BMJ

BMJ 2014;348:g1340 doi: 10.1136/bmj.g1340 (Published 10 March 2014)

RESEARCH

Exclusion of deep vein thrombosis using the Wells rule in clinically important subgroups: individual patient data meta-analysis

G J Geersing *general practitioner*[1], N P A Zuithoff *statistician*[1], C Kearon *professor of medicine*[2], D R Anderson *professor of medicine*[3], A J ten Cate-Hoek *thrombosis specialist*[4], J L Elf *thrombosis specialist*[5], S M Bates *thrombosis specialist*[2], A W Hoes *professor of primary care and clinical epidemiology*[1], R A Kraaijenhagen *thrombosis specialist*[6], R Oudega *general practitioner*[1], R E G Schutgens *thrombosis specialist*[7], S M Stevens *thrombosis specialist*[8], S C Woller *thrombosis specialist*[8], P S Wells *professor of medicine*[9], K G M Moons *professor of clinical epidemiology*[1]

[1]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Universiteitsweg 100, 3584 CG Utrecht, Netherlands; [2]Division of Haematology and Thromboembolism, Department of Medicine, McMaster University Hamilton, Hamilton, Canada; [3]Division of Haematology, Department of Medicine, Dalhousie University, Halifax, Canada; [4]Department of Internal Medicine, Maastricht University Medical Center, Maastricht, Netherlands; [5]Vascular Center, Skane University Hospital, Malmö, Sweden; [6]Department of Medicine, Academic Medical Center Amsterdam, Netherlands; [7]Van Creveld Clinic, University Medical Center Utrecht, Utrecht, Netherlands; [8]Thrombosis Clinic, Intermountain Medical Center, Murray, UT, USA; [9]Department of Medicine, Ottawa Hospital, University of Ottawa, Ottawa, Canada

Abstract

Objective To assess the accuracy of the Wells rule for excluding deep vein thrombosis and whether this accuracy applies to different subgroups of patients.

Design Meta-analysis of individual patient data.

Data sources Authors of 13 studies (n=10 002) provided their datasets, and these individual patient data were merged into one dataset.

Eligibility criteria Studies were eligible if they enrolled consecutive outpatients with suspected deep vein thrombosis, scored all variables of the Wells rule, and performed an appropriate reference standard.

Main outcome measures Multilevel logistic regression models, including an interaction term for each subgroup, were used to estimate differences in predicted probabilities of deep vein thrombosis by the Wells rule. In addition, D-dimer testing was added to assess differences in the ability to exclude deep vein thrombosis using an unlikely score on the Wells rule combined with a negative D-dimer test result.

Results Overall, increasing scores on the Wells rule were associated with an increasing probability of having deep vein thrombosis. Estimated probabilities were almost twofold higher in patients with cancer, in patients with suspected recurrent events, and (to a lesser extent) in males. An unlikely score on the Wells rule (≤1) combined with a negative D-dimer test result was associated with an extremely low probability of deep vein thrombosis (1.2%, 95% confidence interval 0.7% to 1.8%). This combination occurred in 29% (95% confidence interval 20% to 40%) of patients. These findings were consistent in subgroups defined by type of D-dimer assay (quantitative or qualitative), sex, and care setting (primary or hospital care). For patients with cancer, the combination of an unlikely score on the Wells rule and a negative D-dimer test result occurred in only 9% of patients and was associated with a 2.2% probability of deep vein thrombosis being present. In patients with suspected recurrent events, only the modified Wells rule (adding one point for the previous event) is safe.

Conclusion Combined with a negative D-dimer test result (both quantitative and qualitative), deep vein thrombosis can be excluded in patients with an unlikely score on the Wells rule. This finding is true for both sexes, as well as for patients presenting in primary and hospital care. In patients with cancer, the combination is neither safe nor efficient. For patients with suspected recurrent disease, one extra point should be added to the rule to enable a safe exclusion.

Introduction

Doctors regularly encounter patients with leg problems and must decide whether to test for deep vein thrombosis. As signs

Correspondence to: G J Geersing g.j.geersing@umcutrecht.nl

Extra material supplied by the author (see http://www.bmj.com/content/348/bmj.g1340?tab=related#webextra)

Appendices

and symptoms can be non-specific, many patients require additional diagnostic testing; the consequences of missing an event can be fatal (for example, pulmonary embolism).[1] In recent studies, the prevalence of thrombosis in suspected patients was only around 10-15%, suggesting that doctors have a low threshold for diagnostic testing.[2 3] Various clinical decision rules have been developed to improve the clinical investigations for suspected deep vein thrombosis. These rules combine different clinical factors to yield a score, which is then used to estimate the probability of deep vein thrombosis being present. The most widely used clinical decision rule is probably that developed by Wells and colleagues (table 1⇓).[1] Many studies have shown that it is safe to withhold anticoagulant treatment without further diagnostic testing in patients with a low score on the Wells rule combined with a negative D-dimer test result.[4]

The validity of the Wells rule has, however, also been questioned in various subgroups of patients. For example, one study in primary care concluded that the combination of a low score on the Wells rule and a negative D-dimer test result was unsafe, as it was associated with an unacceptably high proportion of missed cases.[5] Also, some argue that the Wells rule is less safe in patients with an active malignancy[6 7] or suspected recurrent deep vein thrombosis,[8] and in male patients,[9] because in all these subgroups the actual prevalence of deep vein thrombosis is higher in the group suspected of having the condition. Therefore, although the Wells rule seems to be a valid tool in the clinical investigation of suspected deep vein thrombosis in unselected patients, its validity in various clinically important subgroups is unclear; most original diagnostic studies on deep vein thrombosis contained few patients in these important subgroups.

To determine whether the Wells rule behaves differently in such subgroups we combined individual patient data from 13 diagnostic studies of patients with suspected deep vein thrombosis (n=10 002).[10-22] Such meta-analyses of individual patient data (data of individual studies combined at patient level) provide a unique opportunity to perform robust subgroup analyses.

