Diagnostic classification in patients with suspected deep venous thrombosis: physicians’ judgement or a decision rule?

Geert-Jan Geersing, Kristel J Janssen, Ruud Oudega, Henk van Weert, Henri Stoffers, Arno Hoes and Karel Moons on behalf of the AMUSE Study Group

ABSTRACT

Background
Clinical decision rules can aid in referral decisions for ultrasonography in patients suspected of having deep venous thrombosis (DVT), but physicians are not always convinced of their usefulness and rely on their own judgement.

Aim
To compare the performance of a clinical decision rule with the probability of DVT presence as estimated by GPs.

Design of study
Cross-sectional survey.

Setting
Primary care practices in The Netherlands.

Method
GPs (n = 300) estimated the probability of the presence of DVT (range 0–100%) and calculated the score for the clinical decision rule in 1028 consecutive patients with suspected DVT. The clinical decision rule uses a threshold of three points and so, for the GP estimates, thresholds were introduced at 10% and 20%. If scores were below these estimates, it was not considered necessary to refer patients for further examination. Differences between the clinical decision rule and the GP estimates were calculated; this is discrimination (c-statistic) and classification of patients.

Results
Data of 1002 patients were eligible for analysis. DVT was observed in 136 (14%) patients. Both the clinical decision rule and GP estimates had good discriminative power (c-statistic of 0.80 and 0.82 respectively). Fewer patients were referred when using the clinical decision rule compared with a referral decision based on GP estimates: 51% versus 79% and 65% (thresholds at 10% and 20% respectively). Both strategies missed a similar and low proportion of patients who did have DVT (range 1.4–2.0%).

Conclusion
In patients suspected of DVT both GP estimates and a clinical decision rule can safely discriminate in patients with and without DVT. However, fewer patients are referred for ultrasonography when GPs rely on a clinical decision rule to guide their decision making.

Keywords
clinical decision rules; clinical gestalt; decision support systems, clinical; deep venous thrombosis; primary care; probability estimations.

INTRODUCTION

In diagnostic research, the development and validation of clinical decision rules has become increasingly popular.1-5 Diagnostic clinical decision rules often combine signs, symptoms, and additional laboratory or imaging test results to estimate the probability of presence (or absence) of a certain disease or to estimate a score that is subsequently related to this disease probability. In order to enhance decision making, a threshold in the probability, or score, is sometimes introduced to categorise patients into groups of, for example, low, moderate, or high risk of having the disease.

Diagnostic clinical decision rules are developed for physicians to objectivate and guide their decision making and reduce unnecessary patient burden and costs.6 Well-known examples are the Ottawa ankle rules,7 and rules for the diagnosis of deep venous thrombosis (DVT)8,9 and pulmonary embolism.10,11 For example, the use of the Ottawa ankle rules to decide...
on further diagnostic management leads to a relative reduction in ankle radiography of around 25%. The Wells rule for excluding DVT can safely reduce the number of referrals for further testing by 30–50% in a secondary-care setting.23

Despite these positive results, a number of obstacles can impede the use of clinical decision rules in daily practice. Several authors have described barriers for effective use of them, varying from a physician’s disdain for ‘cookbook medicine’ or the impractical use of the decision rule instrument itself in busy daily clinical practice.4,14,15 One of the important barriers is ‘the conviction that clinical judgement is superior to the decision rule’. Indeed, some studies report superiority of physicians’ judgement above a formal clinical decision rule.16–19 However, subjective estimation by physicians can also be inaccurate, resulting in both systematic overestimation and underestimation.20,21 In addition, physicians in daily practice, who are provided with risk-calculation tools, sometimes use them inappropriately.22

Although many clinical decision rules have been developed to diagnose DVT, there are almost no studies that compare the performance of physicians’ estimates or empirical assessment with that of a formal clinical decision rule. Current studies available showed that, although both empirical assessment of physicians and assessment of a clinical decision rule had similar overall accuracy in diagnosing DVT, concordance between both methods was poor.23,24 For physicians working in primary care, however, either strategy — empirical assessment or a formal clinical decision rule — will be used to guide patient management.

