Individualized risk prediction for type 2 diabetes

New result: The number needed to treat to prevent 1 diabetes case within the next three years:

Using model recommendation: 7.7776123  
Using lifestyle for everyone: 8.2365394  
Using metformin for everyone: 13.0513984

Discrimination was estimated by concordance (C-) statistic, which measures the probability of assigning higher risk to a case versus a non-case. Calibration was evaluated in terms of agreement between average observed risk and deciles of predicted risk. The index of prediction accuracy (IPA) combines discrimination and calibration in one summary value that ranges from 0 (poor) to 1 (perfect). Net benefit is based on the weighted difference in the true versus false positive rate of a prediction model at a given decision threshold, i.e.,

where the exchange rate is the odds of the threshold probability used for clinical decision making. Net benefit can be interpreted as the number of true-positive cases identified per 100 patients screened.

Table 0: Participants included.

| **Inclusion criteria** | **DPP** | **MESA** |
| --- | --- | --- |
| Study participants | 3,665 | 6,814 |
| Pre-diabetic at baseline1 | 3,152 | 2,104 |
| Randomized to placebo, metformin, or lifestyle | 2,640 | -- |
| *Abbreviations: DPP = Diabetes Prevention Program; and MESA = Multi-Ethnic Study of Atherosclerosis* | | |
| *1Baseline for Multi-Ethnic Study of Atherosclerosis participants was visit 1 or 2 dependent on pre-diabetes status* | | |

Table 1: Descriptive table of Diabetes Prevention Program and Multi-Ethnic Study of Atherosclerosis participants included in the current analysis

| **Characteristic** | **DPP  N = 2,640** | **MESA  N = 2,104** |
| --- | --- | --- |
| Age, years | 51 (11) | 64 (10) |
| Sex |  |  |
| Male | 864 (33%) | 1,058 (50%) |
| Female | 1,776 (67%) | 1,046 (50%) |
| Race/ethnicity |  |  |
| Non-Hispanic White | 1,626 (62%) | 682 (32%) |
| Non-Hispanic Black | 423 (16%) | 684 (33%) |
| Hispanic | 448 (17%) | 467 (22%) |
| Other/Chinese | 143 (5.4%) | 271 (13%) |
| Educational attainment |  |  |
| < High School | 1,352 (51%) | 196 (23%) |
| High School Graduate | 574 (22%) | 168 (20%) |
| Some College or College Graduate | 714 (27%) | 496 (58%) |
| Fasting glucose, mg/dl | 106 (7) | 101 (10) |
| Glycated Hemoglobin, % | 5.78 (0.40) | 5.75 (0.30) |
| HOMA-Insulin Resistance | 6.0 (4.2, 8.6) | --- |
| HOMA-Beta cell function | 199 (136, 272) | --- |
| Body mass index, kg/m2 | 34 (7) | 30 (6) |
| Triglycerides, mg/dl | 144 (101, 206) | 118 (84, 167) |
| Low-density lipoprotein cholesterol, mg/dl | 107 (27) | --- |
| High-density lipoprotein cholesterol, mg/dl | 46 (12) | --- |
| Abbreviations: DPP = Diabetes Prevention Program; HOMA = Homeostatic Model Assessment; and MESA = Multi-Ethnic Study of Atherosclerosis | | |

Table 2: Descriptive table of Diabetes Prevention Program and Multi-Ethnic Study of Atherosclerosis participants excluded from the current analysis

| **Characteristic** | **DPP  N = 1,025** | **MESA  N = 4,710** |
| --- | --- | --- |
| Age, years | 52 (10) | 62 (10) |
| Sex |  |  |
| Male | 364 (36%) | 2,155 (46%) |
| Female | 661 (64%) | 2,555 (54%) |
| Race/ethnicity |  |  |
| Non-Hispanic White | 491 (48%) | 1,941 (41%) |
| Non-Hispanic Black | 328 (32%) | 1,207 (26%) |
| Hispanic | 161 (16%) | 1,029 (22%) |
| Other/Chinese | 45 (4.4%) | 533 (11%) |
| Educational attainment |  |  |
| < High School | 550 (54%) | 821 (17%) |
| High School Graduate | 186 (18%) | 812 (17%) |
| Some College or College Graduate | 289 (28%) | 3,063 (65%) |
| Fasting glucose, mg/dl | 110 (9) | 98 (36) |
| Glycated Hemoglobin, % | 6.23 (0.60) | --- |
| HOMA-Insulin Resistance | 6.7 (4.5, 9.1) | --- |
| HOMA-Beta cell function | 188 (130, 263) | --- |
| Body mass index, kg/m2 | 34 (7) | 28 (5) |
| Triglycerides, mg/dl | 133 (96, 195) | 109 (76, 159) |
| Low-density lipoprotein cholesterol, mg/dl | 108 (28) | --- |
| High-density lipoprotein cholesterol, mg/dl | 45 (12) | --- |
| Abbreviations: DPP = Diabetes Prevention Program; HOMA = Homeostatic Model Assessment; and MESA = Multi-Ethnic Study of Atherosclerosis | | |

Table 3. Cumulative incidence of diabetes over follow-up in the Diabetes Prevention Program and Multi-Ethnic Study of Atherosclerosis

|  | **DPP** | **MESA** |
| --- | --- | --- |
| Over all follow-up time | | |
| Number at risk | 2,640 | 2,103 |
| Total follow-up time, person-years | 7,305 | 18,026 |
| Median (95% CI) follow-up time, years | 3.0 (3.0, 3.0) | 13 (9.6, 13) |
| Up to three years after baseline | | |
| Number of incident cases | 386 | 202 |
| Incidence rate, per 100 years | 5.9 | 3.4 |
| Cumulative incidence (95% CI), % | 17 (16, 19) | 9.8 (8.5, 11) |

Table 4A. Comparison of the individualized model with standard model.

