

Supplementary Material

From Classification to Competing Risks: A Critical Re-Examination of a Widely-Used Cirrhosis Outcome Benchmark

Meseret Assefa Kerga Axel Faes

These supplementary figures support the main text. Each is referenced there by its supplementary label (Fig. S1–S5). All are produced by the same reproducible pipeline released with the paper; see the code repository (`run_advanced.py`, `run_multidataset.py`, `run_competing.py`).

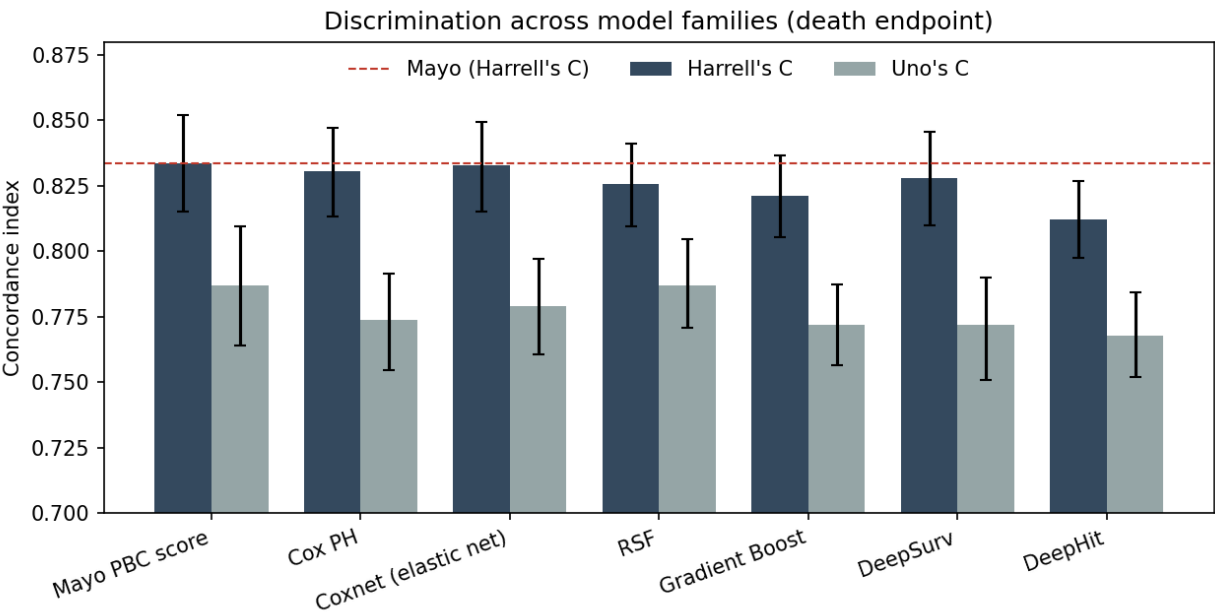


Figure S1: Concordance across model families for the death endpoint. The Mayo score (dashed line) is not exceeded by any ML model. (Companion to main-text Table 2, model comparison.)

