

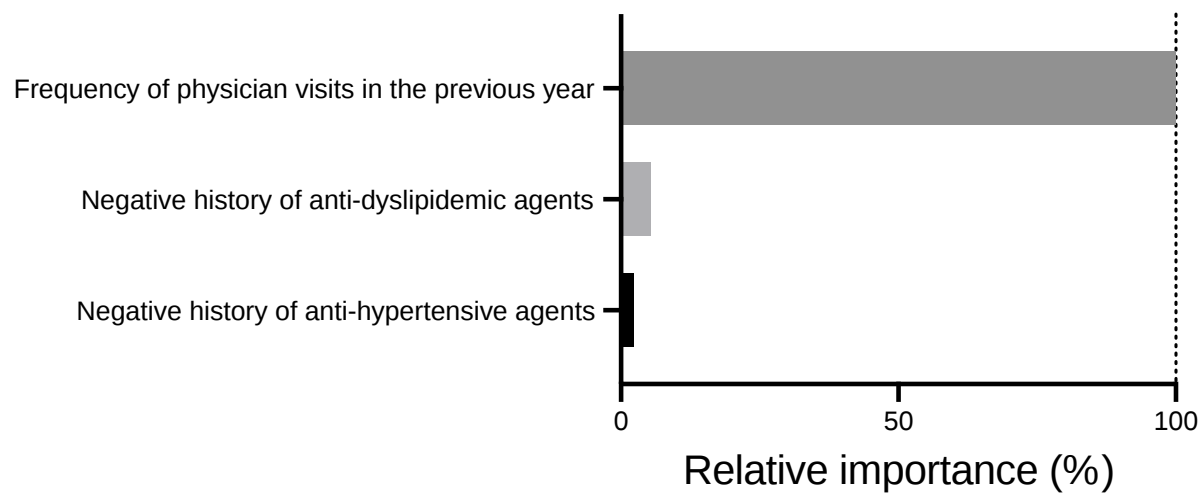
Supplementary Table 1. Results of the previously reported logistic regression model predicting failure to undergo a follow-up visit for diabetes care

Variable	Category	OR	95% CI			P value
Sex	Female	Ref				
	Male	0.94	0.80	–	1.10	0.42
Age	10-year increase	0.87	0.81	–	0.93	<0.001
Body mass index	1 kg/m ² increase	1.02	1.00	–	1.03	0.006
HbA1c, per 1% [11 mmol/mol] decrease		1.20	1.16	–	1.24	<0.001
Glycosuria	Positive	0.99	0.88	–	1.12	0.88
Proteinuria	Positive	0.95	0.83	–	1.09	0.48
Negative history of anti-hypertensive agents		2.50	2.20	–	2.83	<0.001
Negative history of anti-dyslipidemic agents		2.63	2.24	–	3.07	<0.001
Anti-depression prescription		0.67	0.51	–	0.89	0.006
Smoking status	Non-smoker	Ref				
	Smoker	1.21	1.09	–	1.33	<0.001
Liquor intake frequency	Rarely	Ref				
	Occasionally	1.03	0.92	–	1.14	0.65
	Regularly	1.17	1.04	–	1.32	0.010
Hemoglobin (g/dL)		1.02	0.98	–	1.06	0.36
Insured person	Identical person	Ref				
	Dependent	0.85	0.75	–	0.97	0.02

