Introduction: Sudden cardiac death (SCD) is often the result of a sudden loss of blood flow caused by the failure of the heart to effectively pump. To prevent SCD, early prediction of SCD risk is essential. Utilization of electrocardiogram (ECG) markers to predict SCD has a long history. However, no study has been performed to discover the ECG features of SCD patients in different time regions by using 24-hours ECG data.

Methods: In this study, we collected 24 hours ECG data of 24 sudden cardiac death (SCD) samples, 42 high-risk SCD patients and 43 healthy controls. The samples were divided into training and the independent test datasets. In the training dataset, 24 SCD patients and 30 healthy controls were selected, and the remained 39 high-risk patients and eight healthy controls were used for independent test. The ECG data were divided into multiple fragments, and each fragment is one minute long. Within each ECG fragment, LASSO was used to build regression models in the training dataset by 5-fold cross validation. Finally, we used greedy searching method to construct a combine model.

Result: The 24-hours ECG data of individual was divided into 1440 time fragments. After the filtering, 925 time fragments were remained. In each time fragment, we compared 28 ECG features in the SCD patients and the controls by student's t-test, which discovered 13 features significantly different between them. The performance of the individual feature in discriminating SCD patients and controls along 24 hours were examined, which showed that wave duration related features and the R wave area performed the best in 8:00 AM to 12:00 AM, and the S wave area related features achieved the best performance in 2:00 AM to 4:00 AM. Integrating multiple features in each time fragment by the LASSO regression demonstrated area under the receiver operating characteristic curves (AUC) within 8:00 AM-12:00 AM provided the best AUCs. The same conclusion was obtained by the independent tests. By connecting the models within 8:00 AM to 12:00 AM, a unique model was constructed using greedy searching method, which achieved ACC 0.95 in five-fold cross validation test.

Conclusion: This study discovered the S and R wave areas as potential features to significantly discriminate the SCD patients with the controls in 2:00 AM to 4:00 AM and 8:00 AM to 12:00 AM, respectively. A reliable combined model was constructed by integrating multiple ECG features obtained in 8:00 AM to 12:00 AM.