Development and external validation of risk scores for cardiovascular hospitalisation and rehospitalisation in diabetes patients

Dahai Yu, Yamei Cai, Jonathan Graffy, Daniel Holman, Zhanzheng Zhao, David Simmons

The Journal of Clinical Endocrinology & Metabolism
Endocrine Society

Submitted: October 18, 2017
Accepted: January 03, 2018
First Online: January 08, 2018

The Journal of Clinical Endocrinology & Metabolism; Copyright 2018 DOI: 10.1210/jc.2017-02293

Risk score for CV (re) hospitalisation in diabetes

Development and external validation of risk scores for cardiovascular hospitalisation and rehospitalisation in diabetes patients

Dahai Yu [1,2], Yamei Cai [1], Jonathan Graffy [3], Daniel Holman [4], Zhanzheng Zhao [1*], David Simmons [5*]

1. Department of Nephrology, the First Affiliated Hospital, Zhengzhou University, Zhengzhou 450052, China.

2. Arthritis Research UK Primary Care Centre, Research Institute for Primary Care & Health Sciences, Keele University, Keele ST5 5BG, UK.

3.Primary Care Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, Cambridgeshire CB2 0SR, United Kingdom

4. Department of Sociological Studies, University of Sheffield, Sheffield S10 2TU, United Kingdom

5. Western Sydney University, Campbelltown, Sydney NSW 2751, Australia

Received 18 October 2017. Accepted 03 January 2018.

Context

Cardiovascular disease (CVD) is a common and costly reason for hospitalisation and re-hospitalisation among patients with type 2 diabetes.

Objective

This study aimed to develop and externally validate two risk prediction models for cardiovascular hospitalisation and cardiovascular re-hospitalisation.

Design

Two independent prospective cohorts.

Setting

The derivation cohort includes 4,704 patients with type 2 diabetes from 18 general practices in Cambridgeshire. The validation cohort comprises 1,121 patients with type 2 diabetes from post-trial follow-up data.

Main Outcome Measure

Cardiovascular hospitalisation over 2 years and cardiovascular re-hospitalisation after 90 days of the prior CVD hospitalisation.

Results

The absolute rate of cardiovascular hospitalisation and re-hospitalisation was 12.5% and 6.7% in the derivation cohort, and 16.3% and 7.0% in the validation cohort. Discrimination of the models was similar in both cohorts, with C statistics above 0.70, and excellent calibration of observed and predicted risks.

Conclusion

Two new prediction models that quantify risks of cardiovascular hospitalisation and re-hospitalisation have been developed and externally validated. They are based on a small number of clinical measurements that are available for patients with type 2 diabetes in many developed countries in primary care settings and could serve as the tools to screen the population at high risk of cardiovascular hospitalisation and re-hospitalisation.

2 prediction models based on common clinical measurements have been developed and externally validated. They could aid real time decisions on prioritising more intensive diabetes management.

MAIN TEXT

INTRODUCTION

The prevalence and cost of diabetes is growing rapidly worldwide (1). People with diabetes are twice as likely to be admitted to hospital, and at least 10% of those in hospital have diabetes at any one time (2). In some age groups, it is as many as one in five (3). The associated costs of excess admissions, as well as increased costs per admission, are significant contributors to the financial burden borne by healthcare systems from diabetes and often reflect preventable morbidity suffered by patients (4).

Previously, two prediction tools have been developed, both based on secondary care data, to identify those with diabetes, at high risk of either all-cause excessive length of stay or all-cause inpatient mortality over four years (5), or all-cause re-admission within 30 days among hospitalised patients (6). However, the practical application of both prediction models was limited by lack of external validation, non-specificity for people with type 2 diabetes, the use of predictors derived from secondary care rather than primary care data, variations on predictors recorded in different datasets (e.g. comorbidity) and a relative short time-gap between baseline and outcome (30 days' readmission).

Among hospital admissions, cardiovascular events are the major cause for hospitalisation in people with type 2 diabetes (7). Although risk factors such as blood pressure and HbA1c are recognised as warranting intervention on their own (8), (9), there has been no current algorithm to estimate the absolute risk of cardiovascular hospitalisation and rehospitalisation in people with type 2 diabetes.

Using a model to make predictions for individual patients with type 2 diabetes is more comprehensive than using individual risk factors, and is preferred to the risk grouping approach (10), (11).

The aim of our study was to develop and externally validate new prediction models based on reliable clinical measurements in primary care settings for cardiovascular hospitalisation over the next 2 years and cardiovascular re-hospitalisation up to 90 days following a prior cardiovascular hospitalisation.

