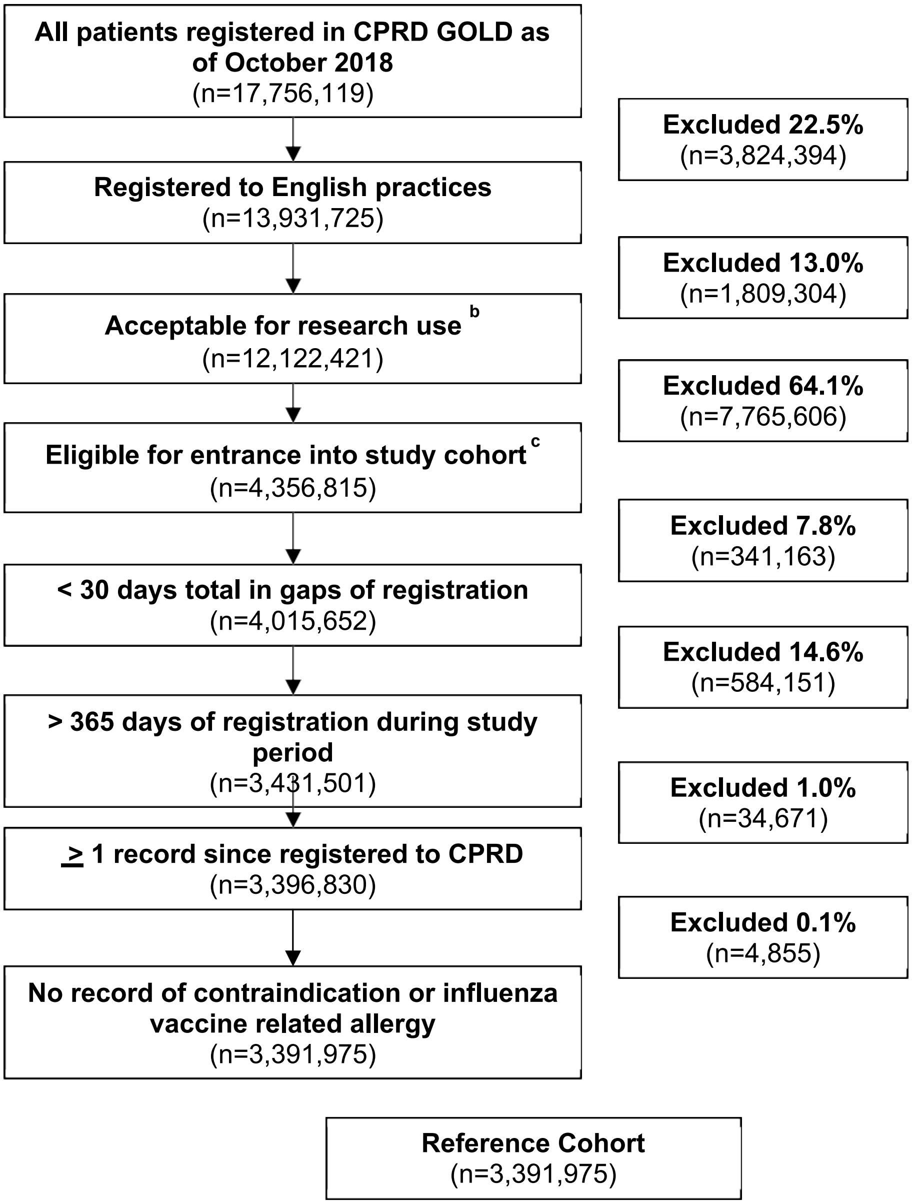
Supplementary Online Content

Loiacono MM, Mitsakakis N, Kwong JC, et al. Development and validation of a clinical prediction tool for seasonal influenza vaccination in England. *JAMA Netw Open*. 2020; 3(6):e207743. doi:10.1001/jamanetworkopen.2020.7743.

eFigure 1. Overview of the selection of the reference cohort [a].



