

Prediction of risk of recurrence of venous thromboembolism following treatment for a first unprovoked venous thromboembolism: systematic review, prognostic model and clinical decision rule, and economic evaluation

Joie Ensor,^{1,2*} Richard D Riley,^{1,2} Sue Jowett,³ Mark Monahan,³ Kym IE Snell,¹ Susan Bayliss,¹ David Moore¹ and David Fitzmaurice⁴ on behalf of the PIT-STOP collaborative group

¹Public Health, Epidemiology and Biostatistics, School of Health and Population Sciences, University of Birmingham, Birmingham, UK

²Research Institute of Primary Care and Health Sciences, Keele University, Staffordshire, UK

³Health Economics, School of Health and Population Sciences, University of Birmingham, Birmingham, UK

⁴Primary Care Clinical Sciences, School of Health and Population Sciences, University of Birmingham, Birmingham, UK

*Corresponding author

Declared competing interests of authors: none

Published February 2016

DOI: 10.3310/hta20120

Scientific summary

Prediction of risk of recurrence of venous thromboembolism

Health Technology Assessment 2016; Vol. 20: No. 12

DOI: 10.3310/hta20120

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Venous thromboembolism (VTE) is a chronic disease which may present in two ways: either a clot in the leg (deep-vein thrombosis) or a clot in the lung (pulmonary embolism). An initial VTE may be termed as either provoked, where a transient risk factor is present (such as prolonged immobility), or unprovoked, in the absence of any known risk factor. Due to the temporary nature of provoking factors, those with a first unprovoked VTE are at much higher risk of recurrent VTE (approaching 30% at 5 years after cessation of therapy) as there is no known cause.

Initial treatment for VTE comprises heparin followed by oral anticoagulation, usually for at least 3 months. However, ideal treatment duration is unclear, particularly for the unprovoked VTE population, as individuals' risk of recurrent VTE is highly heterogeneous. Although anticoagulation effectively prevents recurrent VTE, the patient is at increased risk of bleeding while on therapy. There is therefore a decision problem in balancing the risks between recurrence (off therapy), and bleeding (on therapy). It would therefore be beneficial if therapy decisions could be tailored to patients' risk so that, for example, a patient at higher risk of recurrence off therapy, than their risk of bleeding on therapy, could be recommended to continue anticoagulant therapy.

Prognostic models can be used to predict individuals' risk of recurrence based on clinical, laboratory and demographic patient characteristics. The identification of potentially important predictors which may be associated with recurrence risk is important in the development of such a prognostic model.

Aims

This project primarily aimed to develop and validate a prognostic model for the prediction of individual recurrence risk following cessation of therapy for a first unprovoked VTE. Individual patient data (IPD) were utilised from one large prospective database in order to develop a new prognostic model based on multiple predictors, and to externally validate the developed model. The final developed prognostic model allows individualised recurrence risk prediction, which may help to inform patient care as part of an evidence based approach.

To inform the development of a new prognostic model, the project also aimed to undertake a systematic review of all the evidence on existing prognostic models for VTE recurrence and adverse outcome following cessation of therapy for a first unprovoked VTE. The findings could inform clinical practice and patient care by summarising the current prognostic models and their predictive performance.

Finally, to assess the potential value of the developed model in practice, an economic evaluation was undertaken to assess the cost-effectiveness of a decision rule based on the developed model compared with current practice. Various risk thresholds were assessed, such that if a patient's predicted risk was above the threshold, the patient would be recommended to continue therapy indefinitely, and if predicted risk was below the threshold, therapy could be discontinued. The conclusions of the cost-effectiveness analysis identified an optimum risk threshold at which such a decision rule is cost-effective in various situations.

Systematic review of prognostic models: methods

Bibliographic databases (including MEDLINE, EMBASE, The Cochrane Library and clinical trials registers) were searched using index terms relating to the clinical area and prognosis. Title and abstract reviewing was undertaken by two reviewers independently using pre-defined inclusion/exclusion criteria. Two reviewers further assessed eligible full texts using pre-defined criteria. Included full-text articles were data extracted independently by two reviewers, using piloted data extraction forms. Quality assessment and critical appraisal of included full texts was undertaken using an early version of the Prediction study Risk Of Bias Assessment Tool for risk of bias and applicability in prognostic model studies.