MATERIALS AND METHODS

Data source and study population

We utilised two cohorts from Cambridgeshire, England: one (Derivation) based on the electronic health record data from primary care settings to develop our cardiovascular hospitalisation and re-hospitalisation risk scores and another (Validation) based on post-trial cohort data for external validation.

Derivation cohort

Patient lists from 18 general practices across Cambridgeshire, England, in 2008/2009 were collated and linked with hospital admissions (Secondary Uses Service (SUS)) data as part of an evaluation of diabetes care across the county by the local health board, National Health Service (NHS) Cambridgeshire. This cohort was limited to volunteer practices using the Egton Medical Information Systems (EMIS) general practitioner (GP) software system, from which a predefined set of data could be extracted. There was no systematic selection process for these surgeries, and data extracted were for their entire diabetes population. All patients with diabetes had follow-up hospitalisation data to 2010–2011. Hospital admissions to NHS and private hospitals within and outside Cambridgeshire were followed-up. No personal identifiers were released to researchers, and all subsequent analyses were conducted on anonymised datasets.

Validation cohort

The design and methods of the RAPSID trial have been published previously (12), as have its CONSORT (Consolidated Standards of Reporting Trials) diagram and the results of its primary outcomes (12). Briefly, RAPSID was a 2x2 factorial cluster RCT comparing 4 groups: Controls, 1:1 (individual) peer support, group peer support, and combined 1:1 and group peer support among patients with type 2 diabetes. Participants had their diabetes for at least 12 months and those with dementia or psychotic illness were excluded. Participants were recruited from communities across Cambridgeshire and neighbouring areas of Essex and Hertfordshire. Follow up data were only available for participants in Cambridgeshire and neighbouring areas of Hertfordshire that are served by the Cambridgeshire and Peterborough Clinical Commissioning Group (CCG). Clusters were defined by local government ('parish council') boundaries. The intervention was developed following a pilot (13), using a framework defined by Peers for Progress (14). Peers facilitating peer support were termed peer support facilitators and there selection, training, support and the overall programme are described elsewhere (15). The intervention lasted 8-12 months and was commenced and concluded, cluster by cluster, between 02/06/11 to 12/04/12. Ethics approval was received from the Cambridgeshire REC2 Committee (10/H0308/72), and signed consent included agreement for access to hospital data.

At baseline, demographic data, blood pressure, and HbA1c and lipid profiles information were collected. Each participant was followed up until June 2015 (0.91-4.07 years' follow-up from beginning/entry into the trial). Hospitalisation (NHS hospitals & private hospitals), Accident & Emergency (A&E) and outpatient visits within/outside Cambridgeshire and the included areas of Hertfordshire were completely collected through Cambridgeshire and Peterborough Clinical CCG (16) and the elective/non-elective status, and International Classification of Diseases (ICD-10) codes (8).

Defining cardiovascular hospitalisation and re-hospitalisation

The primary outcome of the study was having at least one hospitalisation with cardiovascular disease (CVD) as the primary diagnosis (ICD-10: I20–I25, I60–I69 and I73 in the first ICD field) over the 2-year follow-up and having at least one CVD re-hospitalisation after 90 days of prior CVD hospitalisation.

Candidate predictors, missing data, and power calculations

To achieve the maximum extrapolation application of our risk algorithm, objective clinical measurements were used as predictors in the model, including body mass index (BMI) , blood pressure (systolic (SBP) and diastolic (DBP)) and the metabolic variables glycated haemoglobin (HbA1c) and lipid profiles. We also included demographic characteristics, (age and gender) and whether the patient was on lipid lowering treatment. Patients with diabetes were invited to have their blood pressure and metabolic variables measured at least once a year after the diagnosis of diabetes and the most recent was taken before 1 April 2009 (a minimum of 50 days before the first admission). Diabetes duration was not universally recorded, and hence was not usefully available for analysis. Diabetes therapy was not included in the dataset. Lipid-lowering treatment was recorded.

Our derivation cohort had missing information on body mass index (3.17%), systolic blood pressure (9.95%), diastolic blood pressure (9.95%), total cholesterol (12.35%), high density lipoprotein (14.56%), and low density lipoprotein (16.27%). We used multiple imputation to replace missing values by using a chained equation approach based on all candidate predictors and outcomes. We created 16 imputed datasets for missing variables that were then combined across all datasets by using Rubin's rule to obtain final model estimates. Limited information

was missing (<1%) in our external validation dataset and the complete dataset was used in our analysis. On the basis of an estimated 588 cardiovascular hospitalisations and 316 cardiovascular re-hospitalisations and 16 predictors or levels in our derivation cohort, we had an effective sample size of 37 cardiovascular hospitalisation and 21 cardiovascular re-hospitalisation per predictor or level, above the minimum requirement suggested by Peduzzi et al (17).

