical Probability. *N Engl J Med* 2019;381:2125–34. doi:10.1056/NEJMoa1909159

37 Geersing GJ, Erkens PMG, Lucassen WAM, *et al*. Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in primary care: prospective cohort study. *BMJ* 2012;345:e6564. doi:10.1136/bmj.e6564

38 Rosendaal FR, Raskob GE. On world thrombosis day. *Lancet* 2014;384:1653–4. doi:10.1016/S0140-6736(14)61652-4

Authorship statement

Emmy M. Trinks-Roerdink and I are joint first authors and equally contributed to defining the research question, performing the systematic literature search, conducting the data analysis, writing the manuscript, and implementing the contribution and feedback of the co-authors and external reviewers up to the final publication.

SUPPLEMENTAL FILES

Supplement 1. Search review diagnostic delay pulmonary embolism
Supplement 2. Risk of bias & applicability (based on QUADAS-2 tool)
Supplement 3. Risk of bias & applicability
Supplement 4. Factors associated with diagnostic delay

SUPPLEMENT 1. Search review diagnostic delay pulmonary embolism

Pubmed
(((((((Pulmonary Embolism*[tiab]) OR (Pulmonary Infarct* [tiab]) OR (Pulmonary Embolism[Mesh]) OR (pulmonary thromboembolism* [tiab])))))) OR "Venous Thromboembolism"[Mesh])) OR ((lung embol*[Title/Abstract] OR lung infarct*[Title/Abstract]))

AND

((((diagnos*[Title/Abstract]) AND (late[Title/Abstract] OR delay*[Title/Abstract] OR missed[Title/Abstract] OR missing[Title/Abstract] OR error*[Title/Abstract] OR inappropriate*[Title/Abstract] OR time[Title/Abstract] OR timing[Title/Abstract] OR timely[Title/Abstract]))) OR (("Delayed Diagnosis"[Mesh]) OR "Diagnostic Errors"[Mesh])) OR (misdiagnos*[tiab] OR undiagnos*[tiab])

Embase
'pulmonary embolism*':ti,ab,kw OR 'pulmonary infarct*':ti,ab,kw OR 'lung embolism*':ti,ab,kw OR 'pulmonary thromboembolism*':ti,ab,kw OR 'venous thromboembolism*':ti,ab,kw OR 'lung infarction*':ti,ab,kw OR 'lung embolism'/exp

AND

('diagnos*':ti,ab,kw AND ('late':ti,ab,kw OR 'delay':ti,ab,kw OR 'missing':ti,ab,kw OR 'missed':ti,ab,kw OR 'error':ti,ab,kw OR 'inappropiate*':ti,ab,kw OR 'time':ti,ab,kw OR 'timely':ti,ab,kw OR 'timing':ti,ab,kw)) OR ('delayed diagnos*':ti,ab,kw OR 'delayed diagnosis'/exp OR 'diagnostic error*':ti,ab,kw OR 'diagnostic error'/exp OR 'misdiagnos*':ti,ab,kw OR 'undiagnos*':ti,ab,kw)

SUPPLEMENT 2. Risk of bias & applicability (based on QUADAS-2 tool)

Risk of bias assessment

Domain	Patient selection	Valid diagnosis	Assessment of delay
Description	Describe methods of patient selection: Describe included patients?	Describe the test used for final diagnosis.	Describe the method of assessment of delay.
Signalling questions	Was a consecutive or random sample of patients enrolled?	Was a CT-scan, V/Q-scan, perfusion scan or ultrasound proven DVT with PE symptoms performed?	- What was the study type? - Risk of recall bias? - Was the delay reported by patients/ doctor/ both? - Was the health record of the patient used?
	Did the study avoid inappropriate exclusions?*		
Risk of bias: High/low/unclear	Could the selection of patients have introduced bias?	Could the test used for diagnosis have introduced bias?	Could the assessment of delay have introduced bias?

Applicability to primary care

Signalling question	Are the included patients in the original studies comparable to patients in primary care?			
Very applicable: Patients included in primary care OR Patients referred by a general practitioner	Likely applicable: Patients partly included in primary care or outpatient clinic	Possibly applicable: Patients included in emergency departments	Not applicable: Patients included in hospital wards during admission	Unclear: Not clearly explained where and how patients are included

* >5% exclusion due to lost to follow-up was classified as 'high' risk of bias

SUPPLEMENT 3. Risk of bias & applicability

Study	Risk of bias			Applicability to primary care
	Patient selection	Valid diagnosis	Assessment of delay	
Ageno 2008	Low	Low	Low	Likely
Alonso-Martinez 2004	Low	Low	Unclear	Not
Alonso-Martinez 2010	Low	Low	Unclear	Not
Aranda 2021	High	Low	Unclear	Not
Aydogdu 2013	Low	Low	Unclear	Possibly
Berghaus 2011	Low	Low	High	Not
Bulbul 2009	Low	Low	High	Unclear
Bulbul 2011	High	Low	Low	Unclear
Chan 2020	Low	Low	High	Not
Den Exter 2013	Low	Low	Unclear	Likely
Elliott 2005	Unclear	Low	Unclear	Unclear
Goyard 2018	Low	Low	Low	Unclear
Hendriksen 2017	Low	Low	High	Very
Ilvan 2015	Low	Low	High	Possibly
Jenab 2014	Low	Low	Low	Possibly
Jimenez 2007	Low	Low	Unclear	Possibly
Kayhan 2012	Low	Low	High	Not
Menéndez 1998	High	Low	High	Unclear
Ozlem 2016	High	Low	High	Possibly
Ozsu 2011	High	Low	High	Likely
Pasha 2014	Low	Low	Low	Likely
Rahimi-Rad 2013	Low	Low	Unclear	Unclear
Walen 2016	High	Low	High	Very
Zycinska 2013	Low	Unclear	Unclear	Likely

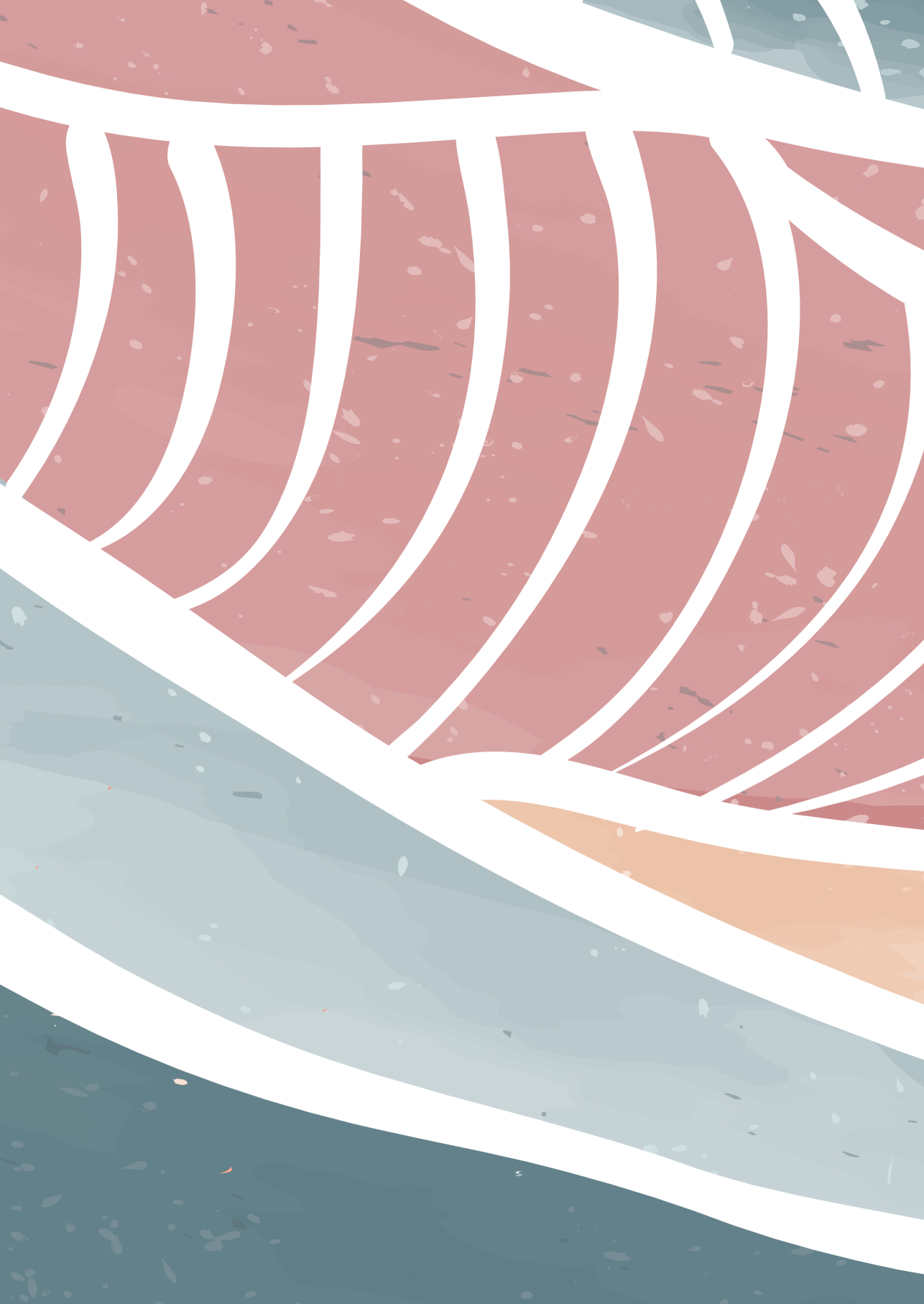
SUPPLEMENT 4. Factors associated with diagnostic delay

		Patient characteristics		Symptoms					Comorbidity				Risk factors					Physical examination	
		Gender (male)	Age	Dyspnoea	Chest pain	Cough	DVT (symptoms)	Syncope	Cancer	Chronic lung disease	Heart Failure	Prior Pulmonary Infection	Previous VTE	Smoking	Immobilization	Recent surgery	Oestrogen use	Hypotension	Tachycardia
Univariable analysis																			
Ageno	2008	= (-)	= (-)	= (+)	= (-)			= (-)	= (-)				= (-)						
Alonso-Mar.	2010		+		= (-)			-					-						
Bulbul	2009	= (+)	= (-)	= (-)	= (-)	+	= (-)											-	= (-)
Bulbul	2011		= (+)	= (+)	= (+)		= (+)	-						= (+)		-		-	
Den Exter	2013	= (+)	+						= (-)	+	+		= (+)		-	-	-		
Goyard	2018	= (-)	= (+)	+	= (-)		= (+)	-	= (-)	= (+)	= (+)		-		-	-	= (+)	= (-)	= (-)
Hendriksen	2017	= (-)	+	= (+)	-	+	-	= (-)	= (+)	+	= (+)	+	= (+)	= (-)	= (-)	= (-)	= (-)		= (-)
Jenab	2014	= (+)	= (+)	+	-		+	= (-)	= (+)		= (+)		= (-)	= (+)	-	-	-	= (-)	= (-)
Jimenez	2007	= (+)	(-)	= (+)	= (+)		= (+)	+	= (+)	= (+)	= (+)		= (+)		= (-)	-		= (-)	= (-)
Ozsu	2011		= (+)				= (+)	-	= (-)		= (+)		= (-)		+	-	= (+)	= (-)	= (+)
Pasha	2014	= (-)	= (+)						= (-)	= (+)	= (+)		= (-)		= (+)				
Walen	2016	= (+)	= (+)	+	-	= (+)	= (-)		= (-)				= (+)						= (-)

SUPPLEMENT 4. (Continued)

		Patient characteristics		Symptoms					Comorbidity				Risk factors					Physical examination	
		Gender (male)	Age	Dyspnoea	Chest pain	Cough	DVT (symptoms)	Syncope	Cancer	Chronic lung disease	Heart Failure	Prior Pulmonary Infection	Previous VTE	Smoking	Immobilization	Recent surgery	Oestrogen use	Hypotension	Tachycardia
Multivariable analysis																			
Bulbul	2009				= (-)	+												-	
Chan	2020		+	-				+		+	+								
Den Exter	2013		+							+					-	-			
Goyard	2018	= (-)	= (+)	+	= (-)		= (+)	-		= (+)			-		-	-	= (-)	= (-)	= (-)
Hendriksen	2017	-	+	-	-					+		+							
Kayhan	2012		= (-)											-					
Ozsu	2011							= (-)	-				= (-)		= (+)	-	= (+)		
Walen	2016			= (+)	-	= (+)	-												

+ Positive association with delay; - Negative association with delay; = No statistically significant association with delay



Real-life impact of clinical prediction rules for venous thromboembolism in primary care: a cross-sectional cohort study

Rosanne van Maanen
Annelieke E.C. Kingma
Ruud Oudega
Frans H. Rutten
Karel G.M. Moons
Geert-Jan Geersing

BMJ Open 2020; 10:e039913

ABSTRACT

Objective Clinical prediction rules (CPRs) followed by D-dimer testing were shown to safely rule-out venous thromboembolism (VTE) in about half of all suspected patients in controlled and experienced study settings. Yet, its real-life impact in primary care is unknown. The aim of this study was to determine the real-life impact of CPRs for suspected VTE in primary care.

Design Cross-sectional cohort study.

Setting Primary care in the Netherlands.

Participants Patients with suspected deep venous thrombosis (DVT, n=993) and suspected pulmonary embolism (PE, n=484).

Interventions General practitioners received an educational instruction on how to use CPRs in suspected VTE. We did not rectify incorrect application of the CPR in order to mimic daily clinical care.

Main outcome measures Primary outcomes were the diagnostic failure rate, defined as the three-month incidence of VTE in the non-referred group, and the efficiency, defined as the proportion of non-referred patients in the total study population. Secondary outcomes were determinants for, and consequences of *incorrect* application of the CPRs.

Results In 267 of the included 1,477 patients VTE was confirmed. When CPRs were correctly applied, the failure rate was 1.51% (95% CI 0.77 to 2.86) and the efficiency 58.1% (95% CI 55.2 to 61.0). However, the CPRs were incorrectly applied in 339 patients, which resulted in an increased failure rate of 3.31% (95% CI 1.07 to 8.76) and a decreased efficiency of 35.7% (95% CI 30.6 to 41.1). The presence of concurrent heart failure increased the likelihood of incorrect application (adjusted OR 3.26; 1.47, 7.21).

Conclusions Correct application of CPRs for VTE in primary care is associated with an acceptable low failure rate at a high efficiency. Importantly, in nearly a quarter of patients the CPRs were incorrectly applied which resulted in a higher failure rate and a considerably lower efficiency.

STRENGTHS AND LIMITATIONS

- A large population of 1,477 patients was included, resulting in an accurate estimate of the failure rate and efficiency of clinical prediction rules for venous thromboembolism.
- The reference standard between referred and non-referred patients suspected of venous thromboembolism differed, which might result in differential verification bias.
- The point-of-care test for D-dimer used during the study had too many false-negative results, leading to an increased failure rate for those patients.
- A sample size calculation was not performed a-priori, however a number of 268 outcome VTE events allowed robust statistical analyses.

INTRODUCTION

When patients visit their general practitioner (GP) with a red and swollen calf, deep venous thrombosis (DVT) may be considered. In case of shortness of breath or thoracic pain, pulmonary embolism (PE) could be the cause. Together, both conditions are part of the spectrum of venous thromboembolic (VTE) diseases, which has an incidence of 1-2 per 1000 person-years.[1,2] It is associated with a considerable global impact on morbidity and mortality, with an estimated number of 370,012 VTE-related deaths yearly in Europe.[2] Prompt referral and initiation of treatment in confirmed cases is thus pivotal. However, for both *suspected* DVT and PE, several alternative diagnoses with mimicking and overlapping signs and symptoms exist, hampering the clinical assessment.[3] Perhaps as a consequence, VTE is also one of the most frequently missed diagnoses in daily clinical care.[4,5] Therefore, to optimise the diagnostic work-up of venous thromboembolism (VTE), clinical prediction rules (CPRs) have been developed. Rigorous validation studies in a controlled research environment showed that with the use of these CPRs, referrals to secondary care were safely avoided in almost half of all patients suspected of either DVT or PE.[6,7] Consequently, the use of CPRs in suspected VTE are recommended in national and international guidelines.[8,9]

However, research on the actual impact of the CPRs when applied in inevitably less 'controlled' day-to-day care is scarce. In fact, to the best of our knowledge, only two prospective studies were published. We previously evaluated the use of the Oudega-rule for patients suspected of having DVT. This study showed that in one third of the patients the CPR was not correctly applied by GPs, commonly because of applying the CPR to patients for whom the strategy should not have been used or because of inappropriate use of the D-dimer test.[10] Clinical outcomes of such incorrect management by GPs were not reported as a main outcome in that study. The other study from Roy and co-workers evaluated the effects of implementing the Wells rule for suspected PE patients, showing that an inappropriate diagnostic assessment in suspected PE was also common (43% of all suspected patients). This inappropriate management was independently associated with a higher rate of preventable thromboembolic occurrences during follow-up.[11]

Hence, ample available evidence suggests that incorrect use of CPRs may occur more frequently than desired, possibly (i) increasing the likelihood of thromboembolic occurrences, but (ii) perhaps also leading to more unneeded, costly and burdensome referrals. Both are worrying outcomes, and thus it is important to understand better what the real-life effects are of implementing these CPRs in the diagnostic work-up for VTE in primary care. The aim of this study was therefore to evaluate the real-life impact of CPRs for both DVT and PE in daily primary care practice. Our secondary aim was to explore determinants and consequences of incorrect use of these CPRs.

METHODS

Study-design

This is (in part) an extension of a previously performed and published diagnostic cohort study that was smaller (619 patients suspected of DVT) and focussed on the implementation outcomes (such as feasibility and sustainability) of the Oudega-rule for DVT in primary care.[10] In the current study, we report on the real-life impact (i.e. the clinical outcomes) of two CPRs (Wells and Oudega rule) in 1017 patients suspected of DVT and 492 patients suspected of PE in primary care.

Participants

From October 2013 until July 2017 patients were recruited from primary care centres in the Netherlands. All patients in whom the GP suspected a diagnosis of DVT or PE (based on clinical symptoms such as calf pain or swelling for DVT, and dyspnoea, coughing or chest pain for PE) were eligible for inclusion. Institutionalized frail elderly patients were not included in this study, given that existing evidences suggests that ruling-out VTE in them with a CPR and D-dimer is unsafe.[12,13]

Study procedures

All GPs received an educational instruction on how to manage their patients according to the CPRs recommended in the primary care guidelines.[8] We explained the use of the CPR as well the patient groups in whom the rule should not be used, i.e.: (i) patients aged <18 years, (ii) pregnant or postpartum women, (iii) current use of oral anticoagulants (vitamin K antagonist, direct oral anticoagulant or low molecular heparin) and (iv) symptoms lasting longer than 30 days. For patients suspected of DVT, the Oudega rule was recommended. This CPR was modified from the original Wells rule and externally validated for the use in primary care given that the original Wells CPR for suspected DVT was shown to be unable to safely rule-out DVT in primary care.[6,14,15] The Wells rule for PE has also been validated for use in primary care, and there was no need for modification or updating.[7] Both CPRs combine seven clinical items into a score ranging from 0 to 8 for DVT and from 0 to 12.5 for PE, which classifies patients in an 'unlikely' or a 'likely' risk category of having VTE. In patients with a score of ≤3 points on the DVT CPR, or ≤4 points on the PE CPR, D-dimer had to be determined. If D-dimer was below the threshold of 500ng/mL, patients were classified as low risk of having VTE and therefore VTE was considered to be safely ruled out without the need for additional investigation. Contrary, patients with a score of ≥4 points for DVT, or ≥4.5 points for PE, or with a D-dimer either above 500ng/mL or a 'positive' result on a qualitative point-of-care

test for D-dimer, were classified as 'high risk' of having VTE. In these patients, all following existing guidelines, referral to the hospital for further diagnostic procedures was recommended. Non-referred patients were instructed to schedule a follow-up appointment with their GP in case of worsening or persistent symptoms. Participating GPs filled out a paper case report form, which consisted of questions about patient clinical characteristics, the items of the CPR, the D-dimer result and whether or not the patient was referred. In this cross-sectional diagnostic study, we used clinical follow-up of three months to assess the final diagnosis. Thus, the reference standard in our study was the clinical follow-up in the non-referred patients and further diagnostic procedures in hospital (most often a compression ultrasound of the leg in case of suspected DVT or a CT-pulmonary angiography in case of suspected PE) in the referred patients. Importantly, the above described strategy was the preferred and recommended approach, yet – after the short educational instruction – we did not rectify incorrect application of the CPR in order to mimic daily clinical care as much as possible. This thus was an assistive recommendation only, with decisions on referral left at the discretion of participating GPs.

Ethical approval

The Medical Research Ethics Committee Utrecht, the Netherlands, judged this study exempt for review according to Dutch law, given that only guideline use was evaluated. A waiver for informed consent was provided, as patient information was encrypted for the researchers. We performed this study according to the World Medical Association's Declaration of Helsinki.[16]

Outcomes

The primary outcome of this study was the impact of the everyday use of the CPRs in primary care, denoted as the diagnostic failure rate and efficiency. The failure rate was defined as the proportion of patients with a VTE diagnosis during the three-month follow-up within the non-referred patients. The efficiency was defined as the proportion of patients not referred to secondary care within the total study population. We first analysed these primary endpoints for the total suspected VTE group, thus regardless of whether or not the actual CPRs were correctly applied. Subsequently, we repeated these analyses for patients in whom the CPR was correctly or incorrectly used, and for patients suspected of having DVT or PE separately. The secondary outcome was incorrect application of the CPRs by GPs. Reasons for incorrect application were defined as (in hierarchical order): (i) the wrong CPR used (i.e. the Oudega rule for PE, or the Wells PE rule for DVT), (ii) applied in inappropriate patients (e.g. patients already on anticoagulants, pregnant or postpartum, or aged <18 years,

see above), (iii) incorrect summation of the CPR points, (iv) inappropriate use of the D-dimer test, and finally, (v) deviation from the standard referral recommendation. Each patient could only be counted once for incorrect CPR use, notwithstanding that in some patients multiple items for incorrect CPR use were applicable. Last, we analysed several possible determinants for incorrect application of both CPRs in the total patients suspected of VTE: age in categories (≤50 years, >50 and ≤75 years and >75 years), sex, heart failure, COPD/asthma, active malignancy, recent surgery or immobilisation, and (for suspected PE patients, as this was only collected for this subgroup) previous VTE. These determinants were selected as the same set of variables was evaluated in the above-mentioned study from Roy and colleagues analysing the appropriateness of the diagnostic management of suspected PE patients, to yield comparable outcomes.

Data analyses

We included only patients with complete follow-up information (i.e. a final diagnosis) in our analysis. Missing values on the items of the CPR were handled by defining these variables as absent, which results in zero points on that variable of the CPR. Baseline characteristics and the presence of all items of the CPRs for suspected DVT and PE patients are described separately. The failure rate and efficiency were quantified with corresponding 95% confidence intervals, both for DVT and PE patients and correct and incorrect CPR use. For the assessment of reasons why the CPR was incorrectly applied, we counted the reasons and described them for DVT and PE separately. To further explore the incorrect application of the CPRs, we analysed the association between the aforementioned determinants and incorrect application of the CPR by performing multivariable logistic regression. Hereto, we defined correct or incorrect use as the binary outcome and the above described potential determinants as independent covariables. This regression analysis yields adjusted odds ratios. All statistical analyses were performed in SPSS (IBM SPSS Statistics software version 25).

Patient and Public Involvement statement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

RESULTS

In total, 1,509 patients with suspected DVT and PE were included. In 32 (2.1%) patients we had missing follow-up information, and thus the study population consisted of 1,477 patients (993 with suspected DVT and 484 with suspected PE). The items of the CPRs had one to six missing values per variable (see table 1 of the appendix). The clinical characteristics of the included patients are shown in table 1. Patients suspected of having DVT were older (64 years versus 49 years) and more frequently male (42.2% versus 31.8%) as compared to patients suspected of having PE. The baseline characteristics of the patients with missing follow-up information were comparable to the study population. The overall prevalence of VTE was 18.1% (23.2% DVT and 7.9% PE).

Failure rate and efficiency of CPRs

The overall failure rate of both CPRs combined in the total study population was 1.8% (95% CI 1.02 to 3.06) and the overall efficiency 53% (95% CI 50.4 to 55.5). The failure rate and efficiency split up for correct and incorrect use of the CPRs in the total study population, suspected DVT and PE group is shown in figure 1. In the total study population the failure rate increased from 1.51% (95% CI 0.77 to 2.86) when the CPR was correctly used to 3.31% (95% CI 1.07 to 8.76) when the CPR was incorrectly used and the efficiency decreased from 58.1% (95% CI 55.2 to 61.0) to 35.7% (95% CI 30.6 to 41.1).(Figure 1) In 787 (79.3%) of the patients suspected of having DVT, the CPR was correctly applied by the GP (Figure 2). Among these patients, 408 were not referred (efficiency of 51.8%) and eight of them had a VTE; failure rate 1.96% (95% CI 0.91% to 3.98%; Figure 1). In the 206 (20.7%) patients in whom the CPR was incorrectly applied, the failure rate was 7.02% (95% CI 2.27 to 17.83); Figure 1), and the efficiency in these patients decreased to 27.7%.

Of the 351 (72.5%) patients suspected of having PE and in whom the GP applied the CPR correctly, 253 (72.1%) patients were not referred (Figure 3). Among these non-referred patients, two were diagnosed with VTE; failure rate 0.79% (95% CI 0.14% to 3.13%; Figure 1). In 133 (27.5%) suspected PE patients, the CPR was incorrectly used by the GP. Sixty-four (48.1%) of these patients were not referred. None of them had a missed VTE.

The 14 (12 DVT, 2 PE) patients in whom a VTE diagnosis was missed are described in detail in Table 2 of the appendix. Most had a low CPR score in combination with a negative D-dimer on the point-of-care assay (8 patients), or a D-dimer <500ng/mL

(3 patients). Three of the undiagnosed DVT patients decided to decline for further diagnostic testing because of high age (89, 93 and 95 years), comorbidities, and insufficient social network.

Table 1. Clinical characteristics with items of the clinical prediction rules of 993 patients suspected of deep venous thrombosis and 484 patients suspected of pulmonary embolism.

Characteristic	**Patients suspected of DVT** (n = 993)	**Patients suspected of PE** (n = 484)
Median age, years (range)	64 (15-96)	49 (13-94)
Male, n (%)	423 (42.6)	155 (32.0)
Active malignancy <6 months, n (%)	64 (6.5)	25 (5.2)
Surgery or immobilisation, n (%)	57 (5.7)	44 (9.1)
Oral contraceptive use, n (%)	59 (5.9)	n.a.
Absence of leg trauma, n (%)	782 (78.8)	n.a.
Distension of collateral veins, n (%)	231 (23.4)	n.a.
Calf swelling >3 cm n (%)	338 (34.0)	n.a.
Clinical signs of DVT, n (%)	n.a.	20 (4.1)
Haemoptysis, n (%)	n.a.	10 (2.1)
PE most likely diagnosis, n (%)	n.a.	152 (31.7)
History of VTE, n (%)	n.a.	99 (20.5)
Heart rate >100 beats/minute, n (%)	n.a.	115 (23.8)
Median score on CPR, points (range)	2 (0-7)	1.5 (0-7)
CPR score 'likely' risk category, n (%)	171 (17.2)	49 (10.1)
Median D-dimer, ng/mL (range)*	660 (100-16900)	370 (15-9000)
D-dimer 'positive' or >500ng/mL, n (%)⌐	354 (42.5)	105 (23.3)
Diagnosis of VTE**, n (%)	230 (23.2)	38 (7.9)

*DVT = deep venous thrombosis, PE = pulmonary embolism, VTE = venous thromboembolism, CPR = clinical decision rule. * Only counted when a quantitative D-dimer was measured. ⌐% of the patients in whom a D-dimer test was performed ** After three months of follow-up.*

Reasons and determinants for incorrect CPR use

The most common reason in suspected DVT and PE patients was inappropriate D-dimer testing when the score on the CPR was high (Figure 2 and 3). The second most common reason for incorrect CPR use was including patients (i) already on anticoagulants, (ii) that were pregnant or postpartum, (iii) aged < 18 years. Thirdly, application of the Oudega rule rather than the Wells rule was the reason in more than a third of patients suspected of PE. The independent risk factors for incorrect use of the CPR and the odds ratios are shown in Table 2. In patients aged between 50

and 75 years and in women, the CPRs were less frequently applied incorrectly (ORs respectively 0.71 (95% CI 0.54 to 0.94) and 0.69 (95% CI 0.54 to 0.89)), while in patients with a history of heart failure and in suspected PE patients with a previous VTE the CPRs were more frequently applied incorrectly (OR respectively 3.26 (95% CI 1.47 to 7.21) and 4.45 (95% CI 2.73 to 7.25)).

