REGULAR ARTICLE

blood advances

Decision theoretical foundations of clinical practice guidelines: an extension of the ASH thrombophilia guidelines

Benjamin Djulbegovic,[1] Iztok Hozo,[2] and Gordon Guyatt[3]

[1]Division of Medical Hematology and Oncology, Department of Medicine, Medical University of South Carolina, Charleston, SC; [2]Department of Mathematics, Indiana University Northwest, Gary, IN; and [3]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

Key Points

- The ASH thrombophilia guidelines modeling was suboptimal.
- The use of appropriate decision methodology leads to more accurate recommendations.

Decision analysis can play an essential role in informing practice guidelines. The American Society of Hematology (ASH) thrombophilia guidelines have made a significant step forward in demonstrating how decision modeling integrated within Grading of Recommendations Assessment, Developing, and Evaluation (GRADE) methodology can advance the field of guideline development. Although the ASH model was transparent and understandable, it does, however, suffer from certain limitations that may have generated potentially wrong recommendations. That is, the panel considered 2 models separately: after 3 to 6 months of index venous thromboembolism (VTE), the panel compared thrombophilia testing (A) vs discontinuing anticoagulants (B) and testing (A) vs recommending indefinite anticoagulation to all patients (C), instead of considering all relevant options simultaneously (A vs B vs C). Our study aimed to avoid what we refer to as the omitted choice bias by integrating 2 ASH models into a single unifying threshold decision model. We analyzed 6 ASH panel's recommendations related to the testing for thrombophilia in settings of "provoked" vs "unprovoked" VTE and low vs high bleeding risk (total 12 recommendations). Our model disagreed with the ASH guideline panels' recommendations in 4 of the 12 recommendations we considered. Considering all 3 options simultaneously, our model provided results that would have produced sounder recommendations for patient care. By revisiting the ASH guidelines methodology, we have not only improved the recommendations for thrombophilia but also provided a method that can be easily applied to other clinical problems and promises to improve the current guidelines' methodology.

Introduction

Current evidence–based clinical practice guidelines suffer from several deficiencies[1] including the following: "black-box" operation, a process with defined inputs and outputs but without complete understanding of its internal workings[2]; and what is referred to as "the integration problem," which entails a lack of a framework for explicit integration of patient preferences and trade-offs between treatment benefits and harms.[2] We have previously argued that the solution of the "black-box" and the "integration" problems is only possible within a decision-analytical framework.[1,3-7] Importantly, such a framework enables not only the logical and transparent integration of patient's values and preferences (V&P) and trade-offs between treatment benefits and harms[8,9] but protects against violation of the principles of rational decision-making.[3,4]

Submitted 14 February 2024; accepted 29 March 2024; prepublished online on *Blood Advances* First Edition 16 April 2024; final version published online 9 July 2024. https://doi.org/10.1182/bloodadvances.2024012931.

The full-text version of this article contains a data supplement.

Recently, the influential American Society of Hematology (ASH) thrombophilia guidelines[10] has made a significant step forward in demonstrating how decision modeling integrated within Grading of Recommendations Assessment, Development, and Evaluation (GRADE)[11] methodology can advance the field of guideline development. The panel developed transparent and easily understandable models that are important to end users.[10] However, perhaps in a desire to simplify its presentation, the panel may have chosen a less-than-optimal decision model, leading to what we refer as the "omitted choice bias"[12] that we illustrate below.

Methods

Using the state-of-the-art guidelines methodology GRADE,[11] the ASH panel developed 23 recommendations (R1-R23) for testing for thrombophilia in various circumstances leading to "provoked" vs "unprovoked" venous thromboembolism (VTE).[10] Figure 1A presents the conceptual thrombophilia model used by the ASH panel.[10] The model represents a truncated version of a decision tree, which we converted into the full decision-analytical model shown in Figure 1B. The panel considered the population of patients who were treated for initial VTE episodes for 3 to 6 months,[13] after which it compared VTE recurrence and major bleeding rates of the following management strategies: perform thrombophilia testing and administer a long-term, indefinite anticoagulation only to those patients who tested positive vs do not conduct testing for thrombophilia (Figure 1A) and discontinue treatment for all (as in the case of "provoked" VTE) or continue treatment indefinitely for all (as in the case of "unprovoked" VTE).

To calculate VTE recurrence and bleeding rates with each strategy, the ASH panel estimated thrombophilia prevalence and the risk ratio (RR) for recurrent VTE in patients with thrombophilia vs patients without thrombophilia (RR_t).[13] The panel relied on the ASH guidelines for the management of VTE[13] to estimate the effects (ie, RR_{rx} and major bleeding RR [RR_{bleed}]) of anticoagulant treatment compared with stopping anticoagulant therapy after completion of primary treatment for the initial VTE. For most recommendations, the ASH thrombophilia panel used the following input parameters: the median prevalence (P) of any thrombophilia was 38.0% (minimum 21.6%; maximum 59.5%); RR_t of 1.65 (95% confidence interval [CI], 1.28-2.47); RR_{rx} of recurrent VTE of 0.15 (95% CI, 0.10-0.23; relative risk reduction = $1-RR_{rx}$); RR_{bleed} of major bleeding on indefinite anticoagulant treatment was estimated at 2.17 (95% CI, 1.40-3.35), with the baseline major bleeding rate of 5 per 1000 patients (0.5%) at low risk and 15 per 1000 patients (1.5%) at high risk of bleeding per year.

To illustrate our approach, we focused on the first 6 thrombophilia panel recommendations in patients with low vs high risk of major bleeding (12 total recommendations). To address the superiority of a given strategy, the panel needed to estimate the overall risk of VTE recurrence without treatment (here denoted as p; not to be confused with P, the prevalence of thrombophilia shown in Figure 1A) after unprovoked VTE (R1), provoked after surgery (R2), provoked after a nonsurgical major transient risk factor, pregnancy, or associated with the use of oral contraceptives (R3-R5) or not specified as provoked or unprovoked VTE (R6). The ASH panel estimated the overall risk (probability) of VTE recurrence without treatment (p) to range from an average of 100 cases per 1000 patients (10%; scenario R1); 10 cases per 1000 patients (1%; R2); 50 cases per 1000 patients (5%; R3-R5); and 75 cases per 1000 patients (7.5%; R6).

After 3 to 6 months of index VTE, the panel modeled a comparison of thrombophilia testing (A) vs treat none (discontinuing anticoagulants; B) and testing (A) vs treat all (recommend indefinite anticoagulation to all patients; C). The panel considered 2 models: test (A) vs treat none (B) and test (A) vs treat all (C), separately instead of considering all relevant options simultaneously (A vs B vs C) in a single model. However, considering the comparisons separately instead of simultaneously can lead to what can be referred as the "omitted choice bias," named after the well-known "omitted variable bias,"[12] when not considering a relevant option can skew the results and lead to incorrect conclusions. A reader must not confuse omission choice bias we are referring to, with tendency of decision makers to prefer nonaction (omission) to action, which is also sometimes referred as "omission choice bias."

Although the authors presented an explicit model structure (Figure 1A), they did not leverage the entire apparatus of decision analysis[3,14] to generate recommendations. Instead, the panel relied on intuitive judgment, presumably informed by their model, to determine the optimal management approach for each clinical situation. Specifically, the panel stated the following[10]: "The following thresholds were used to judge the reduction in VTE (first time or recurrence): trivial, ≤5 events per 1000 patient-years; small, 5 to 20 per 1000 patient-years; and moderate, 20 to 50 per 1000 patient-years." The management strategy resulting in VTE recurrence below these thresholds was considered superior and thus recommended. However, how the panel weighed the trade-offs between VTE and major bleeding rates is unclear.

We converted 2 ASH models into a coherent single decision tree (Figure 1B). Many theoretical frameworks exist for solving a decision tree,[3] but most decision analyses use expected utility theory (EUT),[3] as in this study. EUT is the only theory of choice that satisfies all mathematical axioms of rational decision-making, ensuring that the choices are consistent with the deciders' V&P and trade-offs between treatment benefits and harms.[3] When choosing between different options, the most rational choice is that with the highest expected utility, regardless of statistical significance.[15]

From a decision-making perspective, the task is to determine the probability of VTE recurrence (threshold, P_t) above which we should commit to treatment. By "treatment," we refer to a commitment to a course of action that may include management consisting of treatment or diagnostic testing. We are indifferent between acting in favor of 1 management strategy over another when the net benefits and harms and decision-makers' V&P between these 2 strategies are identical.[3,6,16-18] Importantly, P_t depends only on benefits and harms (and decision-makers' V&P).[3,14,18] When considering only strategies for administering or stopping treatment, the threshold serves in the following way: if the probability of VTE recurrence is greater than P_t, we should give treatment (ie, anticoagulants in our case); if the probability of VTE recurrence is less than P_t, we should not give treatment.[3,14,18] When considering 3 possible strategies, continuing the treatment, administering the test and acting according to the results of the test, or stopping the treatment, we have 2 additional thresholds: the testing threshold P_{tt} and the treatment threshold P_{rx}. So, in total, we consider 5 possible management choices.

