Thrombosis Research 169 (2018) 82–86

Contents lists available at ScienceDirect

Thrombosis Research

journal homepage: www.elsevier.com/locate/thromres

Full Length Article

"HERDOO2" clinical decision rule to guide duration of anticoagulation in women with unprovoked venous thromboembolism. Can I use any D-Dimer?

Marc A. Rodger*, Gregoire Le Gal, Nicole J. Langlois, Barron Gin, Ranjeeta Mallick, Antonio Giulivi, Marisa Freedman, Michael J. Kovacs, for the REVERSE II investigators

The Ottawa Hospital Research Institute, 501 Smyth Rd, Ottawa, ON K1H 8L6, Canada

ARTICLE INFO

Keywords:
D-Dimer
Concordance
Kappa
Risk stratification
Venous thromboembolism

ABSTRACT

Background: The "HERDOO2 rule" is a prospectively validated clinical decision rule used to identify low-risk women who can safely discontinue anticoagulants after completing 5–12 months of anticoagulant treatment for unprovoked venous thromboembolism. The VIDAS® D-Dimer (DD) assay, a component of the rule, was used in the derivation and validation of the rule at half the usual diagnostic cut-point for exclusion of venous thrombosis. It is unknown if other commercial DD assays used at a corresponding cut-point will categorize patients at high concordance with the VIDAS® DD.

Objective: To determine if other available automated quantitative DD assays have high enough concordance with the VIDAS® DD assay to allow their use within the "HERDOO2" clinical decision rule.

Methods: Frozen plasma samples from a sub-set (n = 248) of female participants in the "HERDOO2" validation study were tested using five DD assays: VIDAS®, Innovance®, HemosIL®, Tina-quant® and Liatest®, with duplicate testing for 50 samples. First, using the mean DD for 50 samples with duplicate results, we determined the optimal cut-point values for each test that corresponded with a VIDAS® DD result of 250 μg/L using linear regression analysis. Next, kappa analysis was conducted on the DD results of the remaining 198 samples to determine concordance between each tested DD at the respective optimal cut-point and the VIDAS® DD at 250 μg/L. In a separate analysis we determined the concordance at half the usual venous thrombosis exclusion cut-point.

Results: Regression analysis of the DD results in 50 samples identified the optimal cut-point for each DD assay to match a VIDAS® DD cut-point of 250 μg/L: Innovance® 177 μg/L, Liatest® 233 μg/L, Tina-quant® 48 μg/L and HemosIL® 56 μg/L. Next, in 198 different samples, the concordance of VIDAS® DD (≥ 250 μg/L or < 250 μg/L) was explored at the optimal cut-point of the other DD assays. The concordance was poor for all DD assays: Innovance® (kappa 0.38 (95% CI, 0.26–0.51)), Liatest® (kappa 0.38 (95% CI, 0.25–0.50)), HemosIL® (kappa 0.36 (95% CI, 0.23–0.49)) and Tina-quant® (kappa 0.30 (95% CI, 0.16–0.43)). Similar poor concordance was identified using half of the diagnostic DD cut-point for each tested assay: Innovance® (kappa 0.44 (95% CI, 0.32–0.56)), Liatest® (kappa 0.38 (95% CI, 0.25–0.51)), HemosIL® (kappa 0.04 (95% CI, −0.01—0.08)) and Tina-quant® (kappa 0.04 (95% CI, −0.004–0.07)).

Conclusion: The "HERDOO2 rule" is the only prospectively validated clinical decision rule that can be used to identify low-risk women with unprovoked venous thrombosis who can safely discontinue anticoagulants. An important implementation issue is whether any commercial DD assay can be used in the HERDOO2 rule, and at what cut-point. Our analysis shows that the HemosIL®, Innovance®, Liatest® and Tina-quant® DD assays should not be used in the "HERDOO2" rule due to poor concordance with the VIDAS® DD assay and unacceptable misclassification of women at high and low risk of recurrent venous thrombosis.