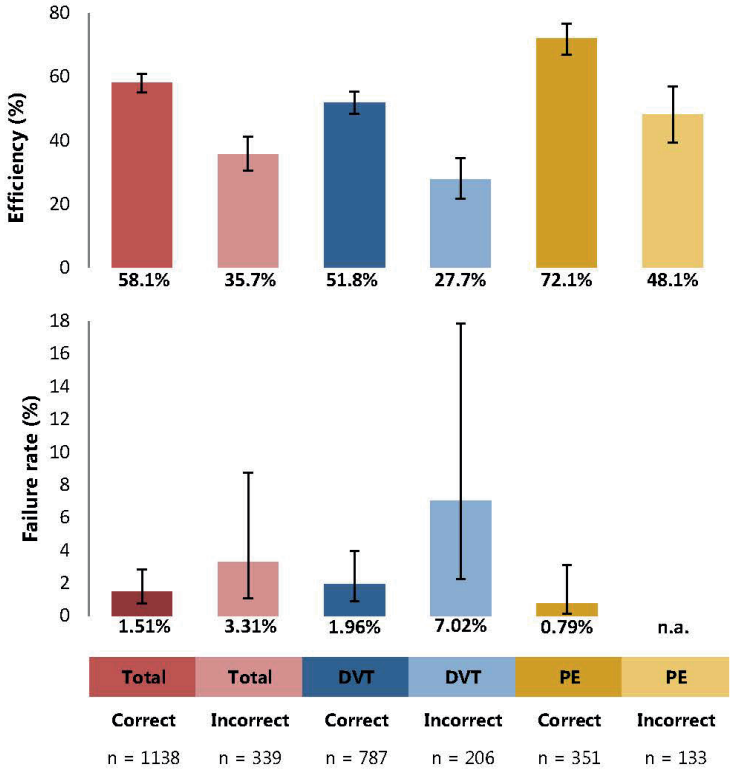


Figure 1. Bar plot of the efficiency and failure rate with corresponding 95% confidence intervals of the evaluated clinical prediction rules, stratified for incorrect and correct use, and in three groups: total included patients, patients suspected of DVT and patients suspected of PE.

DVT = deep venous thrombosis, PE = pulmonary embolism, n.a. = not applicable

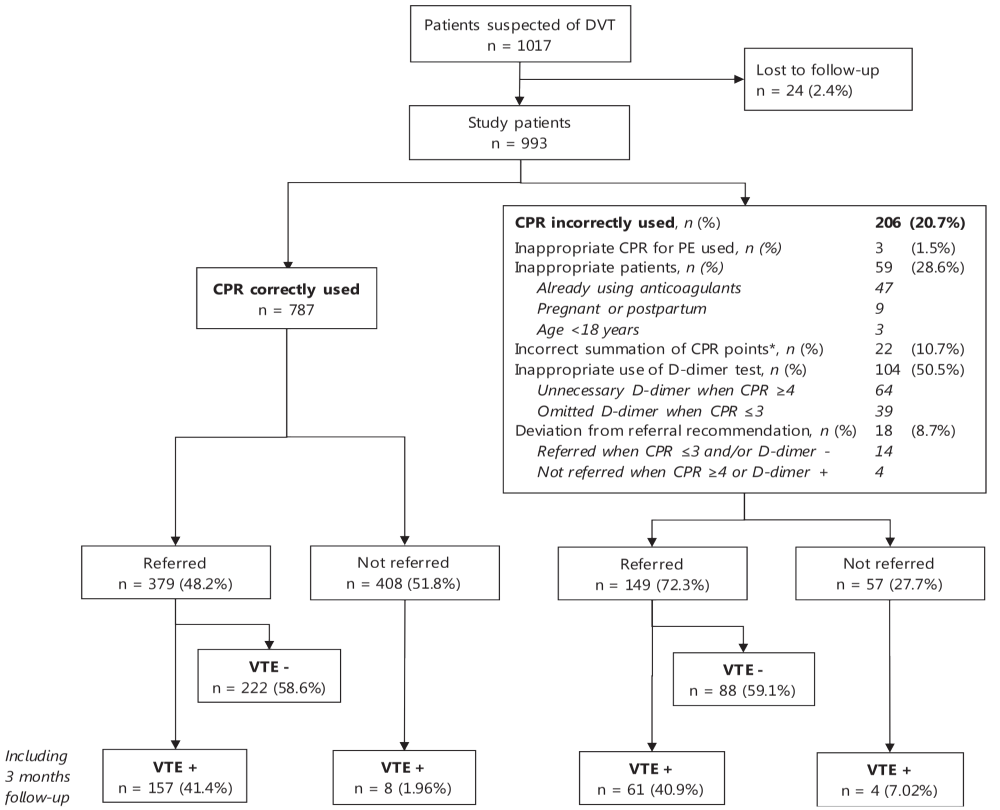


Figure 2. Flowchart of study patients suspected of deep venous thrombosis.

*DVT = deep venous thrombosis, PE = pulmonary embolism, CPR = clinical prediction rule, VTE = venous thromboembolism. * Only counted when incorrect dichotomization in CPR.*

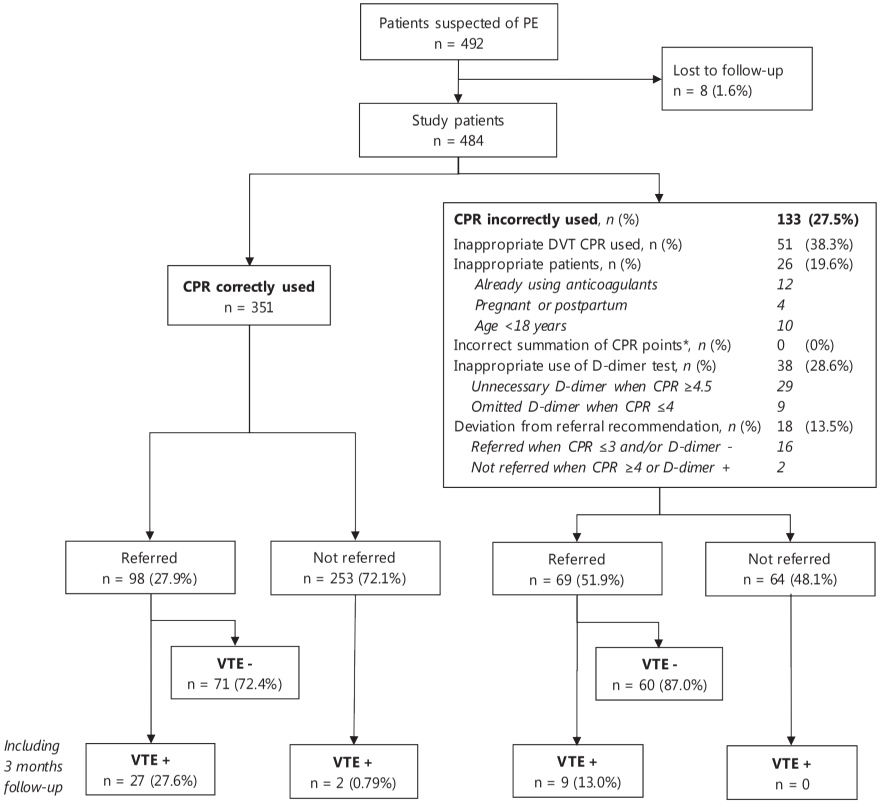


Figure 3. Flowchart of study patients suspected of pulmonary embolism.

*DVT = deep venous thrombosis, PE = pulmonary embolism, CPR = clinical prediction rule, VTE = venous thromboembolism. * Only counted when incorrect dichotomization in CPR.*

Table 2. Determinants for incorrect use of the clinical prediction rule in patients suspected of VTE.

Variable	**All patients*,** n = 1477, *n (%)*	**Correct CPR,** n = 1138 *n (%)*	**Incorrect CPR,** n = 339, *n (%)*	**Univariable analysis,** OR (95% CI)	**Multivariable analysis†,** OR (95% CI)
Age					
≤50 years	543 (36.8)	403 (35.4)	140 (41.3)	1	1
>50 and ≤75 years	651 (44.1)	516 (45.3)	135 (39.8)	0.75 (0.58 – 0.99)	0.71 (0.54 – 0.94)
>75 years	283 (19.2)	219 (19.2)	64 (18.9)	0.84 (0.60 – 1.18)	0.75 (0.53 – 1.07)
Sex					
Men	578 (39.1)	712 (62.6)	187 (55.2)	1	1
Women	899 (60.9)	426 (37.4)	152 (44.8)	0.74 (0.58 – 0.94)	0.69 (0.54 – 0.89)
Heart failure					
No	1449 (98.1)	1123 (98.7)	326 (96.2)	1	1
Yes	28 (1.9)	15 (1.3)	13 (3.8)	2.99 (1.41 – 6.34)	3.26 (1.47 – 7.21)
COPD/ asthma					
No	1307 (88.5)	1018 (89.5)	289 (85.3)	1	1
Yes	170 (11.5)	120 (10.5	50 (14.7)	1.47 (1.03 – 2.09)	1.38 (0.95 – 2.01)
Active malignancy					
No	1380 (93.9)	1066 (94.1)	314 (93.5)	1	1
Yes	89 (6.1)	67 (5.9)	22 (6.5)	1.11 (0.68 – 1.83)	1.11 (0.67 – 1.86)
Recent surgery/mmobilisation					
No	1369 (93.1)	1063 (93.7)	306 (91.1)	1	1
Yes	101 (6.9)	71 (6.3)	30 (8.9)	1.47 (0.94 – 2.29)	1.57 (1.00 – 2.47)
Previous VTE ‡					
No	386 (80.1)	304 (86.9)	82 (62.1)	1	1
Yes	96 (19.9)	46 (13.1)	50 (37.9)	4.03 (2.52 – 6.44)	4.45 (2.73 – 7.25)

CPR = clinical decision rule, VTE = venous thromboembolism, OR = odds ratio, COPD = chronic obstructive pulmonary disease.
**Data were missing for the following variables: active malignancy (8 patients), recent surgery or immobilisation (7 patients), previous VTE (2 patients).*
† Due to missing data in 10 individual patients, 1467 patients were included in the multivariable analysis.
‡ Results based only on 482 suspected PE patients.

DISCUSSION

In this real-world evaluation of the impact of CPRs for VTE we found that, if the Oudega and Wells rule were correctly used, the efficiency was high and the failure rate was acceptably low for patients suspected of DVT and PE. This is a reassuring finding, however, in almost a quarter of the 1,477 patients the CPR was incorrectly applied by GPs. This appears to lead to a considerably lower efficiency and a higher failure rate, especially in patients suspected of DVT. The most common mistakes in applying the CPRs were: D-dimer use when not needed, using the CPRs for inappropriate patients (e.g. already using an anticoagulant), and applying the Oudega rule in suspected PE patients. Incorrect application of the CPRs appeared to occur more frequently in patients with heart failure or in patients with a history of VTE (suspected PE only), whereas increasing age and female sex were associated with a lower risk of incorrect CPR application.

Strengths and limitations

The real-life impact of CPRs for both DVT and PE in primary care, including the effects of incorrect application of the CPRs, has – to the best of our knowledge – never been evaluated before. We included a large population of 1,477 patients suspected of VTE, which results in an accurate estimate of the failure rate and efficiency of the CPRs. Similar as to previous studies, we confirmed that correct application of both CPRs in suspected VTE is associated with an acceptable low failure-rate and a high efficiency. This study, however, also has some limitations. First, there is a difference of the reference standard between referred and non-referred patients. For patients referred to secondary care, the reference standard consisted of further diagnostic procedures, whereas in the non-referred patients it consisted of a three month follow-up period. Differential verification might result in bias towards overestimating the safety.[17] This approach however is routinely applied in management studies in the field of diagnostic VTE research, thus allowing our outcomes to be compared with existing literature. Second, during the inclusion period, the point-of-care test for D-dimer (Clearview Simplify) was withdrawn from the market, because of too many false-negative results likely due to peri-procedural quality-related faults when performing the test (i.e. incorrect withdrawal of capillary blood or not keeping test cold enough until use). Albeit a direct consequence of implementing this point-of-care test in day-to-day practice, its effect surely needs to be incorporated into our main outcome and analyses. These false-negative results likely resulted in more missed VTE diagnoses and therefore an underestimation of the safety of the CPRs. Indeed, 8 of all 14 patients in whom a VTE diagnosis was missed

in our study had false-negative results on this qualitative Clearview Simplify D-dimer. This in part explains the observed failure rate for the stratified DVT and PE analyses that appears to be perhaps slightly higher than was observed in earlier studies, and notably also explains the relatively wide 95% CIs that for some analyses cross the border of the commonly accepted safety threshold of 3.0%. In 357 patients (209 suspected DVT and 148 suspected PE) a quantitative D-dimer test was performed. Nevertheless, if we restrict our analysis to these patients, the main inferences of our analyses showing a higher failure rate in those in whom the CPRs are incorrectly remain the same (data not shown). Furthermore, three of the patients categorized as having a missed diagnosis of DVT were not referred to secondary care at their own request, but did contribute to the calculated failure rate in the group in which the CPR was incorrectly applied. Thus, 'incorrect' use here was intentional. Third, we did not perform a sample size calculation a-priori, given that for diagnostic validation studies (like ours) clear methodological recommendations on how to estimate a reliable sample size calculation are only recently proposed (i.e. after the initiation of our study).[18] Nevertheless, our dataset did include a total number of 1,447 patients suspected of VTE in primary care, with a total number of 268 outcome VTE events (230 DVT; 38 PE), allowing for robust statistical analyses notably for the full population; the stratified sub-analyses for DVT and PE separately though should be interpreted with a little bit more caution, notably for those suspected of PE. Lastly, we could not report on the long-term clinical outcomes. It could be hypothesized that when the CPRs are incorrectly applied, the time to diagnose VTE potentially increases. It has been speculated that such delay in diagnosis could lead to a higher risk of long-term complications, such as the post-thrombotic syndrome or chronic thromboembolic pulmonary hypertension, albeit these effects are still largely uncertain.[3,19]

Comparison with existing literature

The prevalence of DVT (23.2%) in this study corresponds with the previously described prevalence in primary care of 22%.[14] The prevalence of PE was low (7.9%), but roughly comparable with an earlier study in primary care which reported a prevalence of 12.2% and in fact almost similar to the overall prevalence of the recent PeGed study.[7,20] Apparently, the threshold of suspecting PE by physicians has lowered over time. This might be the result of the fact that physicians are more afraid to miss a PE than DVT given the associated morbidity and mortality, as well as the increasing availability of D-dimer testing and CTPA imaging.[21] Hence, it can be argued that the inclusion of more low-risk patients in this real-life observational study, has led to a higher efficiency of using the CPR for PE (65.5%) as compared to the efficiency reported in the validation study of this CPR (45.5%).[7] When the CPR was

correctly applied by the GP, we found a proportion of missed VTE diagnosis of 2.0% for patients suspected of DVT and 0.8% for patients suspected of PE. These failure rates are comparable with previous studies assessing the effects of using CPRs for VTE in primary care.[6,7,10] The incorrect use of the CPR in suspected DVT patients resulted in a high failure rate of 7.0%. Although the CPRs were incorrectly used in a quarter of our included patients (20.7% for DVT and 27.5% for PE suspected patients), this proportion is still lower than reported by previous studies. Namely, the incorrect use of the Oudega rule for DVT in the previous implementation study was 32%.[10] Another study reported that the diagnostic management of patients suspected of PE at emergency departments was inappropriate in 43%.[11] In addition, they identified determinants for inappropriate management and concluded that clinicians are more frequently deviating from the guideline in patients in which contrast media may carry increased risk (e.g. elderly) and in patients in which the symptoms could be ascribed to an alternative diagnosis. The latter might also be the case in our study population: we observed that the CPRs were more frequently applied incorrectly in patients with heart failure. In these patients, the GP might first think of this disease as diagnosis – for instance peripheral oedema mimicking DVT or shortness of breath mimicking PE – and is therefore possibly more prone to (intentionally) deviate from the CPR. Unlike the findings from Roy and colleagues, we could not confirm that increasing age is associated with an increased likelihood of incorrect CPR application. In fact, we observed the contrary; with increasing age the odds of an incorrect application of the CPR seems to decrease. Furthermore, the association between female sex and incorrect application of CPRs is not reported before. Last, the CPR was more frequent incorrectly applied in suspected PE patients with a previous VTE, which is also in contrast with previous findings.[11] Importantly, we identified determinants for incorrect use of the CPR for the total group of patients suspected of both DVT and PE. It could be argued though that some determinants could be more specifically explaining incorrect CPR use in one of these VTE-diseases.

Implications for practice

We believe our study has several implications for clinical practice. First, it is reassuring that correct application of CPRs for both suspected DVT and PE patients leads to a safe and efficient diagnostic management. Ruling-out VTE in primary care in more than half of all suspected patients at an acceptable safety margin would be considered highly attractive by many GPs, and as such our findings strengthen the evidence base of ruling-out VTE in an outpatient, community healthcare setting. However, we showed that incorrect application is common in daily primary care practice and notably is associated with an increased risk of missing VTE in those not-

referred. Of note, VTE prevalence in those referred appears to be similar in those in whom the CPRs were correctly used versus those in whom it was incorrectly applied. Albeit strictly speaking not the objective of our study, we could hypothesize about opportunities to improve the correct implementation of CPRs for VTE in primary care. First, simplification of the CPRs might enhance correct application. The current CPRs for DVT and PE consist of seven different clinical items with scores ranging from 1 to 3 points per item. This could be one of the reasons for the frequent incorrect use of the CPRs, especially since VTE is relatively rare in primary care and GPs do not often use the CPRs.[22] Recently, a simplified CPR for PE has been developed and validated in secondary care: the YEARS algorithm.[23] This algorithm only consists of three clinical items with subsequent D-dimer testing in all patients, which potentially makes it easier to apply. Validation of this algorithm in the hospital setting showed that PE could be safely excluded with a 14% reduction of CT-pulmonary angiographies as compared to the Wells rule for PE with a fixed D-dimer threshold.[24] Incorporating this new and simplified CPR might enhance guideline adherence of GPs, but awaits validation in a primary care setting before using it in daily primary care practice.[25] Second, integration of a CPR in the electronic health system might also result in more correct use of the CPR and thereby adequate management, but further research is needed. We showed that the two most common mistakes were including patients in whom the CPRs should not be used and inappropriate D-dimer testing. So, thirdly, educational training in when and how to use the CPRs plus D-dimer testing might be an opportunity to improve correct application, for example by educational outreach visits since GPs evaluated this as most encouraging.[10]

Conclusion

Correct application of CPRs for VTE in primary care is associated with a high efficiency and an acceptable low failure rate. Importantly, in nearly a quarter of patients the CPRs were incorrectly applied which resulted in a lower efficiency and a higher failure rate. Such incorrect application of CPRs was more common in the presence of concurrent heart failure.

ACKNOWLEDGEMENTS

We are grateful to all GPs who participated in the study.

Contributors
GJG, RO, KGMM and AEC designed the study and were involved in writing the original study protocol and AEC collected most study data. RvM completed the data collection, carried out the statistical analyses and drafted the first version of the manuscript. All authors (RvM, AEC, RO, FHR, KGMM and GJG) critically reviewed and revised the manuscript before providing final approval.

Funding
The Netherlands Organization for Health Research and Development (ZonMw, grant number: 837003003). ZonMw did not interfere in the design and conduct of the trial, nor had input into data collection, analysis and interpretation, or preparation of the manuscript.

Ethical approval
Medical Research Ethics Committee Utrecht, the Netherlands judged this study exempt for review (numbers 12-571/C and 14-507/C).

Competing interests
All authors have completed the ICMJE uniform disclosure form and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Data sharing statement

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

REFERENCES

1 Heit JA, Spencer FA, White RH. Epidemiology of venous thromboembolism. *J Thromb Thrombolysis* 2016;41:3–14. doi:10.1007/s11239-015-1311-6

2 Cohen AT, Agnelli G, Anderson FA, *et al*. Venous thromboembolism (VTE) in Europe - The number of VTE events and associated morbidity and mortality. *Thromb Haemost* 2007;98:756–64. doi:10.1160/TH07-03-0212

3 Kearon C. Natural history of venous thromboembolism. *Circulation* 2003;107:22–31. doi:10.1161/01.CIR.0000078464.82671.78

4 Schiff GD, Hasan O, Kim S, *et al*. Diagnostic Error in Medicine: Analysis of 583 Physician-Reported Errors. *Arch Intern Med* 2009;169:1881–7. doi:10.1007/s10459-009-9187-x

5 Hendriksen JMT, Koster-van Ree M, Morgenstern MJ, *et al*. Clinical characteristics associated with diagnostic delay of pulmonary embolism in primary care: a retrospective observational study. *BMJ Open* 2017;7:e012789. doi:10.1136/bmjopen-2016-012789

6 Buller HR, Cate-hoek AJ, Hoes AW, *et al*. Safely Ruling Out Deep Venous Thrombosis in Primary Care. *Ann Intern Med* 2009;150:229–36. doi:10.7326/0003-4819-150-4-200902170-00003

7 Geersing GJ, Erkens PMG, Lucassen WAM, *et al*. Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in Primary care: Prospective cohort study. *BMJ* 2012;345:1–10. doi:10.1136/bmj.e6564

8 NHG-werkgroep Diepe veneuze trombose en longembolie. NHG-Standaard Diepe veneuze trombose en longembolie (tweede partiële herziening). *Huisarts Wet* 2017;60:460.

9 Konstantinides S V, Meyer G, Becattini C, *et al*. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). *Eur Heart J* 2019;:1–61. doi:10.1093/eurheartj/ehz405

10 Kingma AEC, van Stel HF, Oudega R, *et al*. Multi-faceted implementation strategy to increase use of a clinical guideline for the diagnosis of deep venous thrombosis in primary care. *Fam Pract* 2016;:cmw066. doi:10.1093/fampra/cmw066

11 Roy P-M, Meyer G, Vielle B, *et al*. Appropriateness of Diagnostic Management and Outcomes of Suspected Pulmonary Embolism. *Ann Intern Med* 2006;144:157–64.

12 Schouten HJ, Koek HL, Oudega R, *et al*. Validation of the Oudega diagnostic decision rule for diagnosing deep vein thrombosis in frail older out-of-hospital patients. *Fam Pract* 2015;32:120–5. doi:10.1093/fampra/cmu068

13 Schouten HJ, Geersing GJ, Oudega R, *et al*. Accuracy of the Wells Clinical Prediction Rule for Pulmonary Embolism in Older Ambulatory Adults. *J Am Geriatr Soc* 2014;62:2136–41. doi:10.1111/jgs.13080

14 Oudega R, Hoes AW, Moons KGM. The Wells rule does not adequately rule out deep venous thrombosis in primary care patients. *Ann Intern Med* 2005;143. doi:10.7326/0003-4819-143-2-200507190-00008

15 Oudega R, Moons KGM, Hoes AW. Ruling out deep venous thrombosis in primary care - a simple diagnostic algorithm including D-dimer testing. *Thromb Haemost* 2005;94:853–8. doi:10.1160/TH04

16 World Medical Association. WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects – WMA – The World Medical Association. 2013;:1–8. doi:10.1017/CBO9781107415324.004

17 Naaktgeboren CA, Groot JAH de, Smeden M van, *et al*. Evaluating Diagnostic Accuracy in the Face of Multiple Reference Standards. *Ann Intern Med* 2013;159:195–203. doi:10.7326/0003-4819-159-3-201308060-00009

18 van Smeden M, Moons KG, de Groot JA, *et al*. Sample size for binary logistic prediction models: Beyond events per variable criteria. *Stat Methods Med Res* 2019;28:2455–74. doi:10.1177/0962280218784726

19 Klok FA, Delcroix M, Bogaard J, *et al*. Chronic thromboembolic pulmonary hypertension from the perspective of patients with pulmonary embolism. 2018;:1–12. doi:10.1111/jth.14016

20 Kearon C, de Wit K, Parpia S, *et al*. Diagnosis of Pulmonary Embolism with <scp>d</scp> -Dimer Adjusted to Clinical Probability. *N Engl J Med* 2019;381:2125–34. doi:10.1056/NEJMoa1909159

21 Barais M, Morio N, Cuzon Breton A, *et al*. 'I can't find anything wrong: It must be a pulmonary embolism': Diagnosing suspected pulmonary embolism in primary care, a qualitative study. *PLoS One* 2014;9:1–8. doi:10.1371/journal.pone.0098112

22 Schols AMR, Meijs E, Dinant G-J, *et al*. General practitioner use of D-dimer in suspected venous thromboembolism: historical cohort study in one geographical region in the Netherlands. *BMJ Open* 2019;9:e026846. doi:10.1136/bmjopen-2018-026846

23 Van Es J, Beenen LFM, Douma RA, *et al*. A simple decision rule including D-dimer to reduce the need for computed tomography scanning in patients with suspected pulmonary embolism. *J Thromb Haemost* 2015;13:1428–35. doi:10.1111/jth.13011

24 van der Hulle T, Cheung WY, Kooij S, *et al*. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet* 2017;390:289–97. doi:10.1016/S0140-6736(17)30885-1

25 van Maanen R, Rutten FH, Klok FA, *et al*. Validation and impact of a simplified clinical decision rule for diagnosing pulmonary embolism in primary care: design of the PECAN prospective diagnostic cohort management study. *BMJ Open* 2019;9:e031639. doi:10.1136/bmjopen-2019-031639

Authorship statement

The study was previously set up by the co-authors. I completed the data collection, carried out the statistical analyses, and drafted the first version of the manuscript. During the whole process, I asked for and implemented input and feedback from the other contributors to this study.

SUPPLEMENTAL FILES

Supplement 1. Missing values included patients
Supplement 2. Description of non-referred patients with VTE

SUPPLEMENT 1. Missing values of included patients

Characteristic	**Missing values suspected DVT patients** (n = 993)	**Missing values suspected PE patients** (n = 484)
Median age	0	0
Male	0	0
Active malignancy <6 months	6	2
Surgery or immobilisation	4	3
Oral contraceptive use	0	n.a.
Absence of leg trauma	1	n.a.
Distension of collateral veins	4	n.a.
Calf swelling >3 cm	2	n.a.
Clinical signs of DVT	n.a.	3
Haemoptysis	n.a.	2
PE most likely diagnosis	n.a.	4
History of VTE	n.a.	2
Heart rate >100 beats/minute	n.a.	5
Score on CPR	0	0
D-dimer **	41	11
Diagnosis of VTE*	0	0

*DVT = deep venous thrombosis, PE = pulmonary embolism, VTE = venous thromboembolism, CPR = clinical decision rule. * During three months of follow-up. ** Only defined as missing when a D-dimer measurement should been done according to the CPR.*

SUPPLEMENT 2. Description of non-referred patients with VTE

Patient No.	Age (years)	Sex	DVT / PE	CPR score	D-dimer	Remarks
1	60	Female	DVT	1	Negative	
2	54	Male	DVT	2	Negative	
3	84	Male	DVT	2	Negative	
4	44	Female	DVT	2	Negative	
5	73	Female	DVT	2	Negative	
6	39	Male	DVT	3	Negative	
7	53	Male	DVT	3	Negative	
8	76	Female	DVT	0	280 ng/mL	
9	24	Male	DVT	5	240 ng/mL	
10	89	Female	DVT	2	Positive	Preference of patient to decline further diagnostic procedures
11	93	Female	DVT	3	Positive	Preference of patient to decline further diagnostic procedures
12	95	Female	DVT	2	2540 ng/mL	Preference of patient to decline further diagnostic procedures
13	73	Female	PE	1.5	Negative	
14	51	Male	PE	0	290 ng/mL	

DVT = deep venous thrombosis, PE = pulmonary embolism

3

Accuracy of the physicians' intuitive estimation in the diagnostic management of pulmonary embolism: an international individual patient data meta-analysis

Rosanne van Maanen
Emily S.L. Martens
Toshihiko Takada
Frederikus A. Klok
Jeanet W. Blom
Karel G.M. Moons
Frans H. Rutten
Maarten van Smeden
Geert-Jan Geersing
Kim Luijken

Submitted

ABSTRACT

Background In patients suspected of pulmonary embolism (PE), physicians often rely on an intuitive estimation ('gestalt') of PE presence. Although shown to be predictive, gestalt is also criticized given its lack of standardization and assumed variation among different physicians and healthcare settings.

Objectives To assess the diagnostic accuracy of physician's gestalt in diagnosing PE and gain more insight into its possible variation.

Methods We performed an individual patient data meta-analysis (IPD-MA) including patients suspected of PE. The primary outcome was the diagnostic accuracy of gestalt for diagnosing PE, quantified as a risk ratio (RR) between gestalt and PE from a two-stage random-effect log-binomial meta-analysis regression as well as gestalts' sensitivity and specificity. Variability of these indices was explored across healthcare settings, study year, PE prevalence, and the patient subgroups based on age, sex, heart failure, chronic lung disease, as well as items of the Wells algorithm.

Results We analysed 20,770 patients suspected of PE from 16 individual studies. The prevalence of PE in patients with and without 'gestalt positive' was 28.8% versus 9.1%, corresponding to a pooled RR of 3.02 (95% CI 2.35-3.87), a pooled sensitivity of 0.74 (95% CI 0.68-0.79) and a pooled specificity of 0.61 (95% CI 0.53-0.68). The diagnostic accuracy of gestalt varied across individual studies, yet performance remained stable across all healthcare settings and subgroups.

Conclusions A positive gestalt estimation predicts a three-fold higher risk of PE in suspected patients compared to negative gestalt. The diagnostic accuracy of positive gestalt was consistent across subgroups and healthcare settings.

ESSENTIALS

- Insight into the diagnostic accuracy of 'physicians' gestalt' of pulmonary embolism (PE).
- An individual patient data meta-analysis of 20,770 patients suspected of PE.
- A positive physician's gestalt estimation increased the risk of a PE diagnosis about threefold.
- The diagnostic accuracy was stable across patient subgroups or healthcare settings.