Methods

Study identification

Based on a previous meta-analysis[1] as well as contacts with experts in the specialty, we contacted all known principal investigators of published primary studies on the diagnosis of deep vein thrombosis that met the inclusion criteria. To check for additional papers, we performed an updated search using a validated algorithm for finding diagnostic studies.[23] We restricted the search to only new papers that were published after the previous meta-analysis in 2006 (see appendix I on bmj.com). Apart from papers retrieved by our contacts with experts in the specialty, we found no additional papers. We asked the authors for their original datasets in the context of this meta-analysis. Individual patient data were provided anonymously, and for all original publications the contributors had acquired ethical approval (including written informed consent) for each included dataset. No further ethical approval was needed for this meta-analysis.

Study selection

Studies needed to fulfil several criteria to be included in the individual patient dataset. They had to recruit consecutive outpatients with suspected deep vein thrombosis; contain a dataset with all predictors that form the Wells rule; include categorisation of patients using the Wells rule, before venous imaging (reference test); have the results of any D-dimer testing available before reference testing; and document the presence or absence of proximal deep vein thrombosis by an acceptable reference test. Acceptable such tests were either compression ultrasonography or venography at initial presentation, or, if venous imaging was not performed, an uneventful follow-up for at least three months.

Construction of individual patient level dataset

From each original dataset we identified all diagnostic tests or predictors included in the Wells rule, the results of D-dimer testing (quantitative or qualitative), and the presence or absence of deep vein thrombosis based on any combination of the three reference tests. Furthermore, we documented the patients' age, sex, and presence or absence of previous deep vein thrombosis.

To avoid the bias from excluding patients with missing data (see appendix II on bmj.com for information on the proportion of missing values per study), we used a multivariable regression technique to impute missing values within each study.[24-26] We imputed missing values per dataset only.[27] The proportion of missing values that were imputed in our individual patient dataset ranged from less than 1% for the presence or absence of deep vein thrombosis to 5% for the Wells rule variable "alternative diagnosis as likely as or more possible than deep vein thrombosis." In five studies, D-dimer test results were not available for this individual patient data analysis, either because in the particular study D-dimer testing was not carried out or because D-dimer test results could not be provided by the original study group.[10 12 15 16 20] As we imputed per dataset, these missing D-dimer test results were not imputed. Accordingly, D-dimer test results were available in 7625 patients.

After imputation, we merged all datasets into one individual patient dataset. Construction of the individual patient dataset, including the imputation, was performed in SPSS (version 18.0 (PASW), IBM, Chicago, IL).

Statistical analyses

To analyse the accuracy of the Wells rule to exclude or include the presence of deep vein thrombosis for each subgroup, we used logistic regression modelling with the observed presence or absence of deep vein thrombosis as the outcome. As with each individual patient dataset, patient data were inherently clustered within the 13 different studies. Accordingly, patients within a study are more alike than two randomly chosen patients from different studies, which will possibly affect the accuracy (or the associations) of the predictors in the Wells rule. We accounted for this clustering using multilevel logistic regression techniques, with the levels being patients within studies (a so-called one step model).[28-33] Following methodological recommendations, we included a random effect for the intercept (related to the prevalence of deep vein thrombosis in each original study), whereas we assessed the effect of each covariate using a fixed effect approach.[34 35]

We used various (multilevel) logistic regression models to assess the impact of each a priori defined subgroup on the performance of the Wells rule. The first basic model used the Wells rule score as the only covariate. Using the intercept and regression coefficient of this model, we calculated the mean predicted probabilities of deep vein thrombosis (with corresponding 95% confidence intervals) for each score on the Wells rule. To check the robustness of this model, we compared these mean predicted probabilities of deep vein thrombosis with actual observed rates (including a range of this deep vein thrombosis rate) over the different included studies. To determine if the Wells rule should

be applied differently for each subgroup we extended this basic model with a covariate for each subgroup, plus an interaction term for the Wells rule score with that subgroup variable (separately per subgroup covariate). Again, we used these models to compute the mean predicted probabilities of deep vein thrombosis for each score on the Wells rule, given the presence (or absence) of that particular subgroup variable. We a priori defined several subgroup variables: care setting (primary or hospital), presence of an active malignancy, male or female sex, and a history of deep vein thrombosis.

Next, we quantified to what extent the addition of a negative D-dimer test result (as reported in the original studies) to an unlikely Wells rule score (≤1) enhanced the exclusion of deep vein thrombosis per subgroup. Using a similar approach as with the subgroup analyses of the Wells rule, we estimated the mean predicted probabilities of deep vein thrombosis (with 95% confidence intervals) for patients with a low score on the Wells rule (≤1) combined with a negative D-dimer test result per subgroup; this is the failure rate of the Wells rule in each subgroup. In addition, the proportion of patients in whom deep vein thrombosis can possibly be excluded based on a low score on the Wells rule and a negative D-dimer test result is also an important measure of the value of the Wells rule in each subgroup. This proportion is often called the efficiency of the rule. Hence, in each patient the ability of the Wells rule to exclude deep vein thrombosis—that is, a low score on the Wells rule (≤1) combined with a negative D-dimer test result—as either present or not present. In a similar fashion to the multilevel logistic models that were used to estimate mean predicted probabilities of deep vein thrombosis, we constructed a model using this observed exclusion ability as outcome, and the presence or absence of the subgroup variables as covariates. Accordingly, we calculated the mean predicted efficiency proportions (with corresponding 95% confidence intervals) per subgroup. We repeated the analyses for the overall failure rate and efficiency, adding type of D-dimer assay to the model: a quantitative versus a qualitative assay. For the quantitative assays we used the cut-off for the D-dimer test as reported in the original publication (usually <500 µg/L). Qualitative (point of care) assays only reported a positive or negative test result. Similarly, we assessed the impact of prevalence on these estimates by adding the (logit transformed) prevalence of each study as a covariate to the failure rate and efficiency models.