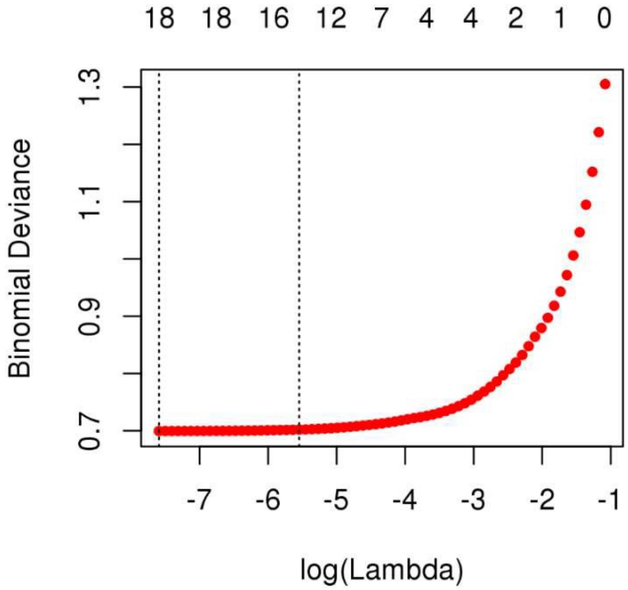
[a] As described in Loiacono, MM, et al. (2020)[1]

[b] "Acceptable" is a variable determined by CPRD that deems the quality of a patient's records high enough for research use

[c] Eligibility for entrance into study cohort consists of:

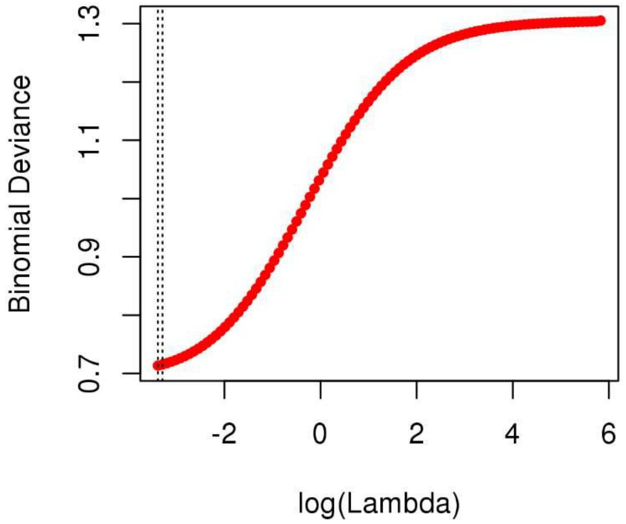
- Sex is "Male" or "Female" only (excluded "Indeterminate")
- 18+ years old within study period
- Patient's first registration date is prior to December 31, 2016
- Patient's death or transfer-out date is after January 1, 2011
- Practice's up-to-standard date is prior to or within the study period
- Practice's last collection date is within or after the study period
- Final enrollment date must be before December 31, 2016
- Final disenrollment date must be after January 1, 2011

eFigure 2. Binomial deviance versus log(Lambda) for 10-fold within-sample cross-validation of LASSO model among patients aged 18-64 years.



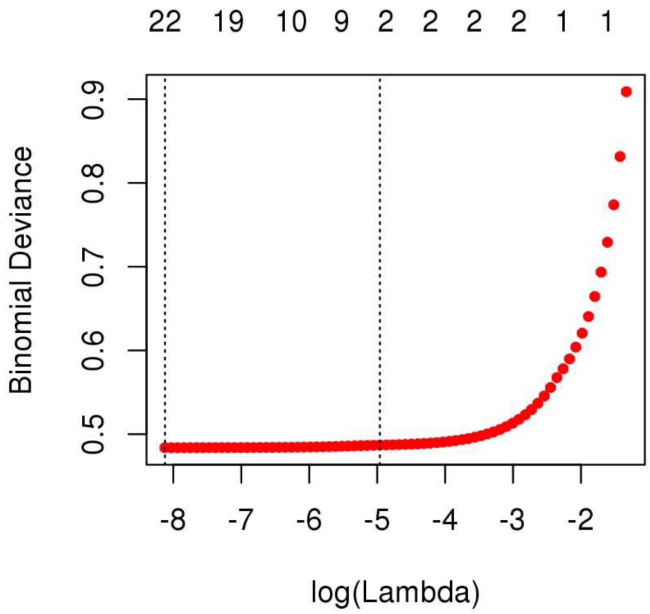
Note: First dotted vertical line indicates the value of log(Lambda) where deviance is minimized, while the second dotted vertical line indicates the value of log(Lambda) where deviance is within 1 standard error of the minimum. Values on top of figure indicate the total number of features included in the model for the specified value of log(Lambda).