INTRODUCTION

Pulmonary embolism (PE) is a potentially fatal disease that warrants early detection and treatment. However, diagnosing PE is challenging and a delayed diagnosis is common.[1] Symptoms of shortness of breath and chest pain may also occur in other, often less severe conditions such as intercostal neuralgia or localised chest myalgia.[2,3] Haemoptysis is more specific but also an uncommon symptom which nevertheless may be due to fulminant coughing due to whatever cause. The classical triad of shortness of breath, pleuritic pain, and haemoptysis is thus only present in 10% of the patients with established PE.[3] Considering the potential severity, physicians have a low threshold for additional testing in patients in whom they suspect PE, either by D-dimer testing (biomarker used for clot detection) or direct (referral for) computed tomography pulmonary angiography (CTPA; the reference standard for diagnosing PE). Historically, the decision-making process for this challenging diagnosis was mainly driven by the clinicians' intuitive judgement called 'gestalt'. Subsequently, this probability estimation has become an important component of clinical decision rules (CDR) for PE diagnosis.[4–6]. Gestalt has since then repeatedly shown to increase PE probability in individual diagnostic prediction studies.[7,8]

Nevertheless – although intuitively attractive – the merit of gestalt in the diagnostic management of patients suspected of PE is also debated. Several studies showed that when physicians only used gestalt in the work-up of suspected PE, the predicted risk of PE was often overestimated compared to the observed risk, resulting in a decreased overall efficiency of the diagnostic process.[9–11] Another,

perhaps even more important concern is that a gestalt estimation is dependent on clinical experience with diagnosing PE in everyday practice, resulting in variable interobserver reproducibility.[12–14] Thus, the diagnostic accuracy of gestalt in patients suspected of PE may be unstable and potentially vary between healthcare settings due to differences in experience among physicians working in that setting, as well as differences in 'case-mix'. This, however, has never been adequately studied. Tackling this knowledge gap is needed to understand the diagnostic 'behavior' of this subjective item in assessing PE probability, and for re-assuring physicians when deciding upon the context in which clinical gestalt may be of merit, and when not.

Therefore, this study aimed to quantify the diagnostic accuracy of gestalt in the diagnostic management of PE in suspected patients across different healthcare settings and patient subgroups. Hereto, we performed an ancillary analysis of a large international individual patient data meta-analysis (IPD-MA) including more than 35,000 patients suspected of PE.[15] The primary aim of this current study was to compare the estimated risk ratio (RR) of PE presence as a function of clinical gestalt across relevant patient subgroups and healthcare settings. Next sensitivity and specificity of gestalt were estimated across the same patient subgroups and healthcare settings.

METHODS

This is an ancillary analysis of a pre-registered IPD-MA (PROSPERO database for systematic reviews number CRD42018089366), of which a protocol has been published.[15] Previous studies using this IPD-MA explored the diagnostic accuracy of existing clinical decision rules (CDRs) for PE across clinically relevant subgroups and healthcare settings, but not of physician's gestalt.[16,17] Ethical approval and informed consent of individual patients were obtained in each included original study. Throughout this IPD-MA, we adhered to the Preferred Reporting Items for Systematic Review and Meta-Analysis for Individual Participant Data (PRISMA-IPD) and the Preferred Reporting Items for Diagnostic Test Accuracy (PRISMA-DTA) guideline on reporting of systematic reviews including individual-patient data.[18,19]

Study eligibility, identification, and selection

The systematic search strategy for this IPD-MA including information sources and the study selection process was described in detail previously.[15] In short, MEDLINE was searched from 1 January 1995 until 1 November 2021. Studies were eligible if they evaluated diagnostic strategies for PE, had a prospective cohort design, included patients suspected of PE, and objectively confirmed a diagnosis of venous thromboembolism (VTE) with either imaging or clinical follow-up of at least one month. For the current analyses, we excluded studies not assessing the variable 'PE most likely diagnosis' and studies selectively including only patients with low clinical pre-test probability. Full-text screening was performed independently by two couples (GJG and NK, and FAK and NvE). Authors from the eligible studies were asked to provide de-identified individual patient data. The risk of bias in the individual studies was independently assessed by three pairs of authors (GJG and TT, NvE and NK, and FAK and MAMS) by using the QUADAS-2 tool for assessment of the risk of bias and applicability of primary diagnostic accuracy studies.[20] Disagreements were solved by discussion within each pair and between pairs.

Variable measurements

Clinical gestalt was defined conform the definitions used in the Wells and other diagnostic PE decision rules: 'whether PE was considered the most likely diagnosis'. If PE was not considered the most likely diagnosis, the gestalt item was defined as negative, while if PE was considered the most likely diagnosis, the gestalt item was defined as positive. The RR was estimated across different patient subgroups, defined by the following variables: male versus female patients, age on a continuous scale, heart failure (present or absent at presentation with suspected PE), chronic lung

disease (defined as chronic obstructive pulmonary disease, asthma, pulmonary fibrosis or any other chronic lung disease present or absent at presentation with suspected PE), subgroups based on the presence or absence of the Wells items (i.e. clinical signs/symptoms of DVT, previous VTE, hear rate >100, haemoptysis, immobilization during >3 day/ surgery in previous 4 weeks, active malignancy) and subgroups of patients without, with one, or with two or more Wells CDR items in addition to the gestalt item. We also estimated the RR across study year, PE prevalence, and categories of healthcare settings. We identified the following healthcare settings based on a previous IPD-MA from our study group: (i) hospital or nursing home care, (ii) referred secondary or emergency care, (iii) primary healthcare, and (iv) self-referral emergency care.[17] Five expert panel members (GJG, FAK, MAMS, NK, and NvE) independently categorized each study into one of the four defined healthcare settings and discussed disagreements until they reached a consensus. When studies were performed in more than one setting, individual patients were categorized based on the information provided by the principal investigators.

Missing data

A summary of missing data in each original study is shown in Supplement 1. Variables were either partially missing (i.e., missing in a certain proportion of patients within the study) or systematically missing (i.e., completely missing in the study). Partially missing values were imputed within each study using multiple imputation techniques with all available variables, including the outcome, using the R-package MICE [21], unless the variables were missing in more than 80% of patients.[22] Ten imputation datasets per study were created. Measures of log-RR, logit-sensitivity, and log-specificity were computed in each imputed set and combined using Rubin's rules.[23]

Data analysis

First, we described the characteristics and the prevalence of PE of the included patients stratified by gestalt positively versus negatively scored. The primary analysis of this study was the diagnostic accuracy of the clinical gestalt estimate for PE presence. We quantified the diagnostic accuracy of gestalt as RR, i.e., the presence (or risk) of having PE in individuals with a positive versus negative gestalt item, as well as its sensitivity and specificity.[24] We expressed the diagnostic accuracy of gestalt as RR rather than the commonly used diagnostic odds ratio because of non-collapsibility issues of the odds ratio derived from a standard logistic regression model.[25] We performed a two-stage meta-analysis. In the first stage, the RR was estimated using a log-binomial regression model in each study. In the second stage, these estimates were pooled using a separate intercept for each study and

a random effect for gestalt using restricted maximum likelihood estimation, which allows studies to differ in the association between gestalt and final PE diagnosis because of real differences in the RR rather than chance variation only.[26] This resulted in an overall RR and 95% prediction interval (PI) for the association between gestalt and PE diagnosis.

To gain insight into the diagnostic accuracy of the gestalt item 'PE most likely diagnosis' in different settings and patient types, we stratified the data into the following subgroups: male versus female patients, patients with versus without heart failure, patients with versus without chronic lung disease, the publication year of the study (before 2010 versus 2010 and later), and subgroups based on the presence or absence of the Wells items and subgroups of patients without, with one, or with two or more Wells CDR items in addition to the gestalt item, and lastly, three different healthcare settings: (i) hospital or nursing home care, (ii) referred secondary or emergency care, (iii) primary healthcare. The random-effects meta-analysis model was fitted in each subgroup separately to estimate the subgroup-specific RR of gestalt and a final PE diagnosis. These analyses yielded RRs with 95% PI, the latter indicating the between-study heterogeneity. Furthermore, we assessed how the RR varied across age on a continuous scale by fitting a log-binomial model on the stacked imputed data with an interaction between the variable gestalt and age, where age was modeled using a restricted cubic spline with five knots (on the percentiles 0.05, 0.275, 0.50, 0.725, and 0.95) for each imputed data set. Then, the risk of PE across ages 18 to 90 years was predicted from this model under gestalt positive and negative using the stacked imputed data. This RR was computed from the ratio of these predicted risks and plotted. We plotted the prevalence of PE in each study against the RR of gestalt. It is previously shown that the efficiency and failure rate of diagnostic strategies are dependent on the prevalence of PE; as PE prevalence increases, the failure rate increases and the efficiency decreases.[17] Therefore, we hypothesized that the diagnostic accuracy of gestalt would also be related to PE prevalence.

Last, we calculated the pooled sensitivity and specificity for the gestalt item on the final PE diagnosis in all above-described subgroups using a bivariate meta-regression model on the logit sensitivity and logit specificity of each study.[24] This yielded an estimate and 95% confidence interval (CI) for sensitivity and specificity. All analyses were performed using R, version 4.0.3 (R Foundation for Statistical Computing, www.R-project.org), particularly using the metafor package.[27]

RESULTS

Study selection and included patients

The systematic literature search retrieved 3,892 unique studies. A total of 23 studies fulfilled the eligibility criteria and the original IPD was retrieved from corresponding authors, resulting in 35,248 unique patients suspected of PE. We excluded three studies that did not assess the variable 'PE most likely' [28–30] and four studies selectively including patients with a low clinical pre-test probability (i.e.. studies evaluating the PERC CDR) [31–34]. Hence, in the current analysis, 16 studies were included with in total 20,770 patients suspected of PE. The risk of bias in each included study was generally scored as low (Supplement 2). Characteristics of the included studies are summarised in Table 1. The prevalence of PE ranged from 7.4% to 40.9% and the percentage of patients in whom PE was scored as the most likely diagnosis ranged from 22.1% to 62.1% in the individual studies.

Patient characteristics stratified by the gestalt item are shown in Table 2. The overall prevalence of PE was 19.8%; 9.1% in the negative gestalt group versus 28.8% in the positive gestalt group. The median age was 56.6 years and 60.1% was female. Patients in whom the gestalt item was positively scored had less frequently concurrent heart failure or chronic lung disease, but had more often risk factors for PE, namely an active malignancy, recent surgery or immobilisation, clinical signs of DVT, and/or a history of VTE. The median D-dimer level was higher in patients with gestalt positively scored compared to patients with gestalt negatively scored (1001 ng/mL (IQR 510; 2421) versus 582 ng/mL (IQR 298; 1200)).

Main outcomes

In patients in whom the gestalt item was positively scored, the risk or probability of having PE was on average 28.8%, higher than those with gestalt negative (9.1%) or the average (18.4%). The point estimates of the RR for the association between gestalt and a final PE diagnosis from the individual studies ranged from 1.46 to 7.71, with a pooled point estimate of 3.02 (95% PI 1.14, 7.94), see figure 1. Heterogeneity across studies was observed, depicted by the relatively wide 95% prediction interval and an I squared of 90.6%. The RRs for each subgroup are shown in figure 2. The RR in females was 3.26 (95% PI 1.37, 7.78) and in males 2.79 (95% PI 0.93, 8.34). Three studies did not report the presence or absence of heart failure, and four studies did not report chronic lung disease. These studies were therefore excluded from the subgroup analysis for comorbidities. The RRs for patients with and without heart failure and with and without chronic lung disease were 1.98 (95%

Table 1. Characteristics of included studies of patients suspected of pulmonary embolism

Author, year [reference]	Continent	Patients and healthcare setting	Number of patients included	PE prevalence (%)	'Gestalt' positively scored (%)
Sanson et al, 2000 [35]	The Netherlands	Referred secondary care and inpatients	517	30.9	60.6
Perrier et al, 2004 [36]	Switzerland	Referred secondary care	965	23.7	29.4
Perrier et al, 2005 [37]	Switzerland	Referred secondary care	755	26.1	38.5
Kearon et al, 2006 [38]	Canada	Primary healthcare and inpatients	1123	15.0	48.4
van Belle et al, 2006 [39]	The Netherlands	Referred secondary care and inpatients	3296	21.2	61.5
Goekoop et al, 2007 [40]	The Netherlands	Referred secondary care	876	12.6	47.4
Righini et al, 2008 [41]	Switzerland	Referred secondary care	1692	21.3	45.9
Douma et al, 2011 [42]	The Netherlands	Referred secondary care and inpatients	807	23.8	56.5
Galipienzo et al, 2012 [43]	Spain	Referred secondary care	240	26.3	22.1
Geersing et al, 2012 [5]	The Netherlands	Primary healthcare	597	12.2	55.6
Schouten et al, 2014 [44]	The Netherlands	Primary healthcare and nursing homes	129	39.8	49.6
Righini et al, 2014 [45]	Switzerland	Referred secondary care	3324	19.2	53.1
Mos et al, 2014 [46]	The Netherlands	Referred secondary care and inpatients	279	40.9	62.1
Penaloza et al, 2017 [47]	France & Belgium	Referred secondary care	705	21.7	45.8
van der Hulle et al, 2017 [4]	The Netherlands	Referred secondary care and inpatients	3448	13.7	50.0
Kearon et al, 2019 [48]	Canada	Primary healthcare and inpatients	2017	7.4	21.0

PI 1.42, 2.76) versus 3.07 (95% PI 1.06, 8.89), and 2.19 (95% PI 0.62, 7.72) versus 3.11 (95% PI 1.03, 9,41), respectively. The RRs in the three different settings were 4.03 (95% PI 0.09, 182.9) for hospital or nursing home care, 2.85 (95% PI 0.90, 8.99) for emergency ward or hospital care, and 3.81 (95% PI 3.39, 4.28) for primary healthcare. There were no studies in our selection from the setting 'self-referral emergency care'. The subgroups defined by the publication year of the study showed a comparable RR: 2.89 (95% PI 1.15, 7.24) for studies performed before 2010 and 3.17 (95% PI 0.76, 13.29) for studies performed in 2010 and later. The RRs in the subgroups based on the presence or absence of any or more of the other Wells items were comparable and ranged between 2.01 and 3.19. The plot of the RR for age on a continuous scale shows that the RR is decreasing with increasing age (figure 3), albeit with a wide confidence interval, especially in the youngest and oldest patients due to fewer observations in these age groups. The sensitivity and specificity for all subgroups are shown in Supplemental table 3. The pooled sensitivity of all studies was 0.74 (95% CI 0.68 to 0.79) and the specificity 0.61 (95% CI 0.53 to 0.68), with similar inferences across all evaluated subgroups (Supplement 3). Lastly, we plotted the RR of PE and the gestalt item against the prevalence of PE in individual studies (Supplement 4). This did not reveal a clear relation between the diagnostic accuracy of gestalt and the prevalence of PE.

Table 2. Clinical characteristics of patients in whom the physician scored PE as most likely diagnosis (Gestalt +) and patients in whom the physician did not score PE as most likely diagnosis (Gestalt -)

Characteristic	**Missing Proportion *,** %	**Gestalt +** (n = 9860)	**Gestalt -** (n = 10910)	**Total** (n= 20770)
Median age, years (IQR)	0.0	57.0 (42.8, 71.0)	56.0 (41.1, 70.0)	56.6 (42.0, 70.0)
Female, n (%)	0.0	5919 (60.0)	6570 (60.2)	12489 (60.1)
Heart failure, n (%)	20.0	483 (5.5)	599 (7.0)	1082 (6.2)
Chronic lung disease, n (%)	12.8	906 (10.6)	1166 (14.3)	2072 (12.4)
Active malignancy <6 months, n (%)	0.0	1266 (12.8)	938 (8.6)	2204 (10.6)
Surgery or immobilisation <4 weeks, n (%)	0.0	1814 (18.4)	1370 (12.6)	3184 (15.3)
Clinical signs of DVT, n (%)	0.0	951 (9.6)	587 (5.4)	1538 (7.4)
Haemoptysis, n (%)	0.0	477 (4.8)	501 (4.6)	978 (4.7)
History of VTE, n (%)	0.0	1653 (16.8)	1244 (11.4)	2897 (13.9)
Heart rate >100 beats/minute, n (%)	0.0	2385 (24.2)	2622 (24.0)	5007 (24.1)
Median D-dimer, ng/mL (IQR)	15.0	1001.0 (510.0, 2421.0)	582.0 (298.0, 1200.0)	780.0 (354.0, 1706.0)
Diagnosis of PE**, n (%)	0.0	2844 (28.8)	988 (9.1)	3832 (18.4)

IQR, interquartile range; DVT, deep venous thrombosis; PE, pulmonary embolism; VTE, venous thromboembolism.

** After imputation in each study; ** After three months of follow-up.*

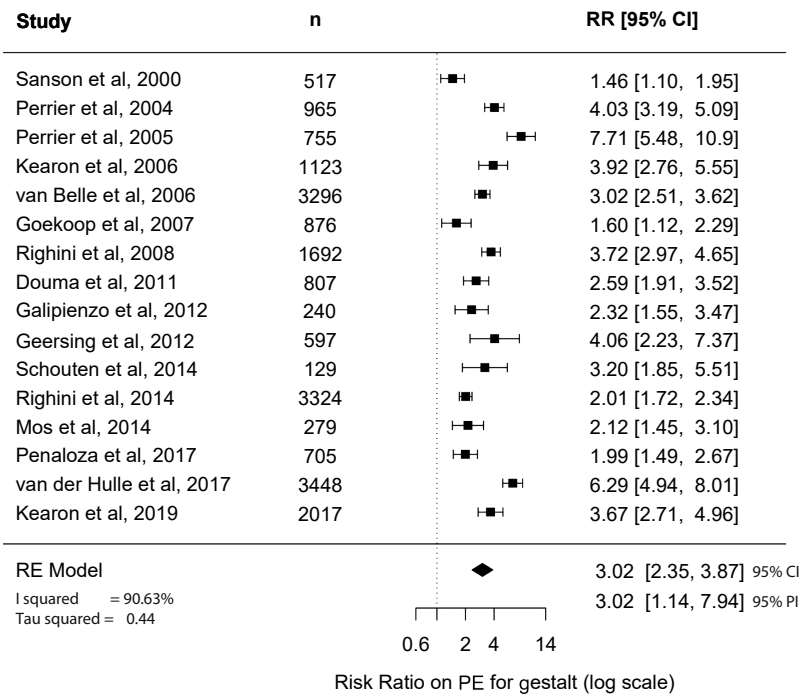


Figure 1. Risk ratio of PE with gestalt in individual studies and the pooled estimate

RR: risk ratio; CI: confidence interval; PI: prediction interval; n: number of included patients

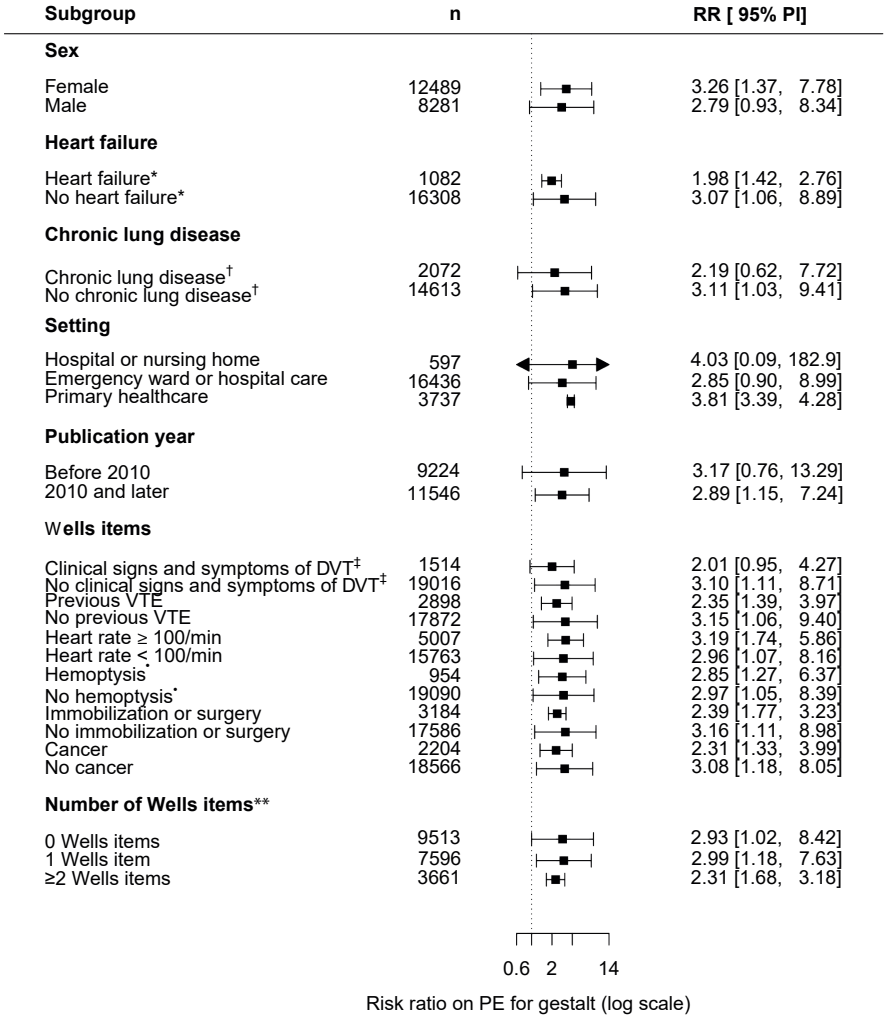


Figure 2. Risk Ratio of PE with gestalt in the predefined subgroups.

*DVT: deep venous thrombosis; VTE: venous thromboembolism; RR: risk ratio; PI: prediction interval. *3 studies excluded because of systematic missing, † 4 studies excluded because of systematic missing, ‡ 1 studie excluded because of convergence issues, •2 studies excluded because of convergence issues. ** number of Wells items in addition to the gestalt item.*

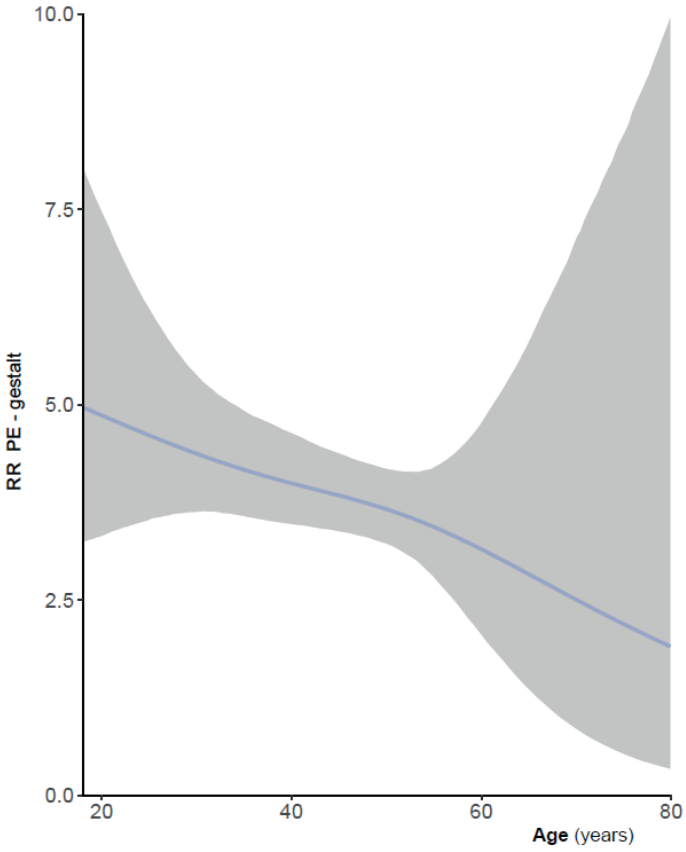


Figure 3. Risk ratio of PE of gestalt against age on a continuous scale

DISCUSSION

In this IPD-MA including 20,770 patients suspected of acute PE, we explored the diagnostic accuracy of gestalt in the diagnostic management of PE. Overall, a positive gestalt estimation (i.e. a positive score on the item 'PE most likely diagnosis') increased the risk (or probability) of having PE on average by threefold. Although there was considerable heterogeneity around this pooled estimate across studies, the diagnostic accuracy of gestalt remained remarkably stable across various patient subgroups, healthcare settings, PE prevalence, and year of study publication. Only the analysis with age showed that with increasing age, the diagnostic accuracy is slowly decreasing, albeit with wide and overlapping prediction intervals.

Consequently, in our extensive subgroup analyses in this large IPD-MA, we could not find any patient subgroup nor clinical setting or underlying PE prevalence where gestalt is not beneficial in the diagnostic management of PE. Our findings reject the previous hypothesis that assumed that large variability and subjectivity of gestalt would severely hamper its diagnostic accuracy in different healthcare populations. As an example, much attention was paid to the report of a previous study that the pre-test probability of PE increases with clinical experience.[14] Interestingly, however, this seems to be in contrast to our finding that the healthcare setting, and thus different physician experience, did not have a substantial impact on the accuracy of gestalt in the diagnosis of PE. Another previous study showed that with every additional point in the Wells rule, patients had a 1.2-fold increased chance of being assigned the subjective 'PE most likely diagnosis'.[8] Indeed, we see that in patients with an active malignancy, recent surgery of immobilisation, clinical signs of DVT, and a history of VTE the gestalt item is more frequently positively scored. However, the diagnostic accuracy of gestalt expressed as RR was comparable across patient subgroups based on the presence of these Wells items when assessed in isolation or even the combined presence of these Wells items. An observation that we in part did observe in our current analyses is the impact of D-dimer on gestalt. D-dimer values indeed were higher in patients with a positive than in a negative gestalt. Interpreting this finding, one has to realise that in some studies – for instance the YEARS study [4] – the D-dimer result was already known before the scoring of the subjective item 'PE most likely diagnosis'. This has likely resulted in an 'overestimation' of the accuracy of gestalt alone in these studies.

Strengths and limitations

We performed a comprehensive IPD meta-analysis including data from many individual patients suspected of PE, thus being the largest study exploring the diagnostic accuracy of gestalt. This allowed us to perform robust subgroup analyses and to provide reliable estimates of the diagnostic accuracy of gestalt across different healthcare settings and patient subgroups. We performed multilevel imputation of missing values and state-of-the-art statistical methods to quantify the diagnostic value of gestalt in suspected PE patients.

Yet, for full appreciation, several limitations must be discussed. The most important limitation is that the subjective gestalt item was scored in various ways in the individual studies. For instance, in some studies (n=12) the gestalt estimation was part of the CDR (i.e. the Wells rule and the YEARS algorithm) and thus was scored in the context of these CDRs. On the other end of the spectrum, we included studies evaluating the Geneva rule in which the gestalt item was scored only for research purposes, thus not being part of a CDR. Finally, in studies performed in the primary care setting it was always scored before knowing the D-dimer result, whereas likely in most studies performed in the hospital setting the gestalt estimate was at least to some extent influenced by D-dimer because the result was available when the CDR was filled out. Nevertheless, maybe counter-intuitive, the highest RR for gestalt was found in a hospital-based study evaluating the Geneva score (7.71, 95% PI 5.48, 10.9)[37] and lowest in a hospital-based study evaluating the Wells rule (1.46, 95% PI 1.10, 1.95)[35]. Thus, we believe that our inference that the overall estimate of a threefold increase in PE risk in patients in whom gestalt is scored positively, seems to be closest to the real effect, although, with a substantial range, but always clearly an RR above 1. Another limitation is that the number of patients in the subgroups with heart failure or chronic lung disease was relatively low, and four studies did not mention these comorbidities. Similarly, for the subgroups with the Wells items 'clinical signs/symptoms of DVT' and 'haemoptysis', the counts in some studies were low, resulting in convergence issues. Therefore, due to empty cells in these studies in our two-stage meta-analytical approach, we had to exclude these studies from the specific subgroup analyses. This could have resulted in less precise estimates of the RRs and wider prediction intervals.

Interpretation of the main findings and the clinical implications

When interpreting our findings, it is important to acknowledge that the focus of this IPD-MA was to explore the variability of the diagnostic accuracy of the gestalt item in the diagnostic management of PE as an alone-standing item in different patient

subgroups and healthcare settings. The goal was *not* to define whether a CDR should or should not include a gestalt estimation. We did not perform such multivariable analyses exploring the incremental diagnostic value of gestalt beyond other CDR items.

From a clinical perspective, we believe the following inferences can be drawn: although heterogeneity across individual studies was observed, the diagnostic accuracy of gestalt remained remarkably stable with on average a threefold increased risk of PE in patients with a positive gestalt across all our evaluated patient subgroups and healthcare settings. Hence, this heterogeneity of gestalt across our studies in this IPD-MA was not explained by differences in case-mix or healthcare settings among these individual studies. Although speculative, our analyses suggest that the diagnostic accuracy of the intuitive gestalt item ('gut feeling' of physicians on PE presence) is thus not substantially related nor influenced by risk factors for VTE, sex, age, comorbidity of patients, or healthcare setting. Rather, it seems to be related to other factors that are harder to define such as the physician's clinical impression of the severity of the disease.[49] This is supported by previous work; if physicians experienced a 'sense of alarm' in patients with shortness of breath, the odds of having a life-threatening disease increased about twofold, also for pulmonary embolism.[50] Based upon our analyses, we might conclude that this intuitive gestalt estimation or 'sense of alarm' holds its merit, albeit with remaining not fully explained heterogeneity, across *all* patients with suspected PE, regardless of the healthcare setting in which they present themselves or to what subgroup they belong. Only the analysis with age showed that the diagnostic accuracy of gestalt might be slowly declining with increasing age. Indeed, diagnosing PE could be extremely challenging in elderly patients given the subtle signs and symptoms and the presence of other cardiac or pulmonary comorbidities which may mimic PE symptoms, as well as that frailty may negatively impact the accuracy of diagnostic tests.[51,52] An area of future research would therefore be the elderly population suspected of PE to evaluate what diagnostic strategy fits them best.