1. Background

Venous thromboembolism (VTE) is a common, potentially fatal yet treatable condition. The duration of anticoagulation for a first unprovoked venous thromboembolism (VTE) is considered one of the most important unanswered questions in VTE management [1].

Current guidelines suggest indefinite anticoagulation in patients with a first unprovoked VTE without a high risk of bleeding [2].

* Corresponding author.
E-mail address: mrodger@ohri.ca (M.A. Rodger).

https://doi.org/10.1016/j.thromres.2018.07.020
Received 1 June 2018; Received in revised form 6 July 2018; Accepted 14 July 2018
Available online 17 July 2018

Box 1
HERDOO2 clinical decision rule to identify low-risk women with unprovoked venous thromboembolism who can discontinue anticoagulants after 5–12 months of anticoagulant therapy.

Men continue and HERDOO2
All men continue oral anticoagulants
Women with 2 or more of the following features should continue oral anticoagulants:

1) HER: any Hyperpigmentation, Edema, Redness of either lower extremity
2) VIDAS D- Dimer: ≥250 μg/L
3) Obesity: BMI ≥ 30 kg/m^2
4) Older age: ≥65 years

However, these are low grade recommendations and do not support individualized treatment decisions based on risk of VTE recurrence.

The HERDOO2 rule (see Box 1) has been derived [3] and validated [4] to identify low risk women who can safely discontinue anticoagulants after completing 5–12 months of anticoagulant therapy. Women with a HERDOO2 score of 0 or 1 have a 3% risk of recurrent VTE in the first year after discontinuing anticoagulants. A VIDAS® D-Dimer (bioMérieux, Marcy L'Etoile, France) is a component of the rule and was used in both the derivation and validation studies. Importantly, the VIDAS® D-Dimer (DD) cut-off level of 250 μg/L, used for the application of the HERDOO2 rule, is half of the level used for the exclusion of VTE in patients with a suspected diagnosis of VTE (500 μg/L). A crucial implementation issue when applying the rule in varied clinical settings is whether other commercial DD assays can be used in the HERDOO2 rule and at what cut-point. Clinicians, patients, policymakers, and regulatory authorities will need to be confident that other DD assays can lead to near identical HERDOO2 classification prior to permitting their use in making this decision about indefinite treatment.

The importance of accurate classification cannot be understated. A single DD measurement is used in the HERDOO2 rule to determine if a patient should receive indefinite anticoagulation or stop anticoagulation. False negatives may lead to inappropriate clinical decisions: stopping anticoagulants and exposing patients to a significant risk of recurrent VTE, which can be fatal. False positives may lead to unnecessary ongoing anticoagulation, exposing patients to a significant risk of major hemorrhage for an extended period, possibly lifelong.

The purpose of the study was to establish the appropriateness of using alternate DD assays when applying the HERDOO2 clinical decision rule. We sought to determine the concordance (kappa) of four commonly used, commercially available, automated quantitative DD assays compared to the VIDAS® D-Dimer Exclusion™ II, at an equivalent cut-point in samples from women with unprovoked VTE.

2. Methods

This D-Dimer concordance study was funded by bioMérieux, the makers of the VIDAS® DD assay. bioMérieux had no role in the design of this study or its execution, analyses, interpretation of the data, writing of the manuscript or decision to submit results. Our objectives were to 1) determine the cut-point for 4 commercially available quantitative DD assays that corresponds to a cut-point of 250 μg/L with the VIDAS® DD assay and then 2) compare the classification performance of four commercially available quantitative DD assays at the optimised cut-point to the VIDAS® D-Dimer Exclusion™ II assay at a cut-point of 250 μg/L.

