PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Real-life impact of clinical prediction rules for venous thromboembolism in primary care: a cross-sectional cohort study
AUTHORS	van Maanen, Rosanne; Kingma, Annelieke; Oudega, Ruud; Rutten, Frans; Moons, Karel; Geersing, Geert-Jan

VERSION 1 – REVIEW

REVIEWER	A Delluc Department of Medicine The Ottawa Hospital Ottawa, ON Canada
REVIEW RETURNED	21-Jul-2020

GENERAL COMMENTS	In this paper, the authors described application of CPRs to safely rule-out venous thromboembolism (VTE) in primary care. They confirmed that following CPRs is the safest management of suspected deep vein thrombosis or pulmonary embolism. This is a valuable confirmation of what is already known from VTE management in the emergency room. Here are the specific issues I would suggest the authors to comment in their manuscript: 1) there is no sample size calculation a-priori 2) workup for VTE is usually considered safe if the upper limit of the 95% CI around the 3-month incidence of VTE in patients in whom VTE is initially ruled out and who are left untreated is lower than 3%. Please comment on that in the manuscript 3) please describe the location of DVTs (i.e. how many were distal, proximal). Were there only DVT counted in the outcomes (no isolated superficial vein thrombosis?) 4) discuss the impact of patients lost to follow-up

REVIEWER	Paolo Ventura University of Modena and Reggio Emilia
REVIEW RETURNED	20-Aug-2020

GENERAL COMMENTS	The paper by van Maanen et al. aims at determining the real-life impact of use of VTE (DVT and PE) Clinical prediction rules (CPRs) for identification of suspected VTE in primary care setting. The study is interesting, because different score and studies have been developed in order to evaluate the role and the efficiency of different CPRs for DVT and PE in the setting of hospitalized patients, but only scarce data are available about CPRs' efficiency

	in primary care settings. Moreover, also 'real-life'-derived data on CPR use in primary care setting are scarce. In their study, the authors aim also at assessing the diagnostic failure rate together with the main determinants for, and consequences of incorrect application of the CPRs by the primary care physicians. The paper is interesting but some issues should be addressed before considering it for publication: Major issues - Many VTE (DVT/PE) CPRs have been proposed so far. Most of them have been built and validated for inpatients. Only few are validated and available for outpatients. A Wells' score CPR is available in a form for DVT and in a form for PE for outpatients. Considering the authors used the Wells score as CPR for PE what is the reason why they did not use the Wells score form for DVT, as well, instead of Oudega CPR ? - Considering that Oudega CPR seem to fail in significant rate in frail patients (i.e.aged patients) (Fam Pract 2015 32:120-125), it would be important to have data about how many patients older than 80 years were included into the study (or, otherwise, if different results in failure rate were observed in more aged patients). - Besides the values of efficiency rates (always higher in subjects having a correct CPR evaluation) and failure rates (always lower in subjects having a correct CPR evaluation) showing an advantage in subjects with correct CPR evaluation, if we consider (fig.2) DVT patients, we notice that , within referred patients, VTE+ or VTE- patients have similar rates in both groups (CPR correctly used vs. CPR incorrectly used). This data should be commented. - A comment should be added regarding the overall results: are efficiency values (especially when considering overall results or DVT results, or PE results for CPR incorrect use) really acceptable ? if yes why ? Is 50% of efficiency a too low value? Minor issues -In table 1 page 9 probably CDR is wrong. Is it CPR ? -In table 1 page 9 a percentage of patients with d-dimer higher than reference range (out of total patients where it was measured) should be added. A percentage of patients with CPR values also may be added (or, at least, a percentage ofpatients with CPR suggesting different risk levels) -Figure 1 should be somehow improved, for its lecture is not so immediate. For example, in the text (section failure rate and efficiency of CPRs, page 10) (first lines) the authors report the failure rate of both combined CPRs about 1.8% and efficiency about 53% referring to figure 1, but in figure 1 these data are not reported, probably because they derive from aggregation of correct/incorrect values. -Figure 2 VTE+ in not referred (failure in Correct DVT with CPR correctly used) is reported 2%, better to indicate 1.96% in order to maintain the same values reported in fig.1 -Figure 3 In the six last squares at the end of the figure, VTE+ or VTE- is reported. Is it an error for PE ?

VERSION 1 – AUTHOR RESPONSE

Comments to reviewers:

Reviewer 1: Aurelien Delluc, The Ottawa Hospital, Department of Medicine, Ottawa, Canada

GENERAL COMMENTS: In this paper, the authors described application of CPRs to safely rule-out venous thromboembolism (VTE) in primary care. They confirmed that following CPRs is the safest management of suspected deep vein thrombosis or pulmonary embolism. This is a valuable confirmation of what is already known from VTE management in the emergency room. Here are the specific issues I would suggest the authors to comment in their manuscript:

1). There is no sample size calculation a-priori.

