The framing of time-dependent machine learning models improves risk estimation among young individuals with acute coronary syndromes

L.S.F. Carvalho[1], A.C.C. Nogueira[1], M.D. Fernandez[2], A.A.M. Soares[3], T.B. Rezende[4], S.E.F. Avila[4], G. Alexim[5], A.C. Sposito[4]

[1]Catholic University of Brasilia, Brasilia, Brazil
[2]Clarity Healthcare Intelligence, Jundiai, Brazil
[3]Aramari Apo Institute, Brasilia, Brazil
[4]State University of Campinas (UNICAMP), Campinas, Brazil
[5]University of Brasilia, Brasilia, Brazil
On behalf of B-CaRe:QCO

Funding Acknowledgements: Type of funding sources: Public grant(s) – National budget only. Main funding source(s): FAPESP and FAPDF

Introduction: Acute coronary syndrome (ACS) is a common cause of death in individuals older than 55 years. Although younger individuals are less frequently seen with ACS, this clinical event shows high recurrence rates and triggers considerable economic burden. Young individuals with ACS (yACS) are usually underrepresented and show idiosyncratic epidemiologic features compared to older subjects. These differences may justify why available risk prediction models usually penalize yACS with higher false positive rates compared to older subjects. We hypothesized that exploring temporal framing structures such as prediction time, observation windows and subgroup-specific prediction, could improve time-dependent prediction metrics. More specifically, we evaluated potential improvements in risk prediction quality among yACS by (i) developing models specifically in yACS versus a global cohort; (ii) splitting predictive rules into short- and long-term prediction windows [STWm and LTWm] versus a global follow-up model (GFm).

Methods: We included 6341 consecutive subjects with ACS (2242 with yACS) admitted into public hospitals in a large city in Brazil who undergone coronarography up to 48h after hospital admission from 2011 to 2020. The observation window in STWm and GFm included the first 48h upon hospital admission, and LTWm included all in-hospital information. yACS cohort was divided into train/validation-set (70%, n=1569) and test-set (30%, n=673); global cohort was divided in training/validation-set (70%, n=4439) and test-set including the 673 yACS subjects. Models were repeated over five cross-validation folds and then assessed in the test-set. C-statistics was the evaluation metric for STWm and time-dependent concordance (Ctd-index) for LTWm. STWm evaluated the occurrence of in-hospital cardiovascular deaths and recurrent ACS (MACE) and LTWm estimated events occurring post-discharge from index ACS hospitalization considering time-to-event with competing risks (MACE versus non-cardiovascular deaths).

Results: Among yACS and older subjects, respectively, in-hospital MACE occurred in 180 and 493 individuals, post-discharge MACE in 454 and 881; and post-discharge non-cardiovascular death in 47 and 285 after 6.67 years (95%CI 5.59-7.24) of follow-up. The best strategy was to design models specifically in yACS individuals combining STWm and LTWm. Among yACS subjects, STWm and LTWm yielded a C-statistics [0.921 (95%CI 0.889-0.953)] and Ctd-index [0.722 (95%CI 0.678-0.760)], respectively, while the best Ctd-index in GFm was 0.681 (95%CI 0.654-0.703). There was very low concordance among top predictors of MACE for prACS versus global cohort, as well as for STWm versus LTWm.

Conclusions: Predictive accuracy for adverse clinical events was optimized by using specific rules for yACS and splitting short-term and long-term prediction windows, leading to the detection of 80% of events, compared to 69% by using a rule designed for the global cohort.