ORIGINAL ARTICLE

Accuracy in Diagnosing Deep and Pelvic Vein Thrombosis in Primary Care

An Analysis of 395 Cases Seen by 58 Primary Care Physicians

Lobna El Tabei, Gernot Holtz, Cornelia Schürer-Maly, Heinz-Harald Abholz

SUMMARY

Background: Ruling out a deep vein thrombosis (DVT) is difficult in general practice because the clinical manifestations of DVT are nonspecific and more often due to other diseases. The aim of diagnostic screening in primary care must be to rule out a DVT with high accuracy in most patients, so that only those who are likely to have a DVT will undergo further testing. In this study, we tested the accuracy of exclusion of DVT by the combination of a clinical score (the Wells score) with either a bedside D-dimer test or selective compression sonography.

Method: This cohort study included all patients who presented to the participating primary care physicians and were suspected of having a DVT on the basis of pre-defined inclusion criteria. To rule out DVT, a Wells score was determined for all patients, and all patients additionally underwent either a D-dimer test or selective compression sonography as required by the clinical algorithm. Patients were seen six weeks later in follow-up to determine whether they had actually had a DVT (gold standard). The negative predictive value (NPV) for the exclusion of DVT in this way was determined, as was the NPV of clinical judgment alone, without knowledge of Wells score or D-dimer results.

Results: 395 patients were evaluated by 58 primary care physicians for suspected DVT; 59 were ultimately found to have had a definite DVT, and 9 a probable DVT. Exclusion of DVT with the study protocol had an NPV of 99.0% (95% CI, 96.3 to 99.8)—i.e. only one case of DVT in 100 patients was missed (maximum: 4, minimum: 0)—while clinical judgment alone had an NPV of 95.0% (95% CI, 90.7 to 97.7).

Conclusion: We recommend the Wells score combined with either a D-dimer test or selective compression sonography according to the algorithm used in this study for use in primary care to rule out DVT. Clinical judgment alone is less effective.

► Cite this as:

El Tabei L, Holtz G, Schürer-Maly C, Abholz HH: Accuracy in diagnosing deep and pelvic vein thrombosis in primary care—an analysis of 395 cases seen by 58 primary care physicians. *Dtsch Arztebl Int* 2012; 109(45): 761–6.
DOI: 10.3238/arztebl.2012.0761

The annual incidence of thromboembolic events in the general population is around 0.1% (1, 17). A small but relevant proportion of cases of deep vein thrombosis (DVT) of the leg have serious consequences: the risk of fatal pulmonary embolism is reported at between 2.6% and 9.4% (2–4). Around 20% of patients with extensive DVT of the leg develop chronic post-thrombotic syndrome (5).

The symptoms of DVT can be very unspecific. In such situations, the essential thing is to be able to rule out DVT with very high accuracy, so as to concentrate expensive technical diagnostic tests on high-risk patients as far as possible.

Because individual findings or even combinations of findings suggesting a DVT are not accurate enough to identify patients who very likely do *not* have a DVT, scoring systems have been developed, which weight combinations of individual findings that occur in the presence of DVT. The most widely used of these is the Wells score (6, 7). However, the Wells score is not enough to exclude a diagnosis of DVT with certainty (15).

To improve the accuracy of excluding a DVT, therefore, such scores have been combined with D-dimer testing. In the primary care setting, however—that is, a setting where patients are unselected—so far only three prospective studies have been carried out on the use of diagnostic scores combined with a D-dimer test, and the data from one study group have been analyzed more than once, with differing results (6–11). However, all these studies were done in healthcare systems that are different from the German one, and therefore with different mechanisms of selection of the patients investigated, and in different medical cultures, with resulting effects on the time at which patients present to a physician. This in turn influences the prevalence of DVT in the population under investigation and hence the negative predictive value (NPV). Differences in the clinical training of primary care physicians in the various countries can also affect the result of the Wells score.

For these reasons, the findings of the studies referred to cannot be uncritically extrapolated to conditions in Germany. Hence, in the present study, the aim was to test the algorithm under realistic conditions of routine clinical practice.

