BMJ Open

Determinants of treatment duration in the prevention of recurrent venous thromboembolism: a protocol for a balanced vignette experiment.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015231
Article Type:	Protocol
Date Submitted by the Author:	18-Nov-2016
Complete List of Authors:	Ten Cate, Vincent; Maastricht University, Epidemiology; Maastricht University CAPHRI School for Public Health and Primary Care Essers, Brigitte; Maastricht University Medical Center, Clinical Epidemiology and Medical Technology Assessment Prins, Martin; Maastricht University Medical Center, Epidemiology
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Epidemiology, Evidence based practice, Cardiovascular medicine
Keywords:	EPIDEMIOLOGY, STATISTICS & RESEARCH METHODS, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Vascular medicine < INTERNAL MEDICINE

Title

Determinants of treatment duration in the prevention of recurrent venous thromboembolism: a protocol for a balanced vignette experiment.

Authors and affiliations

Ten Cate V[1,2], Essers BAB[3], Prins MH[1].

1. Maastricht University Medical Center, Department of Epidemiology Maastricht, the Netherlands.

2. Maastricht University CAPHRI School for Public Health and Primary Care Maastricht, the Netherlands.

3. Maastricht University Medical Center, Department of Clinical Epidemiology and Medical Technology Assessment

Corresponding author

Vincent Ten Cate, MSc
Maastricht University Medical Center, Department of Epidemiology
Maastricht, the Netherlands.
Email: vincent.ten.cate@mumc.nl
Telephone: +31631766450

Keywords

Venous Thromboembolism, Secondary Prophylaxis, Duration of Treatment, Surveys and Questionnaires

Abstract (298 words)

Introduction

Venous thromboembolism (VTE) is a condition that annually occurs in approximately 1 ‰ of the world's population. The incidence of VTE increases with age to approximately 5-6‰ in those over 80 years old. Approximately 6% of DVT cases and 12% of PE cases are lethal within one month of diagnosis. Patients who have already had a VTE are at elevated risk for a recurrent VTE. Recurrent events increase the risk of long-term sequelae and could potentially be fatal. Adequate secondary prophylaxis is thus needed to prevent such events.

VTE patients are often prone to bleeding, and pharmacologic prophylaxis exacerbates bleeding risk. Physicians carefully consider which risk is greater in their patients, based on a number of patient characteristics. Expert opinions on the optimum duration of secondary prophylaxis in VTE still vary substantially. The existence of treatment guidelines has not led to uniformity of VTE secondary prophylaxis strategies, which means that physicians still adhere to individual risk calculi in determining treatment duration.

Methods and Analysis

The aim of this study is to establish what factors lie at the root of this variance in VTE secondary prophylactic treatment strategies, and what risk factors are deemed of particular importance in determining the perceived risks and benefits of variable treatment durations. To do this, we created a survey based on a D and G-efficient balanced experimental vignette design. Results will be analyzed using generalized linear mixed models (GLMM).

Ethics and Dissemination

All data are de-identified, and any identifying characteristics of the respondents will not be reported in a final manuscript or elsewhere.

A paper containing a qualitative summary of the expert interviews is currently being drafted. The results of the experiment described in this protocol will be described in a separate article, to be drafted once the analyses have been conducted.

Strengths and limitations

Strengths

- We conducted thirteen interviews with experts on the subject matter to compose our initial design.
- Our design features a broad range of patient characteristics, all of which are equally represented across our vignettes by virtue of our methodology.
- Our factorial design is D and G-efficient, and the levels are balanced. This means that there is minimal multicollinearity-induced noise in our experiment, allowing for relatively unbiased parameter estimation.
- We conducted a pilot of an initial design of our experiment, enlisting expert participants. We improved our design based on their comments.
- Our use of mixed effects methods will provide highly accurate estimates of regression coefficients, while simultaneously clarifying the mechanisms underlying inter-physician variance.

Limitations

- We could not incorporate all risk factors that were mentioned in the expert interviews in our design, as this design would be too complex.
- Physicians are to evaluate recurrence and bleeding risks based on a list of risk factors. Physician responses may differ from real life, as e.g. face-to-face contact with any of our hypothetical patients is impossible.
- As we simulated a balanced set of patient vignettes, the patient characteristics featured in our experiment do not necessarily occur in real-life distributions.

Word count

3001 words.

Introduction

Venous thromboembolism (VTE) is a condition that annually occurs in approximately 1 ‰ of the world's population [1]. VTE encompasses deep vein thrombosis (DVT) and pulmonary embolism (PE). The elderly are at particular risk, as the incidence of VTE increases with age to approximately 5-6‰ in those over 80 years old [1]. Two-thirds of VTE diagnoses constitute cases of deep vein thrombosis, and the remainder pulmonary embolism. 25-50% of VTE cases are idiopathic, which means that they are of unknown etiology [2].

Approximately 6% of DVT cases and 12% of PE cases are lethal within one month of diagnosis [2]. Recovery rates from a VTE event are high given appropriate pharmacologic anticoagulant therapy [2]. If the embolus in patients with PE is not completely resolved after treatment, long-term complications such as chronic thromboembolic pulmonary hypertension (CTPH) may occur [4]. Similarly, patients with DVT are at elevated risk of post-thrombotic syndrome (PTS) during long term follow up. Moreover, patients who have already had a VTE are at elevated risk for a recurrent VTE [5]. Recurrent events increase the risk of long-term sequelae and are potentially fatal [5]. Hence, adequate secondary prophylactic treatment, i.e. the outpatient preventive treatment that follows the resolution of an acute event, is necessary. However, what exactly constitutes adequate secondary prophylaxis of VTE is still controversial: physicians have yet to achieve consensus on both the ideal type and duration of treatment [6-11].

The difficulty in achieving consensus regarding the duration of secondary thromboprophylaxis in patients with VTE stems from the balance that has to be struck between the thrombotic risk associated with non-treatment, and the reduction in risk of recurrence and the induced bleeding risk associated with anticoagulant therapy. As stated, VTE is a disease that disproportionately affects the elderly, and the elderly are a group particularly prone to bleedings on account of their frailty. Physicians are thus faced with the quandary of prescribing a treatment that itself poses a risk to the health of their patients, while not doing so might incur an even greater risk due to recurrent disease. Their task is then to find the optimum duration of treatment that most effectively mitigates the risk of VTE recurrence in a patient, while minimizing their exposure to an inherently risky treatment.

Organizations such as the American College of Chest Physicians (CHEST, formerly: ACCP) and the European Society of Cardiology (ESC) have attempted to coordinate and formalize the treatment and prevention of VTE with periodical installments of treatment guidelines [12,13]. However, the recommendations outlined in these guidelines are wont to change with every new installment, and it is unclear how many physicians adhere to these guidelines. Moreover, these guidelines represent a combination of evidence-based research and expert opinions, which themselves vary from physician to physician. The ongoing debate on the optimum type and duration of treatment is testament to the existence of many different calculi individual physicians employ in their long-term preventive VTE treatment strategies.

The aim of this study is to establish what factors lie at the root of this variance in VTE secondary prophylactic treatment strategies, and what risk factors are deemed of particular importance in determining the perceived risks and benefits of ceasing or continuing preventive treatment for a particular duration of time. We do this by having physicians evaluate patient vignettes, i.e. hypothetical patient profiles, the attributes of which were derived from expert interviews.

Methods and analysis

Vignette experiment

Vignettes are hypothetical scenarios that mimic real-life situations, enabling researchers to evaluate their phenomenon of interest in a controlled experimental setting [14]. Since there is little noise in a designed experiment, covariates can be assessed in a relatively unbiased manner [14]. Additionally, the use of vignettes precludes the need to recruit a balanced sample of subjects, as such a sample can simply be simulated. In our case, we aim to establish what drives the decision-making process surrounding secondary prophylaxis of VTE. In a real-life setting, it would be difficult to recruit a diverse enough set of patients to take the various risk factors that may influence this process into account. With the vignette experiment described in this paper, we achieved a highly efficient alternative to a much more costly and time-consuming real-life experiment.

Preselection of attributes

From December 2015 through February 2016, we conducted 13 semi-structured interviews with an equal number of senior-level physicians affiliated with Einstein Choice-enrolled medical centers around the world [15]. The purpose of these interviews was to establish which factors are particularly important in determining the duration of secondary prophylaxis of VTE, and, conversely, which factors weigh against a decision to continue treatment past the resolution of the initial event. Our questions were informed by a review of the pertinent literature and guidelines [12,13], which had yielded an initial selection of relevant attributes. Importantly, we decided to exclude cancer as a risk factor, as its impact on the decision to continue or stop anticoagulation is too dominant: presence of an active cancer invariably warrants indefinite anticoagulation, regardless of other factors [12].

In our interviews, we specifically asked our interviewees to identify patient characteristics that significantly impact either the risk of VTE recurrence or the risk of bleeding. Although all physicians weighted individual risk factors differently, common factors and their corresponding levels were identified. Table 1 is a frequency table of all risk factors that were considered to impact either on VTE recurrence risk or on bleeding risk, by number of physicians who mentioned them. This table does not include risk factors that were mentioned by fewer than two interviewees.

Reducing design complexity

The preselection of attributes yielded 18 factors that were indicated to be of importance by at least two physicians. A recommendation for conjoint experiments, which exhibit commonalities with our approach, states to limit the complexity of a design as much as possible, as an overly complex design may confuse respondents and distort the responses to an experiment [16]. We first discarded factors that were unsuitable for the vignette experiment due to their ungeneralizability: uncontrolled hypertension and physical activity. We then removed the 'unstable or high INR' factor on account of it being specific to vitamin K antagonists. We further trimmed down the list of factors by discarding the risk factors that only rarely occur in real life VTE patients: thrombocytopenia (<1%) [17] and liver disease (~1.5%) [18]. Finally, we combined 'signs of PTS' with the 'type of VTE' factor, so as to prevent inappropriate interactions: post-thrombotic syndrome almost exclusively occurs in tandem with DVT, rather than with PE [19]. This leaves 12 factors in our reduced model.

Experimental design

Given 12 factors, of which nine have two levels; two have three levels; and one has six levels, there is a total of $2^9 3^2 6 = 27648$ level permutations in the full factorial model. As having respondents evaluate all 27648 possible scenarios would be prohibitively costly, we created an efficient fractional factorial design via Fedorov's 1972 exchange algorithm [20]. To generate this design, we utilized the R package 'AlgDesign' [21], an implementation of the Fedorov algorithm. The Fedorov algorithm randomly selects rows from full factorial matrix X to create a new matrix ξ, and then exchanges rows from ξ with rows from $X \setminus \xi$ to optimize a specified criterion. In our case, we optimized our reduced design for *D* and *G*-efficiency.