eFigure 3. Binomial deviance versus log(Lambda) for 10-fold within-sample cross-validation of Ridge model among patients aged 18-64 years.



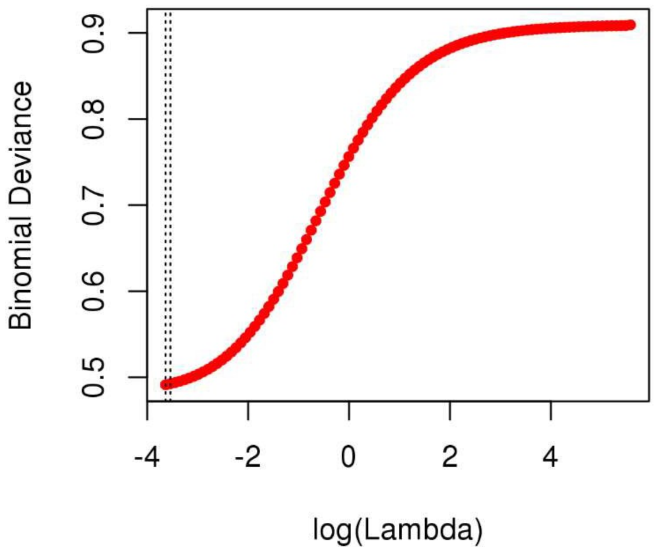
Note: First dotted vertical line indicates the value of log(Lambda) where deviance is minimized, while the second dotted vertical line indicates the value of log(Lambda) where deviance is within 1 standard error of the minimum.

eFigure 4. Binomial deviance versus log(Lambda) for 10-fold within-sample cross-validation of LASSO model among patients aged 65+ years.



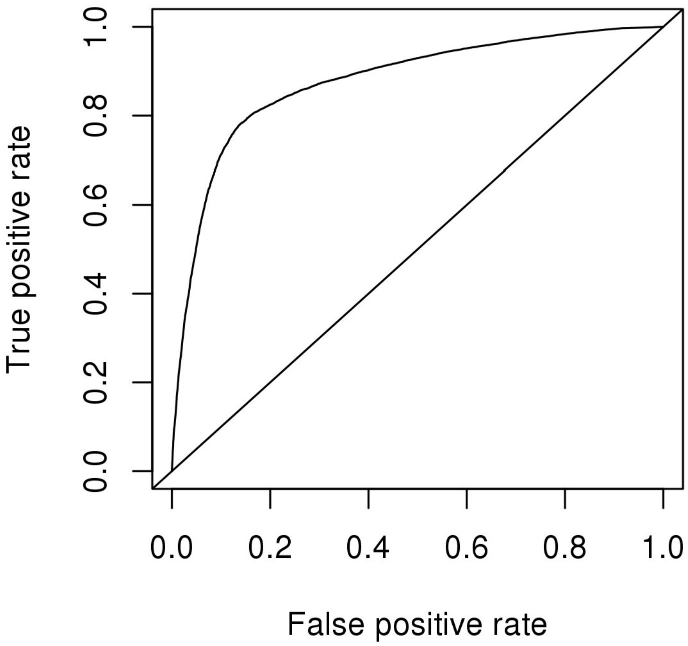
Note: First dotted vertical line indicates the value of log(Lambda) where deviance is minimized, while the second dotted vertical line indicates the value of log(Lambda) where deviance is within 1 standard error of the minimum. Values on top of figure indicate the total number of features included in the model for the specified value of log(Lambda).

eFigure 5. Binomial deviance versus log(Lambda) for 10-fold within-sample cross-validation of Ridge model among patients aged 65+ years.