Finally, we performed additional analyses aimed at illustrating the heterogeneity between the studies included in our individual patient dataset. To do so, we calculated prediction intervals around our model estimates (this is, the probability of deep vein thrombosis) for each Wells rule score.[30] In contrast with a confidence interval (that provides information on the precision of our model estimates), a prediction interval illustrates the heterogeneity of our results and basically can be seen as the range of possible probabilities of deep vein thrombosis (given our individual patient dataset) for each Wells rule score that can be expected in a new study. Following methodological recommendations, we calculated these prediction intervals using the variability around the average intercept (random effects model).[34 35] To explain observed heterogeneity of our results, we added (logit transformed) prevalence to this model as well.

All statistical analyses were performed using R version 2.10.0, in particular the lme4 package (R foundation for Statistical Computing, www.R-project.org).

Results

Included studies

The final individual patient dataset included 10 002 patients who were enrolled in 13 studies in Canada, the Netherlands, the United States, and Sweden (table 2⇓ and figure⇓). Patient management was guided by the Wells rule in seven studies,[10-12 14 15 19 22] whereas four studies validated the Wells rule.[13 17 18 21] Two studies were originally designed to determine whether serial testing was needed in patients with an initial normal ultrasonogram (either by ultrasonography repeated after one week or venography in patients with a positive D-dimer test result).[16 20] Ten studies included hospital care outpatients[10-12 14-17 19 20 22] and three studies included primary care outpatients.[13 18 21] Inpatients were not included in any of the studies. Of the 10 002 patients, 1864 (19%) had proximal deep vein thrombosis. The median age was 59 years and 62% of patients were female.

Overall accuracy of Wells rule and accuracy in various subgroups

Table 3⇓ presents the mean predicted probabilities for each score on the Wells rule. These mean predicted probabilities of deep vein thrombosis are compared with the mean actual observed rate from the included studies (table 4⇓), showing overall good agreement between the predicted and observed probabilities (see appendix III on bmj.com for details of the raw data from our individual patient dataset used to calculate these actual observed rates, per study). Table 4 also illustrates the heterogeneity (and prediction intervals) around these model estimates that were largely explained by differences in prevalence of deep vein thrombosis in our included studies. In both the overall analysis, as in the analyses for all subgroups, an increasing score on the Wells rule was correlated with a higher predicted probability of deep vein thrombosis. However, even with a Wells rule score of −2, the actual probability was around 5% (range 2.0-5.9%). Therefore deep vein thrombosis cannot be excluded using the Wells rule in isolation. Also, notably for lower scores on the Wells rule (roughly up to a score of 1), predicted probabilities of deep vein thrombosis were almost twice as high in patients with a history of deep vein thrombosis, patients with cancer, and, to a lesser extent, males.

Exclusion safety and efficiency using a dichotomised Wells score and D-dimer testing

Table 5⇓ shows the mean predicted probability of deep vein thrombosis in patients with a low score on the Wells rule (≤1) combined with a negative D-dimer test result (that is, the failure rate). In the overall analyses, this failure rate was 1.2% (95% confidence interval 0.7% to 1.8%). Given that even invasive venography cannot find all cases, a failure of up to 2% is often deemed as acceptable.[1 8] This acceptability threshold was only crossed for patients with cancer and for those with a history of deep vein thrombosis.

In addition to safety, the efficiency of a rule-out strategy is important, as an effective strategy must exclude deep vein thrombosis in a substantial proportion of patients. The Wells rule combined with a negative D-dimer test result remained highly efficient among subgroups, except for patients with cancer (efficiency 9.1%, 95% confidence interval 5.5% to 14.7%). Excepting patients with cancer, deep vein thrombosis can be excluded using the Wells rule and D-dimer testing in about 1 in every 3 patients (range of efficiency 23-43%, see table 5). Exclusion efficiency was highest if the Wells rule was

combined with a qualitative D-dimer test in a low prevalence population (prevalence 5%), whereas this did not result in an unacceptably high increased failure rate; table 6⇓. (See appendix IV on bmj.com for a more detailed description of all the models used for the analyses.)

Discussion

We performed a comprehensive diagnostic individual patient data meta-analysis of 13 studies on the diagnostic accuracy of the Wells rule for excluding deep vein thrombosis. Our dataset included 10 002 patients with suspected deep vein thrombosis, with 1864 confirmed cases. This large individual patient data meta-analysis enabled us to perform more robust subgroup analyses on clinically important subgroups than in the individual studies.

An increasing Wells rule score was associated with a higher probability of deep vein thrombosis, yet even with the lowest score, deep vein thrombosis cannot be excluded without adding D-dimer testing to the rule. In additional analyses, this finding was also true in patients in whom an alternative diagnosis than deep vein thrombosis was deemed more likely (data not shown). We found no clinically important differences for the accuracy of the Wells rule in males or females, or in patients presenting in primary or hospital care. For all these subgroups, whether combined with a qualitative or a quantitative D-dimer test, less than 2% of cases were missed; thus enabling a safe exclusion of deep vein thrombosis in these subgroups. In patients with cancer, the mean predicted probability was almost twofold higher, notably for lower scores on the Wells rule. As a result, even when a low score on the Wells rule is combined with a negative D-dimer test result, the probability of deep vein thrombosis crossed the commonly accepted safety threshold of missing a maximum proportion of 2% of deep vein thrombosis cases in these patients.

Strengths and limitations of this study

The strength of this analysis is that it includes individual patient data from more than 10 000 patients with suspected deep vein thrombosis, with almost 2000 confirmed cases. Yet, for full appreciation of our findings, several points should be discussed in more detail.