Conclusion

A positive gestalt estimation in the diagnostic management of PE predicts on average a three-fold higher risk of PE in suspected patients compared to a negative gestalt estimation. Although heterogeneity was observed across individual studies, the diagnostic accuracy of gestalt remains remarkably stable across different subgroups of patients and healthcare settings.

REFERENCES

1 Maanen R van, Trinks-Roerdink EM, Rutten FH, Geersing GJ. A systematic review and meta-analysis of diagnostic delay in pulmonary embolism. *Eur J Gen Pract* Taylor & Francis; 2022; 28: 165–72.

2 Erkens PMG, Lucassen WAM, Geersing GJ, van Weert HCPM, Kuijs-Augustijn M, van Heugten M, Rietjens L, ten Cate H, Prins MH, Büller HR, Hoes AW, Moons KGM, Oudega R, Stoffers HEJH. Alternative diagnoses in patients in whom the GP considered the diagnosis of pulmonary embolism. *Fam Pract* 2014; 31: 670–7.

3 Meyer G, Roy PM, Gilberg S, Perrier A. Pulmonary embolism. *BMJ* 2010; 340: 974–6.

4 van der Hulle T, Cheung WY, Kooij S, Beenen LFM, van Bemmel T, van Es J, Faber LM, Hazelaar GM, Heringhaus C, Hofstee H, Hovens MMC, Kaasjager KAH, van Klink RCJ, Kruip MJHA, Loeffen RF, Mairuhu ATA, Middeldorp S, Nijkeuter M, van der Pol LM, Schol-Gelok S, et al. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet* 2017; 390: 289–97.

5 Geersing GJ, Erkens PMG, Lucassen WAM, Büller HR, Ten Cate H, Hoes AW, Moons KGM, Prins MH, Oudega R, Van Weert HCPM, Stoffers HEJH. Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in Primary care: Prospective cohort study. *BMJ* 2012; 345: 1–10.

6 Le Gal G, Righini M, Roy P-M, Sanchez O, Aujesky D, Bounameaux H, Perrier A. Prediction of Pulmonary Embolism in the Emergency Department: The Revised Geneva Score. *Ann Intern Med* 2006; 144: 165–71.

7 Klok FA, Karami Djurabi R, Nijkeuter M, Huisman M V. Alternative diagnosis other than pulmonary embolism as a subjective variable in the wells clinical decision rule: Not so bad after all [8]. *J Thromb Haemost* 2007; 5: 1079–80.

8 Klok FA, Zidane M, Djurabi RK, Nijkeuter M, Huisman M V. The physician's estimation "alternative diagnosis is less likely than pulmonary embolism" in the Wells rule is dependent on the presence of other required items [5]. *Thromb Haemost* 2008; 99: 244–5.

9 Zarabi S, Chan TM, Mercuri M, Kearon C, Turcotte M, Grusko E, Barbic D, Varner C, Bridges E, Houston R, Eagles D, de Wit K. Physician choices in pulmonary embolism testing. *CMAJ* 2021; .

10 Kline JA, Stubblefield WB. Clinician Gestalt Estimate of Pretest Probability for Acute Coronary Syndrome and Pulmonary Embolism in Patients With Chest Pain and Dyspnea. *Ann Emerg Med* Mosby; 2014; 63: 275–80.

11 Hendriksen JMT, Lucassen WAM, Erkens PMG, Stoffers HEJH, van Weert HCPM, Büller HR, Hoes AW, Moons KGM, Geersing G-J. Ruling Out Pulmonary Embolism in Primary Care: Comparison of the Diagnostic Performance of "Gestalt" and the Wells Rule. *Ann Fam Med* American Academy of Family Physicians; 2016; 14: 227–34.

12 Barais M, Morio N, Cuzon Breton A, Barraine P, Calvez A, Stolper E, Van Royen P, Liétard C. "I can't find anything wrong: It must be a pulmonary embolism": Diagnosing suspected pulmonary embolism in primary care, a qualitative study. *PLoS One* 2014; 9: 1–8.

13 Rodger MA, Maser E, Stiell I, Howley HEA, Wells PS. The interobserver reliability of pretest probability assessment in patients with suspected pulmonary embolism. *Thromb Res* Pergamon; 2005; 116: 101–7.

14 Kabrhel C, Camargo CA, Goldhaber SZ. Clinical gestalt and the diagnosis of pulmonary embolism: Does experience matter? *Chest* The American College of Chest Physicians; 2005; 127: 1627–30.

15 Geersing G-J, Kraaijpoel N, Büller HR, van Doorn S, van Es N, Le Gal G, Huisman M V., Kearon C, Kline JA, Moons KGM, Miniati M, Righini M, Roy P-M, van der Wall SJ, Wells PS, Klok FA. Ruling out pulmonary embolism across different subgroups of patients and healthcare settings: protocol for a systematic review and individual patient data meta-analysis (IPDMA). *Diagnostic Progn Res* Diagnostic and Prognostic Research; 2018; 2: 1–8.

16 Stals MAM, Takada T, Kraaijpoel N, van Es N, Büller HR, Courtney DM, Freund Y, Galipienzo J, Le Gal G, Ghanima W, Huisman M V., Kline JA, Moons KGM, Parpia S, Perrier A, Righini M, Robert-Ebadi H, Roy P-M, van Smeden M, Wells PS, et al. Safety and Efficiency of Diagnostic Strategies for Ruling Out Pulmonary Embolism in Clinically Relevant Patient Subgroups. *Ann Intern Med* 2021; .

17 Geersing G-J, Takada T, Klok FA, Büller HR, Courtney DM, Freund Y, Galipienzo J, Le Gal G, Ghanima W, Kline JA, Huisman M V., Moons KGM, Perrier A, Parpia S, Robert-Ebadi H, Righini M, Roy P-M, van Smeden M, Stals MAM, Wells PS, et al. Ruling out pulmonary embolism across different healthcare settings: A systematic review and individual patient data meta-analysis. *PLOS Med* 2022; 19: e1003905.

18 Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF. Preferred reporting items for a systematic review and meta-analysis of individual participant data: The PRISMA-IPD statement. *JAMA - J Am Med Assoc* 2015; 313: 1657–65.

19 McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, Cohen JF, Deeks JJ, Gatsonis C, Hooft L, Hunt HA, Hyde CJ, Korevaar DA, Leeflang MMG, Macaskill P, Reitsma JB, Rodin R, Rutjes AWS, Salameh JP, Stevens A, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies The PRISMA-DTA Statement. *JAMA - J Am Med Assoc* 2018; 319: 388–96.

20 Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MMG, Sterne JAC, Bossuyt PMM, Group* and the Q-2. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011; 155: 529–36.

21 van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011; 45: 1–67.

22 Janssen KJM, Donders ART, Harrell FE, Vergouwe Y, Chen Q, Grobbee DE, Moons KGM. Missing covariate data in medical research: To impute is better than to ignore. *J Clin Epidemiol* Elsevier Inc; 2010; 63: 721–7.

23 Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009; 9: 57.

24 Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* United States; 2005; 58: 982–90.

25 Pang M, Kaufman JS, Platt RW. Studying noncollapsibility of the odds ratio with marginal structural and logistic regression models. *Stat Methods Med Res* England; 2016; 25: 1925–37.

26 Debray TPA, Moons KGM, van Valkenhoef G, Efthimiou O, Hummel N, Groenwold RHH, Reitsma JB. Get real in individual participant data (IPD) meta-analysis: A review of the methodology. *Res Synth Methods* 2015; 6: 293–309.

27 Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. *J Stat Softw* 2010; 36: 1–48.

28 Kline JA, Nelson RD, Jackson RE, Courtney DM. Criteria for the safe use of D-dimer testing in emergency department patients with suspected pulmonary embolism: a multicenter US study. *Ann Emerg Med* United States; 2002; 39: 144–52.

29 Wicki J, Perneger T V, Junod AF, Bounameaux H, Perrier A. Assessing clinical probability of pulmonary embolism in the emergency ward: a simple score. *Arch Intern Med* United States; 2001; 161: 92–7.

30 Ghanima W, Almaas V, Aballi S, Dörje C, Nielssen BE, Holmen LO, Almaas R, Abdelnoor M, Sandset PM. Management of suspected pulmonary embolism (PE) by D-dimer and multi-slice computed tomography in outpatients: an outcome study. *J Thromb Haemost* England; 2005; 3: 1926–32.

31 Kline JA, Runyon MS, Webb WB, Jones AE, Mitchell AM. Prospective study of the diagnostic accuracy of the simplify D-dimer assay for pulmonary embolism in emergency department patients. *Chest* United States; 2006; 129: 1417–23.

32 Kline JA, Courtney DM, Kabrhel C, Moore CL, Smithline HA, Plewa MC, Richman PB, O'Neil BJ, Nordenholz K. Prospective multicenter evaluation of the pulmonary embolism rule-out criteria. *J Thromb Haemost* England; 2008; 6: 772–80.

33 Kline JA, Hogg MM, Courtney DM, Miller CD, Jones AE, Smithline HA. D-dimer threshold increase with pretest probability unlikely for pulmonary embolism to decrease unnecessary computerized tomographic pulmonary angiography. *J Thromb Haemost* 2012; 10: 572–81.

34 Runyon MS, Beam DM, King MC, Lipford EH, Kline JA. Comparison of the Simplify D-dimer assay performed at the bedside with a laboratory-based quantitative D-dimer assay for the diagnosis of pulmonary embolism in a low prevalence emergency department population. *Emerg Med J* England; 2008; 25: 70–5.

35 Sanson BJ, Lijmer JG, Mac Gillavry MR, Turkstra F, Prins MH, Büller HR. Comparison of a clinical probability estimate and two clinical models in patients with suspected pulmonary embolism. ANTELOPE-Study Group. *Thromb Haemost* Germany; 2000; 83: 199–203.

36 Perrier A, Roy P-M, Aujesky D, Chagnon I, Howarth N, Gourdier A-L, Leftheriotis G, Barghouth G, Cornuz J, Hayoz D, Bounameaux H. Diagnosing pulmonary embolism in outpatients with clinical assessment, D-Dimer measurement, venous ultrasound, and helical computed tomography: a multicenter management study. *Am J Med* United States; 2004; 116: 291–9.

37 Perrier A, Roy P-M, Sanchez O, Le Gal G, Meyer G, Gourdier A-L, Furber A, Revel M-P, Howarth N, Davido A, Bounameaux H. Multidetector-row computed tomography in suspected pulmonary embolism. *N Engl J Med* United States; 2005; 352: 1760–8.

38 Kearon C, Ginsberg JS, Douketis J, Turpie AG, Bates SM, Lee AY, Crowther MA, Weitz JI, Brill-Edwards P, Wells P, Anderson DR, Kovacs MJ, Linkins L-A, Julian JA, Bonilla LR, Gent M. An evaluation of D-dimer in the diagnosis of pulmonary embolism: a randomized trial. *Ann Intern Med* United States; 2006; 144: 812–21.

39 van Belle A, Büller HR, Huisman M V, Huisman PM, Kaasjager K, Kamphuisen PW, Kramer MHH, Kruip MJHA, Kwakkel-van Erp JM, Leebeek FWG, Nijkeuter M, Prins MH, Sohne M, Tick LW. Effectiveness of managing suspected pulmonary embolism using an algorithm combining clinical probability, D-dimer testing, and computed tomography. *JAMA* United States; 2006; 295: 172–9.

40 Goekoop RJ, Steeghs N, Niessen RWLM, Jonkers GJPM, Dik H, Castel A, Werker-van Gelder L, Vlasveld LT, van Klink RCJ, Planken E V, Huisman M V. Simple and safe exclusion of pulmonary embolism in outpatients using quantitative D-dimer and Wells' simplified decision rule. *Thromb Haemost* Germany; 2007; 97: 146–50.

41 Righini M, Le Gal G, Aujesky D, Roy P-M, Sanchez O, Verschuren F, Rutschmann O, Nonent M, Cornuz J, Thys F, Le Manach CP, Revel M-P, Poletti P-A, Meyer G, Mottier D, Perneger T, Bounameaux H, Perrier A. Diagnosis of pulmonary embolism by multidetector CT alone or combined with venous ultrasonography of the leg: a randomised non-inferiority trial. *Lancet* England; 2008; 371: 1343–52.

42 Douma RA, Mos ICM, Erkens PMG, Nizet TAC, Durian MF, Hovens MM, van Houten AA, Hofstee HMA, Klok FA, ten Cate H, Ullmann EF, Büller HR, Kamphuisen PW, Huisman M V. Performance of 4 clinical decision rules in the diagnostic management of acute pulmonary embolism: a prospective cohort study. *Ann Intern Med* United States; 2011; 154: 709–18.

43 Galipienzo J, Garcia de Tena J, Flores J, Alvarez C, Garcia-Avello A, Arribas I. Effectiveness of a diagnostic algorithm combining clinical probability, D-dimer testing, and computed tomography in patients with suspected pulmonary embolism in an emergency department. *Rom J Intern Med* Germany; 2012; 50: 195–202.

44 Schouten HJ, Geersing GJ, Oudega R, Van Delden JJM, Moons KGM, Koek HL. Accuracy of the Wells Clinical Prediction Rule for Pulmonary Embolism in Older Ambulatory Adults. *J Am Geriatr Soc* 2014; 62: 2136–41.

45 Righini M, Van Es J, Den Exter PL, Roy PM, Verschuren F, Ghuysen A, Rutschmann OT, Sanchez O, Jaffrelot M, Trinh-Duc A, Le Gall C, Moustafa F, Principe A, Van Houten AA, Ten Wolde M, Douma RA, Hazelaar G, Erkens PMG, Van Kralingen KW, Grootenboers MJJH, et al. Age-adjusted D-dimer cutoff levels to rule out pulmonary embolism: The ADJUST-PE study. *JAMA - J Am Med Assoc* 2014; 311: 1117–24.

46 Mos ICM, Douma RA, Erkens PMG, Kruip MJHA, Hovens MM, van Houten AA, Hofstee HMA, Kooiman J, Klok FA, Büller HR, Kamphuisen PW, Huisman M V. Diagnostic outcome management study in patients with clinically suspected recurrent acute pulmonary embolism with a structured algorithm. *Thromb Res* United States; 2014; 133: 1039–44.

47 Penaloza A, Soulié C, Moumneh T, Delmez Q, Ghuysen A, El Kouri D, Brice C, Marjanovic NS, Bouget J, Moustafa F, Trinh-Duc A, Le Gall C, Imsaad L, Chrétien J-M, Gable B, Girard P, Sanchez O, Schmidt J, Le Gal G, Meyer G, et al. Pulmonary embolism rule-out criteria (PERC) rule in European patients with low implicit clinical probability (PERCEPIC): a multicentre, prospective, observational study. *Lancet Haematol* England; 2017; 4: e615–21.

48 Kearon C, de Wit K, Parpia S, Schulman S, Afilalo M, Hirsch A, Spencer FA, Sharma S, D'Aragon F, Deshaies J-F, Le Gal G, Lazo-Langner A, Wu C, Rudd-Scott L, Bates SM, Julian JA. Diagnosis of Pulmonary Embolism with <scp>d</scp> -Dimer Adjusted to Clinical Probability. *N Engl J Med* 2019; 381: 2125–34.

49 Stolper E, Van de Wiel M, Van Royen P, Van Bokhoven M, Van der Weijden T, Dinant GJ. Gut feelings as a third track in general practitioners' diagnostic reasoning. *J Gen Intern Med* 2011; 26: 197–203.

50 Barais M, Fossard E, Dany A, Montier T, Stolper E, Van Royen P. Accuracy of the general practitioner's sense of alarm when confronted with dyspnoea and/or chest pain: a prospective observational study. *BMJ Open* 2020; 10: e034348.

51 Righini M, Le Gal G, Perrier A, Bounameaux H. The challenge of diagnosing pulmonary embolism in elderly patients: influence of age on commonly used diagnostic tests and strategies. *J Am Geriatr Soc* United States; 2005; 53: 1039–45.

52 Engbers MJ, van Hylckama Vlieg A, Rosendaal FR. Venous thrombosis in the elderly: incidence, risk factors and risk groups. *J Thromb Haemost* England; 2010; 8: 2105–12.

Authorship statement

I contributed to defining the research question, proposed the experimental design, performed the data analysis together with a statistician/methodologist, and wrote the first draft of the manuscript. During the whole process, I asked for and implemented input and feedback from the other contributors to this study.

SUPPLEMENTAL FILES

Supplement 1. Proportion of missing data in each study included in the IPD-MA
Supplement 2. Risk-of-bias assessment among included studies using the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies 2) tool.
Supplement 3. Sensitivity and specificity of gestalt and PE pooled per subgroup.
Supplement 4. Plot of risk ratio of PE and gestalt against the prevalence of PE per study

SUPPLEMENT 1. Proportion of missing data in each study included in the IPD-MA

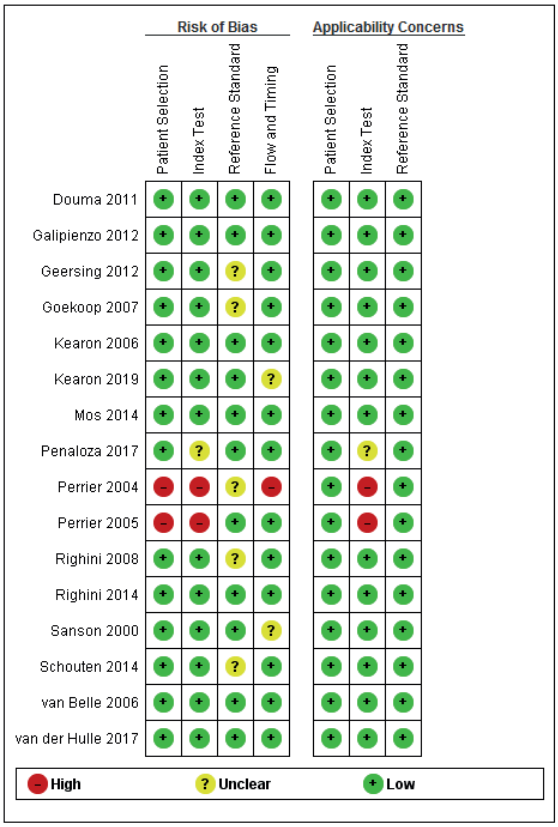
	Sanson Thromb Haemost 2000	Perrier Am J Med 2004	Perrier N Engl J Med 2005	Kearon Ann Intern Med 2006	van Belle JAMA 2006	Goekoop Thromb Haemost 2007	Righini Lancet 2008	Douma Ann Itern Med 2011	Galipienzo Rom J Intern Med 2012	Geersing BMJ 2012	Mos Thromb Res 2014
N	517	965	755	1123	3296	876	1692	807	240	597	279
Age	0	0	0	0	0	0	0	0	0	0	0
Female sex	0	0	0	0	0.3	0	0	0	0	0	0.4
Inpatient	0	0	0	0	0	0	0	0	0	0	2.9
Previous VTE	0.8	0.2	0	0	0.2	0.2	0	0	0	0	0
Heart rate > 100/min	1.2	0.4	0.4	0	0.2	0.1	0.1	0	0	0	61.3
Surgery or immobilization < 4 weeks	0.6	0	0	0	0.2	0	0	0	0	0	24.4
Haemoptysis	0.8	0	0	0	0.2	0.2	0	0	0	0	26.2
Active cancer	6.4	0.3	0	0	0.2	0.1	0	0	0	0	20.8
Clinical signs of DVT	0.2	0	0.1	0	0.2	0.1	0.1	0	0	0	20.8
Alternative diagnosis less likely than PE	0.4	3.5	1.1	0	0.2	0	2.3	0	0	0	17.9
Heart failure	0	0	0	100	0.5	1.4	13.6	2.5	100	0	9.9
Chronic lung disease	0	0	0	100	0.5	1.4	0	1.2	100	0	11.3
Qualitative D-dimer	3.9	0	0.1	0.4	100	100	0.5	100	100	0	100
Quantitative D-dimer	100	0.1	1.2	100	15.6	8.9	0.5	7.4	2.5	100	27.2
PE	0	0	0	0	0	0	0	0	0	0	0

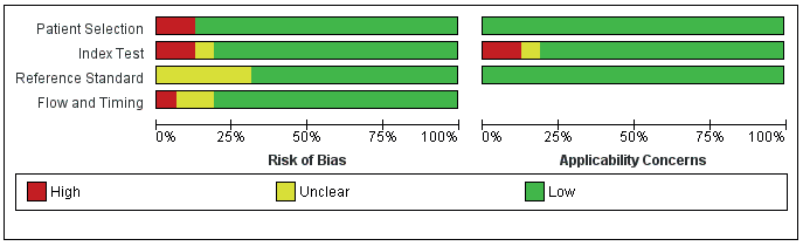
SUPPLEMENT 1. (Continued)

	Righini JAMA 2014	Schouten J Am Geriatr Soc 2014	Penaloza Lancet Haematol 2017	van der Hulle Lancet 2017	Kearon N Engl J Med 2019
N	3324	129	705	3448	2017
Age	0	0	0	0	0
Female sex	0	0	0	0	0
Inpatient	0	0	0	0	0
Previous VTE	0	1.6	0	0.1	0
Tachycardia	4.9	3.9	0	2	0
Surgery or immobilization < 4 weeks	0	0	0	0.1	0
Haemoptysis	0.1	1.6	0	0	0
Active cancer	0	0.8	0	0.1	0
Clinical signs of DVT	2.7	0	0	0	0
Alternative diagnosis less likely than PE	0	0	0	0	0
Heart failure	3.4	1.0	3.1	0.03	100
Chronic lung disease	0	1.7	100	0	100
Qualitative D-dimer	100	7	100	100	100
Quantitative D-dimer	7.3	61.2	13.2	0.2	0.6
PE	0	0.8	0	0	0

Abbreviations: N, number of patients; VTE, venous thromboembolism; DVT, deep vein thrombosis; PE, pulmonary embolism

SUPPLEMENT 2. Risk-of-bias assessment among included studies using the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies 2) tool.

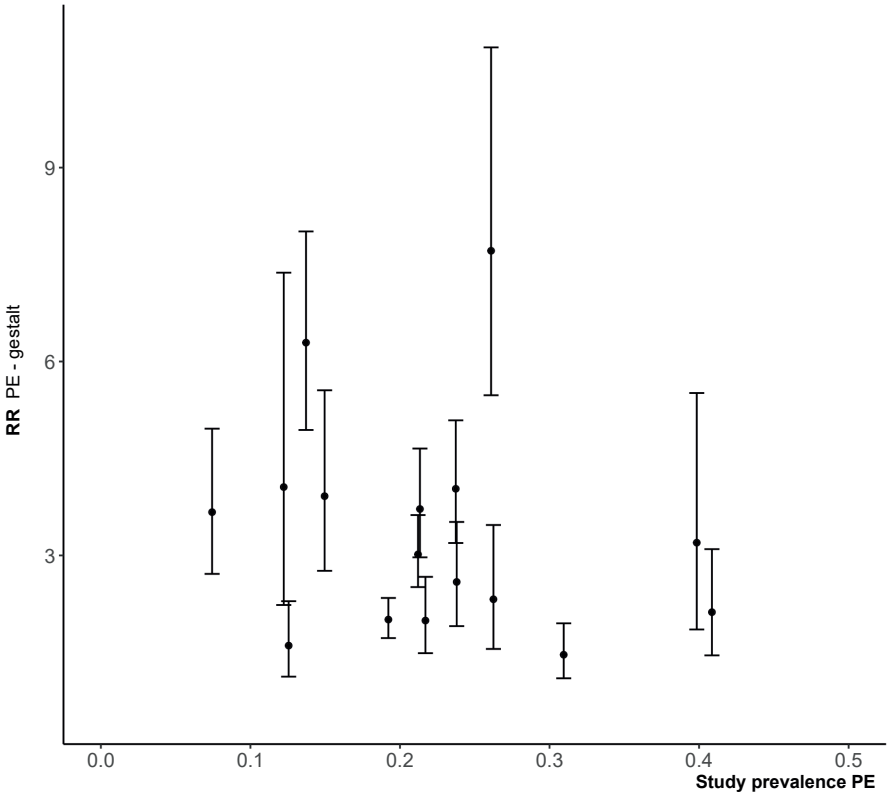




SUPPLEMENT 3. Sensitivity and specificity of gestalt and PE pooled per subgroup.

Subgroup	Sensitivity (95% CI)	Specificity (95% CI)
Male	0.74 (0.68 to 0.79)	0.61 (0.53 to 0.68)
Female	0.70 (0.63 to 0.76)	0.63 (0.55 to 0.70)
No heart failure	0.62 (0.53 to 0.71)	0.64 (0.54 to 0.72)
Heart failure	0.76 (0.71 to 0.80)	0.58 (0.51 to 0.64)
No chronic lung disease	0.61 (0.55 to 0.67)	0.66 (0.54 to 0.75)
Chronic lung disease	0.77 (0.72 to 0.81)	0.57 (0.49 to 0.64)
Primary healthcare	0.79 (0.71 to 0.86)	0.62 (0.57 to 0.66)
Emergency ward or hospital care	0.72 (0.65 to 0.78)	0.61 (0.52 to 0.69)
Hospital or nursing home	0.73 (0.53 to 0.86)	0.64 (0.45 to 0.79)
Before 2010	0.70 (0.60 to 0.79)	0.62 (0.52 to 0.71)
2010 and later	0.74 (0.67 to 0.80)	0.61 (0.50 to 0.71)
Clinical signs and symptoms of DVT	0.80 (0.73 to 0.85)	0.50 (0.43 to 0.57)
No clinical signs and symptoms of DVT	0.73 (0.67 to 0.78)	0.60 (0.53 to 0.67)
Immobilization or surgery	0.74 (0.68 to 0.79)	0.58 (0.49 to 0.65)
No immobilization or surgery	0.72 (0.66 to 0.78)	0.62 (0.55 to 0.69)
Previous VTE	0.75 (0.69 to 0.80)	0.55 (0.50 to 0.60)
No previous VTE	0.71 (0.64 to 0.77)	0.63 (0.55 to 0.71)
Heart rate ≥ 100 bpm	0.75 (0.68 to 0.81)	0.63 (0.56 to 0.70)
Heart rate < 100 bpm	0.71 (0.65 to 0.77)	0.61 (0.54 to 0.69)
Haemoptysis	0.73 (0.64 to 0.81)	0.62 (0.52 to 0.71)
No haemoptysis	0.71 (0.64 to 0.77)	0.62 (0.54 to 0.69)
Cancer	0.75 (0.68 to 0.80)	0.56 (0.48 to 0.64)
No cancer	0.72 (0.65 to 0.77)	0.62 (0.55 to 0.69)
Two or more Wells items	0.76 (0.70 to 0.81)	0.55 (0.48 to 0.61)
One Wells item	0.73 (0.66 to 0.79)	0.61 (0.53 to 0.68)
No Wells items	0.65 (0.57 to 0.72)	0.65 (0.57 to 0.73)
Overall	**0.74 (0.68 to 0.79)**	**0.61 (0.53 to 0.68)**

SUPPLEMENT 4. Plot of risk ratio of PE and gestalt against the prevalence of PE per study.



Validation and impact of a simplified clinical decision rule for diagnosing pulmonary embolism in primary care: design of the PECAN prospective diagnostic cohort management study

Rosanne van Maanen
Frans H. Rutten
Frederikus A. Klok
Menno V. Huisman
Jeanet W. Blom
Karel G.M. Moons
Geert-Jan Geersing

BMJ Open 2019; 9:e031639

ABSTRACT

Introduction Combined with patient history and physical examination, a negative D-dimer can safely rule-out pulmonary embolism (PE). However, the D-dimer test is frequently false positive, leading to many (with hindsight) 'unneeded' referrals to secondary care. Recently, the novel YEARS algorithm, incorporating flexible D-dimer thresholds depending on pre-test risk, was developed and validated, showing its ability to safely exclude PE in the hospital environment. Importantly, this was accompanied with 14% fewer computed tomographic pulmonary angiography than the standard, fixed D-dimer threshold. Albeit promising, in primary care this algorithm has not been validated yet.

Methods and analysis The PECAN (Diagnosing Pulmonary Embolism in the context of Common Alternative diagNoses in primary care) study is a prospective diagnostic study performed in Dutch primary care. Included patients with suspected acute PE will be managed by their general practitioner according to the YEARS diagnostic algorithm and followed-up in primary care for 3 months to establish the final diagnosis. To study the impact of the use of the YEARS algorithm, the primary endpoints are the safety and efficiency of the YEARS algorithm in primary care. Safety is defined as the proportion of false-negative test results in those not referred. Efficiency denotes the proportion of patients classified in this non-referred category. Additionally, we quantify whether C-reactive protein measurement has added diagnostic value to the YEARS algorithm, using multivariable logistic and polytomous regression modelling. Furthermore, we will investigate which factors contribute to the subjective YEARS item "PE most likely diagnosis".

Ethics and dissemination The study protocol was approved by the Medical Ethical Committee Utrecht, the Netherlands. Patients eligible for inclusion will be asked for their consent. Results will be disseminated by publication in peer-reviewed journals and presented at (inter)national meetings and congresses.

Trial registration The PECAN study is registered at the Netherlands Trial Register (NTR 7431).

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This is the first study that prospectively validates the YEARS algorithm for diagnosing pulmonary embolism (PE) in primary care.
- The added diagnostic value of C-reactive protein will be quantified by developing a double biomarker approach with the aim to better differentiate between PE and pneumonia.
- Because this study does not randomise general practices between the current guidelines and the YEARS algorithm, results are compared with existing literature rather than a direct comparison.
- A possible limitation of this study is that different reference tests will be used to determine the presence of PE, which might lead to differential verification bias.

