Hybrid Human and Machine Learning Adjudication Approach for Evaluating Arrhythmia Discriminator Algorithm Performance

Christopher Piorkowski
Fady Dawoud
Fujian Qu
Kyungmoo Ryu

**Introduction:** Implantable cardiac monitors (ICMs) are being increasingly used to diagnose cardiac arrhythmias including ventricular asystole (pause) and bradycardia (brady), leading to a vast increase in the number of episodes triggered and transmitted to remote monitoring clinics. Reviewing large amounts of ICM episodes can be challenging and time consuming for clinics. Machine learning algorithms have the potential to quickly adjudicate large amounts of data. The objective of this analysis is to assess the accuracy of a hybrid human and machine learning approach for adjudicating ICM episodes, and the feasibility of using this approach to evaluate recent detection algorithm enhancements in Confirm Rx™ ICM.

**Methods:** Pause and brady episodes triggered and transmitted to Merlin.net™ patient care network were extracted and used to train a random-forest gradient-boosting machine learning (ML) model. The ML model incorporated 15 relevant features derived from stored electrocardiograms such as R-wave amplitude and R-R intervals. Model convergence used 3-fold cross validation minimizing exponential loss. The human verification data subset was selected according to consistent criteria based on ML and ICM algorithm outcomes. The accuracy of the ML model alone and the hybrid approach combining the model-based result with human verification was evaluated using an independent fully adjudicated dataset.

**Result:** The training dataset contained 34,383 pause and 14,938 brady episodes from 9578 devices. ML model with 150 learners and learning rate of 0.15 achieved combined 99.1% sensitivity and 97.2% specificity in training. The validation dataset contained 2408 pause and 1502 brady episodes randomly sampled from 115 devices. The trained ML model alone achieved sensitivity of 100% and 97.6% with specificity of 96.3% and 91.6% in pause and brady episodes, respectively. Using the hybrid adjudication approach on the same data subset, sensitivity was 100% and 100% while specificity increased to 100% and 99.4%, respectively. Based on the outcomes of the ML model and ICM detection algorithms, the hybrid approach identified 359 (14.9%) pause and 311 (20.7%) brady episodes for human verification, thereby allowing 5.8 times reduction in human effort in comparison to manually adjudicating the entire validation dataset.

**Conclusion:** The hybrid adjudication approach leveraging machine learning techniques and human verification is highly accurate and significantly reduces human adjudication burden. It allows accurate performance evaluation of ICM detection algorithm enhancements.