| **Evaluation statistic1** | **Internal Validation2** | | **External Validation3** | |
| --- | --- | --- | --- | --- |
| **Standard** | **Individualized** | **Standard** | **Individualized** |
| NRI, Overall | 0 (ref) | 2.5 (-0.41, 5.3) | 0 (ref) | 0.60 (-0.53, 2.1) |
| NRI, Positive | 0 (ref) | 0.15 (-2.6, 2.6) | 0 (ref) | 0.96 (0.00, 2.5) |
| NRI, Negative | 0 (ref) | 2.4 (1.3, 3.5) | 0 (ref) | -0.37 (-0.71, -0.10) |
| AUC | 69.8 (66.7 73.0) | 70.7 (67.6 73.9) | 85.8 (83.3 88.3) | 85.6 (83.0 88.1) |
| IPA | 9.9 | 10.4 | 18.8 | 18.6 |
| Race/ethnicity | | | | |
| Demographic parity4 | 90 | 87 | 56 | 55 |
| Equal opportunity5 | 89 | 74 | 70 | 70 |
| Equal odds6 | 89 | 74 | 70 | 70 |
| Sex | | | | |
| Demographic parity | 93 | 90 | 88 | 88 |
| Equal opportunity | 83 | 84 | 82 | 85 |
| Equal odds | 83 | 84 | 82 | 85 |
| Abbreviations: AUC = Area underneath the receiver-operator characteristic curve; IPA = Index of prediction accuracy; and NRI = Net reclassification index | | | | |
| 1Table values are scaled by a factor of 100 for ease of interpretation. | | | | |
| 2Internal validation results are based on 10-fold cross-validation in the Diabetes Prevention Program data | | | | |
| 3External validation results are based on application of models fitted to the Diabetes Prevention Program data to the Multi-Ethnic Study of Atherosclerosis data. | | | | |
| 4Demographic parity is satisfied when a model's predictions have the same predicted positive rate across groups. | | | | |
| 5Equal opportunity is satisfied when a model's predictions have the same true positive and false negative rates across protected groups. | | | | |
| 6Equal odds is satisfied when a model's predictions have the same false positive, true positive, false negative, and true negative rates across protected groups. | | | | |

Table 4B: Subgroup comparison of the individualized model with standard model.

| **Evaluation statistic1** | **Internal Validation2** | | **External Validation3** | |
| --- | --- | --- | --- | --- |
| **Standard** | **Individualized** | **Standard** | **Individualized** |
| Women | | | | |
| AUC | 68.4 (64.6 72.3) | 69.6 (65.8 73.4) | 85.9 (82.1 89.7) | 85.7 (81.8 89.5) |
| IPA | 8.2 | 9.1 | 19.2 | 18.9 |
| Men | | | | |
| AUC | 72.8 (67.3 78.2) | 73.2 (67.7 78.7) | 85.9 (82.5 89.2) | 85.6 (82.3 89.0) |
| IPA | 13.3 | 13.1 | 18.4 | 18.4 |
| NH-Black Race | | | | |
| AUC | 69.1 (60.8 77.5) | 71.6 (63.6 79.7) | 85.4 (80.7 90.1) | 85.2 (80.5 90.0) |
| IPA | 6.7 | 7.1 | 17.5 | 17.4 |
| NH-White Race | | | | |
| AUC | 70.9 (67.0 74.8) | 71.2 (67.3 75.2) | 87.8 (83.9 91.6) | 87.5 (83.6 91.4) |
| IPA | 11.0 | 11.5 | 19.9 | 19.7 |
| Hispanic Race | | | | |
| AUC | 67.4 (59.8 75.1) | 69.4 (62.0 76.7) | 84.1 (78.6 89.5) | 83.8 (78.3 89.4) |
| IPA | 9.1 | 9.6 | 17.8 | 17.5 |
| Other Race | | | | |
| AUC | 65.1 (50.1 80.1) | 69.0 (54.5 83.6) | 85.7 (77.4 94.0) | 85.5 (77.1 93.9) |
| IPA | 7.5 | 8.7 | 21.2 | 20.6 |
| Abbreviations: AUC = Area underneath the receiver-operator characteristic curve; and IPA = Index of prediction accuracy | | | | |
| 1Table values are scaled by a factor of 100 for ease of interpretation. | | | | |
| 2Internal validation results are based on 10-fold cross-validation in the Diabetes Prevention Program data | | | | |
| 3External validation results are based on application of models fitted to the Diabetes Prevention Program data to the Multi-Ethnic Study of Atherosclerosis data. | | | | |

Table 5A: Reclassification matrix comparing individualized versus standard risk classification among participants in the Diabetes Prevention Program (internal) and the Multi-Ethnic Study of Atherosclerosis (external)

| **Standard risk categories** | **Individualized risk categories** | | |
| --- | --- | --- | --- |
| **0 to < 10%** | **10% to < 20%** | **≥ 20%** |
| Internal data | | | |
| 0 to < 10% | 708 (27%) | 157 (5.9%) | 3 (0.11%) |
| 10% to < 20% | 170 (6.4%) | 790 (30%) | 60 (2.3%) |
| ≥ 20% | 0 | 114 (4.3%) | 638 (24%) |
| External data | | | |
| 0 to < 10% | 1,174 (56%) | 37 (1.8%) | 0 |
| 10% to < 20% | 1 (0.05%) | 610 (29%) | 10 (0.48%) |
| ≥ 20% | 0 | 1 (0.05%) | 271 (13%) |

Table 5B: Reclassification matrix comparing individualized versus standard risk classification among men and women.

