COPYRIGHT © BY THE JOURNAL OF BONE AND JOINT SURGERY, INCORPORATED RATLIFF ET AL.

PREDICTING OCCURRENCE OF SPINE SURGERY COMPLICATIONS USING "BIG DATA" MODELING OF AN ADMINISTRATIVE CLAIMS DATABASE

http://dx.doi.org/10.2106/JBJS.15.00301

Page 1

## **Appendix**

## LASSO Regularization

LASSO (least absolute shrinkage and selection operator) is a tool for "regularization" in statistics. Regularization involves tuning or selecting the preferred level of model complexity to improve model prediction. It is effective when working with multiple predictors that may be correlated.

LASSO is almost the same as incorporating a special form of prior knowledge favoring estimates that are close or equal to zero in estimating the unknown regression coefficients. The ordinary maximum likelihood estimators for regression coefficients based on the data are shrunk toward zero by penalizing the parameter estimates. The larger the penalty, the further the estimate is shrunk toward zero. The features with zero coefficients after the shrinking are automatically dropped from the prediction model.

The extent of the applied penalization or degree of shrinkage depends on the size of the penalty parameter. In our analysis, we selected the penalty parameter by  $\kappa$ -fold cross-validation.

## Example of Score Calculation

In this example, a 25-year-old female patient is undergoing an anterior cervical discectomy and fusion. The patient's only comorbidity is hypertension. The preoperative diagnosis is degenerative disease. BMP is not used in the procedure. A single level is treated, with instrumentation.

The risk score for this patient is calculated by adding the numbers in the relevant boxes under "Anterior Cervical Procedure" in Table IX: -1.97 (age 20-29 years) +0.05 (female) -0.69 (degenerative disease) +0.14 (instrumentation) +0.19 (fusion) +0.22 (hypertension) =-2.06. The risk of a complication is calculated as  $\exp(-2.06)/\{1+\exp(-2.06)\}=11.3\%$ .