Predicting the risk of rotator cuff tear using Bayesian networks

Bangyi Yuan[1], Yibin Liu[2], Jinjin Ma[2, 3]
[1]Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, Guangzhou, China, [2]School of Medicine, South China University of Technology, Guangzhou, China, [3]Institute of Future Health, South China University of Technology, Guangzhou International Campus, Guangzhou, China.

Disclosures: None

INTRODUCTION: The rotator cuff muscles are essential for maintaining shoulder joint stability and function, and injuries to the rotator cuff are a major cause of shoulder pain that greatly impacts quality of life. Bayesian networks, increasingly applied in healthcare, integrate expert knowledge with observational data and provide intuitive graphical representations of variable dependencies, with strong interpretability in quantifying probabilistic relationships. Accurate prediction of rotator cuff tears (RCT) is critical for diagnosis, management, and treatment. In this study, we developed a Bayesian network model using large-scale population data and expert knowledge to explore the interrelationships between RCT and its risk factors. The learned structure and conditional probability tables were implemented in GeNIe software to enable probabilistic inference and support clinical decision-making.

METHODS: In this study, a Bayesian network was constructed to model the risk of rotator cuff tears. Thirteen variables encompassing demographic, clinical, and lifestyle factors were included, such as age, sex, body mass index (BMI), Hyperlipidemia (HP), diabetes mellitus (DM), and smoking frequency. The network structure was learned from the health data of 435,093 participants in the UK Biobank using a hill-climbing algorithm combined with the Bayesian Dirichlet equivalent (BDE) scoring metric. Conditional probability tables were estimated through maximum likelihood estimation. The learned structure and parameters were implemented in GeNIe software, enabling probabilistic inference based on observed evidence. Model uncertainty was quantified through posterior distributions..

RESULTS SECTION: The Bayesian network analysis demonstrated that subacromial impingement syndrome (SIS) was the strongest predictor of rotator cuff tear (RCT). Patients diagnosed with SIS had an estimated risk of 28.82%, compared with only 0.95% among those without SIS, representing a nearly 30-fold increase. This underscores the critical role of SIS as a major risk factor and is consistent with previous clinical and epidemiological findings. Age was also an important determinant, with the risk of RCT increasing from 1.28% in individuals younger than 67 years to 2.51% in those older than 77 years, highlighting the contribution of degenerative changes with aging. Metabolic disorders further elevated risk: diabetes mellitus (DM), hyperlipidemia (HP), and hypertension (HT) increased RCT risk from 1.55%, 1.22%, and 1.27% to 2.13%, 2.06%, and 1.69%, respectively. Osteoporosis (OP) was associated with a modest but measurable increase, rising from 1.48% to 1.82%, suggesting that impaired bone quality may affect rotator cuff integrity. Lifestyle factors showed mixed associations. Exercise frequency did not demonstrate a significant relationship with RCT, with risks of 1.60% among those exercising rarely and 1.50% among those exercising almost daily ($p = 0.133$). In contrast, daily alcohol consumption was associated with a higher risk, increasing from 1.48% among non-drinkers to 1.84% among daily drinkers, indicating a potential adverse effect of alcohol on tendon health. Overall, the Bayesian network effectively quantified the relative contributions of demographic, clinical, and lifestyle variables to RCT risk and provided interpretable conditional probabilities that can be readily applied in clinical practice. Importantly, as larger datasets and additional evidence are incorporated, the precision of risk estimates is expected to improve. To enhance accessibility, the model was implemented in GeNIe software, allowing clinicians and even non-specialists to perform probabilistic inference and apply the results to individualized risk assessment and decision-making.

DISCUSSION: To enable risk prediction of RCT, support health management, and assist in clinical decision-making, we developed a risk prediction model based on a discrete Bayesian network. This model integrates expert knowledge with large-scale cohort data and enables posterior inference given observed evidence, thereby providing a quantitative basis for individualized diagnosis, treatment, and prevention strategies. The analysis demonstrated that advancing age, male sex, hyperlipidemia, osteoporosis, hypertension, subacromial impingement syndrome, bursitis, and higher frequency of alcohol consumption were all significantly associated with an elevated risk of RCT, whereas the influence of other variables was relatively minor. These findings suggest that clinical management and decision-making regarding RCT should prioritize these key risk factors. Moreover, the model exhibits strong interpretability, allowing clear quantification of the impact of each risk factor on disease risk and presenting posterior probabilities in a manner that is both intuitive and clinically applicable. For example, the model estimated that a 60-year-old female who consumes alcohol almost daily has a 1.81% risk of RCT, which could be reduced to 1.41% upon alcohol cessation, thereby underscoring its value in guiding health management. Nonetheless, the model has certain limitations, particularly regarding the accuracy of causal interpretations and the objectivity of variable selection, as it remains partially reliant on prior knowledge.

SIGNIFICANCE/CLINICAL RELEVANCE: This study presents a Bayesian network model to predict rotator cuff tear risk, providing an interpretable tool that quantifies the impact of clinical and lifestyle factors and supports individualized decision-making in patient management.

REFERENCES: 1.Bedi et al.Nat. Rev. Dis. Primers, 2024, 2. Ordovas et al.Comput. Methods Programs Biomed, 2023. 3.Shen et al. Exp. Sys. App., 2022. 4.Liu et al. BCM Msk Dis, 2024.

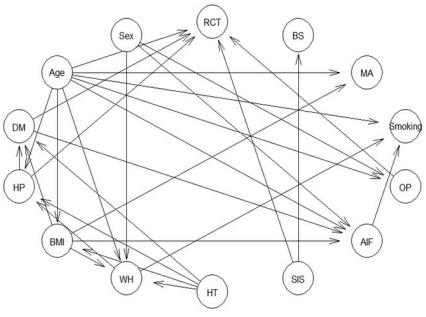

Fig. 1 The Bayesian network structure learned using the UKbiobank data

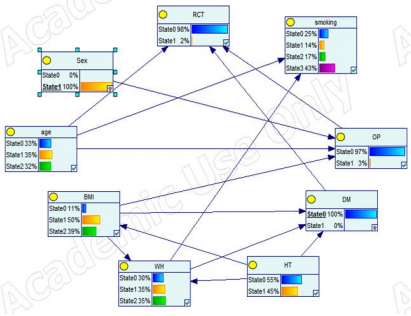

Fig. 2 Implementing RCT risk reasoning and visualization in the GeNIe software(Evidence: Female, No diabetes)

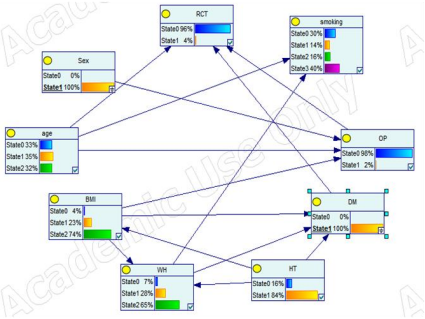

Fig. 3 Implementing RCT risk reasoning and visualization in the GeNIe software(Evidence: Female, Diabetes)