Systematic review of prognostic models: results

The systematic review of existing prognostic models for recurrent VTE in the unprovoked population identified three full-text articles, along with seven abstracts related to these three full-text articles and a further two unique abstracts. The three included studies developed three unique prognostic models: HER DOO 2 [Hyperpigmentation, Edema, Redness/D-dimer, Obese (body mass index > 30 kg/m²), Old (age > 65 years)/2 or more factors should indicate for patients to continue therapy], Vienna and D-dimer, Age, Sex, Hormone therapy models. Quality assessment and critical appraisal highlighted several methodological issues with the development of the models (especially for the HER DOO 2 model), and the applicability of the models to the proposed population. For example, varying definitions of unprovoked VTE were used in the data for model development, making the performance of the model unreliable within a new population in which some patients may be classified as unprovoked using a different definition to the development study.

All three models were classed as at least moderate risk of bias, as none had received external validation. This motivated the current research to ensure any new model developed was externally validated. The review also identified potentially important predictors included consistently in the existing models, which could be investigated within the new prognostic model developed.

Prognostic model development: methods

Individual patient data for three databases were acquired from project collaborators, and the Recurrent VTE Collaborative database (containing seven trial populations) was identified as most appropriate for model development, with the remaining databases planned for external validation of the model if possible.

Exploratory and univariable analyses were performed to prepare the data and identify potential predictors of interest, as well as to investigate the effect of predictors of importance indicated within the review. The set of candidate predictors included age, sex, treatment duration, site of index event, D-dimer and lag time. Sample size exceeded 10 events per candidate predictor. Given the available predictors of importance, two potential models were identified based on the point at which the models could be applied. First, a model which could be used at the cessation of therapy to predict recurrence risk (the pre D-dimer model), and second a model which could only be used after a set 'lag time' (the post D-dimer model). The second model could only be used at some point after cessation of therapy because it included D-dimer as a predictor, which is measured at some 'lag time' after stopping therapy because anticoagulation therapy affects its value.

A flexible parametric survival model was used to model the recurrence outcome and to allow investigation of the baseline hazard within trials. A multivariable fractional polynomial algorithm was used for predictor selection, to consider non-linear associations between continuous predictors and recurrence. Differences in the baseline hazard due to the trial populations were accounted for using a random-effects intercept within the model, producing a weighted mean baseline hazard and an estimate of between-study

variability around this mean. Sensitivity analyses investigated the inclusion of interaction effects of interest, time-dependent effects, and the impact of missing data on the analysis using multiple imputation. Model assumptions were checked.

A novel internal–external cross-validation (IECV) approach was used to utilise the seven distinct trial populations. The IECV approach iteratively selects $N-1$ studies from the N total studies available, and the prognostic model is developed within this subset of studies, leaving the remaining study for external validation of the model. A total of N different models are derived (one for each set of included studies, though one larger study was retained in all cycles due to sample size issues) and each is validated in the omitted study. Model performance can then be summarised across the omitted studies using random-effects meta-analysis. Model performance was measured in terms of both calibration and discrimination.

Prognostic model development: results

Predictor selection identified sex and site of index event as important within the pre D-dimer model, while patient age, D-dimer and lag time were additionally included within the post D-dimer model. Model performance through the IECV approach showed that the post D-dimer model had superior performance in terms of discrimination, with the average c -statistic 0.69 in the external validation of this model compared with 0.56 in the external validation of the pre D-dimer model. This suggests that D-dimer and its associated lag time are important and strong predictors, which add significantly to the discriminatory ability of the model.

For both the post D-dimer model and the pre D-dimer models, on average the calibration across all external validation trials was consistently strong with close agreement between observed and predicted risk of recurrence up to at least 2 years. Interrogation of the model fit in regard to multiple imputation of missing data, interaction terms, non-linear trends, outliers and other advanced aspects did not suggest the final models produced should be modified.

Overall the pre D-dimer model was shown to be inadequate in terms of discrimination, which may be expected given that only sex and site of index event were shown to be statistically significant. Conversely the post D-dimer model showed good performance across the validation trials, for both discrimination and calibration, and may be useful in clinical practice to predict individuals' risk and thereby inform treatment decisions, alongside clinical judgement and patient preference.

Systematic review of cost-effectiveness: methods and results

Similar methods as for the systematic review of existing prognostic models were employed in the systematic review of cost-effectiveness studies. Economic models, trial-based economic evaluations and costing studies were eligible for inclusion. Relevant outcomes were cost-effectiveness, cost estimates, resource utilisation estimates and quality of life/utility estimates. Included studies were assessed using relevant economic checklists.

The review did not identify any studies for inclusion, highlighting the current lack of evidence on the cost-effectiveness of using a prognostic model-based decision rule in patients with a first unprovoked VTE. The conclusions of the review therefore indicated that the development of an economic model and cost-effectiveness analysis needed to be undertaken.