The number of vignettes in the reduced design was selected on the basis of the following criteria:

I. D-efficiency
II. G-efficiency
III. Level balance
IV. Feasibility of the required sample size

I. D-efficiency

The *D*-efficiency, perhaps the most frequently used efficiency criterion in the design of experiments (DoE) [22,23], is a measure of information in a reduced design ξ relative to the full factorial matrix *X*. In the AlgDesign package, the D criterion for a design ξ is computed as the k^{th} root of its generalized variance [21], i.e. the determinant of the normalized covariance matrix $\Sigma(\xi) = (\xi^T \xi / N_\xi)$:

$$D_\xi = |\Sigma(\xi)|^{\frac{1}{k}}$$

Where k is the number of columns and N_ξ the number of rows in ξ. The relative D-efficiency *De* of ξ with respect to X is computed by taking the ratio of D_ξ to the D criterion of the full factorial design:

$$De_\xi = \frac{D_\xi}{D_X}$$

II. G-efficiency

G-efficiency is another optimality criterion, which aims to minimize the prediction variance of a fractional factorial design [23]. In the AlgDesign package, G-efficiency is computed as follows [21]:

$$Ge_\xi = \frac{k}{\max\limits_{x \in X} (x^T \Sigma(\xi)^{-1} x)}$$

Here, x denotes a row of X. In natural language, this means that the number of columns k is divided by the row in X with the highest precision (the inverse of the covariance matrix is the precision matrix [24]) of ξ over X.

III. Level balance

Level balance is an important attribute of experimental designs, which ensures design orthogonality [25]. In order to achieve level balance, the number of rows N_ξ should be an integer multiple of the number of levels for all factors. This ensures that all levels can be represented an equal number of times per factor in the design matrix, creating a balanced set of vignettes [26].

IV. Feasibility of the required sample size

There does not appear to be a uniform standard when it comes to determining an appropriate sample size for experimental designs such as the one described in this protocol. One formula, used in conjoint analyses, dictates a sample size of $n = 500 \times c/(a \times t)$, where c signifies the maximum number of levels for any of the factors, a denotes the number of choice alternatives presented at once, and t denotes the number of unique vignettes presented to each respondent [27-29]. Another recommendation states that all vignettes in a design should be evaluated at least six times [30]. We are currently recruiting respondents in a variety of ways, e.g. emailing physicians connected to the Einstein Choice study and approaching relevant professional networks, in order to meet these criteria.

Number of vignettes in the reduced design

Based on the above considerations, we selected a design ξ with $N_\xi = 72$. This number of vignettes yielded a well-rounded balance between our efficiency criteria, as illustrated in figure 1 and table 2. To reduce workload for our respondents, we divided this design into six blocks of 12 vignettes each, using the optBlock function in the 'AlgDesign' R package. After optimization, we

achieved a block design diagonality of 0.88. Plugging in the values for a single block into our sample size formula returns a required sample size of $n = 500 \times \frac{6}{1 \times 12} = 250$. As there is considerable overlap between blocks, we assume a sample size of 250 to be sufficient to draw reliable inferences with the design as specified above. Following the 'six-times rule', accounting for twelve vignette evaluations per respondent, yields a required sample $n = \frac{72 \times 6}{12} = 36$.

Pilot evaluation of the design

In May 2016, we sent out Microsoft Word-based pilot surveys to twelve thrombosis experts to evaluate our initial design. From the comments we received, it was apparent that some of the factors and levels in our model were not formulated precisely enough. For instance, pilot participants indicated that in patients who had had a previous VTE, it was important to know when this had occurred. For this item, we added a timeframe to improve clarity. Some participants struggled with the unqualified formulation 'renal function: insufficient'. We included the more detailed description '<50 ml/min' in these cases, referring to the corresponding creatinine clearance rate. Additionally, of patients who received concurrent antiplatelet therapy, a number of pilot participants said to be confused as to whether the practitioner treating the patient for VTE recurrence could cease this concurrent therapy. As our intention was to refer to patients with an absolute indication for antiplatelet therapy, we reformulated this item to more explicitly refer to such patients. With these changes, we arrived at our final design, shown in table 3.

Physician characteristics

From our expert interviews, it was apparent that physicians involved in the secondary prevention of VTE all have unique perspectives on thrombosis and bleeding risk in VTE. In order to understand why physicians who are faced with the same clinical vignettes may respond to these differently, we decided to include a number of questions relating to characteristics of the physicians themselves. Before commencing the experiment, physicians are asked to provide their age, gender, country of employment, specialty, years of experience with the treatment of VTE, and an estimate of the number of unique VTE patients they treat per annum. These physician characteristics were selected on the basis of previous research into physician guideline adherence and clinical decision-making [32-36].

By incorporating the 'country of employment' variable as a random effect, we hope to capture the variance resulting from any country-specific aspects of physician's decision-making that would otherwise be difficult to quantify.

We provided six possible answers to choose from in response to the specialty question, i.e.: internal medicine, vascular medicine/angiology, cardiology, hematology, pulmonology or other. The specialty of responding physicians will also be included as a random effect in our analysis.

Outcome variables

As outcome variables, we incorporated two continuous variables to respectively estimate VTE recurrence risk and bleeding risk on a 100-point scale indicating low to high risk, and five discrete variables: continue treatment for three additional months, continue 9 additional months, continue indefinitely, stop treatment, or switch to aspirin. These choices were based on the most recent CHEST and ESC VTE treatment guidelines [12,13], the conducted expert interviews and the comments received on our pilot design.

Implementation of the survey

The survey has been implemented online by independent web developers, and is currently live at www.vte-survey.com (figure 2). Survey respondents are randomized to one of six pre-specified blocks at the start of the experiment.

As mentioned above, one aspect of our experimental design was to make the workload for study participants as slight as possible. For this reason, in the online implementation of the survey the progress of participants is continuously recorded. This allows participants to take breaks during the survey and even to leave the website, without their responses up until that point being forfeited. Participants may return to the website at any time and resume the survey where they left off. In our current sample of respondents, there is a small number of such cases that skew the mean survey completion time considerably upward. A better reflection of the average workload is given by the median survey completion time, which in our current sample is equal to 9 minutes. Based on this number, we conclude that the risk of participant fatigue is negligible for our study.

Data collection was commenced in July 2016, and will be ceased once both sample size criteria have been met. We hope to achieve this goal by or in Q1 2017. We welcome any physicians involved in the preventive treatment of VTE who are reading this protocol to participate in this study.

Analysis of the results

We will conduct several separate analyses of our results (figure 3). A first linear mixed effects regression model will relate all risk factors and physician characteristics to the continuous estimated VTE recurrence risk outcome variable. A second model does the same for bleeding risk. After having estimated the coefficients in these two linear models, we will use generalized linear mixed models (GLMMs) to estimate the associations of all independent variables, including physician characteristics, with the choice of treatment options. We will use a GLMM with a binomial logit link function for grouped outcomes (stop versus continue treatment), and a multinomial logit regression to discover which variables best predict the five separate treatment outcomes.

The choice for mixed effects analyses was informed by the theoretical assumption that there is considerable inter-physician variability, and to control

for random variance associated with aspects of no interest (e.g. block allocation). While analysis of the fixed effects (i.e. the regression coefficients) will provide general insights as to which features are of particular interest in determining a patient's risk of VTE recurrence or bleeding, it is the decomposition of the error term that will show us specifically where and to what extent physicians disagree with one another. To this end, we will specify random slopes that vary by physician for all relevant features. All analyses will be conducted in the statistical software package R.

Implications of this research

The research described in this protocol will provide insight into the mechanisms underlying the decision-making of physicians involved in the outpatient treatment of VTE. Our aims are to identify the relative weights of risk factors; to determine how physicians decide on the duration of preventive treatment, and to uncover what prompts physicians to discontinue further treatment.

Simultaneously, this research places a strong focus on the variance that exists between physicians in determining VTE treatment duration. We will explore which factors contribute to this variance by using a mixed-effects approach, enabling direct analysis of the inter-individual differences in decision-making patterns. In so doing, it is our hope that this research will lead to a better understanding of why physicians may deviate from best-practice guidelines, and what biases may potentially be at play in making this decision.

We believe that the insights acquired through this study will benefit physicians involved in the treatment of VTE, by elucidating the overlaps and discrepancies between treatment strategies and by highlighting the areas potentially prone to oversights. For this same reason, our findings should also be of benefit to the national and international collectives whose aim it is to unify and coordinate VTE treatment. By explicitly showing where physician decision-making coincides and where it varies, we believe that VTE prevention education and policy-making can at once be made more efficient and more effective.

Ethics and dissemination

All data are de-identified, and any identifying characteristics of the respondents will not be reported in a final manuscript or elsewhere.

A paper containing a qualitative summary of the expert interviews is currently being drafted. The results of the experiment described in this protocol will be described in a separate article, to be drafted once the analyses have been conducted. We aim to satisfy our sample size criteria in or before Q1 2017. Both articles will be submitted to peer-reviewed journals.

References

[1] Cushman M. Epidemiology and Risk Factors for Venous Thromboembolism. Semin Hematol. 2007 April 44(2): 62-69.

[2] White RH. The epidemiology of venous thromboembolism. Circ. 2003; 107(23 suppl 1):I4-I8.

[3] Sweet PH et al. Fatal pulmonary embolism update: 10 years of autopsy experience at an academic medical center. J R Soc Med Sh Rep. 2013; 4: 1-5.

[4] Baser O. Prevalence and Economic Burden of Venous Thromboembolism After Total Hip Arthroplasty or Total Knee Arthroplasty. Am J Manag Care. 2011;17:S6-S8.

[5] Zhu T, Martinez I, Emmerich J. Venous Thromboembolism: Mechanisms, Treatment, and Public Awareness. Arterioscl Throm Vas. 2009; 29: 298-310.

[6] Beyer-Westendorf J, Cohen AT, Monreal M. Venous thromboembolism prevention and treatment: expanding the rivaroxaban knowledge base with real-life data. Eur Heart J Supp. 2015;7(17)suppl D:D32-D41.

[7] Sternitzky R, Hochauf S, Schellong SM. Secondary prophylaxis of venous thromboembolism. Hamostaseologie. 2007 Feb;27(1):32-40.

[8] Hibbert PD, Hannaford NA, Hooper TD et al. Assessing the appropriateness of prevention and management of venous thromboembolism in Australia: a cross-sectional study. BMJ Open 2016;6:e008618.

[9] Caprini JA, Tapson VF, Hyers TM et al. Treatment of venous thromboembolism: adherence to guidelines and impact of physician knowledge, attitudes, and beliefs. J Vasc Surg. 2005 Oct;42(4):726-33.

[10] Brenner B. Venous thromboembolism: advances and controversies. Best Pract Res Clin Haematol. 2012 Sep;25(3):233. doi: 10.1016/j.beha.2012.07.009.

[11] Schulman S, Douketis J. Secondary prevention of venous thromboembolism. BMJ. 2013 Sep 9;347:f5440. doi: 10.1136/bmj.f5440.

[12] Kearon C, Akl EA, Ornelas J et al. Antithrombotic Therapy for VTE Disease: CHEST Guideline and Expert Panel Report. Chest. 2016 Feb;149(2):315-52.