In the REVERSE II study [4] we obtained consent from most enrolled patients to collect and store plasma samples for future research. Study participants were patients with unprovoked VTE and all samples were collected during oral anticoagulant treatment, 5–12 months after diagnosis. The samples were frozen and maintained at −80 °C. We retrieved samples from 248 randomly selected female participants who consented to future use of their samples. We limited this sub-study to women as in the HERDOO2 rule, DD results influence risk stratification only in women. Samples were shipped on dry ice. After thawing the samples, five commercially available DD tests (Table 1) were performed per their product inserts. The tests were conducted in three accredited clinical laboratories and one research laboratory in Ontario and Quebec, Canada, by trained laboratory personnel who were blinded to all clinical data, the results of any previous DD tests for the same participant, and the reference standard result. These laboratories were selected because they had the equipment for the DD assays and personnel that were experienced with the test.

To evaluate the reliability of the VIDAS® DD, we compared the results from the REVERSE II study where VIDAS® DD was measured at the individual study sites to the results of the central VIDAS® DD testing that was conducted for this study.

We determined the optimal cut-point for each individual DD assay that matches the VIDAS® DD cut-point of 250 μg/L by plotting separate

Table 1
D-Dimer assays and instrumentation.

D-Dimer assay	Instrument	Lower limit of detection
HemosIL® D-Dimer HS (Instrumentation Laboratory, Bedford Massachusetts, USA)	ACL-TOP 500 System (performed in London, Ontario, Canada)	203 μg/L
Innovance® D-Dimer (Siemens Healthcare Diagnostics, Erlangen, Germany)	CS-2100i System (performed in Ottawa, Ontario, Canada)	170 μg/L
Stago STA®-Liatest® D-Di (Diagnostica Stago, Asnières sur Seine, France)	Stago STA Satellite USB System (performed in Ottawa, Ontario, Canada)	270 μg/L
Tina-quant® D-Dimer Gen. 2 (Roche Diagnostics, Mannheim, Germany)	Roche/Hitachi cobas® c System (performed in Montreal, Quebec, Canada)	150 μg/L
VIDAS® D-Dimer Exclusion™ II (bioMérieux, Marcy-l'Etoile, France)	bioMérieux mini VIDAS® Analyzer (performed in Ottawa, Ontario, Canada)	50 μg/L

graphs for each individual DD result for 50 samples versus the VIDAS® DD results from the same 50 samples. Each DD test was run in duplicate and the mean result was calculated and used in this analysis. Linear Regression analysis was then conducted to identify the cut-point for each DD that corresponds to the VIDAS® DD at a 250 μg/L cut-point.

We then determined the between-assay classification performance of each individual DD, at their optimal cut-point, in comparison to the VIDAS® D-Dimer at a 250 μg/L cut-point using 198 different patient samples from women with unprovoked VTE enrolled in the REVERSE II study. We created a 2 × 2 table (individual DD positive or negative versus VIDAS® DD positive or negative). We then used these data to calculate a two rater un-weighted kappa statistic for each individual DD in comparison to the VIDAS® DD assay (the clinical gold standard).

Recognizing that most assays have higher lower limits of detection than the VIDAS® DD, we also conducted a sensitivity analysis in which we repeated the calibration and concordance exercise excluding data points with VIDAS® DD below the lower limit of detection of the other DD assays (see Table 1).

We also determined kappa for each DD at half their usual diagnostic cut-off compared to the VIDAS® DD at a 250 μg/L cut-point (half of the level used for the exclusion of VTE). We conducted this analysis because many clinicians might intuitively assume that this is a safe approach.

In general, a kappa score above 0.8 is considered excellent reliability, a kappa score between 0.6 and 0.8 is considered good reliability, a kappa score between 0.4 and 0.6 is considered moderate reliability and a kappa score below 0.4 is considered poor reliability [5]. However, more exact standards are necessary for this indication. A priori, we established that if a Kappa score of < 0.9 for a given DD was found, we would conclude that it should not be used in the HERDOO2 score as it will likely lead to a > 5% false positive and/or > 5% false negative rate, which would be clinically unacceptable given the lifelong consequences of the treatment decision guided by this clinical decision rule. Exact binomial 95% confidence intervals were calculated for each kappa.

We also report on the percentage misclassified as high risk compared to the VIDAS® DD and percentage misclassified as low risk compared to the VIDAS® DD for each DD to highlight the potential clinical consequences of misclassification.