| **Standard risk categories** | **Individualized risk categories** | | |
| --- | --- | --- | --- |
| **0 to < 10%** | **10% to < 20%** | **≥ 20%** |
| Men | | | |
| 0 to < 10% | 819 (43%) | 72 (3.7%) | 0 |
| 10% to < 20% | 60 (3.1%) | 559 (29%) | 25 (1.3%) |
| ≥ 20% | 0 | 34 (1.8%) | 353 (18%) |
| Women | | | |
| 0 to < 10% | 1,063 (38%) | 122 (4.3%) | 3 (0.11%) |
| 10% to < 20% | 111 (3.9%) | 841 (30%) | 45 (1.6%) |
| ≥ 20% | 0 | 81 (2.9%) | 556 (20%) |

Table 5C: Reclassification matrix comparing individualized versus standard risk classification among men and women.

| **Standard risk categories** | **Individualized risk categories** | | |
| --- | --- | --- | --- |
| **0 to < 10%** | **10% to < 20%** | **≥ 20%** |
| Caucasian | | | |
| 0 to < 10% | 790 (34%) | 109 (4.7%) | 2 (0.09%) |
| 10% to < 20% | 98 (4.2%) | 707 (31%) | 47 (2.0%) |
| ≥ 20% | 0 | 73 (3.2%) | 482 (21%) |
| African American | | | |
| 0 to < 10% | 543 (49%) | 31 (2.8%) | 1 (0.09%) |
| 10% to < 20% | 33 (3.0%) | 297 (27%) | 9 (0.81%) |
| ≥ 20% | 0 | 20 (1.8%) | 173 (16%) |
| Hispanic | | | |
| 0 to < 10% | 362 (40%) | 39 (4.3%) | 0 |
| 10% to < 20% | 26 (2.8%) | 277 (30%) | 9 (0.98%) |
| ≥ 20% | 0 | 16 (1.7%) | 186 (20%) |
| Other | | | |
| 0 to < 10% | 187 (45%) | 15 (3.6%) | 0 |
| 10% to < 20% | 14 (3.4%) | 119 (29%) | 5 (1.2%) |
| ≥ 20% | 0 | 6 (1.4%) | 68 (16%) |

Table 6: Optimal preventive intervention and 3-year counterfactual risk for type 2 diabetes from an individualized risk prediction model in the Diabetes Prevention Program

| **Optimal intervention** | **N (%)** | **Counterfactual 3-year T2D Predicted Risk, Mean (SD)** | | |
| --- | --- | --- | --- | --- |
| **Lifestyle** | **Metformin** | **Placebo** |
| Internal | | | | |
| Lifestyle | 2,267 (86%) | 10 (6.5) | 17 (9.4) | 22 (16) |
| Metformin | 373 (14%) | 20 (12) | 15 (7.6) | 27 (17) |
| External | | | | |
| Lifestyle | 2,035 (97%) | 6.5 (5.2) | 14 (9.7) | 15 (15) |
| Metformin | 69 (3.3%) | 16 (11) | 13 (9.0) | 23 (20) |

Table 7: Summary of and instructions for using the individualized prediction model for type 2 diabetes.

| **Variable1** | **Hazard ratio (95% CI)2** |
| --- | --- |
| Glycated Hemoglobin | 1.21 (1.09, 1.34) |
| Triglycerides | 1.24 (1.15, 1.35) |
| Age | |
| Lifestyle | 0.95 (0.78, 1.16) |
| Metformin | 1.02 (0.86, 1.23) |
| Placebo | 0.97 (0.84, 1.13) |
| Body mass index | |
| Lifestyle | 1.37 (1.16, 1.60) |
| Metformin | 0.89 (0.74, 1.06) |
| Placebo | 1.04 (0.91, 1.19) |
| Fasting glucose | |
| Lifestyle | 1.47 (1.23, 1.76) |
| Metformin | 1.53 (1.31, 1.78) |
| Placebo | 1.90 (1.67, 2.17) |
| Sex | |
| Female | 1.16 (0.95, 1.43) |
| Male | 1.00 (Reference) |
| 1Predictor variables included in the table were selected apriori based on clinical availability and known associations. | |
| 2Hazard ratios are adjusted for all variables listed in the table. Hazard ratios for continuous variables correspond to a one standard deviation change in the variable. | |
| Instructions for computing 3-year predicted risk:   Step 1: Scale predictors:   - divide age by 10.58244063 - divide fasting glucose by 6.67887385 - divide glycated hemoglobin by 0.39663737 - divide triglycerides by 95.62000286 - divide body mass index by 6.55160219   Step 2: Compute linear predictor (LP):   LP = (sex = female) \* 0.15268123 + (treatment group = metformin) \* 1.73746560 + (treatment group = placebo) \* -2.06808662 + (glycated hemoglobin) \* 0.19224983 + (triglycerides) \* 0.21835857 age \* [-0.04914900 + (treatment group = metformin) \* 0.07302687) + (treatment group = placebo) \* 0.02231801)] body mass index \* [ 0.31117408 + (treatment group = metformin) \* -0.42922833) + (treatment group = placebo) \* -0.26864075)] fasting glucose \* [0.38697814 + (treatment group = metformin) \* 0.03748913) + (treatment group = placebo) \* 0.25646930)]   Step 3: Plug LP into the risk formula:   Risk = 1 - exp(-0.155833 \* exp(LP - 11.216613)) | |

Table 8: Characteristics of participants stratified by recommended treatment to prevent diabetes according to the individualized prediction model.