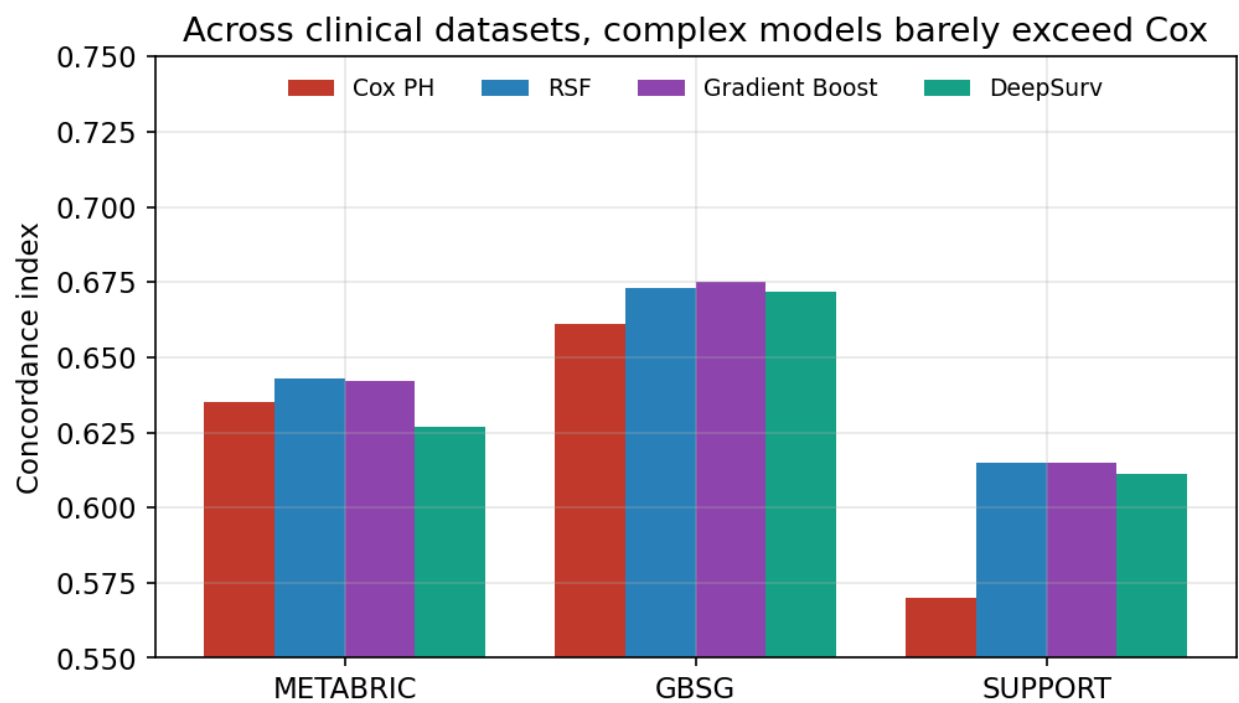


Figure S2: Across external datasets, complex models exceed Cox only marginally on smaller cohorts (METABRIC, GBSG) and meaningfully only on the largest (SUPPORT). (Companion to the external multi-dataset table.)

