

Supplementary Materials

Methods

Measures and procedures

Subjects

Ethnicity was verified by the certificate of identification and categorized as Han and Yugur people. A participant was considered to be Han or Yugur people if he/she and his/her parents were of the same ethnicity.

Body mass index

Height and weight were measured with participants wearing light clothing and without shoes. Height was measured to the nearest 0.1 cm using a fixed stadiometer and weight was measured to the nearest 0.1 kg with a body composition analyzer (BC-420, TANITA, Japan). Body mass index (BMI) was defined as the weight in kilograms divided by the square of height in meters (kg/m^2).

Educational level

Educational level was operationalized as the highest obtained educational degree and divided into three groups: low (primary school or lower), medium (secondary school), and high education (university and higher).

Health-related behavior

Health-related behaviors were measured by smoking status, alcohol consumption, and physical activity. As described in detail before,^[1] a current smoker was defined as smoking at least one cigarette per day and lasting for at least 6 months, and a former smoker was defined as having quit smoking for >6 months preceding the survey. A current drinker was defined as those who drank at least twice per month (>640 mL of beer or 100 mL of Chinese liquor, about 57 g of alcohol) and had lasted for at least 6 months, and a former drinker was defined as a person who had stopped drinking >6 months before the study. The level of physical activity, considered both occupational and leisure-time physical activity, was classified as low (light levels of both occupational and leisure-time physical activity), moderate (moderate or high levels of either occupational or leisure-time physical activity), or high (moderate or high level of both occupational and leisure-time physical activity).^[2]

Statistical analyses

Sample size

Riley *et al*^[3] and van Smeden *et al*^[4] have described how to calculate the minimum

sample size for prediction model development with binary outcomes, conditional on the overall outcome risk in the target population, the number of candidate predictor parameters and the anticipated model performance in terms of overall model fit, as defined by the Cox-Snell R squared statistic (R^2_{cs}). In addition, Riley *et al*^[3] laid out a series of criteria the sample size should meet. The maximum possible value of R^2_{cs} was calculated by an outcome proportion.^[5] For the hypertension proportion of 29.76% in Gansu province, the $\max(R^2_{cs})$ value is 0.7. We conservatively assume that the new model will explain 20% of the variability; thus, the anticipated R^2_{cs} value is $0.20 \times 0.7 = 0.14$. Based on the hypertension proportion (29.76%), the number of candidate predictor parameters including dummy variables (17), the default level of shrinkage (0.9), and the R^2_{cs} estimated $>(0.14)$, the minimum sample size was 1006 with 300 events and 17.61 events per candidate predictor parameter (EPP), by using the “pmsampsize” package. Thus, the sample size in the training set (1161/2792) met the minimum sample size requirements.

GiViTI calibration belt

This approach estimated the relation between the logits of the probability predicted by a model and of the outcome rates observed, which was represented by a polynomial function. If the 80% or 95% confidence interval did not contain the bisector, defective calibration intervals were described for the models.

Relevant packages

Relevant packages in this study included pmsampsize (version 1.0.3), glmnet (version 4.0-2), rms (version 6.0-1), Hmisc (version 4.4-1), givitiR (version 1.3), rmda (version 1.6), pROC (version 1.16.2), DynNom (version 5.0.1), and shiny (version 1.5.0).

