# Effect of lupus on mortality rates

## Mixed effects model with random intercept and slope for lupus within hospitals

In this first model, we fit a generalized linear mixed model with a random intercept and a random slope for lupus within hospital, and adjusting for age category. We plot the histogram of the partially pooled log odds ratio estimates by hospital for lupus. This does not appear to show any systematic effect of lupus on mortality.



## A machine learning approach

The second approach is to mimic the idea of a standardized mortality ratio, which may be defined as the ratio between the observed number of deaths in an study population and the number of deaths would be expected, based on the age- and sex-specific rates in a standard population and the age and sex distribution of the study population. In our case, we’ll take the non-lupus patients to be our “standard” population, and train a machine learning model to predict **death based on age, ventilator use, comorbidities, insurance, and gender**. We also included dummy variables for hospitals in case there was a significant difference in pattern between hospitals. We then used this model to predict the probability of death for lupus patients (noting that by including hospital in the predictive model we are effective nesting within hospital if that hospital was predictive of death). This prediction is what we would expect in the lupus patients if they had the same experience as non-lupus patients. We can then compare this with the actual outcome of the lupus patients.

The machine learning model we used in this analysis is XGBoost (Chen and Guestrin 2016). This popular machine learning model is a regularized version of gradient boosted trees that is meant to have low bias. It also is consistent under general conditions (Biau and Cadre 2017). XGBoost is computationally very fast and has been used in many data science applications and competitions as an optimal prediction engine.

The first plot (below) is a predictive calibration plot. This plot shows, on the x-axis, ranges of predicted probabilities, and plots the proportion of observations with predicted probabilities in that range with a 95% confidence interval. The reference line is to show perfect calibration. The idea is that if you look at individuals with predicted probabilities of death between 0.1 and 0.2, for example, you expect the observed proportion of them dying to also be in that range if you have good calibration. What we are looking for here is the opposite, i.e., evidence of **poor calibration**. Why? If the predictions are well calibrated then that would imply that there is no difference in experience between the non-lupus subjects (who are the basis of the predictions) and the lupus patients. What we see here is that, on an individual, overall level, there probably is not very much difference in experience between the two groups.



Figure 1 Predictive calibration using xgboost

The second plot (below) shows the distribution of the hospital-level SMR-like statistics. It is clear here that there are several hospitals which, in aggregate, do poorer with lupus patients than non-lupus patients. 1.7% had a ratio over 5, while 16.5% had a ratio over 2 (i.e. number of observed deaths was at least twice what was expected if there was no mortality difference between lupus and non-lupus patients). Note that this study is restricted to hospitals which had at least 5 admissions of lupus patients with sepsis.



## References

Biau, G. and B. Cadre (2017). "Optimization by gradient boosting." arXiv preprint arXiv:1707.05023.

Chen, T. and C. Guestrin (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA, ACM**:** 785-794.

Table 1: Individual level (n, age, sex, elix\_score

Plot of SMR