INTRODUCTION

Diagnosing acute pulmonary embolism (PE) is challenging, particularly in primary care. Signs and symptoms are often non-specific and may mimic other cardiopulmonary diseases.[1–3] D-dimer is used as a biomarker to disentangle PE from such other conditions. Yet, notably in patients with other cardiopulmonary diseases, this is far less efficient, with a chance of a false-positive D-dimer reaching 90% in older patients with cardiopulmonary comorbidity.[3,4] Subsequently, many patients suspected of PE are referred for reference testing (computed tomography pulmonary angiography; CTPA) whereas only 10-15% will have a confirmation of the diagnosis.[5] Moreover, CTPA has the inherent risk of contrast nephropathy, that may also occur in up to 10-15% of all CTPAs performed, depending on pre-existing renal impairment.[6] Besides, CTPA is costly and patients are exposed to ionizing radiation which may increase their cancer risk later in life.[7] While most evidence on use of diagnostic tests for suspected PE has been gained in the hospital setting, patients often visit their general practitioner (GP) first. The few studies that focussed on the use of CDR and D-dimer testing in primary care observed that both deep venous thrombosis (DVT) and PE could be safely ruled out in almost 50% of suspected patients. However, of the referred patients only 25% and 30% respectively had a confirmed DVT or PE, a number that is much lower in the elderly.[8–10] Importantly, PE is also

still one of the most frequently missed diagnoses in primary care, underlining the need for improvement of the diagnostic algorithm.[11]

Recent studies in secondary care developed and validated a new algorithm with flexible D-dimer thresholds: the YEARS algorithm.[12–14] This strategy starts with assessing three patient history and physical examination items: (i) clinical signs of DVT, (ii) haemoptysis and (iii) PE considered the most likely diagnosis by the physician. At the same time, D-dimer testing is performed in all patients. If none of the three YEARS items are present, a D-dimer threshold of 1000 ng/ml is applied. In contrast, if one or more YEARS items are present, the conventional threshold of 500 ng/ml is used. If the D-dimer is below the relevant threshold, PE is considered ruled out and patients are not referred for CPTA. As compared to applying a fixed D-dimer threshold of 500 ng/ml, this YEARS algorithm increased the proportion of patients in whom CTPA was not required from 34% to 48%. Importantly, refraining from referral for CTPA was safe with a 3 months failure rate in patients with initial normal tests of 0.61% (95% CI 0.36 – 0.96).[13] Given the substantially lower prior probability of PE in primary care, we hypothesize that the YEARS algorithm can also be safely used when used in the primary care setting. However, considering differences in case-mix of patients and physician experience, prospective validation and impact assessment of the use of the YEARS algorithm applied by GPs is necessary before its wide-scale use in primary care can be recommended. Therefore, the aim of this study is to prospectively validate the YEARS algorithm for ruling-out PE in primary care, with the use of a point-of-care (POC) or rapid D-dimer assay.

OBJECTIVES

The objectives of this study are threefold. The primary objective is to prospectively validate the YEARS algorithm in primary care. We will calculate its calibration (observed versus expected probabilities) and its discriminative ability. Using the previously proposed decision threshold we will also estimate the impact of the use of the YEARS algorithm by estimating its safety (defined as the proportion of missed PE cases in the group of patients not referred for CTPA), and its efficiency (defined as the proportion of patients correctly not referred for CTPA).

A secondary objective is to quantify the added diagnostic value of performing CRP measurement in patients with suspected PE in primary care. GPs seldom only consider a single diagnosis in patients presenting with (sub)acute shortness of breath, and pneumonia is an important alternative diagnosis in these patients.[1,15] Hence, a combined D-dimer and CRP biomarker approach may lead to a better classification of underlying causes of respiratory and/or chest symptoms, and thus better exclusion of PE.

Another secondary objective is to investigate which determinants contribute to a 'yes' answer on the YEARS item "PE most likely diagnosis". We hypothesize that scoring of this item is correlated with several patient-related and physician-related factors. Scoring of this YEARS item may differ in primary and secondary care physicians and may therefore influence the use and interpretation of the YEARS algorithm.

METHODS AND ANALYSIS

Study design

We will perform a multicentre, prospective diagnostic cohort study in the primary care setting including patients with suspected acute PE, defined as (sub)acute onset of unexplained shortness of breath with or without chest symptoms, such as thoracic pain or pain on inspiration. Patients will be managed according to the YEARS algorithm and will be followed-up for 3 months, with an uneventful follow-up period being the diagnostic standard for ruling out PE. The inclusion period for recruiting patients into the study is estimated at 2-3 years.

Clinical setting and participants

Our study will be conducted within the Dutch primary care setting. Patients will be recruited by their GP both during working hours and out-of-hours primary care services. Participating GPs will identify eligible patients and obtain informed consent. Consecutive patients with suspected PE are eligible for inclusion if they are aged 18 years or older and provide written informed consent. Exclusion criteria will be current treatment with therapeutic doses of vitamin K antagonists, low-molecular weight heparin or a direct oral anticoagulant, life expectancy less than 1 month estimated by the GP and pregnancy until 6 weeks after delivery.

Study procedures

The study procedures are shown in figure 1. Patients who are visiting their GP because of suspected PE will be asked for consent for participation and data collection. Next, the GP scores the three YEARS items, age, sex, signs, symptoms, co-morbidity items and performs a quantitative POC or rapid D-dimer test. In this study, different D-dimer assays are used, namely the quantitative assay for point-of-care testing and several different D-dimer assays in the laboratories in the participating regions. When a patient has no YEARS items and a D-dimer below the threshold of 1000 ng/mL, or when a patient has one or more YEARS items and a D-dimer below the threshold of 500 ng/mL, PE will be considered as ruled out. However, referral for other reasons, such as a severe pneumonia, remains appropriate and left at the discretion of the attending GP. In patients with no YEARS items and a D-dimer above 1000 ng/mL, or in patients with one or more YEARS items present and a D-dimer above 500 ng/mL, a referral to secondary care for CTPA will follow. Deviation from the YEARS recommendation is allowed for, if judged clinically needed and as such left at the discretion of the physician. Patients will not be followed up in person, but will be instructed to schedule a follow-up appointment in case of worsening or persistent

symptoms. If the patient has symptoms of PE during this follow-up appointment, the management decision is made based on the standard guidelines and discretion of the physician.

For our secondary objective, additional blood will be drawn to determine CRP during the initial visit with the same venepuncture on the POC or rapid assay. Yet, CRP is only determined for research proposes, i.e. one of the secondary objectives, and therefore GPs are instructed to use the YEARS algorithm primarily for further clinical management. Furthermore, several determinants who may have an association with the YEARS item "PE most likely diagnosis" are collected and reported as such by the participating GP, see table 1.

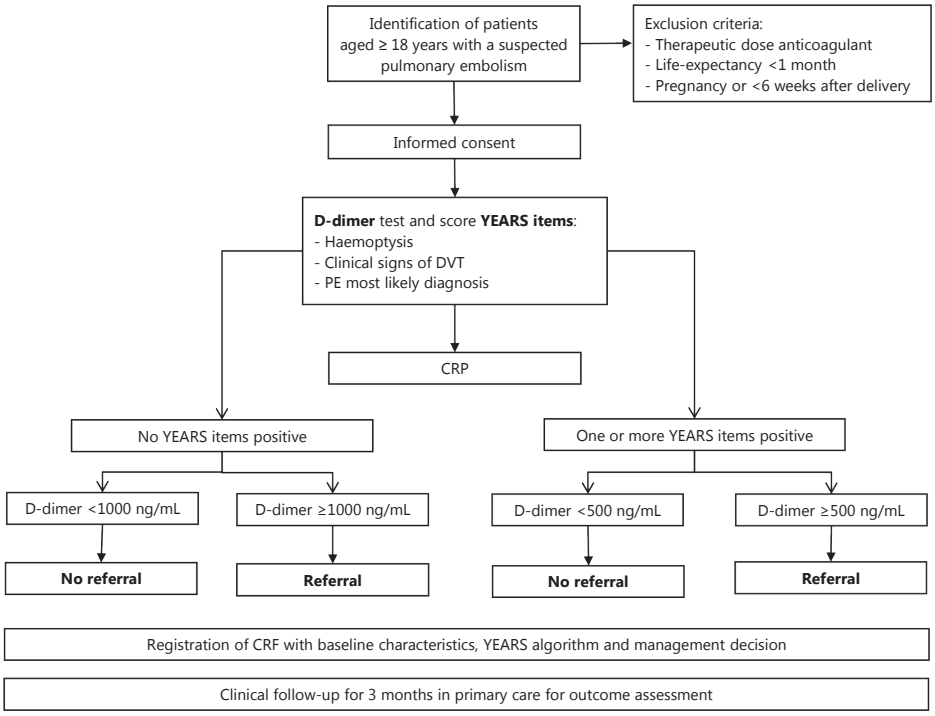


Figure 1. Flowchart study procedures.

DVT: deep venous thrombosis; CRP: C-reactive protein; CRF: case report form.

Table 1. Questions and categorical responses associated with the subjective YEARS item

Questions/ determinants	Categorical responses
Have you ever missed a diagnosis of a pulmonary embolism in your practice?	Yes/ No
Did you have a 'gut feeling' that there was something wrong?	Yes/ No
Does it concern a well-known patient?	Yes/ No
Was there diagnostic delay?	Yes/ No *If yes:* physician delay, patient delay or both
How was the working load the day of consultation?	Low/ medium/ high
What is your implicit probability assessment of pulmonary embolism?	Unlikely, likely, very likely

Reference standard

After 3 months of follow-up, the GP receives a follow-up form with questions about the final diagnosis and treatment. This final diagnosis is the reference standard for this study, similarly as done in previous studies performed in the field of diagnosing PE.[3,13,16,17] In the follow-up form, it is asked whether PE is proven – i.e. by CTPA, VQ scan, ultrasonography showing a DVT, or a combination of these imaging procedures, all as performed and classified by hospital physicians after referral – and (finally) whether there were alternative diagnoses present. Also, the treatment decision (anticoagulation or treatment for alternative diagnosis) is recorded by the GP. Here, it is important to acknowledge that in the Dutch healthcare system GPs are always fully informed of changes in patient status, also from affiliating hospitals, exemplifying the fidelity of this approach which we also found highly reliable in previous diagnostic VTE studies from our group.[8,9]

A diagnosis of PE presence is considered definitive as (i) CTPA demonstrating a filling defect in a central, segmental or lobular pulmonary artery, or a sub-segmental filling defect which requires anticoagulant therapy, or (ii) a high probability ventilation/ perfusion lungscan, or (iii) a pulmonary angiogram demonstrating an intraluminal filling defect, or (iv) PE demonstrated at autopsy in case of death, or (v) DVT confirmed with ultrasonography of the leg in patients with suspected PE. Importantly, PE is considered ruled-out in the absence of any PE-defining items as described above during the initial clinical assessment plus 3 months of uneventful follow-up. In case of absence of PE, the GP will fill in the most likely alternative diagnosis on the follow-up form, including pneumonia, heart failure, chronic obstructive pulmonary

disease or asthma, cardiac disease and myalgia, based on specialist letters and the clinical judgement and management decisions made by the GP, as explained above. All patients with an unexpected death during follow-up will be adjudicated for the presence of possible PE as the cause of death, following definition from the on-going work on defining PE-related death from the International Society on Thrombosis and Haemostasis (ISTH) which will be available before the study is completed and endpoints are adjudicated.

Sample size calculation

According to previous studies performed in primary care, the prevalence of PE in suspected patients with a low PE probability (indicating no referral for CTPA) based on the used decision rule plus negative D-dimer, was 1.0% to 1.5%.[9] Although recently the ISTH proposes a variable diagnostic safety threshold with adjustment for the prevalence of PE in the study population [18], 3% is internationally deemed as an acceptable safety margin of missed PE in the low probability (i.e. non-referred) patients, so-called false negatives, and is widely used in diagnostic studies of PE.[8,9,17,19] Hence, we also use 3% as the upper margin of the 95% CI around the point estimate of our false-negative rate. Assuming a conservative false negative rate of 1.5% with the upper margin of the 95% CI not exceeding 3.0% (one-sided, as any proportion lower than 1.5% is preferable), we need to include 300 patients in the low probability group according to the YEARS algorithm. It should be stressed here that this point-estimate of 1.5% is highly conservative as well, given that in the validation study of the YEARS algorithm in secondary care the point estimate was 0.67%. Moreover, the proportion of patients classified in the non-referred category was 48%[13]. Yet our study is conducted in primary care, with a lower overall PE prevalence of around 12% (table 2). We therefore anticipate that the proportion of patients in the low probability group will be at least as high. Thus, to arrive at 300 patients with a low YEARS probability (i.e. a negative YEARS algorithm), we need to include at least 600 primary care patients suspected of PE. Accounting for 10% of patients with missing follow-up information, we conservatively target to include 750 patients for this study of which at least 300 patients with a negative YEARS algorithm. This full sample of 750 patients would allow us to robustly demonstrate (or reject) the safety of applying the YEARS algorithm in primary care. To arrive at a total study population of 750 patients, we estimated to need approximately 75 full-time working GPs who will include 5 patients per year in the study period of 2 years.

Table 2. Description of the different prospectively validated clinical decision rules for pulmonary embolism.[9,13,17]

Study characteristics			Results		Diagnostic accuracy	
Year	**Clinical decision rule**	**Population**	**Sample size**	**Prevalence PE**	**Safety ***	**Efficiency†**
2012	Wells rule	Primary care	598	12.2%	1.5%	45.5%
2014	Age-adjusted D-dimer threshold	Secondary care	3346	19.0%	0.6%	39.8%
2017	YEARS algorithm	Secondary care	3465	13.2%	0.5%	48.0%

PE=pulmonary embolism.
** Proportion of false-negatives among patients not referred at baseline.*
† Proportion of patients not referred at baseline among all included patients

Data analysis

To quantify the diagnostic accuracy of the YEARS algorithm, we will estimate its discrimination (using the c-statistic) and its calibration (using the calibration plot comparing predicted probability with observed probability). Hereto, the linear predictor of the YEARS algorithm first needs to be estimated for each included patient into our study, using the original regression coefficients of the YEARS items as derived in the original derivation paper, with the intercept refitted to the prospective data as included in our primary care cohort (to best reflect differences between overall PE prevalence across populations).[12] Subsequently, by applying the previously proposed YEARS algorithm threshold (see above), we will also estimate the corresponding false negative proportion and efficiency of the strategy. Besides, results will be stratified for each assay specific, where deemed appropriate and necessary. Lastly, we will perform an additional analysis with and without including the diagnoses of sub-segmental PE because of the clinical unknown significance.[20,21]

Second, we will quantify the added diagnostic value of CRP beyond the YEARS algorithm. Hereto, we will first construct multivariable logistic regression models with PE being the binary outcome and the YEARS items with D-dimer as dependent variables, which model is then extended by addition of CRP. D-dimer and CRP will be included into the model on a continuous scale, if needed using natural cubic splines function if the association between both laboratory markers and pulmonary embolism is non-linear. We will use the likelihood ratio test (using a p-value of 0.15) to quantify the added contribution of CRP. Similarly as above, we will also quantify the calibration

and discrimination of both models using bootstrapping techniques to correct for overfitting.[22]

Additionally we will quantify to what extent the YEARS algorithm with D-dimer and CRP can predict the presence of the differential diagnoses simultaneously, using polytomous regression modelling.[23] The differential diagnosis is hereto divided in three categories: PE, pneumonia and other. The final polytomous regression model will consist of those two sub-models and allows one to estimate the probability of presence of PE, pneumonia and other diagnoses in each patient.

For our third objective, the prevalence of the subjective YEARS item "PE most likely diagnosis" will first be described with a corresponding 95% CI. Then, variables among patients with and without a positive score on this YEARS item will be compared. The variables that will be investigated are described in table 1. These variables are first compared univariably, and then combined in a multivariable logistic regression analysis with the item "PE most likely diagnosis" as a dichotomous outcome. Some of these quantitative data will later be used to complement with qualitative data obtained from another future study entailing interviews with GPs, during a mixed method analysis.

Safety interim analysis

After the first 100 included patients with a negative YEARS algorithm a safety analysis will be performed. Based on previous studies, the expected percentage false-negatives (i.e. patients with PE in the low probability category of the YEARS algorithm) should at least not be higher than 1.5%. This will correspond to an expectation of approximately 1 to 2 missed PE cases in the first 100 non-referred patients. If the proportion of false-negatives in these first 100 patients with a negative YEARS algorithm clearly and beyond reasonable doubt is larger than 1.5%, the study will be put 'on hold' pending additional analyses. Although it is difficult to identify when study continuation in such a diagnostic management study is clearly contra-indicated, we arbitrarily use a difference of at least three standard deviations (P-value » 0.002). For this study, that would mean missing no more than arbitrarily 5 to 6 patients in the first 100 patients with a PE in the low probability category. If this safety analysis is satisfactory, the study will continue as planned with additional safety checks alongside patient accrual into the study where deemed appropriate and necessary.

Handling of missing data

In case of missing data, the researchers will first contact the treating physician to retrieve this information. When this is not possible, or information remains missing, we will use multiple imputation techniques to yield unbiased inferences, if the missing at random assumption is likely.[24] We expect to detect missing data which are missing at random, i.e. that the missing data for that subject is based on other observed patient characteristics. Multiple imputation could be reliably used even if 40% of the data of one variable is missing (as shown by a simulation study) which however is unlikely to occur in our study but exemplifies that we anticipate multiple imputation to provide robust results. Then the missing values will be multiple imputed with a conditional imputation method to minimize bias and increase precision.[22,24]

Patient and Public Involvement statement

There are no patients involved in the development of the study design and protocol. However, a patient representative with a strong network within the field of patient advocates gave insight in the patient experience to our study group.

Ethics and dissemination

This study will be conducted according to the principles of the Declaration of Helsinki and in accordance with the Medical Research Involving Human Subjects Act. The Medical Ethical Committee Utrecht in the Netherlands approved the study protocol. Patients eligible for inclusion will be asked for their written consent by participating GPs before the YEARS algorithm is applied. Results of the PECAN study are expected in 2022/2023 and will be disseminated by publication in peer-reviewed journals and presented at (inter)national meetings and congresses.

DISCUSSION

The PECAN study will evaluate the safety and efficiency of implementing the YEARS algorithm for ruling out PE in primary care. Using the three clinical examination items, plus D-dimer testing and subsequently decide whether referral is necessary, is worldwide routine clinical practice and based upon current (inter)national guidelines.[25] Previous research showed that the YEARS algorithm is safe in a secondary care setting with 14% fewer referrals for CTPAs and is now standard-of-care in many emergency wards in the Netherlands.[13] This decrease of CTPAs could be especially useful for primary care medicine, since GPs are the gatekeepers to secondary care and often need to decide whether the patient could be treated in primary care or has to be referred to secondary care. GPs are constantly balancing between over- and under-referral of patients with suspected PE, given the associated harms related to both over- and under-referral. The YEARS algorithm may safely reduce the need for referral for CTPA, which may at least partly alleviate the diagnostic uncertainty and dilemma whether or not to refer a patient to secondary care. Yet, good performance and safety in secondary care based studies is not always a guarantee that the model also performs well in a primary care healthcare setting, due to inherent differences in prevalence, case-mix and physician experience in primary care. This is the primary argument to embark on this prospective diagnostic validation study. Additionally, we will further explore the ability to refine the diagnostic process by incorporating CRP into the diagnostic model, as well as study determinants for the diagnostic item "PE most likely diagnosis".

Limitations

A possible limitation of this study is that, by design, a combined reference standard will be used to determine the presence or absence of PE. The reference standard for the non-referred patients is a clinical follow-up of 3 months, while in the referred patients imaging techniques are used. This combined reference standard may result in differential verification bias.[26] This may lead to biased estimates of the sensitivity and specificity, but gives reliable and clinically interpretable positive and negative predictive values, as the choice of the reference standard is almost – by design - fully dependent on the outcome of the YEARS algorithm. However, we explicitly designed this validation study as a pragmatic study following routine care to evaluate the accuracy and safety of the YEARS algorithm *as would be performed* in real-world daily practice. Using this combined reference standard is compliable with the practical use when implemented as standard-of-care in primary care centres in the future.

Some GPs participating in our study will perform D-dimer on a specific POC-assay. Yet, not all physicians will have access to this specific POC-assay and a substantial proportion may have to determine D-dimer via the laboratory. This could lead to practical issues when applying the YEARS algorithm, since D-dimer needs to be determined in all patients before it can be decided to refer the patient for CTPA or not. So, when a POC-assay for D-dimer is not available, those patients first have to visit a laboratory. However, including patients from general practices with and without a POC-assay for D-dimer will increase the generalizability of our results and is an advantage when implementing this as standard-of-care when proven safe and efficient.

Also, for our secondary objective, a CRP measurement is done in all patients suspected of PE. Although we instruct GPs to only use the YEARS algorithm without formally interpreting the CRP result, we cannot completely rule-out the possibility that this might influence their management decision. However, this is similar as conducting diagnostic VTE studies in an emergency department where often multiple tests are available and interpreted during the diagnostic work-up of PE.[13,14,17]

Lastly, our study does not include a control group because we do not randomise general practices between the current guidelines and the new YEARS algorithm. Therefore, a direct comparison between the YEARS algorithm and usual care will explicitly not be part of this validation study. Rather, results are compared with existing literature, most notably (albeit not exclusively) the recent validation of the YEARS algorithm in secondary care.

Comparison with literature findings

Recently, another strategy with the aim to reduce unnecessary CTPAs has been prospectively validated in secondary care: the age-adjusted D-dimer threshold.[17] Although this strategy alone would result in a larger proportion of patients in whom PE could be considered ruled out, this is only applicable in patients older than 50 years whereas younger patients benefit most of refraining from CTPA given the long-term radiation effects. Also, a comparison of the YEARS algorithm and the age-adjusted D-dimer threshold showed an absolute reduction of 8.7% of CTPAs in favour of the YEARS algorithm.[13] Furthermore, a post-hoc analysis was performed to investigate the added value of the age-adjusted D-dimer threshold when incorporated with the YEARS algorithm. This study showed that in the patients aged above 50 years, the efficiency of the algorithm was increasing with a 4.7% decrease of CTPAs. However, the safety was jeopardized with four additional missed diagnosis of PE resulting in

a failure rate of 1.2%.[27] A summary of the prospective validation studies of the YEARS algorithm and age-adjusted D-dimer threshold in secondary care, as well as the Wells rule in primary care are shown in table 2.[9,13,17]

To conclude, the PECAN study will prospectively validate and quantify the safety of the YEARS algorithm in patients with suspected acute PE in primary care. If proven safe, this new clinical decision rule could then be implemented in daily care. In addition, the diagnostic value of performing CRP measurement in patients with suspected PE will be quantified.

DECLARATIONS

Ethics approval

The Medical Ethical Committee Utrecht, the Netherlands, provided approval of the study on 1 August 2018.

Author Contributions

GJG, FHR, KGMM, MVH, FAK, JWB and RM designed the study. RM drafted the first version of the manuscript. All authors critically reviewed and revised the manuscript before providing final approval.

Funding

The PECAN study is funded with an unrestricted grant from The Netherlands Organisation for Health Research and Development (ZonMw, projectnumber 839110020). There are no restrictions to the execution of the study or the publication process by this subsidizing party.

Competing interests

RM, JWB and KGMM declare that they have no competing interest. FHR and GJG have received educational institutional grants from Boehringer Ingelheim, Daiichi Sankyo, Bayer Healthcare and Pfizer-BMS and an unrestricted institutional grant from Boehringer Ingelheim. In addition GJG is supported by a research grant on diagnosing pulmonary embolism from the Netherlands Organization for Scientific Research (non-profit, ZonMw 016.166.030). FAK has received research grants from Bayer, BMS/Pfizer, Boehringer Ingelheim, Daiichi Sankyo, MSD, Actelion, Dutch Heart foundation and Dutch Thrombosis association. MVH has received research grants and consultancy fees from Boehringer Ingelheim, Bayer Healthcare, Pfizer-BMS and Aspen.

ACKNOWLEDGEMENTS

We are thankful to the region coordinators affiliated to the Reinier Haga Medical Diagnostic Center and the ELAN Research Network for promoting the study in the participating general practices.

REFERENCES

1 Erkens PMG, Lucassen WAM, Geersing GJ, *et al*. Alternative diagnoses in patients in whom the GP considered the diagnosis of pulmonary embolism. *Fam Pract* 2014;31:670–7. doi:10.1093/fampra/cmu055

2 Geersing GJ, De Groot JA, Reitsma JB, *et al*. The impending epidemic of chronic cardiopulmonary disease and multimorbidity: The need for new research approaches to guide daily practice. *Chest* 2015;148:865–9. doi:10.1378/chest.14-3172

3 Schouten HJ, Geersing GJ, Oudega R, *et al*. Accuracy of the Wells Clinical Prediction Rule for Pulmonary Embolism in Older Ambulatory Adults. *J Am Geriatr Soc* 2014;62:2136–41. doi:10.1111/jgs.13080

4 Schouten HJ, Geersing GJ, Koek HL, *et al*. Diagnostic accuracy of conventional or age adjusted D-dimer cut-off values in older patients with suspected venous thromboembolism: Systematic review and meta-analysis. *BMJ* 2013;346:1–13. doi:10.1136/bmj.f2492

5 Le Gal G, Bounameaux H. Diagnosing pulmonary embolism: Running after the decreasing prevalence of cases among suspected patients. *J Thromb Haemost* 2004;2:1244–6. doi:10.1111/j.1538-7836.2004.00795.x

6 Mitchell AM, Jones AE, Tumlin JA, *et al*. Prospective study of the incidence of contrast-induced nephropathy among patients evaluated for pulmonary embolism by contrast-enhanced computed tomography. *Acad Emerg Med* 2012;19:618–25. doi:10.1111/j.1553-2712.2012.01374.x

7 Zondervan RL, Hahn PF, Sadow CA, *et al*. Body CT Scanning in Young Adults: Examination Indications, Patient Outcomes, and Risk of Radiation-induced Cancer. *Radiology* 2013;267:460–9. doi:10.1148/radiol.12121324

8 Buller HR, Cate-hoek AJ, Hoes AW, *et al*. Safely Ruling Out Deep Venous Thrombosis in Primary Care. *Ann Intern Med* 2009;150:229–36. doi:10.7326/0003-4819-150-4-200902170-00003

9 Geersing GJ, Erkens PMG, Lucassen WAM, *et al*. Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in Primary care: Prospective cohort study. *BMJ* 2012;345:1–10. doi:10.1136/bmj.e6564

10 Geersing GJ, Janssen KJM, Oudega R, *et al*. Excluding venous thromboembolism using point of care D-dimer tests in outpatients: A diagnostic meta-analysis. *BMJ* 2009;339:450. doi:10.1136/bmj.b2990

11 Hendriksen JMT, Koster-van Ree M, Morgenstern MJ, *et al*. Clinical characteristics associated with diagnostic delay of pulmonary embolism in primary care: a retrospective observational study. *BMJ Open* 2017;7:e012789. doi:10.1136/bmjopen-2016-012789

12 Van Es J, Beenen LFM, Douma RA, *et al*. A simple decision rule including D-dimer to reduce the need for computed tomography scanning in patients with suspected pulmonary embolism. *J Thromb Haemost* 2015;13:1428–35. doi:10.1111/jth.13011

13 van der Hulle T, Cheung WY, Kooij S, *et al*. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet* 2017;390:289–97. doi:10.1016/S0140-6736(17)30885-1

14 van der Pol LM, Tromeur C, Bistervels IM, *et al*. Pregnancy-Adapted YEARS Algorithm for Diagnosis of Suspected Pulmonary Embolism. *N Engl J Med* 2019;380:1139–49. doi:10.1056/NEJMoa1813865

15 van Mourik Y, Rutten FH, Moons KGM, *et al*. Prevalence and underlying causes of dyspnoea in older people: a systematic review. *Age Ageing* 2014;43:319–26. doi:10.1093/ageing/afu001

16 Huisman M V., Klok FA. Diagnostic management of acute deep vein thrombosis and pulmonary embolism. *J Thromb Haemost* 2013;11:412–22. doi:10.1111/jth.12124

17 Righini M, Van Es J, Den Exter PL, *et al*. Age-adjusted D-dimer cutoff levels to rule out pulmonary embolism: The ADJUST-PE study. *JAMA - J Am Med Assoc* 2014;311:1117–24. doi:10.1001/jama.2014.2135

18 Dronkers CEA, van der Hulle T, Le Gal G, *et al*. Towards a tailored diagnostic standard for future diagnostic studies in pulmonary embolism: communication from the SSC of the ISTH. *J Thromb Haemost* 2017;15:1040–3. doi:10.1111/jth.13654

19 Van Belle A, Büller H, Huisman M, *et al*. Effectiveness of Managing Suspected Pulmonary Embolism Using an Algorithm. *Jama* 2006;295:172–9. doi:10.1001/jama.295.2.172

20 Carrier M, Righini M, Wells PS, *et al*. Subsegmental pulmonary embolism diagnosed by computed tomography: Incidence and clinical implications. A systematic review and meta-analysis of the management outcome studies. *J Thromb Haemost* 2010;8:1716–22. doi:10.1111/j.1538-7836.2010.03938.x

21 Ikesaka R, Carrier M. Clinical significance and management of subsegmental pulmonary embolism. *J Thromb Thrombolysis* 2015;39:311–4. doi:10.1007/s11239-015-1169-7

22 Moons KGM, Kengne AP, Woodward M, *et al*. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 2012;98:683 LP-690. doi:10.1136/heartjnl-2011-301246

23 Biesheuvel CJ, Vergouwe Y, Steyerberg EW, *et al*. Polytomous logistic regression analysis could be applied more often in diagnostic research. *J Clin Epidemiol* 2008;61:125–34. doi:10.1016/j.jclinepi.2007.03.002

24 Donders ART, van der Heijden GJMG, Stijnen T, *et al*. Review: A gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;59:1087–91. doi:10.1016/j.jclinepi.2006.01.014

25 NHG-werkgroep. NHG-Standaard Diepe veneuze trombose en longembolie. *Huisarts Wet* 2017;60:460.

26 Naaktgeboren CA, Groot JAH de, Smeden M van, *et al*. Evaluating Diagnostic Accuracy in the Face of Multiple Reference Standards. *Ann Intern Med* 2013;159:195–203. doi:10.7326/0003-4819-159-3-201308060-00009

27 van der Pol LM, van der Hulle T, Cheung YW, *et al*. No added value of the age-adjusted D-dimer cut-off to the YEARS algorithm in patients with suspected pulmonary embolism. *J Thromb Haemost* 2017;15:2317–24. doi:10.1111/jth.13852

Authorship statement

I contributed to defining the research questions, the experimental design, and the methodology, and wrote the first draft of the manuscript. During the whole process, I asked for and implemented input and feedback from the other contributors to this study.