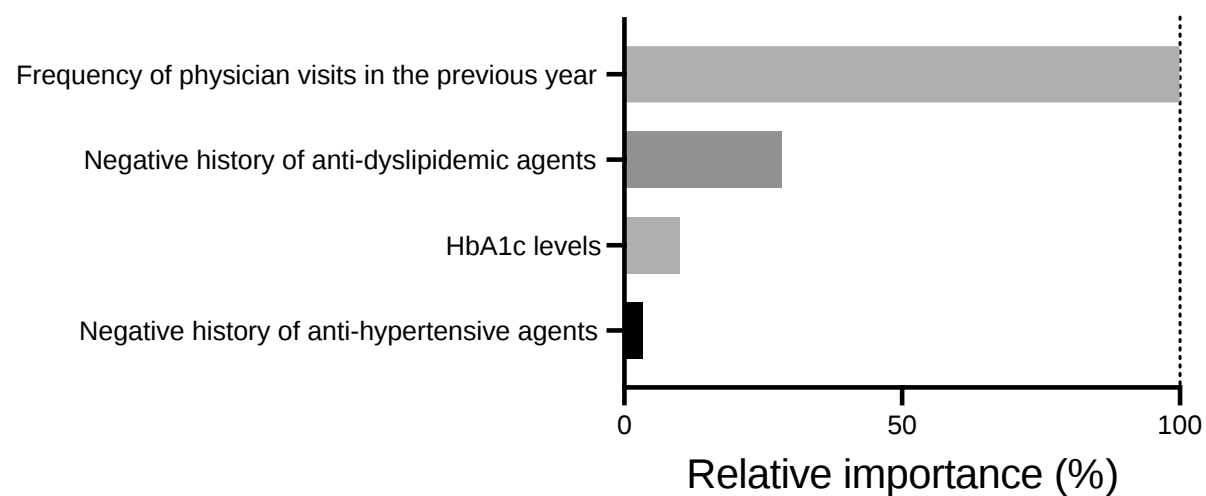
OR, odds ratio; CI, confidence interval

Supplementary Figure 1. Variable importance derived from the Lasso regression model for failure to attend a follow-up visit for diabetes care when stratified by HbA1c levels

(A) Stratified analysis: variable importance in patients with HbA1c of 6.5-7.9% [48-63 mmol/mol]