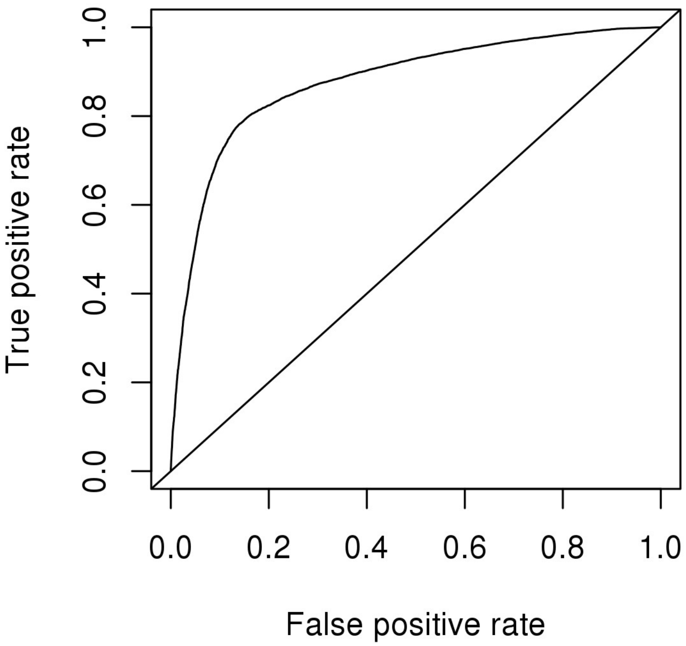


Note: First dotted vertical line indicates the value of log(Lambda) where deviance is minimized, while the second dotted vertical line indicates the value of log(Lambda) where deviance is within 1 standard error of the minimum.

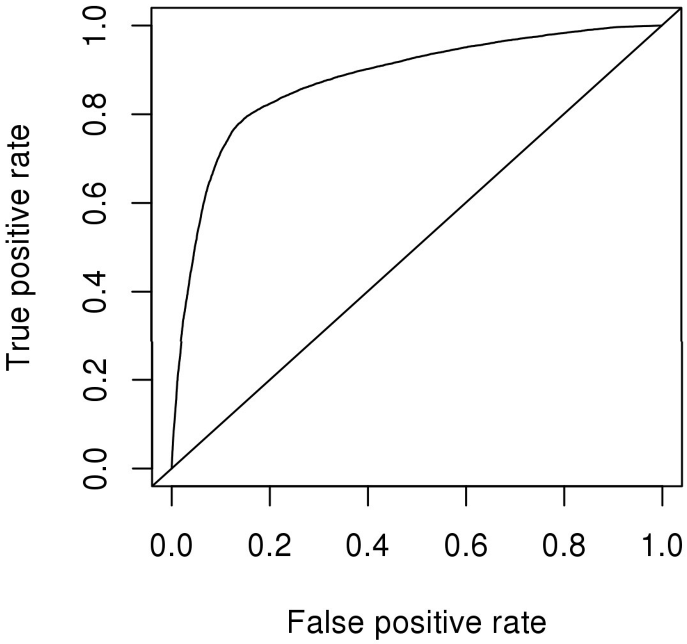
eFigure 6. ROC curve for out-of-sample validation of Stepwise model among patients aged 18-64 years.