FIGURE 1

Study algorithm; DVT, deep vein thrombosis; CS, compression sonography

Study objective

The objective of this study is to determine the diagnostic accuracy of an algorithm made up of Wells score combined with, as appropriate, D-dimer test and/or compression sonography for the exclusion of DVT in German primary care practices; and, as a secondary objective, to compare the resulting NPV with that of the traditional clinical diagnostic process (subjective judgment of the primary care physician).

Design and methods

Recruitment of practices and training of participating primary care physicians

The study was carried out as a prospective cohort study recruiting all consecutive patients with clinical symptoms of DVT in practices in the Düsseldorf and Witten area, over a period of 18 months in each practice (December 2007 to April 2010). Physicians were recruited by fax in two regional recruitment waves.

At the start of the study, the physicians were offered a brief training course in DVT diagnosis. All participants received a short introduction to the study materials and information covering the algorithm employed in this study, the Wells score, and the bedside D-dimer test.

To prevent gaps in recruitment or documentation, participating practices were contacted every 2 to 3 weeks. Practices and physicians were not paid for taking part in the study.

Case documentation

Symptoms, biometric data, risk factors for DVT, the action taken by the physician, and the results of the D-dimer test or any other further diagnostic tests were recorded in the study questionnaire.

Before applying the algorithm, physicians graded their subjective estimate of the degree of certainty as to whether DVT was present using a 5-point Likert scale.

A follow-up questionnaire on the patient's further course was filled out 6 weeks after the initial presentation. This was designed to gather information on whether DVT had in fact occurred in patients in whom it had previously not been diagnosed. It was assumed—in the absence of any data on the subject—that any existing DVT would have become clinically manifest by this time. If the patient had not attended again by the end of 6 to 8 weeks, he or she was telephoned by the physician for an update assessment.

Gold standard for assessment

The findings recorded in the follow-up questionnaire as to the presence or absence of DVT were taken as the gold standard by which it was decided whether the algorithm—or indeed the subjective judgment of the primary care physician—was correct in relation to exclusion of a DVT.

A conscious decision was made not to carry out compression sonography as the gold standard in all patients: This would have impeded the carrying out of an outpatient study with complete capture of all patients, and would have ended in a diagnostic process that was far removed from the daily reality of primary care.

Inclusion and exclusion criteria

The formulation of the patient inclusion criterion was intentionally broad and took particular account of the way in which physicians go about forming a subjective judgment, in order to include as many as possible of the cases that would have been suspected if the tested algorithm came into use in primary care. The instruction was:

“Include every patient who, in your opinion, might have a thrombosis, or for whom a thrombosis should be included in the differential diagnosis.”

Exclusion criteria were previous treatment with heparin/anticoagulants and severe co-morbidity (e.g., tumor, severe hepatic disease, infections), as in these patients elevated D-dimer test values are possible even in the absence of a DVT and the course is unpredictable.

Study design and D-dimer test

The study was based on two steps:

- Determination of suspected cases of DVT according to the operationalizing sentence for their inclusion. The probability of this diagnosis was estimated by the physician using a 5-point Likert scale.
- Next, the Wells score was applied and—depending on its outcome—either the D-dimer test or compression sonography was carried out (Figure 1).

FIGURE 2

Overview of the clinical course of cases of suspected DVT.
DVT, deep vein thrombosis

The algorithm prescribed that the primary care physician should carry out a D-dimer test always and only in patients with a Wells score ≤ 1 (low clinical probability) (16). If the D-dimer test was negative, it was assumed that no DVT was present. However, this was not proven until the gold standard selected for this study—the documented course over 6 weeks—became available.

Patients with a Wells score >1 , on the other hand, were always, and without D-dimer testing, to be referred to a consultant for further investigation by compression venous ultrasound or color duplex ultrasound.

In this study a rapid point-of-care D-dimer test was used (Clearview Simplify D-dimer assay; Inverness Medical, Bedford, UK), which has been validated in several studies (12–14).

Ethical approval

The ethics committee at Düsseldorf University Hospital approved the study. All patients gave written informed consent to their participation.

Results

Primary care practices

Of the practices approached by fax (half of them teaching practices of the University of Düsseldorf and half of the University of Witten-Herdecke), 66% participated.

