Identification of Parameters Predictive of Chromosomally Normal Embryos

Livia Zarnescu Carmichael Ong Apoorva Rajagopal

December 13, 2013

1 Introduction

Infertility affects nearly 1 in 6 couples in the United States, many of whom turn to in vitro fertilization (IVF) to have children. During in vitro fertilization (IVF), embryos are cultured for 5 days after fertilization and a clinician chooses one to transfer back to the patient in hopes that a pregnancy will result. This is a difficult problem because many embryos which appear morphologically normal at day 5 may not be viable and/or may be chromosomally abnormal.

In the past few years time-lapse imaging of embryos has emerged as a technique to aid in viability prediction. In particular, three cell cycle parameters describing the embryo's cell division over time have been found to be predictive of embryo survival to day 5 in culture (green and yellow data points in Figure 1) [Wong 2010]. Although these cell cycle parameters can predict viability to some extent, they are only weakly predictive of aneuploidy (chromosomal abnormality) in embryos. Avoiding the transfer of aneuploid embryos to patients is important because aneuploid embryos may appear morphologically normal at day 5, but the majority are nonviable and ultimately result in miscarriage. Furthermore, while it is possible to test chromosomal composition in the clinic, the procedure is invasive and damaging to the embryo.

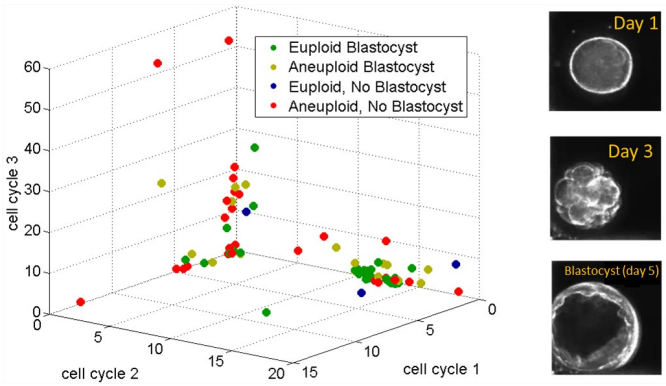


Figure 1. Cell cycle parameters (1st cytokinesis duration, time between 1st and 2nd mitosis, time between 2nd and 3rd mitosis) used in previous work to predict embryo viability.

Our data set contains time-lapse videos of 188 human embryos from days 1 to 5 post-fertilization along with their chromosomal composition. We thank Maria Rodriguez for her help in gathering this data. Our goal is to use non-invasive measures (time lapse parameters and morphological assessments of the embryos) that can be extracted from the video to identify feature combinations that are predictive of euploid (chromosomally normal) embryos by day 5.

2 Methods

To accomplish our goal, we first extracted various features of interest from the time-lapse videos of the embryos. We then used the classification methods of k-Nearest Neighbors (kNN), Support Vector Machine (SVM), and logistic regression and feature selection methods to compare the performance between the classification methods and determine which is best suited for our problem.

2.1 Features Extracted

Features were extracted from the videos manually. In total, 30 features were used which included timing between all cell divisions, cell fragmentation, abnormal divsions, duration of cavitation and compaction, and morphological scores.

2.2 Classification Techniques

We used three methods to classify our data. First, SVM with a Gaussian kernel and l_1 regularization was used. We used forward feature selection on subsets of parameters as described in the next section to find the best combination of features to use. Next, weighted-kNN, using the inverse of the square of the Euclidean distance, was evaluated. For this method, all 30 features were used, and the number of nearest neighbors considered for classification was varied between 1 and 10. Finally, logistic regression using the best set of features found from forward feature selection with SVM was used to classify the data.

To evaluate the performance of each classification method, the area under the receiver operating characteristic (ROC) curve was calculated for the different models within each method. The value of the model with the largest area under the ROC curve was then used as the performance metric for that classification method.

2.3 Feature Selection

Previously published work has found three cell cycle parameters that predict embryo viability: 1^{st} cytokinesis duration, time between 1^{st} and 2^{nd} mitosis, and time between 2^{nd} and 3^{rd} mitosis. We performed forward feature selection with 4 different groups of features. We first compared our methods to literature results by implementing forward feature selection using only the previously mentioned three features. We then compared the performance of this group with 3 groups of features extracted from our data: all 30 features, features before day 3 only (early in development), and features after day 3 only (later in development). Forward feature selection was performed within each of these 4 groups using an SVM classifier. In each iteration of forward feature selection, we used the following method to determine the best feature to add to the feature set:

1. We start with a feature set $\mathcal{F}$ and add a new feature $i \notin \mathcal{F}$ to create a new feature set $\mathcal{F}_i = \mathcal{F} \cup \{i\}$.
2. For $j = 1, ..., N$ {
   (a) Randomly split the data set S into a training set, $S_{j,\text{train}}$, and a hold-out cross-validation set, $S_{j,\text{cv}}$, where $S_{j,\text{train}}$ contains 90% of the data in S.
   (b) Optimize the parameters of the SVM (i.e. the variance of the Gaussian Kernel and the regularization constant (a.k.a. box constraint) for the slack variables on the funcional margin constraint) using $S_{j,\text{train}}$.
   (c) Use the optimized SVM classifier on $S_{j,\text{cv}}$ and calculate a ROC curve based on the orthogonal displacement from the decision boundary.
   }
3. Calculate $AUC_{i,\text{avg}} = \frac{1}{N}\sum_{j=1}^{N} AUC_j$, where AUC_j is the area under the j^{th} ROC curve.