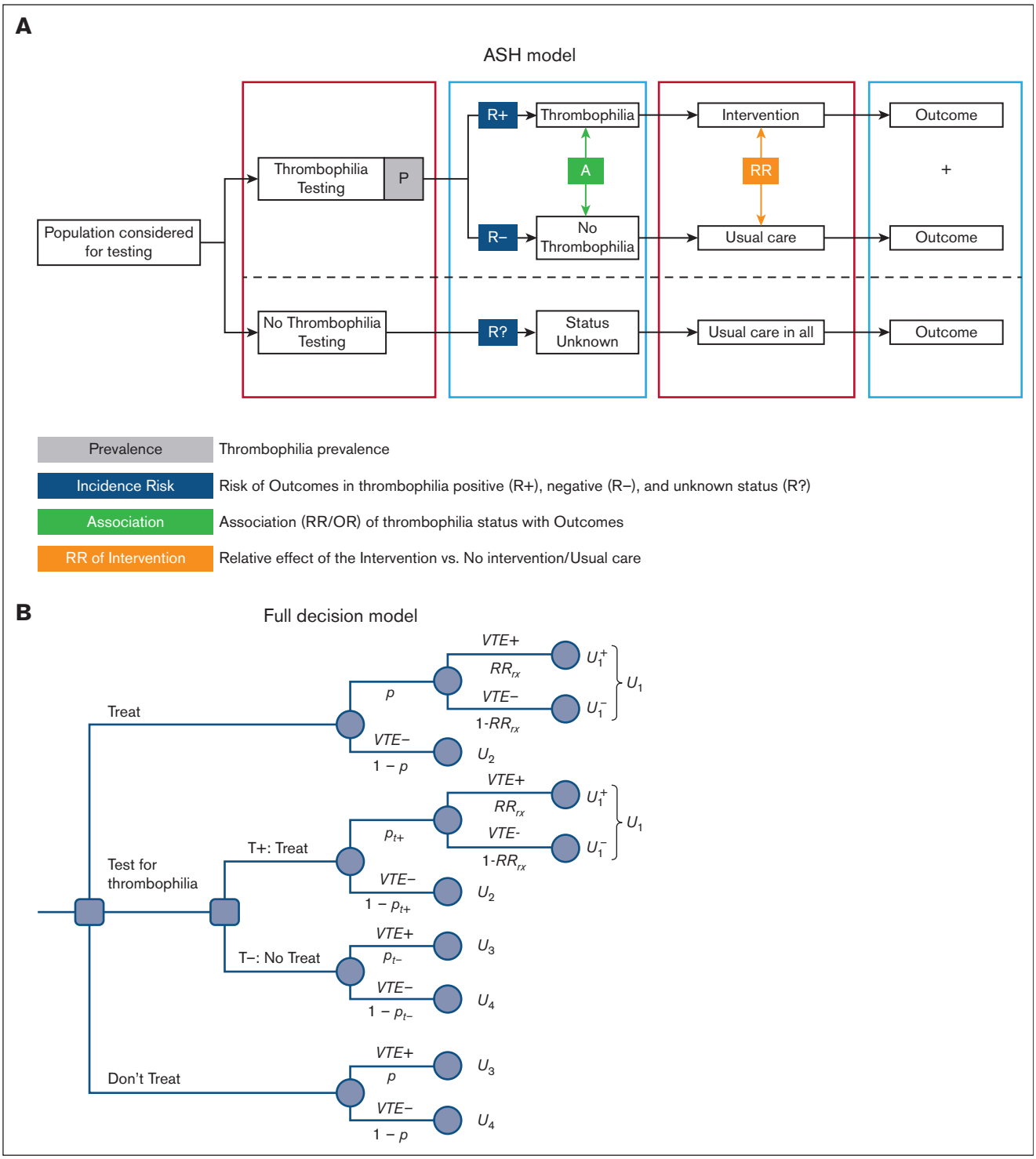


Figure 1. ASH two choices model vs a 3-choices decision model. (A) ASH modeling approach for determining the effect of thrombophilia testing. The model starts with the population considered for thrombophilia testing. Thrombophilia testing refers to testing for any type of thrombophilia or a specific type. Intervention is the course of action other than usual care. Depending on the particular question, this means prescribing thromboprophylaxis, withholding thromboprophylaxis, extending thromboprophylaxis, stopping thromboprophylaxis, withholding birth control pills, or withholding hormone replacement therapy. Usual care typically consists of short-term (3-6 months) anticoagulation (provoked VTE) or indefinite treatment (unprovoked VTE). P-thrombophilia prevalence (denoted in the manuscript as T_p); incidence risks of VTE recurrence is denoted in the manuscript as p_{t+} and p_{t-} for patients with (thrombophilia) positive results and for patients with negative test results, respectively; Association refers to RR for recurrent VTE in patients with thrombophilia vs patients without thrombophilia (RR_t); Relative effects of intervention (anticoagulant) on VTE recurrence (RR_{rx}) and bleeding (RR_{bleed}) compared with no intervention. (B) A decision tree showing a 3-choice clinical dilemma: administer treatment (anticoagulants) vs performing a diagnostic test (T) (thrombophilia testing) vs

Using the threshold decision-analytical model,[3] we can define these 3 thresholds (see Appendix 1 for complete derivations of the threshold equations):

$$P_t = \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}} \quad (1)$$

$$P_{tt} = \frac{RR_t \cdot T_p + (1 - T_p)}{RR_t} \cdot \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}} \quad (2)$$

$$P_{rx} = (RR_t \cdot T_p + (1 - T_p)) \cdot \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}} \quad (3)$$

In the equations above, H_{norx} refers to harms (ie, adverse events such as bleeding) observed in the "no treatment" arm; $T_p = P(T+)$ denotes the probability of (thrombophilia) positive test results and $P(T-) = 1\text{-}T_p$, the probability of negative test results; RR_t, RR_{bleed} and RR_{rx} represent variables as defined above. RV (relative value) refers to the patient's V&P; when $RV < 1$, the patient values avoiding outcomes of VTE more than avoiding harms of bleeding; if $RV > 1$, the patient places more importance on avoiding the harms of treatment than on the consequences of the disease; when the patient is indifferent to treatment harms and the consequences of the disease outcome, RV is 1. When RV is 1, the thresholds are solely determined by empirical evidence. ASH panel has compiled data on patients' V&P regarding VTE and bleeding, but their calculations relied only on empirical evidence. It is unclear how the panel considered integrating V&P in formulating their recommendations.

P_{tt} (test–no treatment threshold) refers to the pretest (prior) probability of VTE recurrence at which we are indifferent between no treatment and testing.[3,14,18] P_{rx} (test-treatment threshold) refers to the pretest (prior) probability of VTE recurrence at which we are indifferent between testing and treatment.[3,14,18] If the prior probability of VTE recurrence is less than P_{tt}, this guarantees that the posttest probability of VTE recurrence will always be less than treatment threshold, P_t, regardless of the test results.[3,14,18] Hence, we should not test and should withhold treatment under these circumstances. If the prior probability of VTE recurrence is greater than P_{rx}, this guarantees that the posttest probability of VTE recurrence will always be greater than treatment threshold, P_t, regardless of the test results.[3,14,18] If this relationship holds, we can give treatment without further testing.

Note how formal threshold equations 1 to 3 effectively capture everyday clinical intuition (Table 1).

Thus, in the case of thrombophilia recommendations, we contrast the overall risk (probability) of VTE recurrences (p) against these thresholds. According to our model, the thrombophilia testing should only be done for $P_{tt} < p < P_{rx}$. No testing/no treatment should be recommended for $< P_{tt}$. Treatment with anticoagulants should be recommended for $p > P_{rx}$.

Results

Reproducing the ASH thrombophilia results

The ASH panel presented its results as VTE and bleeding rates, counted separately. As explained below, such an approach introduces bias. Nevertheless, sometimes, we may wish to count events descriptively. If so, our model can be easily used to this effect.

Appendix 3 Table 1 in Appendix 3 illustrates using our model based on the data from the ASH report[10] to reproduce the panel's calculations for R1 (an identical approach can be used to reproduce the calculations for other recommendations).

Figure 2A displays these results graphically, comparing all management strategies (for R1). Figure 2B-D show the results of VTE and major bleeding rates for the remaining recommendations (R2, R3-R5, and R6) in low-risk bleeding settings. Figure 3A-D shows the same results assuming high-risk bleeding. Notably, even though the ASH report referred to the threshold in issuing its recommendations,[10] it is impossible to derive the VTE threshold from the method used by the ASH panel. The thresholds for the prevalence of VTE recurrence are 0, and the thresholds for the bleeding rates are undefined (see Appendix 2 for proof).

Comparison of the ASH thrombophilia recommendations with the threshold decision model

Table 2 displays calculations of the decision thresholds based on equations 1 to 3. These default calculations reflect the ASH thrombophilia evidence report and assume RV is 1, that is, that patients are equally concerned by the burden of VTE vs major bleeding. We explore this issue in detail below.

As explained above, to determine whether we should recommend thrombophilia testing, we contrast the probability of VTE recurrence (p) against the decision thresholds. The ASH panel estimated the probability of VTE recurrence (p) of 0.01 (1%) in the scenario guiding derivation of R2, 0.05 (5%) for R3 to R5, 0.075 (7.5%) to develop R6, and 0.1 (10%) in the clinical scenario resulting in R1. Figure 4 shows the results. Table 3 shows the ASH panel's R1 to R6 thrombophilia recommendations compared with the recommendations according to our threshold model.

For R1, the probability of VTE recurrence (p) of 10% is greater than P_{rx} = of 0.85% (low bleeding risk) and 2.57% (high bleeding risk), which means that the patients should be offered long-term treatment with anticoagulation without thrombophilia testing. This agrees with the ASH R1 recommendation.

For R2, in patients at low bleeding risk, p of 1% is greater than P_{rx} = of 0.85%. However, in the high bleeding risk, p of 1% is lower than P_{rx} = of 2.57% and P_{tt} of 1.56%. Thus, no thrombophilia testing should be offered in the low-risk bleeding scenario, but extended anticoagulant treatment should be recommended (Figure 4A). In the high–bleeding risk scenario, p of 1% is less than P_{tt} = of 1.56%, and no treatment nor thrombophilia testing should be offered to these patients. Thus, our analysis agrees with the ASH R2 recommendation only in the case of high bleeding risk (Figure 4B).

Figure 1 (continued) withholding therapy. Each treatment consists of the management strategies "treat all patients," "treat none," and "use thrombophilia test" to decide whether to treat. By "treatment," we refer to a commitment to a course of action that may include management consisting of treatment or diagnostic testing. p_{t+}=Pr$(D+|T+)$ refers to the probability of VTE recurrence when the thrombophilia test is positive (T+). U1 to U4, utilities (outcomes; see Appendix 1 for details).

