Predictors on outcomes of cardiovascular disease of male patients in Malaysia using Bayesian network analysis

ABSTRACT

Objectives Despite extensive advances in medical and surgical treatment, cardiovascular disease (CVD) remains the leading cause of mortality worldwide. Identifying the significant predictors will help clinicians with the prognosis of the disease and patient management. This study aims to identify and interpret the dependence structure between the predictors and health outcomes of ST-elevation myocardial infarction (STEMI) male patients in Malaysian setting. Design Retrospective study. Setting Malaysian National Cardiovascular Disease Database-Acute Coronary Syndrome (NCVD-ACS) registry years 2006–2013, which consists of 18 hospitals across the country. Participants 7180 male patients diagnosed with STEMI from the NCVD-ACS registry. Primary and secondary outcome measures A graphical model based on the Bayesian network (BN) approach has been considered. A bootstrap resampling approach was integrated into the structural learning algorithm to estimate probabilistic relations between the studied features that have the strongest influence and support. Results The relationships between 16 features in the domain of CVD were visualised. From the bootstrap resampling approach, out of 250, only 25 arcs are significant (strength value ≥0.85 and the direction value ≥0.50). Age group, Killip class and renal disease were classified as the key predictors in the BN model for male patients as they were the most influential variables directly connected to the outcome, which is the patient status. Widespread probabilistic associations between the key predictors and the remaining variables were observed in the network structure. High likelihood values are observed for patient status variable stated alive (93.8%), Killip class I on presentation (66.8%), patient younger than 65 (81.1%), smoker patient (77.2%) and ethnic Malay (59.2%). The BN model has been shown to have good predictive performance. Conclusions The data visualisation analysis can be a powerful tool to understand the relationships between the CVD prognostic variables and can be useful to clinicians.