It is important to identify patients at low risk of developing DVT so it can be excluded without referral for imaging tests. This is particularly pertinent as most patients suspected of having DVT are not diagnosed and accessibility of imaging tests is often poor in primary care. Discrepancies between both strategies regarding the possibility of excluding DVT without imaging are currently not known. Perhaps, as a consequence, many physicians in daily practice tend to trust their own clinical judgement instead of a formal clinical decision rule when they have to decide whether or not to refer those patients suspected of having DVT for imaging tests.

This study was carried out to learn more about these discrepancies between physicians’ probability estimates and estimations by a formal clinical decision rule. The DVT probability classification of patients suspected of having DVT by a clinical decision rule23,24 was compared with the GPs’ probability estimates in primary care. Consequently, the aim was to provide guidance for physicians working in daily practice on which approach is better to safely and efficiently exclude DVT (without referral for imaging tests) in patients suspected of having it.

METHOD

Patients, setting, and scoring tool

The present study was a secondary analysis of a recently conducted study among 1086 primary care patients suspected of having DVT. Written informed consent was obtained from all patients and the study was approved by the local ethic review boards. The design and main results of the study have been published elsewhere.25

Briefly, the aim of the main study was to quantify the safety and accuracy of a diagnostic clinical decision rule to exclude DVT when used by GPs. This clinical decision rule — the Oudega rule — was specifically derived and validated to determine the presence versus absence of DVT in primary care patients suspected of having DVT,25,26 as the Wells rule (commonly used in a hospital setting) cannot safely exclude DVT in primary care. The Oudega rule includes predictors from patient history, physical examination, and the D-dimer test result (Table 1).25

Subjects were selected from 300 GPs in The Netherlands. Suspicion of DVT was based on the

<table>
<thead>
<tr>
<th>Variable</th>
<th>Points</th>
<th>Regression coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex</td>
<td>1</td>
<td>0.59</td>
</tr>
<tr>
<td>Oral contraceptive use</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>Presence of active malignancy (within last 6 months)</td>
<td>1</td>
<td>0.42</td>
</tr>
<tr>
<td>Major surgery (last 3 months)</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>Absence of leg trauma</td>
<td>1</td>
<td>0.60</td>
</tr>
<tr>
<td>Vein distension</td>
<td>1</td>
<td>0.48</td>
</tr>
<tr>
<td>Calf swelling ≥3cm</td>
<td>2</td>
<td>1.13</td>
</tr>
<tr>
<td>Abnormal D-dimer result</td>
<td>6</td>
<td>3.01</td>
</tr>
</tbody>
</table>

The score could range from 0 to 14: ≤3 = not referred for compression ultrasonography, ≥4 = directly referred for compression ultrasonography. The score was derived from the following original logistic model: probability DVT = 1/(1+exp (-5.47+0.59*male sex + 3.01*abnormal D-dimer result)), DVT = deep venous thrombosis.
presence of at least one of the following symptoms: swelling, redness, or pain of the lower extremity. Patients were excluded if they were <18 years of age; received anticoagulant treatment (that is, vitamin K antagonists or low molecular weight heparin) at presentation; were pregnant; or were unwilling to participate.

Scoring the Oudega rule
Participating GPs systematically documented the eight predictors of the Oudega rule, and calculated the score for an individual patient according to the variables outlined in Table 1. D-dimer testing was performed using a point-of-care D-dimer assay (Clearview® Simplify D-dimer assay, Inverness Medical, Bedford, UK). The score could range from 0 to 14.

Physicians’ probability estimates of DVT presence
The GPs were also asked to estimate the probability of the presence of DVT on a visual analogue scale ranging from 0% to 100%. They could use the same variables as incorporated in the Oudega rule (including D-dimer testing) and were, therefore, not explicitly blinded to the result of the clinical decision rule.