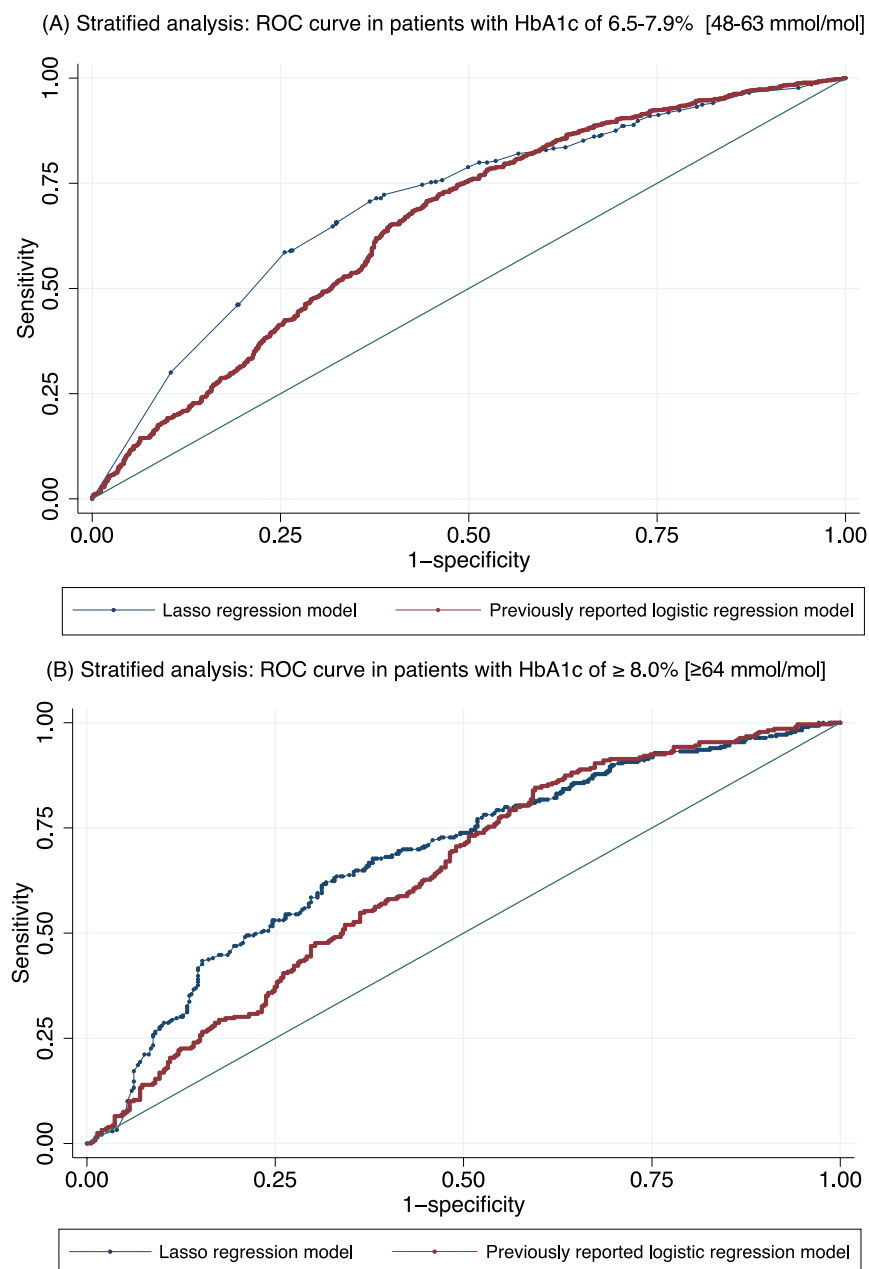


(B) Stratified analysis: variable importance in patients with HbA1c of $\geq 8.0\%$ [≥ 64 mmol/mol]



The variable importance in each stratum is a measure scaled to a maximum value of 100.

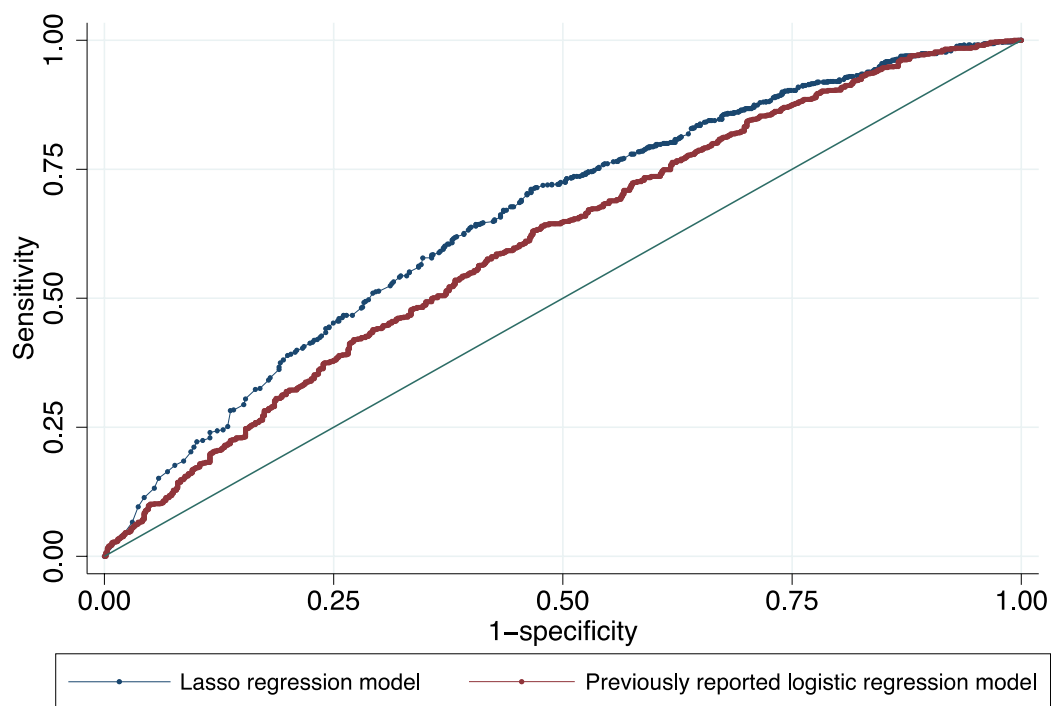
Supplementary Figure 2. ROC curves in the Lasso regression and previously reported logistic regression models when stratified by HbA1c levels