Firstly, one validation study by Oudega and colleagues, also included in our analysis, found that the Wells rule in combination with a D-dimer test was not safe for excluding deep vein thrombosis in primary care: at a threshold of ≤1 or ≤0, both combined with a negative D-dimer test result, the authors found a missed proportion of 2.9% and 2.3% missed cases, respectively.[5 18] Consequently, they decided to develop, validate, and study the impact of a separate rule for primary care (the Oudega rule), with good results.[13 18 21 36] Later, a study performed by van der Velde that indirectly compared this Oudega rule with the Wells rule (using data from a large management study where patient management was guided by the Oudega rule) found similar and safe results for both rules: missed proportion of a low score combined with a negative D-dimer test result 1.4% for the Oudega rule and 1.6% for the Wells rule.[37] Our individual patient data meta-analysis, which included about five times as many patients as these individual studies, confirms that, besides the Oudega rule, the Wells rule seems safe to use for the exclusion of deep vein thrombosis in primary care. Reasons for these conflicting results on the accuracy of the Wells rule in primary care are not entirely clear but could include the fact that in the validation study by Oudega and colleagues a relatively large proportion of patients had suspected recurrent thrombosis: 24% of all included patients compared with 15% suspected recurrent events in the study by van der Velde.[5 37] The reported high failure rate in the original Oudega publication was based on the calculation of the original Wells rule,[5] whereas this individual patient data meta-analysis clearly showed that this original Wells rule is unsafe in patients with suspected recurrent deep vein thrombosis.

Secondly, concerns on the preferred diagnostic strategy in patients with cancer and suspected deep vein thrombosis have been raised. The reasons for these concerns are twofold. Firstly, most patients with an active malignancy have increased D-dimer levels, even in the absence of thrombosis, thereby reducing the efficiency of a rule-out approach.[7] Secondly, the pretest probability of deep vein thrombosis in patients with cancer is higher compared with patients without cancer, reducing the negative predictive value of the Wells rule and D-dimer testing.[6] Our individual patient data meta-analysis confirmed both concerns. The combination of a low score on the Wells rule and a negative D-dimer test result in patients with cancer occurred in only 9% of all patients. Moreover, in these low risk patients, deep vein thrombosis was still present in 2.2% (95% confidence interval 0.5% to 8.6%); thereby crossing the safety margin of 2% of missed cases.

Thirdly, controversy remains on the value of the Wells rule in patients with suspected recurrent deep vein thrombosis. Leg problems in patients with a history of deep vein thrombosis can be clinically a difficult dilemma, as the probable causes include either an exacerbation of post-thrombotic syndrome or recurrent disease. In addition, recurrent deep vein thrombosis is difficult to confirm even with imaging, and this diagnosis often leads to long term anticoagulation treatment.[38] An incorrect diagnosis commits the patient to an unnecessary risk of bleeding complications. An accurate method to exclude recurrent thrombosis is therefore of high clinical importance. An updated version of the Wells rule was created for this purpose that included one extra point given if a suspected patient has a history of confirmed deep vein thrombosis.[1] Yet, this modified model was never properly validated. In fact, in our individual patient data meta-analysis, seven studies even excluded patients with a history of deep vein thrombosis.[7 10 12 14-16 20] Consequently, the ninth American College of Chest Physicians guideline concluded that diagnostic strategies in patients with suspected recurrent deep vein thrombosis were never adequately evaluated.[8] As such, available evidence on this topic was graded as low. Our individual patient data meta-analysis is the largest present dataset available to tackle the problem of how to diagnose deep vein thrombosis in patients with a history of the condition: we included nearly 1000 patients with a history of deep vein thrombosis (n=941) with 220 confirmed cases in that particular sample. Using this dataset, we found that using the original Wells model an unacceptably high proportion of low risk patients combined with a negative D-dimer test result still have deep vein thrombosis: 2.5% (95% confidence interval 1.2% to 5.4%). In a sensitivity analysis, using the updated model, this failure rate was indeed lower. Adding one point to the original Wells score in patients with a history of deep vein thrombosis (the updated model), and defining low risk as a score ≤1 combined with a negative D-dimer test result leads to a failure rate of 1.0% (95% confidence interval 0.6% to 1.6%; see appendix IV on bmj.com for details of the model). This in turn enables a safe exclusion of deep vein thrombosis in (still) around 1 in every 3 patients (efficiency 27.2%, 95% confidence interval 19.2% to 37.0%; see appendix IV on bmj.com).

Fourthly, the American College of Chest Physicians guideline recommends D-dimer testing to rule out deep vein thrombosis

in those with a low score on the Wells rule; a negative D-dimer test result precludes the need for imaging.[8] Our individual patient data meta-analysis confirmed that recommendation, and as such that finding of our analyses is not novel. Yet, different thresholds on the Wells rule have been defined to categorise patients as low risk. The Wells rule can be either trichotomised (three level Wells rule; low score ≤0, moderate score 1 or 2, and high score >2) or dichotomised (two level Wells rule; unlikely score ≤1 and likely score ≥2).[1] The evidence summarised in the American College of Chest Physicians guideline suggested that moderately sensitive D-dimer tests (such as qualitative point of care D-dimer tests) are only safe in patients defined as low risk by the three level Wells rule (score ≤0), whereas highly sensitive laboratory based D-dimer testing (using assay specific cut-offs, usually <500 µg/L) can still safely exclude deep vein thrombosis in patients with a moderate pretest probability using the three level Wells rule (score 1 or 2).[39 40] Yet, evidence on the pooled estimates for the safety of excluding deep vein thrombosis using the two level Wells rule combined with either point of care or laboratory based D-dimer testing was unfortunately lacking.[8] A novel finding of our individual patient data meta-analysis is that we can now provide such pooled estimates on the probability of deep vein thrombosis in patients with an unlikely score on the Wells rule (≤1) combined with either a point of care (qualitative) or a laboratory (quantitative) based negative D-dimer test result. Overall, the probability of deep vein thrombosis in this group of patients is low: 1.2% (95% confidence interval 0.7% to 1.8%). In addition, we found no clinically relevant differences for this probability using either a quantitative or a qualitative D-dimer test (see table 6). Qualitative assays do seem to be more efficient (notably in low prevalence populations, see table 6), which is not surprising given their higher specificity (fewer false positives). To increase the efficiency of quantitative D-dimer testing, notably in older people, an age adjusted cut-off was recently proposed: age×10 in patients aged more than 50 years.[41 42] In our dataset, information on quantitative D-dimer testing was available in 1930 patients. An additional analysis on the safety and efficiency of using this age adjusted cut-off indeed showed that (combined with the Wells rule) this still is a safe approach (failure rate 0.6%, 95% confidence interval 0.2% to 2.3%) with an, albeit small, higher efficiency (21%, 95% confidence interval 13% to 33%), compared with 18% (14% to 23%). Further prospective validation studies are needed on the actual clinical value of using this age adjusted cut-off in daily practice.