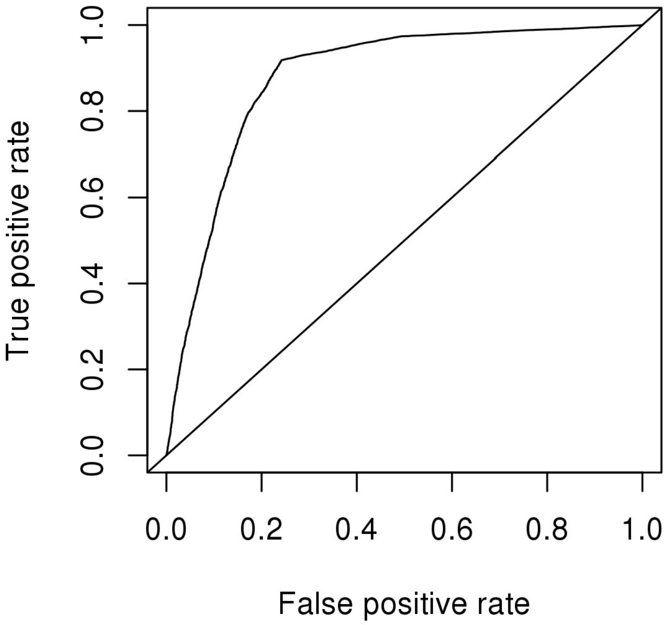
eFigure 7. ROC curve for out-of-sample validation of LASSO model among patients aged 18-64 years.



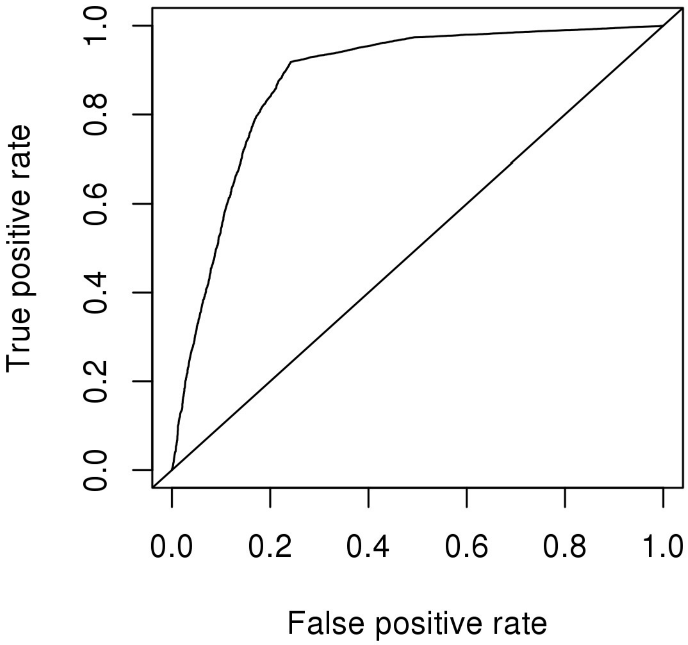
eFigure 8. ROC curve for out-of-sample validation of Ridge model among patients aged 18-64 years.



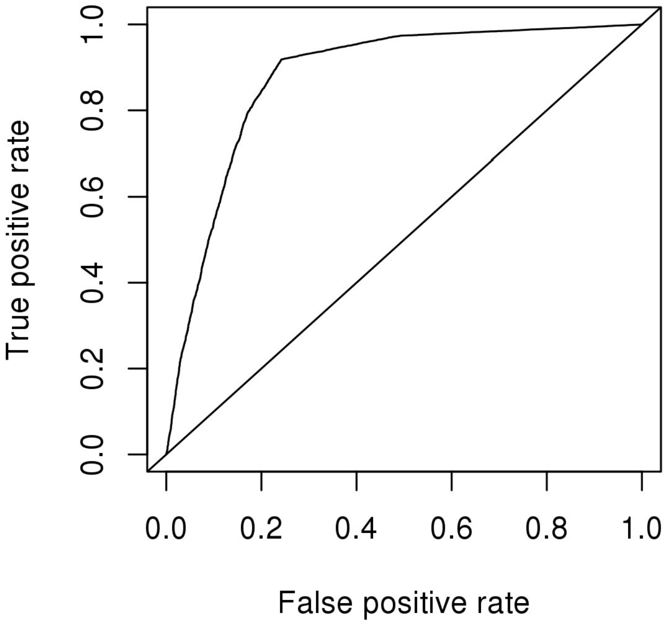
eFigure 9. ROC curve for out-of-sample validation of Stepwise model among patients aged 65+ years.