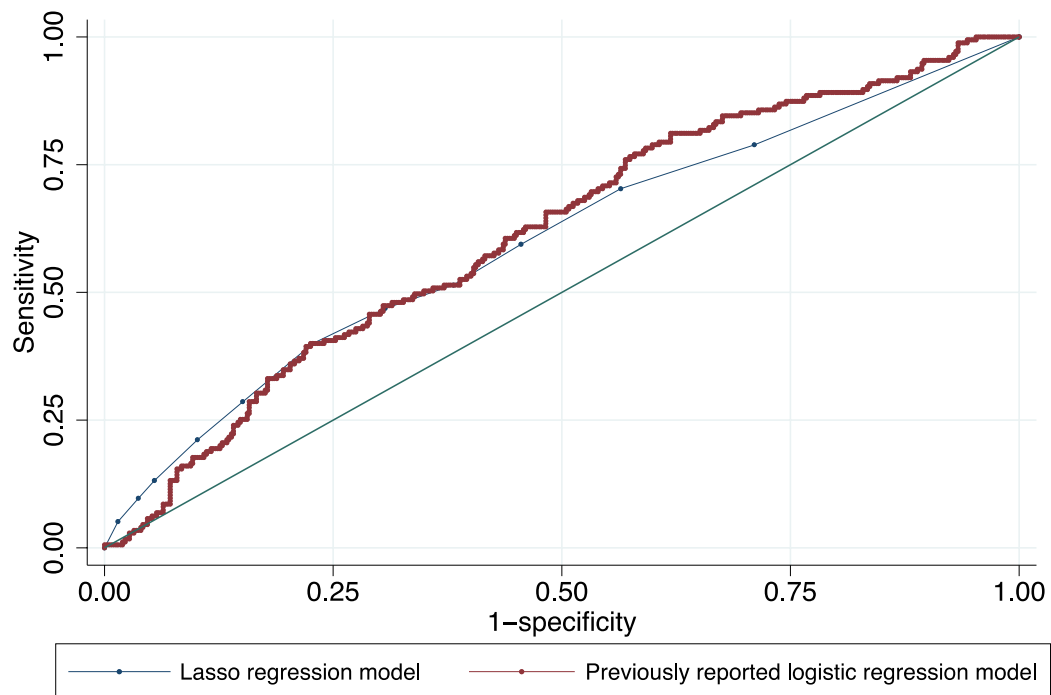
Prediction performances of the Lasso regression and previously reported logistic regression models for failure to attend a follow-up visit for diabetes care within six months after being recommended to consult a physician in an analysis stratified by HbA1c levels. The receiver operating characteristic (ROC) curves for predicting failure to attend a follow-up visit are shown. (A) The Lasso regression model had a better discrimination ability (C-statistic: 0.70 in the Lasso regression model vs. 0.66 in the previously reported logistic regression model) in those whose HbA1c levels were 6.5–7.9% [48–63 mmol/mol]. (B) The Lasso regression model had a better discrimination ability in those whose HbA1c levels were $\geq 8.0\%$ [≥ 64 mmol/mol] (C-statistic: 0.69 in the Lasso regression model vs 0.64 in the previously reported logistic regression model).