Ethical approval

The derivation cohort work had approval from the Cambridgeshire research ethics committee as part of a wider service evaluation. Ethics approval for validation cohort was received from the Cambridgeshire REC2 Committee (10/H0308/72), and signed consent included agreement for access to hospital data.

Statistical analysis for model derivation and external validation

We treated incidence occurrence of cardiovascular hospitalization after the first 90 days since the start of follow-up and the incident occurrence of cardiovascular re-hospitalisation as binary outcome measures. For each of the 15 candidate predictors or levels, we used a univariate logistic regression model to calculate the unadjusted odds ratios. For derivation of the risk prediction model, we initially included all candidate predictors in a multivariable logistic regression model. We used fractional polynomials to model potential non-linear relationships between continuous predictors and outcome.

Through backward elimination, we excluded lower lipid treatment from the multivariate model as it was not statistically significant (P>0.1 based on change in log likelihood). After elimination, we reinserted the excluded predictor into the final model to further check whether it became statistically significant. We also rechecked fractional polynomial terms at this stage and re-estimated them if necessary. We formed the risk equations for predicting the log odds of cardiovascular hospitalisation and cardiovascular re-hospitalisation by using the estimated regression coefficients multiplied by the corresponding predictors included in our models together with the intercepts. This process ultimately led to equations for the predicted risk=$1/(1+e^{-riskscore})$, whether the "risk score" is the predicted log odds of cardiovascular hospitalisation or cardiovascular re-hospitalisation from the developed models.

To facilitate model utilisation in clinical practice, the logistic regression equations were transformed into prognostic score charts. The coefficients in the logistic regression equation were multiplied by 50 and rounded to the nearest integer to obtain the prognostic score per predictor. Multiplication by 50 was chosen to get the majority of the coefficients close to an integer, thereby minimizing the effects of rounding. The sum of all prognostic scores reflects patients' probability of cardiovascular hospitalisation or cardiovascular re-hospitalisation.

We assessed the performance of the models in terms of the C statistics and calibration slope (where 1.00 is ideal). The C statistics represents the probability that for any randomly selected pair of people with type 2 diabetes with and without outcomes, the patient with outcomes had a higher predicted risk (18). A value of o.50 indicated no discrimination and 1.00 represents perfect discrimination. We then undertook internal validation to correct measures of predictive performance for optimism (over-fitting) by bootstrapping 100 samples of the derivation data. We repeated the model derivation process in each bootstrap sample to produce a model, applied the model to the same bootstrap sample to quantify apparent performance, and applied the model to the original dataset to test model performance (calibration slope and C-statistics) and optimism (difference in the test performance and apparent performance). We then estimated the overall optimism across all models.

We applied our risk prediction model to each patient with type 2 diabetes in the external validation cohort on the basis of the presence of one or more predictors. We examined the performance of this final model both in the derivation dataset and then in external validation dataset in terms of discrimination by calculating the C statistics. We examined calibration by plotting agreement between predicted and observed risks across tenth of the predicted risks.

We used Stata V14.0 for all statistical analyses. This study was conducted and reported in line with the Transparent Reporting of a multivariate prediction model for Individual Prediction Diagnosis (TRIPOD) guidelines (19).

Role of the funding source
The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

RESULTS

Study participants

In our derivation cohort, we analysed information on 4,704 type 2 diabetes patients with 588 cardiovascular hospitalisations within 2 years and 316 re-hospitalisations after 90 days since a prior cardiovascular hospitalisation. Our validated cohort had information on 1,121 type 2 diabetes patients with 183 cardiovascular hospitalisations and 78 re-hospitalisations. Table-1 summarises the basic characteristics and potential predictors of the study population. Patients with type 2 diabetes in both cohorts had similar age, gender, blood pressure and total cholesterol. Patients in the derived cohort had a higher level of high density lipoprotein, low density lipoprotein, and HbA1c. Compared with the derivation cohort, those in the validation cohort were more likely to be prescribed lowering lipid medicine and had more cardiovascular hospitalisation and re-hospitalisation.