Reference standard
For the main study, patients with a score of ≥4 on the Oudega rule were referred for compression ultrasonography; for those with a score of ≤3, no compression ultrasonography was performed and patients remained under the surveillance of their GP. All patients returned to their GP after 1 week for evaluation of their symptoms. If suspicion of DVT remained, a compression ultrasonography was performed after all. During 90 days of follow-up, symptomatic thromboembolic events — fatal pulmonary embolism, non-fatal pulmonary embolism, and DVT — were recorded for all patients, as assigned by an independent adjudication committee. Positive outcomes were DVT as diagnosed by compression ultrasonography, venous thromboembolic complications, or death due to a possible thrombo-embolic event during the 90-day follow-up period. Negative outcomes were the absence of these events during the study period.

Comparison of the Oudega rule and GPs’ probability estimates
The Oudega rule and GPs’ probability estimates were compared using various methods. To start, the overall ability to discriminate between patients with and without DVT was estimated for both tools using the area under the receiver operating characteristic (ROC) curve. A ROC area of 0.5 reflects no discriminative ability, while a perfect diagnostic tool has a ROC area of 1.0.28

Both strategies were then compared on their diagnostic classification, as the overall discrimination does not directly reflect a diagnostic’s clinical value, in terms of guiding patient management. For this purpose, the Oudega rule and the GP probability estimates were compared first on their ability to classify a group of patients as low risk. In this low-risk group, DVT is considered excluded and physicians can subsequently refrain from referring the patient for imaging. For the Oudega rule the previously developed and validated threshold of three points9,25,26 was used to classify patients into such a low-risk group. With regard to GP probability estimates, no such validated threshold exists so the probability estimates (0–10%, 10–20%, and >20%) were first trichotomised and, subsequently, thresholds at 10% and 20% (below which patients were considered low risk) were introduced.

Finally, a classification table was constructed in which differences in classification (for patients with and without DVT) between each strategy could be quantified.29,30

All statistical analyses were performed in SPSS (version 15.0 for Windows; SPSS, Chicago, US).

Table 2. Characteristics of study population (n = 1002).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years, mean (SD)</td>
<td>57.7 (17.1)</td>
</tr>
<tr>
<td>Male sex</td>
<td>375 (37%)</td>
</tr>
<tr>
<td>Leg complaints</td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>874 (87%)</td>
</tr>
<tr>
<td>Swelling</td>
<td>784 (78%)</td>
</tr>
<tr>
<td>Redness</td>
<td>371 (37%)</td>
</tr>
<tr>
<td>Absence of leg trauma</td>
<td>737 (74%)</td>
</tr>
<tr>
<td>Varicose veins/venous insufficiency</td>
<td>337 (34%)</td>
</tr>
<tr>
<td>Distension of collateral veins</td>
<td>151 (15%)</td>
</tr>
<tr>
<td>Difference in calf circumference ≥3cm</td>
<td>304 (30%)</td>
</tr>
<tr>
<td>Duration of complaints, days, median (IQR)</td>
<td>5 (3–10)</td>
</tr>
<tr>
<td>Previous episode of DVT</td>
<td>159 (16%)</td>
</tr>
<tr>
<td>Previous episode of PE</td>
<td>51 (5%)</td>
</tr>
<tr>
<td>Paresis</td>
<td>13 (1%)</td>
</tr>
<tr>
<td>Recent surgery in previous month</td>
<td>81 (8%)</td>
</tr>
<tr>
<td>Recent immobilisation</td>
<td>75 (7%)</td>
</tr>
<tr>
<td>Bed rest &gt;3 days</td>
<td>72 (7%)</td>
</tr>
<tr>
<td>Active malignancy in past 6 months</td>
<td>54 (5%)</td>
</tr>
<tr>
<td>Malignancy not treated in the past 6 months</td>
<td>64 (6%)</td>
</tr>
<tr>
<td>Use of hormonal contraceptives</td>
<td>107 (11%)</td>
</tr>
<tr>
<td>Travel (car/bus/plane) &gt;4 hours</td>
<td>90 (9%)</td>
</tr>
</tbody>
</table>

