as median-high risk (table 1). The sensitivities and specificities are as follows for each tool: QRISK2 (19%, 93%), FRS (22%, 93%), mFRS (46%, 83%), QRISK3 (47%, 78%), SLECRE (61%, 63%), respectively. The tools were similar in negative predictive value, ranging from 89% (QRISK2) to 92% (SLECRE). The FRS and mFRS had the highest c-statistics, both equalling 0.73, demonstrating the greatest predictive accuracy amongst the tools, while the QRISK3 and SLECRE had the lowest (0.67).

Conclusions While the mFRS performance was superior to the FRS, the QRISK3 did not outperform the mFRS. Although the SLECRE had the highest sensitivity, it had the lowest specificity, demonstrated by grouping the most cases and controls in the median-high risk category. Several factors are important to consider when deciding which risk assessment tools to utilize clinically: ease of use, sensitivity/specificity, and laboratory data accessibility. Thus, the mFRS continues to be a practical tool with a simple, intuitive scoring system appropriate for the ambulatory clinic setting based on the initial weighting of the FRS while adjusting for SLE.

Funding Source(s): Lupus Foundation of America - Gina M. Finzi Memorial Student Summer Fellowship

Cardiovascular Disease Risk Stratification by Risk Assessment Tool

Abstract 292 Figure 1

Background As the prototype of autoimmune disease, systemic lupus erythematosus (SLE) has complex and diverse clinical manifestations which may be harmful even life-threatening. Its pitiful that we can just passively respond to these serious complications. It will be a great advantage if the high-risk groups...
could be predicted and prevented with pre-treatment. The raising of risk prediction models depends on the collection of patient phenotypes, which are scattered in various forms and very cumbersome.

In this study, we collected the largest database of complete medical record of inpatients of lupus in China. The clinical phenotype database was generated by using natural language processing (NLP) techniques, then lupus nephritis (LN) prediction model was built.

Methods A total of 14,439 SLE patients were collected from the rheumatology and immunology departments of 13 Chinese tertiary hospitals in this study, including 13,062 females (90.46%), with an average age of 33.4 years, and the time span of EMR (Electronic Medical Records) was from October 28, 2001 to March 31, 2017. It includes basic information about patients, physical examination, inspection and diagnostic information, etc. We designed a hybrid NLP system combined NLP technical and expert knowledge at the same time, which was named as Deep Phenotyping System (DPS), to extract all the phenotypic information recorded in EMR. Based on these standard formatted entities, the machine learning and deep learning prediction methods are used to predict the LN in SLE.

Results The DPS efficiently processed EMR data, and its accuracy, precision, and recall were each greater than 93%. It extracted 73,794 entities from 14,439 SLE cases, each with time attributes, and produced 18,785,000,640 entities. Thus, a LN prediction model was raised, which the likelihood of lupus patients without nephritis will develop lupus nephritis within half and one year can be predicted. More than 35,000 phenotypes was used in this model and it was verified with independent samples. The best accuracy ACC and area under the curve (AUC) can be achieved 0.88 and 0.86 respectively.

Conclusions The comprehensive SLE phenotype database constructed by NLP greatly improves the research efficiency of lupus clinical phenotype. We first proposed a predictive model of lupus nephritis, which is high applicability and efficiency. The experimental results of good close and open testing fully demonstrate the authenticity and practicality of this database. The research process and method based on real world data are also applicable to predict other important complications of lupus.

Funding Source(s): None