Model derivation, performance measure, and validation

In the derivation dataset, the absolute risks of cardiovascular hospitalisation within 2 years and re-hospitalisation within 9o days post cardiovascular hospitalisation were 12.5% and 6.7%, respectively. Univariable associations between cardiovascular hospitalisation and cardiovascular re-hospitalisation are listed in supplemental Table-1. Of the 10 candidate predictors (16 categories), 9 predictors (15 categories) were statistically significantly associated with cardiovascular hospitalisation and re-hospitalisation in the final multivariable model (**Table-2**). Table-2 shows apparent and internal validation performance statistics of the risk prediction model. After adjustment for optimism, the final risk prediction model was able to discriminate type 2 diabetes patients with and without cardiovascular hospitalisation with a C statistics of 0.7094 (95% confidence interval 0.7067 to 0.7205), and discriminate type 2 diabetes patients with and without cardiovascular re-hospitalisation with a C statistics 0.7118 (0.7077 to 0.7159). The agreement between the observed and predicted proportion of cardiovascular hospitalisation and re-hospitalisation showed good apparent calibration (**Figure-1**, top left for cardiovascular hospitalisation and top right for cardiovascular re-hospitalisation). The optimism adjusted calibration slope was 1.0301 (0.9856 to 1.0747) and 1.0001 (0.9711 to 1.0247) for cardiovascular hospitalisation and re-hospitalisation, respectively (**Table-3**).

External validation

In the external validation cohort, the absolute risks for cardiovascular hospitalisation and re-hospitalisation were 16.3% and 7.0%, respectively. Applying our final risk prediction model to the independent population gave a C statistic of 0.7092 (0.7033 to 0.7151) for cardiovascular

hospitalisation and 0.7098 (0.7014 to 0.7182) for cardiovascular re-hospitalisation, and good calibration (Figure-1, bottom left for cardiovascular hospitalisation and bottom right for cardiovascular re-hospitalisation), with the calibration slope 1.0001 (0.9807 to 1.0195) and 0.9981 (0.9948 to 1.0482) for cardiovascular hospitalisation and re-hospitalisation, respectively.

Performance at the threshold for 10% and 20% of patients at highest risk

Table-4 shows the sensitivity, specificity, and observed risk for the 5%, 10%, 15%, 20% and 25% of patients at the highest predicted risk of each outcome in the validation cohort shown for illustrative purposes. For example, when a risk threshold of 24.53% for cardiovascular hospitalisation and 7.93% for cardiovascular re-hospitalisation is used to identify the 20% at highest predicted risk, the sensitivity was 33.40% for cardiovascular hospitalisation and 45.20% for cardiovascular re-hospitalisation, the specificity was 84.60% for cardiovascular hospitalisation and 75.90% for cardiovascular rehospitalisation, and the observed risk was 30.09% for cardiovascular hospitalisation and 11.98% for cardiovascular re-hospitalisation, respectively.

Clinical examples

Supplemental Chart-1 gives a clinical example of the application of prognostic score charts with graphical illustrations for cardiovascular hospitalisation and re-hospitalisation risk prediction models to predict 2-year risk of cardiovascular hospitalisation and risk of re-hospitalisation within 90 days of a prior cardiovascular hospitalisation.

DISCUSSION

We have developed two new risk prediction models to estimate the absolute risk of cardiovascular hospitalisation within 2 years and cardiovascular re-hospitalisation after 90 days of prior cardiovascular hospitalisation in a cohort of patients with type 2 diabetes in England. We then externally validated this model in another English cohort. The two prediction models had excellent calibration and useful discrimination, with C statistics of greater than 0.70 both in the derivation cohort and external validation cohort. The two prediction models were built from clinical variables usually recorded and accessible in primary care settings, implying that they can be readily applied in routine primary care.

Strengths and limitations

Our two risk algorithms have several advantages over those in utilisation in many developed countries. Our models are based on absolute risks determined and validated in two independent populations. The models are developed from routinely recorded demographic and clinical measurements in primary care settings, which suggests that they can be straightforwardly applied in general practice and are readily amenable for further external validations in countries that have routine recorded data accessible for such aims. And the two risk algorithms can be easily integrated into online calculators for implementation in general practices.