*The full sample totals 1028 patients but data on 26 were excluded as their management was not completed according to protocol. DVT = deep venous thrombosis. IQR = interquartile range. PE = pulmonary embolism. SD = standard deviation.
RESULTS
A total of 1086 patients with clinically suspected DVT entered the study, of whom 58 met one or more of the exclusion criteria. For 26 patients, management was not completed according to protocol, so data of 1002 patients (97%) were available for analysis of their characteristics (Table 2). The mean age was 58 years and 37% were males. Most common complaints were leg pain (87%) and leg swelling (78%). The overall prevalence of DVT confirmed by objective testing was 14% (136/1002 patients).

ROC analysis
As Figure 1 shows, the Oudega rule and the GP probability estimates both had good and similar overall discriminative ability in diagnosing DVT. The area under the ROC curve was 0.80 (95% confidence interval [CI] = 0.77 to 0.84) and 0.82 (95% CI = 0.78 to 0.86), respectively.

Comparison of classification
If the decision to refer a patient suspected of having DVT for imaging was based on the Oudega rule (using the threshold of three points) 51% of suspected patients were referred for imaging (Table 3). If this decision was based only on the GP’s probability estimate, the vast majority of patients suspected of having DVT were referred for imaging: 79% when using a cut off of 10% and 65% when using a cut off of 20%. In addition, the proportion of negative ultrasounds was higher when using a GP’s probability estimate, compared with using the Oudega rule. Consequently, this results in a lower yield of ultrasonographies when using only a GP’s probability estimate.

Both diagnostic strategies (the Oudega rule and GP probability estimates) resulted in a similar and low number of missed DVT cases in patients not referred (range of missed DVT cases in patients not referred 1.4—2.0% [Table 3]). In six of the seven patients missed by the Oudega rule, the GPs estimated a high probability (>20%) of DVT presence (Table 4). Using a threshold of 10% or 20% for the GP probability estimates, three and seven patients were missed respectively, the Oudega rule assigned a high score to two and six of these patients respectively (Table 4).

In patients without DVT (n = 866), more patients were correctly classified as low risk (score ≤3) by the Oudega rule (n = 493) than by a GP probability estimate (n = 186 using a cut off of 10% and n = 299 using a cut off of 20%). In patients correctly classified as low risk by the Oudega rule (n = 493), GPs estimated an intermediate (10–20%, n = 113) or high (>20%, n = 194) probability of DVT presence.

DISCUSSION
Summary of the main findings
The diagnostic classification of a previously developed and validated clinical decision rule (the Oudega rule) for DVT\textsuperscript{9,25,26} was compared with GP probability estimates in 1002 patients suspected of having DVT. Both the GP estimates and the Oudega rule had good overall

<table>
<thead>
<tr>
<th>Table 3. Diagnostic classification of referral decisions for the Oudega rule and GPs’ probability estimates.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total referrals, n</strong></td>
</tr>
<tr>
<td>% of all patients suspected of having DVT (95% CI)</td>
</tr>
<tr>
<td>Ultrasound negative referrals, n</td>
</tr>
<tr>
<td>% of total referrals (95% CI)</td>
</tr>
<tr>
<td>Total non-referrals, n</td>
</tr>
<tr>
<td>% of all patients suspected of having DVT (95% CI)</td>
</tr>
<tr>
<td>DVT failures in non-referrals, n\dagger</td>
</tr>
<tr>
<td>% of total non-referrals (95% CI)</td>
</tr>
</tbody>
</table>

\textsuperscript{a}The Oudega rule uses a threshold of 3 points: patients with a score ≤3 were not referred for imaging (see Table 1). \textsuperscript{b}In this scenario, patients with a GP probability estimate below 10% or 20% were considered not referred for imaging. \textsuperscript{c}No fatal or near-fatal thrombo-embolic events occurred in any of these DVT failures in non referrals. DVT = deep venous thrombosis.
Table 4. Comparing GPs’ probability estimates with the Oudega rule.