We then add the feature with the greatest value of $AUC_{i,\text{avg}}$. Note that we used the average of the area under the ROC curves from multiple ($N = 100$) hold-out cross-validations because the training was sensitive to the data that was held-out.

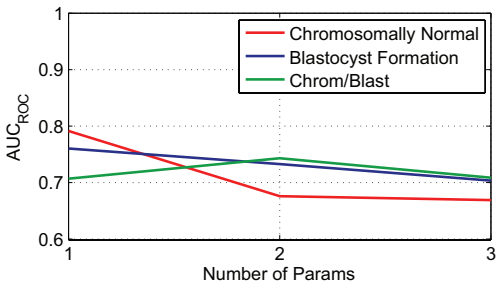
3 Results

3.1 SVM

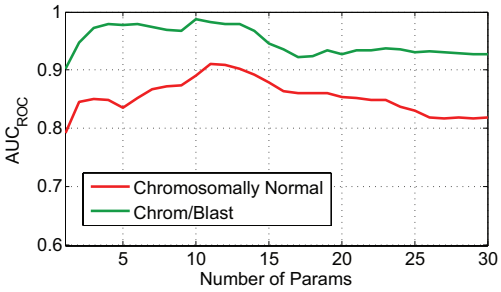
We found that combinations of our features could predict chromosomally normal embryos better than the cell cycle parameters commonly used in clinics (Figure 3.1).

Figure 3.1

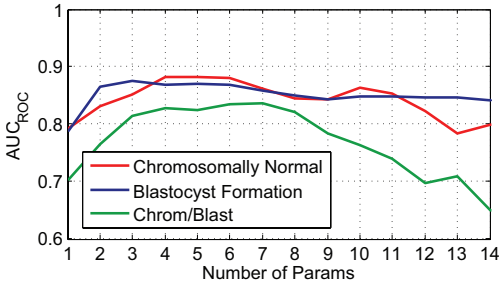
(a) Using cell cycle params



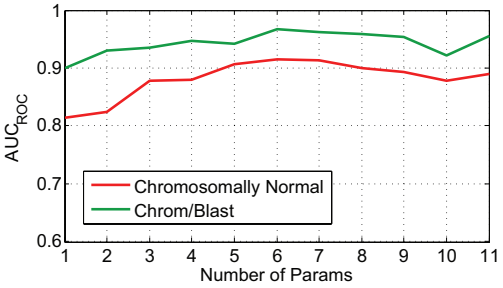
(b) Using all params



(c) Using params before day 3



(d) Using params after day 3



3.2 kNN

Weighted-kNN worked best using the nearest 5 neighbors (Figure 3.2). With $k = 5$, the area under the ROC is 0.82 when only predicting normal chromosomal composition, and 0.86 when predicting chromosomally normal blastocysts.

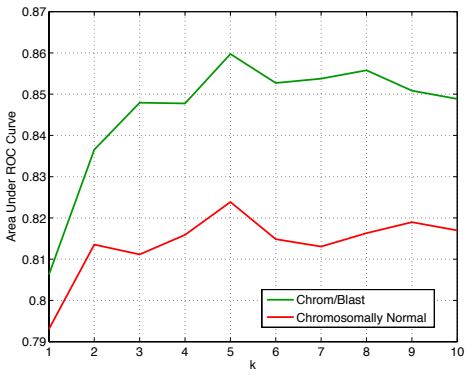


Figure 3.2. Area under ROC curves for different values of k, averaged over 10 iterations of 10-fold-leave-out testing.

3.3 Logistic Regression

Logistic regression was very sensitive to redundant data and ill-conditioning of our feature matrix. This method was best able to predict chromosomally normal blastocysts when using the 5 best features from forward feature selection (average area under ROC curve = 0.96) and was best able to predict chromosomal composition only using the best 17 features from forward feature selection (area under ROC curve = 0.89). (Figure 3.3)

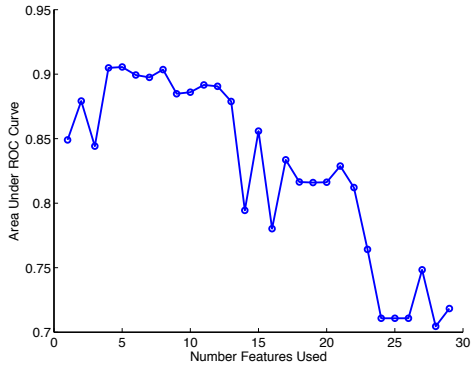


Figure 3.3. Area under ROC curves for different values of k, averaged over 10 iterations of 10-fold leave-out cross-validation.

4 Conclusion

Comparing the results from the three classification methods tested (best combination of all features), we see that SVM performs slightly better than kNN and LR.

Method	Area under ROC Curve	
	Chrom Norm	Chrom Norm & Blast
kNN	.82	.86
SVM	.91	.99
LR	.89	.96

Parameters from later in development (after day 3) achieved significantly better classification performance compared to cell cycle parameters or other parameters from early in development (before day 3). The best parameters were blastocyst timing, time to cavitation, compaction duration, # of cells before compaction, and blastocyst morphology. Of the early parameters, only the second cell cycle parameter was useful. Other combinations of early parameters can achieve significantly better performance if transfer before day 3 is to be performed, and late parameters should be used if day 5 transfer is to be performed.

References

1. Wong, CC et al. Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. Nature Biotechnology 28, 1115-1121 (2010).