Supplementary Figure 3 ROC curves in the Lasso regression and the previously reported logistic regression models when stratified by history of receiving either an anti-hypertensive, anti-dyslipidemic, or anti-hyperuricemic agent

(A) Stratified analysis: ROC curve in patients without history of anti-hypertensive, anti-dyslipidemia, or anti-hyperuricemia prescription



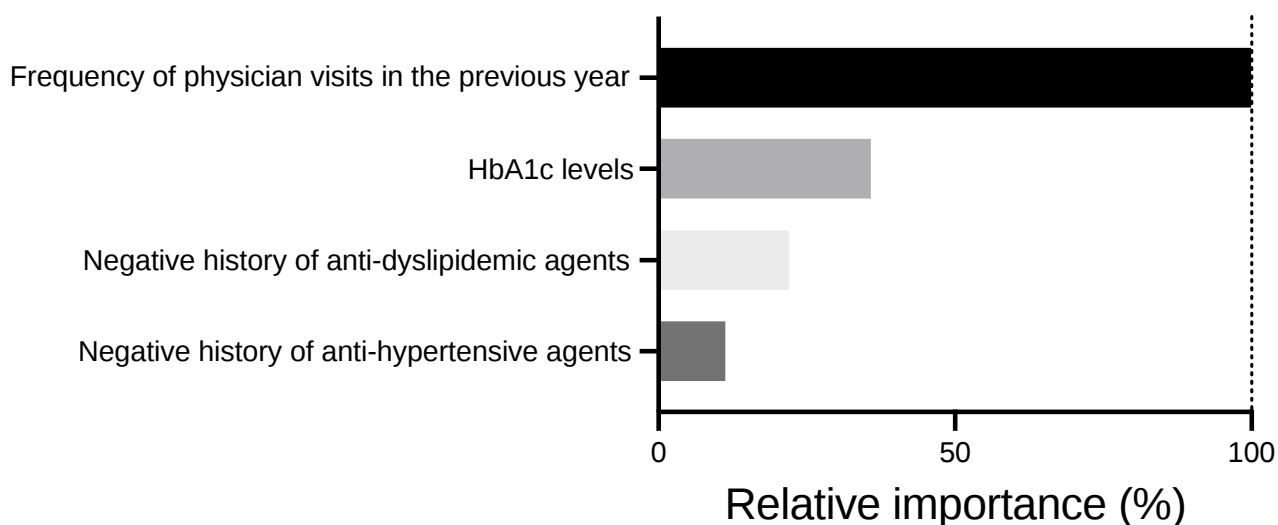
(B) Stratified analysis: ROC curve in patients with history of anti-hypertensive, anti-dyslipidemia, or anti-hyperuricemia prescription



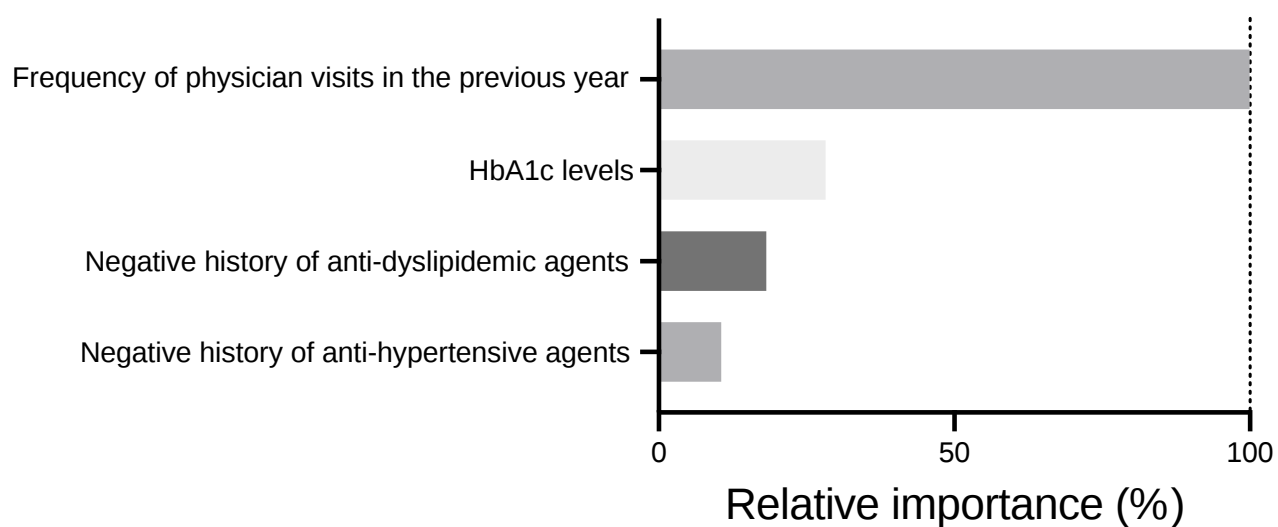
Prediction performances of the Lasso regression and the previously reported logistic regression models for failure to attend a follow-up visit for diabetes care within six months after being recommended to consult a physician, in an analysis stratified by administration of anti-hypertensive, anti-dyslipidemic, or anti-hyperuricemic agents. The ROC curves for predicting failure to attend a follow-up visit are shown. (A) The Lasso regression model had a better discrimination ability (C-statistic: 0.66, Lasso regression model vs. 0.61, previously reported logistic regression model) in those who did not receive either agent. (B) The Lasso regression model had a comparable discrimination ability in those receiving either agent (C-statistic: 0.60, Lasso regression model vs. 0.61, previously reported logistic regression model).