Table 1. Intuitive presentation of the threshold model

Note how the manuscript's formal threshold equations 1-3 effectively capture everyday clinical intuition.
Equation 1 states that the administration of treatment depends on consideration of the benefits and harms of treatment adjusted for the patient's V&P regarding how they feel about the burden of disease (eg, VTE) vs the adverse events of treatment (eg, major bleeding).
Technically, we refer to the effects of treatment on patient outcomes as "utility."

(1) *Treatment threshold* (when tests are not taken into consideration)
Clinically, we are interested in finding out at which probability of disease or outcome we should act ["how high is "high" for us to give treatment; how low is "low" for us not to administer it]. *Intuitively when benefits do not exceed harms, we are uncertain how to proceed.* That is, we are at the treatment "threshold." According to a decision-analytical theory, the threshold is equal to the expected utility* of administering treatment vs not administering treatment. From here, we can obtain a simple formula for the treatment threshold[3]:

$$P_t = \frac{\text{Absolute risk of major bleeding (bleed)}}{\text{Relative risk reduction (RRR)}}$$

in which P_t is the probability of disease or outcome (eg, recurrence of VTE). Treatment should be given if the benefit of treatment exceeds its harms at the given probability of disease (eg, VTE recurrence) and patients' V&P. Thus, if the probability of VTE (pVTE) $>P_t$, we should administer treatment (eg, anticoagulants). If $P_t <$ pVTE, we should not give treatment. Importantly, our decisions whether to test (and act according to the test result) are also contrasted against P_t, which serves as an action threshold against which testing decisions are compared.

(2) Testing thresholds
(a) Test–no treatment threshold (P_{tt})
Equation 2 tells us what every physician intuitively knows: *when the probability of disease is fittingly very small, we can forgo testing and treatment.* It is also self-evident that P_{tt} must be smaller than P_t (ie, $P_{tt} < P_t$).‡ Generally, we can forgo testing when the pretest probability of disease is so low that even with a positive test result, the posttest probability would have always been below the action threshold P_t
(b) Test-treatment threshold (P_{rx})
Equation 3 also agrees with clinicians' intuition: *we don't always need diagnostic confirmation to act.* We can forgo testing when the pretest probability of disease is so high that even with a negative test result, the posttest probability would have always been above the action threshold, P_t. Here too, it is self-evident that P_{rx} must be larger than P_t (ie, $P_{rx} > P_t$).§

These results reflect an old clinical wisdom: "do not order a test that will not change your management."

To decide about thrombophilia testing, we contrast the overall risk (probability) of VTE recurrences (*p*VTE) against these thresholds. According to the threshold model, *the thrombophilia testing should only be done for* $P_{tt} < p\text{VTE} < P_{rx}$. No testing/no treatment should be recommended for $< P_{tt}$. Treatment with anticoagulants should be recommended for $p\text{VTE} > P_{rx}$ (Figure 4; also see the visual abstract).

*Expected utility is the average of all possible outcomes weighted by their corresponding probabilities.
†To simplify exposition, we avoid the consideration of V&P; please refer to the main manuscript and Appendix for full technical details. Note that there are many metrics for treatment benefits and harms that may result in different threshold formulas. Here, we show a formula pertinent to the treatment of patients at risk of VTE recurrence.
‡According to the expected EUT, the most widely used decision-analytical theory and that used in this manuscript.
§See footnote †.

For R3 to R5, *p* of 5% is greater than P_{rx} = of 0.85% (low bleeding risk) and P_{rx} = of 2.57% (high bleeding risk), meaning the patients should be offered long-term treatment with anticoagulation without thrombophilia testing. This result does not agree with the ASH R3 to R5 recommendations, which recommend thrombophilia testing and administering indefinite anticoagulation only to those patients who tested positive for thrombophilia (Table 3).

Finally, for R6, the probability of VTE recurrence (*p*) of 7.5% is greater than P_{rx} = of 0.85% (low bleeding risk) and P_{rx} = of 2.57% (high bleeding risk), indicating that the patients should be recommended long-term treatment with anticoagulation without thrombophilia testing. This conclusion also agrees with the ASH R6 recommendation (Table 3).

The analysis above assumed an RV of 1. When sensitivity analysis for RV was performed (ie, when RV /= 1), the results may differ. Table 4 shows the effect of RV on thrombophilia recommendations. For example, if we assume that the patient places twice as much importance on avoiding major bleeding than VTE recurrence (RV = 2), then recommendation 2 is consistent with the ASH panel's R2 ("do not test for thrombophilia; discontinue anticoagulants"[10]; Table 3). Interestingly, in the high–bleeding risk scenarios, the same conclusion ("don't test/don't treat") holds for RV ≥0.64, whereas testing and treatment strategy are recommended for the patients' with RV <0.641 and <0.388, respectively; that is, when the patient prefers avoiding VTE 1.5 (=1/0.641) and 2.5 (=1/0.388) times over bleeding. The ASH panel reports that the patients generally prefer avoiding VTE recurrence over bleeding.[10] In the low-risk bleeding scenario, we found that RV, in most cases, is unrealistically high (and never <1 for all recommendations) to affect our default recommendations for RV of 1. Nevertheless, it is conceivable that some patients fear the consequences of bleeding far more than they do of VTE. For instance, in case of R3 to R5, in low- and high-risk bleeding scenarios, patients might prefer to avoid major bleeding 5.8 to 9.6 times and 1.9 to 3.2 times more often than avoiding VTE, respectively. Under these conditions, our model aligns with the ASH recommendations (see "Discussion").

Discussion

In this study, we reanalyzed 6 × 2 recommendations made by the ASH thrombophilia panel[10] for various "provoked" vs "unprovoked" VTE clinical scenarios used as an example to illustrate the need to improve the broader field of decision-making and guideline development. By incorporating decision modeling within GRADE methods, the ASH thrombophilia panel has taken a substantial stride in advancing the guidelines development methodology and tackling "black-box" operation and "integration" challenges.[2] Although we believe that the application of decision modeling is the only logical and transparent method to facilitate the integration of all relevant ingredients required to make recommendations,[3] it is also essential to choose the correct model. The thrombophilia testing model developed by the ASH panel[10] generated 4 of 12 recommendations that proved inaccurate when judged against using a fully developed threshold model (Table 3).[3] Unlike the ASH panel, our model, assuming that patients placed equal value on avoiding VTE and bleeding, generated recommendations against thrombophilia testing in all scenarios considered.

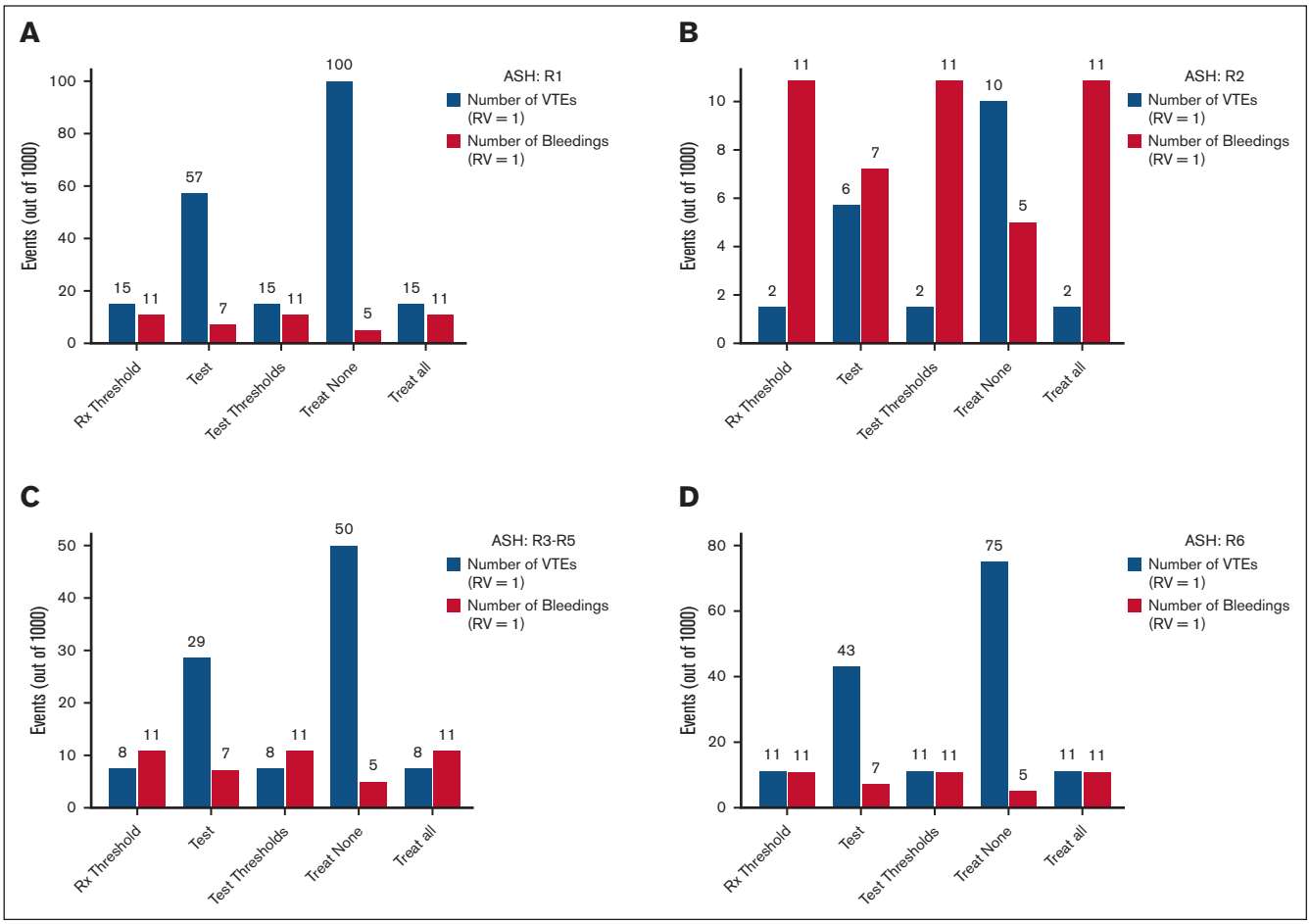


Figure 2. Number of VTE and major bleeding (low bleeding risk scenario). The impact analysis displaying the total number of VTE and major bleeding incurred for ASH panel recommendation R1 (A), R2 (B), R3 to R5 (C), and R6 (D) in the low–bleeding risk scenario. Five decision strategies are shown (from left to right): treat according to the threshold (Rx threshold; equation 3 in the manuscript); perform testing and act accordingly; test according to thresholds (equation 2 in the manuscript); treat none (give anticoagulants to no patient without testing); and treat all (provide anticoagulants for all patients without testing).