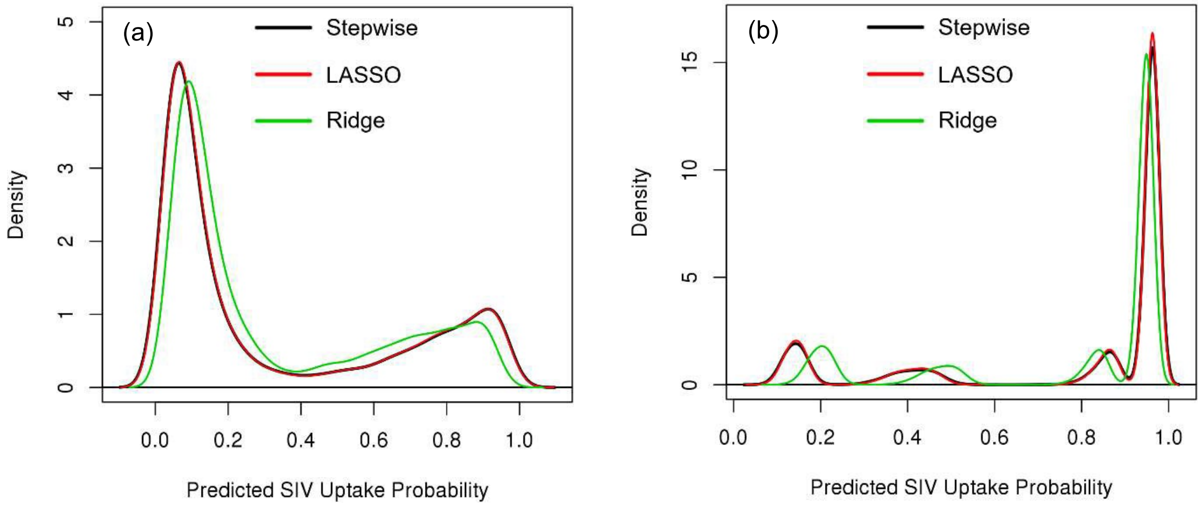
eFigure 10. ROC curve for out-of-sample validation of LASSO model among patients aged 65+ years.



eFigure 11. ROC curve for out-of-sample validation of Ridge model among patients aged 65+ years.



eFigure 12. Density plots of predicted SIV uptake probabilities from for out-of-sample validation for Stepwise, LASSO, and Ridge models among patients aged (a) 18-64 (b) 65+ years.



eTable. Performance measures from sensivitiy anaylsis (10-fold within-sample cross-validation) for Stepwise, LASSO, and Ridge models among patients aged (a) 18-64 and (b) 65+ years.

(a) 18-64

Performance Measure [a]

Model	Misclassification Rate	Brier Score	AUC	Sensitivity	Specificity
Stepwise	0.134 ± 0.0010	0.104 ± 0.0008	0.911 ± 0.0014	0.795 ± 0.0039	0.906 ± 0.0014
LASSO	0.134 ± 0.0019	0.104 ± 0.0011	0.911 ± 0.0011	0.795 ± 0.0031	0.906 ± 0.0021
Ridge	0.134 ± 0.0020	0.106 ± 0.0010	0.911 ± 0.0012	0.787 ± 0.0035	0.910 ± 0.0021

(b) 65+

Performance Measure [a]