Finally, some methodological limitations should be considered. For instance, in many of the included studies no explicit blinding of the assessor of deep vein thrombosis to the results of the Wells rule and D-dimer testing was described. This potentially can lead to the situation that the assessor (who interprets the results of compression ultrasonography) incorporates this information on the Wells rule and D-dimer testing in deciding on the presence or absence of deep vein thrombosis—that is, incorporation bias. This can lead to a biased agreement between both and thus an overoptimistic estimate of the value of the Wells rule and D-dimer testing in assessing deep vein thrombosis.[43] Yet, in clinical practice this situation is almost always present. We hypothesise that the actual impact of such bias is, although inevitably present, small. Also, most studies used compression ultrasonography as (part of) the reference standard. Yet, it is widely acknowledged that this reference standard is less able to identify recurrent events, notably ipsilateral ones.[8] The original studies included in our individual patient dataset often poorly reported if suspected recurrent events were ipsilateral or contralateral, and this can have an (albeit small) impact on our results. Finally, we performed additional analyses on the observed heterogeneity between studies and consequently the uncertainty around the probabilities of deep vein thrombosis for each Wells rule score. In this regard, prediction intervals were calculated (indicating the expected range of probabilities of deep vein thrombosis for each Wells rule score if a new study would be performed) that indeed showed relatively wide intervals and thus heterogeneity around our estimates (see table 4). Yet, this was mainly explained by differences in the prevalence of deep vein thrombosis in the different include studies. In table 4 all analyses are repeated with prevalence set at 15%, as this prevalence best reflects recent studies, particularly for a European or primary care based healthcare setting.[13 40]

Clinical recommendations and conclusions

Based on our large individual patient data meta-analysis, we can give various clinical recommendations. Firstly, the Wells rule can be used to assess the pretest probability of deep vein thrombosis in suspected patients as this probability increases with higher scores. Subsequently, this pretest probability of deep vein thrombosis can drive further diagnostic examinations. As summarised in the ninth American College of Chest Physicians guideline, this generally means compression ultrasonography in patients with a high score on the Wells rule and D-dimer testing in patients with a low score on the Wells rule.[8] Secondly, in terms of ruling-out deep vein thrombosis, an unlikely score (≤1) on the Wells rule combined with a negative D-dimer test result can safely exclude deep vein thrombosis in about 1 in every 3 patients, missing less than 2% of cases. No clinically relevant differences were observed on this safety margin of 2% with either a qualitative point of care or a quantitative laboratory based D-dimer assay. Thirdly, except for patients with an active malignancy or in patients with suspected recurrent thrombosis, these findings are robust across different clinically relevant subgroups. In patients with an active malignancy, exclusion of deep vein thrombosis is not only less safe but also, and perhaps more importantly, less efficient. For patients with suspected recurrent deep vein thrombosis, the use of the original Wells rule is less safe, whereas the modified model (where one extra point is given to patients with a history of deep vein thrombosis) enables the safe exclusion of deep vein thrombosis in suspected patients. Hence we recommend adding one extra point to the original Wells model in patients with a history of deep vein thrombosis.

Contributors: GJG and KGMM wrote the first version of the manuscript. NPAZ and KGMM provided statistical expertise. All analyses were performed by GJG and supervised by KGMM. All authors provided intellectual content and critically revised the manuscript. GJG had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding: This study received no funding.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) 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.

Ethical approval: Not required.

Data sharing: In this study individual patient data from 13 different studies are combined. Requests for data sharing can be sent to the first

What is already known on this topic

The Wells rule is a safe tool for the exclusion of deep vein thrombosis in unselected patients with an unlikely score, combined with a negative D-dimer test result

However, in various clinically important subgroups, such as patients with cancer or those presenting in primary care, the validity of the Wells rule is questioned

The optimal threshold on the Wells rule depending on type of D-dimer assay (qualitative or quantitative) used is currently unknown

What this study adds

The probability of deep vein thrombosis in patients with an unlikely Wells rule score (≤1) combined with a negative D-dimer test result is low (<2%), enabling exclusion of about one in three suspected patients

This finding is consistent in subgroups defined by sex, care setting (primary versus hospital), and type of D-dimer assay (qualitative versus quantitative)

In patients with cancer, the combination is neither safe nor efficient, and in patients with suspected recurrent disease one extra point should be added to the score to enable the safe exclusion of deep vein thrombosis

author of this paper (GJG), and are then discussed with all other coauthors.

Transparency: The lead author (GJG) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