| **Characteristic** | **Optimal predicted treatment** | |
| --- | --- | --- |
| **lifestyle  N = 4,302** | **metformin  N = 442** |
| study |  |  |
| DPP | 2,267 (53%) | 373 (84%) |
| MESA | 2,035 (47%) | 69 (16%) |
| Age, years | 58 (12) | 46 (9) |
| Sex |  |  |
| Male | 1,858 (43%) | 64 (14%) |
| Female | 2,444 (57%) | 378 (86%) |
| Race/ethnicity |  |  |
| Non-Hispanic White | 2,049 (48%) | 259 (59%) |
| Non-Hispanic Black | 991 (23%) | 116 (26%) |
| Hispanic | 854 (20%) | 61 (14%) |
| Other/Chinese | 408 (9.5%) | 6 (1.4%) |
| Educational attainment |  |  |
| < High School | 1,338 (43%) | 210 (52%) |
| High School Graduate | 650 (21%) | 92 (23%) |
| Some College or College Graduate | 1,106 (36%) | 104 (26%) |
| Predicted 3-year risk under metformin intervention | 0.13 (0.09, 0.19) | 0.12 (0.09, 0.18) |
| Predicted 3-year risk under lifestyle intervention | 0.07 (0.05, 0.10) | 0.17 (0.11, 0.26) |
| Fasting glucose, mg/dl | 104 (8) | 107 (8) |
| Glycated Hemoglobin, % | 5.76 (0.37) | 5.81 (0.37) |
| HOMA-Insulin Resistance | 5.6 (3.9, 8.0) | 8.8 (6.6, 11.7) |
| HOMA-Beta cell function | 187 (127, 258) | 273 (213, 368) |
| Body mass index, kg/m2 | 30 (5) | 46 (5) |
| Triglycerides, mg/dl | 132 (92, 188) | 132 (96, 184) |
| Low-density lipoprotein cholesterol, mg/dl | 107 (27) | 104 (26) |
| High-density lipoprotein cholesterol, mg/dl | 46 (12) | 44 (10) |

Figure 1: Decision curve analysis of the individualized and standard model in internal and external validation.

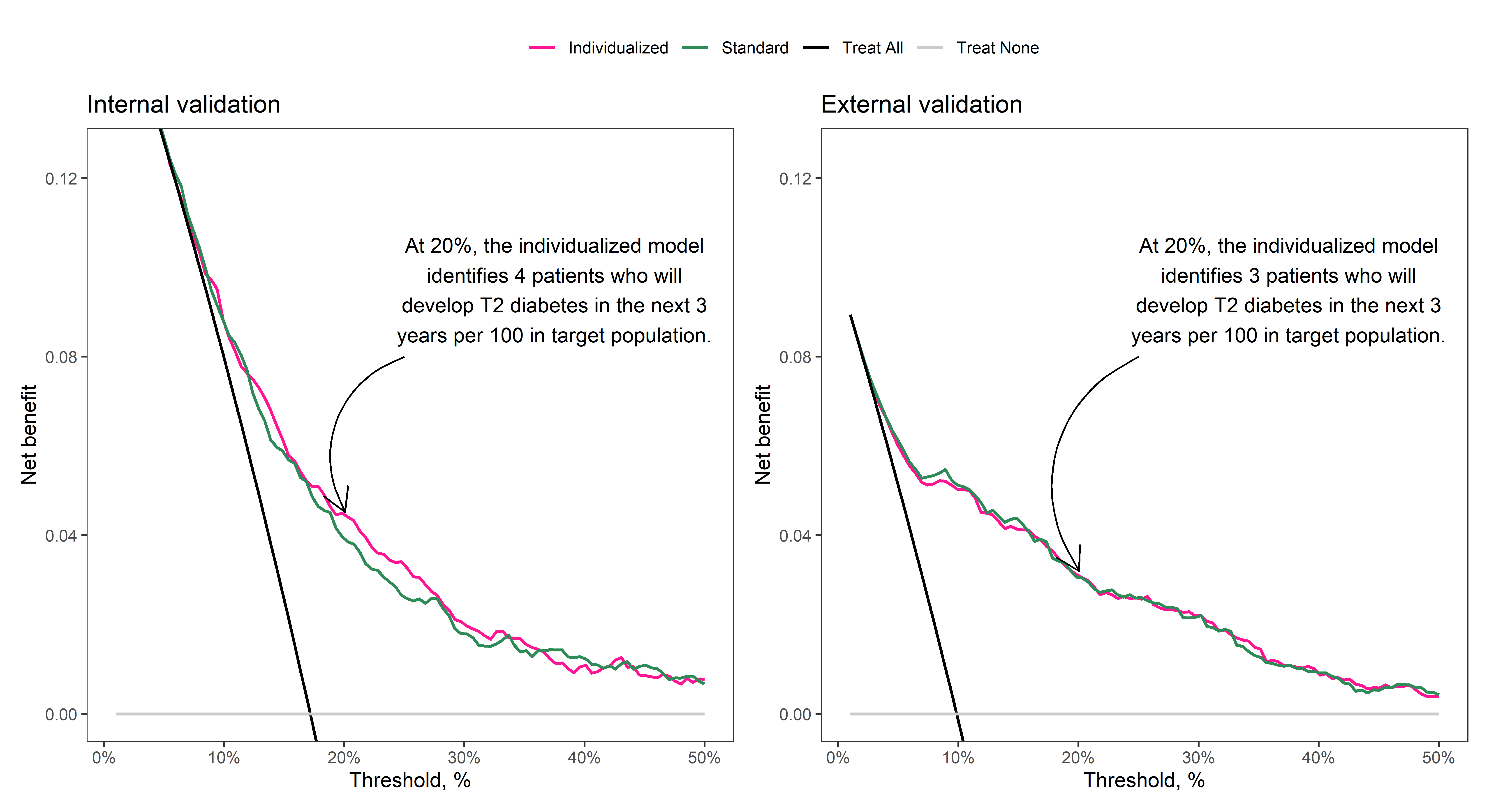


Figure 2: Calibration of the individualized model in internal and external validation.