Supplementary Figure 4. Variable importance derived from the Lasso regression model for failure to attend a follow-up visit for diabetes care when the observation period was set at three and nine months

(A) Sensitivity analysis: variable importance when the observation period is set at 3 months



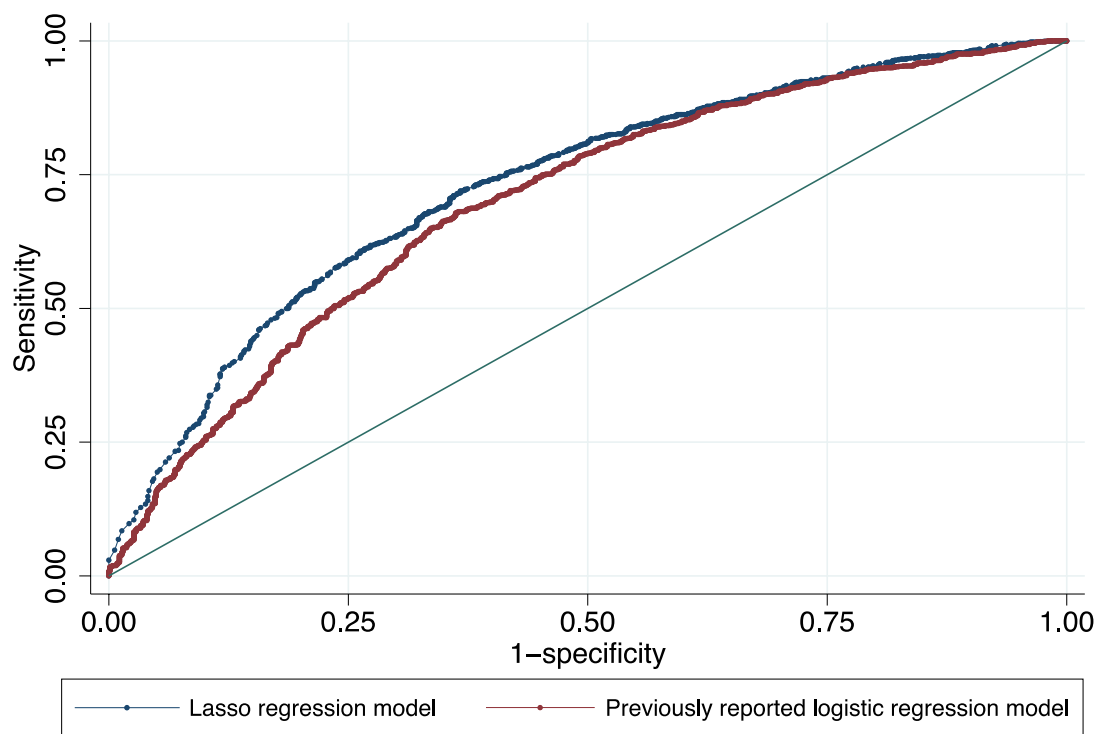
(B) Sensitivity analysis: variable importance when the observation period is set at 9 months



