Machine Learning for Healthcare 2019 – Clinical Abstract Track

Errors in Vital Sign Recording Within a Publicly Accessible ICU Research Database

H O'Halloran B Arts Sc, MD[1], R Veldhoen BSc, MSc, MD[1], DM Maslove MD, MS, FRCPC[1,2]
[1]Department of Medicine, Queen's University, Kingston, Ontario, Canada, [2]Department of Critical Care Medicine, Queen's University, Kingston, Ontario, Canada

Background

As big data becomes increasingly available in the healthcare field, there is great potential for machine learning algorithms. Data integrity, however, is an important consideration. We evaluated the eICU Collaborative Research Database, which offers researchers access to records from over 200,000 ICU stays at 208 hospitals across America. We searched for potential sources of error within vital sign readings within this dataset.

Methods

Using a structured query language, we extracted records the eICU database. One dataset comprised automatically-generated recordings, while the other involved information vetted by bedside nurses. Our analysis included blood pressure, central venous pressure (CVP), temperature, heart rate and oxygen saturation. We estimated clinically realistic ranges for each parameter, and identified constraints outside of which readings would be illogical. We then used the R statistical computing language to investigate values with respect to these parameters. Finally, we compared the rate of potential errors between database tables.

Results

Within the automatically entered data, 15.1% of CVP and 3.1% of temperature measurements fell outside of likely physiologic ranges, with a high probability of being errors. 0.85% of blood pressure values were determined to be likely errors. 99.95% of oxygen saturation values and 99.07% of heart rate values fell within a realistic range. Entries verified at the bedside were more likely than unverified monitor output to be realistic with only 4.31% of CVP values and 0.40% of temperature values being likely errors.

Discussion

The majority of vital sign measurements in the eICU database were clinically plausible. Common errors included the 3.1% of temperature readings that appear to have been entered in Fahrenheit rather than in Celsius. We also note a distribution of CVP values around 270 mmHg, which may be due to measurements being recorded while central venous lines are under pressure from infusions.

Both of the above errors were less common in the manually-verified readings than in the automatic recordings. The added check of bedside validation likely reduces the number of clearly unrealistic records, although this comes at the cost of manual data curation, which is time consuming and could introduce other biases.

With appropriate filtering for errors such as the above, the eICU dataset would be a strong candidate for large scale machine learning exercises.