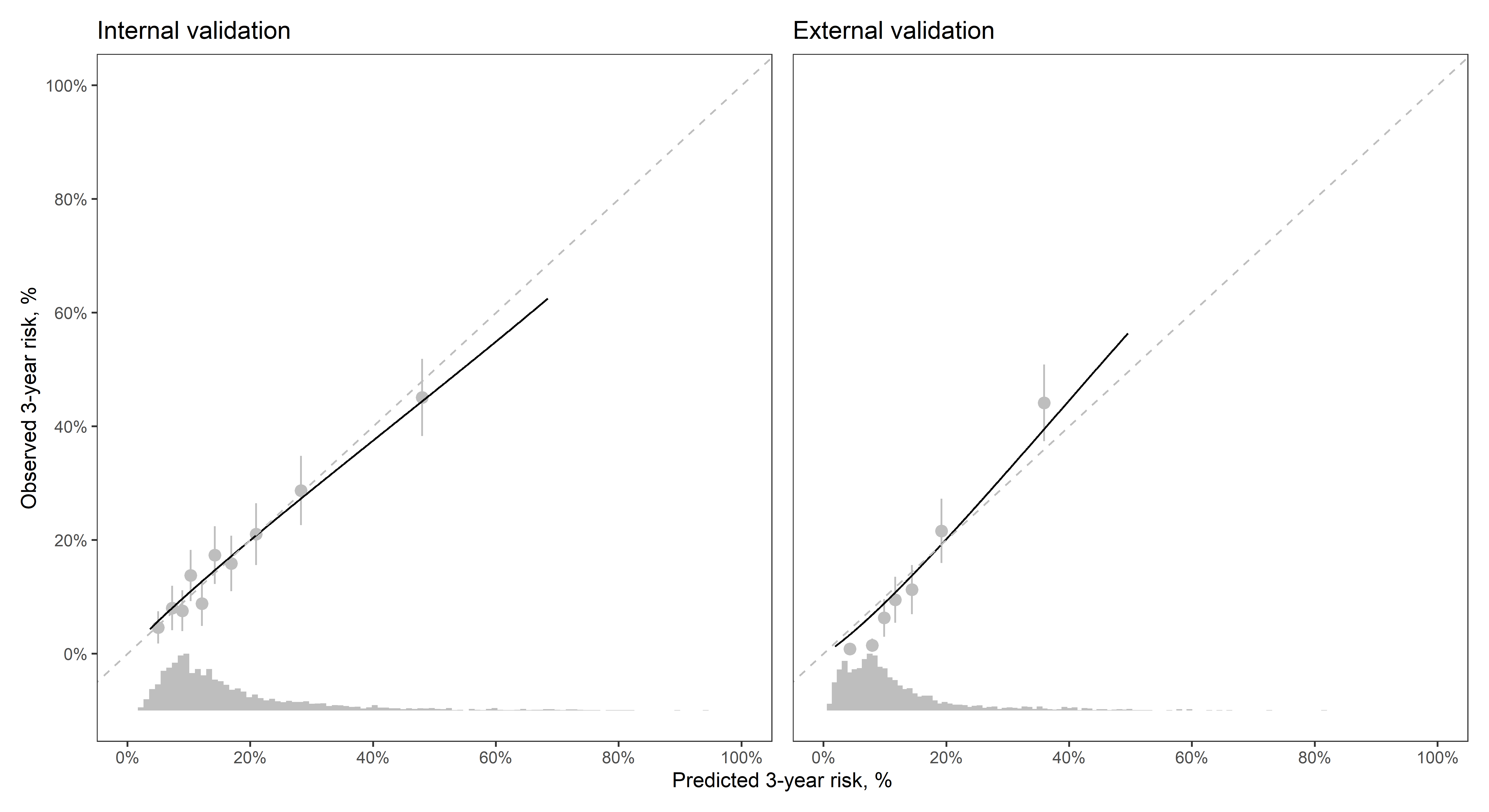


Figure 3: Calibration of the individualized model in internal and external validation among male study participants