The discrepancy occurred because of the omitted choice bias[12]: the panel considered 2 models separately instead of considering all relevant options simultaneously. This resulted in rationally incoherent recommendations when judged against an appropriate decision model. Nevertheless, because the certainty of evidence was judged to be very low, the panel issued conditional (weak) recommendations, meaning that "most individuals in this situation would want the suggested course of action, but many would not."[10] Thus, on the surface, the ASH panel's judgment appears consistent with its recommendations, even if it disagreed with its model. Unfortunately, the panel never explained its deviation from the proposed model that apparently guided all the panel's recommendations. Why engage in modeling or develop guidelines if very low certainty of evidence almost always (barring some exceptions[19]) generates uncertain recommendations that may or may not be coherent with the underlying model structure?

The ASH panel has not performed any sensitivity analyses to assess the uncertainty range at which their recommendations could possibly switch. We have argued that it is precisely in these circumstances that modeling followed by judicious deliberation of the panel is most useful because it combines the explicitness and transparency of decision modeling with the considered panel's judgments.[1]

Somewhat surprisingly, we also concluded that extended anticoagulation should be offered to patients with low bleeding risk who developed surgery-related VTE. This conclusion disagreed with most current guidelines recommending discontinuing oral anticoagulation after 3 months of surgery-related VTE.[20] Nonetheless, the UK NICE guidelines state that "in low bleeding risk patients, the benefits of continuing anticoagulation treatment are likely to outweigh the risks."[21] The reason for this recommendation can be related to the extraordinarily high efficacy (RRR)/major bleeding (0.85/0.00585 = 145.3) ratio in the low-risk scenarios calculated based on the evidence presented in the ASH thrombophilia guidelines.[10] Such a high benefit/harms ratio generated a low test-treatment threshold (P_{rx}) = of 0.85%, which is below the 1% probability of VTE recurrence after surgery estimated by the ASH guidelines.[10] According to EUT, as long as the EU of 1 strategy is higher than the other, regardless of whether the differences are trivial or large, we should select that management option.[15] The earlier studies indicated a 0% risk of VTE after surgery,[22] but more recent studies suggested a risk of ~3%, noting that after a provoked

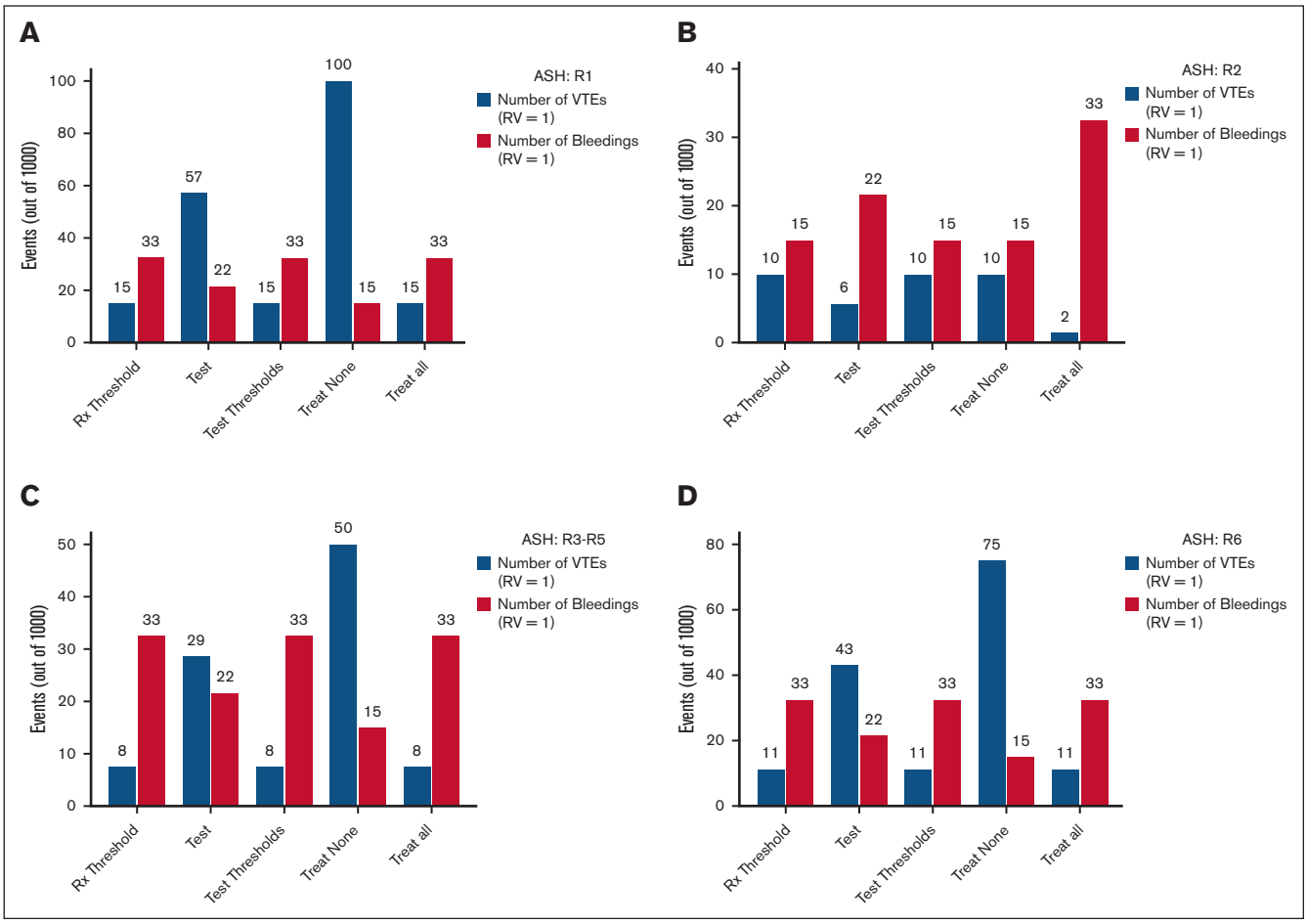


Figure 3. Number of VTE and major bleeding (high bleeding risk scenario). The impact analysis displaying the total number of VTE and major bleeding incurred for ASH panel recommendations R1 (A), R2 (B), R3 to R5 (C), and R6 (D) in the high–bleeding risk scenario. Five decision strategies are shown (from left to right): treat according to the threshold (Rx threshold; equation 3 in the manuscript); perform testing and act accordingly; test according to thresholds (equation 2 in the manuscript); treat none (give anticoagulants to no patient without testing); and treat all (provide anticoagulants for all patients without testing).

Table 2. Calculations of the decision thresholds for thrombophilia testing

(A) Low bleeding risk (5/1000)
Treatment threshold:

$$P_t = \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}} = \frac{1 \cdot (2.17 - 1) \cdot 0.005}{1 - 0.15} = .0069 = 0.69\%$$

Test vs no treatment threshold:

$$P_{tt} = \frac{RR_t \cdot T_p + (1 - T_p)}{RR_t} \cdot \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}} = \frac{1.65 \cdot 0.38 + (1 - 0.38)}{1.65} \cdot \frac{1 \cdot (2.17 - 1) \cdot 0.005}{1 - .15} = .0052 = 0.52\%$$

Test vs treatment threshold:

$$P_{rx} = (RR_t \cdot T_p + (1 - T_p)) \cdot \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}} = (1.65 \cdot 0.38 + (1 - 0.38)) \cdot \frac{1 \cdot (2.17 - 1) \cdot 0.005}{1 - 0.15} = 0.0085 = 0.85\%$$

(B) High bleeding risk (15 per 1000)
Treatment threshold:

$$P_t = \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}} = \frac{1 \cdot (2.17 - 1) \cdot 0.015}{1 - 0.15} = .0206 = 2.06\%$$

Test vs no treatment threshold:

$$P_{tt} = \frac{RR_t \cdot T_p + (1 - T_p)}{RR_t} \cdot \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}} = \frac{1.65 \cdot 0.38 + (1 - 0.38)}{1.65} \cdot \frac{1 \cdot (2.17 - 1) \cdot 0.015}{1 - .15} = .01560417 = 1.56\%$$

Test vs treatment threshold:

$$P_{rx} = (RR_t \cdot T_p + (1 - T_p)) \cdot \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}} = (1.65 \cdot 0.38 + (1 - 0.38)) \cdot \frac{1 \cdot (2.17 - 1) \cdot 0.015}{1 - 0.15} = .0257 = 2.57\%$$