[13] Konstantinides S et al. 2014 ESC Guidelines on the diagnosis and management of acute pulmonary embolism. Eur Heart J. 2014; 35: 3033–3080.

[14] Evans SC, Roberts MC, Keeley JW et al. Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies. Int J Clin Hlth Psyc. 2015; 15(2):160-170.

[15] Bayer. Reduced-dosed Rivaroxaban in the Long-term Prevention of Recurrent Symptomatic VTE (Venous Thromboembolism) (Einstein Choice). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000- [cited 2016 Oct 27]. Available from:

https://clinicaltrials.gov/ct2/show/NCT02064439 NLM Identifier: NCT02064439.

[16] DeShazo JR, Fermo G. Designing choice sets for stated preference methods: the effects of complexity on choice consistency. J Environ Econ Manag 2002;44:123–43.

[17] Stein PD, Hull RD, Matta F et al. Incidence of thrombocytopenia in hospitalized patients with venous thromboembolism. Am J Med. 2009 Oct;122(10):919-30.

[18] Saleh T, Matta F, Alali F et al. Venous thromboembolism with chronic liver disease. Am J Med. 2011 Jan;124(1):64-8.

[19] Bova C, Rossi V, Ricchio R et al. Incidence of post-thrombotic syndrome in patients with previous pulmonary embolism. A retrospective cohort study. Thromb Haemost. 2004 Nov;92(5):993-6.

[20] Fedorov VV. Theory of optimal experiments. Elsevier; 1972.

[21] Wheeler RE. Comments on algorithmic design 2004–2009. http://cran.r-project.org/web/packages/AlgDesign/vignettes/AlgDesign.pdf. (accessed April 2016)

[22] Kuhfeld WF. Efficient experimental designs using computerized searches. Research Paper Series, SAS Institute, Inc. 1997 Nov.

[23] Rodríguez M, Jones B, Borror CM, Montgomery DC. Generating and assessing exact G-optimal designs. Journal of quality technology. 2010 Jan 1;42(1):3.

[24] Wasserman L. All of statistics: a concise course in statistical inference. Springer Science & Business Media; 2013 Dec 11.

[25] Collins LM, Dziak JJ, Li R. Design of Experiments with Multiple Independent Variables: A Resource Management Perspective on Complete and Reduced Factorial Designs. Psychol Methods. 2009 Sep;14(3):202-224.

[26] Kuhfeld WF, Tobias RD, Garratt M. Efficient Experimental Design with Marketing Research Applications. J Mark Res. 1994;31(4):545-557.

[27] Rose JM, Bliemer MCJ. Sample size requirements for stated choice experiments. Transportation 2013;40:1021–41.

[28] Vilikus O. Optimization of Sample Size and Number of Tasks per Respondent in Conjoint Studies Using Simulated Datasets. Studia Ekonomiczne/Uniwersytet Ekonomiczny w Katowicach. 2012(120 Survey sampling in economic and social research):77-83.

[29] Engel J, van der Wulp I, Poldervaart JM et al. Clinical decision-making of cardiologists regarding admission and treatment of patients with suspected unstable angina or non-ST-elevation myocardial infarction: protocol of a clinical vignette study. BMJ Open. 2015 Apr 8;5(4):e006441.

[30] Louviere JJ, Hensher DA, Swait JD. Stated choice methods. Analysis and application. New York: Cambridge University Press, 2000:83–111.

[31] Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. Ann Intern Med. 2005 Feb 15;142(4):260-73.

[32] Solomon DH, Brookhart MA, Gandhi TK et al. Adherence with osteoporosis practice guidelines: a multilevel analysis of patient, physician, and practice setting characteristics. Am J Med. 2004 Dec 15;117(12):919-24.

[33] Kumar A, Fonarow GC, Eagle KA et al. Regional and practice variation in adherence to guideline recommendations for secondary and primary prevention among outpatients with atherothrombosis or risk factors in the United States: a report from the REACH Registry. Crit Pathw Cardiol. 2009 Sep;8(3):104-11. doi: 10.1097/HPC.0b013e3181b8395d.

[34] McKinlay JB, Lin T, Freund K, Moskowitz M. The unexpected influence of physician attributes on clinical decisions: results of an experiment. J Health Soc Behav. 2002 Mar;43(1):92-106.

[35] Berthold HK, Gouni-Berthold I, Bestehorn KP, Böhm M, Krone W. Physician gender is associated with the quality of type 2 diabetes care. J Intern Med. 2008 Oct;264(4):340-50. doi: 10.1111/j.1365-2796.2008.01967.x.

Acknowledgments

We extend our gratitude to the thirteen interviewees whose input helped us select risk factors and outcome variables for our vignette experiment. Our thanks also go out to Maastricht University Medical Center's thrombosis department, Dr. Avi Leader of Israel and Dr. Michael Nagler of Switzerland for their valuable comments and evaluation of our pilot survey. We further thank Petro van Bergen for his help in recruiting survey participants, and the International Society for Thrombosis and Haemostasis (ISTH) for featuring this experiment in their newsletter (October 2016). We are grateful to Dr. David McAneny of Boston University and Dr. Charles Mahan of the University of New Mexico, whose comments have inspired substantial revision of this manuscript. Lastly, we are indebted to all survey participants for contributing to our research.

Author contributions

VTC and MHP designed the protocol. VTC wrote the protocol. BABE provided helpful input throughout.

Funding statement

This work was supported by Bayer AG, grant number BAY59-7939/16416add.

Competing interests

VTC and BABE declare no competing interests. MHP has received research support and honoraria, and participated in advisory boards for Bayer HealthCare, Sanofi-Aventis, Boehringer Ingelheim, GlaxoSmithKline, Daiichi Sankyo, LeoPharma, ThromboGenics, and Pfizer.

Data sharing statement

The R code that was used to generate the described design and figure 1 is available upon request. Please contact the first author, VTC, if you would like to receive the script containing said code.

Appendix

Table 1. Frequency table of risk factors mentioned by interviewed physicians

Factors (k)	Levels	Type of risk factor	Freq. (%)
Idiopathic	no – yes	TR	13 (100)
VTE type	distal DVT – proximal DVT – non-massive PE – massive PE	TR	13 (100)
Thrombophilia	none – acquired – inherited	TR	11 (85)
History of bleeding	no – yes	BR	10 (77)
On concurrent antithrombotic medication (e.g. aspirin)	no – yes	BR	9 (69)
Age	<65 – >65	BR/TR	8 (62)
Renal function	normal – insufficient	BR	7 (54)
Liver disease	no – yes	BR	7 (54)
Unstable or high INR	no – yes	BR	7 (54)
History of thrombosis	no – yes	TR	6 (46)
BMI	underweight – normal weight – overweight	BR/TR	5 (38)
Uncontrolled hypertension	no – yes	BR	5 (38)
Alcohol or drug abuse	no – yes	BR	5 (38)
Family history of VTE	no – yes	TR	4 (31)
Physical activity	sedentary – active	TR/BR	4 (31)
Thrombocytopenia	no - yes	BR	3 (23)
Signs of PTS	no – yes	TR	3 (23)
Sex	female – male	TR	2 (15)

BR denotes bleeding risk; INR denotes international normalized ratio; PTS denotes post-thrombotic syndrome; TR denotes thrombotic risk (i.e. VTE recurrence risk).

Figure 1. D and G efficiencies of design matrix ξ.

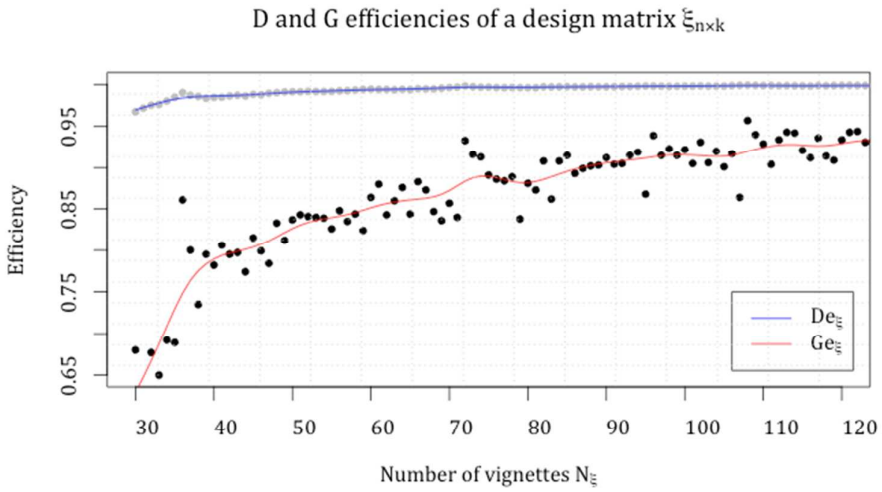


Table 2. Efficiency criteria of design matrix ξ.

N_ξ	De_ξ	Ge_ξ
72	0.998	0.94

Table 3. Final experimental design.

Factor	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Sex	female	male				
Age	<65	>65				
BMI	underweight	normal weight	overweight			
Idiopathic	no	yes				
Type of VTE	distal DVT	distal DVT with signs of PTS	proximal DVT	proximal DVT with signs of PTS	non-massive PE	massive PE
Previous VTE	no	yes (2 years ago)				
Family history of VTE	no	yes				
History of major bleeding	no	yes				
Thrombophilia	none	acquired	inherited			
Renal function	normal	insufficient (<50 ml/min)				
Alcohol or substance abuse	no	yes				
Absolute indication for aspirin	no	yes				

Figure 2. Screen capture of the web-based survey (www.vte-survey.com).

