Predicting Necrotizing Enterocolitis in Premature Infants
Peter Kang, Jonathan Palma, Stephen Pan, Mark Smith

Introduction
Necrotizing enterocolitis (NEC) is characterized by death of intestinal tissue, and affects about 7% of very low birth weight (VLBW) infants (defined by birth weight <1500 grams). The inflammatory process in the bowel can have profound systemic effects, which lead to the significant morbidity and mortality (up to 30%) of the disease. NEC typically presents at 8 to 10 days of age, and its severity is described by a staging system, where stage 1 includes nonspecific clinical findings; stage 2 has specific radiographic findings; and stage 3 is defined by bowel perforation, a serious complication with increased morbidity and mortality that requires surgical intervention. Despite advances in the field of newborn intensive care, the likelihood of a VLBW infant developing NEC has remained nearly constant, likely because the pathophysiology of this multifactorial disorder is incompletely understood.

De-identified laboratory and physiologic (vital signs) data from premature infants were used to develop classification models for the disease. In addition, NEC cases of varying clinical severity were visualized using principal component analysis (PCA). An algorithm to predict onset of NEC could aid in earlier diagnosis, prevention, and/or mitigation of the disease. Understanding the features important to the model’s prediction may enhance our understanding of the disease’s biology.

Methods
De-identified clinical data, including diagnosis codes, laboratory results, and vital signs were obtained from the Stanford Translational Research Integrated Database Environment (STRIDE). These data represent patient encounters since 2008 in the 40 bed neonatal intensive care unit at Lucile Packard Children’s Hospital. Clinical data collected during the 5 days prior to NEC diagnosis were compared to 5 day sliding windows of control data which matched the distribution of age at NEC diagnosis.

Sliding Window Selection in Controls

![Diagram](image)
The 18 case patients were those with diagnosis codes for any of the three specific clinical stages of NEC. The 44 controls were those with codes for prematurity but not NEC. Although data existed for more than 700 labs, the vast majority were only available for a small subset of the patient population. The following graph conveys that for more than 600 of the different labs, data was available for fewer than 40 of the total 350 unfiltered patients.

We examined the test data using domain expertise, and after eliminating infrequent, unimportant, or redundant lab tests, 39 distinct tests remained. In order to incorporate temporal trends prior to diagnosis, we took the daily average of each feature and calculated slope values between each day. We then calculated descriptive statistics (min, mean, max, SD, quartiles, medians) for all features (labs, vitals, and their respective slopes). This process yielded a total of 645 features for our 62 examples. Missing values (approximately 2.8% in control group, none in the cases) were imputed by assigning mean values from cases or controls.

First, we visualized a group of controls and specific cases of stage 1, 2, and 3 NEC using PCA. We then applied multiple machine learning approaches with various parameter adjustments to our data, including k-nearest neighbor (knn), support vector machines, logistic regression, and lasso, but ultimately chose to use elastic net regularization and feature selection with an alpha parameter = 0.5. This was implemented using the glmnet package in R. We divided our data into equal-sized training and test sets, then performed 10-fold cross validation during generation of the model. The cross-validation data was used to select a cutoff above which we designated cases positive.
Results

K-Nearest-Neighbor
Our initial attempt to predict NEC onset using a knn algorithm did not perform as well as we hoped. The knn pipeline was as described above. After filtering patients we were left with 219 controls and 21 cases. We pruned the lab values using data from a preliminary run of elastic net (see below) in conjunction with expert knowledge in the medical domain to determine features that would be more likely to predict NEC. We established a “neighborhood” using a randomly selected 80% of the data set, and tested with the remaining 20%. We varied the value of k from 1 to 31, and because of the large class imbalance, we required that 219/240 (or 91.25%) of the neighbors must be control for a given data point to be classified as control.

![Graph showing errors on test set](image)

The knn algorithm was not well-suited to this data, and we were not able to achieve satisfactory classification. False positive rate and false negative rate are shown in the above graph.

PCA
We then turned to principle components analysis to further visualize the data and to help give some sense of the likelihood of successful classification. Cases and controls were further filtered down to a set of 21 confirmed cases of NEC of varying severities and 54 confirmed controls whom had the most features available in the raw data set. Features used for the PCA analysis were selected similar to what was described above with the addition that only features present in >30% of cases/controls were used. Mean values of either cases or controls were used to impute missing values in their respective cohort.
We found that the first three principle components explained ~30% of the total variance in this data set. When plotted, the results were quite encouraging, with clear segregation of cases and controls when PC2 is plotted against PC1. While the segregation of severities of NEC is less distinct, there appears to be some progression in NEC severity going from upper left to bottom right in the figure above. In fact, the three cases of NEC stage III closest to the bottom right and denoted by the arrows all suffered from intestinal perforation, arguably the most severe complication of NEC.
Elastic Net
We thus aimed to use an elastic net classification model to both predict NEC as well as to determine what features were most important for predicting this disease. For the training set, our resulting model had 100% sensitivity and 100% specificity and 0% misclassification error. When applied to our test set, one case was misclassified, yielding a sensitivity of 88.9% and specificity of 100%; misclassification error on the test set was 0.018%.

Training Set Classification

Test Set Classification
The model utilized 19 features, which represented 11 distinct laboratory test and one vital sign. The distinct features included alkaline phosphatase, alanine transaminase, aspartate aminotransferase, total and direct bilirubin, globulin, hematocrit (HCT), mean corpuscular hemoglobin (MCH), mean corpuscular volume (MCV), red blood cell count (RBC), total protein, and heart rate.

**Discussion**

We were successful in classifying cases of necrotizing enterocolitis based upon clinical data collected as a byproduct of routine neonatal care during the 5 days prior to the development of disease.

After training several machine learning methods, we ultimately used elastic net as a means for both feature selection and regularization. Because elastic net performs better than other methods such as lasso in datasets where the number of predictors is greater than then number of examples, it was well suited to the relatively small number of patients in our clinical dataset. In addition, because elastic net tends to either include or exclude groups of strongly correlated predictors from the model together, this method of feature selection might be more useful when seeking a biologic explanation for the model.

Given the small size of our data set, further validation of the model will be important. Nonetheless, it is interesting to consider patterns in the features selection and their possible relationship to NEC. Globulin levels relate to immune system function, and can be altered in the setting of infection or inflammation, both of which are thought to play a role in NEC. Several hematologic labs (e.g. HCT, MCV) were present, and altered blood flow to the intestines may also be a factor in NEC. Recent research has highlighted the utility of using physiologic data to predict outcomes in premature infants. Indeed, heart rate, which is also linked to tissue perfusion, was a feature selected in our model.

This work could serve as a foundation for predicting NEC before it develops, potentially allowing the disease to be mitigated or prevented altogether.