Patients with possible DVT

Out of 395 recruited patients (suspected cases), 59 (14.9%) had a confirmed DVT; in addition, there were 9 probable cases of DVT (17.2% together). In all, 51% of the patients were women. The mean age of patients in the study was 61.3 years. The main symptoms leading to inclusion in the study were: dragging pain (66%), a feeling of tension and heaviness (59%), and acute swelling (50%) (eFigures 1 and 2).

Figure 2 shows the distribution of cases of suspected DVT in relation to diagnosis at the end of the diag-

TABLE 1

Weighting of clinical features for calculation of the Wells score for clinical probability of presence of DVT (16)

Clinical feature	Points
Active or treated cancer within the past 6 months	1
Paralysis, paresis, or recent plaster immobilization of the lower extremity	1
Bedridden recently for >3 days or major surgery within the past 12 weeks	1
Pain/induration along deep venous system	1
Swelling of entire leg up to upper thigh	1
Unilateral calf swelling >3 cm compared to other side	1
Pitting edema in the symptomatic leg	1
Extended superficial (nonvaricose) collateral veins	1
Previously documented DVT	1
Alternative diagnosis as likely as DVT	-2
Clinical probability of DVT	
Cut-off score for low clinical probability of DVT (low-risk group)	≤ 1
Cut-off score for high clinical probability of DVT (high-risk group)	>1

DVT, deep vein thrombosis

nostic work-up or confirmation of diagnosis at the end of 6 weeks.

“Probable DVT” cases are those that showed a Wells score greater than 1 and a negative compression ultrasound scan, but a positive D-dimer test result. However, the further compression sonography exam 1 week later that is envisaged by the algorithm was not carried out in this patient group. For the purposes of the study analysis, these patients were grouped with the “confirmed DVT” cases according to the 6-week-course questionnaire, since for most of them initial antithrombotic treatment was started to prevent progression of findings and pulmonary embolism in the same way as in patients with a confirmed DVT.

In contrast to those patients were 310 in whom a DVT was excluded. These cases were defined by the fact that neither the primary care physician nor other treating institutions identified the presence of a DVT by the end of 6 weeks.

The “thrombosis excluded” cases included three patients with pulmonary embolism. In two of these, DVT was ruled out by means of compression sonography, carried out in these cases because they belonged to the high-risk group. In the case of the third patient, the physician went directly to pulmonary embolism therapy, without any venous diagnostic procedure. None of these patients died during the observation period.

Distribution of Wells score values for all cases of suspected DVT (N = 395)

FIGURE 3

A total of 17 of the above-mentioned 395 suspected cases had to be excluded from the analysis because they were “visiting” patients from other practices, most of which were unknown, and from which it would not be possible to get follow-up data about the 6-week course. Thus, this study found an 18% incidence of DVT among all evaluable cases of suspected DVT. A total of 16 cases of pulmonary embolism were reported in the 6-week follow-up documentation, 11 in the confirmed DVT and 2 in the probable DVT group. In the remaining 3 patients, following the algorithm had previously led to DVT being regarded as ruled out. There were also 3 cases of DVT of the upper extremity, all of them in patients with confirmed DVT.

Wells score

Table 1 and Figure 3 show how the Wells score is calculated and the distribution of Wells score results for all cases of suspected DVT. Most suspected cases have a score of 1.

If suspected cases are dichotomized according to Wells score, 288 patients included on the basis of suspected DVT had a low probability (Wells score ≤ 1) and 107 a high probability (Wells score > 1).

Diagnostic accuracy of the algorithm

The D-dimer test is intended to serve as an additional aid to definitively rule out a DVT in the group of patients with a low-risk Wells score, while patients in the high-risk group are anyway required by the algorithm to be referred for further investigation (compression sonography).

The upper half of Table 2 shows the calculated diagnostic accuracy values (specificity, sensitivity, positive predictive value [PPV], and negative predictive value [NPV]) for the whole group of evaluable patients (N = 370). Five patients could not be included in the calculation because—in contravention of the study instructions—they did not undergo the D-dimer test. Another 7 patients could not be included because there was no follow-up information for them.

The NPV calculated for the algorithm investigated in this study was 99.0% (95% CI: 96.3 to 99.8). It must be pointed out here that the algorithm is not suitable for positive diagnosis of a DVT (the PPV is low), but this is already known from the literature.

Physicians' subjective judgment

Figure 4 shows the analysis of physicians' subjective clinical judgment. The blue bars relate to all suspected cases. There are five categories, from “very probable” to “very improbable.” This distribution is compared to the red bars showing how many of these cases were later shown to be confirmed or probable cases of DVT.

Physicians' subjective judgment markedly overestimated the number of cases of DVT. Moreover, the categories “fairly improbable” and “very improbable” — those for which in clinical routine no further diagnostic procedures would be carried out—contained no less

TABLE 2

Diagnostic accuracy of the chosen algorithm compared to physician's subjective judgment^{*1}

Study algorithm ^{*1}		Total number of patients	370	
True-positive 62	False-positive 115	Specificity	62.4%	95% CI (%): 56.7–67.9
		Sensitivity	96.9%	95% CI (%): 89.1–99.5
False-negative 2	True-negative 191	Pos. predictive value	35.0%	95% CI (%): 28.0–42.5
		Neg. predictive value	99.0%	95% CI (%): 96.3–99.8

Physician's subjective judgment		Total number of patients	370	
True-positive 55	False-positive 135	Specificity	55.9%	95% CI (%): 50.1–61.5
		Sensitivity	85.9%	95% CI (%): 75.0–93.4
False-negative 9	True-negative 171	Pos. predictive value	29.0%	95% CI (%): 22.6–36.0
		Neg. predictive value	95.0%	95% CI (%): 90.7–97.7

^{*1} Here N = 370, because the dataset for both procedures was only evaluable for this number; accordingly, the number of DVT cases was only 64

than 10 of the 68 confirmed or probable cases of DVT: This would have meant roughly a 15% rate of missed cases. In the evaluable 370 cases for the two parameters in Table 2 (NPV for the algorithm and for subjective judgement), the corresponding figures are 9 out of 64.

To calculate the NPV of physicians' subjective judgment, we treated all cases described on subjective judgment as "unclear," "probable," or "very probable" cases of DVT as though DVT had been detected in 100% of them. Moreover, for the categories "probable" and "very probable" it was assumed that the physician in routine clinical practice would not have referred the patient for any further diagnostic investigations (D-dimer test or compression sonography). This resulted in a "best-case scenario" as a basis for the NPV calculation in relation to subjective judgment. Despite this, the calculated NPV (Table 2) was only 95.0% (95% CI: 90.7 to 97.7).

Discussion and conclusions

The study shows that use of the employed algorithm leads to a high degree of accuracy in ruling out DVT. The NPV in the whole group of suspected cases is 99.0% (95% CI: 96.3 to 99.8). According to this, one patient in 100 would be missed; considering the confidence interval it would be four at most and at the lowest it would be none. On the basis of subjective judgment alone, the NPV—even though based on the best-case scenario—would be only 95.0% (95% CI: 90.7 to 97.7); in this case, anywhere between two and nine patients would be missed.

The obvious conclusion from this is to choose to diagnose according to the algorithm in future, even though the two 95% confidence intervals do overlap slightly,—an indication that the results could still be due to coincidence, although with low probability.

In comparing the NPVs of algorithm use versus the use of subjective judgment, it should also be remarked that, because the physicians were using the algorithm, they probably also increasingly learned to incorporate it into their subjective judgment. This would mean that the clinical judgment of the participating physicians tended to be better than that of physicians who were not trained in the algorithm. It may therefore be assumed that the difference between following the algorithm and following normal clinical judgment is larger in reality than indicated by the NPVs in the study. Hence, the NPV for following clinical judgment has been boosted twice over: by being based on the best-case scenario, and by the assumed positive influence of the Wells score on participating physicians.

To definitively determine the true difference between the two routes (algorithm or subjective judgment) would require a randomized study, in which one group of physicians was given the algorithm and the other group was not. Since the superiority of the algorithm has been shown by several studies (even if performed in other settings and countries than Germany), we considered it unethical, however, to perform a randomized trial.

FIGURE 4

Distribution of clinical estimates of the probability of DVT by the primary care physician (blue = suspected cases, n = 395; red = confirmed and probable cases, n = 68). DVT, deep vein thrombosis

It should also be pointed out that relying on clinical judgment alone leads to a greater number of further diagnostic tests. According to Table 2, further tests would have been carried out in 190 cases on the basis of clinical judgment, and in 177 cases on the basis of the algorithm. This means that using the algorithm leads to more certainty and is also associated with less (and less expensive) diagnostic work-up.