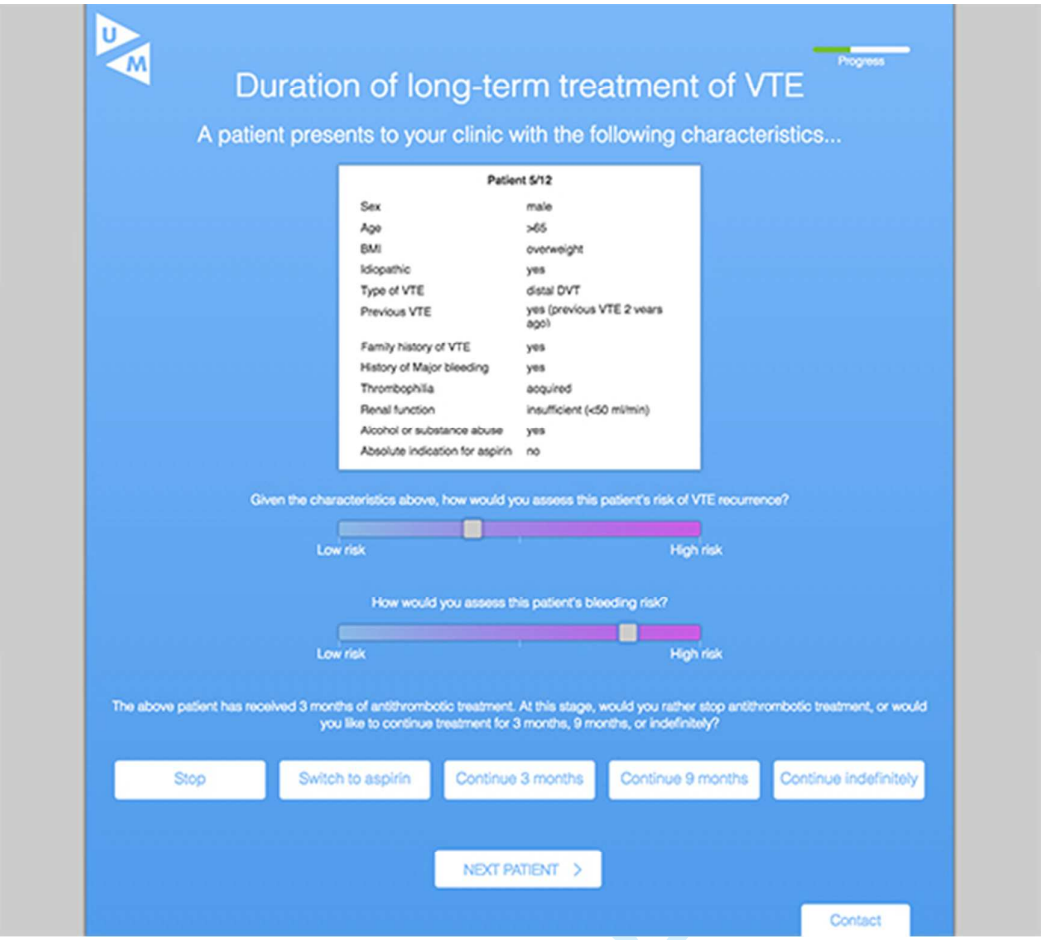
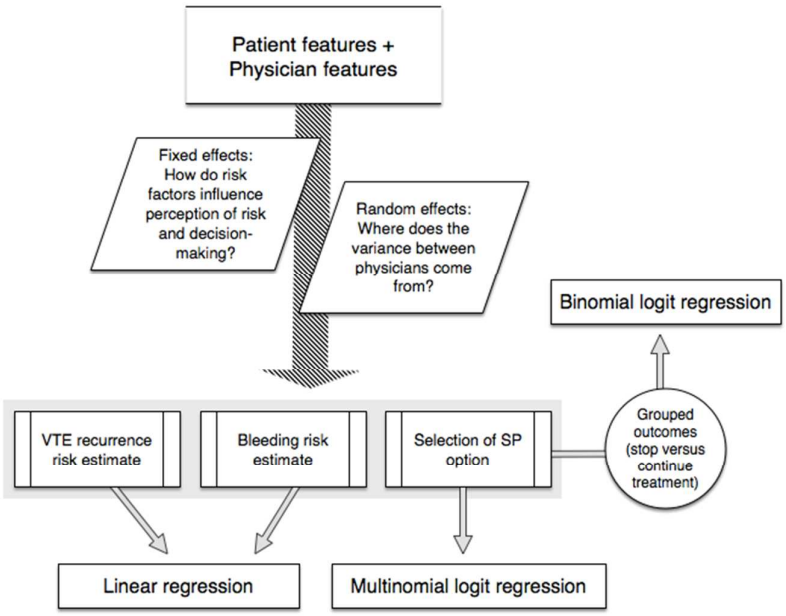


Figure 3. Analysis flow chart.



SP denotes secondary prophylaxis.

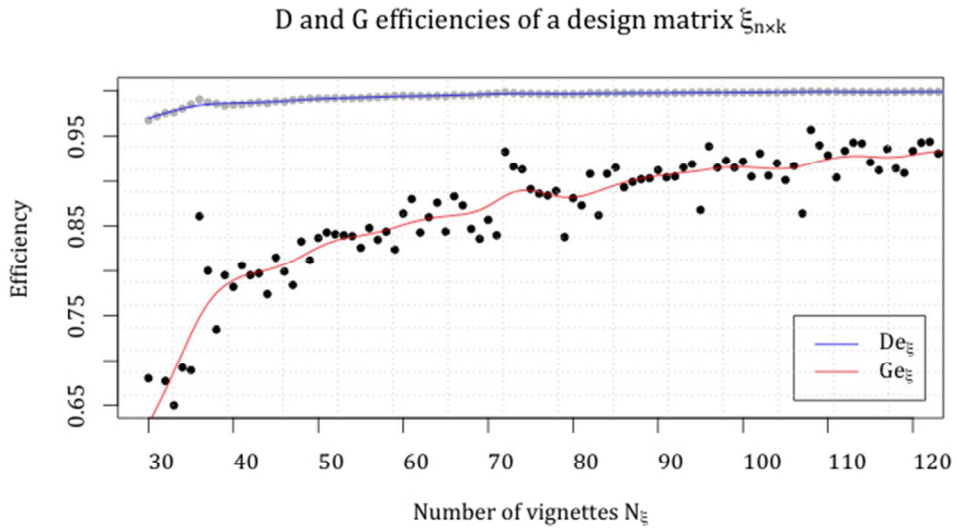


Figure 1: D and G efficiencies of design matrix ξ.
D and G efficiencies of design
238x143mm (72 x 72 DPI)

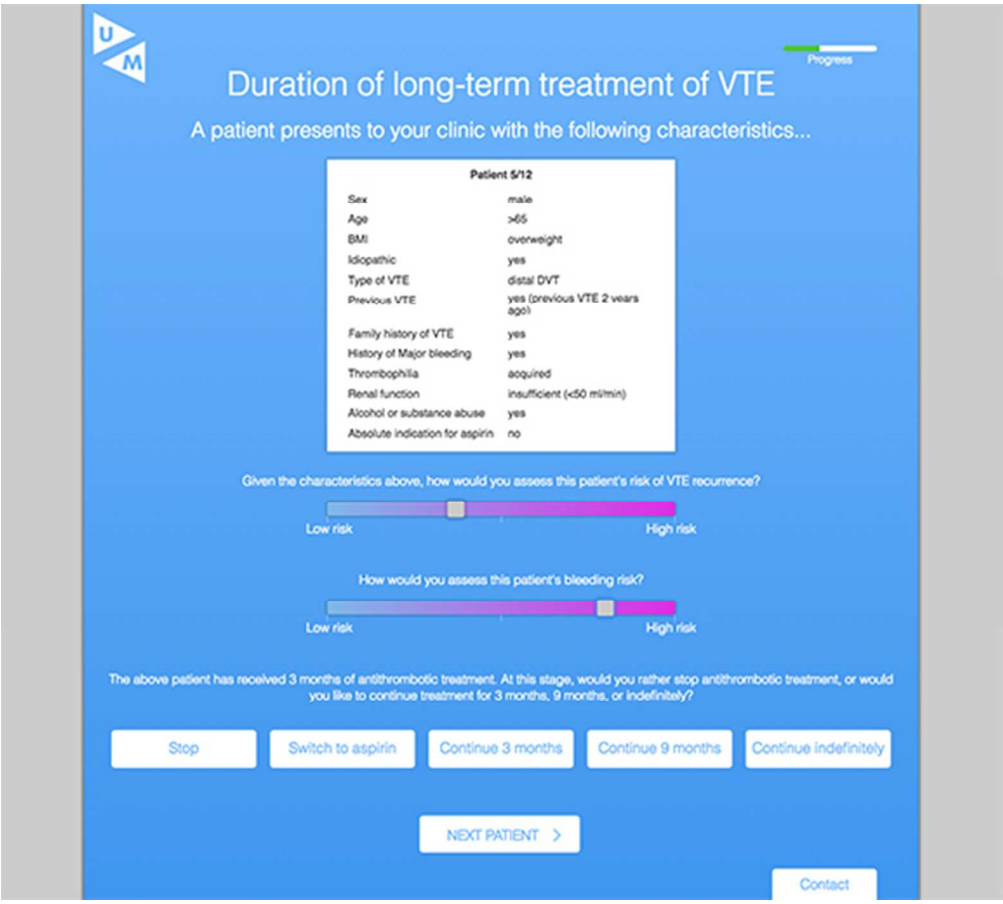


Figure 2. Screen capture of the web-based survey (www.vte-survey.com).
Figure 2. Screen capture of th
203x182mm (72 x 72 DPI)

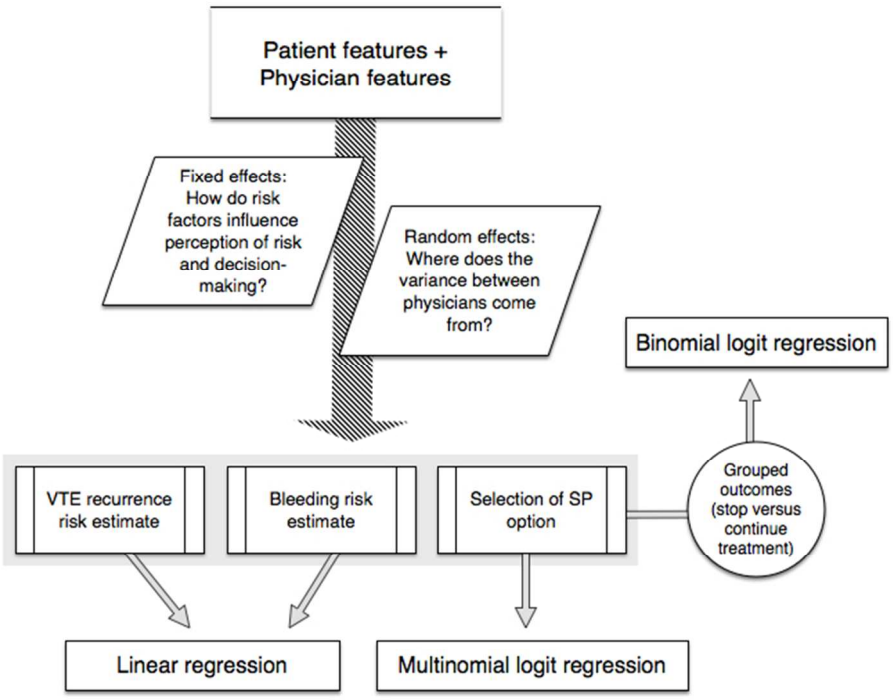


Figure 3. Analysis flow chart.
Figure 3. Analysis flow chart.
255x193mm (72 x 72 DPI)

BMJ Open

Determinants of treatment duration in the prevention of recurrent venous thromboembolism: a protocol for a balanced vignette experiment.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015231.R1
Article Type:	Protocol
Date Submitted by the Author:	09-Mar-2017
Complete List of Authors:	Ten Cate, Vincent; Maastricht University, Epidemiology; Maastricht University CAPHRI School for Public Health and Primary Care Essers, Brigitte; Maastricht University Medical Center, Clinical Epidemiology and Medical Technology Assessment Prins, Martin; Maastricht University Medical Center, Epidemiology
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Epidemiology, Evidence based practice, Cardiovascular medicine
Keywords:	EPIDEMIOLOGY, STATISTICS & RESEARCH METHODS, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Vascular medicine < INTERNAL MEDICINE

Title

Determinants of treatment duration in the prevention of recurrent venous thromboembolism: a protocol for a balanced vignette experiment.