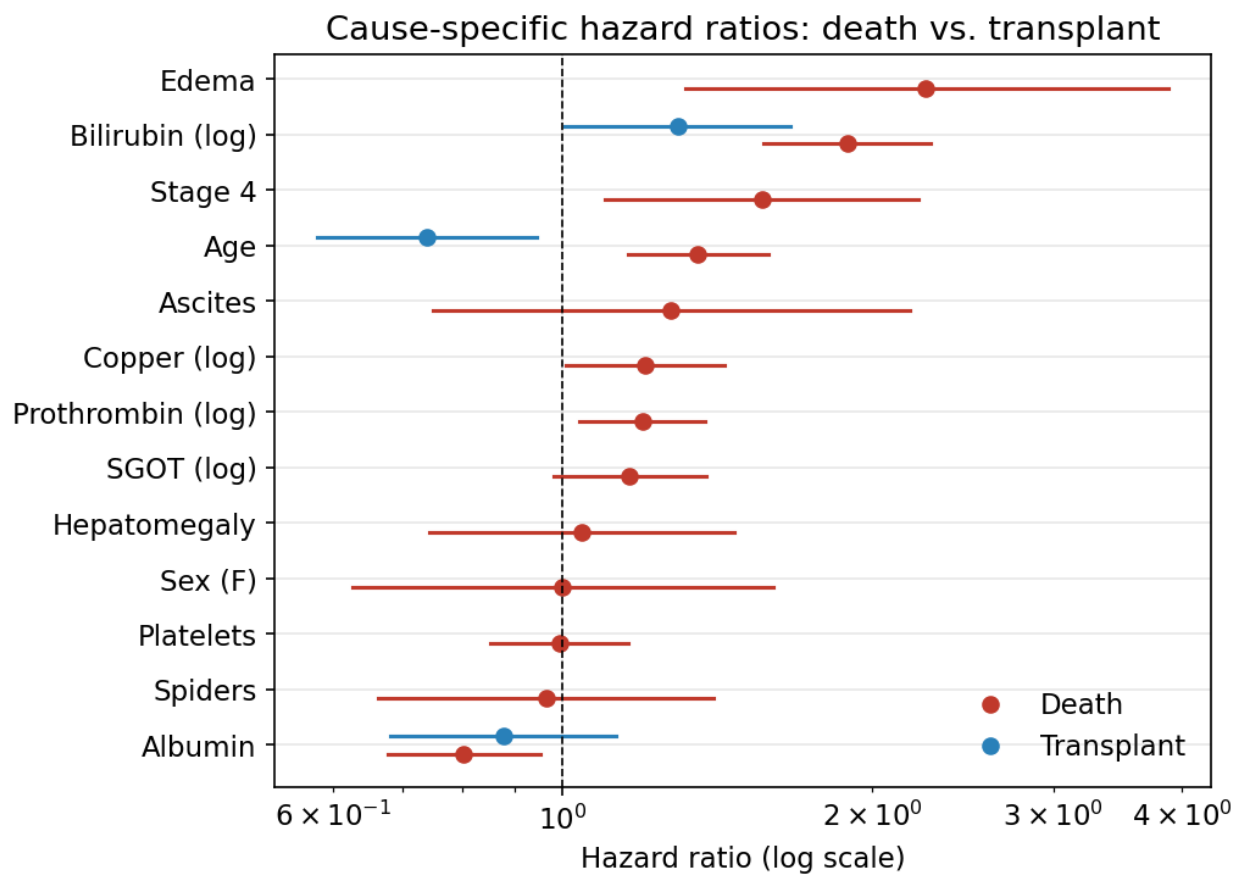


Figure S3: Cause-specific hazard ratios for death (red) and transplantation (blue). Age sits above 1 for death but below 1 for transplantation, the opposite-direction effect. (Companion to the cause-specific Cox table.)

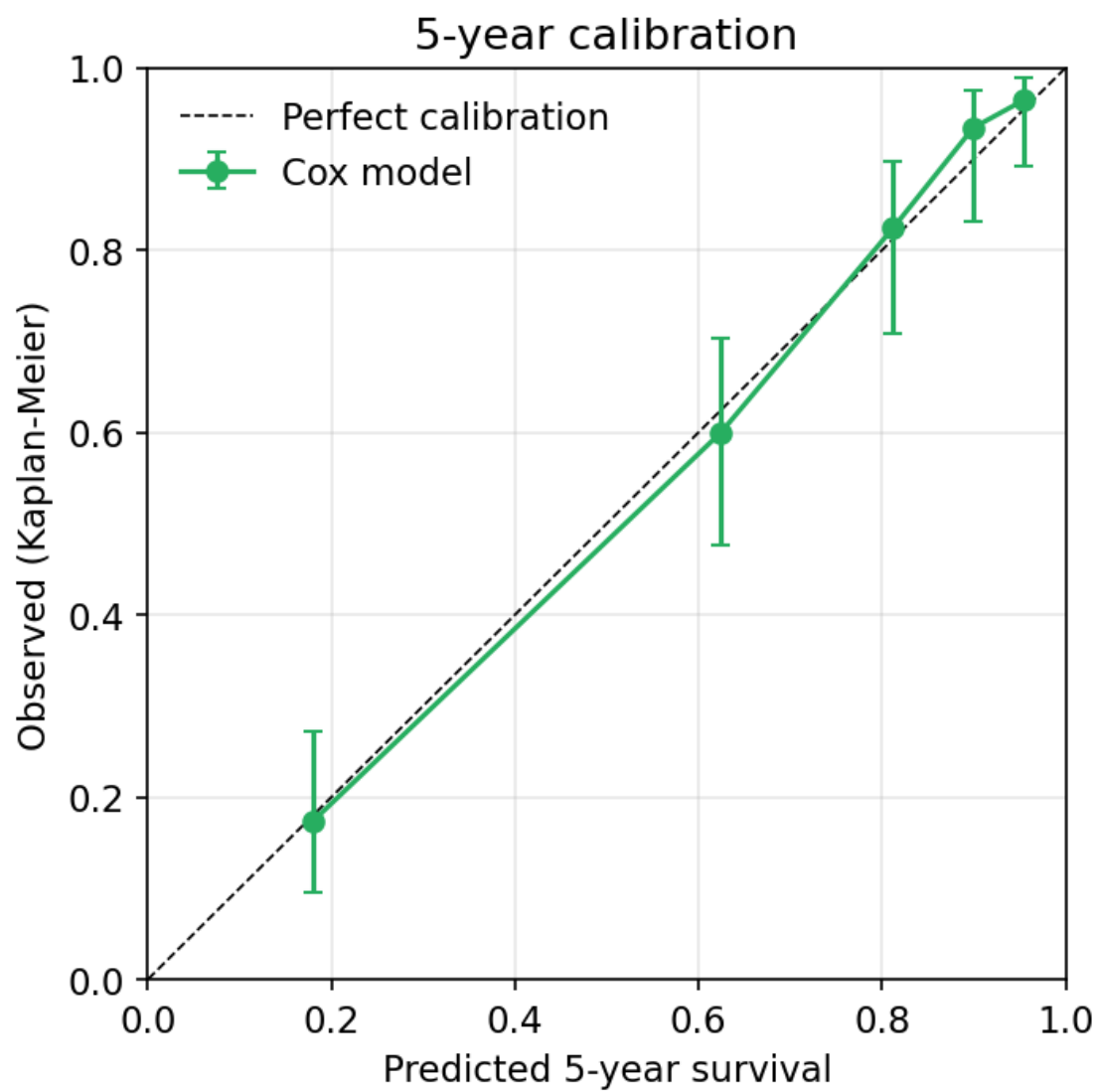


Figure S4: Five-year calibration: predicted survival vs. Kaplan–Meier observed, by risk quintile.

