O3

DEEP ANAESTHESIA: PREDICTING RISK OF CRITICAL PERIOPERATIVE EVENTS USING THE APRICOT DATASET AND DEEP NEURAL NETWORK MACHINE LEARNING

H. Lonsdale[1], A. Jalali[1], L. M. Ahumada[1], H. Yates[1], M. A. Rehman[1], The APRICOT Group[2]
[1]Johns Hopkins All Children's Hospital, St Petersburg, USA
[2]European Society of Anesthesiology Clinical Trial Network

Introduction

Although anaesthesia for children is very safe, rare severe perioperative adverse events (SPAEs) still have the potential to cause long term morbidity or mortality. Accurately predicting the risk of SPAE for each individual child could support decision making and perioperative planning. The APRICOT (Anaesthesia PRactice In Children Observational Trial) study was the largest observational trial to examine the incidence, nature and outcome of SPAEs in children. Two hundred and sixty one participating centers from 33 European countries contributed detailed information about 31 127 anaesthetics in 30 874 children. The study team reported the incidence of SPAEs- defined as an event requiring immediate intervention and that led (or could have led) to major disability or death- as 5.2% (95% CI 5.0-5.5). In this project, we aimed to use all available variables from the full APRICOT dataset to produce a deep neural network machine learning model that can predict the risk of any SPAE for every child undergoing an anaesthetic procedure.

Methods

Machine learning modelling consists of: pre-processing and feature extraction, data profiling, data exploration, model selection, model training and testing. We divided the full dataset into 70% for model training and 30% for model testing. We addressed class imbalance by using up-sampling of the minority (SPAE) class to increase the positive incidence rate to 25%. Hyperparameters were tuned using Bayesian optimisation. We developed multiple deep learning neural networks and selected the most accurate model by comparing model specificity and sensitivity in Area Under the ReceIver Operating Characteristic Curve (AUROC) charts.

Results

The model with the best results for the individual risk of SPAE demonstrated 82% accuracy, recall of 0.53, F-score of 0.24 and AUROC 0.75. The F-score is the harmonic mean of precision (positive predictive value) and recall (sensitivity) and is considered the most important performance measure for a machine learning model as it incorporates a measure of precision and recall values.

Discussion and conclusion

Machine learning offers clinicians the future potential to generate a precision medicine risk assessment for each individual child, based on a much wider range of parameters than a human brain can quickly and accurately assess. The rare nature of critical perioperative events means that very large datasets are required to power studies of causative factors. Machine learning is ideally suited to extract information from large datasets but the small proportion of SPAE within the study population represent a challenge in the form of class imbalance. This currently leads to reduced recall and subsequent impact on F-score. With ongoing refinement of deep learning techniques and methodology for managing class imbalance, we will further improve upon our ability to detect the children at most risk from anaesthesia, better informing the perioperative decision-making process.

Reference

Habre, W., Disma, N., Virag, K., Becke, K., Hansen, T.G., Jöhr, M., Leva, B., Morton, N.S., Vermeulen, P.M., Zielinska, M. and Boda, K., 2017. Incidence of severe critical events in paediatric anaesthesia (APRICOT): a prospective multicentre observational study in 261 hospitals in Europe. The Lancet Respiratory Medicine, 5(5), pp.412-425.