Appendices

1   Physiological Attributes and Preprocessing

Physiological attributes, treatments and corresponding preprocessing methods

Preprocessing	Attributes
Standardization	`age,Weight_kg,GCS,HR,SysBP,MeanBP,DiaBP,RR,Temp_C,FiO2_1,` `Potassium,Sodium,Chloride,Glucose,Magnesium,Calcium,Hb,` `WBC_count,Platelets_count,PTT,PT,Arterial_pH,paO2,paCO2,` `Arterial_BE,HCO3,Arterial_lactate,SOFA,SIRS,Shock_Index,PaO2_FiO2,` `cumulated_balance_tev, Elixhauser, Albumin, CO2_mEqL, Ionised_Ca`
Log transformation	`max_dose_vaso,SpO2,BUN,Creatinine,SGOT,SGPT,Total_bili,` `INR,input_total_tev,input_4hourly_tev,output_total,output_4hourly`

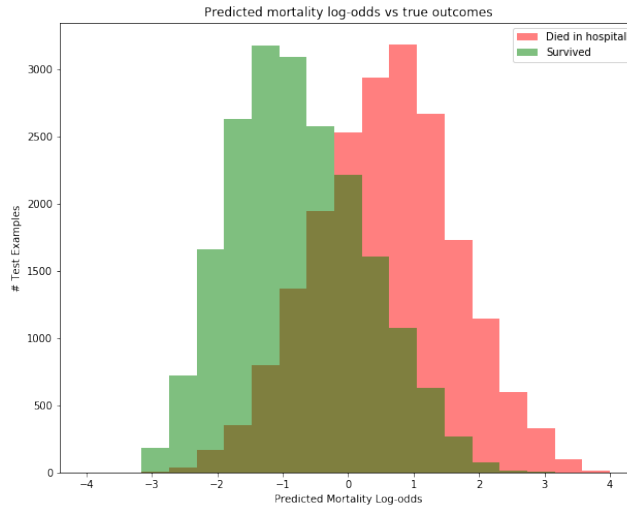## 2   Predicted vs True Log-odds Mortality

Figure 1: Mortality log-odds distribution for mortality and survivor classes.

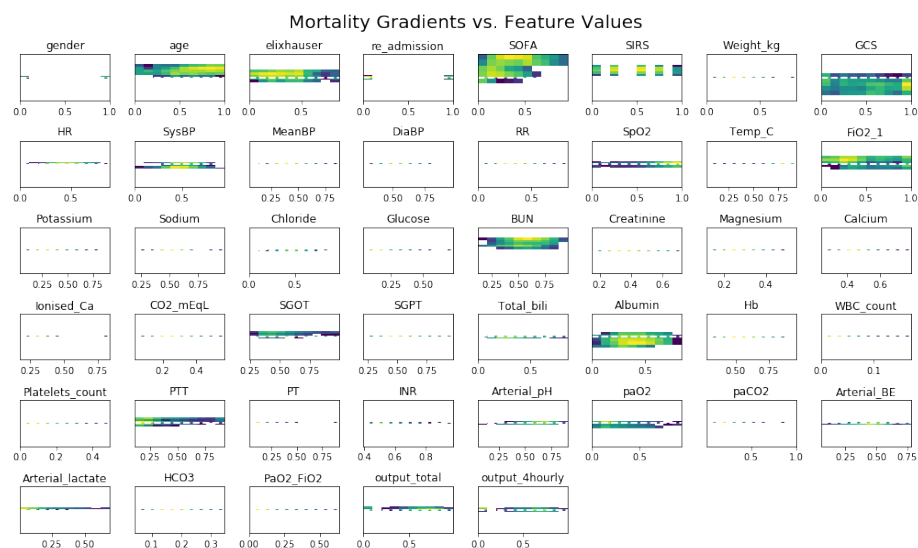# 3   Mortality Gradients vs. Feature Values

Figure 2: Mortality gradients versus feature values. No single observation value is highly correlated with mortality.

4 Variance of Policy Estimates

The following plots show the distribution of the WDR estimator difference between MoE and three other policies based on 1000 bootstraped test datasets. The red vertical lines indicate the difference calculated for the original test dataset.
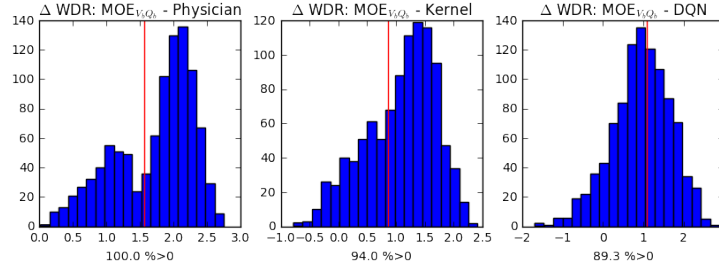
On average, the WDR estimator predicts the MoE policy to outperform all other proposed policies. We note, however, that for some bootstraped datasets the difference is negative. This is to be expected, as Gottesman et al. [24] demonstrate that high variability of IS based estimators is common in healthcare-data. Gottesman et al. [24] further demonstrate that IS based estimators can have a high selection bias when evaluating policies which are significantly different from the behavior policy.