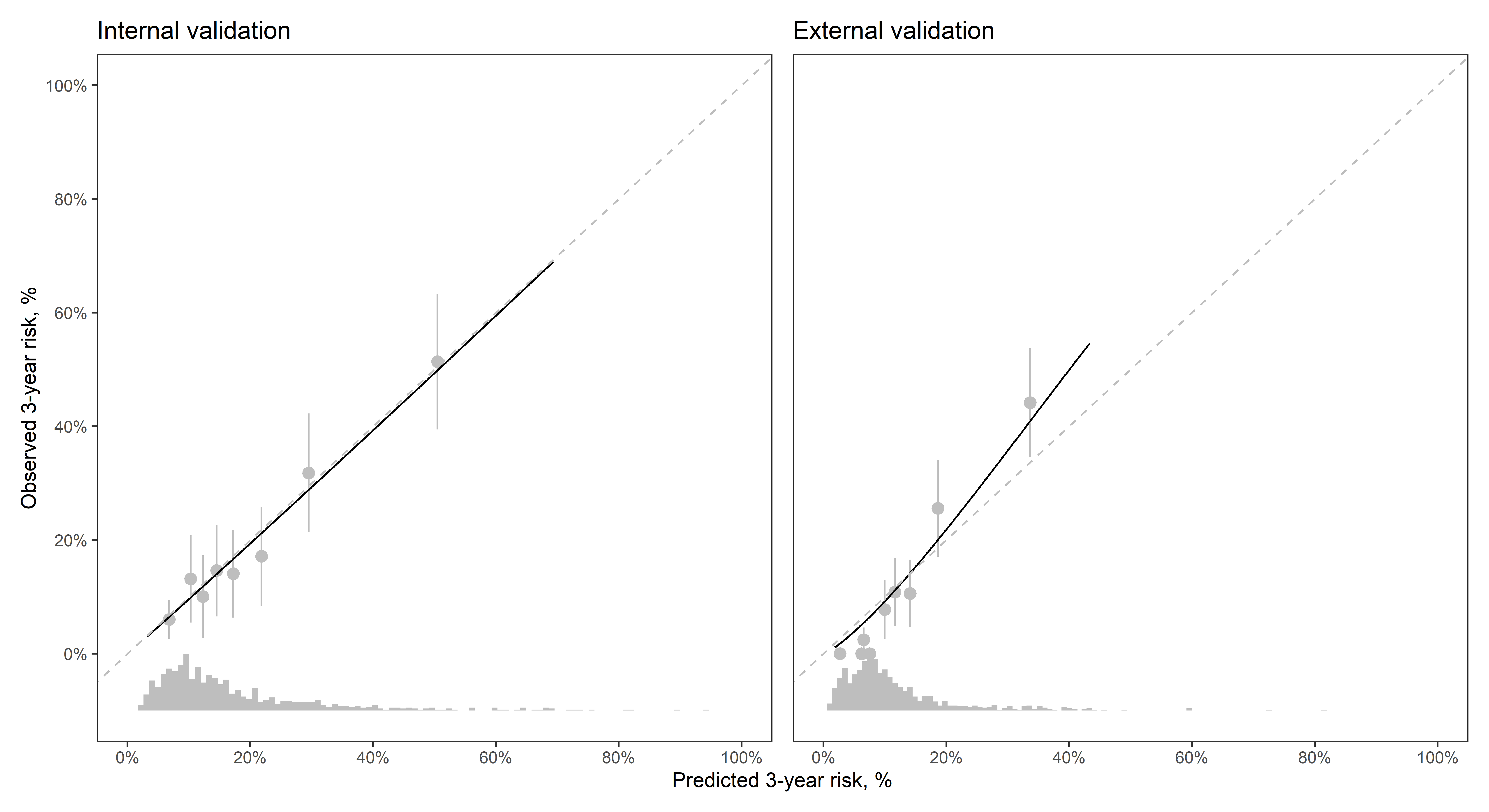


Figure 4: Calibration of the individualized model in internal and external validation among female study participants

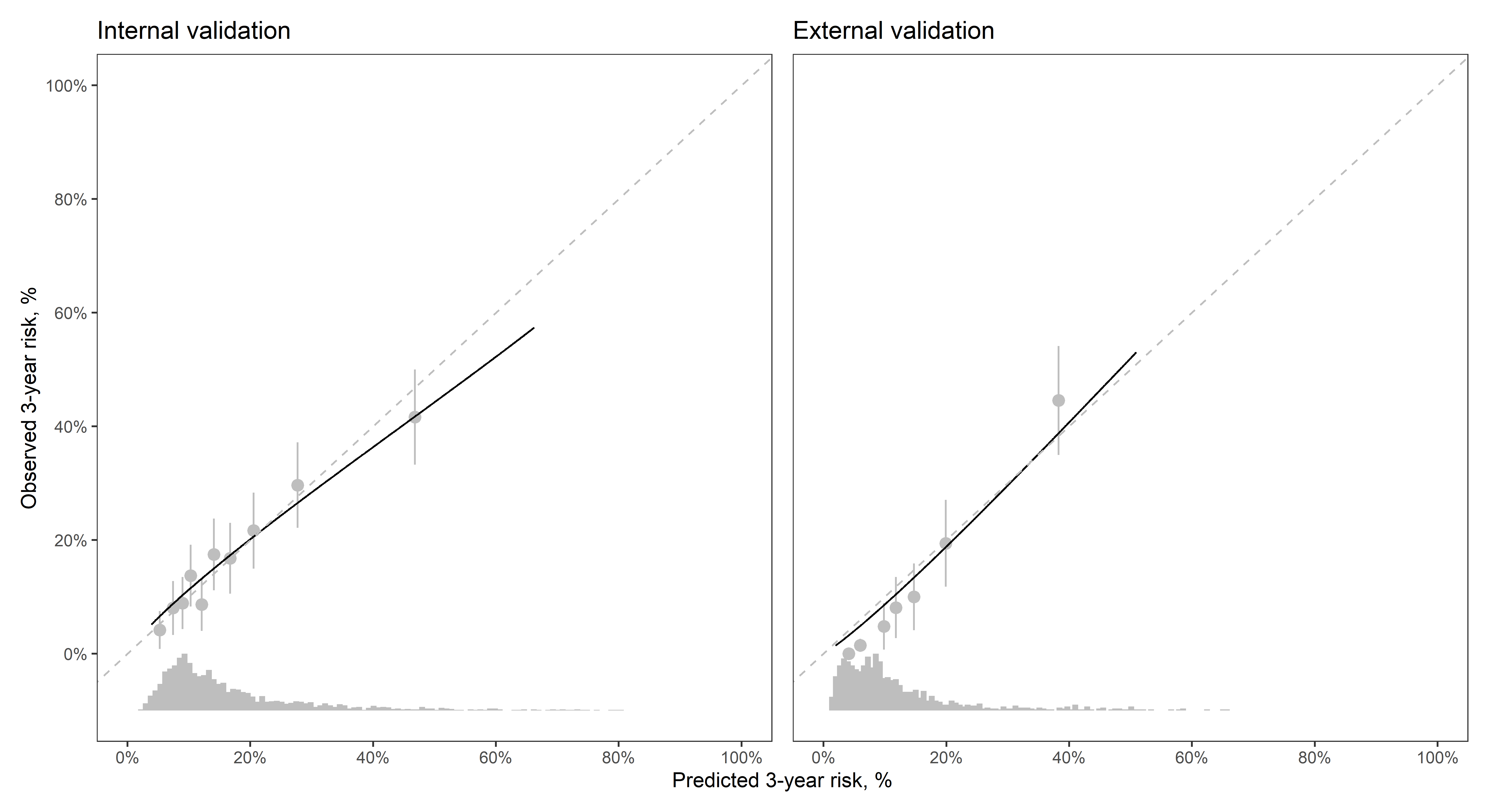


Figure 5: Calibration of the individualized model in internal and external validation among African American study participants