Answer: Indeed, we thank the reviewer for this valuable and good comment. Calculating a sample size for diagnostic validation studies, like ours, unfortunately is not straightforward. In fact, methods to derive a valid sample size for prediction studies have only recently been proposed, i.e. after the initiation of our study.[1] As such we did not perform a sample size calculation a-priori for this study. Nevertheless, we agree with this reviewer that this topic does deserve more attention. Our total dataset includes 1,477 patients suspected of VTE with 268 VTE outcome events (230 DVT; 38 PE), and as such we believe that would allow for robust and valid statistical inferences. Prompted by this reviewers' comment, we now elaborated on this topic in the discussion, section of 'strengths and limitations'. We added the following sentences:

'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.'

2) Workup for VTE is usually considered safe if the upper limit of the 95% CI around the 3-month incidence of VTE in patients in whom VTE is initially ruled out and who are left untreated is lower than 3%. Please comment on that in the manuscript.

Answer: This is a very good point. Indeed, for some sub-analyses, the 95% CI crosses the border of 3.0%, which typically is the agreed safety margin for excluding VTE. Nevertheless, in our view our main inference remains valid, namely: that in those in whom CPRs are correctly applied this appears to be safe, and subsequently that the failure rate becomes unacceptably high in those in whom the CPRs are incorrectly applied, i.e. increasing from 1.51% (95% CI 0.77 to 2.86) to 3.31% (95% CI 1.07 to 8.76). Nevertheless, we agree that this is an important topic, relating we believe to two topics: (i) indeed the sample size in particular for the sub-analyses stratified for DVT and PE (see also the previous question), and (ii) the higher likelihood for missing VTE in patients where a Clearview Simplify point-of-care (POC) assay was used. Both topics have now been discussed in more detail in the limitation section of the manuscript, highlighting the uncertainty for these sub-analyses as well as the problems related to the Clearview Simplify POC D-dimer assay.

3) Please describe the location of DVTs (i.e. how many were distal, proximal). Were there only DVT counted in the outcomes (no isolated superficial vein thrombosis?)

Answer: Thank you for this comment. It would be interesting to describe the location of DVTs but unfortunately we do not have these data. This was a real-life study, and we used clinical data on how patients were actually managed in primary care. The participating GPs filled in a follow-up form after three months with questions about the final diagnosis and treatment. They only filled in whether there was a diagnosis of DVT (or PE) or not, without further details. Hence, our definition of the outcome was based on these forms and not on the radiology report or hospital records. Isolated superficial vein thrombosis, however, was not counted in our outcomes.

4) Discuss the impact of patients lost to follow-up.

Answer: of the total 1509 enrolled suspected DVT and PE patients in our study, only 32 (2.1%) were lost to follow-up. The baseline characteristics of the included patients and the patients lost to follow-up were comparable. Therefore, we can state that this loss of follow-up information is missing completely at random (MCAR) and thus will not result in selection bias. Moreover, current methodological guidance do not recommend imputing these missing values under the MCAR assumption. We added a sentence for clarification in our manuscript on page 9.

Reviewer 2: Paolo Ventura, University of Modena and Reggio Emilia

The paper by van Maanen et al. aims at determining the real-life impact of use of VTE (DVT and PE) Clinical prediction rules (CPRs) for identification of suspected VTE in primary care setting. The study is interesting, because different score and studies have been developed in order to evaluate the role and the efficiency of different CPRs for DVT and PE in the setting of hospitalized patients, but only scarce data are available about CPRs' efficiency in primary care settings. Moreover, also 'real-life'-derived data on CPR use in primary care setting are scarce. In their study, the authors aim also at assessing the diagnostic failure rate together with the main determinants for, and consequences of incorrect application of the CPRs by the primary care physicians.
The paper is interesting but some issues should be addressed before considering it for publication:

We thank the reviewer for his positive remarks on our paper; below we will address the remaining issues for further clarification.

Major issues:
1). Many VTE (DVT/PE) CPRs have been proposed so far. Most of them have been built and validated for inpatients. Only few are validated and available for outpatients. A Wells' score CPR is available in a form for DVT and in a form for PE for outpatients. Considering the authors used the Wells score as CPR for PE what is the reason why they did not use the Wells score form for DVT, as well, instead of Oudega CPR ?

Answer: This indeed is very good question. The main reason for this is that we previously demonstrated that the Wells CPR for suspected DVT was shown to be less suitable for use in primary care medicine.[2] In fact, the point estimate of missed DVT cases in patients with a low Wells score and a negative D-dimer was 2.9%. We now added this for clarification in the Methods section.

2). Considering that Oudega CPR seem to fail in significant rate in frail patients (i.e. aged patients) (Fam Pract 2015 32:120-125), it would be important to have data about how many patients older than 80 years were included into the study (or, otherwise, if different results in failure rate were observed in more aged patients).