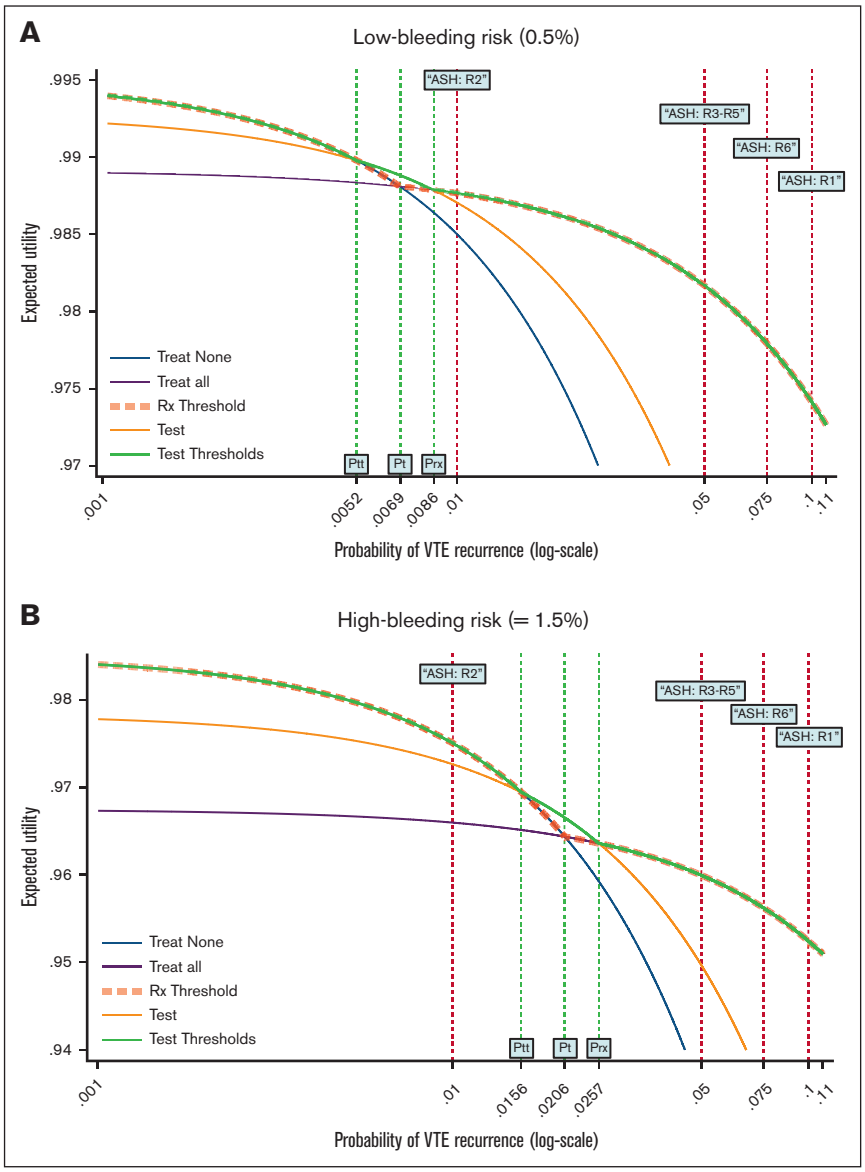


Figure 4. Threshold decision model analysis. The results of the threshold decision model analysis in the setting of the low bleeding risk (A) and high bleeding risk (B). The vertical lines (ASH R*) refer to the recommendations 1 to 6 by the ASH thrombophilia panel. Theoretical thresholds above or below which treatment vs thrombophilia testing vs no anticoagulant treatment should be given are denoted by P_{tt}, P_t, and P_{rx}, respectively (see equations 1-3). Note that because all ASH R lines are to the right side of P_{rx} that is, larger than the treatment threshold P_{rx} in a low-risk bleeding scenario, offering indefinite anticoagulant treatment to all patients represents the best management strategy (A). The same holds for ASH R1, R3 to R5, and R6 in the setting of high-risk bleeding. Because the vertical line ASH R2 is to the left, that is, lower than P_{tt} (test–no treatment threshold), discontinuing anticoagulation after 3 months of treatment after VTE due to surgery is recommended (see manuscript for details, Table 3, R2).

VTE, the risk may not return to the population baseline of 0.1 to 0.2 per 100 patients per year.[20,23] Therefore, under the assumed ASH risk of VTE of 1% after 3 months of anticoagulation for VTE provoked by surgery,[10] recommending extended anticoagulation is logically justifiable. Of course, the recommendations depend on the trustworthiness of these estimates. The ASH panel judged that the evidence used in their (and consequently our calculations) is of very low certainty because they were based on calculations with serious indirectness and imprecision of the estimates.[10] Such evidence can be equally right or wrong.[24] It is quite possible that different assumptions would have resulted in different recommendations. However, the panel presented the best evidence on the topic to date, often with a range of estimates. Unfortunately, in the case of overall risk for VTE recurrence, the panel was able to provide only point estimates (eg, the risk of unprovoked VTE was 100 per 1000 patients in the first year).[10]

Another disagreement between our models relates to recommendations R3 to R5 (Figure 4; Table 3). The reason that the strategy "test and treat only positive" remains inferior to the strategy of "treating all" patients according to the threshold model is because patients with false negative tests would have incorrectly not received treatment.

Our model is based on the well-known threshold model,[3] but equations 2 and 3 are novel derivations using the input parameters specified by the ASH thrombophilia panel.[10] The unavailability of these threshold formulas may have been a reason why the ASH thrombophilia panel did not use the full scope of decision modeling as outlined in our paper. If so, we urge the panel to update its recommendations accordingly.

One limitation that affected both our and ASH thrombophilia models is that it is based on the "average" data obtained from the

Table 3. Comparisons of the ASH thrombophilia panel's recommendations with the recommendations based on decision modeling

(R) numbers	Populations	ASH Considered strategies (after 3-6 mo of treatment)	ASH panel recommendations	Recommendations based on decision model with all 3 strategies
R1	Unprovoked VTE	Test vs treat all	Do not test for thrombophilia; recommend indefinite anticoagulant treatment to all patients	Do not test for thrombophilia; recommend indefinite anticoagulation (regardless of assumed bleeding risk)
R2	VTE provoked by surgery	Test vs treat none	Do not test for thrombophilia; recommend discontinuing anticoagulant treatment	Do not test for thrombophilia; recommend extended anticoagulation (low bleeding risk) Do not test for thrombophilia; discontinue anticoagulant treatment (high bleeding risk)
R3	VTE provoked by nonsurgical major transient risk factor	Test vs treat none	Test for thrombophilia; recommend indefinite anticoagulant treatment for patients with thrombophilia with stopping anticoagulant treatment for patients without thrombophilia	Do not test for thrombophilia; recommend indefinite anticoagulation (regardless of assumed bleeding risk)
R4	VTE provoked by pregnancy or postpartum	Test vs treat none	Test for thrombophilia; recommend indefinite anticoagulant treatment for patients with thrombophilia with stopping anticoagulant treatment for patients without thrombophilia	Do not test for thrombophilia; recommend indefinite anticoagulation (regardless of assumed bleeding risk)
R5	VTE associated with use of COC	Test vs treat none	Test for thrombophilia; recommend indefinite anticoagulant treatment for patients with thrombophilia with stopping anticoagulant treatment for patients without thrombophilia	Do not test for thrombophilia; recommend indefinite anticoagulation (regardless of assumed bleeding risk)
R6	An unspecified type of VTE (ie, not specified as provoked or unprovoked VTE)	Test vs treat all	Do not test for thrombophilia; recommend indefinite anticoagulant treatment to all patients	Do not test for thrombophilia; recommend indefinite anticoagulation (regardless of assumed bleeding risk)

literature, including the overall average and estimates of the probability of VTE recurrence. Indeed, individual patients are at different risks of VTE recurrence. As a result, we have called for developing more individualized recommendations by the guidelines panels.[1,25] This can be accomplished by integrating the best evidence from systematic reviews/meta-analyses on the average treatment effects with predictive models to estimate individualized disease risks or outcomes and threshold decision models.[1,25,26] Despite the plethora of models, some better validated than others,[27] the use of predictive models to help individualize the guideline recommendations has not been widely promoted in the VTE field. Indeed, some experts favor recommendations based on intuitive, holistic assessment over predictive models.[20] The GRADE method refers not only to the benefits and harms of the management and patients' V&P but also to resource use, feasibility, acceptability, and equity.[11]

However, all thrombophilia recommendations, both ASH's and ours, were driven by decision modeling without formally considering these other factors. Clearly considering, articulating, and transparently displaying such issues would be desirable. Indeed, the whole idea of combining decision analysis with the GRADE methodology is to generate recommendations using explicit, easily understood decision models based on the best existing evidence following time-honored cognitive scientists' advice: "to value formal principles of rationality" but then reflect on the appropriateness of further adjustments consistent with our explicit and implicit reasoning.[28-30] Still, the experience from other fields suggests that statistical rules typically outperform experts who rely on intuitive judgments.[31] Intuitive approaches to integrating complex elements such as synthesis of treatment benefits and harms with patients' V&P often do not agree with EUT models.[4,32] Under these

Table 4. Impact of patient's V&P on recommendations

				Probability of VTE recurrence			
				R1	R2	R3-R5	R6
Best strategy for RV range			Best strategy	0.1	0.01	0.05	0.075
Bleeding risk	Low risk	0.005	Treat none	RV ≥ 19.226	RV ≥ 1.923	RV ≥ 9.613	RV ≥ 14.419
			Test	11.652 ≤ RV ≤ 19.226	1.165 ≤ RV ≤ 1.923	5.826 ≤ RV ≤ 9.613	8.739 ≤ RV ≤ 14.419
			Treat all	RV ≤ 11.652	RV ≤ 1.165	RV ≤ 5.826	RV ≤ 8.739
	High risk	0.015	Treat none	RV ≥ 6.409	RV ≥ 0.641	RV ≥ 3.204	RV ≥ 4.806
			Test	3.884 ≤ RV ≤ 6.409	0.388 ≤ RV ≤ 0.641	1.942 ≤ RV ≤ 3.204	2.913 ≤ RV ≤ 4.806
			Treat all	RV ≤ 3.884	RV ≤ 0.388	RV ≤ 1.942	RV ≤ 2.913