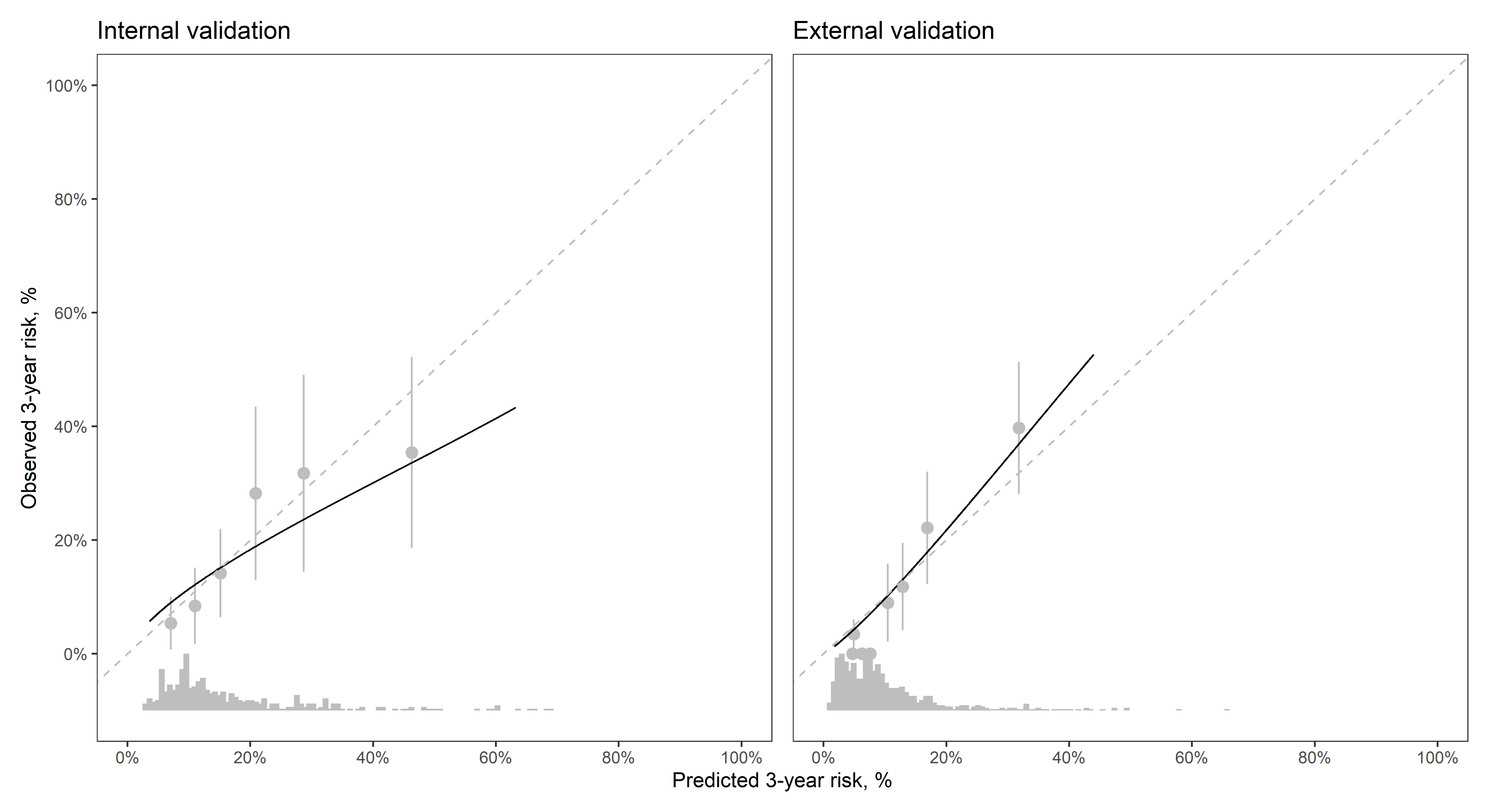


Figure 6: Calibration of the individualized model in internal and external validation among Caucasian study participants

