Abstract: P5705

Machine learning for phenotyping and risk prediction in cardiovascular diseases: a systematic review

Authors:
A Banerjee¹, S Chen¹, G Fatemifar¹, H Hemingway¹, T Lumbers¹, S Denaxas¹, ¹University College London, Institute of Health Informatics - London - United Kingdom of Great Britain & Northern Ireland,

INTRODUCTION: Heart failure (HF), acute coronary syndromes (ACS) and atrial fibrillation (AF) are among the commonest cardiovascular diseases (CVD), frequently co-exist and share pathophysiology. Definitions of diagnosis and prognosis are suboptimal. Machine learning (ML) is increasingly used in subtype definition and risk prediction, but the design, methods and results of studies have not been appraised.

PURPOSE: To conduct a systematic review of ML for discovery of new subtypes and risk prediction in HF, ACS and AF.

METHODS: PubMed, MEDLINE, and Web of Science databases were searched (January 2000-August 2018) for English language publications with agreed search terms pertaining to machine learning, clustering, CVD, subtype and risk prediction. The baseline characteristics of the study population, the method of ML, covariates and results were extracted for each study.

RESULTS: Of 5012 identified studies, 43 met inclusion criteria. Of the 33 studies of unsupervised ML for disease clustering (mean n=2354; min 117, max 44886), there were 22 in HF, 9 in ACS and 2 in AF. 22/33 studies involved <1000 individuals and 24 were based in North America. Across diseases, 27 studies were in outpatients, and 5 used trial data. The mean number of covariates used was 26; most commonly demographic and symptom variables. The ML methods used were partitional (n=12), hierarchical (n=4), self-organising map (n=1) and hidden Markov model (n=1). Most studies used only one ML method (n=25). Only 15 studies validated or replicated findings. 20/33 studies found 2 or 3 disease clusters, Most studies found 2-3 clusters (20/33) and most clusters were based on physical or physiological characteristics (30/33).

Of the 10 studies of supervised ML for risk prediction (mean n=43003; min 228, max 378256), 4 were in HF, 5 in ACS and 1 in AF. 2/11 studies involved <1000 individuals and most were from North America (n=6). All studies had an observational design, used at least 2 ML methods and validated or replicated findings. The setting was varied: primary care (n=2), emergency department (n=2), inpatient (n=4) and mixed (n=2). The mean number of covariates was 102. The commonest ML methods were neural networks (n=5), random forest (n=4) and support vector machine (n=4). All studies showed positive finding, i.e. ML approaches improved risk prediction.

CONCLUSIONS: Studies to-date of ML in HF, ACS and AF have focused on North America (68.2%), and 50% included less than 1000 individuals. Moreover, there is heterogeneity in clinical setting, study designs for data collection and ML methods used. Comparison between methods of ML and validation are common to studies of risk prediction but not disease clustering. There is likely to be a publication bias of ML studies in HF, AF and ACS. ML may improve data-driven characterisation of CVD but consensus guidelines for reporting of research using ML are urgently needed to ensure the internal and external validity and applicability of study findings.
Abstract: Machine learning for phenotyping and risk prediction in cardiovascular diseases: a systematic review

INTRODUCTION: Heart failure (HF), acute coronary syndromes (ACS) and atrial fibrillation (AF) are among the commonest cardiovascular diseases (CVD), frequently co-exist and share pathophysiology. Definitions of diagnosis and prognosis are suboptimal. Machine learning (ML) is increasingly used in subtype definition and risk prediction, but the design, methods and results of studies have not been appraised.

PURPOSE: To conduct a systematic review of ML for discovery of new subtypes and risk prediction in HF, ACS and AF.

METHODS: PubMed, MEDLINE, and Web of Science databases were searched (January 2000–August 2018) for English language publications with agreed search terms pertaining to machine learning, clustering, CVD, subtype and risk prediction. The baseline characteristics of the study population, the method of ML, covariates and results were extracted for each study.

RESULTS: Of 5012 identified studies, 43 met inclusion criteria. Of the 33 studies of unsupervised ML for disease clustering (mean n=2354; min 117, max 44886), there were 22 in HF, 9 in ACS and 2 in AF. 22/33 studies involved <1000 individuals and 24 were based in North America. Across diseases, 27 studies were in outpatients, and 5 used trial data. The mean number of covariates used was 26; most commonly demographic and symptom variables. The ML methods used were partitional (n=12), hierarchical (n=4), self-organising map (n=1) and hidden Markov model (n=1). Most studies used only one ML method (n=25). Only 15 studies validated or replicated findings. 20/33 studies found 2 or 3 disease clusters, most studies found 2–3 clusters (20/33) and most clusters were based on physical or physiological characteristics (30/33).

Of the 10 studies of supervised ML for risk prediction (mean n=43003; min 228, max 378256), 4 were in HF, 5 in ACS and 1 in AF. 2/11 studies involved <1000 individuals and most were from North America (n=6). All studies had an observational design, used at least 2 ML methods and validated or replicated findings. The setting was varied: primary care (n=2), emergency department (n=2), inpatient (n=4) and mixed (n=2). The mean number of covariates was 102. The commonest ML methods were neural networks (n=5), random forest (n=4) and support vector machine (n=4). All studies showed positive finding, i.e. ML approaches improved risk prediction.

CONCLUSIONS: Studies to-date of ML in HF, ACS and AF have focused on North America (68.2%), and 50% included less than 1000 individuals. Moreover, there is heterogeneity in clinical setting, study designs for data collection and ML methods used. Comparison between methods of ML and validation are common to studies of risk prediction but not disease clustering. There is likely to be a publication bias of ML studies in HF, AF and ACS. ML may improve data-driven characterisation of CVD but consensus guidelines for reporting of research using ML are urgently needed to ensure the internal and external validity and applicability of study findings.