<table>
<thead>
<tr>
<th>Patients with DVT</th>
<th>Oudega score ≤3</th>
<th>Oudega score &gt;4</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP estimate &lt;10%</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>GP estimate 10-20%</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>GP estimate &gt;20%</td>
<td>6</td>
<td>123</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patients without DVT</th>
<th>Oudega score ≤3</th>
<th>Oudega score &gt;4</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP estimate &lt;10%</td>
<td>186</td>
<td>25</td>
</tr>
<tr>
<td>GP estimate 10-20%</td>
<td>113</td>
<td>23</td>
</tr>
<tr>
<td>GP estimate &gt;20%</td>
<td>194</td>
<td>325</td>
</tr>
</tbody>
</table>

*DVT = deep venous thrombosis.*

discriminative ability and performed equally. If the decision to refer a patient was based on the Oudega rule, fewer patients were referred for imaging.

As previously described (using the Oudega rule) GPs can safely refrain from referring a patient in nearly half of all cases where DVT is suspected (\(n = 500\)), missing only seven patients with DVT (that is, 1.4%). Six of the seven patients in the current study whose DVT was missed were identified by the GP as having a high estimated probability (>20%) of DVT presence. However, using only the GP probability estimate also resulted in a similar number of missed DVT cases (three and seven patients, depending on the chosen threshold for the estimate). Of these patients, most had a high score on the Oudega rule. Furthermore, GPs assigned scores indicating intermediate or high probability of DVT presence to a large group of patients, despite the vast majority of subjects not having DVT (Table 4); hence, it is possible that GPs may be uncertain of adequately differentiating between patients at low versus intermediate or high risk of DVT. Given the potential risk of fatal pulmonary embolism, GPs seem to prefer to assign an intermediate or high probability to most patients and refer them for subsequent testing.

**Strengths and limitations of the study**

To appreciate these results a few issues need to be discussed. Although both the GP estimates and the Oudega rule are able to discriminate between patients with and without DVT, it is of higher clinical relevance to compare the classification of both approaches using thresholds that are frequently used in practice to decide on further patient management. The threshold used was previously determined following the generally accepted proportion of 2% missed cases by compression ultrasonography, which is the current reference standard. Hence, using any diagnostic strategy for DVT that misses <2% of DVT cases is considered safe.

Although the decision threshold for the Oudega rule is based on previous validation studies and a generally acceptable failure percentage of ≤2%, no such validated threshold(s) exist for GP probability estimates. Therefore, the GPs’ estimates were trichotomised using cut-off points at 10% and 20%. As in risk stratification, physicians often consider three groups of patients: those who probably have the disease, those who probably do not have the disease, and those with persisting doubts. This study’s thresholds were arbitrarily chosen and different thresholds result in a different number of patients categorised as low risk. For example, increasing the threshold to 30% for the GP probability estimates will yield a similar number of patients in the low-risk category as compared with the clinical decision rule (50% for both tools, data not shown). However, this will subsequently result in more missed DVT cases as compared with the clinical decision rule (2.4% versus 1.4% respectively). Hence, based on this study, a clinical decision rule seems a better tool to safely exclude DVT in primary care, regardless of the chosen threshold for GP probability estimates.

As the Oudega rule scores and probability estimates were recorded simultaneously, it cannot be completely ruled out that the knowledge of the clinical items of the Oudega rule, as well as the result of a D-dimer test, may have influenced the GP probability estimates; this is incorporation bias. Such incorporation bias may lead to a dilution of the difference between two diagnostic tests: the accuracy of the GP probability estimates and the Oudega rule become more alike. However, the study design does reflect clinical practice where physicians apply any risk stratification and decide on further patient management. Moreover, at the time of the study, the Oudega rule had already been included in the current guidelines for GPs in The Netherlands; therefore, avoidance of any incorporation bias, by for example blinding the GPs to the result of the clinical decision rule, was almost impossible.

