This is the preprint of the following article: Stiglic G, Wang F, Sheikh A, Cilar L. Development and validation of the type 2 diabetes mellitus 10-year risk score prediction models from survey data. Primary Care Diabetes. 2021, which has been published in final form at https://doi.org/10.1016/j.pcd.2021.04.008.

Development and validation of the type 2 diabetes mellitus 10-year risk score prediction models from survey data

Gregor Stiglic*[,1,2], Fei Wang[3], Aziz Sheikh[4], Leona Cilar[1]

[1]University of Maribor, Faculty of Health Sciences, Zitna ulica 15, 2000 Maribor, Slovenia

[2]University of Maribor, Faculty of Electrical Engineering and Computer Science, Koroska cesta 46, 2000 Maribor, Slovenia

[3]Weill Cornell Medicine Graduate School of Medical Sciences, 1300 York Ave, New York, NY 10065, United States

[4]Usher Institute, University of Edinburgh, Old Medical School, Teviot Place, Edinburgh EH8 9AG, UK

*Corresponding author: gregor.stiglic@um.si, tel.: +386 2 300 4731, Zitna ulica 15, 2000 Maribor, Slovenia.

Funding

This work was supported by the Slovenian Research Agency [grant numbers ARRS-N2-0101, ARRS-BI-US/18-19-064, ARRS-P2-0057].

Author Contributions

G.S., F.W., A.Z. and L.C. designed the study. G.S. performed the data preparation and analysis. G.S., F.W., A.Z. and L.C. interpreted the results, wrote and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Data Availability

The data that support the findings of this study are available from SHARE Research Data Center (http://www.share-project.org). Some restrictions apply to the availability of these data as the user needs to register in order to obtain the access to the data. After the registration the data are freely available for research purposes.

Abstract

Aims: In this paper, we demonstrate the development and validation of the 10-years type 2 diabetes mellitus (T2DM) risk prediction models based on large survey data.

Methods: The Survey of Health, Ageing and Retirement in Europe (SHARE) data collected in 12 European countries using 53 variables representing behavioural as well as physical and mental health characteristics of the participants aged 50 or older was used to build and validate prediction models. To account for strongly unbalanced outcome variables, each instance was assigned a weight according to the inverse proportion of the outcome label when the regularized logistic regression model was built.

Results: A pooled sample of 16,363 individuals was used to build and validate a global regularized logistic regression model that achieved an area under the receiver operating characteristic curve of 0.702 (95% CI: 0.698 – 0.706). Additionally, we measured performance of local country-specific models where AUROC ranged from 0.578 (0.565 - 0.592) to 0.768 (0.749 - 0.787).

Conclusions: We have developed and validated a survey-based 10-year T2DM risk prediction model for use across 12 European countries. Our results demonstrate the importance of re-calibration of the models as well as strengths of pooling the data from multiple countries to reduce the variance and consequently increase the precision of the results.

Keywords: type 2 diabetes mellitus, 10-years risk, prediction models, model calibration

Introduction

The number of people with diabetes mellitus is increasing rapidly. Nowadays, about 1 in 11 adults have diabetes mellitus and about 90% of them have type 2 diabetes mellitus (T2DM) [1]. T2DM is a leading cause of mortality and morbidity worldwide and presents a major economic burden to health systems and society [2]. Its prevalence is rapidly increasing, and it is predicted that T2DM will increase by 48% and will affect about 629 million people worldwide by 2045 [3]. Another risk factor for the development of functional decline and disability besides T2DM is ageing [4, 5]. With the increase of the population age worldwide, there are also more and more people living with T2DM. At the start of 2019 there were 101.1 million of people older than 65 years in the European Union (EU). This number is expected to rise at 149.2 million by 2050 [6]. With older age, changes such as memory problems, poorer functionality, visual and hearing impairment, and reductions in manual dexterity can make management of the T2DM complex and challenging [7]. Because of non-specific symptoms, T2DM is often diagnosed when complications appear. Older people with T2DM are more likely to be hospitalized than those without T2DM. Moreover, T2DM is significantly correlated to mortality in terms of socioeconomic status [8]. In a study conducted among 71,483 Swedish adults, the authors found that T2DM is a strong risk factor for myocardial infarction, ischemic stroke, aortic valve stenosis, and heart failure as well as blindness, chronical kidney disease or amputations. The prevention activities and interventions are needed to be implemented among those who are at high risk of developing T2DM [9]. Thus, populations younger than 65 years must be included in the research.

The Survey of Health, Ageing and Retirement in Europe (SHARE) is a multidisciplinary and cross-national database of data on health of about 140,000 individuals aged 50 or older [10]. It is widely used due to easy access to data from different European Union (EU) countries and the wide range of included explanatory variables. The SHARE database allows researchers to use a wide range of variables that may have impact on the onset of self-reported T2DM.

Prevention activities focusing on lifestyle changes can delay the onset of the T2DM in individuals with pre-diabetes [11]. Recent studies showed that the healthy diet habits such as consummation of whole grains, fruits and low-fat dairy products can lower the possibility of developing T2DM. On the other hand, high intakes of red meat, processed meat and sweetened beverages consumption are the risk factors for developing T2DM [12-16]. Another protective factor in adults is regular physical activity [17], specifically participation in leisure-time running or moderate physical activity [18, 19]. Exercise is also effective at improving blood glucose in individuals with established T2DM [20, 21]. American Diabetes Association [22] recommended at least annual monitoring of people with prediabetes and their inclusion in intensive behavioural lifestyle intervention, and usage of technology-assisted tools to prevent diabetes and promote healthy lifestyle. Moreover, in older populations who already have T2DM it is important to consider medical, psychological, functional, and social domains to determine targets and therapeutic approaches for diabetes management.

Predictive models have demonstrated their usability in exploring the risk factors for developing T2DM for many years [23-30]. A prediction model is a mathematical equation using patient risk factor data to estimate or predict the probability of patients experiencing a healthcare outcome [31]. As such, clinical prediction models can estimate the potential risk of developing disease for an individual on the conditional values of multiple predictors, such as gender, age, biomarkers [32], etc. Predictive modelling can be conducted using different approaches including rules, logistic regression, machine learning algorithms such as neural network, decision tree, support vector machine, Bayes network, random forest, boosted regression trees or deep learning techniques. Despite several potential advantages, artificial intelligence (AI) based methods are not widely deployed in the field of healthcare [26, 33]. For understanding and interpreting these models, healthcare practitioners need a basic understanding of the methodology of clinical prediction modelling [34].

The aim of this study was to develop and validate a T2DM risk prediction model using 53 variables representing behavioural, physical and mental health characteristics of the participants older than 50 years. In contrast to other similar studies [35], the data used in this study were collected in a longitudinal study of aging (SHARE). Since SHARE represents data collected in studies from different countries, we were also interested how the prediction performance varies across countries when local country-specific data are used compared to a global model build on data collected across the countries.

Methods

Study design and research data

Data from SHARE survey [36] waves 1 to 7 collected between 2004 and 2017 were used in this study to develop and validate prognostic models for 10-year T2DM prediction in European countries. Waves 1 (2004-2005) and 2 (2006-2007) were used as baseline data, while waves 6 (2015) and 7 (2017) were used as endpoints of the 10-year observation period [37]. The only exception was Israel, where wave 2 data were collected in 2009-2010 period and therefore the prediction window was shortened by approximately two years. SHARE is a cross-national longitudinal study on aging in Europe that allows simultaneous and consistent international comparisons of many factors including healthcare related items. Participants of the SHARE survey from Austria, Belgium, Denmark, France, Germany, Greece, The Netherlands, Israel, Italy, Spain and Sweden, Switzerland were included in this study. The data are freely available to registered researchers from the SHARE Research Data Center (http://www.share-project.org).

Study setting and sample

Populations in the SHARE study dataset consist of older adults who are at least 50 years old when they are invited for participation. Since the spouses of invited participants also participate in the interview, there might be some participants younger than 50 years included in the study. Two overlapping cohorts of data were used to select participants in the study. SHARE wave 1 included data for 43,969 participants from 12 countries while wave 2 included survey data for 55,295 participants from 14 countries. Age of the participants was limited to 50 or older and complete case analysis was performed without missing value imputation. We decided against missing value imputation due to many participants with large proportions of missing data where imputational approaches would result in most of the values being imputed. Since there was more than 50% of values missing in most of the samples with missing data, we removed all samples with missing data. More specific data preparation steps with more exact numbers of participants included in the final sample of 16,363 participants are described in the Results section under Data pre-processing.

Predictor variables

The variables included in model development and validation datasets consisted four groups of variables shown in Table 1. Some of the variables were removed from the initial set of variables as either they were not present in both waves or the questions significantly differed between waves 1 and 2.

Table 1. Overview of variables used as predictors in the model development and validation

Variable group	SHARE dataset	Number of variables	Examples / Sets of variables
Demographic	cv_r	2	Gender, Age.
Physical health	ph	15	Long-term illness, Heart attack (ever diagnosed), High blood pressure or hypertension, High blood cholesterol, Stroke, Chronic lung disease, Asthma, Arthritis, Osteoporosis, Cancer, Stomach or duodenal ulcer, peptic ulcer, Parkinson disease, Cataracts, Hip fracture or femoral fracture.
Aggregated health data	gv_health	31	Activities of daily living / quality of life (13 variables), BMI (2 variables), CASP (Control, Autonomy, Self-Realization and Pleasure) score, EURO-D depression (14 variables), Cognitive abilities (2 variables).
Behavioral data	br	5	Smoking (ever smoked), Alcohol consumption (days a week in the last 6 months), Vigorous sports activities, Activities requiring moderate level of energy, Question for participants on their ability to answer questions by themselves.

Outcome

The outcome used in model development was the presence of T2DM diagnosis in the follow-up period of 10 years for all participants who reported no T2DM diagnosis at the baseline interview at wave 1 or 2. To check for T2DM status we used the question: “Diabetes or high blood sugar: ever diagnosed?” from the SHARE physical health subset of data. In case of participants who were included in the study in wave 1, we checked for presence of self-reported T2DM in all consecutive waves up to wave 6. Wave

7 was used as an endpoint for participants who entered the study in wave 2. The mean difference between the first (wave 1 or 2) and the last (wave 6 or 7) interview was 10.5 ± 0.7 years. Out of 16,363 participants eligible for the model development phase 1,360 (8.3 %) answered positive to diabetes or high blood sugar question in waves 6 or 7. This percentage is comparable to the prevalence of T2DM in most European countries [35].

Statistical analysis and model validation

Due to a high number of variables with a high risk of multicollinearity, we decided to use a regularized logistic regression model. More specifically, an implementation of least absolute shrinkage and selection operator (LASSO) regression model from *glmnet* package [38, 39] in R statistical language [40] was adopted. To account for strongly unbalanced outcome variables, each instance was assigned a weight according to the inverse proportion of the outcome label. The lambda value defining the optimal number of variables to be included in the model was determined by 5-fold cross-validation. The global model was built using a set of additional dichotomous variables representing 12 countries while local country-specific models were developed using an initial set of 53 predictor variables.

Validation of the models was performed using repeated 10-fold cross-validation with 10 runs to estimate the variance of the results. The same validation approach was used for a global model as well as local country specific datasets. Basic prediction model metrics like Area Under the Receiver Operator Characteristic Curve (AUROC), Sensitivity, Specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV) were calculated in each run and each fold of repeated 10-fold cross-validation. Additionally, we calculated mean number of selected variables (by LASSO regression) representing the complexity of the final models and percentage of positively classified cases that can be used to estimate the cost of sending the individuals at risk to additional screening. 95% confidence intervals were calculated for all performance metrics.

Calibration of the models was assessed using visual comparison of predicted versus actual values. Slope and “calibration-in-the-large” (difference between mean predicted probabilities with mean observed outcomes) were calculated and presented along with the visualization [41]. Calibration of the models in this study was done using a regression based calibration approach where output values are transformed using the regression function that aims to balance the observed and predicted probability in the training set of data [42]. AUROC was calculated for basic and calibrated model to show that there were no significant differences in terms of performance after the calibration of the model output.

Ethical considerations and reporting

Secondary analysis of data was conducted under the SHARE Data Access Rules [37] allowing free access to the data for research purposes. During the SHARE study, all participants gave informed consent prior to taking part in the study. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) reporting guideline for reporting results of multivariate prediction models was followed to report the results in this study [43].

Results

Data pre-processing

Preparation of the data in this study was performed by extracting baseline variable values at wave 1 for participants who entered the study in 2004 or 2005. Wave 1 data consisted of 43,969 participants with 13,743 participants where no missing values were recorded for any of the 54 variables including the outcome where data for the follow-up waves (up to wave 6) had to be present. After this step we removed participants with diabetes present in wave 1 (n = 1,244) and participants who were younger than 50 years (n = 518) to obtain a final set of 11,981 participants from wave 1. We repeated the

procedure for all eligible participants who joined the study in wave 2 (n = 4382) and obtained a final set of 16,363 participants. Complete case analysis was performed as we removed all participants with missing data in any predictor variable (n = 4,208) thus reducing the number of potential maximal number of participants from 20,571 to 16,363. No bias due to imputation of the variable values was introduced, especially given the fact that a large proportion of the participants with missing values were missing almost all predictor variable values. A summary table with mean values or frequencies for the basic variables by T2DM status is displayed in Table 2.

Table 2. Summary table for the basic SHARE variables by T2DM status

		No T2DM	T2DM	p
Age	Mean (SD)	64.9 (9.2)	65.4 (8.8)	0.003
Gender	Male	6674 (90.3)	717 (9.7)	<0.001
	Female	8329 (92.8)	643 (7.2)	
BMI	Mean (SD)	26.1 (4.0)	28.8 (4.6)	<0.001
BMI Class	Underweight (below 18.5)	159 (96.4)	6 (3.6)	<0.001
	Normal (18.5-24.9)	6125 (96.1)	248 (3.9)	
	Overweight (25.0-29.9)	6439 (91.2)	623 (8.8)	
	Obese (30.0 and above)	2280 (82.5)	483 (17.5)	
Alcohol consumption in the last 6 months (br010)	Almost every day	3469 (92.7)	274 (7.3)	<0.001
	Five or six days a week	433 (91.5)	40 (8.5)	
	Three or four days a week	1260 (93.9)	82 (6.1)	
	Once or twice a week	3020 (93.1)	224 (6.9)	
	Once or twice a month	1776 (92.2)	150 (7.8)	
	Less than once a month	1236 (91.4)	117 (8.6)	
	Not at all in the last 6 months	3809 (89.0)	473 (11.0)	
Vigorous physical activity (br015)	More than once a week	6127 (93.0)	463 (7.0)	<0.001
	Once a week	2315 (92.6)	184 (7.4)	
	One to three times a month	1449 (91.3)	138 (8.7)	
	Hardly ever, or never	5112 (89.9)	575 (10.1)	

Additional to the global model built on data from all available participants, we also build country-specific models on country level data. Table 3 presents number of available participants for each country with percentage of positive T2DM cases in each country.

Model performance and variable importance

All 16,363 eligible participants were included in the global model with repeated cross-validation to obtain global prediction model validation results. Additional to AUROC of 0.702 (95% CI: 0.698 – 0.706), we also measured sensitivity of 0.642 (0.633 – 0.650) and specificity of 0.648 (0.643 – 0.652) that were well balanced.

Table 3. Percentage of participants who reported onset of T2DM in 10-year follow-up period

		No T2DM	T2DM
Country	All	15003 (91.7)	1360 (8.3)
	AT	879 (90.8)	89 (9.2)
	BE	2112 (93.4)	149 (6.6)
	CH	591 (95.3)	29 (4.7)
	DE	1107 (92.6)	88 (7.4)
	DK	1009 (93.8)	67 (6.2)
	ES	1031 (84.9)	184 (15.1)
	FR	1251 (91.9)	110 (8.1)
	GR	1660 (92.8)	128 (7.2)
	IL	929 (86.9)	140 (13.1)
	IT	1262 (89.6)	146 (10.4)
	NL	1470 (93.6)	101 (6.4)
	SE	1702 (93.0)	129 (7.0)

From an economic perspective, especially in prediction models used for screening of population, it is also important to know how many individuals were classified as high-risk patients. In our case, 37.8 % of participants were classified as positive – i.e. candidates for further testing. We were also interested in the number of variables selected by the LASSO logistic regression model where 40.10 (38.10 – 42.10) were selected on average. It should be noted that the global model consisted of 14 additional dichotomous variables representing country of participants compared to local models where only 53 basic variables were used.

Global model development and validation was followed by local, country specific model development and validation using the same classification metrics and cross-validation settings. The results for 100 runs in each country including a comparison to a global (All) model are presented in Figure 1. AUROC results ranged from 0.578 (0.565 - 0.592) in Greece to 0.768 (0.749 - 0.787) in Denmark. One can observe large difference in variance of the results with much wider confidence intervals in case of country specific models when compared to a global model built on much larger sample of data. There

were 3 countries (Denmark, Sweden and Switzerland) with mean AUROC higher than the AUROC of the global model.