Authors and affiliations

Ten Cate V[1,2], Essers BAB[3], Prins MH[1].

1. Maastricht University Medical Center, Department of Epidemiology Maastricht, the Netherlands.

2. Maastricht University CAPHRI School for Public Health and Primary Care Maastricht, the Netherlands.

3. Maastricht University Medical Center, Department of Clinical Epidemiology and Medical Technology Assessment

Corresponding author

Vincent Ten Cate, MSc
Maastricht University Medical Center, Department of Epidemiology
Maastricht, the Netherlands.
Email: vincent.ten.cate@mumc.nl
Telephone: +31631766450

Keywords

Venous Thromboembolism, Secondary Prophylaxis, Duration of Treatment, Surveys and Questionnaires

Abstract (292 words)

Introduction

Venous thromboembolism (VTE) is a condition that annually occurs in approximately 1 ‰ of the world's population. Patients who have already had a VTE are at elevated risk for a recurrent VTE. Recurrent events increase the risk of long-term sequelae and can be fatal. Adequate secondary prophylaxis is thus needed to prevent such events.

VTE patients are often prone to bleeding, and pharmacologic prophylaxis exacerbates bleeding risk. Expert opinions on the optimum duration of secondary prophylaxis in VTE still vary substantially. The existence of treatment guidelines has not led to uniformity of VTE secondary prophylaxis strategies, which means that physicians still adhere to individual risk calculi in determining treatment duration.

Methods and Analysis

The aim of this study is to establish what factors lie at the root of this variance in VTE secondary prophylactic treatment strategies, and what risk factors are deemed of particular importance in determining the perceived risks and benefits of variable treatment durations. To do this, we created a survey based on a D and G-efficient balanced experimental vignette design. This protocol covers all aspects of how this survey was set up and how it was implemented.

The analysis of the experimental data will be carried out using mixed-effects methods, which are beneficial in scenarios with high interindividual variance and correlated (e.g. repeated-measures) responses. We propose the use of maximal random effects structures insofar possible.

Ethics and Dissemination

All data are de-identified, and any identifying characteristics of the respondents will not be reported in a final manuscript or elsewhere.
A paper describing the expert interviews is currently under peer review. A manuscript that contains the analysis of the results of the experiment described in this protocol is being drafted, and will also be submitted to a peer-reviewed journal.

Strengths and limitations

Strengths

- The experiment uses a D and G-efficient balanced fractional factorial design.
- The design was inspired by and modified based on interactions with several specialists in the field.
- Mixed effects methods are used to interpret and account for interindividual variance.

Limitations

- Not all relevant risk factors were incorporated in the model due to technical constraints.
- Due to the nature of the experiment, the risk factors featured in it do not necessarily occur in real-life distributions.

Word count

3895 words.

Introduction

Venous thromboembolism (VTE) is a condition that annually occurs in approximately 1 ‰ of the world's population [1]. VTE encompasses deep vein thrombosis (DVT) and pulmonary embolism (PE). The elderly are at particular risk, as the incidence of VTE increases with age to approximately 5-6‰ in those over 80 years old [1]. Two-thirds of VTE diagnoses constitute cases of deep vein thrombosis, and the remainder pulmonary embolism. 25-50% of VTE cases are idiopathic, which means that they are of unknown etiology [2].

Approximately 6% of DVT cases and 12% of PE cases are lethal within one month of diagnosis [2]. Recovery rates from a VTE event are high given appropriate pharmacologic anticoagulant therapy [2]. If the embolus in patients with PE is not completely resolved after treatment, long-term complications such as chronic thromboembolic pulmonary hypertension (CTPH) may occur [3]. Similarly, patients with DVT are at elevated risk of post-thrombotic syndrome (PTS) during long term follow up. Moreover, patients who have already had a VTE are at elevated risk for a recurrent VTE [4]. Recurrent events increase the risk of long-term sequelae and are potentially fatal [4]. Hence, adequate secondary prophylactic treatment, i.e. the outpatient preventive treatment that follows the resolution of an acute event, is necessary. However, what exactly constitutes adequate secondary prophylaxis of VTE is still controversial: physicians have yet to achieve consensus on both the ideal type and duration of treatment [5-10].

The difficulty in achieving consensus regarding the duration of secondary thromboprophylaxis in patients with VTE stems from the balance that has to be struck between the thrombotic risk associated with non-treatment, and the reduction in risk of recurrence and the induced bleeding risk associated with anticoagulant therapy. As stated, VTE is a disease that disproportionately affects the elderly, and the elderly are a group particularly prone to bleedings on account of their frailty. Physicians are thus faced with the quandary of prescribing a treatment that itself poses a risk to the health of their patients, while not doing so might incur an even greater risk due to recurrent disease. Their task is then to find the optimum duration of treatment that most effectively mitigates the risk of VTE recurrence in a patient, while minimizing their exposure to an inherently risky treatment.

Organizations such as the American College of Chest Physicians (CHEST, formerly: ACCP) and the European Society of Cardiology (ESC) have attempted to coordinate and formalize the treatment and prevention of VTE with periodical installments of treatment guidelines [11,12]. However, the recommendations outlined in these guidelines are wont to change with every new installment, and it is unclear how many physicians adhere to these guidelines. Moreover, these guidelines represent a combination of evidence-based research and expert opinions, which themselves vary from physician to physician. The ongoing debate on the optimum type and duration of treatment is testament to the existence of many different calculi individual physicians employ in their long-term preventive VTE treatment strategies.

The aim of this study is to establish what factors lie at the root of this variance in VTE secondary prophylactic treatment strategies, and what risk factors are deemed of particular importance in determining the perceived risks and benefits of ceasing or continuing preventive treatment for a particular duration of time. We do this by having physicians evaluate patient vignettes, i.e. hypothetical patient profiles, the attributes of which were derived from expert interviews.

Methods and analysis

Vignette experiment

Vignettes are hypothetical scenarios that mimic real-life situations, enabling researchers to evaluate their phenomenon of interest in a controlled experimental setting [13]. Since there is little noise in a designed experiment, covariates can be assessed in a relatively unbiased manner [13]. Additionally, the use of vignettes precludes the need to recruit a balanced sample of subjects, as such a sample can simply be simulated. In our case, we aim to establish what drives the decision-making process surrounding secondary prophylaxis of VTE. In a real-life setting, it would be difficult to recruit a diverse enough set of patients to take the various risk factors that may influence this process into account. With the vignette experiment described in this paper, we achieved a highly efficient alternative to a much more costly and time-consuming real-life experiment.

Preselection of attributes

From December 2015 through February 2016, we conducted 13 semi-structured interviews with an equal number of senior-level physicians affiliated with medical centers around the world that participate in the Einstein Choice study [14]. The purpose of these interviews was to establish which factors are particularly important in determining the duration of secondary prophylaxis of VTE, and, conversely, which factors weigh against a decision to continue treatment past the resolution of the initial event. Our focus during these interviews was on specialist decision-making in the period directly following the resolution of the initial event, and so patient opinion and risk factors relating to period following the discontinuation of treatment (e.g. D-dimer levels) were not taken into account. Our questions were informed by a review of the pertinent literature and guidelines [11,12], which had yielded an initial selection of relevant attributes. Importantly, we decided to exclude cancer as a risk factor, as its impact on the decision to continue or stop anticoagulation is too dominant: presence of an active cancer invariably warrants indefinite anticoagulation, regardless of other factors [11].

In our interviews, we specifically asked our interviewees to identify patient characteristics that significantly impact either the risk of VTE recurrence or the risk of bleeding. Although all physicians weighted individual risk factors differently, common factors and their corresponding levels were identified. Table 1 is a frequency table of the risk factors that were considered to

significantly impact either on VTE recurrence risk or on bleeding risk, by number of physicians who mentioned them. This table does not include risk factors that were mentioned by fewer than two interviewees.

Reducing design complexity

The preselection of attributes yielded 18 factors that were indicated to be of importance by at least two physicians. A recommendation for conjoint experiments, which exhibit commonalities with our approach, states to limit the complexity of a design as much as possible, as an overly complex design may confuse respondents and distort the responses to an experiment [15]. We first discarded factors that were unsuitable for the vignette experiment due to their ungeneralizability: uncontrolled hypertension and physical activity. We then removed the 'unstable or high INR' factor on account of it being specific to vitamin K antagonists. We further trimmed down the list of factors by discarding the risk factors that only rarely occur in real life VTE patients: thrombocytopenia (<1%) [16] and liver disease (~1.5%) [17]. Finally, we combined 'signs of PTS' with the 'type of VTE' factor, so as to prevent inappropriate interactions: post-thrombotic syndrome almost exclusively occurs in tandem with DVT, rather than with PE [18]. This leaves 12 factors in our reduced model.

Experimental design

Given 12 factors, of which nine have two levels; two have three levels; and one has six levels, there is a total of $2^9 3^2 6 = 27648$ level permutations in the full factorial model. As having respondents evaluate all 27648 possible scenarios would be prohibitively costly, we created an efficient fractional factorial design via Fedorov's 1972 exchange algorithm [19]. To generate this design, we utilized the R package 'AlgDesign' [20], an implementation of the Fedorov algorithm. The Fedorov algorithm randomly selects rows from full factorial matrix X to create a new matrix ξ, and then exchanges rows from ξ with rows from $X \setminus \xi$ to optimize a specified criterion. In our case, we optimized our reduced design for *D* and *G*-efficiency.

The number of vignettes in the reduced design was selected on the basis of the following criteria:

I. D-efficiency
II. G-efficiency
III. Level balance
IV. Feasibility of the required sample size

I. D-efficiency

The *D*-efficiency, perhaps the most frequently used efficiency criterion in the design of experiments (DoE) [21,22], is a measure of information in a reduced design ξ relative to the full factorial matrix *X*. In the AlgDesign package, the D criterion for a design ξ is computed as the k^{th} root of its generalized variance [20], i.e. the determinant of the normalized covariance matrix $\Sigma(\xi) = (\xi^T \xi / N_\xi)$:

$$D_\xi = |\Sigma(\xi)|^{\frac{1}{k}}$$

Where k is the number of columns and N_ξ the number of rows in ξ. The relative D-efficiency *De* of ξ with respect to X is computed by taking the ratio of D_ξ to the D criterion of the full factorial design:

$$De_\xi = \frac{D_\xi}{D_X}$$

II. G-efficiency

G-efficiency is another optimality criterion, which aims to minimize the prediction variance of a fractional factorial design [22]. In the AlgDesign package, G-efficiency is computed as follows [20]:

$$Ge_\xi = \frac{k}{\max\limits_{x \in X} (x^T \Sigma(\xi)^{-1} x)}$$

Here, x denotes a row of X. In natural language, this means that the number of columns k is divided by the row in X with the highest precision (the inverse of the covariance matrix is the precision matrix [23]) of ξ over X.