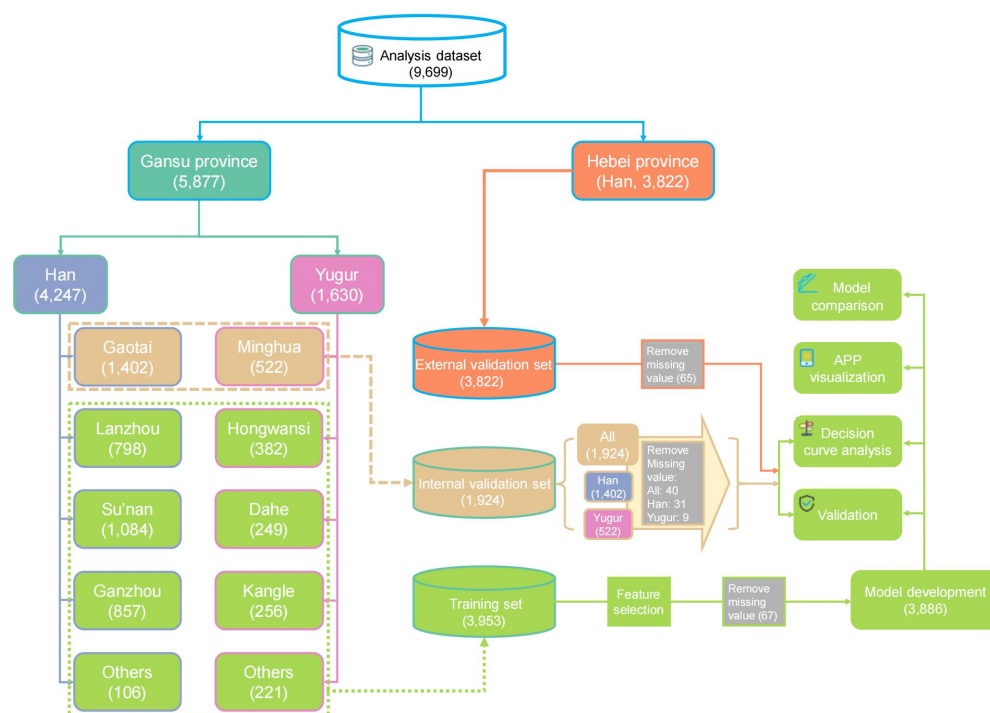
References

1. He H, Pan L, Pa L, Cui Z, Ren X, Wang D, *et al*. Data resource profile: The China National Health Survey (CNHS). *Int J Epidemiol* 2018;47:1734–1735. doi: 10.1093/ije/dyy151.
2. Shan G, Wei D, Wang C, Zhang J, Wang B, Ma M, *et al*. Trends of overweight and obesity in Yi people between 1996 and 2007: An Yi migrant study. *Biomed Environ Sci* 2011;24:467–474. doi: 10.3967/0895-3988.2011.05.003.
3. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE Jr, Moons KG, *et al*. Minimum

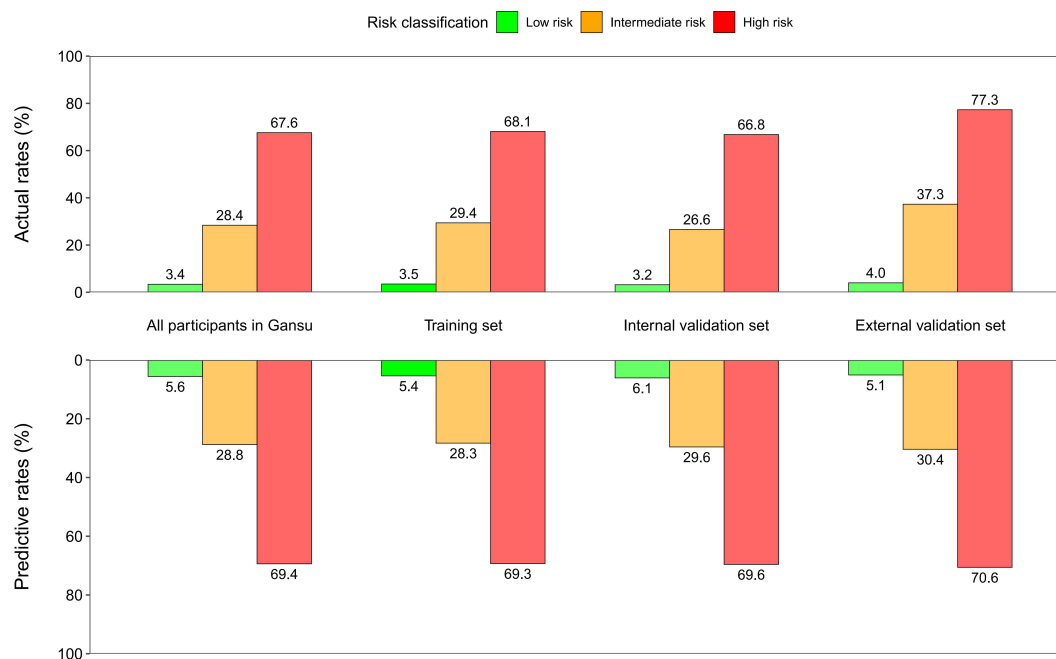
sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med* 2019;38:1276–1296. doi: 10.1002/sim.7992.