The methods used to derive and validate the model are similar to those for other risk prediction algorithms derived from the CPRD and QResearch databases (20), (21). The majority of predictors in our final model are accurate and reliable clinical measurements (22) routinely recorded in primary care settings and updated and reviewed for patients with type 2 diabetes, and are less varied than in other datasets. Moreover, the proportion of missing values was low, which would lead to little variation in external applications, although multiple imputation was still applied in our study. We acknowledge that our prediction models do not take into account diabetes duration, antidiabetes treatments, anti-hypertensive treatments, prior history of

cardiovascular diseases, other diabetes complications (e.g. renal failure), lifestyle risk factors (like smoking), and other comorbidities due to limitations in the original data, but we feel that the clinical measurements included in our models could be proxies for missing predictors. Data limitations also prevented extending our model to all diabetes complications rather than those relating to cardiovascular hospitalisation. The relatively low sensitivities of our models to identify individuals at high risk of cardiovascular hospitalisation and re-hospitalisation is another limitation of the study. Due to the similarity between the derivation and validation cohorts, further external validation (e.g. cohorts from other countries) are warranted.

Comparison with other studies

Nirantharakumar et al. developed a prediction model among patients with diabetes to estimate adverse events (either excessive length of stay or inpatient mortality) over 4 years using a secondary care dataset in Birmingham, England (5). The predictors applied in this model covered demographic characteristics, clinical pathological test results, and use of insulin, recorded within 72 hours of hospitalisation. That population represented the people with at least previous inpatient hospitalisation, and probably reflects a cohort with more severe conditions, and likely higher prior probabilities of an event. The ranges of clinical measurements during a hospital admission would tend to be greater than in the community, as patients would be sicker and e.g. blood glucose control could be the reason for hospitalisation, or exacerbated by acute illness, making the dataset difficult to use as a basis for a prediction tool in routine care. Most importantly, this prediction model has not been externally validated and the model performance needs to be further evaluated in external populations before its application in clinical practices.

Rubin et al developed a tool to predict the risk of all-cause re-admission within 30 days among hospitalised patients with diabetes using hospitalised data (6). The short time-gap between predictor measurements and outcome made the tool less useful for clinical practice. The reasons for hospitalisation could be quite mixed, with different pathway and potential interventions. Therefore, using the all-cause hospitalisation risk as the outcome provides different information and allows less targeted interventions. As with Nirantharakumar et al's model (5), this model has also not been externally validated in any independent population.

Previous studies have not focussed on cardiovascular disease as both a major cause and cost for hospital admission among patients with diabetes. To understand the potential risk of cardiovascular hospitalisation in the next year, and the risk of a new episode (within 90 days) of a cardiovascular event (re-hospitalisation) could be helpful for clinicians to facilitate tailored, more intensive care to those with high risk profiles and to reduce hospitalisation inpatient cost.

Conclusion and policy implication

As far as we are aware, our study is the first study to develop prediction tools to estimate the 2-year risk of cardiovascular hospitalisation and re-hospitalisation within 90 days of a previous hospitalisation. Our two prediction models have two important implications for clinical practice. First, they can be used as tools to screen populations at high risk of cardiovascular hospitalisation and re-hospitalisation. Both algorithms are based on readily accessible clinical data routinely recorded in primary care and reviewed by diabetes management teams. They can be readily integrated into primary care computer systems or developed into an app for a handheld device for ease of use. Secondly, our risk prediction models could be used to establish new treatment thresholds in clinical practice through consensus development of national guidelines.

ACKNOWLEDGMENTS
We would like to thank Toby Prevost, Chris Bunn, Simon Cohn, Sarah Donald, Charlotte Paddison, Candice Ward, Peers for Progress, West Anglia CLRN, Cambridgeshire & Peterborough PCT, Primary Care Research Network – East of England, Eastern Diabetes Research Network, MRC Epidemiology Unit, participating general practices, Jackie Williams, Caroline Taylor, Kym Mercer, Kevin Baker, Ben Bowers, Kalsoom Akhter (CUH Wolfson Diabetes & Endocrinology Clinic), James Brimicombe (Cambridge University), Kim Birch of Trumpington St General Practice, CUH Wolfson Diabetes & Endocrinology Clinic Educators, The RAPSID Patient Committee (Phillip Jones, Liz Carvlin, Roger Smith) and the peers and peer support participants.

***Correspondence 1:** Professor David Simmons, Macarthur Clinical School, School of Medicine, Western Sydney University, Locked Bag 1797, Campbelltown NSW 2751, AUSTRALIA, Email: dsworkster@gmail.com, TEL: (61+2) 4620 3899, FAX: (61+2) 4620 3890

***Correspondence 2:** Professor Zhanzheng Zhao, Department of Nephrology, the First Affiliated Hospital, Zhengzhou University, Zhengzhou 450052, CHINA, Email: zhanzhengzhao@zzu.edu.cn, TEL: +86 139 3852 5666, FAX: +86 371 6698 8753

Address all correspondence and requests for reprints to: Professor David Simmons, Macarthur Clinical School, School of Medicine, Western Sydney University, Locked Bag 1797, Campbelltown NSW 2751, AUSTRALIA. Email: dsworkster@gmail.com

This work was supported by the National Institute for Health Research under its Research for Patient Benefit (RfPB) Programme (Grant Reference Number PB-PG-0808-17303) and National Institute for Health Research for Patient Benefit Programme (Ref PB-PG-0610-22311).