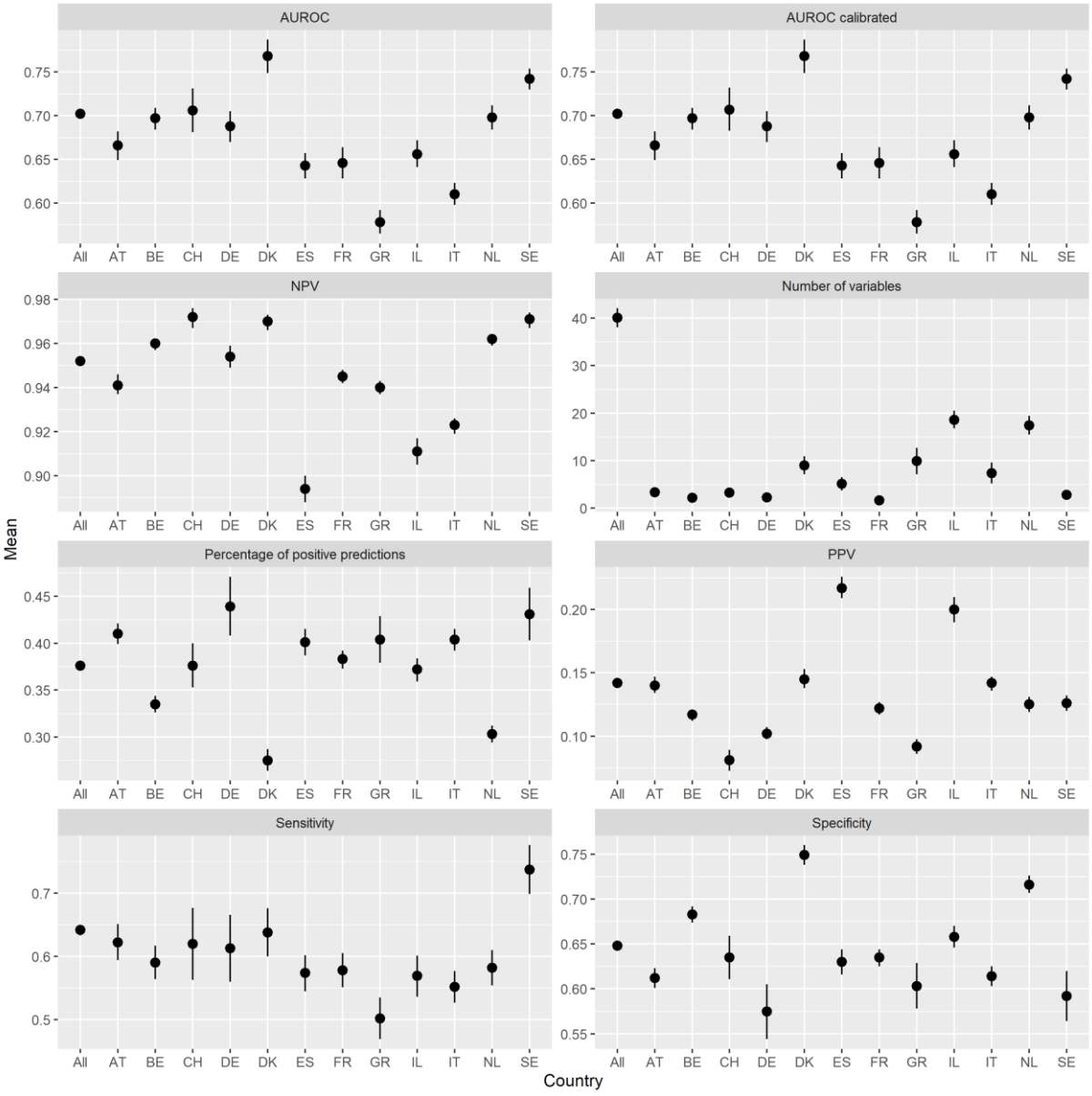


Figure 1. Overview of validation results for local and global models represented as mean values with corresponding 95% confidence intervals.

We also compared the complexity of models by country in terms of number of variables selected. The number of variables used in local models was significantly lower than number of variables in global model even if we consider additional dichotomous country related variables in global models. The simplest models in repeated 10-fold cross-validation were obtained on data from France with average number of variables at 1.61 (1.26 – 1.96). On the other hand, the most complex models were developed for data from Israel with average number of variables used in the model at 18.60 (16.80 – 20.50).

In Supplement 1 we provide information on selected features for all country specific as well as global model validation runs. Additional to SHARE variable names, we provide the frequency of positive and

negative coefficients for a specific selected variable that explains the influence of the variable on the outcome where positive outcome corresponds to higher ten-year risk for T2DM. There were 14 variables that were selected in all (n = 100) cross-validation runs for the global model. Those variables represented the following characteristics of the participants: BMI, alcohol consumption, physical activity, CASP (Control, Autonomy, Self-Realization and Pleasure) score, orientation in time score, presence of any chronical health problems, high blood pressure or hypertension, high blood cholesterol, self-rated health and three dichotomous variables representing country of the interview (Spain, Israel and Italy).

In country specific models a set of selected variables varies from country to country. In countries with extremely small number of selected variables, we observed a high frequency of variables like BMI, BMI Class or Gender. In the Danish model with the best prediction performance, BMI was the only variable that was selected in all cross-validation runs, followed by Alcohol consumption in the last six months (br010) and Smoking status (br001) which were selected in 48% and 47% of the models, respectively.

Model calibration

Performance of the model, often measured as accuracy or AUROC (also known as c-statistics), is one of the most important aspects when considering which prediction model can be applied in clinical settings. However, the models need to be well calibrated as well to avoid over- (predicted probability is much higher than observed probability) or under-prediction (predicted probability is much lower than observed probability). One of the possibilities to observe the calibration of the model is to plot a probability curve where we group and plot individuals according to their predicted probability against the observed probability.

Figure 2 presents a typical example demonstrating that even with a very similar slope of the calibration line we can observe very different calibration of the model [44]. Therefore a measure that can be used to detect consistent deviance between predicted probability and observed proportion like “calibration-in-the-large” gives us an important information on the calibration of the model. We can observe the strong overprediction in the uncalibrated model with “calibration-in-the large” value of -2.39. The AUROC (c-stiatic) remains the same – not only in case of a global model recalibration, but also for the local models (Figure 1).

In supplemental material (Supplement 2) we provide calibrated vs. uncalibrated plots for all 12 country-specific models. Similar to the global model, the uncalibrated models demonstrate strong overprediction in all local models as well. However, the problem of small sample size is much more evident here as even after the calibration, some models show large deviations and variance from the optimal predicted vs. observed probability line.

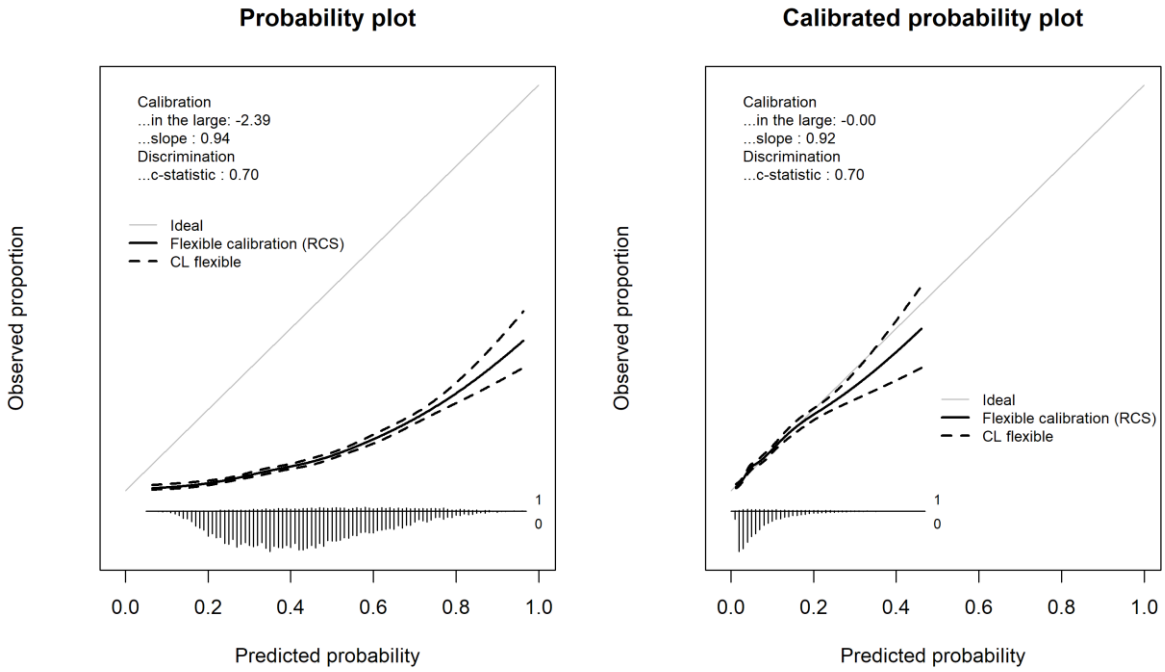


Figure 2. Comparison of probability plots for basic and calibrated model.

Discussion

Main findings

We developed and validated prediction models for 10-year T2DM risk score prediction using SHARE survey data. A global model from all available country-specific data was used and compared to local models to demonstrate the strength of collecting the data from multiple countries in a single database. A prediction model built on this database achieved an AUROC of 0.70 (95% CI: 0.70 – 0.71). Twelve country specific models were also developed and validated on the locally collected data to expose the problems related to small sample size due to limited local level data collection. Additionally, we were able to show large variance in the quality of collected data and results obtained in different countries.

The existing models use different follow-up periods for T2DM prediction, but most frequently five or ten-year prediction window is used. Most prediction models can successfully identify people at high risk of developing T2DM in a time frame of 5 to 10 years [45], but the performance of the model varies among different studies and datasets. [35] updated the widely used Finnish diabetes risk questionnaire, a simple tool for identification of those at risk for drug-treated type 2 diabetes. They used clinically diagnosed T2DM as an outcome on a data sample of 18,301 participants from different centres in different European countries with a five-year prediction window. The overall prediction performance was 0.74 (0.73 - 0.76), but the AUROC on the level of healthcare centres where the data was collected ranged from 0.63 to 0.78.

Kengne et al. [46] compared 12 prediction models for incident diabetes on data from eight European countries. Most of the analysed prediction models have used the following variables when predicting T2DM: age, gender and waist circumference. The AUROC ranged from 0.69 (0.68 - 0.70) to 0.85 (0.83 - 0.87). It is interesting to note that the lowest performance was recorded on the dataset from Denmark, while our results using SHARE data achieved the best performance in the same country with

AUROC of 0.77 (0.75 - 0.79). Overall performance measured in AUROC for twelve different models tested on pooled data from all eight European countries (n = 27,779) ranged from 0.76 (0.74 – 0.79) to 0.81 (0.77 – 0.84) for the QDscore model [47]. Both studies also emphasized the importance of model recalibration, especially when adapting models to a specific population.

Limitations

There are multiple limitations related to the nature of the data that was collected in the scope of a large cross-national study on aging. One of the limitations is related to the questions asked by the interviewers including the question that was used to define the outcome of our models. The question whether the participants were ever diagnosed with Diabetes or high blood sugar is not specific enough to clearly identify the onset of T2DM. Therefore, it can only be used as a proxy to define T2DM patients and can often be interpreted differently in the local context. It can also include patients with pre-diabetes status as it also includes self-reported high blood sugar status. However, the percentage of the positive cases in each country corresponds to the actual values of the T2DM prevalence in the European countries. Another limitation includes the follow-up time from the initial interview for specific participants. On average, the follow-up time exceeded 10 years, but there was also a case where the initial wave of the study was conducted two years later as in other countries. Therefore, the results from Israel should not be generalized and should be interpreted separately from other country specific models.

Implications of the study

To our knowledge, this is the first study to explore the possibilities in building prediction models for T2DM risk based on a large cross-national survey dataset. It offers policymakers the insight into the possibilities, advantages and limitations on using the large survey studies to build screening prediction models from data collected across different countries. In addition to prediction, we also assessed calibration of the models that should represent an important characteristic for all prediction models introduced in the clinical environment. In this respect, we show that one should observe both, the offset and the slope of the probability plot curve as noted by Siregar et al. [48].

As discussed in the previous section on limitations, the outcome variable represents one of the limitations in building even better prediction models from large survey datasets. In case of SHARE dataset, the dried blood spot samples were already collected in waves 6 and 7 in most participating countries [49]. It was found that we can reliably measure Haemoglobin A1c which would represent an important step towards obtaining more reliable status of T2DM in participants of the SHARE study. It is expected that the first results of the SHARE study including the A1c results will be available in 2020.

Multiple studies have already demonstrated the effectiveness of using electronic health record (EHR) data to build T2DM screening models [50]. Despite many problems related to collecting the data that can be used in building prediction models [51], we believe that a combination of data from EHR and survey data could represent an important step toward reliable and effective screening prediction models. Data from the Danish SHARE participants (n = 3,400) were already linked to registry data (education records, dementia diagnoses, and mortality) to show important results in relation between education and cognitive decline in older population [52]. This use case could point at a great potential of linking data from the EHR and survey data for the future of screening model development at the primary healthcare level.

References

1. Zheng, Y., Ley, S. H. & Hu, F. B. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nature Reviews Endocrinology, 14, 88-98 (2018).
2. 2. Upadhyay, J, Polyzos, S. A., Perakakis, N., Thakkar, B., Paschou, S. A., Katsiki, N., Underwood, P., Park, K-H., Seufert, J., Kang, E. S., Sternthal, E., Karagianniss, A., Mantzoros, C. S. Pharmacotherapy of type 2 diabetes: An update. Metabolism, 78, 13-42 (2018).
3. International Diabetes Federation. IDF Diabetes Atlas - 8th Edition. https://diabetesatlas.org/resources/2017-atlas.html (2017).
4. 4. Strain, W. D., Hope, S. V., Green, A., Kar, P., Valabhji, J., Sinclair, A. J. Type 2 diabetes mellitus in older people: a brief statement of key principles of modern day management including the assessment of frailty. A national collaborative stakeholder initiative. Diabetic Medicine, 35, 838-845 (2018).
5. 5. Sinclair, A. J., Abdelhafiz, A. H., Forbes, A., Munshi, M. Evidence-based diabetes care for older people with Type 2 diabetes: a critical review. Diabetic Medicine, 36, 399-413 (2019).
6. 6. Eurostat. Aging Europe: Looking at the lives of older people in the EU. Luxembourg: Publications Office of the European Union. https://ec.europa.eu/eurostat/documents/3217494/10166544/KS-02-19-681-EN-N.pdf/c701972f-6b4e-b432-57d2-91898ca94893 (2019).
7. Hill, J. The older person with diabetes: considerations for care. British Journal of Community Nursing, 24(4), 160-164 (2019).
8. Rodriguez-Sanchez, B., & Cantarero-Prieto, D. Socioeconomic differences in the associations between diabetes and hospital admission and mortality among older adults in Europe. Economics & Human Biology, 33, 89-100 (2019).
9. Larsson, S. C., Wallin, A., Hakansson, N., Stackelberg, O, Back, M., Wolk, A. Type 1 and type 2 diabetes mellitus and incidence of seven cardiovascular diseases. International Journal of Cardiology, 262, 66-70 (2018).
10. SHARE. SHARE - Survey of Health, Ageing and Retirement in Europe. http://www.share-project.org/ (2019).
11. Dagogo-Jack, S. Prevention begets prevention — lessons from the Da Qing Study. Nature Reviews Endocrinology, (2019).
12. Yu, D., Zheng, W., Cai, H., Xiang, Y-B., Li, H., Gao, Y-T., Shu, X-O. Long-term Diet Quality and Risk of Type 2 Diabetes Among Urban Chinese Adults. Diabetes Care, 41(4), 723-730 (2018).
13. Palacios, O. M., Kramer, M., & Maki, K. C. Diet and prevention of type 2 diabetes mellitus: beyond weight loss and exercise. Expert Review of Endocrinology & Metabolism, 14(1), 1-12 (2018).
14. Weickert, M. O., Pfeiffer, A. F. H. Impact of Dietary Fibre Consumption on Insulin Resistance and the Prevention of Type 2 Diabetes. J Nutr, 148, 7–12 (2018).
15. Schwingshackl, L., Hoffmann, G., Lampousi, A.M., Knüppel, S., Iqbal, K., Schwedhelm, C., et al. Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies. Eur J Epidemiol, 32, 363–375 (2017).
16. Aune, D., Norat, T., Romundstad, P., & Vatten, L.J. Whole grain and refined grain consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of cohort studies. Eur J Epidemiol, 28, 845–858 (2013).
17. Blatt, D., & Gostic, C. L. Reducing the Risk of Diabetes and Metabolic Syndrome with Exercise and Physical Activity. Nutritional and Therapeutic Interventions for Diabetes and Metabolic Syndrome, 2nd ed., 315-327 (2018).