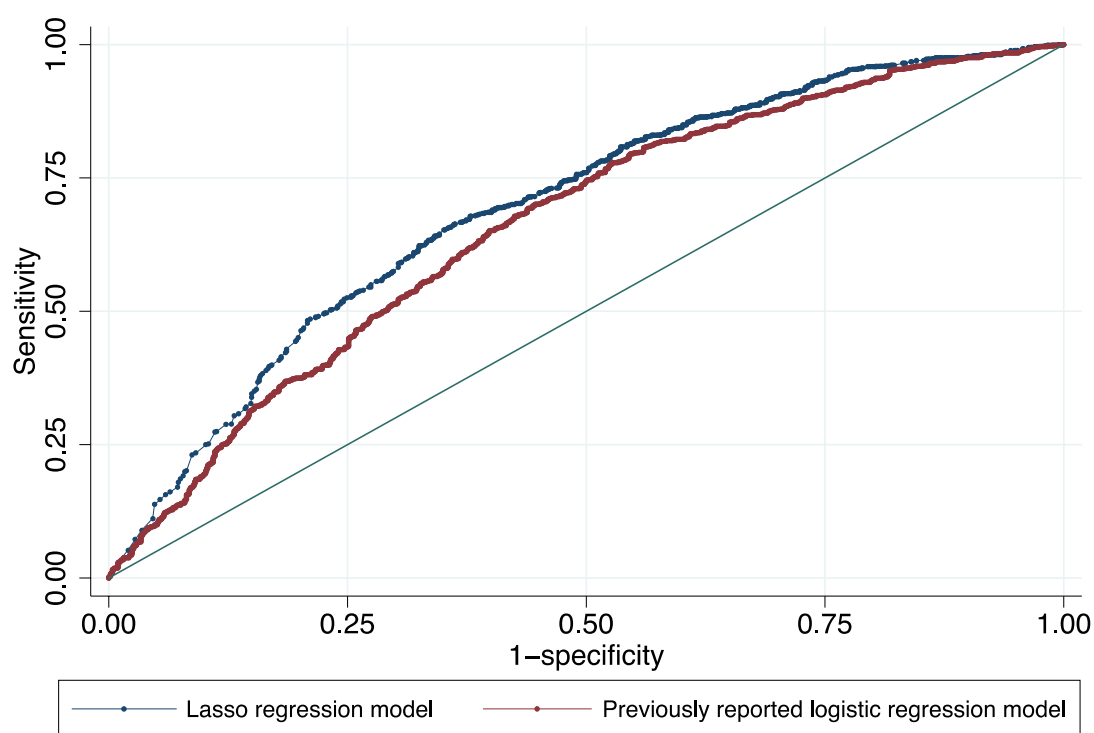
The variable importance in each stratum is a measure scaled to have a maximum value of 100

Supplementary Figure 5 ROC curves in the Lasso regression and previously reported logistic regression models when the observation period was set at three and nine months

(A) Sensitivity analysis: ROC curves when the observation period is set at 3 months



(B) Sensitivity analysis: ROC curves when the observation period is set at 9 months



Prediction performances of the previously reported logistic regression model and Lasso regression model for failure to attend a follow-up visit for diabetes care within three or nine months after being recommended to consult a physician, as a sensitivity analysis. The receiver operating characteristic ROC curves for predicting failure to attend a follow-up visit are shown. (A) In the model with an observation period of three months, the Lasso regression model had a better discrimination ability (C-statistic: 0.73 in the Lasso regression model vs 0.70 in the previously reported logistic regression model). (B) In the model with an observation period of nine months, the Lasso regression model had a slightly better discrimination ability (C-statistic: 0.69 in the Lasso regression model vs. 0.66 in the previously reported logistic regression model).