RV (relative values) refers to patients' V&P; when $RV < 1$, the patient values avoiding outcomes of disease more than avoiding treatment harms; if $RV > 1$, the patient places more importance on avoiding the harms of treatment than on the consequences of the disease. When the patient is indifferent between treatment harms and the consequences of the disease outcome, RV is 1. When RV is 1, the thresholds are solely determined by empirical evidence (see Method section for details).

conditions, people may rely on non-EUT decision strategies,[4,32] such as anticipating regret of being wrong to drive their decisions.[33-37] This is also true for guideline panels. For example, we have previously demonstrated that ASH guideline panel for the management of pulmonary embolism relied on several decision theoretical approaches to formulate their recommendations, some of which were based on non-EUT constructs.[4] Which approach to take will largely depend on the type of problem, the consequences and the likelihood of being wrong, and contextual issues such as V&P, time, and available resources, etc.[38] However, there is a general consensus that we should always start with the EUT threshold model based on the best available evidence and further adjust it depending on the other elements considered essential for decision-making.[38] As we showed, our model clearly indicates the importance of consulting the patient's V&P, which may include consideration of costs and other burdens specified within the GRADE system. Nevertheless, it can be argued that "the most optimal decisions may be those that achieve coherence at both the normative and intuitive levels."[29] But, when these 2 types of knowledge do not agree, maximum efforts should be undertaken to reconcile the differences by exploring various theoretical approaches. This may be the most crucial reason why the ASH thrombophilia model should be updated. Providing a transparent and explicit explanation of the reasoning and analytical process through decision modeling, a solution to overcoming the challenges of integration and the "black-box" issue, when closely integrated with the GRADE methodology, can likely produce more coherent and accurate recommendations than relying solely on either decision models or the GRADE process alone. Adding the methodology described in this and other papers[1,3-5,16,25,26] can bring us to the goal of transparent, trustworthy, easily accessible, understandable, and highly accurate guidelines. The method we described in response to the ASH thrombophilia guidelines can be easily applied to other clinical problems and holds promise to improve the current guidelines' methods without requiring additional resources that complex decision modeling does.

Acknowledgment

This project was supported in part by grant number R01HS024917 from the Agency for Healthcare Research and Quality (PI: B.D.).

Authorship

Contribution: B.D. developed the conceptual idea and wrote the first draft of the manuscript; I.H. solved the model and developed program code; and G.G. revised the manuscript and improved its intellectual and research content

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profile: I.H., 0000-0003-2349-5707.

Correspondence: Benjamin Djulbegovic, Division of Medical Hematology and Oncology, Department of Medicine, Medical University of South Carolina, 39 Sabin St, MSC 635, Charleston, SC 29425; email: djulbegov@musc.edu.

References

1. Djulbegovic B, Hozo I, Cuker A, Guyatt G. Improving methods of clinical practice guidelines: from guidelines to pathways to fast-and-frugal trees and decision analysis to develop individualized patient care. *J Eval Clin Pract*. 2024;30(3):393-402.
2. Mercuri M, Baigrie B, Upshur REG. Going from evidence to recommendations: can GRADE get us there? *J Eval Clin Pract*. 2018;24(5):1232-1239.
3. Djulbegovic B, Hozo I. *Threshold Decision-making in Clinical Medicine: With Practical Application to Hematology and Oncology*. Springer Nature; 2023.
4. Djulbegovic B, Hozo I, Lizarraga D, Guyatt G. Decomposing clinical practice guidelines panels' deliberation into decision theoretical constructs. *J Eval Clin Pract*. 2023;29(3):459-471.
5. Djulbegovic B, Hozo I, Lizarraga D, et al. Evaluation of a fast-and-frugal clinical decision algorithm ('pathways') on clinical outcomes in hospitalized patients with COVID-19 treated with anticoagulants. *J Eval Clin Pract*. 2023;29(1):3-12.
6. Djulbegovic B, Hozo I, Mandrola J. Sorites paradox and persistence in overuse and underuse in healthcare delivery services. *J Eval Clin Pract*. 2023; 29(6):877-879.
7. Djulbegovic B, Hozo I, Lyman GH. Linking evidence-based medicine therapeutic summary measures to clinical decision analysis. *MedGenMed*. 2000; 2(1):E6.
8. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet*. 2017;390(10092):415-423.
9. Djulbegovic B, Hozo I, Li S-A, Razavi M, Cuker A, Guyatt G. Certainty of evidence and intervention's benefits and harms are key determinants of guidelines' recommendations. *J Clin Epidemiol*. 2021;136:1-9.
10. Middeldorp S, Nieuwlaat R, Baumann Kreuziger L, et al. American Society of Hematology 2023 guidelines for management of venous thromboembolism: thrombophilia testing. *Blood Adv*. 2023;7(22):7101-7138.
11. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008; 336(7650):924-926.
12. Wooldridge JM. "*Omitted Variable Bias: The Simple Case.*" *Introductory Econometrics: A Modern Approach*. Cengage Learning; 2009:89-93.
13. Ortel TL, Neumann I, Ageno W, et al. American Society of Hematology 2020 guidelines for management of venous thromboembolism: treatment of deep vein thrombosis and pulmonary embolism. *Blood Adv*. 2020;4(19):4693-4738.

14. Sox HC, Higgins MC, Owens D. *Medical Decision Making*. 2nd Ed. Wiley-Blackwell; 2013.
15. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ*. 1999; 18(3):341-364.
16. Djulbegovic B, Hozo I, Mayrhofer T, van den Ende J, Guyatt G. The threshold model revisited. *J Eval Clin Pract*. 2019;25(2):186-195.
17. Pauker SG, Kassirer J. The threshold approach to clinical decision making. *N Engl J Med*. 1980;302(20):1109-1117.
18. Pauker SG, Kassirer JP. Therapeutic decision making: a cost benefit analysis. *N Engl J Med*. 1975;293(5):229-234.
19. Yao L, Ahmed MM, Guyatt GH, et al. Discordant and inappropriate discordant recommendations in consensus and evidence based guidelines: empirical analysis. *BMJ*. 2021;375:e066045.
20. Becattini C, Cimini LA. Provoked vs minimally provoked vs unprovoked VTE: does it matter? *Hematology*. 2023;2023(1):600-605.
21. *Venous thromboembolic diseases: diagnosis, management and thrombophilia testing*. NICE Clinical Guidelines. National Institute for Health and Care Excellence; 2023.
22. Baglin T, Luddington R, Brown K, Baglin C. Incidence of recurrent venous thromboembolism in relation to clinical and thrombophilic risk factors: prospective cohort study. *Lancet*. 2003;362(9383):523-526.
23. Chua CC, Lim HY, Tacey M, Nandurkar H, Ho P. Retrospective evaluation of venous thromboembolism: are all transient provoking events the same? *Eur J Haematol*. 2017;99(1):18-26.
24. Djulbegovic B, Ahmed MM, Hozo I, et al. High quality (certainty) evidence changes less often than low-quality evidence, but the magnitude of effect size does not systematically differ between studies with low versus high-quality evidence. *J Eval Clin Pract*. 2022;28(3):353-362.
25. Hozo I, Djulbegovic B. Generalized decision curve analysis for explicit comparison of treatment effects. *J Eval Clin Pract*. 2023;29(8):1271-1278.
26. Hozo I, Guyatt G, Djulbegovic B. Decision curve analysis based on summary data. *J Eval Clin Pract*. 2024;30(2):281-289.
27. Cella CA, Knoedler M, Hall M, et al. Validation of the ONKOTEV risk prediction model for venous thromboembolism in outpatients with cancer. *JAMA Netw Open*. 2023;6(2):e230010.
28. Stanovich KE. How to think rationally about world problems. *J Intell*. 2018;6(2):25.
29. Djulbegovic B, Hozo I. Evidence and decision-making. *Cancer Treat Res*. 2023;189:1-24.
30. Djulbegovic B, Hozo I, Beckstead J, Tsalatsanis A, Pauker SG. Dual processing model of medical decision-making. *BMC Med Inform Decis Mak*. 2012; 12(1):94.
31. Clark DA. Human expertise, statistical models, and knowledge-based systems. In: Wright G, Bolger F, eds. *Expertise and Decision Support*. Springer; 1992:227-249.
32. Stanovich KE. *Rationality and the Reflective Mind*. Oxford University Press; 2011.
33. Cucchetti A, Djulbegovic B, Crippa S, et al. Regret affects the choice between neoadjuvant therapy and upfront surgery for potentially resectable pancreatic cancer. *Surgery*. 2023;173(6):1421-1427.
34. Djulbegovic B, Hozo I. Making decisions when no further diagnostic testing is available (expected regret theory threshold model). *Cancer Treat Res*. 2023;189:39-52.
35. Hozo I, Djulbegovic B. When is diagnostic testing inappropriate or irrational? acceptable regret approach. *Med Decis Making*. 2008;28(4):540-553.
36. Hozo I, Djulbegovic B. Clarification and corrections of acceptable regret model. *Med Decis Making*. 2009;29:323-324.
37. Djulbegovic B, Hozo I. When should potentially false research findings be considered acceptable? *PLoS Med*. 2007;4(2):e26.
38. Djulbegovic B, Hozo I. Which threshold model? *Cancer Treat Res*. 2023;189:93-99.