Knowing how incorporation bias usually works, it was expected that this bias would have improved the GP probability estimates; that is, lower estimates in patients who did not have DVT and higher estimates in those who did have DVT. However, the GPs systematically overestimated the actual DVT risk in all patients. As such, it is believed that the influence of this incorporation bias was less than expected a priori. It would likely not have influenced the main inference of this study; that is, that more patients would be referred for imaging and additional work up if the referral decision was based only on a GP’s probability estimate.

Finally, in contrast to a clinical decision rule, clinical judgement alone is very contextual and may change over time depending on a GP’s professional learning...
and feedback. Hence, a comparison between both tools may be different in different groups of GPs who have varying levels of experience. This may reduce the generalisability of the results. However, this study was conducted in three different regions in The Netherlands and included 300 different GPs with different levels of experience. Accordingly, it is possible that this large group of GPs can be regarded as a representative sample of (Dutch) GPs, thus making these results, — in the authors’ view — generalisable for primary care physicians in The Netherlands. As many countries share similar vocational training programmes for GPs (including the UK) these results may be also applicable for GPs in these countries.36

Comparison with existing literature

Although previous studies have demonstrated that concordance between a clinical decision rule and physicians’ probability estimates is poor,23,33 to the best of the authors’ knowledge this study is the first to compare GP probability estimates with a formal clinical decision rule (in the domain of primary care) in terms of discrimination and classification of patients. Therefore, results can not be easily compared with existing literature.

Knowledge about the performance of a formal clinical decision rule, compared with physicians’ judgements is important because physicians often feel reluctant to use clinical decision rules for a number of reasons. One of these is the conviction that their own clinical judgement is at least as good as, or even superior to, a decision rule.4 Indeed, a recent survey demonstrated that, although >90% of physicians were aware of the existence of clinical decision rules for assessing patients suspected of pulmonary embolism, only 44% reported using them in daily practice.36 The same may well apply to GPs’ management of patients suspected of having DVT.

The authors and Campbell et al,34 believe that implementation of a clinical decision rule in formal guidelines can improve quality of care for patients suspected of having DVT. Making GPs aware of the patients for whom there is added value of using such a rule improves their implementation and, as such, the quality of care of patients in whom DVT is suspected to be present.

Implications for clinical practice

The present study yielded various clinical implications. The Oudega rule is a safe diagnostic instrument that can be used in patients suspected of DVT as it misses an acceptably small proportion (<2%) of DVT cases. In addition, clinical judgement alone often leads to GPs overestimating the actual DVT risk and, subsequently, results in more unnecessary referrals for compression ultrasonography in a large group of patients. Taking this into account, physicians should carefully weigh the patient burden and costs involved with these unnecessary referrals or use the Oudega rule.

Funding body

The study was funded by the Netherlands Organization for Scientific Research (945–04–009) and the Netherlands Heart Foundation (2006B237). These funding sources had no role in the design, conduct, or reporting of the study, or in the decision to submit the manuscript for publication.

Ethical approval

This study was approved by the Medical Ethical Committee of the Maastricht University Medical Center, The Netherlands.

Competing interests

The authors have stated that there are none.

Acknowledgements

We gratefully acknowledge the enthusiastic participation of GPs and other healthcare professionals who contributed to the study. The steering and writing committee of the AMUSE study group: Academic Medical Center Amsterdam, the Netherlands: Harry R Bûller PhD, MD, Eit F van der Velde MD, Henk GPM van Weert MD; Maastricht University Medical Center, the Netherlands: Arina J ten Cate-Hoek PhD, MD, Manuela A Joore PhD, Martin H Prins PhD, MD, Henri EJH Stoffers PhD, MD; University Medical Center Utrecht, the Netherlands: Arno W Hoes PhD, MD, Ruud Oudega PhD, MD, Diane B Toll PhD.

Discuss this article

Contribute and read comments about this article on the Discussion Forum: http://www.rcgp.org.uk/bjgp-discuss

REFERENCES

14. Kawamoto K, Houtkin CA, Balas EA, Lobach DE. Improving clinical practice using clinical decision support systems: a systematic review of