Model	Misclassification Rate	Brier Score	AUC	Sensitivity	Specificity
Stepwise	0.085 ± 0.0019	0.066 ± 0.0015	0.896 ± 0.0035	0.944 ± 0.0019	0.771 ± 0.0081
LASSO	0.085 ± 0.0019	0.066 ± 0.0015	0.896 ± 0.0036	0.944 ± 0.0020	0.773 ± 0.0079
Ridge	0.088 ± 0.0020	0.067 ± 0.0015	0.896 ± 0.0034	0.955 ± 0.0017	0.705 ± 0.0076

a Performance measures reported as mean ± SD. Misclassification rate: proportion of patients with an incorrectly predicted SIV uptake status (based upon a cut-off of 0.5). Brier Score: measure of the accuracy of a probabilistic prediction, ranging from 0 to 1, where 0 indicates perfect accuracy. AUC: measure of the model’s discrimination power, ranging from 0 to 1, where 0.5 indicates an inability to appropriately classify a patient’s SIV uptake and 1 indicates perfect prediction. Sensitivity: true positive rate. Specificity: is the true negative rate.

eMethods

Measures of Uncertainty

In order to calculate the appropriate measures of uncertainty (95% Confidence Intervals) for the performance measures from the out-of-sample model validations, we utilized normal approximation methods. These methods were applicable in our study given the large sizes of our validation datasets. For Misclassification Rate, Sensitivity, and Specificity, we used a normal approximation for the binomial confidence interval.[2] For Brier Score, we calculated the standard error of the respective sample of Brier Scores and used the normal approximation formula as described in Bradley, Schwartz, and Hashino (2007).[3] For AUC, we calculated the standard error using the formula given by Hanley and Mcneil (1982)[4] and used the standard normal approximation, as the distribution of AUC is approximately normal in large samples.[5]

Statistical Software

All analyses were performed in R 3.4.3 using the following packages: *ROCR*, *glmnet*, *MASS*, *InformationValue*, and *ModelGood.*[6-10] Specifically, *stepAIC* from *MASS* was used to train Stepwise models, while *glmnet* was used to train LASSO and Ridge models. Cross-validation to determine the optimal value of *lambda* for the LASSO and Ridge models was performed using the *cv.glmnet* function with default parameters. Misclassification rate, sensitivity, and specificity were calculated using the *misClassError, sensitivity,* and *specificity* functions from the *InformationValue* package. AUC values were calculated, and ROC curves were constructed using the *prediction* and *performance* functions from the *ROCR* package. Calibration plots were constructed using the *calPlot2* function from the *ModelGood* package.

References

1. Loiacono MM, Mahmud SM, Chit A, et al. Patient and practice level factors associated with seasonal influenza vaccine uptake among at-risk adults in England, 2011 to 2016: An age-stratified retrospective cohort study. *Vaccine X.* 2020;4:100054.
2. NCSS. Chapter 273: Confidence Intervals for One-Sample Sensitivity and Specificity. 2019; https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/PASS/Confidence_Intervals_for_One-Sample_Sensitivity_and_Specificity.pdf. Accessed Mar 20, 2020.
3. Bradley AA, Schwartz SS, Hashino T. Sampling uncertainty and confidence intervals for the Brier score and Brier skill score. *Weather Forecasting.* 2008;23(5):992-1006.
4. Hanley JA, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143(1):29-36.
5. NCSS. Chapter 261: Confidence Intervals for the Area Under an ROC Curve. 2019; https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/PASS/Confidence_Intervals_for_the_Area_Under_an_ROC_Curve.pdf.
6. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. *Bioinformatics.* 2005;21(20):3940-3941.
7. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw.* 2010;33(1):1-22.
8. Venables WN, Ripley BD. *Modern applied statistics with S-PLUS.* Springer Science & Business Media; 2013.
9. Prabhakaran S. InformationValue: Performance Analysis and Companion Functions for Binary Classification Models. R package version 1.2. 3. In:2016.
10. Gerds TA, Andersen PK, Kattan MW. Calibration plots for risk prediction models in the presence of competing risks. *Stat Med.* 2014;33(18):3191-3203.