A sample size of 200 was required to obtain a width of 0.1 around our point estimates of kappa for two observations [5]. Tight confidence intervals are required given the importance of this single measurement to make a lifelong decision.

3. Results

The women in the calibration and concordance study datasets were similar, with a mean age of 50 years and a mean BMI of 32 and 31, respectively (Table 2). Both groups had a mix of women classified as being at high and low risk of recurrent venous thromboembolism, according to the HERDOO2 Score, with a slightly higher proportion of low-risk women in both the calibration study (58%) and the concordance study (56.6%). D-Dimer is known to be affected by many factors including age and co-morbid conditions. The HERDOO2 rule was derived and validated with a D-Dimer cut-off of 250μg/L in all enrolled patients with unprovoked VTE.

Table 2
Participant characteristics.

Characteristic	Calibration study (*N* = 50)	Concordance study (*N* = 198)
Age, mean (SD)	50.06 (16.83)	49.83 (17.22)
BMI, mean (SD)	31.69 (8.18)	31.17 (7.99)
HERDOO score, mean (SD)	1.30 (1.09)	1.31 (1.10)
HERDOO2 ≥ 2, n (%)	21 (42.0%)	86 (43.4%)
HERDOO2 < 2, n (%)	29 (58.0%)	112 (56.6%)

3.1. Calibration results

The randomly selected samples used in the linear regression were from 50 women with unprovoked VTE. The average DD result from duplicate tests were used in the linear regression analysis to determine the relationship between each individual DD and the VIDAS® DD at 250 μg/L. The linear regression results are shown in Fig. 1 and summarized in Table 3.

3.2. Concordance results

The concordance between the optimized cut-points for the 4 commercially available DDs that corresponded to a VIDAS® DD at 250 μg/L and the VIDAS DD at 250 μg/L are shown in Table 4. According to the preselected criteria, concordance between the non-VIDAS® DDs at their equivalent cut-points and the VIDAS® DD at the 250 μg/L cut-point was poor. Use of these 4 tests would lead to important misclassification of risk with the HERDOO2 score, ranging from 14.29% misclassified with the Innovance® DD to 20.20% with the Tina-quant® DD (Table 4).

An additional analysis compared each of the four commercially available DDs at half their usual diagnostic cut-point to VIDAS® DD at 250 μg/L (Table 5). Similarly, at half their usual cut-point, three of the tests had poor concordance with the VIDAS® DD and the forth (Innovance®) had moderate concordance. We also measured concordance between the results of the VIDAS® DD tests conducted at local study facilities for the REVERSE II study to the results for the same women using the frozen samples collected at the same time, that were tested centrally using the same kits (VIDAS® D-Dimer Exclusion™ II) for this study (Table 5); with a kappa of 0.92, concordance is classified as excellent.

3.3. Sensitivity analysis

We could not perform the sensitivity analysis excluding data points below the lower limit of detection for the Tina-quant® assay due to insufficient data points above the Tina-quant® lower limit of detection (1 of 50 data points above the lower limit of detection), the Liatest® lower limit of detection (11 of 50 data points above the lower limit of detection) and the HemosIL® lower limit of detection (0 of 50 data points above the lower limit of detection). In a sensitivity calibration analysis with Innovance® DD, we had sufficient data points above the Innovance® lower limit of detection to conduct a linear regression (25 data points above the lower limit of detection). This sensitivity regression analysis yielded an optimal cut-point of 223 μg/L. However, concordance analysis at this cut-point still demonstrated only moderate concordance (Kappa 0.48 (95% CI 0.36–0.60).

4. Discussion

The VIDAS® DD was used to derive and validate the HERDOO2 rule. Use of the VIDAS® DD at a cut-point of 250 μg/L within the context of the HERDOO2 rule has been clearly shown to permit identification of a low risk group of women with unprovoked VTE who can safely discontinue anticoagulants. Our current study suggests that the other widely-used, commercially available DD assays, at an optimised cut-point of half of their usual diagnostic cut-point, should not be used in the HERDOO2 rule.