1 Wells PS, Owen C, Doucette S, Fergusson D, Tran H. Does this patient have deep vein thrombosis? *JAMA* 2006;295:199-207.
2 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.
3 Reilly BM, Evans AT. Much ado about (doing) nothing. *Ann Intern Med* 2009;150:270-1.
4 Ten Cate-Hoek AJ, Prins MH. Management studies using a combination of D-dimer test result and clinical probability to rule out venous thromboembolism: a systematic review. *J Thromb Haemost* 2005;3:2465-70.
5 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.
6 Lee A. VTE in patients with cancer—diagnosis, prevention, and treatment. *Thromb Res* 2008;123:(Suppl 1):S50-4.
7 Schutgens RE, Beckers MM, Haas FJ, Biesma DH. The predictive value of D-dimer measurement for cancer in patients with deep vein thrombosis. *Haematologica* 2005;90:214-9.
8 Bates SM, Jaeschke R, Stevens SM, Goodacre S, Wells PS, Stevenson MD, et al. Diagnosis of DVT: antithrombotic therapy and prevention of thrombosis, 9 ed. American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2012;141(2 Suppl):e351S-418S.
9 Andreou ER, Koru-Sengul T, Linkins L, Bates SM, Ginsberg JS, Kearon C. Differences in clinical presentation of deep vein thrombosis in men and women. *J Thromb Haemost* 2008;6:1713-9.
10 Anderson DR, Kovacs MJ, Kovacs G, Stiell I, Mitchell M, Khoury V, et al. Combined use of clinical assessment and d-dimer to improve the management of patients presenting to the emergency department with suspected deep vein thrombosis (the EDITED Study). *J Thromb Haemost* 2003;1:645-51.
11 Anderson DR, Wells PS, Stiell I, MacLeod B, Simms M, Gray L, et al. Management of patients with suspected deep vein thrombosis in the emergency department: combining use of a clinical diagnosis model with D-dimer testing. *J Emerg Med* 2000;19:225-30.
12 Bates SM, Kearon C, Crowther M, Linkins L, O'Donnell M, Douketis J, et al. A diagnostic strategy involving a quantitative latex D-dimer assay reliably excludes deep venous thrombosis. *Ann Intern Med* 2003;138:787-94.
13 Buller HR, Ten Cate-Hoek AJ, Hoes AW, Joore MA, Moons KG, Oudega R, et al. Safely ruling out deep venous thrombosis in primary care. *Ann Intern Med* 2009;150:229-35.
14 Elf JL, Strandberg K, Nilsson C, Svensson PJ. Clinical probability assessment and D-dimer determination in patients with suspected deep vein thrombosis, a prospective multicenter management study. *Thromb Res* 2009;123:612-6.
15 Kearon C, Ginsberg JS, Douketis J, Crowther M, Brill-Edwards P, Weitz JI, et al. Management of suspected deep venous thrombosis in outpatients by using clinical assessment and D-dimer testing. *Ann Intern Med* 2001;135:108-11.
16 Kearon C, Ginsberg JS, Douketis J, Crowther MA, Turpie AG, Bates SM, et al. A randomized trial of diagnostic strategies after normal proximal vein ultrasonography for suspected deep venous thrombosis: D-dimer testing compared with repeated ultrasonography. *Ann Intern Med* 2005;142:490-6.
17 Kraaijenhagen RA, Piovella F, Bernardi E, Verlato F, Beckers EA, Koopman MM, et al. Simplification of the diagnostic management of suspected deep vein thrombosis. *Arch Intern Med* 2002;162:907-11.
18 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.
19 Schutgens RE, Ackermark P, Haas FJ, Nieuwenhuis HK, Peltenburg HG, Pijlman AH, et al. Combination of a normal D-dimer concentration and a non-high pretest clinical probability score is a safe strategy to exclude deep venous thrombosis. *Circulation* 2003;107:593-7.
20 Stevens SM, Elliott CG, Chan KJ, Egger MJ, Ahmed KM. Withholding anticoagulation after a negative result on duplex ultrasonography for suspected symptomatic deep venous thrombosis. *Ann Intern Med* 2004;140:985-91.
21 Toll DB, Oudega R, Bulten RJ, Hoes AW, Moons KG. Excluding deep vein thrombosis safely in primary care. *J Fam Pract* 2006;55:613-8.
22 Wells PS, Anderson DR, Rodger M, Forgie M, Kearon C, Dreyer J, et al. Evaluation of D-dimer in the diagnosis of suspected deep-vein thrombosis. *N Engl J Med* 2003;349:1227-35.
23 Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. *PLoS One* 2012;7:e32844.
24 Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;59:1087-91.
25 Groenwold RH, Donders AR, Roes KC, Harrell FE Jr, Moons KG. Dealing with missing outcome data in randomized trials and observational studies. *Am J Epidemiol* 2012;175:210-7.
26 Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychol Methods* 2002;7:147-77.
27 Koopman L, van der Heijden GJ, Grobbee DE, Rovers MM. Comparison of methods of handling missing data in individual patient data meta-analyses: an empirical example on antibiotics in children with acute otitis media. *Am J Epidemiol* 2008;167:540-5.
28 Abo-Zaid G, Guo B, Deeks JJ, Debray TP, Steyerberg EW, Moons KG, et al. Individual participant data meta-analyses should not ignore clustering. *J Clin Epidemiol* 2013;66:865-73 e4.
29 Austin PC, Goel V, van Walraven C. An introduction to multilevel regression models. *Can J Public Health* 2001;92:150-4.
30 Debray TP, Moons KG, Abo-Zaid GM, Koffijberg H, Riley RD. Individual participant data meta-analysis for a binary outcome: one-stage or two-stage? *PLoS One* 2013;8:e60650.
31 Greenland S. Principles of multilevel modelling. *Int J Epidemiol* 2000;29:158-67.
32 Marcucci M, Smith CT, Douketis JD, Tosetto A, Baglin T, Cushman M, et al. Patient-level compared with study-level meta-analyses demonstrate consistency of D-dimer as predictor of venous thromboembolic recurrences. *J Clin Epidemiol* 2013;66:415-25.
33 Urbach DR, Austin PC. Conventional models overestimate the statistical significance of volume-outcome associations, compared with multilevel models. *J Clin Epidemiol* 2005;58:391-400.
34 Greenland S. When should epidemiologic regressions use random coefficients? *Biometrics* 2000;56:915-21.
35 Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;342:d549.
36 Toll DB, Oudega R, Vergouwe Y, Moons KG, Hoes AW. A new diagnostic rule for deep vein thrombosis: safety and efficiency in clinically relevant subgroups. *Fam Pract* 2008;25:3-8.
37 Van der Velde EF, Toll DB, Ten Cate-Hoek AJ, Oudega R, Stoffers HE, Bossuyt PM, 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.
38 Palareti G, Cosmi B, Legnani C, Tosetto A, Brusi C, Iorio A, et al. D-dimer testing to determine the duration of anticoagulation therapy. *N Engl J Med* 2006;355:1780-9.
39 Fancher TL, White RH, Kravitz RL. Combined use of rapid D-dimer testing and estimation of clinical probability in the diagnosis of deep vein thrombosis: systematic review. *BMJ* 2004;329:821.
40 Geersing GJ, Janssen KJ, Oudega R, Bax L, Hoes AW, Reitsma JB, et al. Excluding venous thromboembolism using point of care D-dimer tests in outpatients: a diagnostic meta-analysis. *BMJ* 2009;339:b2990.
41 Schouten HJ, Geersing GJ, Koek HL, Zuithoff NP, Janssen KJ, Douma RA, 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:f2492.
42 Douma RA, le Gal G, Sohne M, Righini M, Kamphuisen PW, Perrier A, et al. Potential of an age adjusted D-dimer cut-off value to improve the exclusion of pulmonary embolism in older patients: a retrospective analysis of three large cohorts. *BMJ* 2010;340:c1475.
43 Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061-6.

