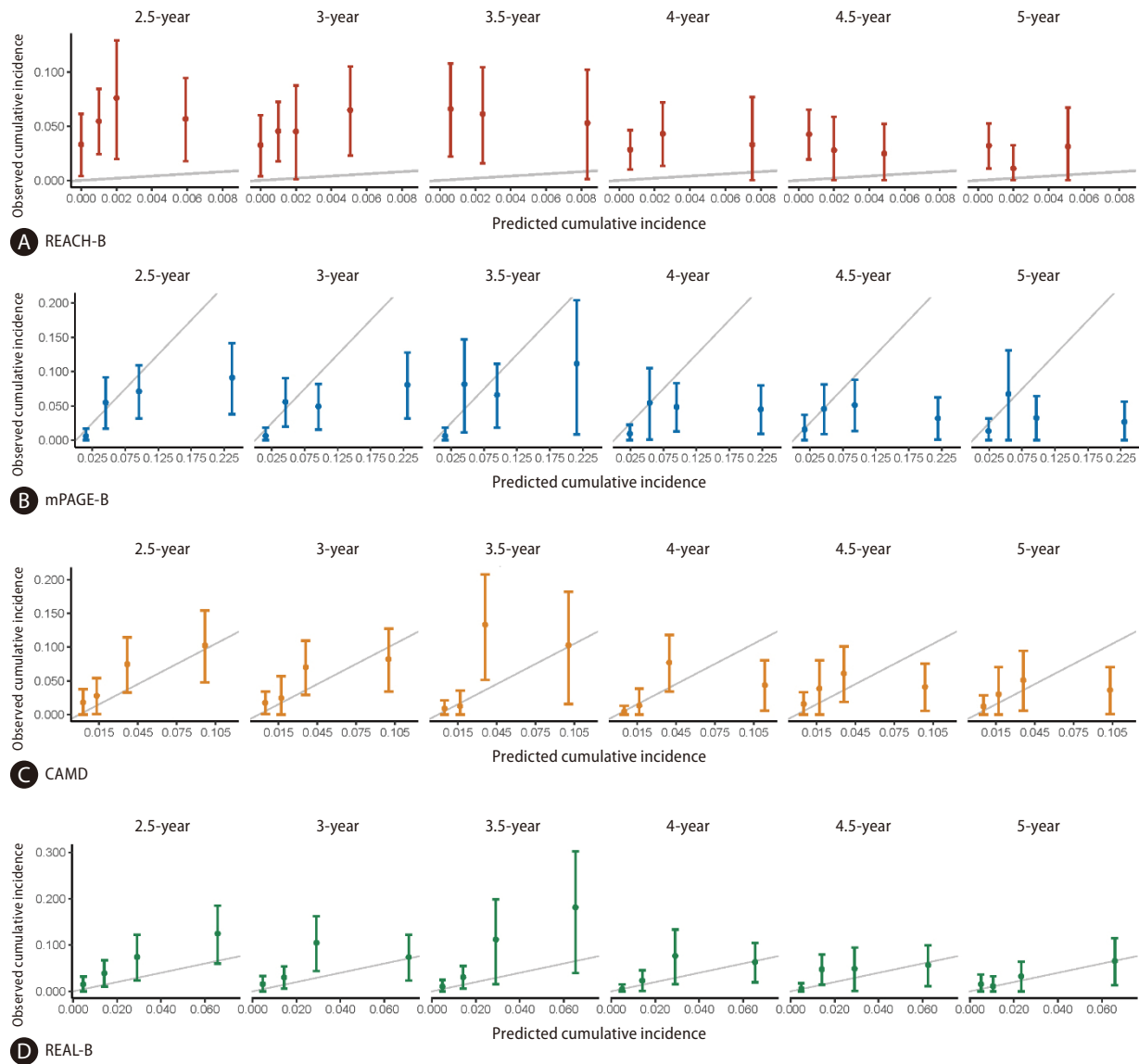


Appendix 7. Calibration of REACH-B, mPAGE-B, CAMD and REAL-B models at different on-treatment timepoints

Supplementary Table 7-1. Brier score of Risk prediction model scores

MODEL	Year 2.5 (n=660)	Year 3 (n=640)	Year 3.5 (n=622)	Year 4 (n=603)	Year 4.5 (n=589)	Year 5 (n=562)
REACH-B	0.048	0.044	0.043	0.031	0.032	0.023
mPAGE-B	0.052	0.046	0.047	0.039	0.040	0.035
CAMD	0.045	0.041	0.041	0.031	0.032	0.024
REAL-B	0.046	0.041	0.041	0.030	0.031	0.022



Supplementary Figure 7-2. Calibration plot of REACH-B, mPAGE-B, CAMD and REAL-B models at different on-treatment timepoints. The calibration plot revealed that REACH-B continuously underestimated HCC risks at all timepoints. For mPAGE-B, CAMD and REAL-B, the HCC risks were only well calibrated for patients at the low- and relatively low-risk quantiles predicted by serial on-treatment model scores.