III. Level balance

Level balance is an important attribute of experimental designs, which ensures design orthogonality [24]. In order to achieve level balance, the number of rows N_ξ should be an integer multiple of the number of levels for all factors. This ensures that all levels can be represented an equal number of times per factor in the design matrix, creating a balanced set of vignettes [25].

IV. Feasibility of the required sample size

There does not appear to be a uniform standard when it comes to determining an appropriate sample size for experimental designs such as the one described in this protocol. One formula, used in conjoint analyses, dictates a sample size of $n = 500 \times c/(a \times t)$, where c signifies the maximum number of levels for any of the factors, a denotes the number of choice alternatives presented at once, and t denotes the number of unique vignettes presented to each respondent [26-28]. Another recommendation states that all vignettes in a design should be evaluated at least six times [29].

Number of vignettes in the reduced design

Based on the above considerations, we selected a design ξ with $N_\xi = 72$. This number of vignettes yielded a well-rounded balance between our efficiency criteria, as illustrated in figure 1 and table 2. To reduce workload for our respondents, we divided this design into six blocks of 12 vignettes each, using

the optBlock function in the 'AlgDesign' R package. After optimization, we achieved a block design diagonality of 0.88. Plugging in the values for a single block into our sample size formula returns a required sample size of $n = 500 \times \frac{6}{1 \times 12} = 250$. As there is considerable overlap between blocks, we assume a sample size of 250 to be sufficient to draw reliable inferences with the design as specified above. Following the 'six-times rule', accounting for twelve vignette evaluations per respondent, yields a required sample $n = \frac{72 \times 6}{12} = 36$.

Confirmation of the required sample size via a priori *power analysis*

As we intend to analyze the experimental results with mixed-effects models (see: *Analysis of the results*), it is necessary to validate the sufficiency of the above sample size calculation [30]. We performed a series of simulations, using the R command 'simulate' and the 'makeLmer' and 'makeGlmer' commands in the R package 'simr' [31], to construct an *a priori* power analysis. We conducted two linear mixed effects models and four generalized linear mixed effects models with logit link functions using the R package 'lme4' [32]. We used pilot data (see: *Pilot evaluation of the design*) to construct these models. Specifically, we regressed outcome variables 'VTE recurrence risk' and 'bleeding risk' (see: *Outcome variables*) on a vector of covariates, consisting of the selected patient risk factors and participating specialist characteristics (see: *Physician characteristics*). We then systematically reduced these models via AIC-based forward and backward stepwise variable selection, and by eliminating variables with clinically insignificant effect sizes (<|0.01| on a scale 0-10) or particularly unpromising p-values (>0.6), as determined via Type II Wald χ^2 tests. We created four binomial generalized linear mixed models to separately analyze the four treatment outcomes in our original pilot survey (continue 3 months; continue 6-12 months; continue indefinitely; cease treatment or switch to aspirin). For the 'continue' treatment options we adapted and modified the 'recurrence risk' model's equation until it exhibited adequate fit and model convergence. For the 'stop treatment' option we used the same process, this time modifying the 'bleeding risk' model's equation. Random slopes and intercepts were included as much as possible; the final selection of mixed effects was based on model improvement and convergence. The recurrence risk model includes 10 covariates, one random slope and three random intercepts. The bleeding risk model features 7 covariates, two random slopes and three random intercepts. The model corresponding to treatment 1 features 8 covariates and a random intercept for participants. Logistic model 2 includes 5 covariates and three random intercepts (participant, block identifier, country). Logistic model 3 incorporates 7 covariates and random intercepts for participant and block identifier. Logistic model 4 features six covariates and three random intercepts (participant, block identifier, country).

From these models created on the basis of pilot data, we extracted the covariate coefficients $(\widehat{\beta})$, the variance and correlation components and the residual standard deviation $(\widehat{\sigma})$. We then generated covariate (X) matrices corresponding to varying sample sizes by concatenating blocks of vignettes and simulating plausible participant characteristics. Finally, we simulated responses for each of

the models at variable sample sizes. We defined the power of all models as the proportion of incorporated covariates with p-values <0.05, assessed with Type II Wald χ^2 tests. The power analysis is summarized in table 3 and figure 2. The power analyses indicate that the lower bound of the sample size should be adjusted to N=100. The original upper bound of N=250 was validated as amply sufficient to estimate coefficients with reliable power.

Pilot evaluation of the design

In May 2016, we sent out Microsoft Word-based pilot surveys to twelve thrombosis experts to evaluate our initial design. From the comments we received, it was apparent that some of the factors and levels in our model were not formulated precisely enough. For instance, pilot participants indicated that in patients who had had a previous VTE, it was important to know when this had occurred. For this item, we added a timeframe to improve clarity. Some participants struggled with the unqualified formulation 'renal function: insufficient'. We included the more detailed description '<50 ml/min' in these cases, referring to the corresponding creatinine clearance rate. Additionally, of patients who received concurrent antiplatelet therapy, a number of pilot participants said to be confused as to whether the practitioner treating the patient for VTE recurrence could cease this concurrent therapy. As our intention was to refer to patients with an absolute indication for antiplatelet therapy, we reformulated this item to more explicitly refer to such patients. With these changes, we arrived at our final design, shown in table 4.

Physician characteristics

From our expert interviews, it was apparent that physicians involved in the secondary prevention of VTE all have unique perspectives on thrombosis and bleeding risk in VTE. In order to understand why physicians who are faced with the same clinical vignettes may respond to these differently, we decided to include a number of questions relating to characteristics of the physicians themselves. Before commencing the experiment, physicians are asked to provide their age, gender, country of employment, specialty, years of experience with the treatment of VTE, and an estimate of the number of unique VTE patients they treat per annum. These physician characteristics were selected on the basis of previous research into physician guideline adherence and clinical decision-making [33-37].

By incorporating the 'country of employment' variable as a random effect, we hope to capture the variance resulting from any country-specific aspects of physician's decision-making that would otherwise be difficult to quantify.

We provided six possible answers to choose from in response to the specialty question, i.e.: internal medicine, vascular medicine/angiology, cardiology, hematology, pulmonology or other. The specialty of responding physicians will also be included as a random effect in our analysis.

Outcome variables

As outcome variables, we incorporated two continuous variables to respectively estimate VTE recurrence risk and bleeding risk on a 100-point visual analogue scale indicating low to high risk, and five discrete variables: continue treatment for three additional months, continue 9 additional months, continue indefinitely, stop treatment, or switch to aspirin. These choices were based on the most recent CHEST and ESC VTE treatment guidelines [11,12], the conducted expert interviews and the comments received on our pilot design.

Implementation of the survey

The survey has been implemented online by independent web developers, and is currently live at www.vte-survey.com (figure 3). Survey respondents are randomized to one of six pre-specified blocks at the start of the experiment.

As mentioned above, one aspect of our experimental design was to make the workload for study participants as slight as possible. For this reason, in the online implementation of the survey the progress of participants is continuously recorded. This allows participants to take breaks during the survey and even to leave the website, without their responses up until that point being forfeited. Participants may return to the website at any time and resume the survey where they left off. In our current sample of respondents, there is a small number of such cases that skews the mean survey completion time considerably upward. A better reflection of the average workload is given by the median survey completion time, which in our current sample is equal to 9 minutes. Based on this number, we conclude that the risk of participant fatigue is negligible for our study.

Data collection was commenced in July 2016, and is still ongoing. We have reached our stated sample size requirements in January 2017, but will leave the website live for at least the remainder of this year, so that more data may be recorded for future analysis. We welcome any specialists involved in the preventive treatment of VTE to participate in this experiment.

Participant recruitment strategy

As we are interested in specialist decision-making, our recruitment strategy focuses exclusively on specialists. This group of subjects is particularly difficult to reach in a non-systematic or stochastic manner. This section details how we approached the recruitment of participants in a way that we believe maximized the chance of reaching a broad and representative sample.

In order to meet the sample size requirements, as determined by the vignette-based sample size formulae as well as the power analysis described above, we have leveraged several channels to recruit study participants so far. First, we contacted a list of several hundred specialists who participate or have participated in the Einstein Choice study [14]. Participation in this context is defined as having recruited at least one patient for the Einstein Choice study, and the list features specialists of variable ages, professional stature and levels of

experience. We also contacted professional collectives of specialists, such as the International Society of Thrombosis and Haemostasis (ISTH), the European Society of Cardiologists (ESC) and the Asian Pacific Society on Thrombosis and Hemostasis (APSTH), which led to calls for participation in the ISTH (October 2016) and APSTH newsletters (January 2017) and on their respective websites, and a call for participation on the youth community (EAPC) of the ESC on LinkedIn. Finally, we asked each specialist who responded favorably to our request to participate to disseminate the survey among colleagues and residents.

Analysis of the results

We will conduct several separate analyses of our results (figure 4). A first linear mixed effects regression model will relate all risk factors and physician characteristics to the continuous estimated VTE recurrence risk outcome variable. A second model does the same for bleeding risk. After having estimated the coefficients in these two linear models, we will use generalized linear mixed models (GLMMs) to estimate the odds ratios for all independent variables, including physician characteristics, in relation to the choice of treatment options. We will use a GLMM with a binomial logit link function for grouped outcomes (stop versus continue treatment), and a multinomial logit regression to discover which variables best predict the five separate treatment outcomes.

In the linear models, we will evaluate model fit via diagnostic plots (AV plots, Q-Q plots, standardized residuals vs fitted plots), the adjusted R^2 and relative measures such as the Akaike Information Criterion (AIC). In the logistic models, we will employ diagnostic plots, the area under the receiver operating characteristic curve (AUROC), McFadden's pseudo-R^2 (ρ^2) [38] and the predictive accuracy of the models on a test set (kept separate from the development set) to evaluate model fit.

The choice for mixed effects analyses was informed by the theoretical assumption that there is considerable inter-physician variability, and to control for random variance associated with aspects of no interest (e.g. block allocation). While analysis of the fixed effects (i.e. the regression coefficients) will provide general insights as to which features are of particular interest in determining a patient's risk of VTE recurrence or bleeding, it is the decomposition of the error term that will show us specifically where and to what extent physicians disagree with one another. To this end, we will specify random slopes that vary by physician for as many relevant features as possible. Our intention in this regard is to keep random effects structures 'maximal' [39], albeit within the realistic constraints imposed by the data [30]. All analyses will be conducted in the statistical software package R.

Implications of this research

The research described in this protocol will provide insight into the mechanisms underlying the decision-making of physicians involved in the outpatient treatment of VTE. Our aims are to identify the relative weights of risk factors; to

determine how physicians decide on the duration of preventive treatment, and to uncover what prompts physicians to discontinue further treatment.

Simultaneously, this research places a strong focus on the variance that exists between physicians in determining VTE treatment duration. We will explore which factors contribute to this variance by using a mixed-effects approach, enabling direct analysis of the inter-individual differences in decision-making patterns. In so doing, it is our hope that this research will lead to a better understanding of why physicians may deviate from best-practice guidelines, and what biases may potentially be at play in making this decision.