Accepted: 28 January 2014

Cite this as: *BMJ* 2014;348:g1340

Tables

Table 1| Items included in Wells rule, and patient characteristics of individual patient dataset included in meta-analysis. Values are numbers (percentages) unless stated otherwise

Variables	Point	Values
Items in Wells rule*:		
Active cancer	1	834 (8.3)
Paralysis, paresis, or recent immobilisation of leg with plaster	1	613 (6.1)
Recently bedridden >3 days or major surgery <4 weeks	1	1111 (11.1)
Localised tenderness of deep venous system	1	5098 (51.0)
Entire leg swollen	1	2767 (27.7)
Calf swelling >3 cm	1	3015 (30.1)
Pitting oedema	1	4784 (47.8)
Collateral superficial veins	1	1216 (12.2)
Alternative diagnosis as likely as or more likely than deep vein thrombosis	–2	4762 (47.6)
History of deep vein thrombosis†	1	941 (9.9)
Median (SD) age (years)	NA	59 (17)
Females	NA	6155 (61.5)

NA=not applicable.

*Scores can be trichotomised into low (≤0), moderate (1 or 2) and high (≥3), or dichotomised into deep vein thrombosis unlikely (≤1) or likely (>1).

†Only awarded one point in modified Wells rule. Also, data on history were available for 9461 patients (95% of total dataset of 10 002 patients).

Table 2| Characteristics of 13 included studies in individual patient data meta-analysis

Reference	Country	No of patients	Reference used	No (%) with DVT
Anderson 2000[11]	Canada	153	Compression ultrasonography or venography	26 (17)
Kearon 2001[15]	Canada	429	Compression ultrasonography or venography	61 (14)
Kraaijenhagen 2002[17]	The Netherlands	1756	Compression ultrasonography	411 (23)
Anderson 2003[10]	Canada	1075	Compression ultrasonography or venography	190 (18)
Bates 2003[12]	Canada	550	Compression ultrasonography	55 (10)
Schutgens 2003[19]	The Netherlands	814	Compression ultrasonography	318 (39)
Wells 2003[22]	Canada	541	Compression ultrasonography	121 (22)
Stevens 2004[20]	USA	436	Compression ultrasonography	42 (10)
Oudega 2005[18]*	The Netherlands	1295	Compression ultrasonography	289 (22)
Kearon 2005[16]	Canada	809	Compression ultrasonography or venography	42 (5)
Toll 2006[21]*	The Netherlands	791	Compression ultrasonography	126 (16)
Elf 2009[14]	Sweden	325	Compression ultrasonography or venography	52 (16)
AMUSE study 2009[13]*	The Netherlands	1028	Compression ultrasonography	131 (13)

DVT=deep vein thrombosis.

*Included only primary care outpatients; all other studies included secondary care outpatients.

Table 3| **Mean predicted probabilities for deep vein thrombosis (DVT) for each score on Wells rule in various subgroups. Values are percentages (95% confidence intervals)**

Wells score	Overall	Care setting		Malignancy	Sex		History of DVT
		Primary	Secondary or hospital		Male	Female	
–2	2.8 (2.1 to 3.8)	2.2 (1.5 to 3.0)	4.6 (2.6 to7.8)	5.9 (3.6 to 9.7)	4.5 (3.3 to 6.2)	2.0 (1.4 to 2.8)	5.7 (3.6 to 8.8)
–1	4.8 (3.6 to 6.3)	4.0 (2.9 to 5.4)	6.4 (3.8 to 10.7)	8.9 (5.8 to 13.4)	7.1 (5.3 to 9.5)	3.5 (2.6 to 4.7)	8.6 (5.8 to 12.4)
0	7.9 (6.1 to 10.2)	7.2 (5.4 to 9.6)	9.0 (5.5 to 14.5)	13.1 (9.2 to 18.4)	11.0 (8.4 to 14.2)	6.1 (4.6 to 8.0)	12.8 (9.2 to 17.5)
1	12.8 (10.2 to 16.2	12.8 (9.9 to 16.6)	12.5 (7.8 to 19.5)	19.0 (14.2 to 24.9)	16.6 (13.0 to 20.9)	10.4 (8.0 to 13.3)	18.6 (14.0 to 24.3)
2	20.0 (15.9 to 24.8)	21.7 (17.1 to 27.2)	17.1 (10.9 to 25.8)	26.6 (20.9 to 33.3)	24.2 (19.4 to 29.8)	17.1 (13.5 to 21.4)	26.3 (20.4 to 33.2)
3	30.0 (24.5 to 36.1)	34.4 (28.0 to 41.5)	22.9 (14.9 to 33.4)	36.0 (29.3 to 43.4)	34.0 (27.9 to 40.7)	26.9 (21.6 to 32.8)	35.8 (28.3 to 44.0)
4	42.2 (35.5 to 49.3)	49.8 (42.1 to 57.5)	30.0 (20.1 to 42.1)	46.6 (38.7 to 54.7)	45.4 (38.1 to 52.9)	39.5 (32.8 to 46.8)	46.5 (37.3 to 55.9)
5	55.6 (48.3 to 62.6)	65.2 (57.5 to 72.1)	38.2 (26.4 to 51.5)	57.5 (48.5 to 66.1)	57.3 (49.3 to 64.8)	53.8 (46.1 to 61.4)	57.5 (46.9 to 67.9)
6	68.1 (61.2 to 74.4)	77.9 (71.5 to 83.3)	47.0 (33.7 to 60.8)	67.7 (58.0 to 76.1)	68.3 (60.5 to 75.3)	67.5 (59.9 to 74.3)	67.9 (56.5 to 77.5)
7	78.5 (72.6 to 83.4	87.0 (82.2 to 90.6)	56.2 (41.8 to 69.5)	76.5 (66.7 to 84.1)	77.7 (70.6 to 83.5)	78.8 (72.2 to 84.1)	76.7 (65.5 to 85.2)
8	86.2 (81.2 to 89.8)	92.6 (89.5 to 94.9)	64.9 (50.3 to 77.1)	83.5 (74.3 to 89.8)	84.9 (78.9 to 89.4)	86.9 (81.1 to 90.7)	83.7 (73.4 to 90.6)