18. Wang, Y. W., Lee, D., Brellenthim, A. G., Eijsvogels, T. M. H., Sui, X., Church, T. S., Lavie, C. J., Blair, S. N. Leisure-Time Running Reduces the Risk of Incident Type 2 Diabetes. The American Journal of Medicine, in press (2019).
19. De Souto Barreto, P., Cesari, M., Andrieu, S., Vellas, B., & Rolland, Y. Physical Activity and Incident Chronic Diseases: A Longitudinal Observational Study in 16 European Countries. American Journal of Preventive Medicine, 52(3), 373–378 (2017).
20. Savikj, M., Gabrial, B. M., Alm, P. S., Smith, J., Caidahl, K., Bjornholm, M., Fritz, T., Krook, A., Zeirath, J. R., Wallberg-Henriksson, H. Afternoon exercise is more efficacious than morning exercise at improving blood glucose levels in individuals with type 2 diabetes: a randomised crossover trial. Diabetologia, 62(2), 233-237 (2019).
21. Jang, J-E., Cho, Y., Lee, B. W., Shin, E-S., Lee, S. H. Effectiveness of Exercise Intervention in Reducing Body Weight and Glycosylated Hemoglobin Levels in Patients with Type 2 Diabetes Mellitus in Korea: A Systematic Review and Meta-Analysis. Diabetes & Metabolism, 43(3), 302-318 (2019).
22. American Diabetes Association. Prevention or Delay of Type 2 Diabetes: Standards of Medical Care in Diabetes – 2018. Diabetes Care, 41(1), 51-54 (2018).
23. Kaur, H., & Kumari, V. Predictive modelling and analytics for diabetes using a machine learning approach. Applied Computing and Informatics, in press (in press).
24. Lai, H., Huang, H., Keshavjee, K., Guergachi, A., Gao, X. Predictive models for diabetes mellitus using machine learning techniques. BMC Endocrine Disorders, 19(101), 1-9 (2019).
25. Wu, H., Yang, S., Huang, Z., He, J., Wang, X. Type 2 diabetes mellitus prediction model based on data mining. Informatics in Medicine Unlocked, 10, 100-107 (2018).
26. Luo, G. Automatically explaining machine learning prediction results: a demonstration on type 2 diabetes risk prediction. Health Information Science and Systems, 4(2), 1-9 (2016).
27. Han, L., Luo, S., Yu, J., Pan, L., Chen, S. Rule extraction from support vector machines using ensemble learning approach: an application for diagnosis of diabetes. IEEE J Biomed Health Inform, 9(2), 728–734 (2015).
28. Razavian, N., Blecker, S., Schmidt, A. M., Smith-McLallen, A., Nigam, S., Sontag, D. Population-Level Prediction of Type 2 Diabetes From Claims Data and Analysis of Risk Factors. Big Data, 3(4), 277-287 (2015).
29. Habibi, S., Ahmadi, M., Alizadeh, S. Type 2 diabetes mellitus screening and risk factors using decision tree: results of data mining. Global J Health Sci, 7(5), 304–310 (2015).
30. Iyer, A., Jeyalatha, S., Sumbaly, R. Diagnosis of diabetes using classification mining techniques. Int J Data Min Knowl Manage Process (IJDKP), 5(1), 1–14 (2015).
31. Grant, S. W., Collins, G. S., Nashef, A. M. Statistical Primer: developing and validating a risk prediction model. Eur J Cardiothorac Surg, 54, 203–208 (2018).
32. Bonnett, L. J., Snell, K. I. E., Collins, G. S., Riley, R. D. Guide to presenting clinical prediction models for use in clinical settings. BMJ, 365(737), 1-8 (2019).
33. Stiglic, G., Watson, R., Cilar, L. R you ready? Using the R programme for statistical analysis and graphics. Research in Nursing & Health, 1-6 (2019).
34. Mijderwijk, H., Steyerberg, E. W., Steiger, H.J., Fischer, I., Kamp, M. A. Fundamentals of Clinical Prediction Modeling for the Neurosurgeon. Neurosurgery, 85, 302–311 (2019).
35. Alssema, M., Vistisen, D., Heymans, M. W., Nijpels, G., Glümer, C., Zimmet, P. Z., ... & Colagiuri, S. The Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance (DETECT-2) update of the Finnish diabetes risk score for prediction of incident type 2 diabetes. Diabetologia, 54(5), 1004-1012 (2011).
36. Börsch-Supan, A. Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7. Release version: 7.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w7.700 (2019).

37. Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., Zuber, S. Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). International Journal of Epidemiology (2013).
38. Friedman, J., Hastie, T., & Tibshirani, R. glmnet: Lasso and elastic-net regularized generalized linear models. R package version, 1(4) (2009).
39. Hastie, T., Tibshirani, R., & Wainwright, M. Statistical learning with sparsity: the lasso and generalizations. Chapman and Hall/CRC (2015).
40. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ (2019).
41. Van Calster, B., Nieboer, D., Vergouwe, Y., De Cock, B., Pencina, M.J., Steyerberg E.W. A calibration hierarchy for risk models was defined: from utopia to empirical data. Journal of Clinical Epidemiology, 74, 167-176 (2016).
42. Steyerberg, E.W. Clinical prediction models. New York: Springer (2019).
43. Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMC medicine, 13(1), 1 (2015).
44. Stevens, R.J., Poppe, K.K., Validation of clinical prediction models: what does the "calibration slope" really measure?, Journal of Clinical Epidemiology (2019).
45. Abbasi, A. et al. Prediction models for risk of developing type 2 diabetes: systematic literature search and independent external validation study. BMJ, 345, 1-16 (2012).
46. Kengne, A. P., et al. Non-invasive risk scores for prediction of type 2 diabetes (EPIC-InterAct): a validation of existing models. Lancet Diabetes Endocrinol, 2, 19-29 (2014).
47. Hippisley-Cox, J., Coupland, C., Robson, J., Sheikh, A., Brindle, P. Predicting risk of type 2 diabetes in England and Wales: Prospective derivation and validation of QDScore. BMJ, 338, 880 (2009).
48. Siregar, S., Nieboer, D., Versteegh, M. I., Steyerberg, E. W., & Takkenberg, J. J. Methods for updating a risk prediction model for cardiac surgery: a statistical primer. Interactive cardiovascular and thoracic surgery, 28(3), 333-338 (2019).
49. Börsch-Supan, A., Börsch-Supan, M., & Weiss, L. M. 36 Dried blood spot samples and their validation. Health and socio-economic status over the life course: First results from SHARE Waves 6 and 7, 349 (2019).
50. Stiglic, G., Kocbek, P., Cilar, L., Fijacko, N., Stozer, A., Zaletel, J., Sheikh, A., Povalej Brzan, P. Development of a screening tool using electronic health records for undiagnosed Type 2 diabetes mellitus and impaired fasting glucose detection in the Slovenian population. Diabetic Medicine, 35(5), 640-649 (2018).
51. Stiglic, G., Kocbek, P., Fijacko, N., Sheikh, A., & Pajnkihar, M. Challenges associated with missing data in electronic health records: a case study of a risk (2017).
52. Foverskov, E., Glymour, M. M., Mortensen, E. L., Holm, A., Lange, T., & Lund, R. Education and Cognitive Aging: Accounting for Selection and Confounding in the Danish Registry-SHARE Data Linkage. American Journal of Epidemiology (2018).

Supplementary Materials

Supplement 1. Lists of selected variables for all country specific models and a global model

Supplement 2. Probability plots for all country specific uncalibrated and calibrated models

Supplement 1. Selection of variables by LASSO regression

Detailed descriptions for specific SAHRE variables can be found at: https://g2aging.org/?section=survey&surveyid=6&display=flowchart (under Survey modules).

Variables selected by LASSO: AT

Num	Variable	Positive	Negative
1	bmi	100	0
2	gender	0	62
3	euro3	0	30
4	eurodcat	0	21
5	casp	0	18
6	euro2	13	0
7	euro5	9	0
8	ph006d7	0	9
9	br010_	7	0
10	br016_	7	0
11	maxgrip	7	0
12	ph004_	0	6
13	ph006d6	0	5
14	euro4	0	4
15	ph006d13	0	4
16	br017_	0	3
17	euro8	0	3
18	ph006d2	3	0
19	euro1	0	2
20	euro7	0	2
21	gali	0	2
22	ph006d1	2	0
23	ph006d4	0	2
24	adl2	1	0
25	bmi2	1	0
26	br001_	0	1
27	br015_	1	0
28	euro10	1	0

Num	Variable	Positive	Negative
29	euro9	0	1
30	mobilit3	0	1
31	numeracy	1	0
32	orienti	1	0
33	ph006d10	1	0
34	ph006d12	0	1
35	ph006d3	1	0
36	ph006d9	0	1
37	phactiv	1	0
38	sphus	0	1
39	sphus2	0	1

Variables selected by LASSO: BE

Num	Variable	Positive	Negative
1	bmi	99	0
2	bmi2	99	0
3	sphus2	7	0
4	br010_	3	0
5	euro9	3	0
6	ph006d8	3	0
7	br001_	0	1
8	br017_	1	0
9	ph006d12	0	1
10	ph006d9	1	0

Variables selected by LASSO: CH

Num	Variable	Positive	Negative
1	bmi2	99	0
2	bmi	39	0
3	euro12	0	27
4	ph006d2	25	0
5	euro2	0	20
6	mobilit3	13	0
7	ph004_	0	12
8	ph006d9	0	12
9	sphus	12	0
10	ph006d13	0	10
11	sphus2	10	0
12	ph006d3	9	0
13	euro4	0	5
14	iadl2	0	5
15	br015_	4	0
16	ph006d4	0	4
17	euro7	3	0
18	euro8	0	3
19	phactiv	3	0
20	br001_	0	2
21	euro5	0	2
22	br010_	1	0
23	casp	0	1
24	euro6	0	1
25	numeracy	1	0
26	ph006d1	0	1
27	ph006d8	1	0

Variables selected by LASSO: DE

Num	Variable	Positive	Negative
1	bmi2	100	0
2	bmi	61	0
3	br001_	0	9
4	gali	7	0
5	euro9	0	6
6	ph006d2	6	0
7	age	4	0
8	mobilit3	0	4
9	orienti	0	4
10	ph006d6	0	4
11	casp	0	3
12	euro8	0	3
13	euro10	2	0
14	gender	0	2
15	ph004_	0	2
16	br010_	0	1
17	br017_	1	0
18	euro3	0	1
19	ph006d10	0	1
20	ph006d12	1	0
21	ph006d14	0	1

Variables selected by LASSO: DK

Num	Variable	Positive	Negative
1	bmi	100	0
2	br010_	48	0
3	br001_	0	47
4	euro5	45	0
5	ph006d7	45	0
6	ph006d2	41	0
7	adl2	40	0
8	euro2	0	38
9	orienti	0	38
10	ph004_	0	38
11	ph006d4	0	38
12	bmi2	37	0
13	br015_	36	0
14	ph006d8	0	35
15	br016_	0	32
16	euro4	0	32
17	phactiv	29	0
18	ph006d3	27	0
19	iadl2	20	0
20	euro10	0	14
21	gender	0	14
22	ph006d6	11	0
23	ph006d9	10	1
24	mobilit2	10	0
25	ph006d14	0	10
26	euro6	8	1
27	eurodcat	7	0
28	numeracy	7	0
29	ph006d11	5	1
30	euro11	0	5
31	euro3	0	5
32	euro9	4	1
33	ph006d13	5	0
34	ph006d12	3	1
35	euro7	1	2

Num	Variable	Positive	Negative
36	ph006d10	0	3
37	sphus	3	0
38	br017_	0	2
39	ph006d1	0	1
40	sphus2	0	1

Variables selected by LASSO: ES

Num	Variable	Positive	Negative
1	bmi2	100	0
2	bmi	97	0
3	orienti	0	95
4	mobilit2	16	0
5	euro4	0	13
6	br001_	0	12
7	age	10	0
8	br010_	10	0
9	br017_	0	10
10	casp	0	9
11	ph006d3	9	0
12	ph006d4	9	0
13	euro3	8	0
14	ph006d10	0	8
15	euro9	0	7
16	mobilit3	0	7
17	ph006d2	7	0
18	sphus2	7	0
19	euro10	0	6
20	euro11	6	0
21	ph006d9	0	6
22	br015_	5	0
23	gender	0	5
24	ph006d12	4	1
25	ph006d8	0	5
26	br016_	4	0
27	ph006d14	4	0
28	ph006d13	3	0
29	adl	0	2
30	adl2	2	0
31	numeracy	1	1
32	ph004_	2	0
33	ph006d11	0	2
34	ph006d6	0	2
35	euro1	0	1

Num	Variable	Positive	Negative
36	euro12	1	0
37	euro2	0	1
38	euro5	0	1
39	euro6	0	1
40	euro8	0	1
41	eurodcat	1	0
42	gali	1	0
43	iadl	1	0
44	iadl2	0	1
45	maxgrip	1	0
46	mobility	1	0
47	ph006d1	1	0
48	ph006d7	1	0
49	phactiv	0	1
50	sphus	0	1

Variables selected by LASSO: FR

Num	Variable	Positive	Negative
1	bmi	100	0
2	br017_	12	0
3	sphus	12	0
4	ph006d2	10	0
5	br015_	7	0
6	euro3	0	7
7	euro2	4	0
8	gali	0	3
9	bmi2	1	0
10	euro1	1	0
11	euro12	0	1
12	gender	0	1
13	maxgrip	1	0
14	mobility	1	0

Variables selected by LASSO: GR

Num	Variable	Positive	Negative
1	sphus	100	0
2	ph006d3	69	0
3	ph006d1	68	0
4	ph006d9	0	50
5	euro6	0	34
6	maxgrip	33	0
7	euro3	32	0
8	iadl2	5	27
9	casp	0	27
10	euro11	0	23
11	euro5	0	23
12	bmi	21	0
13	ph006d11	0	21
14	phactiv	0	21
15	iadl	0	20
16	ph006d4	0	20
17	adl2	19	0
18	br016_	0	19
19	ph006d10	0	19
20	euro9	18	0
21	euro7	0	16
22	ph006d6	0	16
23	euro2	0	14
24	numeracy	0	14
25	euro12	13	0
26	ph006d13	0	13
27	br015_	12	0
28	mobilit3	8	4
29	ph006d12	11	1
30	ph006d14	10	2
31	ph006d8	0	12
32	age	11	0
33	bmi2	0	11
34	euro1	0	11
35	euro8	9	2

Num	Variable	Positive	Negative
36	orienti	11	0
37	br001_	9	1
38	euro4	6	4
39	mobility	10	0
40	adl	6	3
41	br010_	0	9
42	mobilit2	2	7
43	ph004_	9	0
44	ph006d2	9	0
45	ph006d7	7	2
46	sphus2	3	6
47	gender	7	1
48	br017_	0	6
49	eurodcat	4	2
50	euro10	3	2
51	gali	1	4

Variables selected by LASSO: IL

Num	Variable	Positive	Negative
1	bmi2	100	0
2	casp	0	100
3	gender	0	100
4	ph006d10	0	91
5	maxgrip	90	0
6	br010_	87	0
7	euro11	82	0
8	euro2	77	0
9	euro7	0	77
10	br017_	71	0
11	ph006d1	0	71
12	br001_	69	0
13	ph006d4	66	0
14	ph006d6	62	0
15	euro3	60	0
16	ph006d11	59	0
17	iadl2	0	49
18	ph006d2	48	0
19	ph006d8	45	0
20	mobility	44	0
21	sphus2	40	0
22	br016_	37	0
23	euro6	0	32
24	phactiv	0	27
25	ph006d14	0	25
26	ph006d9	0	25
27	bmi	22	0
28	euro10	20	0
29	euro5	0	20
30	age	17	1
31	euro12	0	18
32	ph006d7	18	0
33	adl2	14	0
34	mobilit3	14	0
35	orienti	0	14

Num	Variable	Positive	Negative
36	ph004_	11	0
37	iadl	0	9
38	br015_	8	0
39	sphus	7	0
40	ph006d12	5	0
41	ph006d13	3	2
42	ph006d3	5	0
43	euro1	2	2
44	eurodcat	0	4
45	mobilit2	4	0
46	euro4	1	1
47	euro8	1	1
48	euro9	0	2
49	gali	2	0
50	numeracy	2	0
51	adl	0	1

Variables selected by LASSO: IT

Num	Variable	Positive	Negative
1	bmi	100	0
2	bmi2	61	0
3	age	54	0
4	ph004_	0	46
5	gali	42	0
6	br016_	33	0
7	br015_	26	0
8	ph006d10	22	0
9	ph006d2	21	0
10	ph006d6	21	0
11	sphus	21	0
12	ph006d4	0	16
13	euro9	0	11
14	ph006d12	0	11
15	ph006d8	0	11
16	adl2	10	0
17	br017_	10	0
18	euro12	10	0
19	euro4	0	10
20	euro8	0	10
21	gender	0	10
22	mobility	0	10
23	ph006d14	0	10
24	br010_	9	0
25	euro6	9	0
26	maxgrip	0	9
27	numeracy	9	0
28	phactiv	0	9
29	adl	0	8
30	euro2	0	8
31	ph006d3	2	6
32	br001_	0	7
33	ph006d7	7	0
34	ph006d9	0	7
35	casp	0	6

Num	Variable	Positive	Negative
36	euro10	6	0
37	euro7	0	6
38	mobilit3	0	6
39	ph006d11	0	6
40	euro11	2	3
41	eurodcat	5	0
42	ph006d13	0	5
43	sphus2	3	2
44	euro5	2	2
45	iadl2	1	3
46	mobilit2	2	2
47	orienti	3	1
48	ph006d1	0	4
49	iadl	0	2
50	euro3	0	1

Variables selected by LASSO: NL

Num	Variable	Positive	Negative
1	bmi2	100	0
2	sphus	98	0
3	age	96	0
4	bmi	94	0
5	ph006d4	94	0
6	euro10	93	0
7	gender	0	89
8	ph006d1	87	0
9	ph006d11	83	0
10	ph006d10	77	0
11	ph004_	0	76
12	br010_	61	0
13	euro9	50	0
14	ph006d6	44	0
15	br015_	43	0
16	eurodcat	40	0
17	numeracy	0	36
18	iadl	0	35
19	euro4	0	33
20	br001_	0	32
21	orienti	0	32
22	gali	0	30
23	euro12	0	27
24	ph006d2	25	0
25	phactiv	0	25
26	euro3	0	24
27	euro8	23	0
28	iadl2	0	18
29	euro6	15	1
30	mobilit2	0	15
31	casp	1	13
32	euro2	0	14
33	ph006d9	0	14
34	ph006d13	0	13
35	br017_	0	11

Num	Variable	Positive	Negative
36	ph006d7	0	10
37	ph006d8	10	0
38	euro7	9	0
39	euro1	0	8
40	ph006d14	0	7
41	ph006d3	7	0
42	adl	6	0
43	adl2	0	4
44	euro11	1	3
45	mobilit3	0	4
46	ph006d12	3	1
47	sphus2	4	0
48	br016_	0	2
49	maxgrip	1	1
50	euro5	1	0
51	mobility	0	1