4. van Smeden M, Moons KG, de Groot JA, Collins GS, Altman DG, Eijkemans MJ, *et al.* Sample size for binary logistic prediction models: Beyond events per variable criteria. *Stat Methods Med Res* 2018;28:2455–2474. doi: 10.1177/0962280218784726.

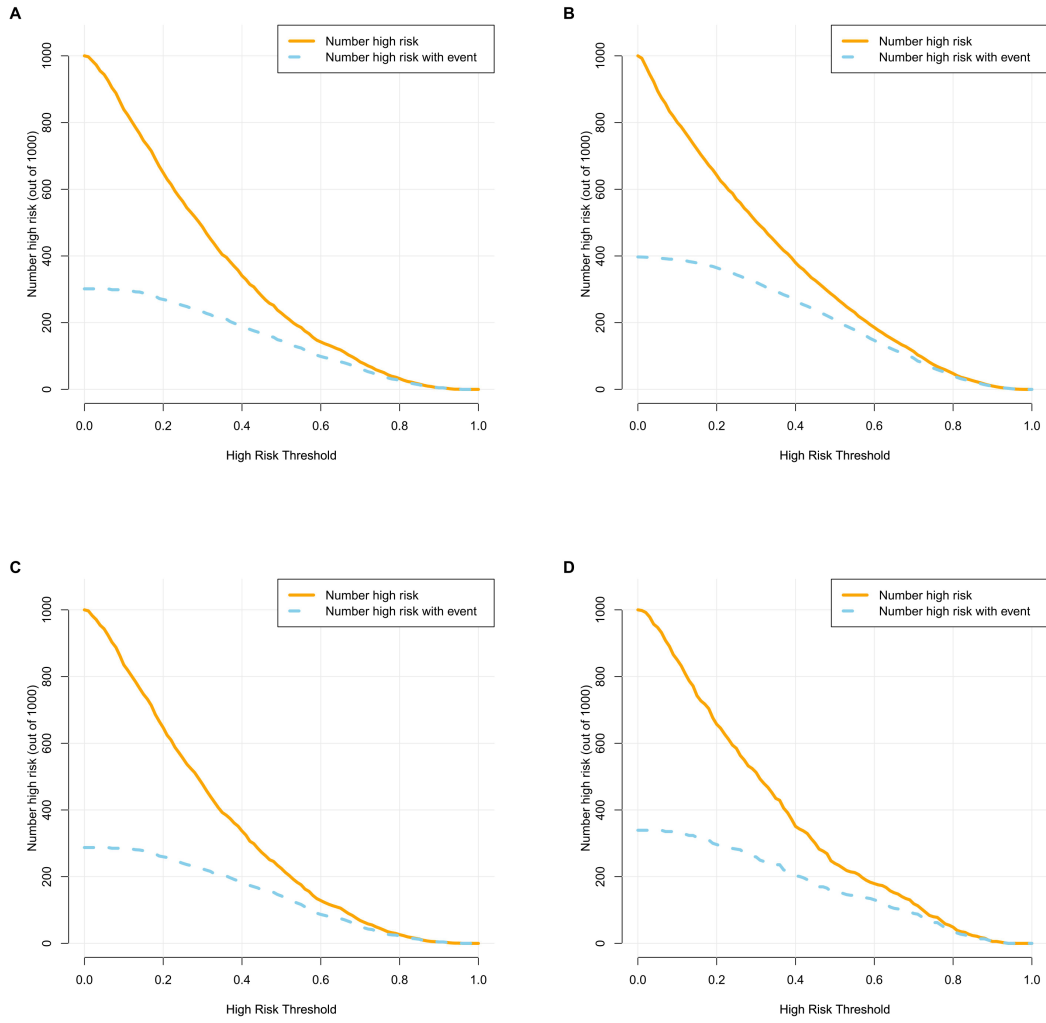
5. Riley RD, Ensor J, Snell KIE, Harrell FJ, Martin GP, Reitsma JB, *et al.* Calculating the sample size required for developing a clinical prediction model. *BMJ* 2020;368:m441. doi: 10.1136/bmj.m441.