We believe that the insights acquired through this study will benefit physicians involved in the treatment of VTE, by elucidating the overlaps and discrepancies between treatment strategies and by highlighting the areas potentially prone to oversights. For this same reason, our findings should also be of benefit to the national and international collectives whose aim it is to unify and coordinate VTE treatment. By explicitly showing where physician decision-making coincides and where it varies, we believe that VTE prevention education and policy-making can at once be made more efficient and more effective.

Ethics and dissemination

All data are de-identified, and any identifying characteristics of the respondents will not be reported in a final manuscript or elsewhere. This study does not make use of real patients, and participant data is not released to third parties. For these reasons, this study is exempt from the requirement to gain approval from a medical research ethics committee (MREC).

A paper containing a qualitative summary of the expert interviews is currently under peer review. We are currently working on a manuscript that contains the analysis of the results of the experiment described in this protocol. When completed, this article will also be submitted to a peer-reviewed journal.

References

[1] Cushman M. Epidemiology and Risk Factors for Venous Thromboembolism. Semin Hematol. 2007 April 44(2): 62-69.

[2] White RH. The epidemiology of venous thromboembolism. Circ. 2003; 107(23 suppl 1):I4-I8.

[3] Baser O. Prevalence and Economic Burden of Venous Thromboembolism After Total Hip Arthroplasty or Total Knee Arthroplasty. Am J Manag Care. 2011;17:S6-S8.

[4] Zhu T, Martinez I, Emmerich J. Venous Thromboembolism: Mechanisms, Treatment, and Public Awareness. Arterioscl Throm Vas. 2009; 29: 298-310.

[5] Beyer-Westendorf J, Cohen AT, Monreal M. Venous thromboembolism prevention and treatment: expanding the rivaroxaban knowledge base with real-life data. Eur Heart J Supp. 2015;7(17)suppl D:D32-D41.

[6] Sternitzky R, Hochauf S, Schellong SM. Secondary prophylaxis of venous thromboembolism. Hamostaseologie. 2007 Feb;27(1):32-40.

[7] Hibbert PD, Hannaford NA, Hooper TD et al. Assessing the appropriateness of prevention and management of venous thromboembolism in Australia: a cross-sectional study. BMJ Open 2016;6:e008618.

[8] Caprini JA, Tapson VF, Hyers TM et al. Treatment of venous thromboembolism: adherence to guidelines and impact of physician knowledge, attitudes, and beliefs. J Vasc Surg. 2005 Oct;42(4):726-33.

[9] Brenner B. Venous thromboembolism: advances and controversies. Best Pract Res Clin Haematol. 2012 Sep;25(3):233. doi: 10.1016/j.beha.2012.07.009.

[10] Schulman S, Douketis J. Secondary prevention of venous thromboembolism. BMJ. 2013 Sep 9;347:f5440. doi: 10.1136/bmj.f5440.

[11] Kearon C, Akl EA, Ornelas J et al. Antithrombotic Therapy for VTE Disease: CHEST Guideline and Expert Panel Report. Chest. 2016 Feb;149(2):315-52.

[12] Konstantinides S et al. 2014 ESC Guidelines on the diagnosis and management of acute pulmonary embolism. Eur Heart J. 2014; 35: 3033–3080.

[13] Evans SC, Roberts MC, Keeley JW et al. Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies. Int J Clin Hlth Psyc. 2015; 15(2):160-170.

[14] Bayer. Reduced-dosed Rivaroxaban in the Long-term Prevention of Recurrent Symptomatic VTE (Venous Thromboembolism) (Einstein Choice). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000- [cited 2016 Oct 27]. Available from: https://clinicaltrials.gov/ct2/show/NCT02064439 NLM Identifier: NCT02064439.

[15] DeShazo JR, Fermo G. Designing choice sets for stated preference methods: the effects of complexity on choice consistency. J Environ Econ Manag 2002;44:123–43.

[16] Stein PD, Hull RD, Matta F et al. Incidence of thrombocytopenia in hospitalized patients with venous thromboembolism. Am J Med. 2009 Oct;122(10):919-30.

[17] Saleh T, Matta F, Alali F et al. Venous thromboembolism with chronic liver disease. Am J Med. 2011 Jan;124(1):64-8.

[18] Bova C, Rossi V, Ricchio R et al. Incidence of post-thrombotic syndrome in patients with previous pulmonary embolism. A retrospective cohort study. Thromb Haemost. 2004 Nov;92(5):993-6.

[19] Fedorov VV. Theory of optimal experiments. Elsevier; 1972.

[20] Wheeler RE. Comments on algorithmic design 2004–2009. http://cran.r-project.org/web/packages/AlgDesign/vignettes/AlgDesign.pdf. (accessed April 2016)

[21] Kuhfeld WF. Efficient experimental designs using computerized searches. Research Paper Series, SAS Institute, Inc. 1997 Nov.

[22] Rodríguez M, Jones B, Borror CM, Montgomery DC. Generating and assessing exact G-optimal designs. Journal of quality technology. 2010 Jan 1;42(1):3.

[23] Wasserman L. All of statistics: a concise course in statistical inference. Springer Science & Business Media; 2013 Dec 11.

[24] Collins LM, Dziak JJ, Li R. Design of Experiments with Multiple Independent Variables: A Resource Management Perspective on Complete and Reduced Factorial Designs. Psychol Methods. 2009 Sep;14(3):202-224.

[25] Kuhfeld WF, Tobias RD, Garratt M. Efficient Experimental Design with Marketing Research Applications. J Mark Res. 1994;31(4):545-557.

[26] Rose JM, Bliemer MCJ. Sample size requirements for stated choice experiments. Transportation 2013;40:1021–41.

[27] Vilikus O. Optimization of Sample Size and Number of Tasks per Respondent in Conjoint Studies Using Simulated Datasets. Studia Ekonomiczne/Uniwersytet Ekonomiczny w Katowicach. 2012(120 Survey sampling in economic and social research):77-83.

[28] Engel J, van der Wulp I, Poldervaart JM et al. Clinical decision-making of cardiologists regarding admission and treatment of patients with suspected unstable angina or non-ST-elevation myocardial infarction: protocol of a clinical vignette study. BMJ Open. 2015 Apr 8;5(4):e006441.

[29] Louviere JJ, Hensher DA, Swait JD. Stated choice methods. Analysis and application. New York: Cambridge University Press, 2000:83–111.

[30] Matuschek H, Kliegl R, Vasishth S, Baayen H, Bates D. Balancing Type I error and power in linear mixed models. J Mem Lang. 2017 Jun;94:305-15.

[31] Green P, MacLeod J. simr: an R package for power analysis of generalised linear mixed models by simulation. Methods Ecol Evol. 2016; 7(4):493-8.

[32] Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015; 67(1):1-48. doi:10.18637/jss.v067.i01.

[33] Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. Ann Intern Med. 2005 Feb 15;142(4):260-73.

[34] Solomon DH, Brookhart MA, Gandhi TK et al. Adherence with osteoporosis practice guidelines: a multilevel analysis of patient, physician, and practice setting characteristics. Am J Med. 2004 Dec 15;117(12):919-24.

[35] Kumar A, Fonarow GC, Eagle KA et al. Regional and practice variation in adherence to guideline recommendations for secondary and primary prevention among outpatients with atherothrombosis or risk factors in the United States: a report from the REACH Registry. Crit Pathw Cardiol. 2009 Sep;8(3):104-11. doi: 10.1097/HPC.0b013e3181b8395d.

[36] McKinlay JB, Lin T, Freund K, Moskowitz M. The unexpected influence of physician attributes on clinical decisions: results of an experiment. J Health Soc Behav. 2002 Mar;43(1):92-106.

[37] Berthold HK, Gouni-Berthold I, Bestehorn KP, Böhm M, Krone W. Physician gender is associated with the quality of type 2 diabetes care. J Intern Med. 2008 Oct;264(4):340-50. doi: 10.1111/j.1365-2796.2008.01967.x.

[38] McFadden D. Conditional logit analysis of qualitative choice behavior. 1973.

[39] Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J Mem Lang. 2013 Apr;68(3). doi: 10.1016/j.jml.2012.11.001.

Acknowledgments

We extend our gratitude to the thirteen interviewees whose input helped us select risk factors and outcome variables for our vignette experiment. Our thanks also go out to Maastricht University Medical Center's thrombosis department, Dr. Avi Leader of Israel and Dr. Michael Nagler of Switzerland for their valuable comments and evaluation of our pilot survey. We further thank Petro van Bergen for his help in recruiting survey participants, and the International Society for Thrombosis and Haemostasis (ISTH) and the Asian Pacific Society on Thrombosis and Hemostasis (APSTH) for featuring this experiment in their newsletters (October 2016; January 2017). We thank the youth community of the European Society of Cardiology (ESC) for featuring a call for participation on their youth community's LinkedIn page. We are grateful to Dr. David McAneny of Boston University, Dr. Charles Mahan of the University of New Mexico and Dr. José A. López-López of the University of Bristol, whose comments have inspired substantial revision of this manuscript. Lastly, we are indebted to all survey participants for contributing to our research.

Author contributions

VTC and MHP designed the protocol. VTC wrote the protocol. BABE provided helpful input throughout.

Funding statement

This work was supported by Bayer AG, grant number BAY59-7939/16416add.

Competing interests

VTC and BABE declare no competing interests. MHP has received research support and honoraria, and participated in advisory boards for Bayer HealthCare, Sanofi-Aventis, Boehringer Ingelheim, GlaxoSmithKline, Daiichi Sankyo, LeoPharma, ThromboGenics, and Pfizer.

Data sharing statement

The R code and data that were used to generate the described design, the simulation-based power analysis and the figures are available upon request. Please contact the first author, VTC, if you would like to receive the relevant scripts and datasets.

Tables

Table 1. Frequency table of risk factors mentioned by interviewed physicians

Factors (k)	Levels	Type of risk factor	Freq. (%)
Idiopathic	no – yes	TR	13 (100)
VTE type	distal DVT – proximal DVT – non-massive PE – massive PE	TR	13 (100)
Thrombophilia	none – acquired – inherited	TR	11 (85)
History of bleeding	no – yes	BR	10 (77)
On concurrent antithrombotic medication (e.g. aspirin)	no – yes	BR	9 (69)
Age	<65 – >65	BR/TR	8 (62)
Renal function	normal – insufficient	BR	7 (54)
Liver disease	no – yes	BR	7 (54)
Unstable or high INR	no – yes	BR	7 (54)
History of thrombosis	no – yes	TR	6 (46)
BMI	underweight – normal weight – overweight	BR/TR	5 (38)
Uncontrolled hypertension	no – yes	BR	5 (38)
Alcohol or drug abuse	no – yes	BR	5 (38)
Family history of VTE	no – yes	TR	4 (31)
Physical activity	sedentary – active	TR/BR	4 (31)
Thrombocytopenia	no - yes	BR	3 (23)
Signs of PTS	no – yes	TR	3 (23)
Sex	female – male	TR	2 (15)

BR denotes bleeding risk; INR denotes international normalized ratio; PTS denotes post-thrombotic syndrome; TR denotes thrombotic risk (i.e. VTE recurrence risk).