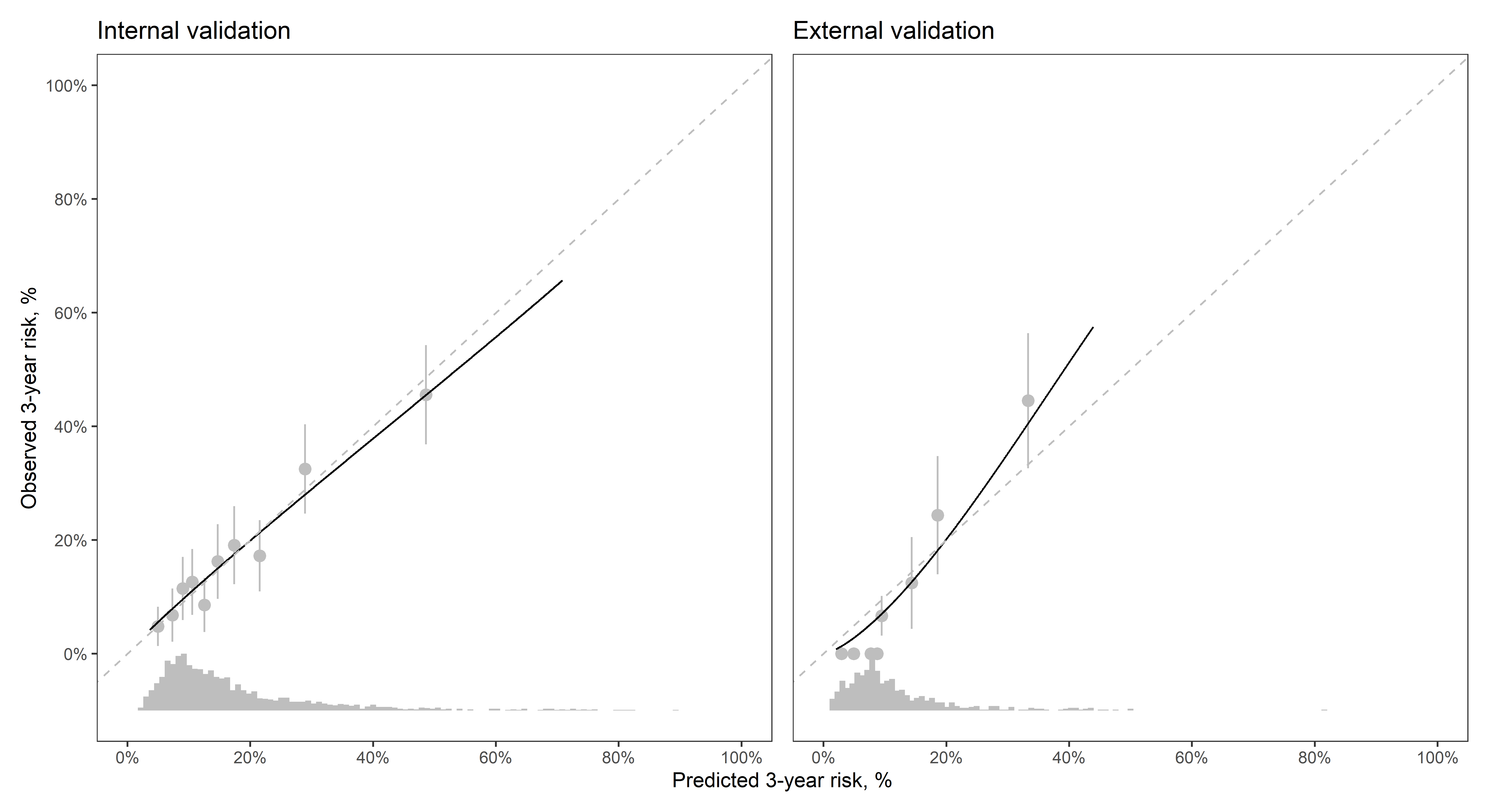


Figure 6: Calibration of the individualized model in internal and external validation among Hispanic study participants

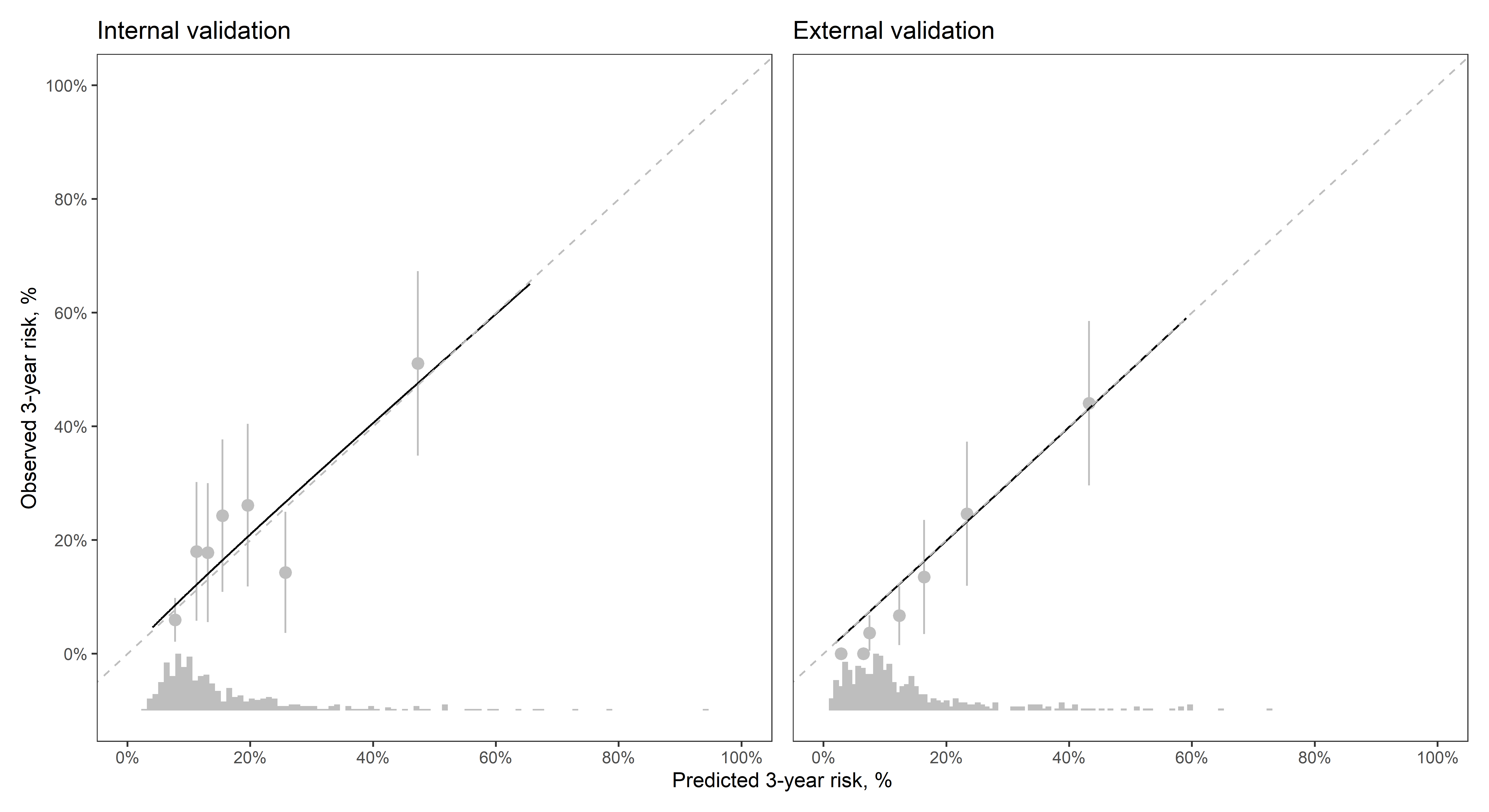


Figure 6: Calibration of the individualized model in internal and external validation among study participants who identified their race/ethnicity as a category other than African American, Caucasian, or Hispanic.