Using any of the 4 alternate DD assays studied may result in important misclassification of women with the potential that 1) some women would be falsely labeled as high risk (low HERDOO2 score with VIDAS® < 250 μg/L but with use of another DD assay, classified as high HERDOO2 score) and recommended to unnecessarily continue anticoagulants indefinitely or 2) falsely labeled as low risk (high HERDOO2 score with VIDAS® but with use of another DD assay, classified as low HERDOO2 score) and erroneously recommended to discontinue anticoagulants. Indeed, in our misclassification percentage analysis we

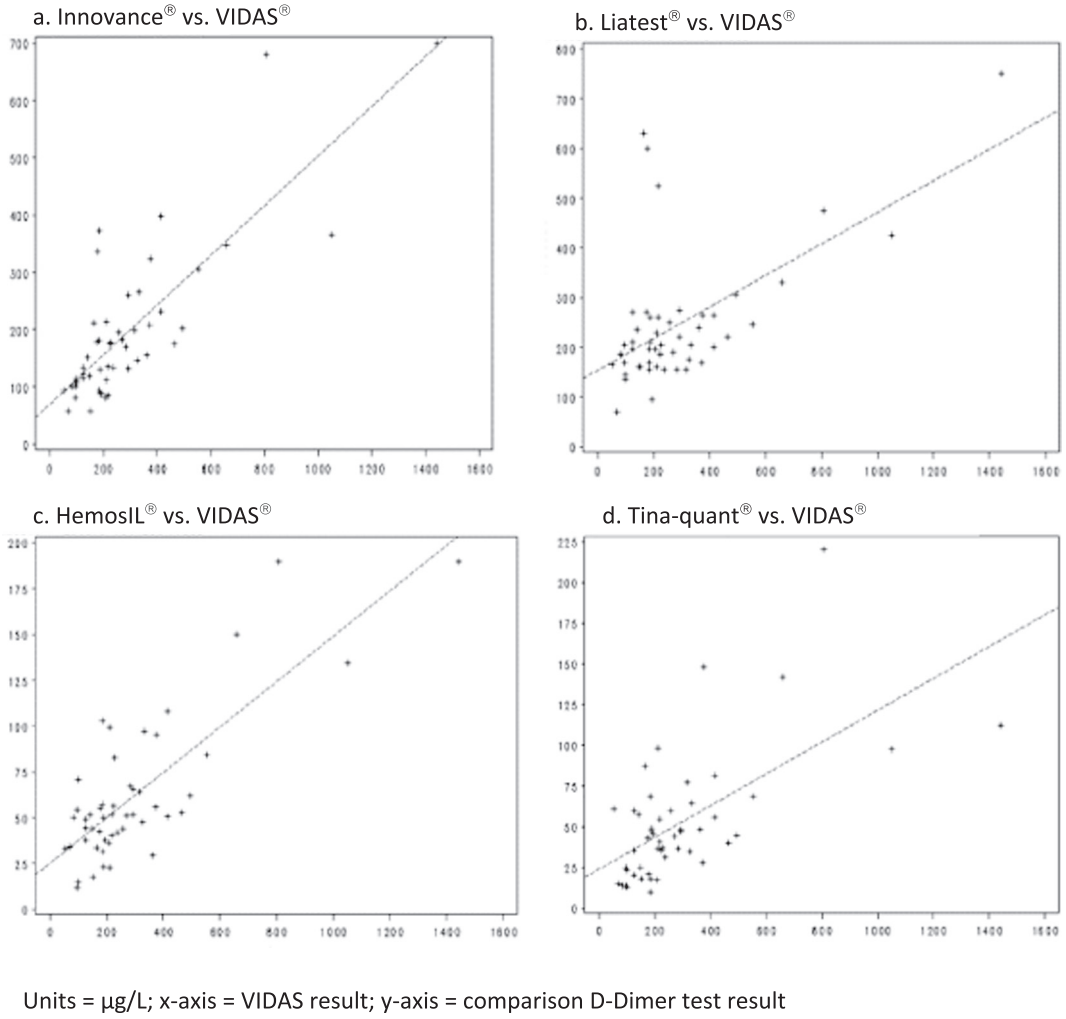

Fig. 1. Regression analysis scatterplots - four commercially available D-Dimers versus VIDAS® D-Dimer.