Answer: Very good point. The reference where the reviewer is referring to nevertheless is a study we previously performed explicitly and exclusively in nursing home patients, thus in institutionalized frail elderly patients. In fact, this is not only true for the Oudega CPR in suspected DVT, but also for the Wells CPR in suspected PE in that setting.[3] Such patients were not included in this current study, which exclusively included patients seen in community healthcare clinics in the Netherlands in an outpatient setting. We now made this more clear in the Methods section. Yet, we do agree that information related to the age of patients with a missed VTE diagnosis by any of the CPRs and D-dimer is important. In the supplementary table 2 we described the clinical characteristics of the non-referred patients with VTE after three months of follow-up. The mean age of this patient group was 64.9 years (suspected DVT patients 65.3 years, suspected PE patients 62.0 years). Hence, in our datawe do not see a clear relationship between an increasing age and a higher proportion of missed VTE cases.

3). Besides the values of efficiency rates (always higher in subjects having a correct CPR evaluation) and failure rates (always lower in subjects having a correct CPR evaluation) showing an advantage in subjects with correct CPR evaluation, if we consider (fig.2) DVT patients, we notice that , within referred patients, VTE+ or VTE- patients have similar rates in both groups (CPR correctly used vs. CPR incorrectly used). This data should be commented.

Answer: Indeed, we see that in de patients suspected of having DVT, the proportion of referred patients actually having a DVT after follow-up is comparable in both the “CPR correctly used group” and the “CPR incorrectly used group” (41.4% versus 40.9%). However, this finding does not inform us about the efficiency of the CPR as defined in our study. Moreover, we feel that being not referred based upon an incorrect use of the CPR may be more important clinically then being referred despite incorrect CPR use. This is, the clinical consequence of incorrect CPR use in our view is more prominent if the consequence is non-referral where referral should have been indicated. Prompted by this remark, we changed our sentences in the section of clinical implications now as follows:

“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.”

4). A comment should be added regarding the overall results: are efficiency values (especially when considering overall results or DVT results, or PE results for CPR incorrect use) really acceptable ? if yes why ? Is 50% of efficiency a too low value?

Answer: We agree that this is an important topic to discuss. We do believe that the ability to rule-out VTE in a community healthcare setting, thus without referring all patients to hospital clinics, would be considered highly attractive for many GPs. To elaborate more on this topic, we added the following sentences to the paragraph of clinical implications, in the Discussion:

‘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 outpatients, community healthcare setting.’

Minor issues:
5). In table 1 page 9 probably CDR is wrong. Is it CPR ?

Answer: We agree, this indeed should be CPR. We now changed it accordingly and thank the reviewer for pointing this out.

6). In table 1 page 9 a percentage of patients with d-dimer higher than reference range (out of total patients where it was measured) should be added. A percentage of patients with CPR values also may be added (or, at least, a percentage of patients with CPR suggesting different risk levels).

Answer: Thank you for this suggestion. Of the 832 patients suspected of having DVT in whom a D-dimer test was performed (either with a dichotomous or dichotomous outcome) 354 (42.5%) had a D-dimer result higher than the reference range (either a 'positive' result or a D-dimer above 500ng/mL). Of the 451 patients suspected of PE in whom a D-dimer test was performed, 105 (23.3%) had a D-dimer result higher than the reference range. In the DVT suspected group 171 patients (17.2%) had a score on the CPR above 3 and in the PE suspected group 49 patients (10.1%) had a score on the CPR above 4, classifying them in the 'likely' risk category. We added this information to table 1.

7). Figure 1 should be somehow improved, for its lecture is not so immediate. For example, in the text (section failure rate and efficiency of CPRs, page 10) (first lines) the authors report the failure rate of both combined CPRs about 1.8% and efficiency about 53% referring to figure 1, but in figure 1 these data are not reported, probably because they derive from aggregation of correct/incorrect values.

Answer: We agree that the percentages as mentioned in the first paragraph of the result section can be somewhat confusing, since they do not correspond with the percentages as mentioned in figure 1. Indeed, figure 1 describes the failure rate and efficiency of both CPRs combined, but split up for correct and incorrect CPR use. We deleted the reference to figure 1 in the first sentence and added a sentence in the paragraph for clarification. This section now reads as follows:

"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)

8). Figure 2 VTE+ in not referred (failure in Correct DVT with CPR correctly used) is reported 2%, better to indicate 1.96% in order to maintain the same values reported in fig.1.

Answer: Thank you for this comment. We changed the percentage to 1.96% in figure 2. Furthermore, we did also add the two decimal numbers for the failure rate in the incorrect DVT group (figure 2) and for the failure rate of the correct PE group (figure 3).

9). Figure 3 In the six last squares at the end of the figure, VTE+ or VTE- is reported. Is it an error for PE?

Answer: The outcome definition in our study was the number of VTE diagnoses in patients suspected of having DVT or PE. Thus, we did define DVT and PE diagnoses to a combined outcome of VTE. The squares in figure 3 indicating VTE + and VTE – are therefore not an error.

VERSION 2 – REVIEW

REVIEWER	Aurelien Delluc The Ottawa Hospital Canada
REVIEW RETURNED	17-Nov-2020

GENERAL COMMENTS	The authors have replied to my comments according

REVIEWER	Paolo Ventura University of Modena and Reggio Emilia Italy
REVIEW RETURNED	01-Nov-2020

GENERAL COMMENTS	None more