Funding: This work was supported by the National Institute for Health Research under its Research for Patient Benefit (RfPB) Programme (Grant Reference Number PB-PG-0808-17303) and National Institute for Health Research for Patient Benefit Programme (Ref PB-PG-0610-22311). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Research for Patient Benefit Programme http://dx.doi.org/10.13039/501100009128, PB-PG-0610-22311, David Simmons; Research for Patient Benefit Programme http://dx.doi.org/10.13039/501100009128, PB-PG-0808-17303, David Simmons
Disclosure Summary: The authors have nothing to disclose.

Disclosure Summary: The authors have nothing to disclose. There are no potential conflicts of interest relevant to this article to report.

References

1. **American Diabetes Association** 2012 Standards of medical care in diabetes--2012. Diabetes Care 35 Suppl 1:S11-63
2. **Sampson MJ, Dozio N, Ferguson B, Dhatariya K** 2007 Total and excess bed occupancy by age, specialty and insulin use for nearly one million diabetes patients discharged from all english acute hospitals. Diabetes Res Clin Pract 77(1):92-98

3. **Simmons D, English P, Robins P, Craig A, Addicott R** 2011 Should diabetes be commissioned through multidisciplinary networks, rather than practice based commissioning? Prim Care Diabetes 5(1):39-44
4. **Vamos EP, Millett C, Parsons C, Aylin P, Majeed A, Bottle A** 2012 Nationwide study on trends in hospital admissions for major cardiovascular events and procedures among people with and without diabetes in england, 2004-2009. Diabetes Care 35(2):265-272
5. **Nirantharakumar K, Hemming K, Narendran P, Marshall T, Coleman JJ** 2013 A prediction model for adverse outcome in hospitalized patients with diabetes. Diabetes Care 36(11):3566-3572
6. **Rubin DJ, Handorf EA, Golden SH, Nelson DB, McDonnell ME, Zhao H** 2016 Development and validation of a novel tool to predict hospital readmission risk among patients with diabetes. Endocr Pract 22(10):1204-1215
7. **Khalid JM, Raluy-Callado M, Curtis BH, Boye KS, Maguire A, Reaney M** 2014 Rates and risk of hospitalisation among patients with type 2 diabetes: Retrospective cohort study using the UK general practice research database linked to english hospital episode statistics. Int J Clin Pract 68(1):40-48
8. **Yu D, Simmons D** 2014 Association between blood pressure and risk of cardiovascular hospital admissions among people with type 2 diabetes. Heart 100(18):1444-1449
9. **Yu D, Simmons D** 2013 Relationship between HbA1c and risk of all-cause hospital admissions among people with type 2 diabetes. Diabet Med 30(12):1407-1411
10. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, Riley RD, Hemingway H, Altman DG, PROGRESS Group 2013 Prognosis research strategy (PROGRESS) 3: Prognostic model research. PLoS Med 10(2):e1001381
11. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, Briggs A, Udumyan R, Moons KG, Steyerberg EW, Roberts I, Schroter S, Altman DG, Riley RD, PROGRESS Group 2013 Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. BMJ 346:e5595
12. **Simmons D, Prevost AT, Bunn C, Holman D, Parker RA, Cohn S, Donald S, Paddison CA, Ward C, Robins P, Graffy J** 2015 Impact of community based peer support in type 2 diabetes: A cluster randomised controlled trial of individual and/or group approaches. PLoS One 10(3):e0120277
13. **Simmons D, Cohn S, Bunn C, Birch K, Donald S, Paddison C, Ward C, Robins P, Prevost AT, Graffy J** 2013 Testing a peer support intervention for people with type 2 diabetes: A pilot for a randomised controlled trial. BMC Fam Pract 14:5-2296-14-5
14. Fisher EB, Ayala GX, Ibarra L, Cherrington AL, Elder JP, Tang TS, Heisler M, Safford MM, Simmons D, Peers for Progress Investigator Group 2015 Contributions of peer support to health, health care, and prevention: Papers from peers for progress. Ann Fam Med 13 Suppl 1:S2-8
15. **Dale JR, Williams SM, Bowyer V** 2012 What is the effect of peer support on diabetes outcomes in adults? A systematic review. Diabet Med 29(11):1361-1377
16. **Simmons D, Yu D, Wenzel H** 2014 Changes in hospital admissions and inpatient tariff associated with a diabetes integrated care initiative: Preliminary findings. J Diabetes 6(1):81-89
17. **Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR** 1996 A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49(12):1373-1379