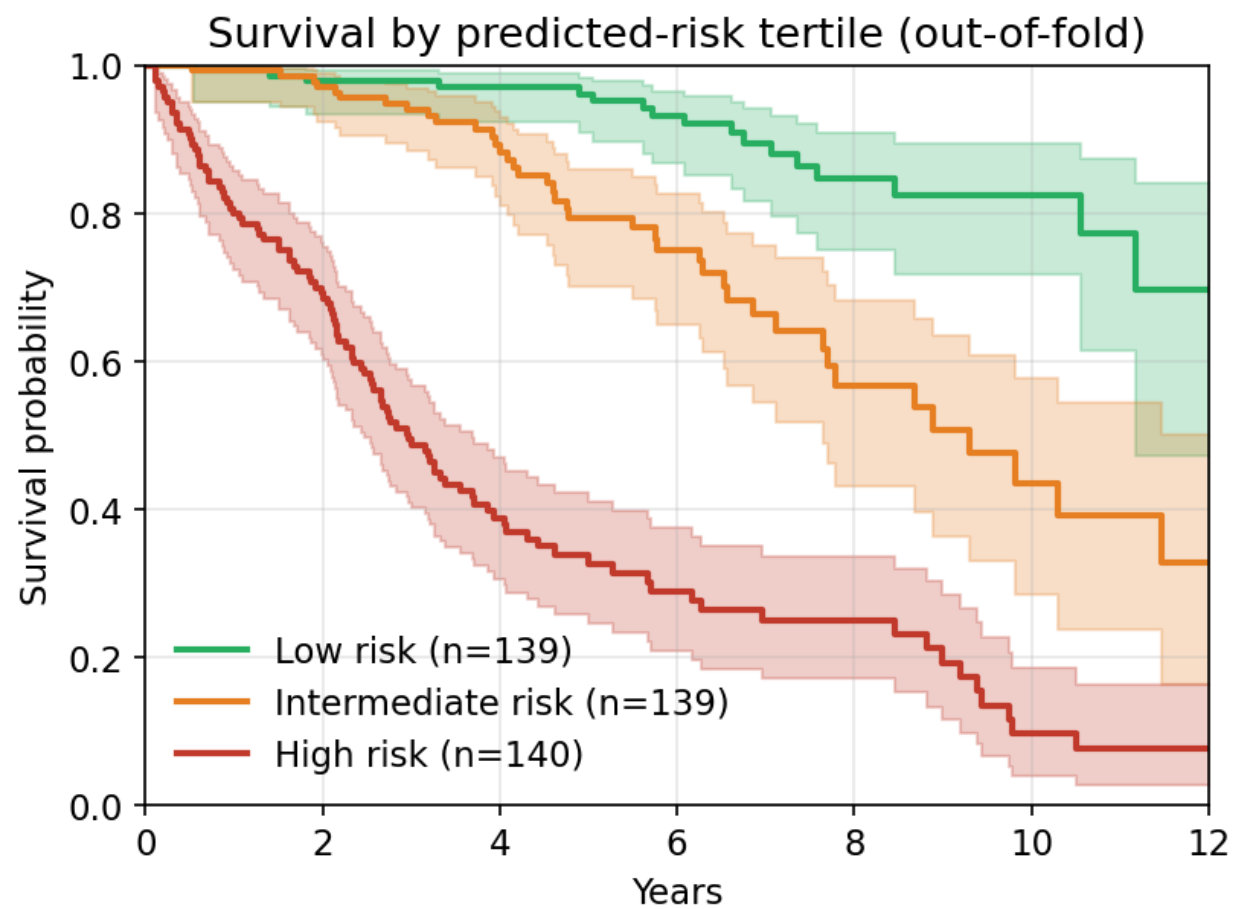


Figure S5: Kaplan–Meier survival by out-of-fold predicted-risk tertile.