Validation and impact of the YEARS clinical decision rule for diagnosing pulmonary embolism in primary care: interim analysis of the PECAN study

Rosanne van Maanen
Frans H. Rutten
Melchior C. Nierman
Frederikus A. Klok
Menno V. Huisman
Jeanet W. Blom
Karel G.M. Moons
Geert-Jan Geersing

ABSTRACT

Background A low clinical probability combined with a negative D-dimer can safely rule-out pulmonary embolism (PE) in suspected patients, also in primary care. The threshold of 500 ng/mL for D-dimer results frequently in false positives, leading to unnecessary referral rates for computed tomographic pulmonary angiography (CTPA; the reference standard). The YEARS algorithm was developed to overcome this problem, and is advocated for the hospital setting to safely reduce the need of CTPAs when either the D-dimer threshold of 500 ng/mL or 1000 ng/mL was used, depending on pre-test probability. There is a need to validate this diagnostic strategy also in primary care where the prevalence of PE is not only lower, but on average also encounters less severe PE cases.

Methods We designed a multi-angled diagnostic validation study of the YEARS algorithm in Dutch primary care. Patients were followed for three months to establish the final diagnosis of PE presence or absence. The primary analysis focused on estimating the failure rate and efficiency of the YEARS algorithm in primary care. The failure rate was defined as the proportion of false negative tests in those where the YEARS algorithm did not recommend referral for CTPA. Efficiency denoted the proportion of all patients classified in this 'non-referral' category by the algorithm. Different routes of inclusion were applied, including a novel designed learning healthcare system (LHS) route that subsequently can be easily transformed into a formal implementation program should validation be successful. Inclusion into this LHS route is still ongoing. Here, we present an interim analysis of currently included patients.

Results This interim analysis includes 482 primary care patients with suspected acute PE. Seven (1.5%) were lost to follow-up, while in the remaining 475 patients, 36 had a PE (prevalence 7.6%). In total, 376 patients were classified as 'low-risk' by the YEARS algorithm (297 with zero YEARS items and a D-dimer <1000 ng/mL, and 79 with ≥1 positive YEARS item and a D-dimer <500 ng/mL) resulting in an efficiency of 79.2% (95% CI 75.2-82.7). Of these patients, four had a (non-fatal) PE during 3-months follow-up, of which two had zero YEARS items and two had ≥1 YEARS item, resulting in an overall diagnostic failure rate of 1.06% (95% CI 0.03-2.89%).

Conclusions This interim analysis confirms that a low-probability estimation according to the YEARS algorithm can safely and efficiently exclude PE in suspected primary care patients. To obtain even more robust estimates of the outcomes of the YEARS algorithm in various subgroups of suspected primary care PE patients, patient enrolment is still taking place using our developed LHS route.

INTRODUCTION

The diagnostic management of pulmonary embolism (PE) is always challenging given the non-specific symptoms that may mimic other, less severe diagnoses such as myalgia or a respiratory tract infection, but also other life-threatening events such as acute coronary syndrome (ACS).[1,2] Considering the potential morbidity and mortality of PE, prompt and adequate decisions about further testing (e.g. D-dimer and radiological imaging) and referral to hospital care of those suspected of PE are pivotal.[3,4] Therefore, general practitioners (GPs), but also other physicians working in primary care (i.e. general internists) have to stratify patients as either at *high risk* for PE – warranting swift referral for subsequent diagnostic testing – or at *low risk* for PE in which case watchful waiting is adequate, or treatment for a mimicking disorder if indicated.

In the past decades, several clinical decision rules have been developed to guide this diagnostic management in patients suspected of PE.[5–7] In Dutch primary care, the Wells rule is now implemented as standard-of-care for suspected PE patients.[8] This clinical decision rule has been previously validated in primary care and applying this rule resulted in a referral rate of 54% (i.e. an efficiency of 46%) with an adequately safe false negative rate of 1.5%.[9] Still, most referred patients were ultimately not diagnosed with PE, and this was most likely because of a false-positive D-dimer result above 500 ng/mL. Notably elderly and patients with comorbidities may have D-dimer values between 500 to 1000 ng/mL without having a PE according to the reference standard; computed tomographic pulmonary angiography (CTPA).[10] Besides, the Wells rule is frequently incorrectly applied by GPs, leading to lower efficiency and, importantly, also a higher failure rate.[11]

Recently, a simplified diagnostic model for PE was introduced recommending either of two D-dimer thresholds (500 ng/mL or 1000 ng/mL) depending on pre-test probability: the YEARS algorithm. This clinical decision rule was developed and validated in the hospital setting.[12–14] Physician have to score three YEARS items, namely (i) haemoptysis, (ii) signs or symptoms of deep venous thrombosis (DVT), and (iii) whether PE is considered the most likely diagnosis. Next, a D-dimer test is ordered for all suspected patients. For patients without any YEARS item present, the D-dimer threshold is elevated to 1000 ng/mL, while for patients with one or more of the YEARS items present, the D-dimer threshold is 500 ng/mL. In patients with a D-dimer below either threshold, PE is considered ruled out, whereas in patients with a D-dimer above either threshold, diagnostic imaging (CPTA) is recommended.

Validation of this algorithm in the hospital setting showed that PE could be ruled out with a reduction of CTPAs of 14% as compared to a fixed D-dimer threshold.[13] Importantly, in only 0,6% of patients a PE diagnosis was missed, which is below the internationally accepted and widely used safety threshold of 3%.[9,15]

A systematic literature review followed by individual patient data meta-analyses showed that evaluating the YEARS algorithm retrospectively in primary care patients might be safe and more efficient than the Wells rule.[16] Nevertheless, before wide scale advocating and incorporating the diagnostic strategy using the YEARS algorithm in primary care, validation in this setting is needed.

METHODS

Study design and procedures

We performed a large-scale, multi-angled prospective cohort study including primary care patients suspected of acute PE. Detailed methods are described in the previously published study protocol.[17] Originally, this study was set up as a multicentre study in GP practices spread across the Netherlands with patients being managed by their GP according to the YEARS algorithm. Because of lagging patient accrual largely caused by difficulties in performing research during the COVID-19 pandemic (inclusion period October 2018 till August 2022), we decided to expand our study by adding additional routes for patient inclusion. Hereto, the following routes were developed, namely 1) a new 'easier-to-use' inclusion route for GPs with the additional advantage that it can be transformed into an implementation program, and by 2) re-using existing, prospectively collected data of suspected PE patients in primary care. The three study routes for patient inclusion are explained below.

First, in the (original) inclusion route, the 'prospective management route', consecutive patients visiting their GP because of suspected PE were asked for informed consent and data collection. Next, the GP scored the three YEARS items and registered age, sex, symptoms, and comorbidities, and performed a rapid D-dimer test. Different D-dimer assays were used via the laboratories of the participating regions. GPs were instructed that when a patient had no YEARS item positive and a D-dimer below 1000 ng/mL, or one or more YEARS items and a D-dimer below 500 ng/mL, PE could be considered as ruled out and patients did not need to be referred to secondary care for further diagnostic testing for PE. The other way around, patients without YEARS items and a D-dimer above 1000 ng/mL, or with one or more YEARS items and a D-dimer above 500 ng/mL, should be referred to the hospital for diagnostic CTPA imaging.

Secondly, the new inclusion route was established in October 2021. Via this route, prospective data collection of the YEARS items scored by GPs in consecutive patients with suspected PE was blended into the digital system that GPs use for D-dimer test ordering at the medical diagnostic center (MDC). In case the GPs considered ordering a D-dimer, the digital system automatically prompted a question whether this D-dimer test was performed for a patient with suspected PE. If the GPs ticked 'yes' as an answer to this question, the system then automatically presented the YEARS-items that could be filled-in by the GP. This information was subsequently gathered by the MDC and eligible patients received a letter to ask for consent for data collection. If patient consent was obtained, sex, age, YEARS items, and D-dimer result were transferred by the MDC

to the researchers. During the execution of this current study, decisions on hospital referral were still left at the discretion of the attending GP, thus not formally dictated by the YEARS algorithm as was the case in the original developed route. This inclusion route is ongoing as part of our learning healthcare system (LHS) in collaboration with our primary care laboratory partners, and will be referred to as our LHS route.

Thirdly, we re-used existing data from patients included in a previously completed prospective diagnostic cohort study. This study was performed between 2013 and 2017 and focussed on implementing the Wells rule for both DVT and PE in primary care.[18] Since the three YEARS items are also included in the Wells rule, we could re-use these data for our current study on evaluating the YEARS algorithm, allowing us to obtain more robust and generalizable estimates on the failure rate and efficiency of the YEARS algorithm. For our current study, we only included patients from this cohort study suspected of PE in combination with a quantitative D-dimer test result. For this route, a waiver for informed consent was provided by ethical committee from the University Medical Center of Utrecht.

Participants

Irrespective of the inclusion route used, participating GPs identified eligible patients aged 18 years or older and suspected of acute PE, defined as (sub)acute onset of unexplained shortness of breath with or without chest discomfort such as chest pain with or without fixation to inspiration. Exclusion criteria were current treatment with therapeutic doses of vitamin K antagonists, low-molecular-weight heparin or a direct oral anticoagulant, life expectancy less than 1 month estimated by the GP, and pregnancy until 6 weeks after delivery.

Reference standard

In all three inclusion routes, the same reference standard was applied. Included patients were followed up prospectively for three months to assess the PE presence or absence to establish a uniform outcome assessment. This reference standard for diagnosing PE is typical for studies performed in the field of diagnosing PE.[10,13] A diagnosis of PE during the three months of follow-up was considered definitive if (1) CTPA demonstrating a filling defect in a central, segmental or lobular pulmonary artery, or a subsegmental filling defect that requires anticoagulant therapy, or (2) in case of a high probability ventilation/perfusion lung scan, or (3) a pulmonary angiogram demonstrating an intraluminal filling defect, or (4) PE was demonstrated during autopsy in case of death, or (5) DVT was confirmed with ultrasonography of the leg in patients (initially) suspected of PE. Thus,

in a patient with an uneventful follow-up of three months, and without fulfilling any of the five above described PE diagnosing bullets, PE was considered absent.

Sample size calculation

The sample size was calculated based on the internationally accepted and widely used safety margin of 3% for missed PE cases (false negatives).[9,19,20] Based on previous studies performed in primary care, we assumed a false negative rate of 1.5% with the upper margin of the 95% confidence interval (CI) not exceeding 3.0% (one-sided, as any proportion lower than 1.5% is preferable).[9] To have adequate power for this analysis, we needed to include 300 patients in the low probability group ((i) patients without any YEARS items and a D-dimer below 1000 ng/mL, and (ii) patients with one or more YEARS items and a D-dimer below 500 ng/mL). In the YEARS validation study in the hospital setting, the proportion of patients in this low probability group was 48%.[13] Given a lower prevalence of PE in primary care, we anticipated that the proportion of low probability patients would be as least as high in our study, which would result in a total sample size of 600 patients suspected of PE. If in addition a 10% loss to follow-up was considered (missing follow-up information on a final diagnosis), we conservatively targeted to include 750 patients of which at least 300 patients should be categorized in the low probability group.

Data analysis

In the current paper, we describe the interim results of currently included patients with the aim to present validation findings of the YEARS algorithm in patients suspected of PE in primary care. Patient characteristics, the presence or absence of the three YEARS items (as well as the remaining Wells items if collected), the median D-dimer value, and PE prevalence were reported for the total study group and stratified by inclusion route. Our primary analysis outcomes, the failure rate and efficiency of the YEARS algorithm, were quantified with corresponding 95% CIs. The efficiency was the proportion of patients in the low-probability category according to the YEARS algorithm among the total study population. The failure rate was the proportion of patients with a diagnosis of venous thromboembolism (VTE) after 3 months of follow-up among the patients with a low probability of PE according to the YEARS algorithm. We calculated the failure rate and efficiency of the YEARS algorithm among the total study population, as well as stratified by each the three different inclusion routes. Furthermore, we calculated the sensitivity and specificity of the YEARS algorithm. Lastly, we described in detail the events of missed diagnoses of PE at baseline and deceased patients. Because missing data was limited (the follow-up information about a final diagnosis was missing in only 7 patients (1.5%)) we proceeded with complete case analysis. All data analyses were performed in R version 4.0.3.

RESULTS

From November 2018 until April 2022, we prospectively included 100 consecutive patients suspected of PE via the prospective management route and 232 patients via the LHS route. In addition, 150 patients were included via the previously completed prospective cohort study route, resulting in a total inclusion of 482 patients (see Figure 1). Of 7 (1.5%) patients we did not retrieve the final diagnosis of the GP and they were excluded from our analysis. The total study population thus consisted of 475 suspected PE patients from primary care. In two patients, one or more YEARS items were missing (clinical signs of DVT and/or haemoptysis). However, in both patients the third YEARS item (PE most likely diagnosis) was scored positively, therefore we could classify them in the group with one or more YEARS items positive with a corresponding D-dimer threshold of 500 ng/mL.

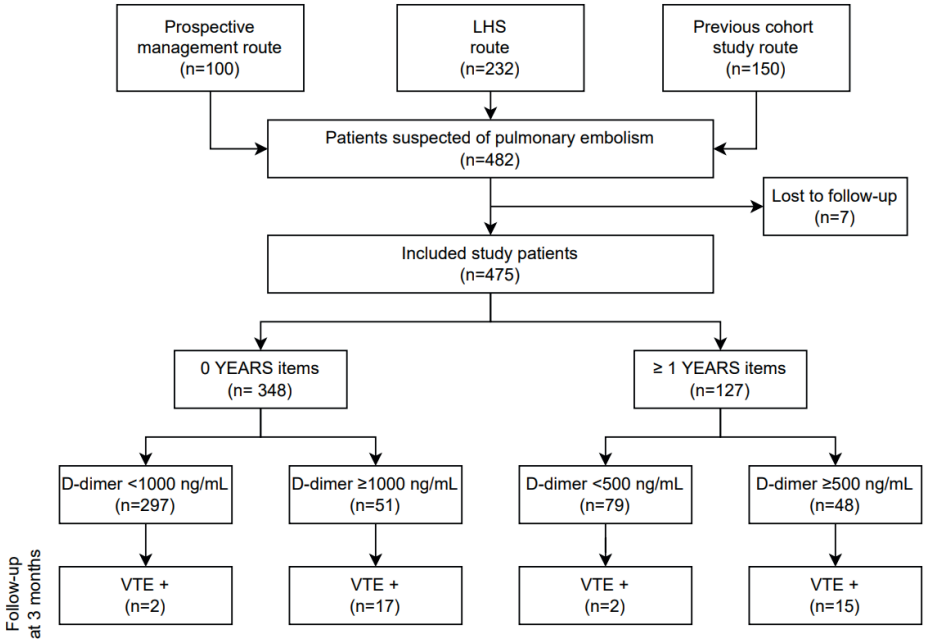


Figure 1. Flowchart of included patients

LHS: learning healthcare system; VTE: venous thromboembolism

Patient characteristics

The clinical characteristics of the 475 included patients are shown in Table 1. The median age was 51 (IQR 36.5-63.0) years, 68% were female, and the prevalence of PE was 7.6%. The YEARS item, 'PE most likely diagnosis' was most often scored positive in the previous cohort study route (30.2%), followed by the prospective management route (27.0%) and the data collection route (11.5%). The other two YEARS items (i.e. clinical signs of DVT and haemoptysis) were scored positively in 5.0% and 6.0% in the prospective management route, 2.7% and 4.4% in the data collection route, and 3.4% and 2.0% in the previous cohort study group, respectively.

Table 1. Clinical characteristics of 475 included patients stratified per inclusion route.

	Missing proportion	**Prospective management route** (n=100)	**LHS route** (n=226)	**Previous cohort study route** (n=149)	**Total study population** (n=475)*
Median age, years (IQR)	0	52.5 (34.0, 63.0)	52.0 (39.0, 63.0)	47.0 (35.0, 64.0)	51.0 (36.5, 63.0)
Women, n (%)	4 (0.8%)	72 (74.2%)	153 (67.7%)	98 (65.8%)	323 (68.0%)
Median D-dimer, ng/mL (IQR)	0	397 (205, 598)	311 (189, 591)	380 (190, 790)	340 (199, 648)
PE most likely diagnosis, n (%)	0	27 (27.0%)	26 (11.5%)	45 (30.2%)	98 (20.6%)
Clinical signs of DVT, n (%)	2 (0.4%)	5 (5.0%)	6 (2.7%)	5 (3.4%)	16 (3.4%)
Haemoptysis, n (%)	1 (0.2%)	6 (6.0%)	10 (4.4%)	3 (2.0%)	19 (4.0%)
Heart rate >100/minute, n(%)	229 (48.0%)	30 (30.0%)	n.a.	28 (19.2%)	58 (23.6%)
Immobilisation/ surgery <1 month, n (%)	229 (48.0%)	6 (6.0%)	n.a.	12 (8.1%)	18 (7.3%)
Active malignancy, n (%)	227 (47.6%)	0 (0%)	n.a.	13 (8.7%)	13 (5.2%)
History of VTE, n (%)	227 (47.6%)	6 (6.0%)	n.a.	21 (14.1%)	27 (10.9%)
Final PE diagnosis, n (%)	0	8 (8.0%)	13 (5.8%)	15 (10.1%)	36 (7.6%)

*IQR: interquartile range; LHS: learning healthcare system; DVT: deep venous thrombosis; VTE: venous thromboembolism; PE: pulmonary embolism; n.a. = not assessed *Percentages were calculated in patients without missing data on that variable.*

Failure rate and Efficiency

In total, 376 patients were categorized in the low probability group according to the YEARS algorithm, with either no YEARS items and a D-dimer below 1000 ng/mL (n=297), or one or more YEARS items and a D-dimer below 500 ng/mL (n=79). In these 376 patients, PE would be considered ruled-out and referral would not be recommended according to the YEARS algorithm. This results in an efficiency of 79.2% (95% CI 75.2–82.7). In the patients with a high probability of PE (n=99), a total of 32 (32.3%) were diagnosed with PE. Of the patients categorized as low probability, four had a (non-fatal) PE after three months of follow-up, resulting in a failure rate of 1.06% (95% CI 0.03-2.89). Three of these four patients were diagnosed with PE at baseline because the GP referred them despite a D-dimer below the threshold. Almost all patients with PE were diagnosed by CTPA, except for two patients included via the LHS route in whom the GP decided to start an anticoagulant because of a suspected PE diagnosis and without referring the patient for further diagnostic imaging. The sensitivity was 0.89 (95% CI 0.73-0.96) and the specificity 0.85 (95% CI 0.81-0.88). See table 2.

Table 2. Contingency table with total number of included patients.

	PE diagnosis +	**PE diagnosis -**	**Total**
YEARS high probability	32	67	99
YEARS low probability	4	372	376
Total	36	439	475

PE: pulmonary embolism

Table 3 shows the failure rate and efficiency stratified by the three different inclusion routes. The failure rate was lowest and the efficiency was highest in the LHS route (0.52% and 85.0%, respectively), followed by the management route (1.28% and 78.0%, respectively) and the previous cohort study route (1.89% and 71.1%, respectively), yet with overlapping 95% confidence intervals.

A description of the events of patients in whom a PE diagnosis was missed at baseline can be found in Table 4. Three patients had a D-dimer below the standard and fixed threshold of 500 ng/mL.

Table 3. Stratified analysis of primary outcomes in different inclusion routes.

Inclusion route	**Prevalence PE**	**Failure rate** (95% CI)	**Efficiency** (95% CI)
Management route (n = 100)	8.0%	1.28% (0.07 – 7.91)	78.0% (68.4 – 85.4)
LHS route (n = 226)	5.8%	0.52% (0.03 – 3.31)	85.0% (79.5 – 89.2)
Previous cohort study route (n = 149)	10.1%	1.89% (0.33 – 7.32)	71.1% (63.1 – 78.1)
Total study population (n = 475)	7.6%	1.06% (0.03 – 2.89)	79.2% (75.2 – 82.7)

LHS: learning healthcare system; PE: pulmonary embolism; CI: confidence interval

Of the 475 included patients, two died during follow-up; both had a high probability of PE and were therefore referred for diagnostic imaging, of which one was diagnosed with PE at baseline. According to the proposed classification by the International Society of Thrombosis and Haemostasis, the causes of death were classified as B2 (insufficient clinical information available to determine the cause of death) and C (cause of death other than PE).[21] Therefore, these death cases were classified as not being PE-related.

Table 4. Diagnostic failures in patients with a negative YEARS algorithm

	Inclusion route	**Sex**	**Age** (years)	**YEARS score**	**D-dimer** (ng/mL)	**Outcome**	**Description of event**
Patient 1	Management route	Female	71	1 (PE most likely)	189	PE	Prophylactic LMWH was stopped 1 week prior to symptoms (hip surgery 5 weeks earlier). On baseline referred by GP despite negative D-dimer.
Patient 2	Data collection route	Male	55	0	712	PE	PE diagnosed by CTPA on baseline
Patient 3	Previous cohort study	-	51	0	290	PE	PE occurred during follow-up
Patient 4	Previous cohort study	Male	53	1 (PE most likely)	390	PE	Referred by the GP on baseline despite negative D-dimer

CTPA: computed tomographic pulmonary angiography; LMWH: low-molecular-weight heparin; PE: pulmonary embolism; GP: general practitioner

DISCUSSION

In this interim analysis of our on-going diagnostic validation study of the YEARS algorithm, almost 80% of patients suspected of PE in primary care were classified as low probability according to the YEARS algorithm and in these patients a hospital referral for CTPA could be safely avoided. The diagnostic VTE failure rate in this low-risk category was 1.06% (95% CI 0.03 – 2.89). These findings were consistent across all three different inclusion routes. Thus, the YEARS algorithm already appears to safely rule-out PE in suspected patients in primary care with a very high efficiency.

Comparison with existing literature

The 79% efficiency of the YEARS algorithm in the primary care setting in our study is much higher than observed in previous primary care studies. In the primary care validation study of the Wells rule for PE the efficiency was 46%.[9] The currently shown efficiency was also higher than the 55% efficiency of applying the YEARS algorithm in a recent individual patient data meta-analysis including more than 3,000 primary care patients suspected of PE.[16] The higher efficiency in our study might partially be caused by the rather low prevalence of PE (8%) in our study compared to previous primary care studies (prevalence of PE ranging from 12% to 28%).[9,10] Interestingly, there seems to be a worldwide tendency to a lower threshold of suspicion of PE, leading to a steadily decreasing prevalence of PE in the last decade in those suspected.[22] A reflection of this lowering threshold of suspicion is also observed in the proportion of patients in whom the subjective item 'PE most likely diagnosis' was positively scored. In our current dataset, this was 20.6%, which is substantially lower than in the YEARS validation study in the hospital setting (50.0%) or a study performed between 2014 and 2017 in primary care (55.6%).[9,13] Over the last years, GPs are more 'defensive' and are more afraid of a missed or delayed diagnosis of PE.[23] A recent qualitative study confirmed that most doctors are indeed concerned about missing a PE diagnosis.[24] It is important to acknowledge and monitor this lowered PE suspicion closely as it will unintentionally lead to over-testing for PE and thus increased unnecessarily referral rates. Nevertheless, in our study, one in three referred patients were diagnosed with PE, which is notably higher than what was historically observed when using the Wells rule (one in four referred patients were diagnosed with PE).[9] As such, the YEARS algorithm may aid both patients and doctors in both directions, i.e. safely reassuring that PE is not present whilst at the same time also accurately identifying a higher risk group in need for prompt referral for CTPA. Even more importantly, the failure rate of the YEARS algorithm in our current study is comparable to other primary care studies

with patients suspected of PE as well as with the prospective validation study of the YEARS algorithm in secondary care.[9,13,25] Interestingly, three of the four patients in whom a PE diagnosis was missed by the YEARS algorithm had a D-dimer below 500 ng/mL. Thus, when the Wells rule would have been applied with a fixed D-dimer threshold of 500 ng/mL, these cases would most definitely also have been missed.

Strengths and limitations

To the best of our knowledge, this is the first study that has prospectively validated the YEARS algorithm in primary care patients suspected of PE. We included 475 patients in this interim analysis, with 376 patients in the low probability group, thus reaching our predefined sample size of including at least 300 patients with a low probability of PE based on the YEARS algorithm. Therefore, we can robustly demonstrate the outcomes of the YEARS algorithm. However, this study also has some limitations.

First, although we initially planned to *manage* all patients by the YEARS algorithm, this turned out to be not feasible from a research perspective due to lagging patient accrual notably enhanced by difficulties to perform diagnostic research during the COVID19 pandemic. Therefore, it was deemed necessary to (i) set up another inclusion route with prospective data collected using a routine healthcare infrastructure and (ii) to re-use existing data from a previously performed prospective cohort study from our group. The inclusion of patients from different recruiting routes could be seen as a strength because it enlarges the generalisability and robustness, but it also has some limitations. The GPs of the patients included in these two latter cohorts did not actually manage patients by applying the YEARS algorithm and decided their subsequent management on the available current guidelines. Consequently, patients from these two cohorts with a D-dimer above 500 ng/mL were referred to secondary care. This could have resulted in a higher rate of subsegmental PEs (with unknown clinical significance) as compared to when the YEARS algorithm would have been applied.[26] Also, different D-dimer assays were used per inclusion route, except for all patients included via the LHS route in which the Innovance® D-dimer assay was used. Furthermore, our exclusion criteria were not always strictly followed in these two routes, since this is real-world data collected from our LHS. As a result, five patients aged below 18 years were included in the previous cohort route and some patients might have been pregnant or might have used anticoagulants; all reasons to refrain from using the YEARS algorithm. We previously demonstrated that such incorrect application of clinical decision rules by GPs leads to a decreasing efficiency and an increasing failure rate.[11] However, our current study was a pragmatic study

and these outcomes reflect daily practice, and most importantly, still only 1.06% of PE cases was missed. Lastly, by combining the data from the three different inclusion routes, we could have introduced heterogeneity and clustering in our data. In the patients included via the LHS route the prevalence of PE was lower (5.8%), and the YEARS item '*PE most likely diagnosis*' was scored less often positive (11.5%) as compared to the management route (7.6% and 27.0%, respectively) and the previous cohort study group (10.1% and 30.2%, respectively). This clearly reflects the inclusion of more low-risk patients when patients are included via the LHS route. Probably most patients included via this route would have also a low risk on the Wells rule, because otherwise they would, conform the GP guidelines on VTE, have been referred to hospital care directly, without D-dimer testing. Of note, GPs will not apply the YEARS algorithm in patients at very high risk of PE or if haemodynamically unstable, because the waiting for the D-dimer result from the laboratory takes hours and thus is not acceptable in these patients. Yet, importantly, the failure rate and efficiency of the YEARS algorithm were comparable among all three routes, indicating that this new diagnostic model remains safe and efficient across different time frames and in different primary care patient groups.

Secondly, inherent to the design of our study, we must deal with differential verification because we used a combined reference standard of 3-month follow-up and diagnostic CTPA imaging. Consequently, (subsegmental) PEs might have been missed in patients who were not referred for diagnostic imaging and this may cause differential verification bias. However, predictive values (i.e. the outcomes of our study, namely failure rate and efficiency) are barely affected by differential verification, as the choice of the reference standard is almost fully dependent on the outcome of the YEARS algorithm.[27] Besides, since this was a pragmatic study, the combined reference standard is compliable with the practical use when the YEARS algorithm would be implemented as standard-of-care in primary care.

The third and last limitation is the absence of a control group and thus the impossibility to directly compare the outcomes of the YEARS algorithm with the outcomes of the Wells rule which is still recommended in the current GP guidelines on VTE. Importantly, however, our results show that with the YEARS algorithm applied in primary care the failure rate is as low as achieved by the Wells rule in primary care, but with the YEARS algorithm having a much higher efficiency.[9]

Implications for practice

Our study shows that applying the YEARS algorithm in primary care safely and efficiently rules out PE. When using the YEARS algorithm, referral to hospital care for diagnostic imaging could be withheld in almost 80% of suspected PE primary care patients. Importantly, the probability of PE in patients who should be referred is relatively high; one-third of patients in this group were eventually diagnosed with PE. Although we do continue patient enrolment into our LHS route, we believe we can already conclude that the YEARS algorithm should be considered for implementation in primary care, based upon this current interim analysis. Continued enrolment into our LHS route allows for such implementation as this can easily be transformed into an implementation program with feedback loops on e.g. D-dimer ordering and final diagnoses to participating GPs. For example, after GPs have filled out the three YEARS items, the laboratory will send the D-dimer result with the corresponding threshold based on the presence or absence of the YEARS items scored by the GP. This would improve the use and correct implementation of the YEARS algorithm and thus guideline use in primary care, thereby overcoming potential barriers of incorrect guideline application that – based upon previous work – may both contribute to an increased failure rate of the strategy as a whole, but also a decline in efficiency.[11]

Conclusion

This interim analysis confirms that a low-probability estimation according to the YEARS algorithm can safely and efficiently exclude PE in suspected primary care patients. To obtain even more robust estimates of the outcomes of the YEARS algorithm in various subgroups of suspected primary care patients, patient enrolment is still taking place.