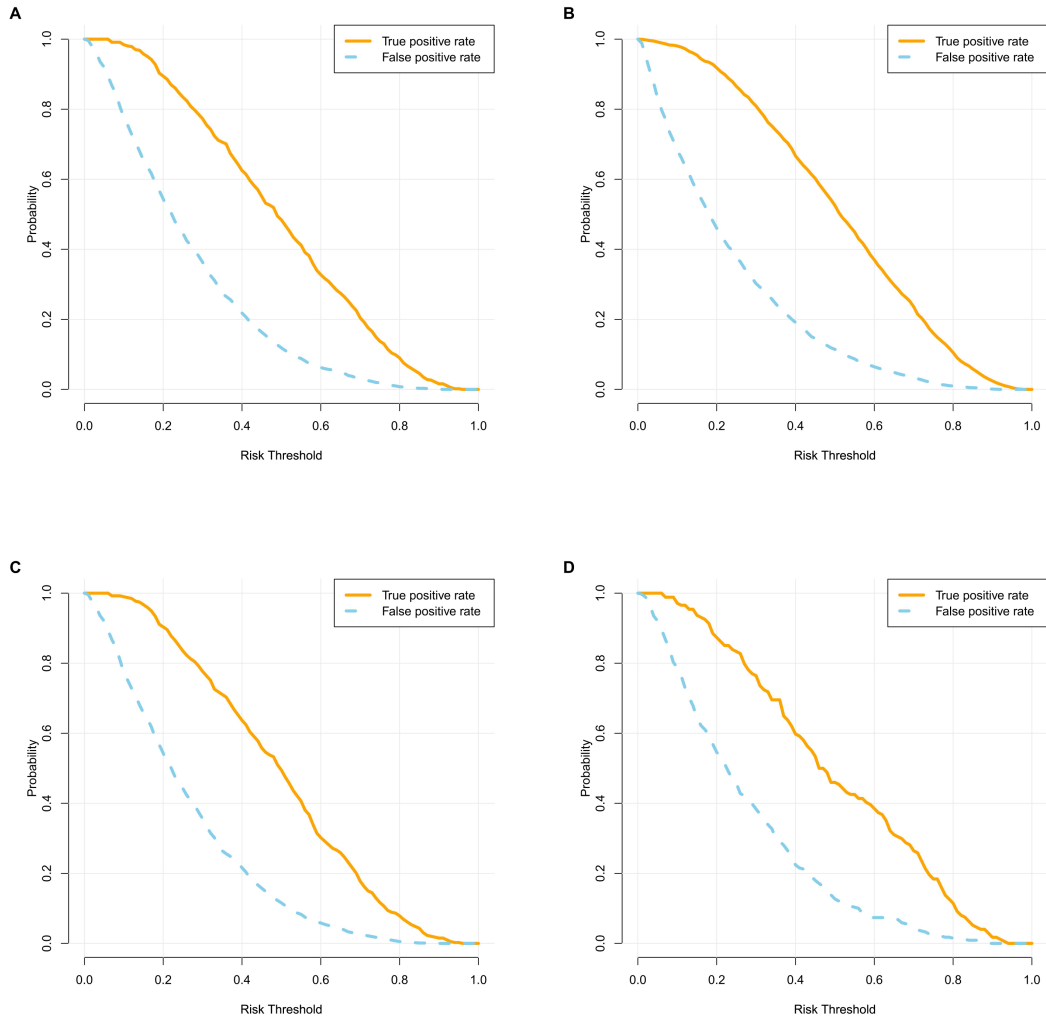
Supplementary Figure 1: Flow chart of data analysis process.



Supplementary Figure 2: The actual and predictive hypertension rates in different risk classifications and different datasets.



Supplementary Figure 3: Clinical impact curve analysis for the diagnostic model in validation sets. Of 1000 patients, the orange solid lines show the total number who would be deemed high risk for each risk threshold, and the sky-blue dashed lines show how many of those would be true positives (cases). (A) Clinical impact curve in internal validation set, (B) Clinical impact curve in external validation set, (C) Clinical impact curve in Han people of internal validation set, (D) Clinical impact curve in Yugur people of internal validation set.



Supplementary Figure 4: True- and false-positive rates for the diagnostic model in validation sets. The figures show information similar to that of a ROC and also show the risk threshold corresponding to each true- and false-positive rate. (A) True- and false-positive rates in internal validation set, (B) True- and false-positive rates in external validation set, (C) True- and false-positive rates in Han people of internal validation set, (D) True- and false-positive rates in Yugur people of internal validation set. ROC: Receiver operating characteristic.