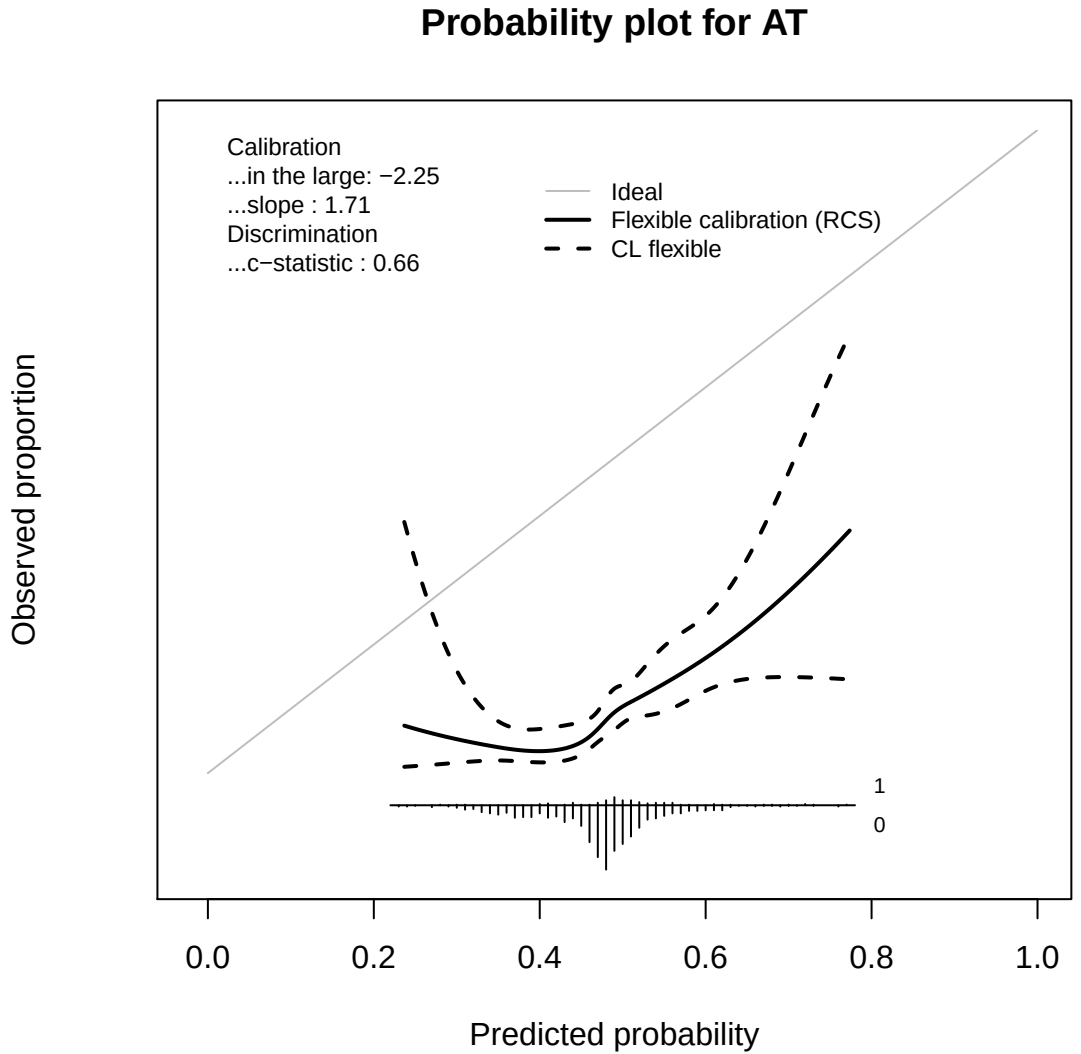
Variables selected by LASSO: SE

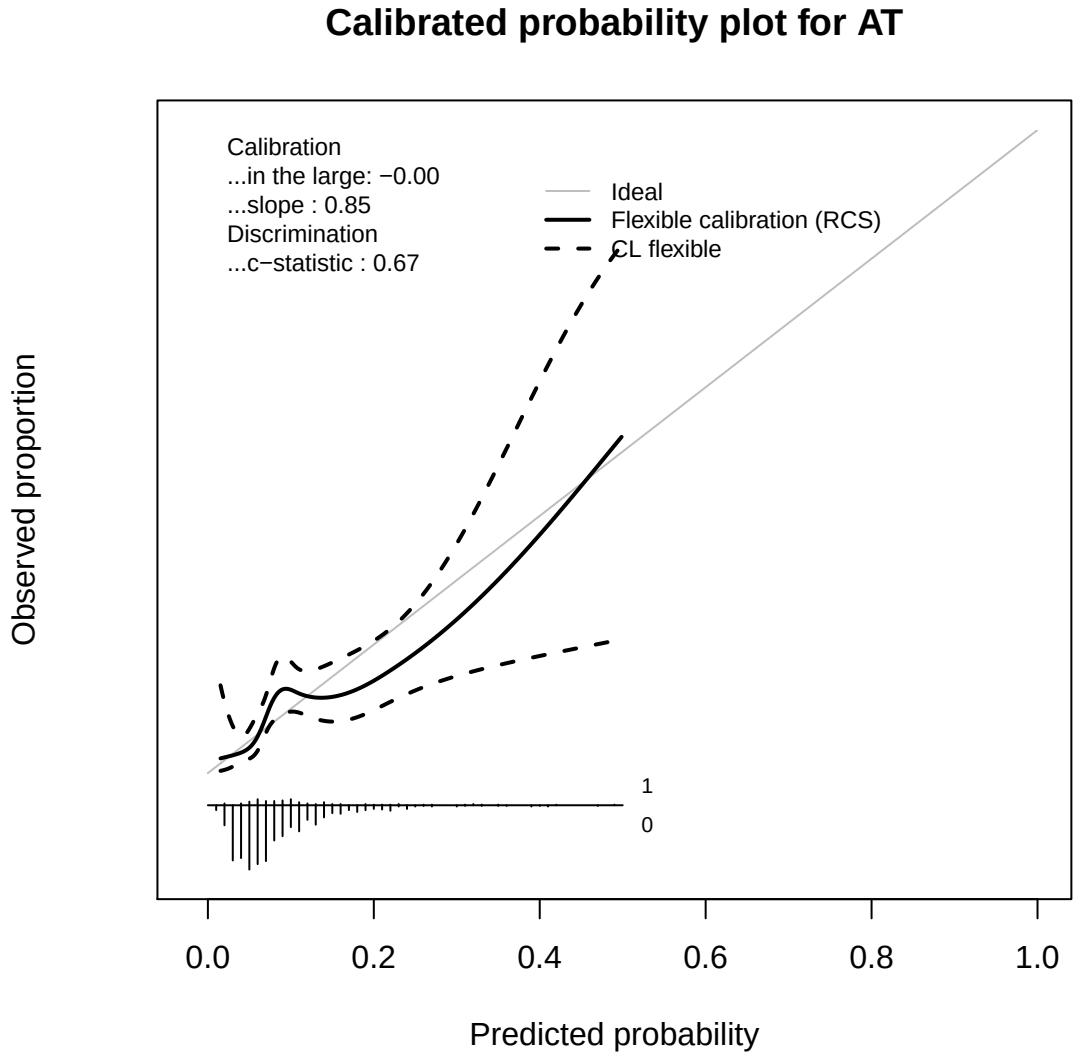
Num	Variable	Positive	Negative
1	bmi2	100	0
2	ph006d3	47	0
3	ph006d2	31	0
4	gender	0	23
5	ph006d10	0	17
6	br016_	14	0
7	ph006d11	12	0
8	maxgrip	7	0
9	ph006d13	0	5
10	euro7	0	4
11	sphus2	4	0
12	bmi	3	0
13	br001_	0	2
14	euro8	0	2
15	br015_	1	0
16	euro2	1	0
17	euro5	0	1
18	iadl	0	1
19	mobilit2	1	0
20	ph004_	0	1
21	ph006d14	0	1
22	ph006d8	0	1

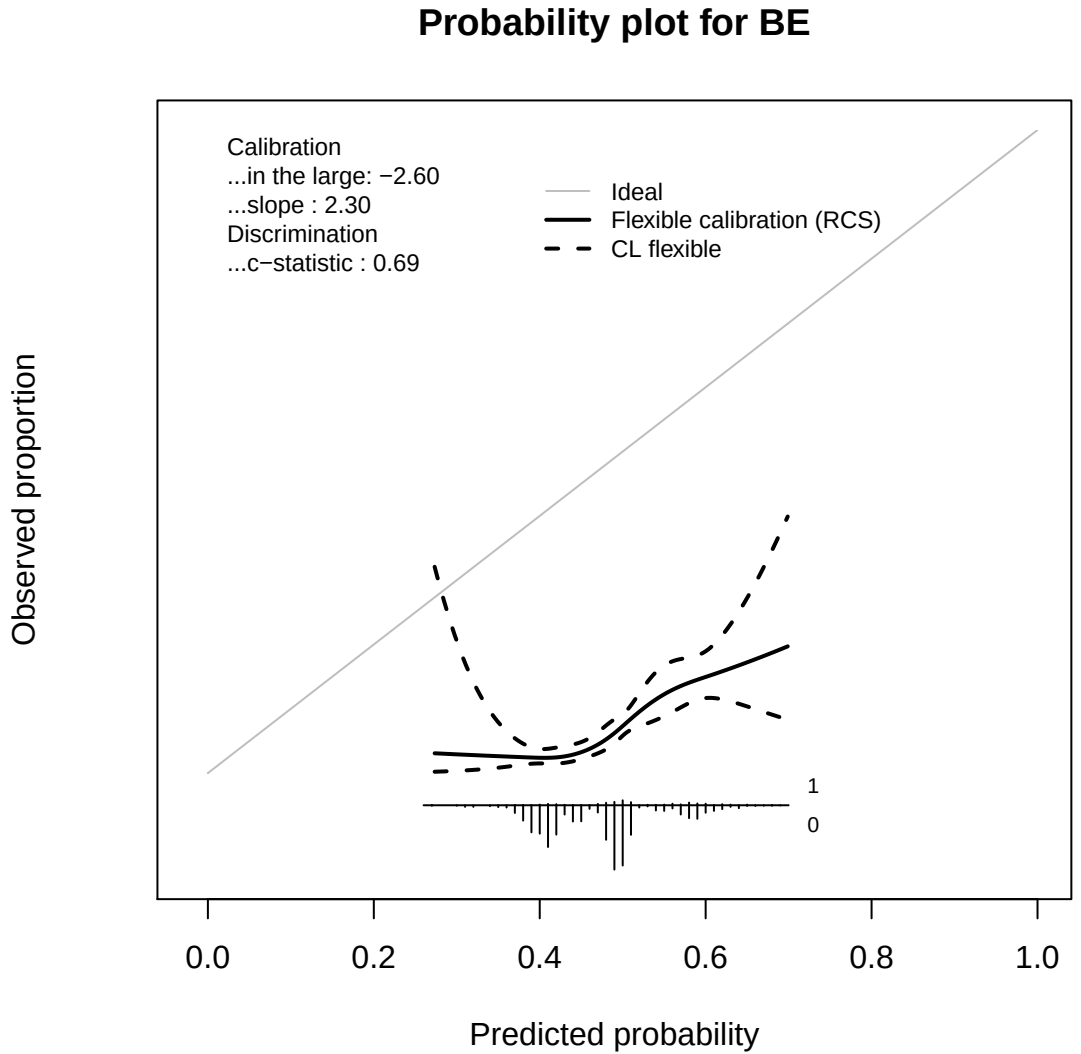
Variables selected by LASSO: All

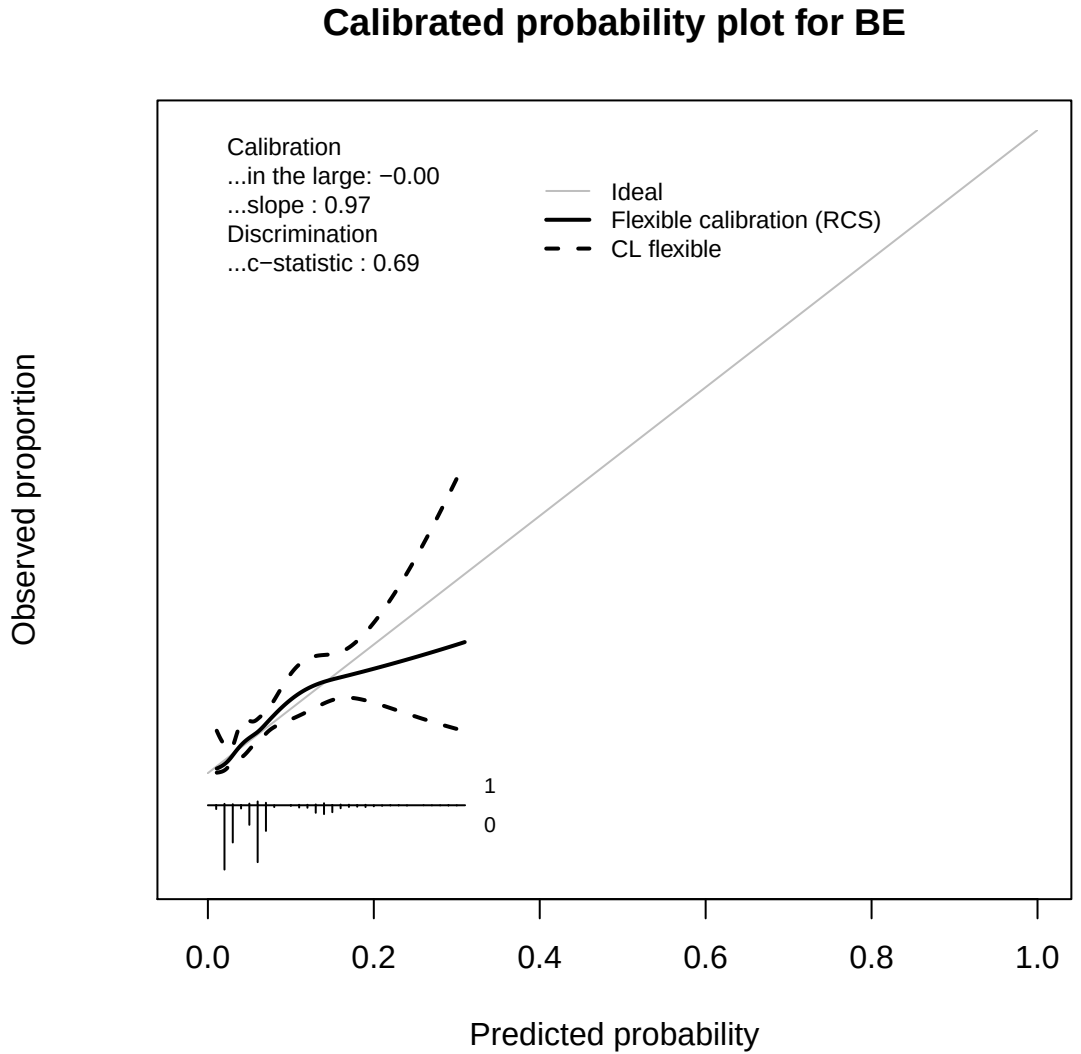
Num	Variable	Positive	Negative
1	bmi	100	0
2	bmi2	100	0
3	br010_	100	0
4	br015_	100	0
5	casp	0	100
6	country_ES	100	0
7	country_IL	100	0
8	country_IT	100	0
9	gender	0	100
10	orienti	0	100
11	ph004_	0	100
12	ph006d2	100	0
13	ph006d3	100	0
14	sphus2	100	0
15	br001_	0	99
16	sphus	96	0
17	country_BE	0	95
18	country_CH	0	95
19	iadl2	0	95
20	age	93	0
21	br017_	92	0
22	euro10	92	0
23	euro7	0	91
24	country_AT	89	0
25	country_FR	88	0
26	ph006d13	0	85
27	adl2	78	0
28	ph006d9	0	78
29	euro4	0	76
30	phactiv	0	74
31	mobilit3	0	73
32	euro2	71	0
33	mobilit2	71	0
34	euro8	0	68
35	ph006d8	0	67