Table 3
Linear regression results – four commercially available D-Dimers and VIDAS® D-Dimer.

	R^2	Intercept	Beta	D-Dimer result corresponding to VIDAS® 250 µg/L
HemosIL®	0.64	25.1	0.12	56 µg/L
Innovance®	0.66	68.8	0.43	177 µg/L
Liatest®	0.35	153.6	0.32	233 µg/L
Tina-quant®	0.39	23.8	0.10	48 µg/L

showed an unacceptably high risk of misclassification with use of other DDs.

A complex mixture of products prevents standardization of DD results across tests and, therefore, makes it impossible to make direct comparisons of results obtained with different methods [6]. Systematic reviews of diagnostic accuracy studies using DD have also acknowledged the limitations of comparing estimated accuracy between DDs in patients with suspected VTE in published reports. These between study comparisons are challenged by differences in study design, patient characteristics such as age and co-morbid conditions, and the quality of reporting [7]. The recommended study design is a full comparative design where the same patients are tested using all DD methods; thereby eliminating between study differences. In this study we

Table 4
Concordance results – four commercially available D-Dimers at optimized cut-point and VIDAS® D-Dimer.

D-Dimer test (optimised cut-point[a])	Concordance with VIDAS® 250 µg/L Kappa (95% CI)	Low-risk women[b] misclassified as high risk (n/N, %)	High-risk women[c] misclassified as low risk (n/N, %)	All HERDOO2 misclassifications (n/N, %)
HemosIL® (56 µg/L)	0.35 (0.22, 0.48)	16/112 (14.29%)	16/86 (18.60%)	32/198 (16.16%)
Innovance® (177 µg/L)	0.38 (0.26, 0.51)	17/111 (15.32%)	11/85 (12.94%)	28/196 (14.29%)
Liatest® (233 µg/L)	0.38 (0.25, 0.50)	21/112 (18.75%)	16/86 (18.60%)	37/198 (18.69%)
Tina-quant® (48 µg/L)	0.30 (0.16, 0.43)	15/112 (13.39%)	25/86 (29.07%)	40/198 (20.20%)

[a] Optimised cut-point identified in the calibration exercise.
[b] Classified as low-risk with HERDOO2 using VIDAS® DD.
[c] Classified as high-risk with HERDOO2 using VIDAS® DD.

Table 5
Concordance results – four commercially available D-Dimers at 250 μg/L cut-point and VIDAS® D-Dimer.

D-Dimer test (half usual diagnostic cut-point)	Concordance with VIDAS® 250 μg/L Kappa (95% CI)
HemosIL® (250 μg/L)	0.04 (−0.01, 0.08)
Innovance® (250 μg/L)	0.44 (0.32, 0.56)
Liatest® (250 μg/L)	0.38 (0.25, 0.51)
Tina-quant® (250 μg/L)	0.04 (−0.004, 0.07)
VIDAS® (250 μg/L)	0.92 (0.86, 0.97)
Testing at local study centres[a]	

[a] Compared to central testing using thawed samples obtained at the same time.

implemented the recommended design. Another strength is that we optimised the classification performance of the other DDs by rigorously conducting an optimal cut-point analysis. This rigorous analysis featured the use of mean results of duplicate tests in a large sample of patients in whom the test is meant to be used (women with unprovoked VTE) and at the time it is meant to be applied (on anticoagulants after completion of short-term therapy, when a decision is being made regarding indefinite anticoagulation therapy). Despite optimising conditions for accurate classification, the other DDs performed poorly at their optimised cut-point.

