

Early risk-prediction models of Gestational diabetes mellitus in Chinese women using support vector machine model

D Ao (CN) [1], Y Wang (CN) [2], H Wang (CN) [3], X Zheng (CN) [4]

Context: A recent study reported a Gestational diabetes mellitus (GDM) prevalence of 19.7% in Beijing. This constitutes a huge health burden for health services. A simple, accurate early screening is thus central to effective prevention strategies. However, the studies on the risk factors of GDM in China started late. It was also lack of research on GDM prediction model. The support vector machine (SVM) has the potential to perform better than traditional statistical methods like logistic regression. We should study an effective early risk-prediction model of GDM in Chinese using SVM. **Objective:** To identify risk factors associated with GDM in Chinese and study predictive models of GDM using SVM that can help the clinician to identify women at high and low risk for GDM early in pregnancy. **Design and Setting:** We conducted a case-control study at a university hospital. **Participants:** 562 GDM cases and 453 controls were enrolled. **Methods:** The data was collected from the hospital computer database and self-administered questionnaires. Logistic regression models was used to analyze the demographic factors and life-style factors respectively. Then we selected the statistically significant risk factors according to the logistic regression results to study predictive models of GDM using SVM. **Main Outcome Measure?**GDM. **Results?** The demographic factors included age (≥ 30) (OR (95%CI) =1.42(1.10-1.85)), family history of diabetes (OR (95%CI) =1.80(1.30-2.48)) and pre-pregnancy overweight/obesity (OR (95%CI) =1.72(1.20-2.29)) was strongly associated with GDM. In life-style factors, 'less vegetable and fruits or much meat consumption' (OR (95%CI) =2.20(1.37-3.52)), 'sweets consumption' per week (OR (95%CI) =2.86(1.98-4.12)) and 'sedentary time at least 5h per day' before pregnancy (OR (95%CI) =2.13(1.45-3.12)) were strongly associated with GDM. 'Less vegetable and fruits or much meat consumption' (OR (95%CI) =2.73(1.65-4.51)) and 'less light-intensity activity per week' (OR (95%CI) =2.28(1.61-3.22)) during pregnancy were also associated with GDM. Subsequently, we adopted eight variables from former study results. In predictive models of GDM including the demographic factors and lifestyle factors, the highest test accuracy was 74.3%. **Conclusions:** Early screening for GDM in Chinese women can be provided by a combination of maternal demographic factors and lifestyle factors using the SVM modeling.

[1] Peking University , [2] Peking University , [3] Peking University , [4] Peking University