Table 2. Efficiency criteria of design matrix ξ.

N_ξ	De_ξ	Ge_ξ
72	0.998	0.94

Table 3. Simulated a priori power analysis: power by sample size.

	Power								
Model	N=10	N=20	N=30	N=40	N=50	N=60	N=100	N=200	N=250
Recurrence risk (LMER)	0.50	0.70	0.70	0.80	0.80	0.90	0.90	0.90	0.90
Bleeding risk (LMER)	0.71	0.86	0.86	0.86	0.86	1.00	1.00	1.00	1.00
Treatment 1 (GLMER)	0.00	0.13	0.38	0.50	0.75	0.75	0.88	1.00	1.00
Treatment 2 (GLMER)	0.00	0.00	0.20	0.20	0.60	0.40	1.00	1.00	1.00
Treatment 3 (GLMER)	0.00	0.50	0.50	0.60	0.90	0.90	0.90	1.00	1.00
Treatment 4 (GLMER)	0.17	0.50	0.67	1.00	1.00	1.00	1.00	1.00	1.00

Table 4. Final experimental design.

Factor	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Sex	female	male				
Age	<65	>65				
BMI	underweight	normal weight	overweight			
Idiopathic	no	yes				
Type of VTE	distal DVT	distal DVT with signs of PTS	proximal DVT	proximal DVT with signs of PTS	non-massive PE	massive PE
Previous VTE	no	yes (2 years ago)				
Family history of VTE	no	yes				
History of major bleeding	no	yes				
Thrombophilia	none	acquired	inherited			
Renal function	normal	insufficient (<50 ml/min)				
Alcohol or substance abuse	no	yes				
Absolute indication for aspirin	no	yes				

.

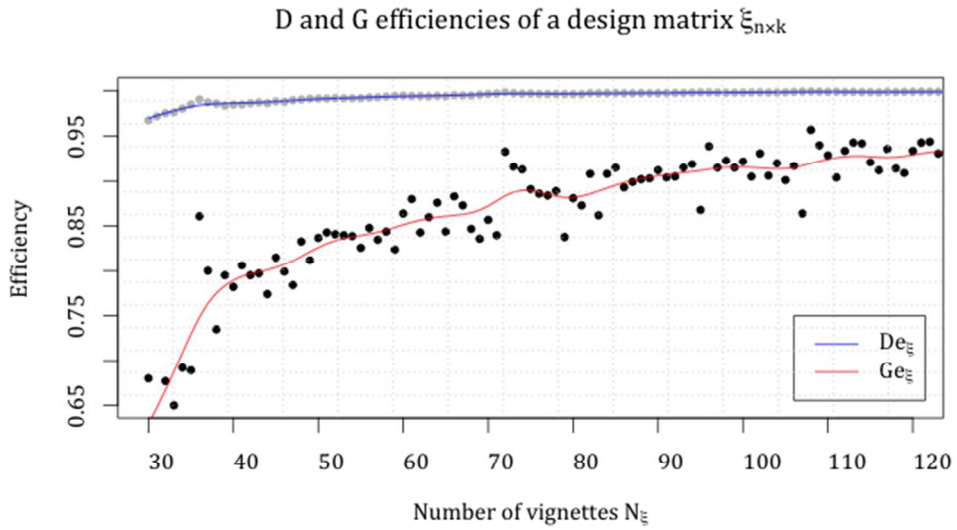


Figure 1. D and G efficiencies of design matrix ξ.
figure 1
238x143mm (72 x 72 DPI)

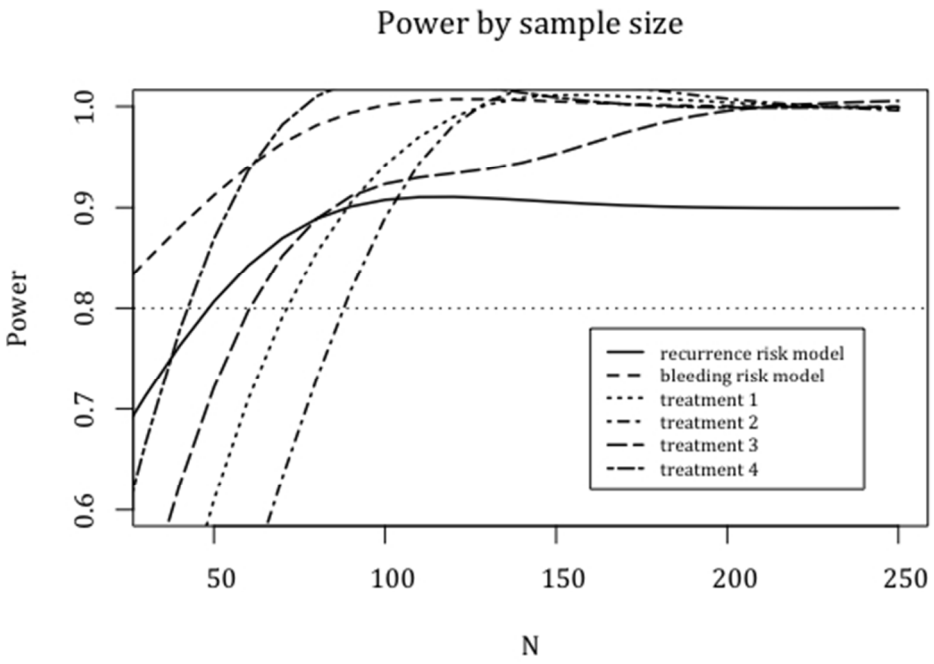


Figure 2. Power by sample size.
figure 2
203x150mm (72 x 72 DPI)

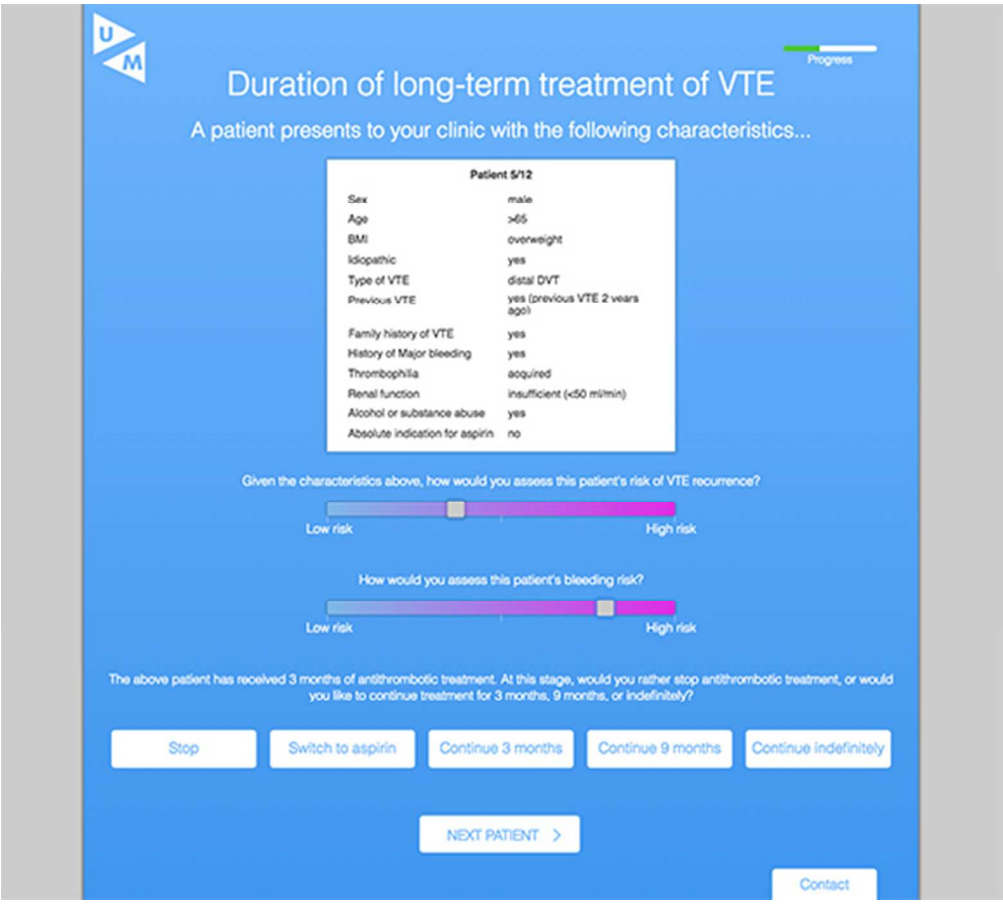


Figure 2. Screen capture of the web-based survey (www.vte-survey.com).
Figure 2. Screen capture of th
203x182mm (72 x 72 DPI)

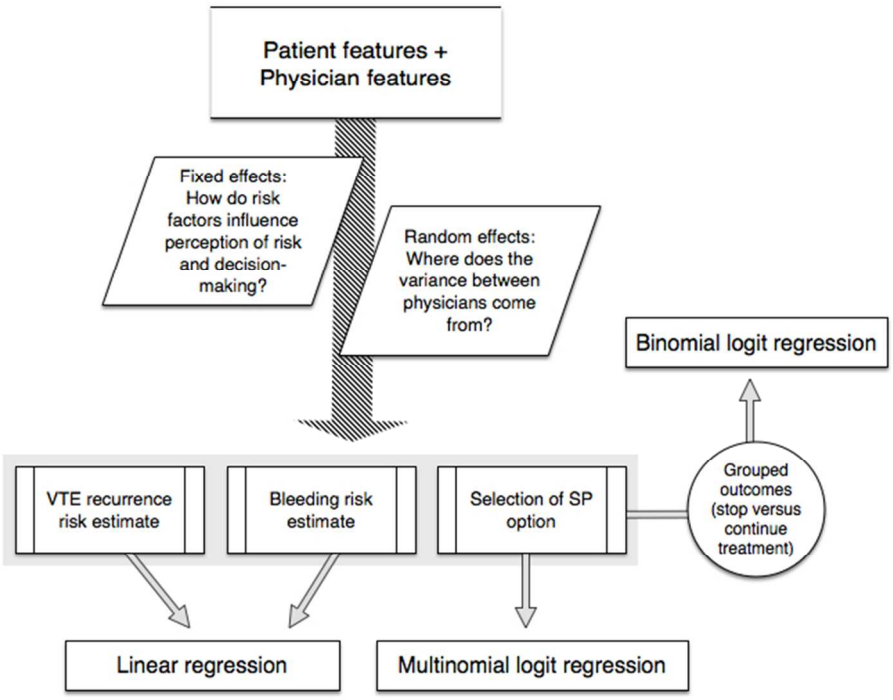


Figure 4. Analysis flowchart.
figure 4
255x193mm (72 x 72 DPI)