It is notable and important that the lower limit of detection for the non- VIDAS® DD assays are significantly higher than the VIDAS® DD. Considering this fact, the findings of our study may not be surprising given that the DD cut-off in the HERDOO2 rule is low (250 μg/L), half the diagnostic cut-off, and near the lower limit of detection for the other DD assays. Indeed, the lower limit of detection of one assay (Liatest®) is above the 250 μg/L cut-point used in the HERDOO2 rule. These differences in lower limit of detection may be the principal reason the other DD assays had poor concordance with the VIDAS® DD in this indication given that this indication requires accuracy at lower levels of DD than in the exclusion of VTE diagnosis indication.

Many clinicians might assume that when applying the HERDOO2 rule, any DD can be used at half of its cut-point used in the diagnostic management of VTE, given that this was the optimal cut-point used with the VIDAS® assay in the derivation and validation studies. Our analysis suggests this approach would lead to a high risk of misclassification.

We acknowledge that it is disappointing that the other DDs did not perform well in the context of the HERDOO2 score. Using only the VIDAS® DD with the HERDOO2 rule creates a barrier to implementation that may limit use of the "Men continue and HERDOO2" rule. However, as we have shown in this study, plasma samples many be frozen, shipped and stored for later central testing with a high concordance with onsite VIDAS® DD testing, permitting labs not equipped with VIDAS® DD testing to collaborate with labs offering this test.

In conclusion, only VIDAS® DD should be used in determining a HERDOO2 score to guide duration of anticoagulant therapy in women with unprovoked VTE. Use of other DDs could result in significant HERDOO2 score misclassification with potentially serious clinical consequences including unnecessarily continuing anticoagulants indefinitely in women at low risk of recurrent VTE and stopping anticoagulants in women at high risk of recurrent VTE.

Competing interests

Dr. Rodger received grant funding from bioMérieux for the HERDOO2 derivation study, validation study and the current project. The other authors have no competing interests to declare.

References

[1] C. Kearon, Indefinite anticoagulation after a first episode of unprovoked venous thromboembolism: yes, J. Thromb. Haemost. 5 (2007) 2330–2335.
[2] C. Kearon, E.A. Akl, J. Ornelas, A. Blaivas, D. Jimenez, H. Bounameaux, M. Huisman, C.S. King, T.A. Morris, N. Sood, S.M. Stevens, J.R.E. Vintch, P. Wells, S.C. Woller, L. Moores, Antithrombotic therapy for VTE disease: CHEST guideline and expert panel report, Chest 149 (2016) 315–352.
[3] M.A. Rodger, S.R. Kahn, P.S. Wells, D.A. Anderson, I. Chagnon, G. Le Gal, S. Solymoss, M. Crowther, A. Perrier, R. White, L. Vickars, T. Ramsay, M.T. Betancourt, M.J. Kovacs, Identifying unprovoked thromboembolism patients at low risk for recurrence who can discontinue anticoagulant therapy, CMAJ 179 (2008) 417–426.
[4] M.A. Rodger, G. Le Gal, D.R. Anderson, J. Schmidt, G. Pernod, S.R. Kahn, M. Righini, P. Mismetti, C. Kearon, G. Meyer, A. Elias, T. Ramsay, T.L. Ortel, M.V. Huisman, M.J. Kovacs, Validating the HERDOO2 rule to guide treatment duration for women with unprovoked venous thrombosis: multinational prospective cohort management study, BMJ 356 (2017) j1065.
[5] V.F. Flack, A.A. Afifi, Sample size determinations for the two rater kappa statistic, Psychometrika 53 (1988) 321–325.
[6] M. Righini, A. Perrier, P. de Moerloose, H. Bounameaux, D-Dimer for venous thromboembolism diagnosis: 20 years later, J. Thromb. Haemost. 6 (2008) 1059–1071.
[7] M. Di Nisio, A. Squizzato, A.W. Rutjes, H.R. Buller, A.H. Zwinderman, P.M. Bossuyt, Diagnostic accuracy of D-dimer test for exclusion of venous thromboembolism: a systematic review, J. Thromb. Haemost. 5 (2007) 296–304.