REFERENCES

1 Erkens PMG, Lucassen WAM, Geersing GJ, *et al*. Alternative diagnoses in patients in whom the GP considered the diagnosis of pulmonary embolism. *Fam Pract* 2014;31:670–7. doi:10.1093/fampra/cmu055

2 Geersing GJ, De Groot JA, Reitsma JB, *et al*. The impending epidemic of chronic cardiopulmonary disease and multimorbidity: The need for new research approaches to guide daily practice. *Chest* 2015;148:865–9. doi:10.1378/chest.14-3172

3 Cohen AT, Agnelli G, Anderson FA, *et al*. Venous thromboembolism (VTE) in Europe - The number of VTE events and associated morbidity and mortality. *Thromb Haemost* 2007;98:756–64. doi:10.1160/TH07-03-0212

4 Swan D, Hitchen S, Klok FA, *et al*. The problem of under-diagnosis and over-diagnosis of pulmonary embolism. *Thromb Res* 2019;177:122–9. doi:10.1016/J.THROMRES.2019.03.012

5 Righini M, Van Es J, Den Exter PL, *et al*. Age-Adjusted D-Dimer Cutoff Levels to Rule Out Pulmonary Embolism. *JAMA* 2014;311:1117. doi:10.1001/jama.2014.2135

6 Goekoop RJ, Steeghs N, Niessen RWLM, *et al*. Simple and safe exclusion of pulmonary embolism in outpatients using quantitative D-dimer and Wells' simplified decision rule. *Thromb Haemost* 2007;97:146–50.

7 Kline JA. Utility of a Clinical Prediction Rule to Exclude Pulmonary Embolism Among Low-Risk Emergency Department Patients. *JAMA - J Am Med Assoc* 2018;319:559–66. doi:10.1001/jama.2017.21904

8 Dutch College of General Practitioners' Guideline deep venous thrombosis and pulmonary embolism. *Huisarts Wet* 2017.

9 Geersing GJ, Erkens PMG, Lucassen WAM, *et al*. Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in Primary care: Prospective cohort study. *BMJ* 2012;345:1–10. doi:10.1136/bmj.e6564

10 Schouten HJ, Geersing GJ, Oudega R, *et al*. Accuracy of the Wells Clinical Prediction Rule for Pulmonary Embolism in Older Ambulatory Adults. *J Am Geriatr Soc* 2014;62:2136–41. doi:10.1111/jgs.13080

11 Van Maanen R, Kingma AEC, Oudega R, *et al*. Real-life impact of clinical prediction rules for venous thromboembolism in primary care: A cross-sectional cohort study. *BMJ Open* 2020;10. doi:10.1136/bmjopen-2020-039913

12 Van Es J, Beenen LFM, Douma RA, *et al*. A simple decision rule including D-dimer to reduce the need for computed tomography scanning in patients with suspected pulmonary embolism. *J Thromb Haemost* 2015;13:1428–35. doi:10.1111/jth.13011

13 van der Hulle T, Cheung WY, Kooij S, *et al*. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet* 2017;390:289–97. doi:10.1016/S0140-6736(17)30885-1

14 van der Pol LM, Tromeur C, Bistervels IM, *et al*. Pregnancy-Adapted YEARS Algorithm for Diagnosis of Suspected Pulmonary Embolism. *N Engl J Med* 2019;380:1139–49. doi:10.1056/NEJMoa1813865

15 Righini M, Le Gal G, Aujesky D, *et al*. Diagnosis of pulmonary embolism by multidetector CT alone or combined with venous ultrasonography of the leg: a randomised non-inferiority trial. *Lancet* 2008;371:1343–52. doi:10.1016/S0140-6736(08)60594-2

16 Geersing G-J, Takada T, Klok FA, *et al*. Ruling out pulmonary embolism across different healthcare settings: A systematic review and individual patient data meta-analysis. *PLOS Med* 2022;19:e1003905. doi:10.1371/journal.pmed.1003905

17 van Maanen R, Rutten FH, Klok FA, *et al*. Validation and impact of a simplified clinical decision rule for diagnosing pulmonary embolism in primary care: design of the PECAN prospective diagnostic cohort management study. *BMJ Open* 2019;9:e031639. doi:10.1136/bmjopen-2019-031639

18 Kingma AEC, van Stel HF, Oudega R, *et al*. Multi-faceted implementation strategy to increase use of a clinical guideline for the diagnosis of deep venous thrombosis in primary care. *Fam Pract* 2016;:cmw066. doi:10.1093/fampra/cmw066

19 Righini M, Van Es J, Den Exter PL, *et al*. Age-adjusted D-dimer cutoff levels to rule out pulmonary embolism: The ADJUST-PE study. *JAMA - J Am Med Assoc* 2014;311:1117–24. doi:10.1001/jama.2014.2135

20 van Belle A, Büller HR, Huisman M V, *et al*. Effectiveness of managing suspected pulmonary embolism using an algorithm combining clinical probability, D-dimer testing, and computed tomography. *JAMA* 2006;295:172–9. doi:10.1001/jama.295.2.172

21 Trischler T, Kraaijpoel N, Girard P, *et al*. Definition of pulmonary embolism-related death and classification of the cause of death in venous thromboembolism studies: Communication from the SSC of the ISTH. *J Thromb Haemost* 2020;18:1495–500.

22 Le Gal G, Bounameaux H. Diagnosing pulmonary embolism: Running after the decreasing prevalence of cases among suspected patients. *J Thromb Haemost* 2004;2:1244–6. doi:10.1111/j.1538-7836.2004.00795.x

23 Maanen R van, Trinks-Roerdink EM, Rutten FH, *et al*. A systematic review and meta-analysis of diagnostic delay in pulmonary embolism. *Eur J Gen Pract* 2022;28:165–72. doi:10.1080/13814788.2022.2086232

24 Zarabi S, Chan TM, Mercuri M, *et al*. Physician choices in pulmonary embolism testing. *CMAJ* Published Online First: 2021. doi:10.1503/cmaj.201639

25 Hendriksen JMT, Geersing GJ, Lucassen WAM, *et al*. Diagnostic prediction models for suspected pulmonary embolism: Systematic review and independent external validation in primary care. *BMJ* 2015;351. doi:10.1136/bmj.h4438

26 van der Pol LM, Bistervels IM, van Mens TE, *et al*. Lower prevalence of subsegmental pulmonary embolism after application of the YEARS diagnostic algorithm. *Br J Haematol* 2018;183:629–35. doi:10.1111/bjh.15556

27 Naaktgeboren CA, Groot JAH de, Smeden M van, *et al*. Evaluating Diagnostic Accuracy in the Face of Multiple Reference Standards. *Ann Intern Med* 2013;159:195–203. doi:10.7326/0003-4819-159-3-201308060-00009

Authorship statement

I contributed to defining the research question, performed the data collection and data management, conducted the data analysis, and wrote the first version of the manuscript. During the whole process, I asked for and implemented input and feedback from the other contributors to this study.

General discussion

Part of this discussion is based on:

Performing diagnostic studies in primary care - a recipe for Lasagna's law?

Rosanne van Maanen
Frans H. Rutten
Jeanet W. Blom
Pieter Langers
Jochen W.L. Cals
Melchior C. Nierman
Geert-Jan Geersing

Provisionally accepted in Huisarts & Wetenschap 2022

MAIN FINDINGS OF STUDIES INCLUDED IN THIS THESIS

As already outlined in the introduction of this thesis, venous thromboembolism (VTE) is a challenging disease, notably, it is a diagnostic challenge. Diagnostic delay of pulmonary embolism (PE) is common and patients may have symptoms for almost one week before PE is diagnosed (**Chapter 1**). The correct application of clinical decision rules (CDRs) for VTE results in a high efficiency and acceptable failure rate. However, we also demonstrated that in nearly a quarter of patients the CDRs were incorrectly applied which resulted in a lower efficiency and a higher failure rate than would occur if correctly used (**Chapter 2**). We showed that a positive gestalt estimation on average increases the probability of PE threefold and that the diagnostic accuracy of gestalt is relatively stable across various healthcare settings and subgroups of patients (**Chapter 3**). To further optimize the diagnostic management of PE and for balancing between preventing over-diagnosing (with unnecessary testing and referrals) and also under-diagnosing (with a missed or delayed VTE diagnosis) we performed the PECAN study. This diagnostic study performed in primary care aimed at evaluating the YEARS algorithm, a simplified algorithm with variable D-dimer thresholds that already showed to be efficient (less unnecessary imaging) in the hospital setting, without an increase in missed diagnoses as compared to the application of the Wells rule (**Chapter 4**). The interim results of this ongoing study showed that applying the YEARS algorithm in primary care seems to be very efficient with a low failure rate (**Chapter 5**).

The PECAN study also provided us important lessons learned. It became evident during the conduct of this study that recruitment of primary care patients suspected of PE is difficult. Studies on incident diseases in primary care are complicated by the tight schedule of GPs (often only 10 minutes per patient) while study procedures have become more time-consuming over the last decades and are therefore not feasible to complete. In addition, the low prevalence of VTE leads to problems regarding remembering there is an ongoing study and a barrier to include patients because there is no routine in the inclusion process.

In this General Discussion we describe the difficulties, but also new solutions for patient recruitment in diagnostic event-driven clinical research in primary care, based on our lessons learned in conducting the PECAN study. We present a new way of performing diagnostic studies that is less time-consuming and embedded in daily routine primary care, which could also be used in future studies on many more types of incident medical disorders.

PERFORMING DIAGNOSTIC STUDIES IN PRIMARY CARE - A RECIPE FOR LASAGNA'S LAW?

Lasagna's Law

Louis Lasagna was an American clinical pharmacologist who described the phenomenon that significantly fewer patients turn out to be eligible for inclusion in clinical studies than estimated.[1] The incidence of the disease of interest seems to have suddenly dropped. This phenomenon is particularly common in research on incidental diseases, where the patient may be asked to participate in the study at the time of the consultation. For example, only 28% of primary care studies in which patients with incidental diseases were included, managed to complete patient recruitment within the planned period.[2] Failure to enrol adequate numbers leads to an increased inclusion period and thus higher costs and workload. In the worst case, a study must even stop because it is impossible to include an adequate amount of participants to ensure enough power and thus robust results. We conducted a prospective diagnostic study about pulmonary embolism (PE) in primary care and encountered the same problems because of 'Lasagna's law'.

Project description

In the fall of 2018, a study started with the aim to validate a new decision rule for PE (the YEARS algorithm, see Figure 1).[3] In secondary care, this strategy was already widely used resulting in a decrease in unnecessary CT scans while keeping a low failure rate (less than 3% missed PE).[4] In the current study, we examined whether the implementation of this algorithm by general practitioners (GPs) results also in fewer unnecessary referrals to hospital care while pertaining to the necessary low failure rate. GPs could participate by including patients with suspected PE and after retrieving written informed consent, subsequently use the YEARS algorithm. After three months of follow-up, the GPs should fill out a short form about the final diagnosis (presence/absence of PE). Based on the incidence of PE in previous studies, it was estimated that we needed at least 75 participating GPs to be able to include 750 patients in two years. In other words, it was estimated that each participating GP could include five patients with suspected PE per year. Although 50 general practices (with in general multiple GPs per practice) consented to participate, the inclusion of participants was very slow, with the additional complication of the restrictions in research in the first wave of the COVID-19 pandemic at the beginning of 2020. After two years, only 90 (12%) participants had been prospectively included. This made us decide to search for other ways to include eligible patients; we started a new inclusion strategy that is closely linked to the routine care of GPs.

When a patient is suspected of PE, the GP often orders a D-dimer test, in the Netherlands through a digital laboratory request form via 'ZorgDomein'. We wanted to use this routine care procedure for our study, and this was made possible through collaboration with ZorgDomein and the medical diagnostic center (MDC) Atalmedial who performs primary care laboratory tests. GPs affiliated with the MDC are immediately shown additional information and questions when they request a D-dimer test via ZorgDomein. The first question is: (i) are you requesting D-dimer because of a suspicion of PE? If yes: (ii) may the MDC approach your patient in the context of this D-dimer determination for research? If these questions are positively answered, information about the study is shown to the GP who is asked to score the three YEARS items. The researchers send a letter to the registered patients with information about the study and ask for informed consent for data collection via a paper consent form, email, or QR code. When the patient gives consent, the inclusion takes place and information gathered by ZorgDomein and the MDC, including the D-dimer test result, is transferred to the researchers. Three months after inclusion, the researchers retrieve the final diagnosis from the participating patients' GP.

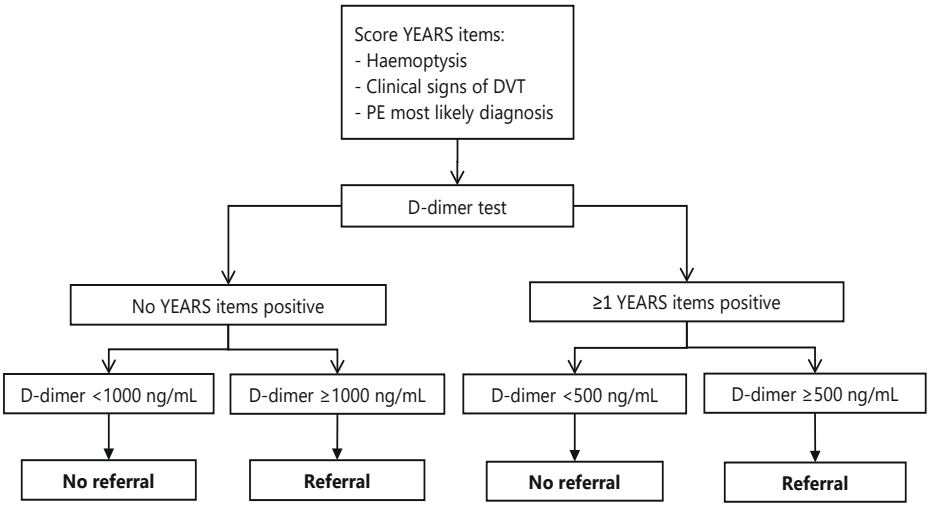


Figure 1. YEARS algorithm

This new inclusion route started at the end of August 2021 in all GP practices affiliated with the MDC Atalmedial and it was successful from the start. Of the 1780 D-dimer tests requested between September 2021 and May 2022, 1157 (65%) were ordered because of a suspicion of PE. Of these D-dimer orders, GPs gave for 581 (50%) patients consent to be contacted and completed the YEARS questions via ZorgDomein. Finally, 270 patients (46%) gave consent for data collection. This corresponds to an inclusion rate of 34 patients per month on average and this is tenfold the previous inclusion rate (see Figure 2). Both inclusion routes continue to date, with the expectation that a total of 750 patients will be included by the end of 2022.

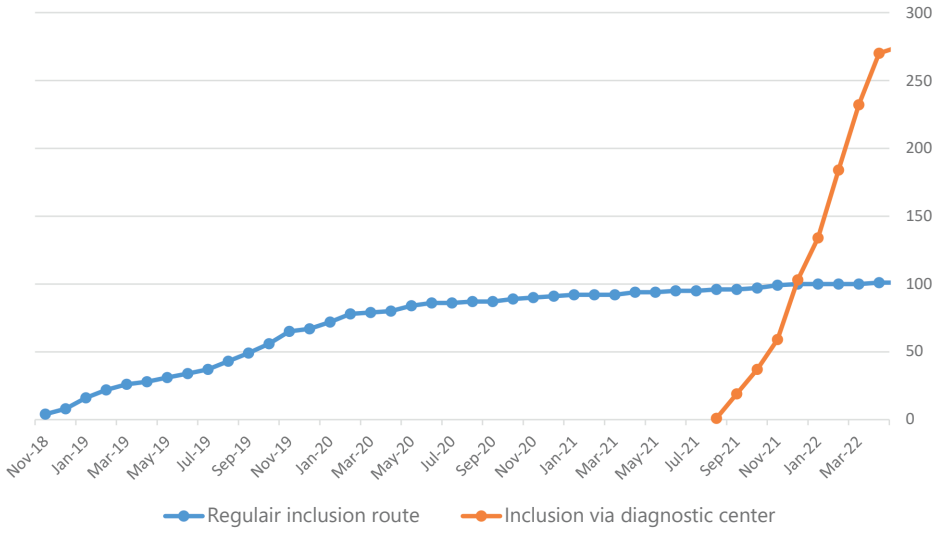


Figure 2. Inclusion rate of a diagnostic study on pulmonary embolism in primary care

Reflection

Performing scientific research in primary care is important for the rationale of our medical practice, but as a GP there are many hurdles to take before a patient is included in a study. First, the GP must be motivated to participate, next he/she must think about the study the moment a patient complies with the inclusion criteria, and then there must be time for an explanation, asking for consent, and filling out several forms. This is often impossible during an overloaded consultation hour. In short, participating in scientific research in primary care takes time and is often not a priority in the context of a high workload, even if the GP is strongly aware of the importance of research.[5] A new inclusion route in which the GP can easily register

a participant during a routine care procedure will benefit the inclusion rate, but more importantly; it will contribute to valid primary care research and thus further scientific evidence for our profession.

EXPERIENCES OF A PARTICIPATING GP AND PATIENT

GP: "*To be honest, I don't even remember the questions in 'ZorgDomein' very well, but I think that underpins the ease with which you can take part in this study as a GP. So I would say: A positive experience!*"

Patient: "*I first thought the letter was a bill from the laboratory but when I read it, I was happy to help. Participation takes little effort, I just filled out the form.*"

A recipe for Lasagna's law

The success of the new inclusion route is related to several factors. First, and perhaps most important, is that study actions for the GPs match routine care actions. GPs did not have to actively think about inclusion but were reminded of it during their D-dimer request. The route for additional data collection described in this article is possible since the MDC approaches the patient and asks for consent after the laboratory test was requested by the GP. The GP only asks the patient if the MDC may approach the patient. Thus, the often time- and regulation-intensive consent procedure is taken out of the GP's consulting room. This 'ZorgDomein route' could of course also be suitable for other (primary care) research. For example, a study about gout when requesting uric acid or a study about sexually transmitted diseases when requesting a chlamydia test. But not only diagnostic studies can be performed via this route. It could also be used to answer questions where a national overview is important. An example of this is the study of COVID-19 via ZorgDomein intending to gain insight into the number of COVID-19 patients who received intensive and palliative COVID care from GPs at the beginning of the pandemic.[6] More than 90% of GPs use ZorgDomein, which allows for nationwide scientific research. It takes little time for GPs to participate in research in this way since they do not have to fill out (consent) forms but data is collected with just a few extra clicks. By collaborating with a local MDC, this data can be transferred to the researchers after the patients' consent. In addition to ZorgDomein, there are also opportunities for embedding scientific research in the routine care of the GP. For example, when a GP enters a particular

ICPC code into the health record, a pop-up could appear to remind him/her of an ongoing study. In our opinion, the above-described factors are the ingredients for the best recipe for (or rather against) Lasagna's Law.

Future developments

The Consortium for Research in General Practice (COH) is currently working on several modules to improve the infrastructure for GPs to facilitate participation in scientific research.[7] The COH is now setting up a generic module with ZorgDomein that will enable GPs throughout the Netherlands to register their patients via ZorgDomein for currently open research studies. This module will be placed under the 'report' button on the ZorgDomein main screen. For a pilot study on shoulder complaints in the Rotterdam region, the functionality is already being tested and used. For a trial of COVID-19 treatment, the functionality will be available nationwide in the autumn of 2022. The aim is to lower the threshold for GPs to participate in primary care research because the administrative effort is greatly reduced. Furthermore, it can help to include future patients nationwide for studies, and not only for studies of the academic GP department in the region.

Conclusion

Patient recruitment for clinical research in primary care is very important but often is complicated, notably for incident disorders. Irrespective of this major hurdle, research in primary care is certainly necessary to scientifically substantiate GP actions and thus achieve better patient care. A new inclusion route that is closely linked to routine care activities of GPs will ensure that participating in research and conducting research is easier and more efficient. This is a win-win situation for everyone: researcher, GP, and last but not least the patient.

REFERENCES

1 Gorringe JAL. Initial preparation for clinical trials. In: Harris EL, Fitzgerald JD, eds. *Princ Pract Clin Trials* 1970;Edinburgh,:41–46.

2 van der Wouden JC, Blankenstein AH, Huibers MJH, *et al*. Survey among 78 studies showed that Lasagna's law holds in Dutch primary care research. *J Clin Epidemiol* 2007;60:819–24. doi:10.1016/j.jclinepi.2006.11.010

3 van Maanen R, Rutten FH, Klok FA, *et al*. Validation and impact of a simplified clinical decision rule for diagnosing pulmonary embolism in primary care: design of the PECAN prospective diagnostic cohort management study. *BMJ Open* 2019;9:e031639. doi:10.1136/bmjopen-2019-031639

4 van der Hulle T, Cheung WY, Kooij S, *et al*. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet* 2017;390:289–97. doi:10.1016/S0140-6736(17)30885-1

5 van der Worp H, Schuch GA, Loohuis AMM, *et al*. Intrinsic motivation of GPs was not related to recruitment success, whereas interest in the study topic was. *J Clin Epidemiol* 2020;125:158–60. doi:https://doi.org/10.1016/j.jclinepi.2020.06.009

6 Cals JWL, Derckx R, Blanker MH. Intensieve en palliatieve covid-19-zorg door huisarts. *Ned Tijdschr Geneeskd* 2021;164:1–9.

7 Blanker MH. COVID-19 behandelingen in de huisartsgeneeskunde: evaluatie huidige pandemie en versterking onderzoeksinfrastructuur voor toekomstige pandemieën. [Internet]. 2021;:https://www.zonmw.nl/nl/over-zonmw/coronavirus/pro.

Authorship statement

I wrote the text of the first part of the general discussion and revised the text once after the comments of my supervisors. In the second part of the general discussion, I performed the data analysis, wrote the first version of the manuscript, and implemented the contribution of the co-authors. During the whole process, I asked for input and feedback from the other contributors to this study.

Summary

INTRODUCTION

Diagnosing venous thromboembolism (VTE) is challenging, both for deep venous thrombosis (DVT) and pulmonary embolism (PE). On the one hand, there is fear among clinicians of missing a VTE diagnosis, knowing that this may have serious consequences. On the other hand, not every patient suspected of VTE can undergo CT pulmonary angiography (CTPA, the reference standard) as referring every suspected patient will not only overwhelm already strained healthcare resources, it also puts the patient at risk for iatrogenic contrast-induced harm (e.g. nephropathy). This dilemma easily results in either underdiagnosing leading to a missed or delayed VTE diagnosis, or overdiagnosing with unnecessary testing and referrals. To help physicians with this diagnostic dilemma, several clinical decision rules (CDRs) have been developed and validated in the past decades. Applying CDRs leads to fewer referrals and is safe in excluding VTE in patients with a low estimated risk as indicated by the CDR in combination with a normal D-dimer level.[1–3] However, despite the use of CDRs, the majority of patients who are referred by the general practitioner (GP) because of the suspicion of VTE do not have a confirmative diagnosis. Hence, it is still debated what the optimal trade-off is in diagnosing VTE. The objective of this thesis is to contribute to improving the diagnostic management of patients suspected of VTE.

CHALLENGES IN DIAGNOSING VTE

Diagnostic delay in patients with PE is common, yet the proportion of patients with PE who in retrospect have experienced delay, and for how many days, is less well described, nor are determinants for such delay. In **Chapter 1** we describe the findings of a systematic review and meta-analysis of diagnostic delay in PE; the prevalence, its extent, and the determinants of such delay. We performed a systematic literature search and identified 10,933 studies, of which eventually 24 fulfilled our inclusion criteria and thus could be included in the final analysis. After random effect modelling, we calculated a pooled estimate of the mean diagnostic delay of 6.3 days (95% prediction interval 2.5 to 15.8). The percentage of patients having >7 days of delay varied between 18% and 38%. The presence of coughing, chronic lung disease, and/or heart failure had a positive association with diagnostic delay, while recent surgery and hypotension, and in most studies also chest pain had a negative association with diagnostic delay of PE.

Although CDRs for VTE are widely used by GPs, the failure rate and efficiency of these CDRs, as well as determinants for, and consequences of incorrect application in real-life primary care are currently unknown. In **Chapter 2** we evaluated the real-life impact of the use of CDRs for both PE and DVT in primary care. We performed a cross-sectional cohort study and included patients suspected of DVT or PE by their GP. The outcomes were the failure rate and the efficiency, as well as determinants for, and the consequences of *incorrect* application of the CDRs. VTE was confirmed in 267 (18.1%) of the included 1,477 patients. If CDRs were correctly applied, the failure rate was 1.51% (95% CI 0.77 to 2.86%) and the efficiency was 58.1% (95% CI 55.2 to 61.0%). However, the CDRs were incorrectly applied in 339 (23.0%) patients, which resulted in a higher failure rate of 3.31% (95% CI 1.07 to 8.76%) and in addition a lower efficiency of 35.7% (95% CI 30.6 to 41.1%) in this subgroup. Concurrent heart failure increased the likelihood of incorrect application (adjusted OR 3.26; 95% CI 1.47 to 7.21). In conclusion, correct application of CDRs for VTE in primary care is associated with an acceptable low failure rate and a high efficiency. However, incorrect application, which occurred in nearly a quarter of patients, results in a failure rate point estimate above the internationally accepted threshold of 3%, and in addition a considerably lower efficiency. Thus, the optimal balance between over- and underdiagnosing is disturbed when CDRs are incorrectly applied in patients suspected of VTE.

THE DIAGNOSTIC ACCURACY OF GESTALT

In the diagnostic management of patients suspected of PE, physicians often rely on an intuitive estimation ('gestalt') of PE presence. Although repeatedly shown to be predictive for PE, gestalt is also criticized given its lack of standardization and generalisability and assumed unstable diagnostic accuracy. In **Chapter 3**, we assessed the diagnostic accuracy of physician's gestalt in diagnosing PE by performing an ancillary analysis of an individual patient data meta-analysis including patients suspected of acute PE. We explored the variability of gestalt, expressed as risk ratio (RR), across different healthcare settings, year of study, PE prevalence, and the following suspected PE patient subgroups: age, sex, heart failure, chronic lung disease, and items of the Wells algorithm for diagnosing PE. Among the included 20,770 patients from 16 individual studies, the prevalence of PE in patients with and without gestalt positively scored was 28.8% versus 9.1%, leading to a pooled RR of 3.02 (95% confidence interval 2.35-3.87). This diagnostic accuracy of gestalt varied across individual studies (range RR 1.46 to 7.71), yet performance remained stable

across subgroups. Our study thus showed that the gestalt item was of similar merit in the diagnostic management of *all* patients suspected of PE.

VALIDATION OF THE YEARS CLINICAL DECISION RULE IN PRIMARY CARE

Recently, a novel CDR for PE was developed and validated in secondary care: the YEARS algorithm.[4,5] This strategy starts with the assessment of three items; (i) are there clinical signs of DVT, (ii) is there haemoptysis, and (iii) is PE the most likely diagnosis according to the physician. In addition, D-dimer testing should be performed in all suspected patients. In patients without any of the 'YEARS items', the D-dimer threshold to be used is 1000 ng/mL, while in patients with one or more 'YEARS items' the threshold remains the standard cut-off of 500 ng/mL. Compared to applying a fixed D-dimer threshold of 500 ng/mL, the YEARS algorithm increased the efficiency: the proportion of patients in whom diagnostic imaging was not required was higher than when the standard cut-off 500 ng/mL would be applied. Importantly, this was achieved with a failure rate that remained low (0.61%). Nowadays, this strategy is implemented as the diagnostic strategy in patients suspected of PE seen at emergency departments in the Netherlands. It seems tempting to use this new algorithm in primary care as well given the higher efficiency with a remaining low failure rate. However, validation in the primary care setting is needed given the lower pre-test probability, a different case mix of patients with relatively less severe cases of PE, and fewer GP exposure to PE cases as compared to the hospital specialist. Hence, we set up a prospective diagnostic study among patients with suspected acute PE who were managed by their GP according to the YEARS diagnostic algorithm: the PECAN study. **Chapter 4** describes the rationale and design of the PECAN study.

In **Chapter 5** we present the interim results of the PECAN study. Because of lagging patient accrual mainly caused by difficulties in performing research during the COVID-19 pandemic, we expanded the PECAN study by adding two additional routes for patient inclusion: 1) a novel designed learning healthcare system (LHS) route, and 2) re-using existing, prospectively collected data of suspected PE patients in primary care, not yet analysed for the purpose of YEARS validation. The presented interim analysis included 482 primary care patients suspected of acute PE. In total, 376 patients were classified as 'low-risk' by the YEARS algorithm (297 with zero YEARS items and a D-dimer <1000 ng/mL, and 79 with ≥1 positive YEARS item and a D-dimer <500 ng/mL) resulting in an efficiency of 79% (95% CI 75 – 83%). Of

these patients, four had a (non-fatal) PE during the 3-months of follow-up, resulting in a failure rate of 1.06% (95% CI 0.03 - 2.89%). This interim analysis confirms that a low-probability estimation according to the YEARS algorithm can safely and efficiently exclude PE in suspected patients in primary care. To obtain even more robust estimates of the outcomes of the YEARS algorithm in various subgroups of suspected primary care PE patients, patient enrolment is ongoing in the LHS route.