Appendix 1

Application of threshold decision model to the ASH Thrombophilia Guidelines

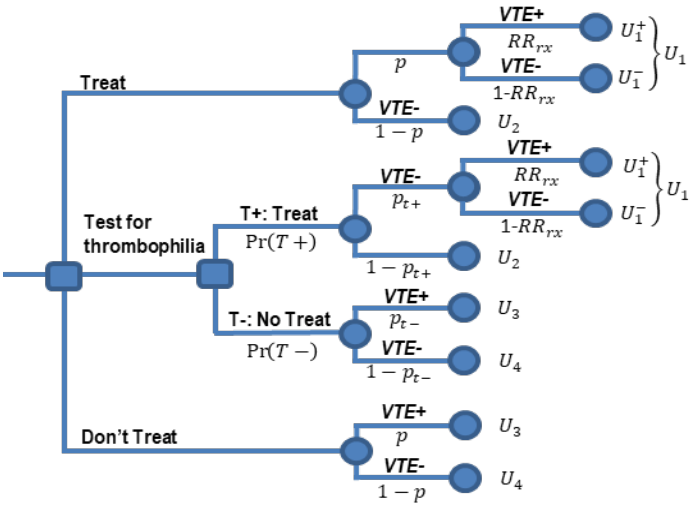
Clinical scenario

We have N = 1000 patients with an identical probability of VTE recurrence. We are faced with a dilemma on whether to administer anticoagulant treatment (Rx) to ALL patients (Treat all), whether to Test for thrombophilia and administer the anticoagulant treatment only to those patients who test positive (Test) or withhold anticoagulant treatment from all patients (Treat none). We want to select the strategy that minimizes the re-occurrence of VTE (event = VTE+) while controlling the treatment's harmful effects (bleeding).
The median % (min-max) for the prevalence of thrombophilia in our population is 38 (21.6 – 59.5). The Patients with thrombophilia are 1.65 more likely to have a recurrence of VTE than those who are thrombophilia-free ($RR_t = 1.65$ with 95% confidence interval $(1.28 - 2.47)$.
The patients receiving anticoagulant therapy are only 15% as likely to experience VTE recurrence event as those who did not receive the treatment, $RR_{rx} = 0.15$ with 95% CI $(0.10 - 0.23)$.
Bleeding occurs in all patients in our population, ranging from low risk (5 per 1000 patients) to high risk (15 per 1000 patients). Patients undergoing anticoagulant treatment have an increased risk of bleeding $RR_b = 2.17$ (95% CI, $1.40 - 3.35$).
The baseline (prior) probability of VTE recurrence (with no anticoagulants) in our population (p) ranges from a typical average of 100 cases per 1000 (10%) patients (recommendation#1, R1); 10 cases per 1000 patients (1%) (R2); 50 cases per 1000 patients (5%) (R3-R5); and 75 cases per 1000 (7.5%) patients (R6).

Expected Utility and the Decision Tree Model

If the baseline prevalence of VTE recurrence is denoted by p, we can calculate the posterior probabilities of VTE recurrence after administering the test for thrombophilia. We are given the following variables: $T^1{}_p = P(T+)$ and the risk ratio (RR_t) for a ratio of incidences of the event between patients with positive and negative test results. Then, we denote the posterior probabilities for the recurrence of VTE by $p_{t+} = \Pr(D+ \mid T+)$, $p_{t-} = \Pr(D+ \mid T-)$; the risk ratio relates these two posterior probabilities by the equation $p_{t+} = RR_t \cdot p_{t-}$.
The prevalence ($p = \Pr(D+)$) is then connected by

$$p = \Pr(D+) = \Pr(D+|T+) \cdot P(T+) + \Pr(D+|T-) \cdot P(T-)$$



[1] Note that the ASH panel refers to Tp (tested positive) as prevalence of thrombophilia P (see Fig 1 and example https://guidelines.ash.gradepro.org/profile/RPlrtP9SOqQ)

$$= p_{t+} \cdot T_p + p_{t-} \cdot (1 - T_p)$$
$$= p_{t+} \cdot T_p + \frac{p_{t+}}{RR_t} \cdot (1 - T_p) \qquad (A)$$

Solving for p_{t+} and p_{t-}, we have:

$$p_{t+} = \frac{p \cdot RR_t}{RR_t \cdot T_p + (1 - T_p)} \quad \text{and} \quad p_{t-} = \frac{p}{RR_t \cdot T_p + (1 - T_p)} \qquad (B)$$

If we denote the utilities as in the figure and use the combined utility $U_1 = RR_{rx}U_1^+ + (1 - RR_{rx}) \cdot U_1^-$, we have:

$$EU[Treat\ all] = p \cdot U_1 + (1 - p) \cdot U_2$$
$$EU[Test] = T_p \cdot (p_{t+} \cdot U_1 + (1 - p_{t+}) \cdot U_2) + (1 - T_p) \cdot (p_{t-} \cdot U_3 + (1 - p_{t-}) \cdot U_4)$$
$$EU[Treat\ none] = p \cdot U_3 + (1 - p) \cdot U_4$$

Thresholds

Therefore, we have three thresholds, i.e., three values for the **baseline (no treatment) probability** of VTE recurrence at which the expected utilities of two strategies in consideration will be equal.

Treat all vs. Treat none strategies (Test is not an option)

$$EU[Treat\ all] = EU[Treat\ none]$$
$$P_t \cdot U_1 + (1 - P_t) \cdot U_2 = P_t \cdot U_3 + (1 - P_t) \cdot U_4$$

$P_t = \frac{U_4 - U_2}{(U_1 - U_3) + (U_4 - U_2)}$	(1)

Test vs. Treat none:

$$EU[Test] = EU[Treat\ none]$$

Note that if we use equation (A) in the expression for $EU[Treat\ none]$, we have:

$$EU[Treat\ none] = \left(p_{t+} \cdot T_p + p_{t-} \cdot (1 - T_p)\right) \cdot U_3 + U_4 - \left(p_{t+} \cdot T_p + p_{t-} \cdot (1 - T_p)\right) \cdot U_4$$
$$= U_4 + p_{t+} \cdot T_p \cdot U_3 + p_{t-} \cdot (1 - T_p) \cdot U_3 - p_{t+} \cdot T_p \cdot U_4 - p_{t-} \cdot (1 - T_p) \cdot U_4$$

Comparing that with

$$EU[Test] = T_p \cdot (p_{t+} \cdot U_1 + (1 - p_{t+}) \cdot U_2) + (1 - T_p) \cdot (p_{t-} \cdot U_3 + (1 - p_{t-}) \cdot U_4)$$
$$= p_{t+} \cdot T_p \cdot U_1 + T_p \cdot U_2 - p_{t+} \cdot T_p \cdot U_2 + p_{t-} \cdot (1 - T_p) \cdot U_3 + (1 - T_p) \cdot U_4 - p_{t-} \cdot (1 - T_p) \cdot U_4$$

we see that the terms that do not cancel are:

$$p_{t+} \cdot T_p \cdot U_3 - p_{t+} \cdot T_p \cdot U_4 = p_{t+} \cdot T_p \cdot U_1 + T_p \cdot U_2 - p_{t+} \cdot T_p \cdot U_2 + -T_p \cdot U_4$$

Rearranging the terms and dividing with T_p, we have

$$U_4 - U_2 = p_{t+} \cdot (U_1 - U_3) + p_{t+} \cdot (U_4 - U_2)$$
$$p_{t+} = \frac{U_4 - U_2}{(U_1 - U_3) + (U_4 - U_2)}$$

So the threshold for baseline probability is then obtained using equation (B)

$P_{tt} = \frac{RR_t \cdot T_p + (1 - T_p)}{RR_t} \cdot \frac{U_4 - U_2}{(U_1 - U_3) + (U_4 - U_2)} = \frac{RR_t \cdot T_p + (1 - T_p)}{RR_t} \cdot P_t$	(2)

Test vs. Treat all:

$$EU[Test] = EU[Treat\ all]$$

If we use equation (A) in the expression for $EU[Treat\ all]$, we have:

$$EU[Treat\ all] = \left(p_{t+} \cdot T_p + p_{t-} \cdot \left(1 - T_p\right)\right) \cdot U_1 + U_2 - \left(p_{t+} \cdot T_p + p_{t-} \cdot \left(1 - T_p\right)\right) \cdot U_2$$
$$= p_{t+} \cdot T_p \cdot U_1 + p_{t-} \cdot \left(1 - T_p\right) \cdot U_1 + U_2 - p_{t+} \cdot T_p \cdot U_2 - p_{t-} \cdot \left(1 - T_p\right) \cdot U_2$$

Comparing that with $EU[Test]$, as above:

$$EU[Test] = p_{t+} \cdot T_p \cdot U_1 + T_p \cdot U_2 - p_{t+} \cdot T_p \cdot U_2 + p_{t-} \cdot \left(1 - T_p\right) \cdot U_3 + \left(1 - T_p\right) \cdot U_4 - p_{t-} \cdot \left(1 - T_p\right) \cdot U_4$$

we see that the terms with p_{t+} cancel and we can rearrange the rest as:

$$\left(1 - T_p\right) \cdot U_4 - \left(1 - T_p\right) \cdot U_2 = p_{t-} \cdot \left(1 - T_p\right) \cdot (U_1 - U_3) + p_{t-} \cdot \left(1 - T_p\right) \cdot (U_4 - U_2)$$

After dividing by $\left(1 - T_p\right)$, and solving for p_{t-}

$$p_{t-} = \frac{U_4 - U_2}{(U_1 - U_3) + (U_4 - U_2)}$$

So the threshold for baseline probability is then obtained using equation (B)

$P_{rx} = \left(RR_t \cdot T_p + \left(1 - T_p\right)\right) \cdot \frac{U_4 - U_2}{(U_1 - U_3) + (U_4 - U_2)} = \left(RR_t \cdot T_p + \left(1 - T_p\right)\right) \cdot P_t$	(3)

Note: From equations (2) and (3), we can see that $P_{rx} = R_t \cdot P_{tt}$.