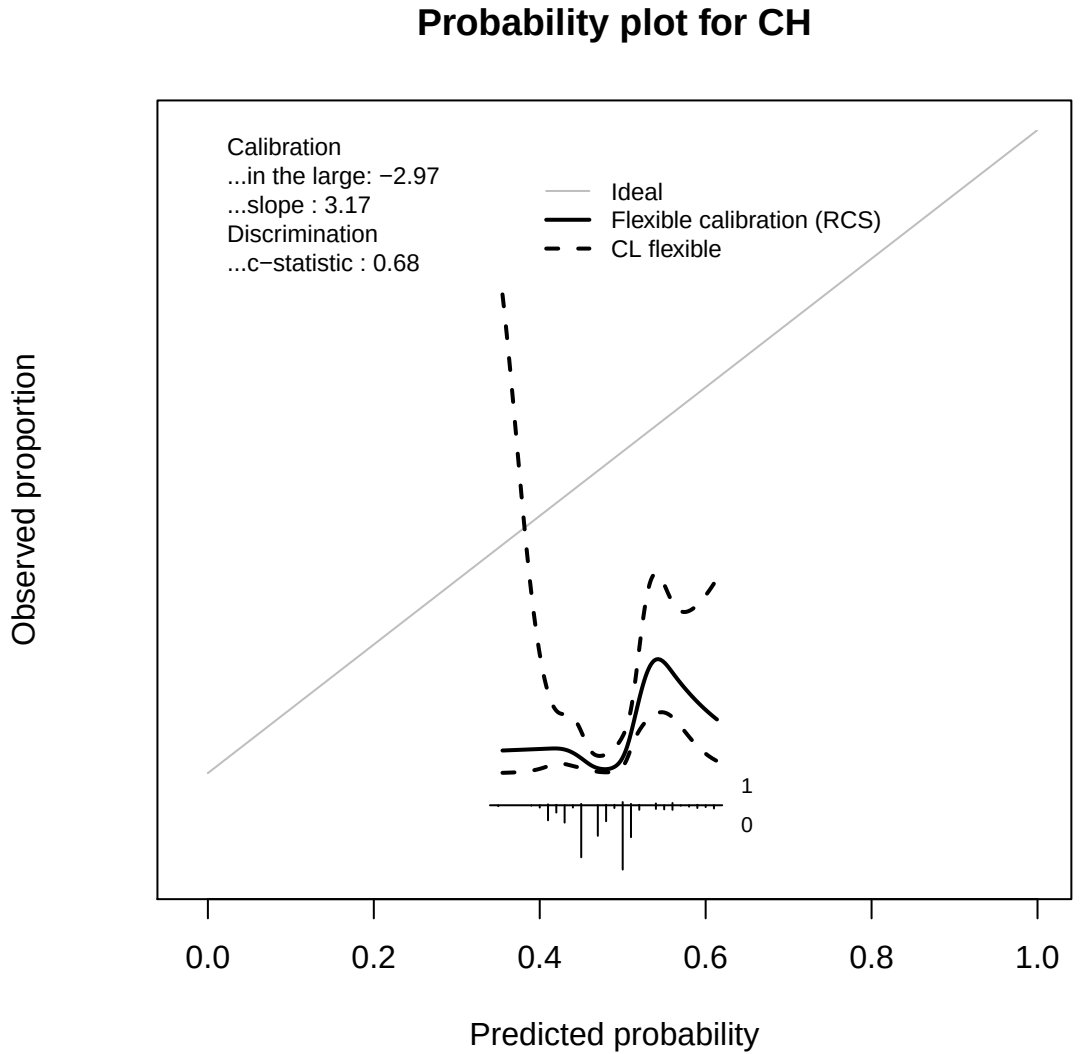
Num	Variable	Positive	Negative
36	adl	0	63
37	country_DK	0	58
38	ph006d6	58	0
39	euro12	56	0
40	numeracy	0	53
41	euro3	0	51
42	ph006d7	49	0
43	country_DE	0	47
44	euro1	47	0
45	country_NL	0	41
46	ph006d11	33	0
47	mobility	31	0
48	euro9	0	26
49	ph006d12	24	1
50	br016_	23	0
51	iadl	0	23
52	country_GR	0	20
53	euro5	0	20
54	ph006d14	1	19
55	euro11	0	16
56	ph006d10	10	6
57	ph006d1	8	7
58	ph006d4	10	5
59	country_SE	0	9
60	euro6	6	3
61	gali	3	4
62	eurodcat	1	4
63	maxgrip	3	2

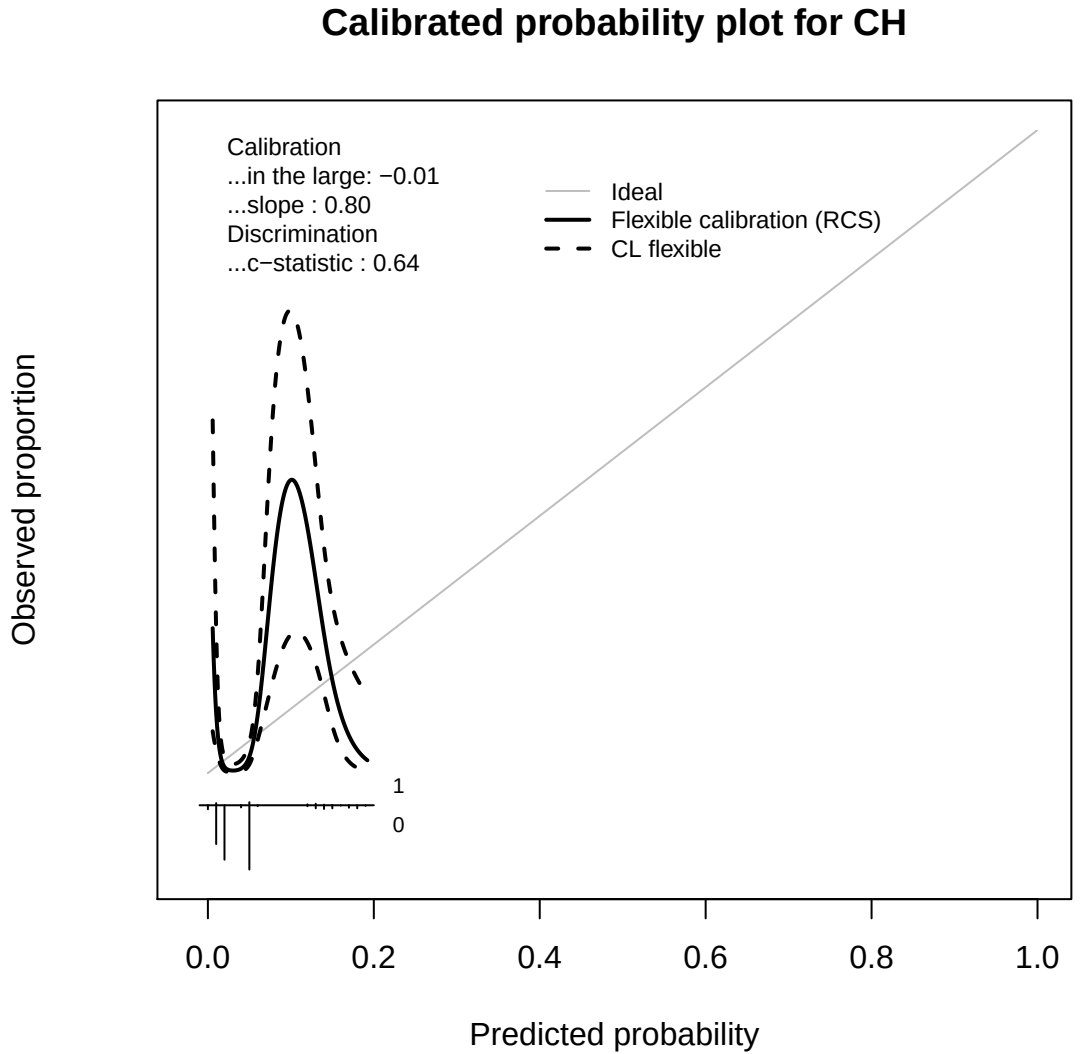

Probability plot for AT
Calibration
...in the large: −2.25
...slope : 1.71
Discrimination
...c−statistic : 0.66
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability


Calibrated probability plot for AT
Calibration
...in the large: −0.00
...slope : 0.85
Discrimination
...c−statistic : 0.67
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability

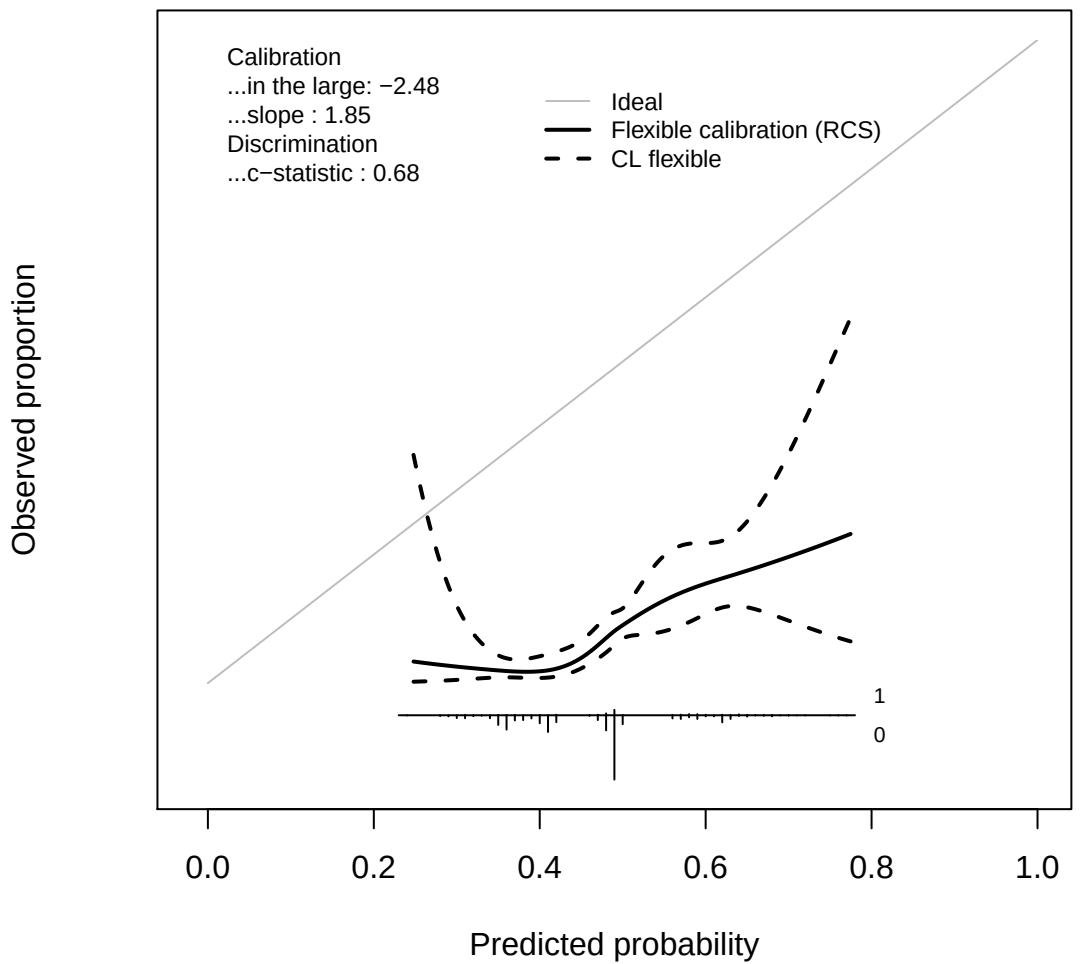

Probability plot for BE
Calibration
...in the large: −2.60
...slope : 2.30
Discrimination
...c−statistic : 0.69
Ideal
Flexible calibration (RCS)
CL flexible
1
0
Observed proportion
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability


Calibrated probability plot for BE
Calibration
...in the large: −0.00
...slope : 0.97
Discrimination
...c−statistic : 0.69
Ideal
Flexible calibration (RCS)
CL flexible
1
0
Observed proportion
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability

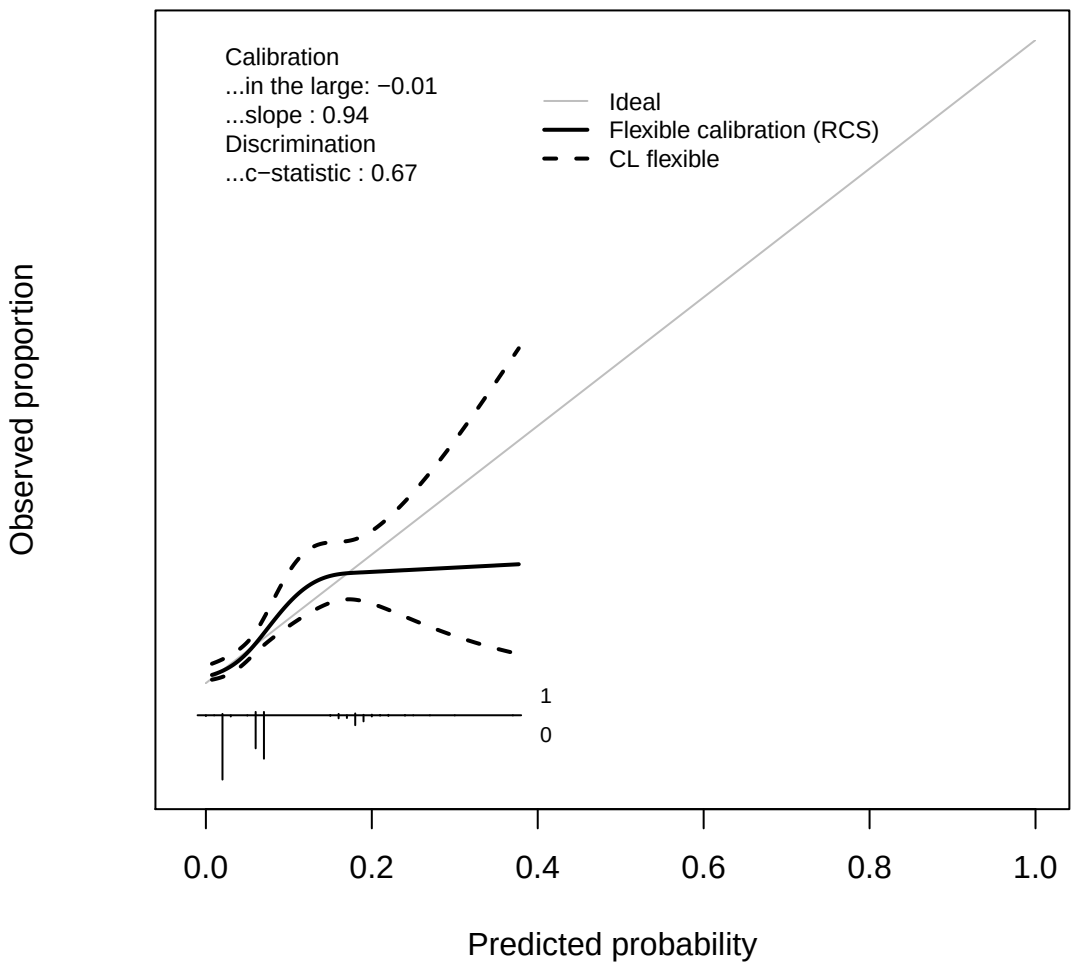

Probability plot for CH
Calibration
...in the large: −2.97
...slope : 3.17
Discrimination
...c−statistic : 0.68
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability

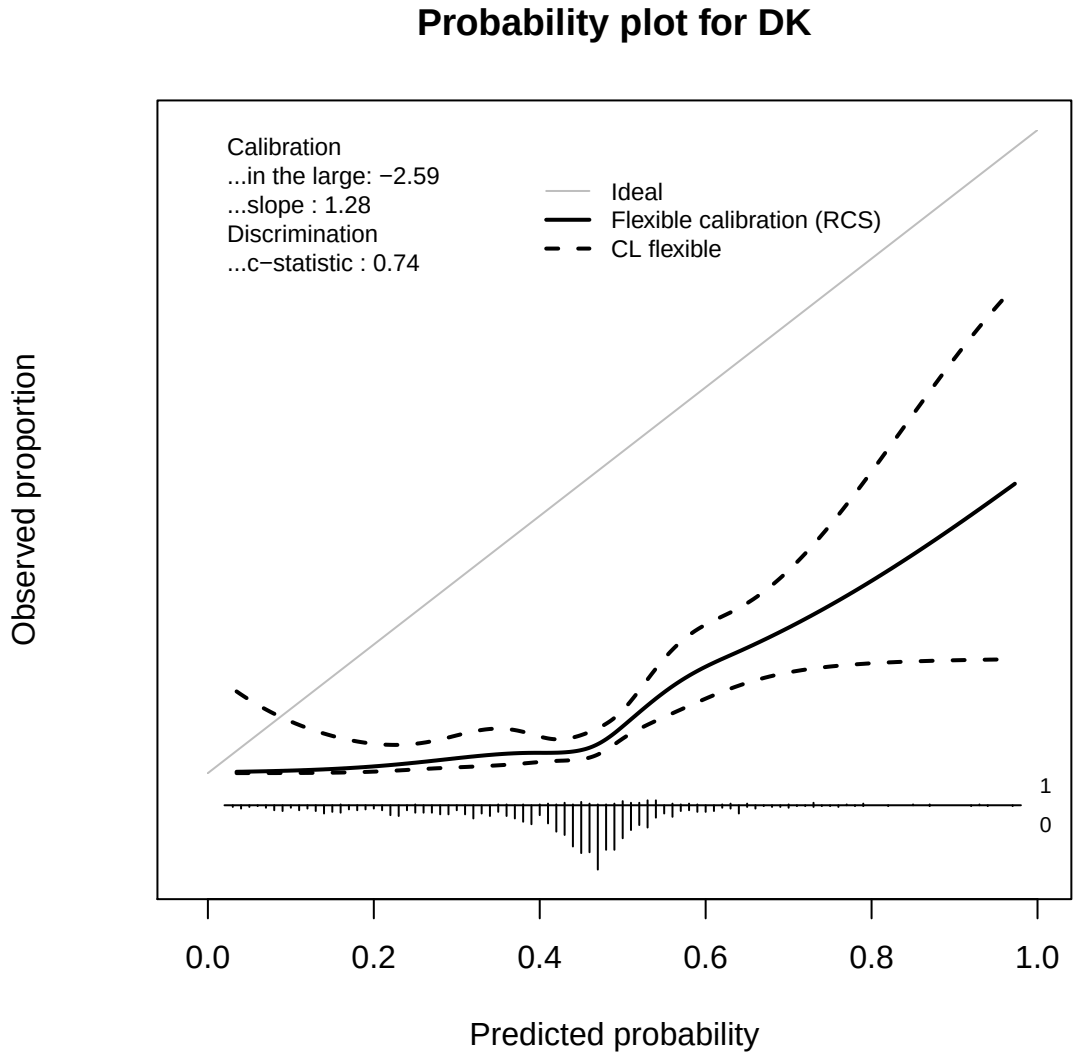

Calibrated probability plot for CH
Calibration
...in the large: −0.01
...slope : 0.80
Discrimination
...c−statistic : 0.64
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability

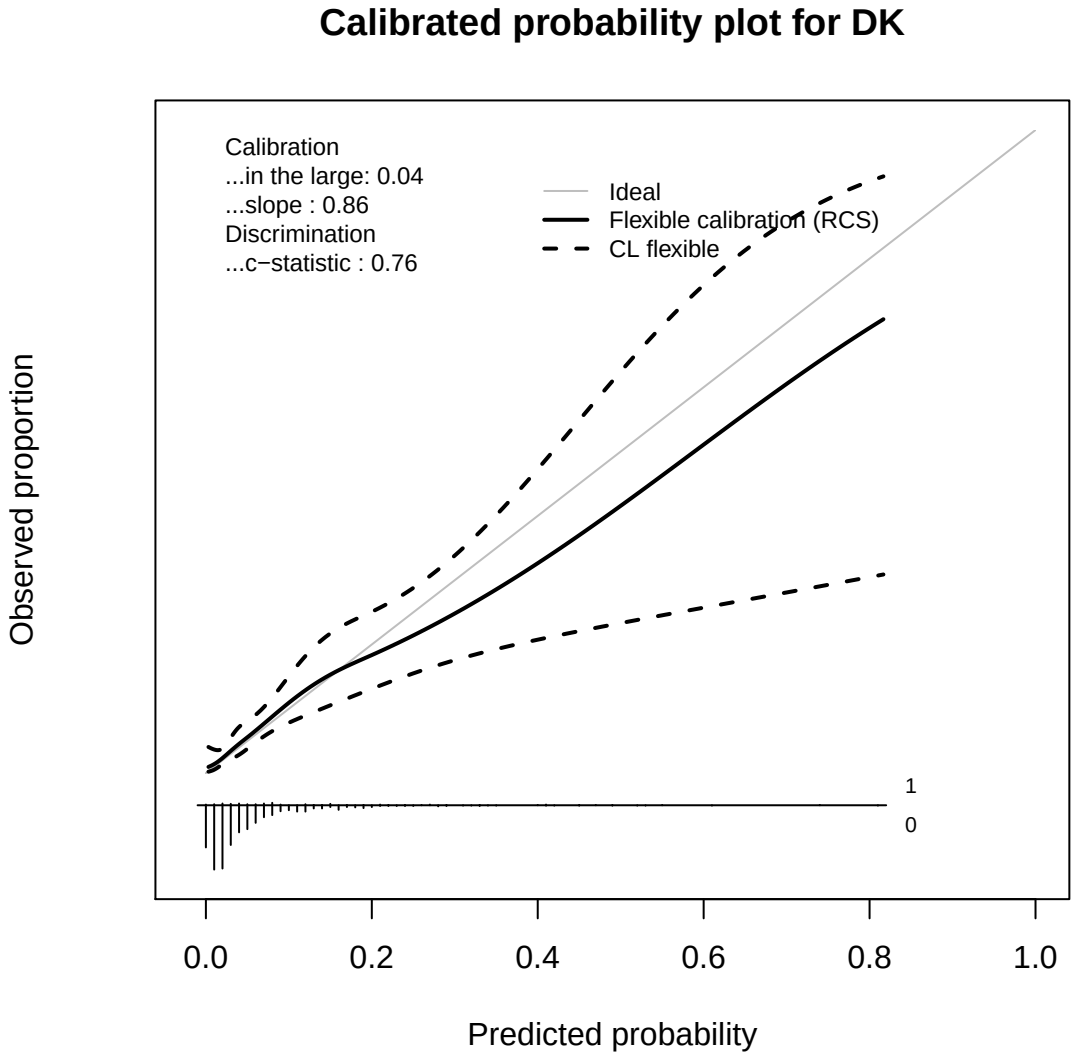
Probability plot for DE

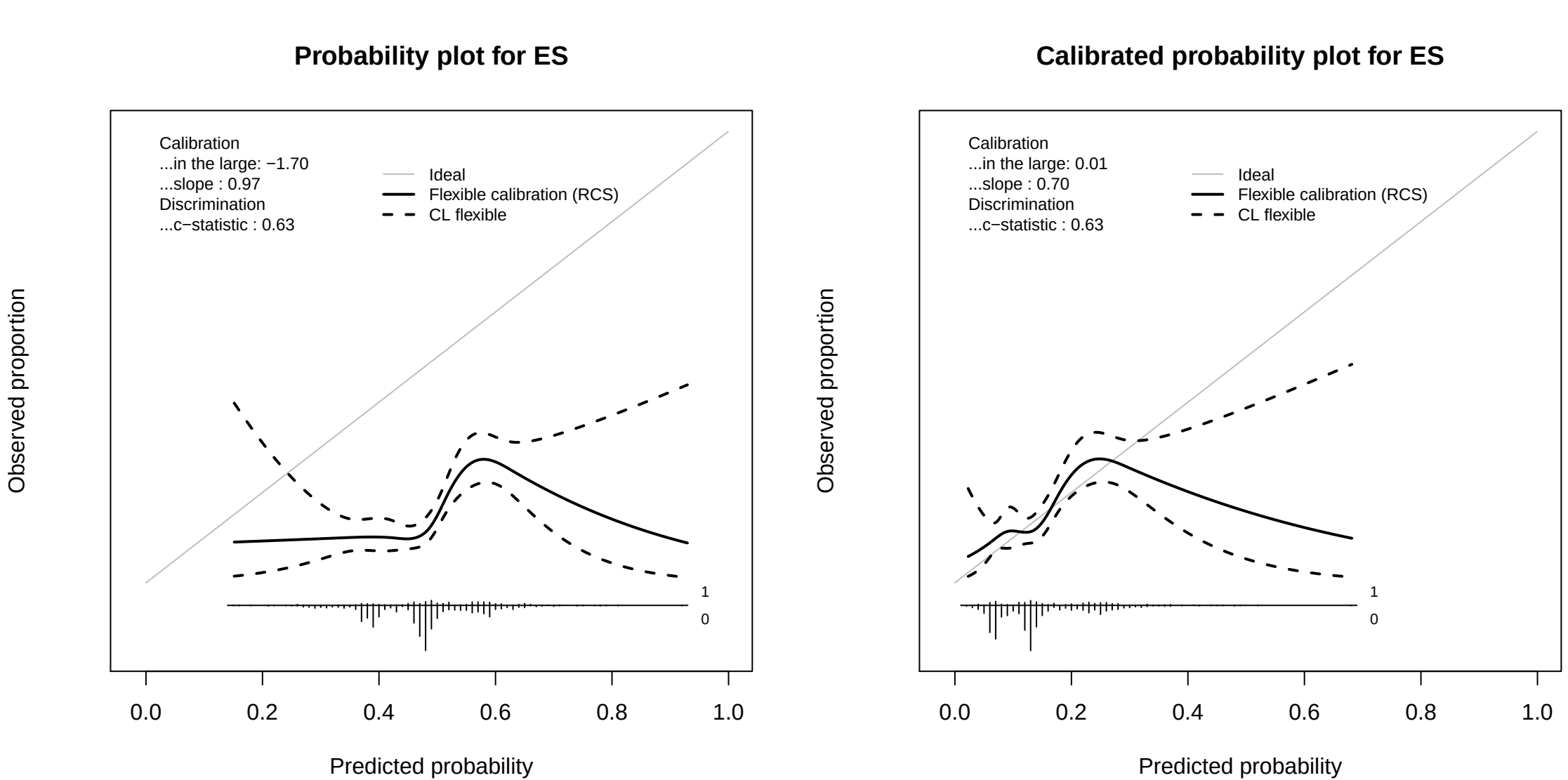


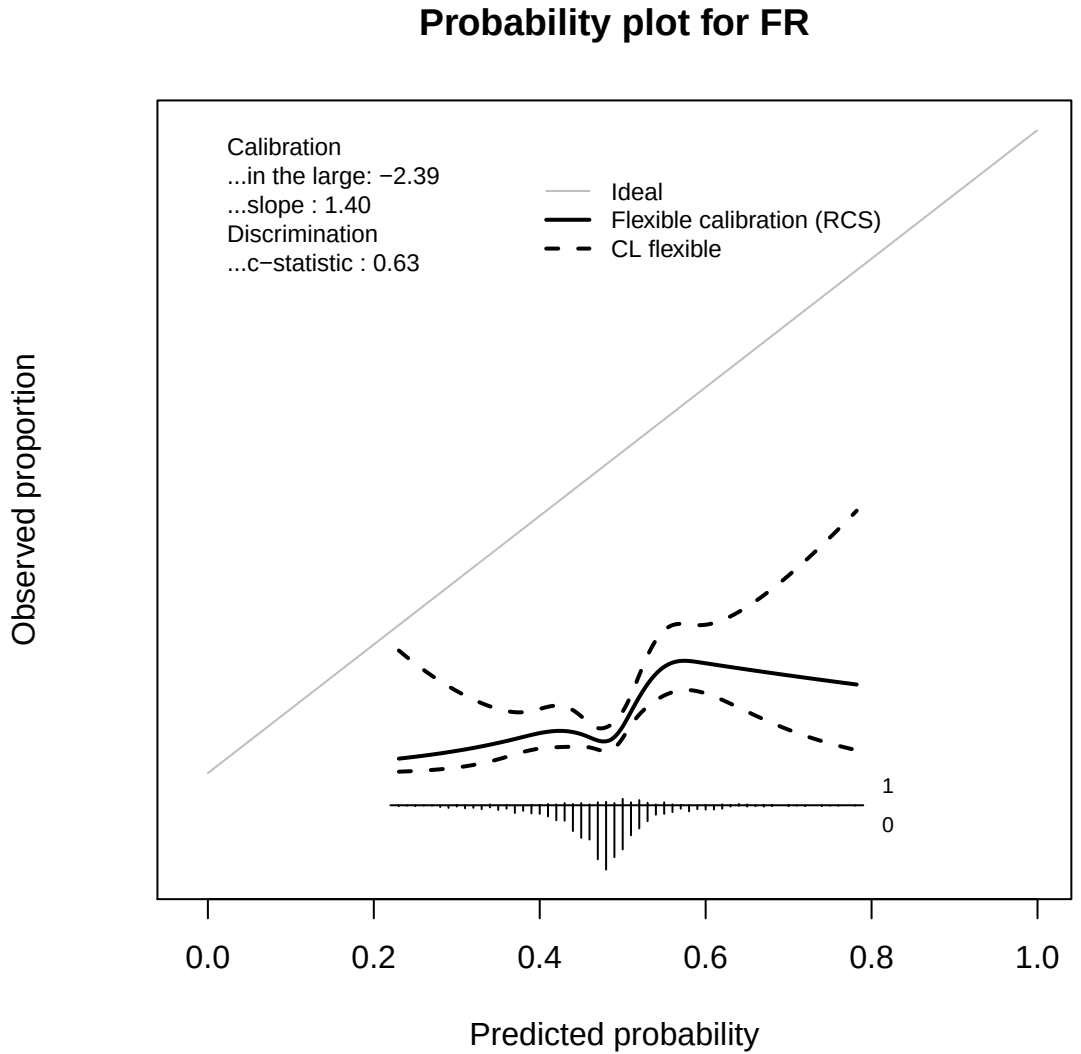
Calibrated probability plot for DE

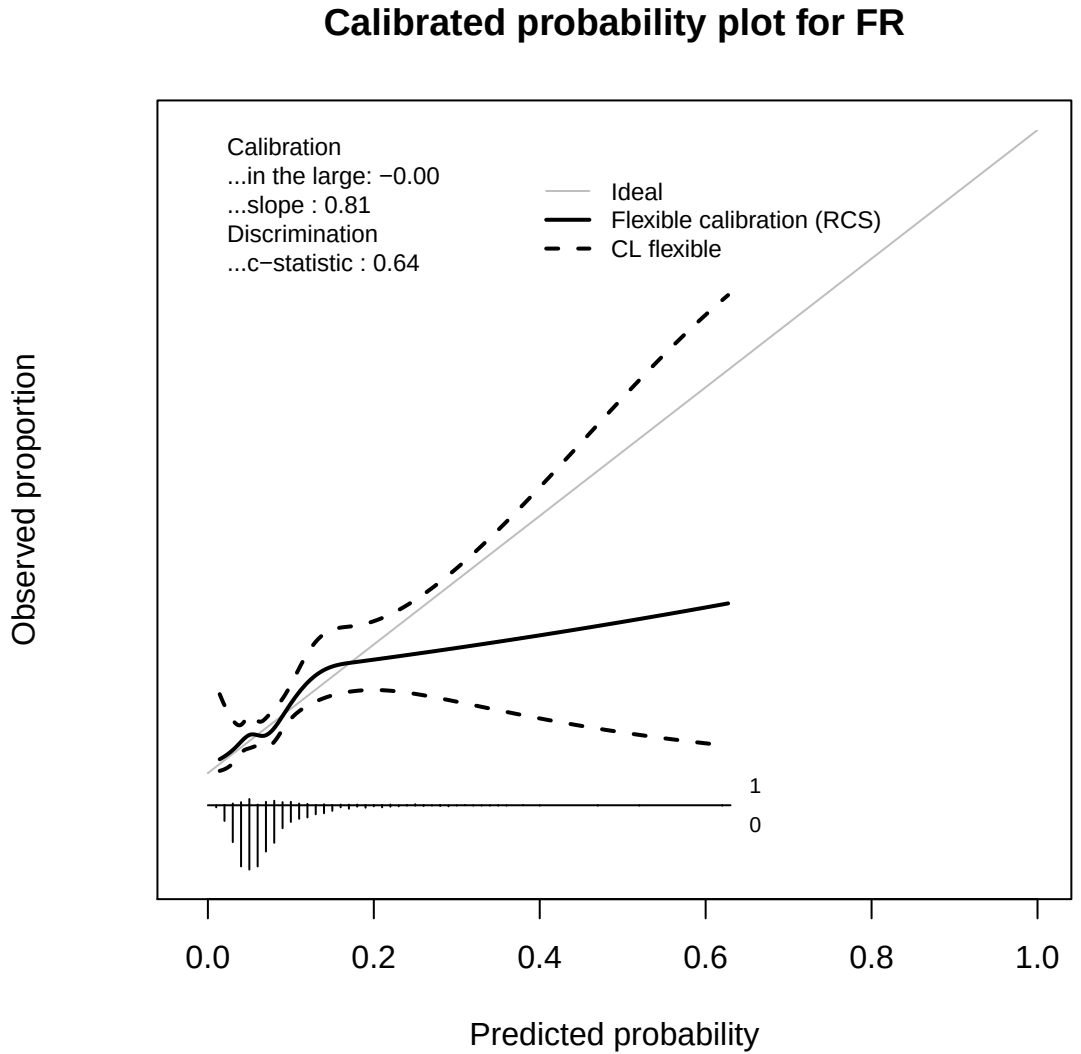


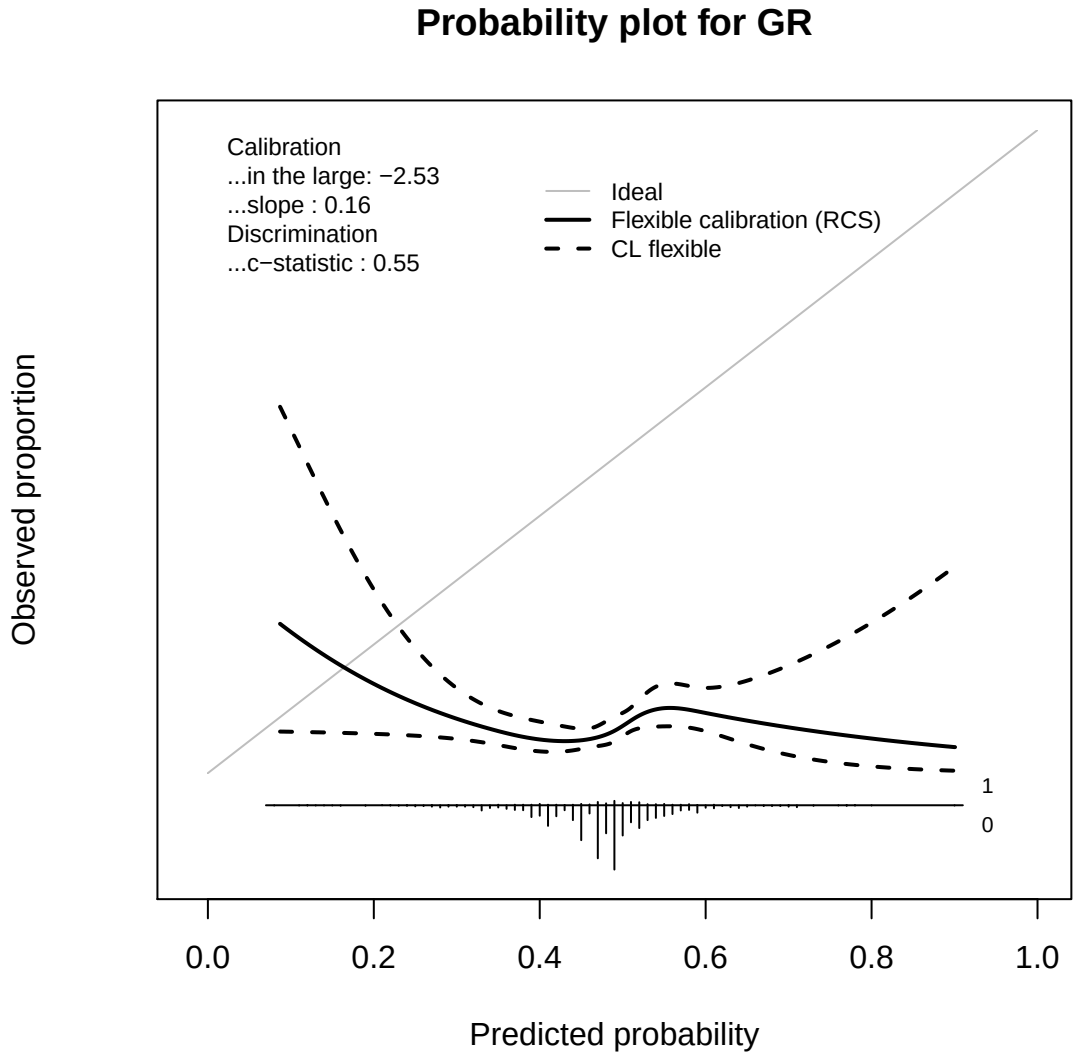

Probability plot for DK
Calibration
...in the large: −2.59
...slope : 1.28
Discrimination
...c−statistic : 0.74
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability

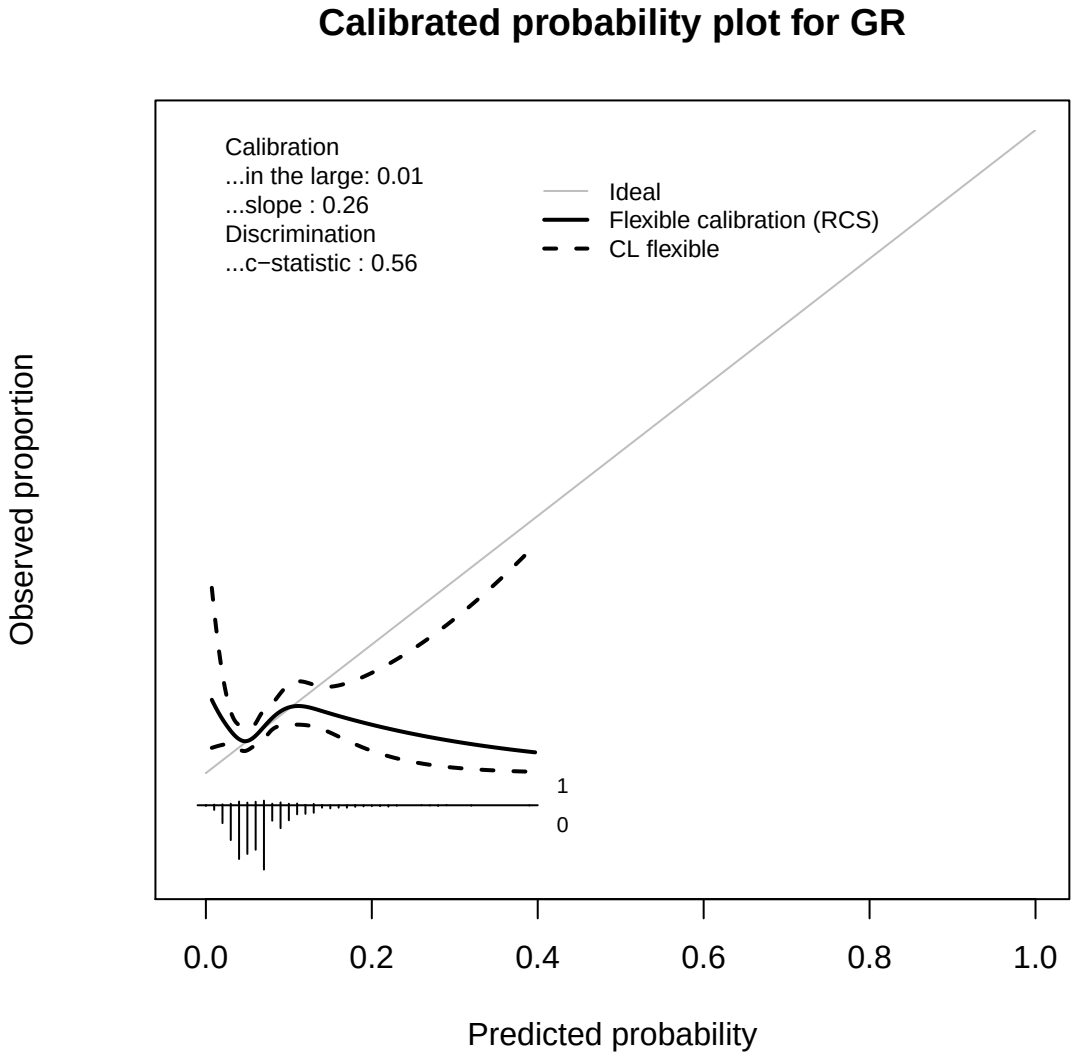

Calibrated probability plot for DK
Calibration
...in the large: 0.04
...slope : 0.86
Discrimination
...c−statistic : 0.76
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability


Probability plot for ES
Calibration
...in the large: −1.70
...slope : 0.97
Discrimination
...c−statistic : 0.63
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
Predicted probability
0.0
0.2
0.4
0.6
0.8
1.0
1
0
Calibrated probability plot for ES
Calibration
...in the large: 0.01
...slope : 0.70
Discrimination
...c−statistic : 0.63
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
Predicted probability
0.0
0.2
0.4
0.6
0.8
1.0
1
0

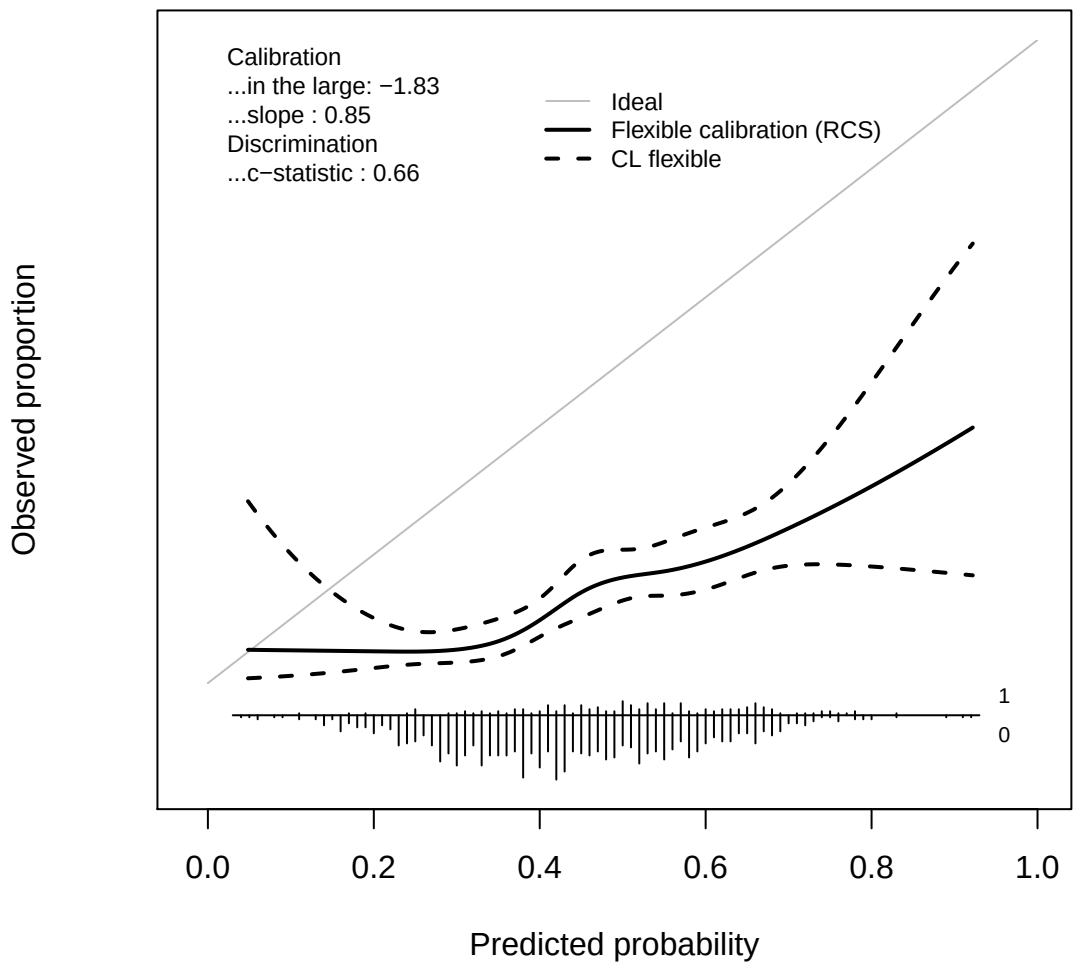

Probability plot for FR
Calibration
...in the large: −2.39
...slope : 1.40
Discrimination
...c−statistic : 0.63
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability


Calibrated probability plot for FR
Calibration
...in the large: −0.00
...slope : 0.81
Discrimination
...c−statistic : 0.64
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability

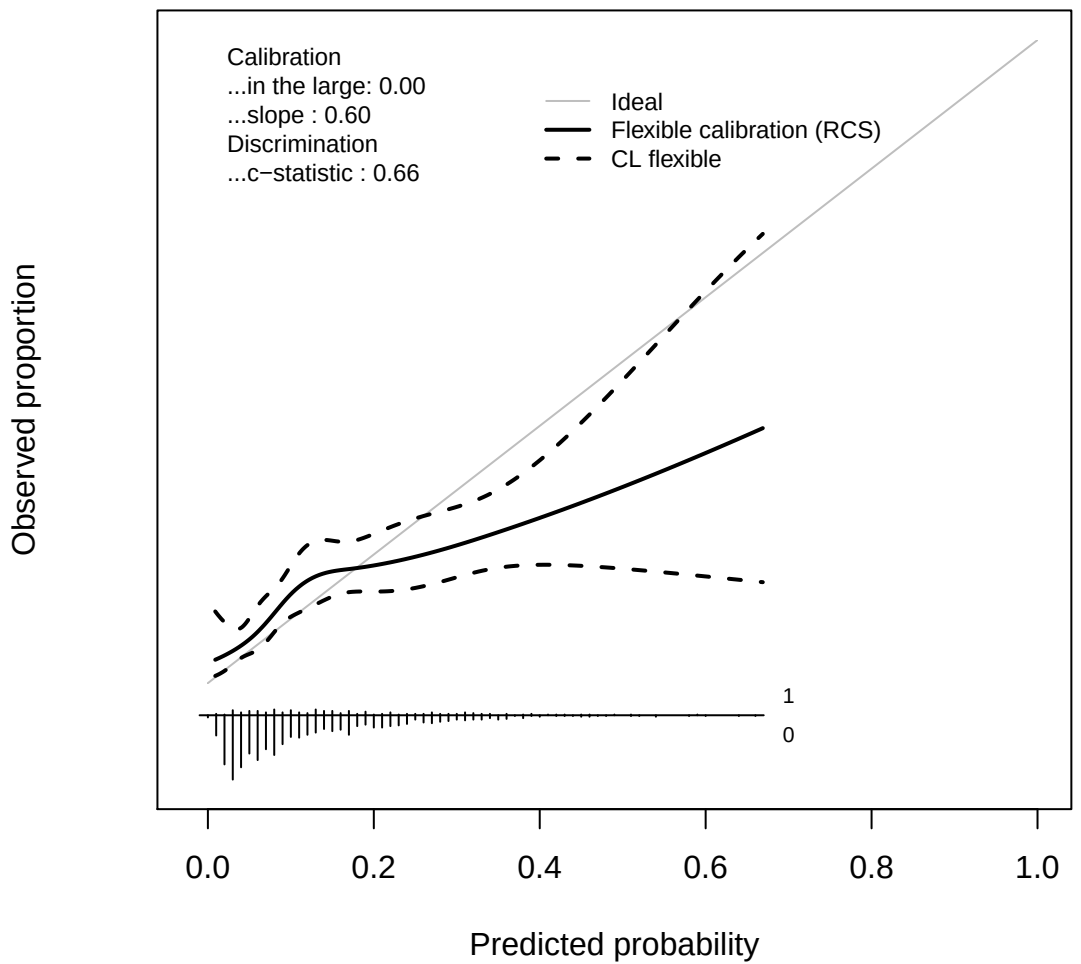

Probability plot for GR
Calibration
...in the large: −2.53
...slope : 0.16
Discrimination
...c−statistic : 0.55
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability

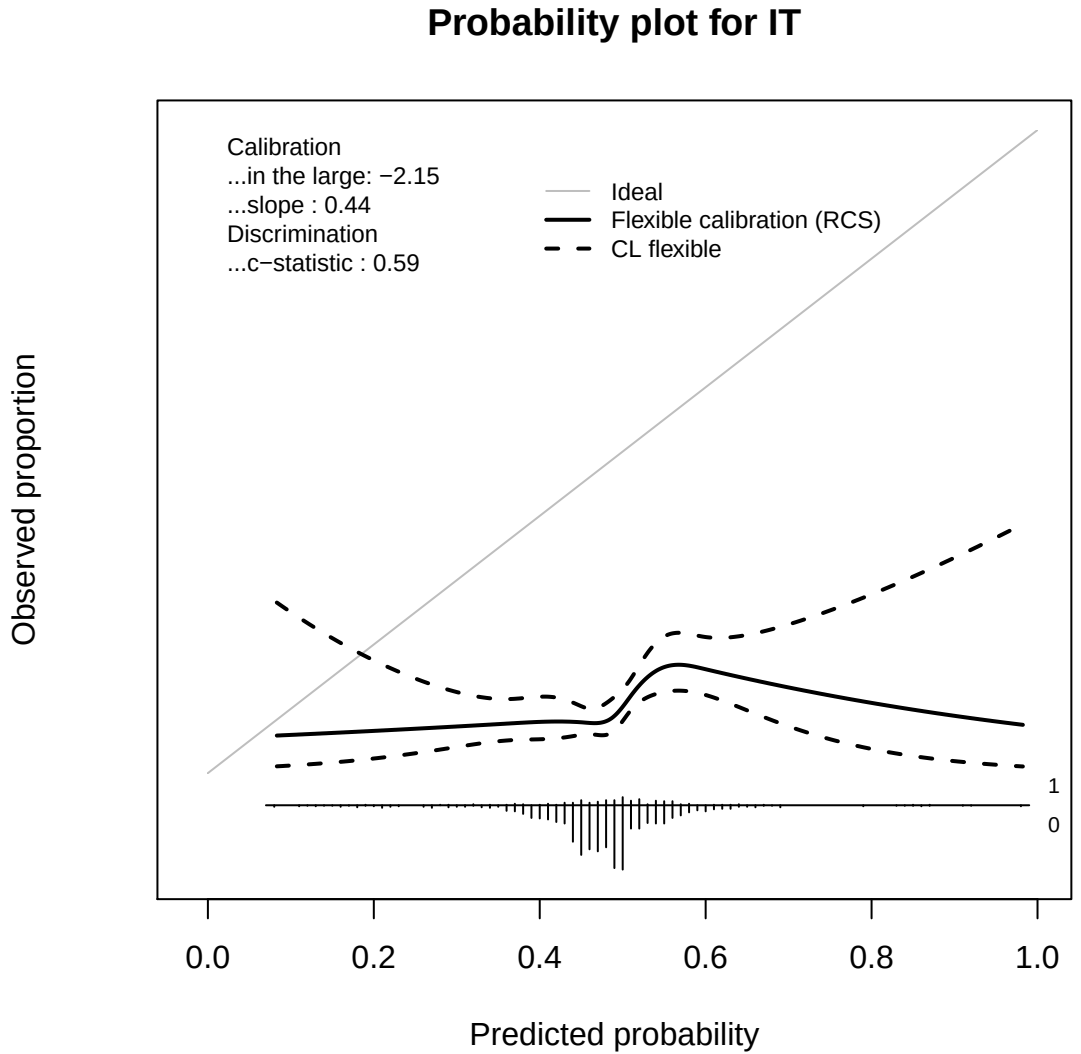

Calibrated probability plot for GR
Calibration
...in the large: 0.01
...slope : 0.26
Discrimination
...c−statistic : 0.56
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability

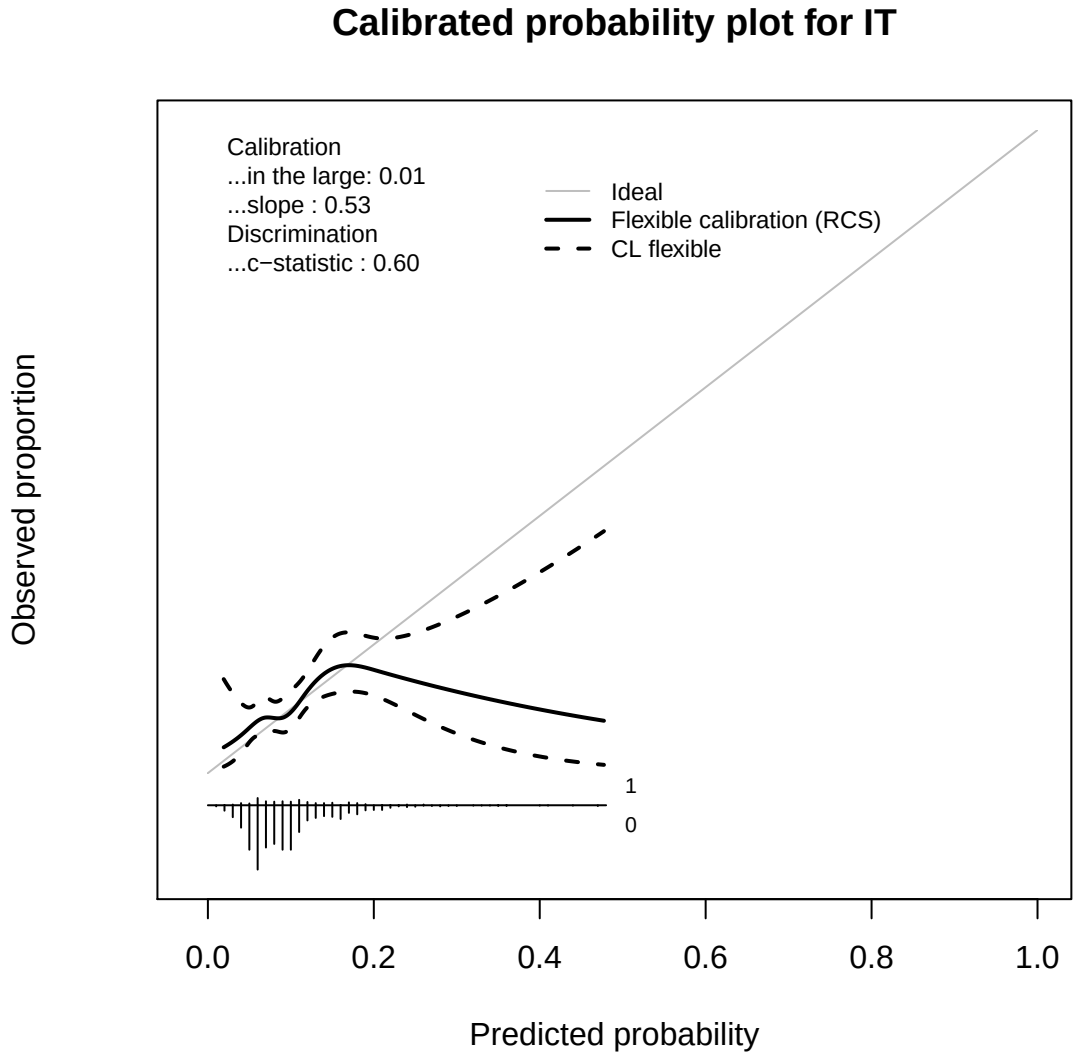
Probability plot for IL

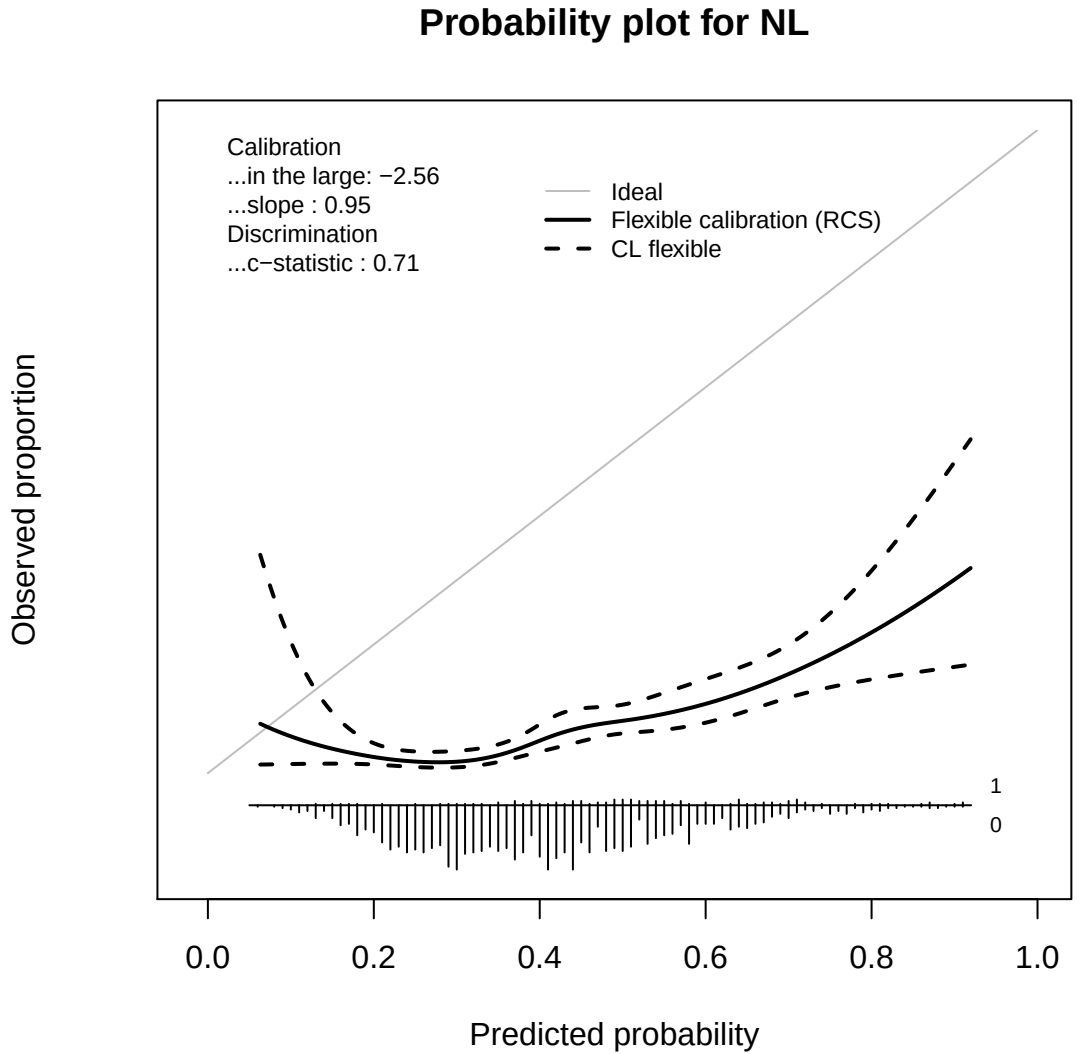


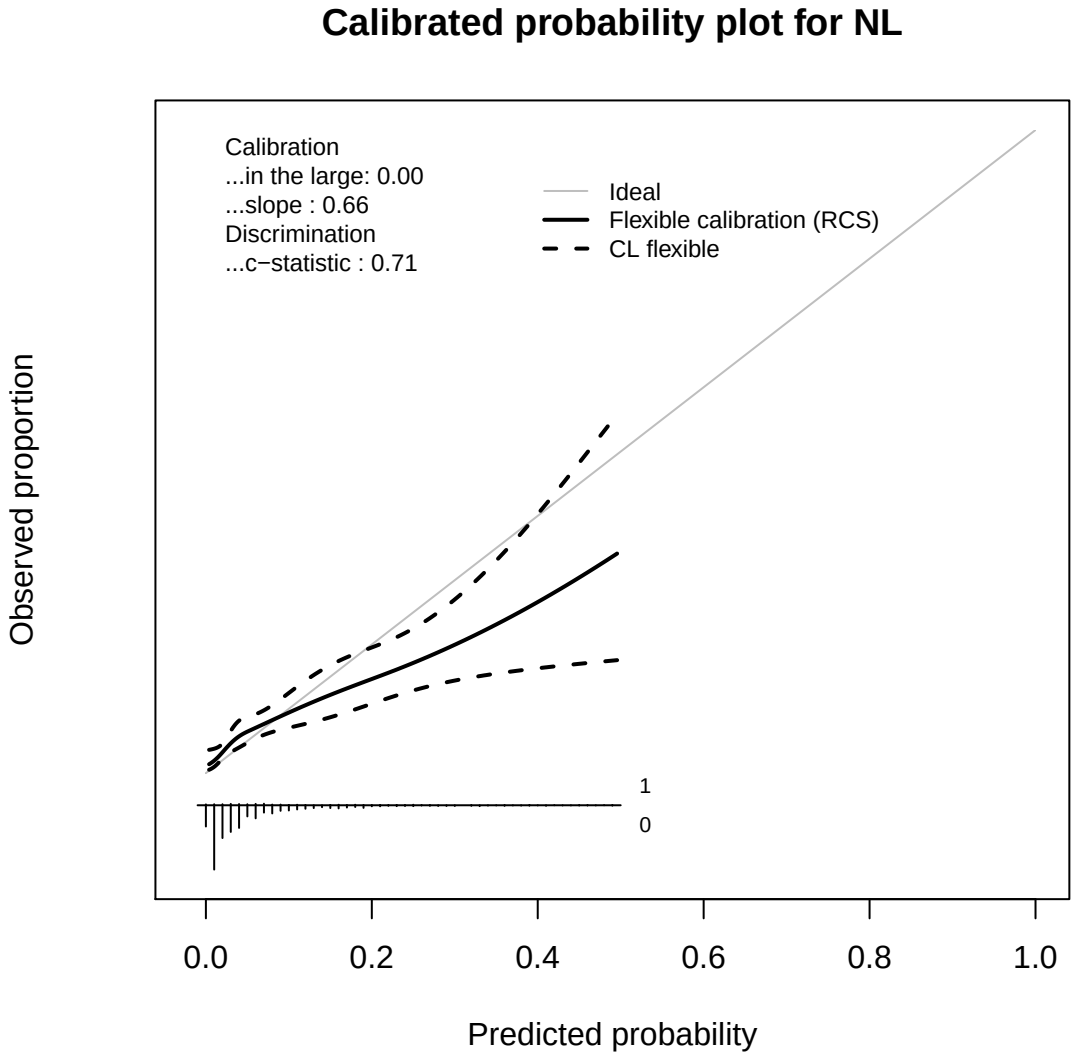
Calibrated probability plot for IL

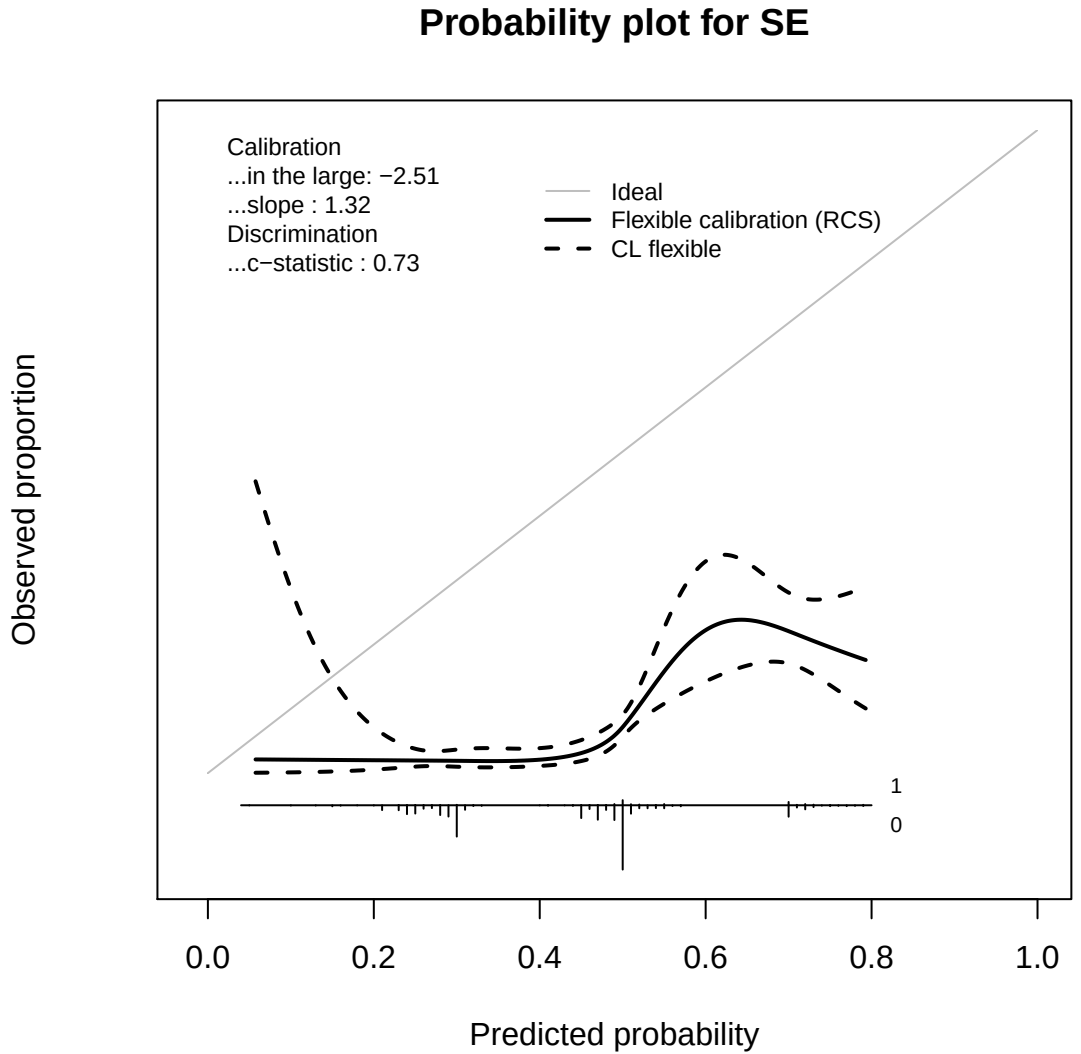


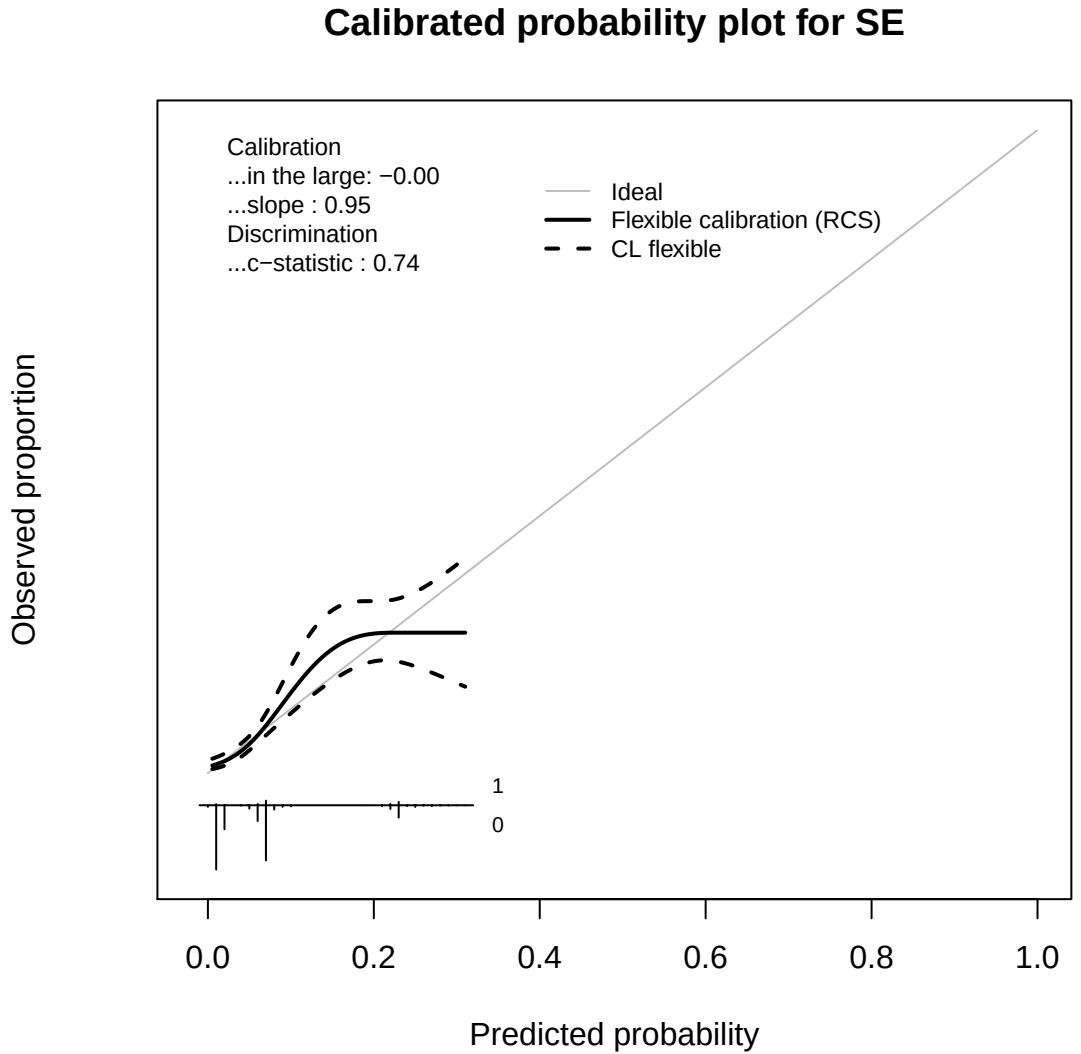

Probability plot for IT
Calibration
...in the large: −2.15
...slope : 0.44
Discrimination
...c−statistic : 0.59
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
Predicted probability
0.0
0.2
0.4
0.6
0.8
1.0
1
0


Calibrated probability plot for IT
Calibration
...in the large: 0.01
...slope : 0.53
Discrimination
...c−statistic : 0.60
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
Predicted probability
0.0
0.2
0.4
0.6
0.8
1.0
1
0


Probability plot for NL
Calibration
...in the large: −2.56
...slope : 0.95
Discrimination
...c−statistic : 0.71
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability


Calibrated probability plot for NL
Calibration
...in the large: 0.00
...slope : 0.66
Discrimination
...c−statistic : 0.71
Ideal
Flexible calibration (RCS)
CL flexible
Observed proportion
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability


Probability plot for SE
Calibration
...in the large: −2.51
...slope : 1.32
Discrimination
...c−statistic : 0.73
Ideal
Flexible calibration (RCS)
CL flexible
1
0
Observed proportion
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability


Calibrated probability plot for SE
Calibration
...in the large: −0.00
...slope : 0.95
Discrimination
...c−statistic : 0.74
Ideal
Flexible calibration (RCS)
CL flexible
1
0
Observed proportion
0.0
0.2
0.4
0.6
0.8
1.0
Predicted probability