18. **Yu D, Simmons D** 2015 Association between pulse pressure and risk of hospital admissions for cardiovascular events among people with type 2 diabetes: A population-based case-control study. Diabet Med 32(9):1201-1206
19. **Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS** 2015 Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med 162(1):W1-73
20. **Sultan AA, West J, Grainge MJ, Riley RD, Tata LJ, Stephansson O, Fleming KM, Nelson-Piercy C, Ludvigsson JF** 2016 Development and validation of risk prediction model for venous thromboembolism in postpartum women: Multinational cohort study. BMJ 355:i6253
21. **Hippisley-Cox J, Coupland C** 2012 Derivation and validation of updated QFracture algorithm to predict risk of osteoporotic fracture in primary care in the united kingdom: Prospective open cohort study. BMJ 344:e3427
22. **Hemke AC, Heemskerk MB, van Diepen M, Dekker FW, Hoitsma AJ** 2015 Improved mortality prediction in dialysis patients using specific clinical and laboratory data. Am J Nephrol 42(2):158-167

Figure-1. Assessing calibration in the derivation cohort (left) and the validation cohort (right) for cardiovascular hospitalisation (above panel) and cardiovascular re-hospitalisation (below panel)

Table-1. Baseline Characteristics of study populations.

	Derivation cohort	External validation cohort
N	4,704	1,121
Cardiovascular hospitalisation, n (%)	588 (12.5)	183 (16.3)
Cardiovascular rehospitalisation, n (%)	316 (6.7)	78 (7.0)
Age, years	65.0±16.3	65.5±11.4
Female, n (%)	1,919 (40.8)	444 (39.6)
Systolic blood pressure, mmHg	134.5±16.0	139.7±20.2
Diastolic blood pressure, mmHg	76.3±10.0	75.5±11.5
Total cholesterol, mmol/L	4.3±1.2	4.2±1.7
High density lipoprotein, mmol/L	1.3±0.6	1.1±1.2
Low density lipoprotein, mmol/L	2.5±1.4	1.4±3.0
Body mass index, kg/m^2	30.8±6.9	32.2±6.0
HbA1c, mmol/mol	61.5±17.2	56.2±15.1
Lipid Lowering treatment, n (%)	3,342 (71.4)	731 (65.2)

Table-2. Final multivariate analysis for cardiovascular hospitalisation and re-hospitalisation risk among people with type 2 diabetes in derivation cohort

Predictors	Coefficient	95% Confidence Interval
Cardiovascular Hospitalisation		
Age ≥ 70 years	0.815914	(0.793045 to 0.838784)
Male gender	0.228943	(0.206719 to 0.251168)
HbA1c ≥ 57 mmol/mol (7.4%)	-0.03967	(-0.06088 to -0.01846)
(Body mass index/10)^-2	-1.85384	(-2.39533 to -1.31235)
(Body mass index/10)^0.5	0.690585	(0.551284 to 0.829887)
(Systolic blood pressure/100)^2	-0.40302	(-0.58492 to -0.22111)
(Systolic blood pressure/100)^2*ln(Systolic blood pressure/100)	0.966205	(0.758028 to 1.174381)
(Diastolic blood pressure/100)^-2	0.474014	(0.387498 to 0.56053)
(Diastolic blood pressure/100)^-2*ln(Diastolic blood pressure/100)	0.2724	(0.188226 to 0.356575)
ln(Total cholesterol/10)	0.514695	(0.27381 to 0.75558)
(Total cholesterol/10)^0.5	-1.05803	(-1.86382 to -0.25223)
ln(High density lipoprotein)	0.073489	(0.04377 to 0.103208)