Utilities: Event-free probabilities (counting combined VTE and bleeding events)

When the utilities are expressed through event-free probabilities, we have the following values of expected utilities and thresholds.

$$U_1^+ = 0 - RV \cdot H_{rx};\ U_1^- = 1 - RV \cdot H_{rx};$$
$$U_1 = RR_{rx} \cdot U_1^+ + (1 - RR_{rx}) \cdot U_1^-$$
$$= RR_{rx} \cdot (-RV \cdot H_{rx}) + (1 - RR_{rx}) \cdot (1 - RV \cdot H_{rx})$$
$$= 1 - RR_{rx} - RV \cdot H_{rx}$$
$$U_2 = U_1^- = 1 - RV \cdot H_{rx};\ U_3 = 0 - RV \cdot H_{norx} \text{ and } U_4 = 1 - RV \cdot H_{norx}$$
$$U_1 - U_3 = 1 - RR_{rx} - RV \cdot (H_{rx} - H_{norx});\ \text{and } U_4 - U_2 = RV \cdot (H_{rx} - H_{norx})$$

Then, the expected utilities and thresholds are

$$EU[Treat\ all] = 1 - p \cdot RR_{rx} - RV \cdot H_{rx}$$
$$EU[Test] = T_p \cdot (p_{t+} \cdot U_1 + (1 - p_{t+}) \cdot U_2) + \left(1 - T_p\right) \cdot (p_{t-} \cdot U_3 + (1 - p_{t-}) \cdot U_4)$$
$$= T_p \cdot (p_{t+} \cdot (U_1 - U_2) + U_2) + \left(1 - T_p\right) \cdot (p_{t-} \cdot (U_3 - U_4) + U_4)$$
$$= T_p \cdot (1 - p_{t+} \cdot RR_{rx} - RV \cdot H_{rx}) + \left(1 - T_p\right) \cdot (1 - p_{t-} - RV \cdot H_{norx})$$
$$= 1 - T_p \cdot p_{t+} \cdot RR_{rx} - \left(1 - T_p\right) \cdot p_{t-} - T_p \cdot RV \cdot (H_{rx} - H_{norx}) - RV \cdot H_{norx}$$
$$= 1 - \frac{T_p \cdot RR_t \cdot RR_{rx} + \left(1 - T_p\right)}{RR_t \cdot T_p + \left(1 - T_p\right)} \cdot p - T_p \cdot RV \cdot (H_{rx} - H_{norx}) - RV \cdot H_{norx}$$
$$EU[Treat\ none] = 1 - p - RV \cdot H_{norx}$$

In our specific example, $H_{rx} = RR_{bleed} \cdot H_{norx}$ so $H_{rx} - H_{norx} = (RR_{bleed} - 1) \cdot H_{norx}$.

$P_t = \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}}$	(4)
$P_{tt} = \frac{RR_t \cdot T_p + \left(1 - T_p\right)}{RR_t} \cdot \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}}$	(5)
$P_{rx} = \left(RR_t \cdot T_p + \left(1 - T_p\right)\right) \cdot \frac{RV \cdot (RR_{bleed} - 1) \cdot H_{norx}}{1 - RR_{rx}}$	(6)

Appendix 2

Utilities (outcomes): Number of VTE cases

$$U_1^+ = 1;\ U_1^- = 0;$$
$$U_1 = RR_{rx} \cdot U_1^+ + (1 - RR_{rx}) \cdot U_1^- = RR_{rx}$$
$$U_2 = U_1^- = 0;\ U_3 = 1 \text{ and } U_4 = 0$$

The expected number of cases for each strategy are then:

$$EU[Treat\ all] = p \cdot U_1 + (1-p) \cdot U_2 = p \cdot RR_{rx}$$
$$EU[Test] = T_p \cdot (p_{t+} \cdot U_1 + (1 - p_{t+}) \cdot U_2) + \left(1 - T_p\right) \cdot (p_{t-} \cdot U_3 + (1 - p_{t-}) \cdot U_4)$$
$$= T_p \cdot \left(\frac{p \cdot RR_t}{RR_t \cdot T_p + \left(1 - T_p\right)} \cdot RR_{rx}\right) + \left(1 - T_p\right) \cdot \left(\frac{p}{RR_t \cdot T_p + \left(1 - T_p\right)}\right)$$
$$= p \cdot \left(\frac{T_p \cdot RR_t \cdot RR_{rx} + \left(1 - T_p\right)}{RR_t \cdot T_p + \left(1 - T_p\right)}\right)$$
$$EU[Treat\ none] = p \cdot U_3 + (1-p) \cdot U_4 = p$$

Note that whenever $RR_{rx} < 1$, we have $\frac{T_p \cdot RR_t \cdot RR_{rx} + (1 - T_p)}{RR_t \cdot T_p + (1 - T_p)} < 1$ so $EU[Test]$ is *always* smaller than $EU[Treat\ none]$. Similarly, $\frac{T_p \cdot RR_t \cdot RR_{rx} + (1 - T_p)}{RR_t \cdot T_p + (1 - T_p)} > RR_{rx}$ and, therefore, $EU\ [Treat\ all]$ is *always* smaller than $EU[Test]$. Therefore, we will always have a strategy "Treat none" with the most VTE recurrence events, a strategy "Treat all" with the least VTE recurrences, and a strategy "Test" between them. And therefore, all the **thresholds are zero.**

Utilities (outcomes): Number of Bleeding cases

$$U_1^+ = H_{rx};\ U_1^- = H_{rx} = RR_{bleed} \cdot H_{norx};$$
$$U_1 = RR_{rx} \cdot U_1^+ + (1 - RR_{rx}) \cdot U_1^- = H_{rx} = RR_{bleed} \cdot H_{norx}$$
$$U_2 = H_{rx} = RR_{bleed} \cdot H_{norx};\ U_3 = H_{norx} \text{ and } U_4 = H_{norx}$$

For these utilities, the number of bleeding cases for "Treat all" strategy will *always be larger* than a number of cases for the "Testing" strategy, which, in turn, is *always larger* than the number of cases for the "Treat none" strategy. Therefore, **all the thresholds are undefined**.

However, we can still calculate the expected number of cases as expected utility for each strategy.

$$EU[Treat\ all] = p \cdot U_1 + (1-p) \cdot U_2 = H_{rx} = RR_{bleed} \cdot H_{norx}$$
$$EU[Test] = T_p \cdot (H_{rx}) + \left(1 - T_p\right) \cdot (H_{norx}) = \left(T_p \cdot RR_{bleed} + \left(1 - T_p\right)\right) \cdot H_{norx}$$
$$EU[Treat\ none] = p \cdot U_3 + (1-p) \cdot U_4 = H_{norx}$$

As we assume that $RR_{bleed} > 1$, we know that the "Treat all" strategy always has more bleeding events than the "Treat none". We also have $RR_{bleed} = T_p \cdot RR_{bleed} + \left(1 - T_p\right) \cdot RR_{bleed} > T_p \cdot RR_{bleed} + \left(1 - T_p\right)$.We know that the "Treat all" strategy always has more bleeding events than the "Test" strategy. Finally, since$T_p \cdot RR_{bleed} + \left(1 - T_p\right) > T_p + \left(1 - T_p\right) = 1$, we know that the "Test" strategy *always* has more bleeding events than the "Treat none".

Therefore, we will *always* have a strategy "Treat all" with the most bleeding events, a strategy "Treat none" with the least bleeding events, and a strategy "Test" between them. As the prevalence of VTE recurrence doesn't appear in these calculations, the **thresholds are undefined.**

Appendix 3

Table 1

Reproducing the ASH thrombophilia results using the threshold decision model: an illustration

The risk (probability) of venous-thromboembolism (VTE recurrences) in patients with unprovoked VTE is given as $p = \frac{100}{1000} = 0.1$; the treatment effect on VTE recurrence is given by $RR_{rx} = 0.15$, the baseline bleeding without any treatment is $H_{norx} = \frac{5}{1000} = 0.005$ for low-risk cases, and the effect of treatment on the bleeding risk is $RR_{bleed} = 2.17$. For recommendation #1, we want to compare strategy "Treat all "[i.e., "Continue indefinite anticoagulation) without thrombophilia testing in all patients with unprovoked VTE"] vs. "Test" ["Perform thrombophilia testing and treat only those patients with positive test results"]

Using our model, we can calculate the number of *VTE recurrences* in each strategy as:

$$EU[Treat\ all] = p \cdot RR_{rx} = (0.10) \cdot (0.15) = 0.015 = \mathbf{\frac{15}{1000}}$$

$$EU[Test] = p \cdot \left(\frac{T_p \cdot RR_t \cdot RR_{rx} + \left(1 - T_p\right)}{RR_t \cdot T_p + \left(1 - T_p\right)}\right) = 0.10 \cdot \left(\frac{0.38 \cdot 1.65 \cdot 0.15 + (1 - 0.38)}{1.65 \cdot 0.38 + (1 - 0.38)}\right) = 0.05726 = \mathbf{\frac{57}{1000}}$$

Thus, and identical to the ASH panel's report, using our formulas, we can calculate that the "Testing" strategy would have resulted in 42 (57-15) more VTE recurrences per 1000 patients than "Treat all" strategy**.**

Let us now calculate *bleeding rate* for each strategy (we will only illustrate our methods in patients at low risk of bleeding, but a reader can use our formulas to calculate event rates for any value of risk):

$$EU[Treat\ all] = RR_{bleed} \cdot H_{norx} = 2.17 \cdot 0.005 \approx 0.011 = \mathbf{\frac{11}{1000}}$$

$$EU[Test] = \left(T_p \cdot RR_{bleed} + \left(1 - T_p\right)\right) \cdot H_{norx} = \left(0.38 \cdot 2.17 + (1 - 0.38)\right) \cdot 0.005 \approx 0.007 = \mathbf{\frac{7}{1000}}$$

Thus, and *identical to the ASH panel's report,* in the patients at a low risk of bleeding, the "Treat all" strategy is associated with 4 (11-7) more bleeding episodes compared with the "Testing" strategy.