Table 4| **Comparison of model estimates with actual observed rates of deep vein thrombosis (DVT). Values are percentages unless stated otherwise**

		Model estimates				Observed data	
		Overall analysis		Heterogeneity analysis, prevalence set at 15%			
Wells score	No of patients	Model estimate	Prediction interval	Model estimate	Prediction interval	Mean observed DVT rate	Range * of observed data
–2	970	2.8	1.0-7.4	1.9	1.2-3.0	3.5	0-13.1
–1	1440	4.8	1.8-12.0	3.1	2.0-4.8	5.4	0-10.4
0	1870	7.9	3.1-18.9	5.2	3.3-8.1	8.1	1.6-23.0
1	1867	12.8	5.1-28.4	8.6	5.6-13.0	13.3	4.9-39.4
2	1583	20.0	8.5-40.4	13.8	9.1-20.2	23.9	8.2-42.5
3	1110	30.0	13.6-53.7	21.3	14.5-30.2	36.3	15.4-50.5
4	763	42.2	21.2-66.5	31.6	22.3-42.5	45.5	22.4-93.5
5	304	55.6	31.5-77.3	43.9	32.7-55.9	57.2	28.6-100
6	80	68.1	43.9-85.4	57.1	45.0-68.4	50.0	0-86.7
7	13	78.5	57.1-90.9	69.4	57.9-78.8	61.5	42.9-100
8	2	86.2	69.3-94.5	79.4	69.8-86.5	50.0	0-100

*Observed in individual studies (see Appendix III on bmj.com for raw data). Heterogeneity was observed in our model estimates, as demonstrated by relatively wide prediction intervals and range of actual observed DVT rates for each Wells score. Adding prevalence as a covariate to the model resulted in smaller prediction intervals, demonstrating that much heterogeneity was explained by differences in prevalence of DVT over the included studies in our individual patient dataset (range 5-39%). In this table, prevalence is set at 15%, as this best reflects the prevalence of DVT in recent studies.

Table 5| Failure rate and efficiency of excluding deep vein thrombosis (DVT) using Wells rule and D-dimer testing in various subgroups. Values are percentages (95% confidence intervals)

Accuracy measures	Overall	Care setting		Malignancy	Sex		History of DVT
		Primary	Secondary or hospital		Male	Female	
Failure rate*	1.2 (0.7 to 1.8)	1.4 (0.9 to 2.3)	0.9 (0.0 to 1.9)	2.2 (0.5 to 8.6)	1.4 (0.8 to 2.6)	1.0 (0.6 to 1.8)	2.5 (1.2 to 5.4)
Efficiency†	28.9 (20.3 to 39.5)	32.8 (21.8 to 46.1)	23.1 (12.8 to 38.3)	9.1 (5.5 to 14.7)	24.2 (16.5 to 34.1)	32.0 (22.6 to 43.2)	30.0 (20.2 to 42.2)

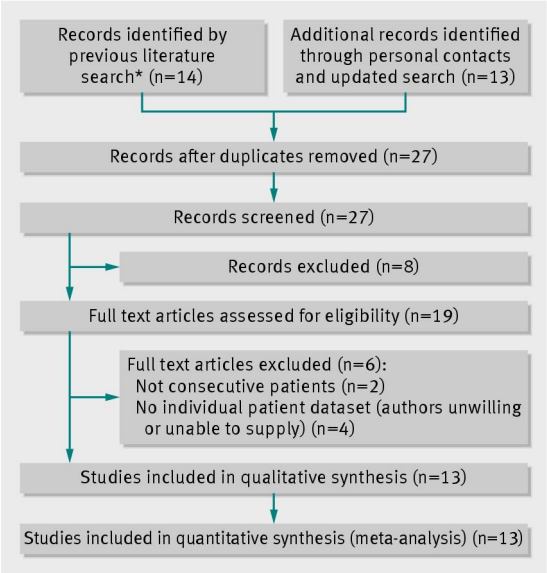
*Defined as mean predicted probability of DVT in patients with an unlikely score on the Wells rule (≤1), combined with a negative D-dimer test result.
†Defined as proportion of patients in this low risk group.

Table 6| Failure rate and efficiency of excluding deep vein thrombosis (DVT) using the Wells rule and D-dimer testing, depending on type of D-dimer assay and prevalence of DVT. Values are percentages (95% confidence intervals)

	Prevalence 5%		Prevalence 15%		Prevalence 40%	
Accuracy measures	**Quantitative D-dimer**	**Qualitative D-dimer**	**Quantitative D-dimer**	**Qualitative D-dimer**	**Quantitative D-dimer**	**Qualitative D-dimer**
Failure rate*	0.1 (0 to 0.2)	0.2 (0.1 to 0.4)	0.4 (0.3 to 0.6)	0.6 (0.4 to 0.9)	0.8 (0.6 to 1.1)	1.2 (0.8 to 1.7)
Efficiency†	23 (10 to 45)	49 (29 to 70)	21 (13 to 32)	46 (35 to 57)	19 (14 to 24)	42 (36 to 50)

*Defined as mean predicted probability of DVT in patients with an unlikely score on the Wells rule (≤1), combined with a negative D-dimer test result.
†Defined as proportion of patients in this low risk group.

Figure



Flow diagram of included studies. *Based on literature search performed by Wells et al[1]