The other studies come to exactly the same conclusion. Since the 1990s there have been 18 individual studies and 8 meta-analyses investigating a combined approach of prediction rules (usually Wells or Oudega) and D-dimer testing. Only three of these studies were carried out at the primary care level; these three studies will now be briefly described.

The results of Oudega et al. (6, 7) and Büller et al. (9) in the Netherlands allow numerical comparison with the results of the present study: The algorithms are the same, or very similar. Oudega et al. use the Wells score, supported by D-dimer testing for the low-risk patients. Büller et al., on the other hand, use a different score, the Oudega rule, according to which a D-dimer test is mandatory in all cases of suspected DVT.

Geersing et al. (10) and van der Velde et al. (11) analyzed the same study data as Büller et al. but used different algorithms, so a direct comparison with our data is not practicable.

In comparing the results of Oudega et al. and Büller et al. with those of the present study, it must be remembered that the DVT prevalence within the group of suspected cases differs: In Oudega et al. it is 22%, in Büller et al. 13.5%, and in the present study it is 17%. This affects the NPV of the algorithm. In addition, various D-dimer tests of differing sensitivity/specificity,

different scores, and various different cut-off levels of the Wells score were employed.

Notwithstanding, the results are very similar. The authors of the present study found a NPV of 99.0% (95% CI: 96.3 to 99.8). Oudega et al. report a NPV of 97.1% (95% CI: 95.0 to 99.2) with the same cut-off value (Wells score ≤ 1) and the identical algorithm to the present study. In Büller et al., the corresponding value for the combination of Oudega score and D-dimer test in all cases (that is, not just for low-risk patients) is 98.6% (95% CI: 97.1 to 99.4). This shows that the percentage of missed thromboses lies between 2.9% (Oudega et al.), 2.4% (Büller et al.), and 1.0% (present study). Whether the differences between the Dutch studies and the present study are due to differences in healthcare systems, physician skills, or patient behavior, or whether they are due to differences in study design, is impossible to determine. The results of the present study together with those of comparable studies make it clear that an algorithm-supported diagnostic process is superior to one that relies on routine clinical decision making.

Strengths and weaknesses

It may be assumed that all cases of suspected DVT were identified, since our regular contact with the practices reminded physicians about the study and ensured that, if a patient had remained undocumented, this could be corrected within a short time.

Missing data for a few patients are a limitation that cannot be avoided in a study of this kind. This limitation occurs with hospital studies too, and appears to be acceptable at the order of magnitude seen in this study. The other weakness, not to have carried out a randomized controlled study, is addressed above.

Acknowledgments

The authors express their thanks to all physicians and practice personnel who collaborated in the study.

KEY MESSAGES

- It is essential to rule out deep vein thrombosis (DVT) with certainty.
- Studies show that clinical examination and history alone are insufficient for this.
- Using the algorithm investigated in this study, comprising Wells score together with either a D-dimer test (in patients identified by their score as low-risk) or compression sonography (as the first step only in high-risk patients), the presence of DVT can be ruled out with high probability.
- In the present study of patients in primary care, only 1% of true-positive DVT cases are missed; with normal clinical subjective judgment, 5% are missed.
- When the algorithm employed in this study is used, fewer patients need to be investigated using compression sonography.

Conflict of interest statement

Inverness Medical made the D-dimer test kits available at a reduced price, but had no influence on the conception, design, performance, or evaluation of the study.

The authors declare that no conflict of interest exists.

Manuscript received on 30 January 2012, revised version accepted on 26 July 2012.

Translated from the original German by Kersti Wagstaff, MA.