PERFORMING DIAGNOSTIC RESEARCH IN PRIMARY CARE

It became apparent during the conduct of the studies described in this thesis, that performing prospective diagnostic VTE studies in primary care is challenging because of the high workload for GPs, time-consuming study procedures, and low prevalence of VTE; all factors leading to poor patient accrual as we indeed observed in the studies described in Chapter 5. In **Chapter 6** (General Discussion) we discuss how and why the aforementioned 'LHS route' of the PECAN study facilitates recruitment for cross-sectional diagnostic studies in daily routine primary care. These lessons learned can also be used in future studies in which incident medical disorders are the domain of study interest. It helps to ensure sufficient participation of eligible patients in research and makes conducting research easier and more efficient.

REFERENCES

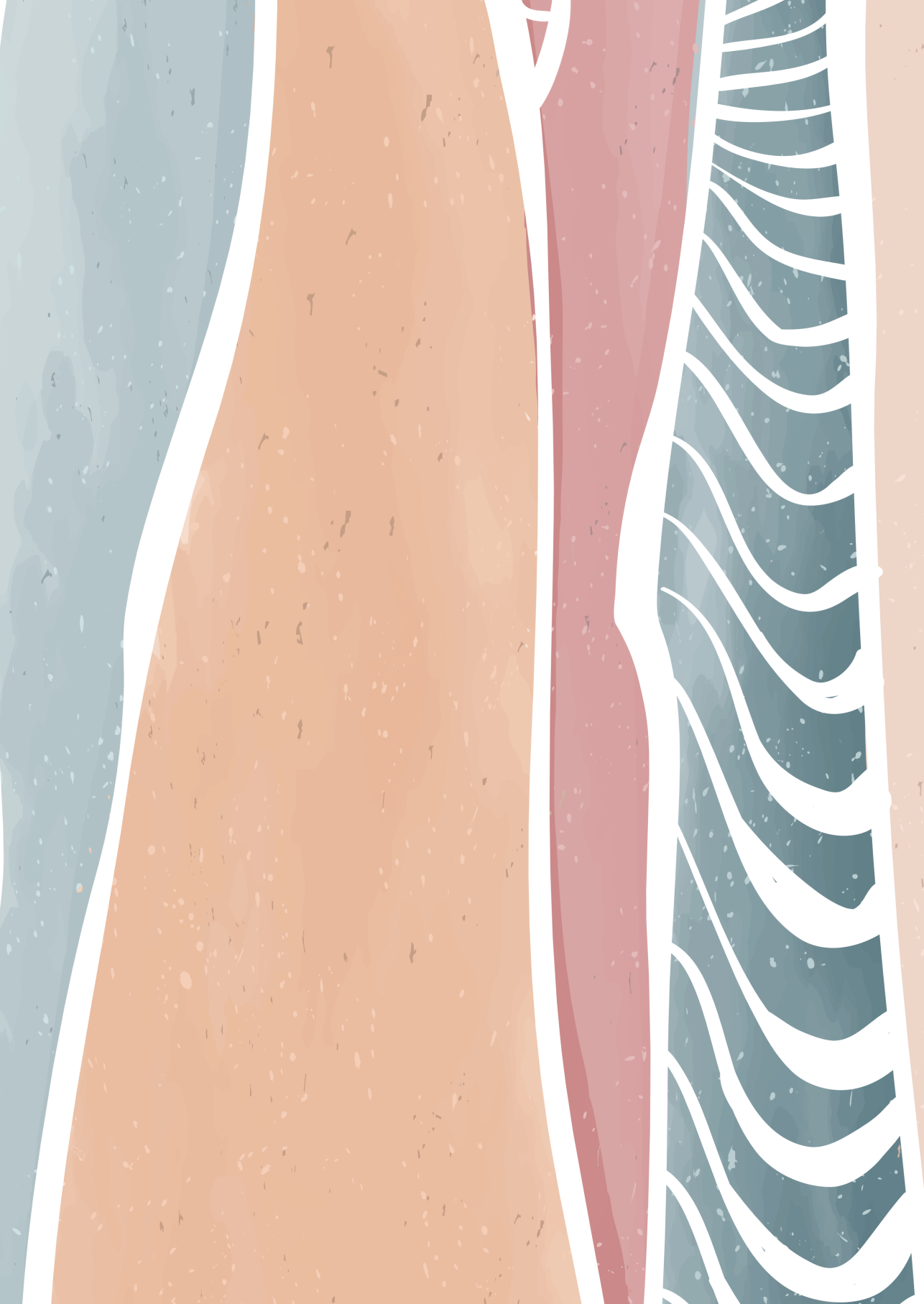
1 Geersing GJ, Erkens PMG, Lucassen WAM, *et al*. Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in Primary care: Prospective cohort study. *BMJ* 2012;345:1–10. doi:10.1136/bmj.e6564

2 Buller HR, Cate-hoek AJ, Hoes AW, *et al*. Safely Ruling Out Deep Venous Thrombosis in Primary Care. *Ann Intern Med* 2009;150:229–36. doi:10.7326/0003-4819-150-4-200902170-00003

3 Righini M, Van Es J, Den Exter PL, *et al*. Age-Adjusted D-Dimer Cutoff Levels to Rule Out Pulmonary Embolism. *JAMA* 2014;311:1117. doi:10.1001/jama.2014.2135

4 Van Es J, Beenen LFM, Douma RA, *et al*. A simple decision rule including D-dimer to reduce the need for computed tomography scanning in patients with suspected pulmonary embolism. *J Thromb Haemost* 2015;13:1428–35. doi:10.1111/jth.13011

5 van der Hulle T, Cheung WY, Kooij S, *et al*. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet* 2017;390:289–97. doi:10.1016/S0140-6736(17)30885-1



Samenvatting

INTRODUCTIE

Bij de ziekte veneuze trombo-embolie (VTE) wordt een ader in het lichaam afgesloten door een bloedstolsel (trombus). Bij een diepveneuze trombose (DVT) is er sprake van een trombus in de diepe venen van het been. Een longembolie is vaak het gevolg van het losschieten van een trombus uit het been, die vastloopt in de bloedvaten van de longen. Dit zorgt voor een afsluiting van een longslagader waardoor een deel van de long niet of slechts gedeeltelijk van bloed wordt voorzien. Zodra de diagnose VTE is gesteld, wordt een behandeling met antistolling gestart waardoor het bloedstolsel weer oplost.

Echter, het stellen van de diagnose VTE is vaak lastig. Symptomen zijn niet-specifiek en kunnen ook passen bij andere aandoeningen, zoals wondroos van het been bij patiënten verdacht van DVT, of luchtweginfecties of hartklachten bij patiënten verdacht van een longembolie. Omdat VTE ook ernstige gevolgen en nadelige langetermijneffecten kan hebben, zijn artsen vaak bezorgd om een diagnose te missen. Om zeker te weten of een patiënt DVT of longembolie heeft, kan er een echo van het been of een CT-scan van de longen gemaakt worden. Maar het is natuurlijk niet mogelijk om iedere patiënt hiervoor door te verwijzen gezien de (over)belasting van het zorgsysteem en mogelijke schadelijke effecten van de onderzoeken, zoals nierschade door een CT-scan met contrastvloeistof. Artsen ervaren dus regelmatig een diagnostisch dilemma met aan de ene kant het risico op onderdiagnostiek wat kan leiden tot een gemiste VTE diagnose, en aan de andere kant overdiagnostiek met onnodige onderzoeken en verwijzingen. Om artsen te helpen bij dit diagnostische dilemma zijn de afgelopen decennia verschillende klinische beslisregels voor VTE ontwikkeld. Bij een beslisregel wordt de aanwezigheid van patiëntkenmerken en risicofactoren gescoord, en wordt een bloedtest voor trombose (een D-dimeer test) gedaan. Dit classificeert patiënten in een laag of hoog risico op VTE en afhankelijk hiervan wordt het advies gegeven om de patiënt wel of niet te verwijzen voor verder onderzoek. Wanneer huisartsen een beslisregel voor VTE gebruiken, leidt dit tot minder doorverwijzingen waarbij er nauwelijks diagnoses gemist worden.[1-3] Maar, ondanks het gebruik van een beslisregel wordt bij slechts 25-30% van de patiënten die door de huisarts zijn verwezen, uiteindelijk de diagnose DVT of longembolie gesteld. Er is dus ruimte voor verbetering rondom het diagnostische dilemma van VTE. Het doel van dit proefschrift is om hier aan bij te dragen.

DIAGNOSTISCHE DILEMMA'S BIJ VTE

Vertraging in de diagnostiek van longembolie komt regelmatig voor, maar het is niet bekend hoe vaak dit precies voorkomt en om hoeveel dagen het gaat, en welke patiënten meer risico lopen op deze diagnostische vertraging. In **Hoofdstuk 1** beschrijven we de bevindingen van een systematische review en meta-analyse van diagnostische vertraging bij longembolie. We hebben de wetenschappelijke literatuur systematisch onderzocht en 10.933 studies gevonden, waarvan er uiteindelijk 24 voldeden aan onze inclusiecriteria en dus konden worden meegenomen in de uiteindelijke analyse. We berekenden een gemiddelde schatting van de duur van het begin van de klachten van een patiënt tot het stellen van de diagnose longembolie; dit bleek 6,3 dagen te zijn. Het percentage patiënten met meer dan zeven dagen diagnostische vertraging varieerde in de studies tussen de 18% en 38%. Diagnostische vertraging bij longembolie kwam vaker voor bij hoesten, chronische longziekte en/of hartfalen, terwijl het minder vaak voorkwam bij een recente operatie en hypotensie, en in de meeste studies ook pijn op de borst.

Hoewel beslisregels in de diagnostiek van VTE al veel gebruikt worden, is het niet bekend wat de veiligheid en efficiëntie, maar ook de redenen en gevolgen van *onjuiste* toepassing van beslisregels zijn in de dagelijkse praktijk. In **Hoofdstuk 2** evalueerden wij de impact van het dagelijkse gebruik van beslisregels voor zowel longembolie als DVT in de eerstelijnszorg. Wij voerden een cohortstudie uit met patiënten waarbij de huisarts dacht aan DVT of longembolie. De uitkomsten waren het faalpercentage (het percentage gemiste diagnoses) en de efficiëntie (het percentage patiënten dat niet verwezen werd voor vervolgonderzoek), alsmede determinanten voor, en de gevolgen van onjuiste toepassing van de beslisregels. Van de 1.477 patiënten in de studie hadden uiteindelijk 267 (18,1%) een VTE. Bij correct gebruik van de beslisregels, werd 1,5% van de diagnoses gemist, met een efficiëntie van 58,1%. In 23,0% van de patiënten werd de beslisregel onjuist toegepast, wat resulteerde in een hoger faalpercentage van 3,3% en bovendien een lagere efficiëntie van 35,7%. Bij patiënten die ook hartfalen hadden, was de kans op onjuiste toepassing van de beslisregel ongeveer drie keer hoger. De conclusie van dit onderzoek is dat correcte toepassing van beslisregels voor VTE door huisartsen gepaard gaat met een aanvaardbaar laag faalpercentage en een hoge efficiëntie. Echter, onjuiste toepassing van beslisregels voor VTE komt in bijna een kwart van de patiënten voor en leidt tot een faalpercentage boven de internationaal aanvaarde drempelwaarde van 3%, met een aanzienlijk lagere efficiëntie.

DE DIAGNOSTISCHE WAARDE VAN 'GESTALT'

Bij de diagnostiek van longembolie vertrouwen artsen vaak op een intuïtieve inschatting van de aanwezigheid van longembolie, wat ook wel 'gestalt' wordt genoemd. Hoewel herhaaldelijk is aangetoond dat gestalt voorspellend is voor longembolie, wordt het ook bekritiseerd vanwege het gebrek aan standaardisatie en generaliseerbaarheid, en de veronderstelling dat de diagnostische waarde erg variabel is. In **Hoofdstuk 3** hebben wij de diagnostische waarde van gestalt in de diagnostiek van longembolie beoordeeld door een analyse uit te voeren in een grote dataset met gegevens van patiënten die verdacht werden van een longembolie uit eerder uitgevoerde studies (Individuele Patiënt Data Meta-Analyse, IPD-MA). Wij onderzochten de variabiliteit van gestalt, uitgedrukt als relatief risico (RR), onder andere in verschillende zorgomgevingen en in subgroepen van patiënten gebaseerd op leeftijd, geslacht, hartfalen, chronische longziekte, en items van de meest gebruikte beslisregel voor longembolie. We includeerden 20.770 patiënten uit 16 individuele studies. Er was sprake van longembolie in 28,8% van de patiënten waarbij gestalt positief was gescoord, en in 9,1% van de patiënten waarbij gestalt negatief was gescoord. Dit resulteert in een RR van 3,02: het risico op longembolie is drie keer hoger wanneer de arts gestalt positief scoort. De diagnostische waarde van gestalt varieerde tussen de individuele studies (de RRs wisselden tussen de 1,46 en 7,71), maar de waarde bleef stabiel in de geanalyseerde subgroepen. Onze studie toont dus aan dat het gestalt item dezelfde waarde heeft in alle typen patiënten en in alle zorgomgevingen.

DE YEARS BESLISREGEL IN DE EERSTELIJNSZORG

Recent is een nieuwe beslisregel voor longembolie ontwikkeld en gevalideerd in de tweedelijnszorg: de YEARS strategie.[4,5] Deze beslisregel begint met de beoordeling van drie punten (YEARS-items) door de arts: (i) zijn er klinische tekenen van DVT, (ii) is er sprake van bloed ophoesten, en (iii) is longembolie de meest waarschijnlijke diagnose. Daarnaast wordt bij alle verdachte patiënten een D-dimeer bloedtest uitgevoerd. Bij patiënten waarbij de YEARS-items afwezig zijn, is de te gebruiken D-dimeer afkapwaarde 1000 ng/mL, terwijl bij patiënten waarbij een of meer YEARS-items aanwezig zijn de afkapwaarde 500 ng/mL is. Bij patiënten met een D-dimeer waarde onder een van deze afkapwaarden is een longembolie voldoende uitgesloten, maar patiënten met een D-dimeer waarde boven een van deze afkapwaarden moeten verwezen worden voor vervolgonderzoek, vaak een CT-scan van de longen. Uit

eerder onderzoek bleek dat wanneer de YEARS strategie in het ziekenhuis wordt toegepast, het percentage patiënten bij wie geen vervolgonderzoek nodig was (de efficiëntie) hoger was dan wanneer een vaste D-dimeer afkapwaarde van 500 ng/mL werd gebruikt. Belangrijk is dat het aantal gemiste diagnoses laag bleef (0,61%). Tegenwoordig wordt deze strategie daarom in ziekenhuizen in Nederland toegepast bij patiënten die verdacht worden van een longembolie. Het lijkt logisch om de YEARS strategie ook direct door huisartsen in de eerstelijnszorg te laten gebruiken gezien de hogere efficiëntie met een nog steeds laag faalpercentage. Validatie in de eerstelijnszorg is echter nodig omdat het niet zeker is dat deze beslisregel ook goed werkt in deze andere setting met een lagere vooraf kans op longembolie, in andere patiënten met relatief minder ernstige longembolieën, en bij andere artsen (huisartsen) die minder vaak een patiënt met longembolie zien. Daarom hebben wij een prospectieve diagnostische studie opgezet waarin patiënten die mogelijk een longembolie hebben door hun huisarts werden behandeld volgens de YEARS strategie: de PECAN-studie. **Hoofdstuk 4** beschrijft de rationale en opzet van de PECAN-studie.

In **Hoofdstuk 5** presenteren wij de tussentijdse resultaten van de PECAN-studie. Vanwege het achterblijvend aantal geïncludeerde patiënten in de studie hebben we twee extra inclusie-routes toegevoegd aan de PECAN-studie: 1) een route voor een lerend gezondheidszorgsysteem (LHS), en 2) hergebruik van bestaande data van patiënten verdacht van longembolie in de eerstelijnszorg. De tussentijdse analyse van de PECAN studie omvatte 482 patiënten waarbij de huisarts dacht aan longembolie. In totaal werden 376 patiënten door de YEARS strategie geclassificeerd als ‘laag risico op longembolie’ (297 zonder YEARS-items en een D-dimeer <1000 ng/mL, en 79 met ≥1 positief YEARS-item en een D-dimeer <500 ng/mL), hetgeen resulteerde in een efficiëntie van 79%. Van deze patiënten bleken er vier toch een longembolie te hebben, wat resulteerde in een faalpercentage van 1,06%. Deze tussentijdse analyse bevestigt dat een laag-risico schatting volgens de YEARS strategie de diagnose longembolie veilig en efficiënt kan uitsluiten bij patiënten in de eerstelijnszorg. Om nog meer robuuste resultaten te verkrijgen van de uitkomsten van de YEARS strategie in verschillende subgroepen van patiënten in de eerstelijnszorg, worden er tot op heden patiënten geïncludeerd via de nieuwe LHS-route.

HET UITVOEREN VAN DIAGNOSTISCH ONDERZOEK IN DE EERSTELIJNSZORG

Tijdens de uitvoering van de in dit proefschrift beschreven studies werd duidelijk dat wetenschappelijk onderzoek naar VTE in de eerstelijnszorg een uitdaging is vanwege de hoge werkdruk voor huisartsen, de tijdrovende onderzoeksprocedures en het weinig voorkomen van VTE. Deze factoren dragen allemaal bij aan een trage inclusiesnelheid van patiënten, zoals we inderdaad hebben waargenomen in de in Hoofdstuk 5 beschreven PECAN-studie. In **Hoofdstuk 6** (General Discussion) bespreken we hoe en waarom de eerder genoemde nieuwe LHS-route van de PECAN-studie de werving van patiënten voor diagnostische studies in de dagelijkse eerstelijnszorg vergemakkelijkt. De LHS-route kan ook worden gebruikt in toekomstige studies naar aandoeningen die weinig voorkomen en waarin snel gehandeld moet worden. Deze nieuwe manier van onderzoek doen, helpt om voldoende patiënten in een wetenschappelijk onderzoek te kunnen includeren en maakt het uitvoeren van diagnostische studies in de eerstelijnszorg eenvoudiger en minder tijdrovend.

REFERENTIES

1 Geersing GJ, Erkens PMG, Lucassen WAM, *et al*. Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in Primary care: Prospective cohort study. *BMJ* 2012;345:1–10. doi:10.1136/bmj.e6564

2 Buller HR, Cate-hoek AJ, Hoes AW, *et al*. Safely Ruling Out Deep Venous Thrombosis in Primary Care. *Ann Intern Med* 2009;150:229–36. doi:10.7326/0003-4819-150-4-200902170-00003

3 Righini M, Van Es J, Den Exter PL, *et al*. Age-Adjusted D-Dimer Cutoff Levels to Rule Out Pulmonary Embolism. *JAMA* 2014;311:1117. doi:10.1001/jama.2014.2135

4 Van Es J, Beenen LFM, Douma RA, *et al*. A simple decision rule including D-dimer to reduce the need for computed tomography scanning in patients with suspected pulmonary embolism. *J Thromb Haemost* 2015;13:1428–35. doi:10.1111/jth.13011

5 van der Hulle T, Cheung WY, Kooij S, *et al*. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet* 2017;390:289–97. doi:10.1016/S0140-6736(17)30885-1

Dankwoord

Ruim vijf jaar geleden begon ik aan mijn promotietraject en nu is mijn proefschrift af! Deze mijlpaal had ik niet kunnen behalen zonder alle mensen die mij deze jaren hebben geholpen. Ik wil hen dan ook enorm bedanken.

Allereerst wil ik mijn promotieteam bedanken: mijn promotoren prof. dr. Rutten en prof. dr. Moons, en mijn copromotoren dr. Geersing en dr. Blom.

Beste Frans, toen jij hoogleraar werd, kreeg ik er een promotor bij. Ik ben blij dat jij in die rol deel hebt uitgemaakt van mijn promotieteam. Bedankt voor je goede adviezen, duidelijke feedback en praatjes op de gang. Ondanks je drukke agenda had ik nooit het gevoel dat je druk was. Je was altijd betrokken, zowel bij het onderzoek als ook op persoonlijk vlak.

Beste Carl, bedankt voor je enthousiasme en kritische blik. Ook al was je de laatste tijd meer op afstand betrokken, je aanvullingen, inspiratie en nieuwe ideeën waren altijd waardevol. Met jouw opbouwende feedback werd ieder stuk beter.

Beste Geert-Jan, jouw gedrevenheid en passie voor het trombose-onderzoek (en in het bijzonder je favoriete biomarker D-dimeer) werken aanstekelijk. Ik bewonder je wetenschappelijke ambitie en onuitputtelijke ideeën voor nieuwe projecten. Als ik even door de bomen het bos niet meer zag, maakte een overleg met jou dit weer goed. Bedankt voor je vertrouwen en begeleiding.

Beste Jeanet, als copromotor uit Leiden was jouw bijdrage in de projecten altijd van meerwaarde. Aan het einde van mijn traject heb je mij enorm geholpen bij het nabellen van potentiele PECAN-patiënten. Dankzij jou hebben we heel wat meer inclusies kunnen behalen!

Graag wil ik ook alle leden van de beoordelingscommissie, prof. dr. Damoiseaux, dr. Koek, prof. dr. Middeldorp, prof. dr. van der Schouw, en prof. dr. Schutgens, bedanken voor het kritisch lezen en beoordelen van mijn proefschrift.

Bedankt collega's van het 'tromboseteam', Geert-Jan, Frans, Sander, Maarten, Carline, Linda, Emmy, Florien en Hannah. Ik ben blij dat ik deel uit mocht maken van deze leuke groep. Ik heb genoten van de samenwerking, gezamenlijke projecten, en natuurlijk de congressen in Berlijn, Krakau, Melbourne en Londen. Lieve Emmy, naast een fijne samenwerking met als resultaat een gezamenlijke review, hebben we ook veel lol gehad. Of het nou was omdat je een 'likelihoedje' op had, we per ongeluk in dezelfde

kleren op werk verschenen, of kletsten over de fratsen van onze dochters. Bedankt dat jij straks naast mij staat als paranimf.

Het succes van de PECAN studie is natuurlijk mede te danken aan alle huisartsen en patiënten die mee wilden doen met het onderzoek. Bedankt voor jullie inzet, samenwerking en deelname! Beste Nelly, bedankt voor al je hulp met het bijhouden van contacten met huisartsen, het versturen van de follow-up brieven en de monitorvisites. Beste Melchior, met jouw hulp is de nieuwe inclusieroute van de PECAN studie een enorm succes geworden. Bedankt voor je enthousiasme en gedrevenheid, ik ging altijd met plezier naar Amsterdam om brieven te versturen naar PECAN-patiënten. Beste Daan, wat fijn dat jij de praktische taken voor de PECAN studie van mij hebt kunnen overnemen. Ik heb er alle vertrouwen in dat de studie met jouw hulp goed zal worden afgerond!

Beste collega's uit Leiden, bedankt voor de fijne samenwerking! Menno Huisman en Erik Klok, bedankt voor jullie betrokkenheid vanuit de tweedelijnszorg rondom de PECAN studie. Emily, ik heb met plezier samengewerkt en hoop op een mooie publicatie binnenkort!

Beste Kim, ons gezamenlijke IPD-MA project was het laatste af te ronden hoofdstuk van mijn proefschrift. Enorm bedankt voor je inzet en hulp bij alle (voor mij) ingewikkelde analyses met soms strakke deadlines.

Lieve collega's van 6.101, ik kijk met heel veel plezier terug op mijn tijd in de leukste kamer van het Stratenum. Bedankt voor de cake-starten, lunchwandelingen in de Botanische Tuinen en alle leuke uitjes. Escaperooms waren favoriet, we hebben er zelfs één met succes ontworpen in onze kamer!

Beste oud-collega's van huisartsenpraktijk Sagenhoek, beste Chantal en Anne To, wat fijn dat ik het laatste jaar van de huisartsopleiding met jullie als opleiders heb kunnen doorlopen. Ik denk nog regelmatig terug aan de leerzame, maar ook gezellige tijd in Amersfoort. Zonder jullie zou ik niet de huisarts zijn die ik nu ben.

Lieve Anna, Anne, Danielle en Merel. Ook al zie ik jullie niet meer zo vaak als vroeger op de WP, ik vind het bijzonder dat we na al die jaren nog steeds de 'vette mensen' zijn. Bedankt voor de gezellige etentjes, jullie betrokkenheid en humor. Op naar nog veel meer jaren!

Lieve ploeggenoten van G&D, lieve Anne-Mieke, Heleen, Lisanne, Rianne, Selma en Sofie, wat fantastisch dat wij bijna 15 jaar geleden bij elkaar in een roeiploeg terecht kwamen en dat dit is uitgegroeid tot een hechte vriendschap. We matchen natuurlijk goed vanwege onze 'golden' en 'delicious' looks, maar ik ben ook heel dankbaar voor jullie interesse, humor en gezelligheid. Het is altijd een feest om jullie te zien!

Lieve Suzanne en Mariese, bedankt voor de fijne studententijd die wij samen hebben gehad! Lieve Mariese, wat leuk dat we naast vriendinnen nu ook collega-huisartsen zijn. De eerste gezamenlijke nascholingen hebben we al gevolgd, hopelijk volgen er nog meer leuke, ook niet werk-gerelateerde, uitjes samen met jou. Bedankt dat jij mijn paranimf wil zijn!

Lieve Marieke, twee jaar geleden zaten we in hetzelfde zwangerschaps-clubje. Wat ben ik blij dat dit is uitgegroeid tot een vriendschap! Bedankt voor je interesse en gezelligheid. Ik kijk uit naar alle vrije woensdagen samen met onze meiden.

Lieve Marlijn, als nicht ken ik je al mijn hele leven, maar ik heb ook enorm genoten van onze studententijd waarin we huisgenoten waren in het Verwende Nest. Ook al zijn we allebei naar een andere kant van het land verhuisd, ik hoop dat we elkaar nog vaak zullen gaan zien!

Lieve schoonfamilie, lieve Harry, Anneke, Patricia, Michael, Mees, Eveline en Derek, jullie zorgen ervoor dat Noordwijk nog meer als mijn thuis voelt. Bedankt voor jullie betrokkenheid, gezellige familiebijeenkomsten en middagen op het strand. Wat fijn dat jullie zo dichtbij wonen en altijd klaar staan om te helpen; van klusjes in huis tot het wekelijks oppassen op Lize. Bedankt!

Lieve papa en mama, wat bof ik met zulke lieve ouders! Jullie onvoorwaardelijke steun en vertrouwen in mij hebben mij gebracht waar ik nu ben. Ondanks dat mijn promotieonderzoek soms wat abracadabra was (is het nou een pee-ha-dee?), zijn jullie altijd betrokken en geïnteresseerd. Bedankt dat ik alle leuke momenten, maar ook tegenslagen en uitdagingen, met jullie kan bespreken. Ik vind het heel mooi om te zien dat jullie nu ook zo'n trotse opa en oma zijn en ik geniet ervan om dit met jullie te kunnen delen. Lieve Rogier, ik vind het knap dat jij altijd kan denken 'het komt wel goed' en zou soms willen dat ik dit ook wat meer had. Wat fijn om zo'n nuchtere, maar ook lieve broer met veel humor te hebben.

Lieve Jeroen, wat hebben wij tijdens mijn promotietraject samen veel mooie dingen beleefd. Een reis naar Australië, verhuizen naar Noordwijk, de geboorte van Lize en in mei een volgende bijzondere gebeurtenis om naar uit te kijken! Dankjewel voor al je liefde, humor, betrokkenheid en luisterend oor. Het hielp mij altijd om met jou te sparren over waar ik in het onderzoek mee bezig was en tegenaan liep, waarbij je vaak goede kritische vragen stelde die mij aan het denken zetten. Ik ben dankbaar dat ik mijn leven met jou kan delen en verheug me op de toekomst samen met jou!

Lieve Lize, wat ben ik blij dat jij er bent! Al kun je dit nu nog niet lezen en vind je het later misschien te saai, ik wil je bedanken voor je vrolijkheid, enthousiasme en lieve kusjes en knuffels. Ik kijk uit naar nog veel meer mooie momenten samen. Lieve baby, ook al ben je er nog niet echt bij, je hebt nu al geholpen met het afronden van mijn proefschrift door een hele duidelijke deadline te stellen. Ik kan niet wachten om je te ontmoeten.

About the author

ABOUT THE AUTHOR

Rosanne van Maanen was born on July 5th 1989 in Utrecht, the Netherlands. After graduating from secondary school 'De Werkplaats' in Bilthoven in 2007, she studied Psychobiology at the University of Amsterdam (UvA). In 2008 she started to study medicine at the Academic Medical Centre/ UvA, where she obtained her medical degree in December 2014.

From 2015-2016 she worked as a resident in Internal Medicine at the Tergooi Hospital in Hilversum and Blaricum. In March 2016 she started her training in General Practice. Since September 2017 she combined this training with a PhD trajectory at the Julius Center for Health Sciences and Primary Care of the University Medical Center Utrecht/ Utrecht University, under the supervision of prof. dr. F.H. Rutten, prof. dr. K.G.M. Moons, dr. G.J. Geersing and dr. J.W. Blom, which resulted in the articles presented in this thesis. During her PhD trajectory, she followed the postgraduate master Clinical Epidemiology at Utrecht University, which she finished in 2020. In September 2021 she completed her training in General Practice.

Rosanne lives in Noordwijk with her partner Jeroen and their daughter Lize (2021) and currently works as a general practitioner at different primary care practices in the 'Bollenstreek'.