(High density lipoprotein)^3	-0.02384	(-0.02699 to -0.02069)
(Low density lipoprotein/10)^0.5	-0.55634	(-0.67239 to -0.44028)
ln(Low density lipoprotein/10)* (Low density lipoprotein/10)^0.5	-0.83161	(-1.01001 to -0.65322)
Constant	-3.80246	(-4.67529 to -2.92963)
Cardiovascular Re-hospitalisation		
Age ≥ 70 years	0.90054	(0.86384 to 0.93724)
Male	0.22328	(0.188299 to 0.258261)
HbA1c ≥ 57 mmol/mol (7.4%)	0.004076	(-0.0294 to 0.037547)
(Body mass index/10)^-2	-4.17347	(-4.62492 to -3.72202)
(Body mass index/10)^3	0.001821	(0.001318 to 0.002324)
(Systolic blood pressure/100)^2	-1.16118	(-1.46728 to -0.85507)
(Systolic blood pressure/100)^3	0.773551	(0.637616 to 0.909486)
(Diastolic blood pressure/100)^-2	0.5875	(0.439237 to 0.735763)
(Diastolic blood pressure/100)^-2*ln(Diastolic blood pressure/100)	0.4095	(0.260667 to 0.558332)
(Total cholesterol/10)^-2	-0.00798	(-0.01031 to -0.00565)
(Total cholesterol/10)^2	-0.02734	(-0.23117 to 0.176482)
ln(High density lipoprotein/10)	0.051443	(0.004285 to 0.0986)
(High density lipoprotein/10)^3	-0.02718	(-0.03277 to -0.02159)
Low density lipoprotein/10	-1.34491	(-1.56307 to -1.12675)
ln(Low density lipoprotein/10)	-0.88347	(-1.28497 to -0.48196)
Constant	-4.55873	(-4.8866 to -4.23086)

Table-3. Model diagnostics (with 95% CI)

	Derivation				Validation
Measure	Apparent performance	Test performance	Average optimism	Optimism corrected	
	Cardiovascular Hospitalisation				
C statistic	0.7163 (0.7136 to 0.7190)	0.7027 (0.6996 to 0.7058)	+0.0069	0.7094 (0.7067 to 0.7205)	0.7092 (0.7033 to 0.7151)
Calibration slope	1.0000 (0.9806 to 1.0194)	0.9933 (0.9899 to 0.9966)	+0.0067	0.9933 (0.9739 to 1.0127)	1.0001 (0.9807 to 1.0195)
	Cardiovascular Re-hospitalisation				
C statistic	0.7154 (0.7113 to 0.7195)	0.7136 (0.7105 to 0.7167)	+0.0036	0.7118 (0.7077 to 0.7159)	0.7098 (0.7014 to 0.7182)
Calibration slope	1.0000 (0.9766 to 1.0234)	0.9976 (0.9949 to 1.0003)	+0.0024	0.9976 (0.9742 to 0.9796)	0.9981 (0.9948 to 1.0482)

Table-4. Predicted risk of cardiovascular hospitalisation and re-hospitalisation the validation cohort based on various cut-offs.

	Cut-off (%) for risk	Mean predicted risk (%)	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Observed risk %
Cardiovascular hospitalisation						
Top 5%	38.17	51.96	10.30 (9.70 to 10.90)	97.40 (97.20 to 97.50)	43.50 (41.50 to 45.50)	43.48
Top 10%	31.73	43.35	17.50 (16.80 to 18.30)	94.60 (94.40 to 94.80)	38.60 (37.20 to 40.10)	38.62
Top 15%	27.54	37.71	24.70 (23.90 to 25.60)	90.10 (89.80 to 90.40)	32.80 (31.80 to 33.90)	32.83
Top 20%	24.53	33.77	34.00 (33.10 to 35.00)	84.60 (84.20 to 84.90)	30.10 (29.20 to 31.00)	30.09
Top 25%	22.22	31.05	42.80 (41.80 to 43.80)	78.40 (78.00 to 78.70)	27.90 (27.20 to 28.60)	27.89
Cardiovascular re-hospitalisation						
Top 5%	11.34	15.86	26.20 (24.90 to 27.50)	91.20 (91.00 to 91.50)	18.30 (17.40 to 19.30)	18.33
Top 10%	9.67	13.63	34.50 (33.10 to 36.00)	84.30 (84.00 to 84.60)	14.20 (13.50 to 14.90)	14.22
Top 15%	8.69	12.59	40.50 (39.00 to	79.10 (78.80 to	12.70 (12.20 to 13.30)	12.73

			42.00)	79.50)		
Top 20%	7.93	12.02	45.20 (43.70 to 46.70)	75.90 (75.50 to 76.30)	12.40 (11.90 to 12.90)	12.37
Top 25%	7.16	11.46	50.00 (48.50 to 51.50)	72.40 (72.00 to 72.70)	12.00 (11.50 to 12.50)	11.98