REFERENCES

1. White RH: The epidemiology of venous thromboembolism. *Circulation* 2003; 107: 14–8.
2. Hansson PO, Sorbo J, Eriksson H: Recurrent venous thromboembolism after deep vein thrombosis: Incidence and risk factors. *Arch Intern Med* 2000; 160: 769–74.
3. Silverstein MD, Heit JA, Mohr DN, Petterson TM, O'Fallon WM, Melton LJ: Trends in the incidence of deep vein thrombosis and pulmonary embolism. *Arch Intern Med* 1998; 158: 585–93.
4. Cushman M, Tsai AW, White RH, et al.: Deep vein thrombosis and pulmonary embolism in two cohorts: The longitudinal investigation of thromboembolism etiology. *Am J Med* 2004; 117: 19–25.
5. Prandoni P, Lensing AW, Cogo A, et al.: The long-term clinical course of acute deep venous thrombosis. *Ann Intern Med* 1996; 125: 1–7.
6. Oudega R, Moons KG, Hoes AW: Ruling out deep venous thrombosis in primary care. A simple diagnostic algorithm including D-Dimer testing. *Thromb Haemost* 2005; 94: 200–5.
7. Oudega R, Hoes AW, Moons KG: The Wells rule does not adequately rule out deep venous thrombosis in primary care patients. *Ann Intern Med* 2005; 143: 100–7.
8. Tagelagi M, Elley CR: Accuracy of the Wells rule in diagnosing deep vein thrombosis in primary health care. *N Z Med J* 2007; 120: 2705.
9. Büller HR, Ten Cate-Hoek AJ, Hoes AW, et al.: Safely ruling out deep venous thrombosis in primary care. *Ann Intern Med* 2009; 150: 229–35.
10. Geersing GJ, Janssen K, Oudega R, et al.: Diagnostic classification in patients with suspected deep venous thrombosis: physicians' judgement or a decision rule? *Br J Gen Pract* 2010; 60: 742–8.
11. Van der Velde EF, Toll DB, Ten Cate-Hoek AJ, et al.: Comparing the diagnostic performance of 2 clinical decision rules to rule out deep vein thrombosis in primary care patients. *Ann Fam Med* 2011; 9: 31–6.
12. Dempfle CE, Zips S, Ergul H, et al.: The Fibrin Assay Comparison Trial (FACT): evaluation of 23 quantitative D-dimer assays as basis for the development of D-dimer calibrators. FACT study group. *Thromb Haemost* 2001; 85: 671–8.
13. Cini M, Legnani C, Bettini F, Cavallaroni K, Palareti D: A new rapid bedside assay for D-dimer measurement (Simplify D-dimer) in the diagnostic work-up for deep vein thrombosis. *J Thromb Haemost* 2003; 1: 2681–3.
14. Van Der Velde EF, Wichers IM, Toll DB, Van Weert HC, Büller HR: Feasibility and accuracy of a rapid 'point-of-care' D-dimer test performed with a capillary blood sample. *J Thromb Haemost* 2007; 5: 1327–30.
15. Wells PS, Owen C, Doucette S, Fergusson D, Tran H: Does this patient have deep vein thrombosis? *JAMA* 2006; 295: 199–207.
16. Wells PS, Anderson PR, Rodger M, Forgie M, Kearon C, Drier J, et al.: Evaluation of D-dimer in the diagnosis of suspected deep-vein thrombosis. *N Engl J Med* 2003; 349: 1227–35.
17. Spencer FA, Emery C, Joffe SW, Pacifico L, Lessard D, Reed G, et al.: Incidence rates, clinical profile, and outcomes of patients with venous thromboembolism. The Worcester VTE study. *J Thromb Thrombolysis* 2009; 28: 401–9.

Corresponding author

Prof. Dr. med. Heinz-Harald Abholz
Institut für Allgemeinmedizin, Heinrich-Heine-Universität, Universitätsklinikum
Moorenstr. 5, 40225 Düsseldorf, Germany
abholz@med.uni-duesseldorf.de

eFigures:
www.aerzteblatt-international.de/12m0761

ORIGINAL ARTICLE

Accuracy in Diagnosing Deep and Pelvic Vein Thrombosis in Primary Care

An Analysis of 395 Cases Seen by 58 Primary Care Physicians

Lobna El Tabei, Gernot Holtz, Cornelia Schürer-Maly, Heinz-Harald Abholz

Documented symptoms at study inclusion of patients with suspected DVT

(the following additional items were recorded: feeling of heaviness and tension and/or dragging pain [12×]; symptoms associated with pulmonary embolism [4×]; other data recorded only once each, e.g., pregnancy associated with leg swelling [7×]. DVT, deep vein thrombosis)

eFIGURE 2

Distribution of clinical features in all cases of suspected DVT; multiple entries exist