Economic modelling: methods

A Markov patient-level simulation was used to consider the economic cost-effectiveness of using a decision rule (based on the prognostic model) to decide on resumption of oral anticoagulant (OAC) therapy (or not). Individual patient characteristics were drawn from distributions based on IPD used in the development of the prognostic model; clinical parameters within the model were obtained from two of the collaborators databases, with remaining parameters informed by the literature, or clinical consensus. Incremental cost-effectiveness ratios were calculated based on an average of the costs and quality-adjusted life-years (QALYs) gained from 50,000 simulated patients.

Economic modelling: results

Results from the economic modelling suggested that a base-case threshold risk of 8% or higher for therapy with warfarin would be cost-effective if decision-makers were willing to pay up to £20,000 per QALY gained, when compared with treat no-one. This indicates that it may be cost-effective to treat all patients with predicted annual risk of recurrence > 8%, and to cease therapy for patients with lower than 8% predicted risk, as opposed to treating no patients.

The model was sensitive to changes in utility and mortality estimates that either solely favoured the no therapy comparator or the decision rule strategy. In order to better assess economic value of such a decision rule further information is required in relation to the long-term bleeding risks on therapy in the unprovoked patient population.

Conclusions

This project has developed a prognostic model which can be used in clinical practice to aid decision-making with regards to the duration of OAC therapy for patients suffering a first unprovoked VTE. The prognostic model was developed using robust methodology and a novel IECV approach allowing external validation in multiple trials. A systematic review of existing prognostic models in this area identified methodological issues to be addressed when developing any new model. In particular, the three existing models had not been externally validated to date, which is an essential step to confirm the performance of the model in a new population. The developed post D-dimer model showed good calibration and discrimination on average across all external validation data sets within the IECV approach. An economic evaluation was undertaken suggesting that a decision rule based on the post D-dimer model would be cost-effective for patients with predicted risk of recurrence of over 8% annually, suggesting continued therapy for risks above 8% and cessation of therapy for risks below 8%. Although the health economic model relies on many assumptions due to lack of routinely collected data, it provides a platform for evaluating further prognostic models once these data are available, particularly in regard to bleeding risks on therapy. This will be useful also for evaluating cost-effectiveness of treatment strategies based on the new generation of OACs. Further work is required to confirm the performance of the model within routine clinical practice (further external validation in non-trials data), and in improving our ability to predict severe bleeding events for patients taking long-term OACs.

Future research recommendations

Further research could build on the pre D-dimer model by seeking additional predictors that may improve discrimination further. Another option is to adapt the model to be used at the exact time of cessation of therapy. This would be beneficial in order to predict patients' recurrence risk at the time of stopping therapy, and thereby negate the need for a lag time period in which the patient is not on therapy (and so is at unnecessarily higher risk of recurrence in this period). Such a model should include predictors for sex and site of index event as these were found to have important associations with recurrence risk in the pre D-dimer model. Future research should aim to incorporate the effect of D-dimer either at the time of cessation of therapy, or on therapy, as D-dimer was shown to be a strong predictor within the post D-dimer model. There is ongoing research investigating the predictive ability of D-dimer levels measured on therapy.

It is also important that further external validation of the post D-dimer model be performed, especially within non-trial populations. Trial populations available within the development database may have been a select group of individuals, and therefore the post D-dimer model requires validation in other populations (e.g. from cohort studies or large databases). Such data sets may not currently be available that contain D-dimer values, and so further observational studies are needed that enrol new patients, measure their predictors following cessation of therapy (including D-dimer and lag time), and record recurrent VTE and adverse outcomes.

Finally there is an essential need for further research to develop and validate a prognostic model for bleeding events on therapy. The current research developed a model which can predict individuals' recurrence risk at some time after cessation of therapy, but this does not account for the subsequent risk of bleeding for patients put on therapy based on their predicted risk. The economic evaluation incorporates the risk of bleeding on therapy, but this could not be individualised as no bleeding event model appropriate to the population exists. There is a need for patient data on both recurrence and subsequent bleeding events which may allow prognostic models to be built and/or validated for both these outcomes simultaneously.

Study registration

This study is registered as PROSPERO CRD42013003494.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.027

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nhredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/hta>

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 10/94/02. The contractual start date was in October 2012. The draft report began editorial review in July 2014 and was accepted for publication in July 2015. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2016. This work was produced by Ensor *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Health Technology Assessment Editor-in-Chief

Professor Hywel Williams Director, HTA Programme, UK and Foundation Professor and Co-Director of the Centre of Evidence-Based Dermatology, University of Nottingham, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Professor Elaine McColl Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research and Development Group, University of Winchester, UK

Professor John Norrie Health Services Research Unit, University of Aberdeen, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk