Proceedings of the 2005 IEEE
Engineering in Medicine and Biology 27th Annual Conference
Shanghai, China, September 1-4, 2005

Predictive Models for Long Term Survival after Premature Rupture of Membranes

Lieveke Ameye, Jos De Brabanter, Johan A. K. Suykens, *Member, IEEE* , Isabelle Cadron, Roland Devlieger, Dirk Timmerman, Bernard Spitz, Sabine Van Huffel, *Member, IEEE*

Abstract— **Premature rupture of the membranes complicates 30 to 40% of the preterm deliveries. The purpose of this study is to assess the outcome of preterm prelabor rupture of the membranes (PPROM) before 26 weeks of gestation and to develop a predictive model for survival. We apply four statistical methods in order to construct a predictive model: Cox's Proportional Hazards model, Logistic Regression, Least Squares – Support Vector Machines and Generalized Estimating Equations. The obtained predictive models are tested on a prospective set and have a very good performance of > 95% for the area under the ROC-curve.**

I. Introduction

PREMATURE rupture of membranes (PROM) is defined as rupture of the membranes prior to the onset of labor. This is a fairly common event, occurring in approximately 10% of all pregnancies. However, it is estimated that PROM complicates 30 to 40% of the preterm deliveries. In this article only preterm prelabor rupture of the membranes (PPROM) at or before 26 weeks completed of gestation is studied. PPROM before 26 weeks of gestation is uncommon, but has a major impact on fetal and maternal health and challenges the process of medical decision making. The management of PPROM, especially at early gestational ages, is complex. The purpose of this study is to develop a predictive model for survival.

The medical records of 119 consecutive pregnancies of PPROM at or before 26 completed weeks of gestation, including 148 babies, were retrospectively analyzed. All women were admitted with a living fetus at the university hospital Gasthuisberg Leuven between 1988 and 1996. For each baby 25 explanatory variables are determined, e.g. the gestational age at the time of rupture of the membranes (in weeks), time interval between rupture and delivery (in days), maternal age, administration of steroids or not, multiple pregnancy or not, sex of the baby. The outcome is the number of days the baby survived. In case the baby died before he/she was born, the outcome specifies whether it was antenatal or intranatal. 7% of the 148 babies died antenatally, 20% intranatally, 24% on the first day, 6% from day 2 till day 7, 4 % from day 8 till day 28, 6% after the first month and 35% of the babies survived.

In 1999 a data analysis has already been performed on this dataset of 148 babies [1]. For that analysis the outcome was summarized into a 0 - 1 variable: 0 if the baby died before the end of the first month, and 1 if the baby survived the first month. Corrections were made for the fact that members of twin or triplet pregnancies do have the same mother by using Generalized Estimating Equations (GEE) [7]. The variables of the obtained model determining the survival chances were: the birth weight of the baby, the sex of the baby, the gestational age at the time of rupture of the membranes, and administration of steroids or not.

In the period ranging from 1997 until 2000 a prospective dataset, containing the medical records of 52 consecutive pregnancies (64 babies) with PPROM, has been constructed. The goal of the study is to develop a prediction model for survival based on the variables measured during the pregnancy using the records in the period 1998-1996 and to test the performance of these prediction models on the prospective set (period 1997-2000).

II. Materials and Methods

The analysis of this dataset is quite complicated. Firstly, not all the variables are measured at the same time. For example, the gestational age at the time of PPROM is already known at the first medical examination, while the administration of steroids is a process variable measured during the time interval between PPROM and delivery. The purpose of this study is to construct a prediction model for survival as early as possible during the pregnancy, i.e. a model based on process variables measured at the first medical examination is clinically more useful than a model based on factors measured after the baby is born.

Secondly, the dataset contains also records of twins and triplets, where one baby may influence the other. Out of the 119 pregnancies, there are 93 (78.2%) single pregnancies and 26 (21.8%) multiple pregnancies, of which 23 twin pregnancies and 3 triplet pregnancies. For babies of multiple pregnancies, the factors concerning the mother and the

Manuscript received April 7, 2005. This work was supported by the Research Council of the K.U.Leuven (GOAMEFISTO666, GOA-AMBioRICS), the Flemish Government (FWO: projects G.0407.02 / G.0269.02 / G.0360.05, research communities ICCoS & ANMMM), the Belgian Federal Government (DWTC: IUAP IV-02 (1996-2001) and IUAP V-22 (2002-2006)), and the EU (PDT-COIL: contract NNE5/2001/887; BIOPATTERN: contract FP6-2002-IST 508803; eTUMOUR: contract FP6-2002-LIFESCIHEALTH 503094).

Lieveke Ameye, Jos De Brabanter, Johan Suykens and Sabine Van Huffel are with the Electrical Engineering Department, K.U.Leuven, Belgium (e-mail: Lieveke.Ameye@esat.kuleuven.ac.be).

Isabelle Cadron, Roland Devlieger, Dirk Timmerman and Bernard Spitz are with the Department of Obstetrics and Gynecology, K.U.Leuven, Belgium.

0-7803-8740-6/05/$20.00 ©2005 IEEE.

process variables measured during the time interval between PPROM and delivery are the same but the amniotic fluid volume after PPROM, the birth weight, sex and apgar score 1 and 5 minutes after birth (the apgar score is a value from 0 to 10 for the general health of the baby) can differ for the babies of the same pregnancy. Remark that the survival chances are higher in the group of single pregnancies than in the group of multiple pregnancies. This is illustrated in Fig. 1. If we look at this figure, we see that 44.1% of the babies of a single pregnancy have survived, while in the group of multiple pregnancies this is only 20%. In this last group significantly more babies die during the birth process. Fig. 2 illustrates that babies of a twin or triplet do not have the same survival chance. For example: twin 1: one baby has died on day 1 and one baby has died after the first month, twin 5: one baby has died on day 1 and one baby has survived, triplet 25: one baby has died during the first month and two babies have survived.

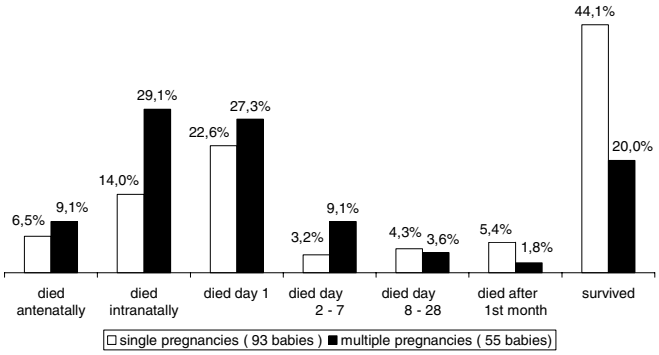


Fig. 1 Comparison of the survival chances between babies of single pregnancies and babies of multiple pregnancies

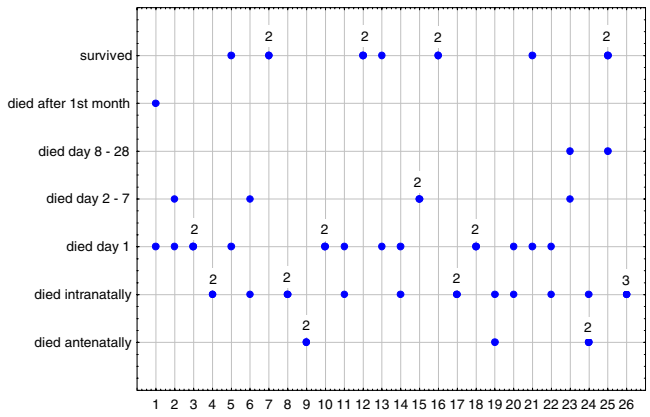


Fig. 2 Outcome of 55 babies of 26 multiple pregnancies. The numbers 1 – 23 indicate twins, 24 – 26 indicate triplets.

In a first part of the analysis we apply a survival model: Cox's Proportional Hazards model [2,3]. When constructing a survival model, a time origin has to be specified. We take time of birth of the baby as time origin. Therefore, we omit the records of babies who died before their birth (approximately 20% of the data).

In a second part we focus on predicting the long term survival chance, i.e. predicting the survival of the first month, using classification models. We follow two ways in order to construct prediction models for this classification problem: 1) The most common way to construct a predictive model based on the records of the 148 babies, where we have to take into account the dependency between babies of twins and triplets, is by using Generalized Estimating Equations (GEE) [7]. We can get robust variance estimates by specifying the working correlation matrix. 2) Another way to look at the problem is to see the data as records of 119 pregnancies. Now, we can consider all observations as independent and just apply ordinary Logistic Regression (LR) [5,6,15]. In case of a multiple pregnancy, we define the outcome as the outcome of the baby that survived the longest. Therefore, in case of single pregnancies we only examine whether the baby survives the first month, while in case of multiple pregnancies we look whether at least one of the babies has a chance of long term survival. Note that for this second point of view, we have to drop the baby-specific explanatory variables, e.g. birth weight and sex of the baby. The number of explanatory variables is then reduced from 25 to 20.

We apply Least Squares Support Vector Machine (LS-SVM) classifiers [8,9,10,11] with RBF kernel, 10-fold cross validation for the hyperparameter selection, and using the same inputs as LR and GEE, to check the quality of LR and GEE as linear models in relation to a nonlinear classifier. Logistic Regression and Generalized Estimating Equations are performed using the SAS System, release 8.02. Least Squares – Support Vector Machines are performed using the LS-SVM toolbox [12] in Matlab 6.5, release 13.

To compare the several classification models, we use the area under the Receiver Operating Characteristic (ROC) curve [14,15] as a measure of the predictive performance of the model. If a model has no discriminatory power, the area under the ROC curve be equal to 0.5. When there is a perfect separation of the two groups, the area be equal to 1. The closer the area is to 1, the better the performance of the model. A rough guide for classifying the accuracy of a model is the following: 0.90 – 1 = excellent, 0.80 – 0.90 = good, 0.70 – 0.80 = fair, 0.60 – 0.70 = poor, 0.50 – 0.60 = fail.

III. Results

A. *Survival Analysis*

We apply Cox's Proportional Hazards model for survival analysis. Since there is a lot of tied data in our dataset (a lot of babies die on the same day, e.g. 24% die on the first day) we have to correct for it. In order to obtain robust variance estimates that allow for dependence between babies of twins and triplets, we use the method of Wei, Lin, and Weissfeld [4].Cox's Proportional Hazards model obtains a generalized R^2 < 10% on the records of 1988 to 1996, which is insufficient to construct a reliable predictive model for

clinical practice.

B. Classification Analysis

1) GEE and LS-SVM applied on the 148 babies

When we apply the GEE model, stepwise selection reveals the following significant variables: the gestational age at the time of rupture of the membranes (in weeks), multiple pregnancy or not, amniotic fluid after PPROM or not, administration of steroids or not, administration of thyreotropin releasing hormone or not, birth weight and sex of the baby. No interactions among the variables were significant. With these seven explanatory variables we get an area under the ROC-curve of 0.88, a sensitivity of 80% and a specificity of 79%. Note that sensivity equals the percentage of babies that are predicted to survive the first month in the group of babies that effectively survived the first month and specificity equals the percentage of babies that are predicted not to survive the first month in the group of babies that died.

To get an idea of the predictive power of this classification model on a new set of records, we randomly divide the original set of 148 babies into two sets: a training set, which contains 70% of the data, and a test set, containing the remaining 30%. The obtained seven factors are used in the GEE model. We use the training set to find the parameter values of the model and the test set to see how good a model is able to classify a new set of observations. We repeat this procedure 20 times. We calculate the parameter values for 20 GEE models based on 20 randomly selected training sets, and apply those models on the 20 test sets. So, we can get an idea of the mean value and the variance of the area under the ROC-curve. Note that GEE is a linear model. In order to know if it is necessary to consider nonlinear terms in our model, we apply the nonlinear black box model LS-SVM with RBF kernel. We use the same seven factors as input variables and apply the LS-SVM model to the randomly selected training sets to train the model. Then we apply the trained models to the 20 corresponding test sets. In case of GEE the mean value of the area under the ROC-curve is 0.85 (stdev 0.04). For the 20 LS-SVM models the mean is 0.87, stdev 0.05. The difference between those two models is not significant [14,15], so there is no need to consider nonlinear terms in our GEE model.

2) LR and LS-SVM applied on the 119 pregnancies

In this section, we look at the data as records of 119 pregnancies. These are all independent of each other, so we can apply ordinary logistic regression (LR). As explained in section II, the outcome in case of a multiple pregnancy is defined as the outcome of the baby that survived the longest.

Stepwise selection applied to the LR model retains only four variables: the gestational age at the time of rupture of the membranes (in weeks), time interval between rupture and delivery (in days), administration of steroids or not, administration of thyreotropin releasing hormone or not. Also here no interactions between the variables were significant. The area under the ROC-curve is 0.86, sensitivity 0.95 and specificity 0.63.

Now, we apply LR and LS-SVM to the same 20 randomizations. In case of LR the mean value of the area under the ROC-curve is 0.85 (stdev 0.04). For the 20 LS-SVM models the mean is 0.87 (stdev 0.05), which is not significantly different from the LR model.

IV. Prospective Analysis

The prospective dataset contains medical records of 64 consecutive cases (52 pregnancies) of PPROM at or before 26 completed weeks of gestation. All women were admitted with a living fetus at the university hospital Gasthuisberg Leuven between 1997 and 2000. The aim is to test prospectively the predictive models that we have developed based on records in the period 1988 – 1996 (Section III B) on the new set of records.

Note that the survival rate is higher in the prospective set (period 1997 – 2000) than in the original one (period 1988 – 1996): while in the original dataset only 35% of the babies survived (31% in the period '88 – '91 and 41% in the period '92 – '95), 52% of the babies survived in the prospective dataset. So, a lot more babies have survived in the prospective dataset than in the original one.

For testing the predictive models prospectively, we follow the same order as in Section III B. The prospective dataset contains records of 64 babies (52 pregnancies). So, we first test the GEE model, developed on 148 babies (period 1988 – 1996), on the 64 cases of the prospective dataset. Then we apply LR, developed on the 119 pregnancies in the period 1988 – 1996, on the prospective set of 52 records (period 1997 – 2000). Here, we also apply Least Squares - Support Vector Machine classifiers using the same inputs as for GEE and LR, to check the quality of GEE and LR as linear models in relation to a nonlinear classifier (see Table 4a, and Table 4b).

TABLE VI a
PERFORMANCE MEASURES ON THE PROSPECTIVE SET OF 64 BABIES

	GEE	LS-SVM
Area under the ROC-curve	0.98	0.98
Correct classification	89%	89%
Sensitivity	79%	79%
Specificity	100%	100%

TABLE VI b
PERFORMANCE MEASURES ON THE PROSPECTIVE SET OF 52 PREGNANCIES

	LR	LS-SVM
Area under the ROC-curve	0.998	0.997
Correct classification	98%	96%
Sensitivity	97%	93%
Specificity	100%	100%

There is no significant difference in the performance

between the linear and the nonlinear model. Note that we obtain a specificity of 100% on the prospective set, i.e. all the babies that have effectively died after one month have been predicted to die.

V. Discussion

Now, we can compare the performance of the two points of view: 1) GEE based on the records of 148 babies and 2) Logistic Regression based on the records of 119 pregnancies.

Concerning the prediction of the survival chances of the first month, the Logistic Regression model is better than the GEE model: LR reaches 98% correct classification, sensitivity 97%, specificity 100%, while GEE only reaches 89% correct classification with a sensitivity of 79% and a specificity of 100%. So, the point of view for analyzing the data as explained in Section III B-2) has its benefits.

Each time we calculated the intrinsically nonlinear LS-SVM model, we did not obtain a significantly better performance. Therefore, there is no need to include nonlinear terms in the Logistic Regression or GEE model.

We can also compare our GEE model for survival of the first month with the GEE model in paper [1]. That model reaches an area under the ROC-curve of 0.96, which is not significantly different from our GEE model for predicting the survival chances of the first month. But, the percentage correctly classified data is higher using our model: we obtain 89% correct classification, while in the paper [1] the GEE model reaches only 81% (sensitivity: 100% and specificity 66%).

Note that the performance in the prospective set is higher than the performance in the original set where the different models were developed on, e.g., Generalized Estimating Equations reach an area under the ROC-curve of 0.88 on the original set and 0.98 on the prospective set of observations. As stated in Section IV there is a significant difference in the survival rate between the original and the prospective dataset. While in the original dataset (period 1988 – 1996) only 35% of the babies survived, 52% of the babies survived in the period 1997 – 2000. It might be that the predictive power of >95% area under the ROC-curve is an overestimation when in the future the survival rate is not as high as the survival rate in our prospective set.

VI. Conclusions

An extended data analysis has been performed in order to construct a predictive model of survival for babies of which the mother has a preterm prelabor rupture of the membranes before 26 weeks of gestation. Using a dataset with records of 148 babies and 25 explanatory variables we developed predictive models using Logistic Regression, Least Squares – Support Vector Machines and Generalized Estimating Equations. The performance of the obtained predictive models has been examined on a prospective set of 64 babies. We reached an excellent performance of >95% as area under the ROC-curve on the prospective set.

References

[1] B. Spitz, C. Vossen, R. Devlieger, F. A. Van Assche, "Rupture of the membranes before 26 weeks of gestation," *Journal of Perinatal Medicine*, vol 27, pp 451-457, 1999.

[2] P.D. Allison. Survival Analysis using the SAS System: A Practical Guide. SAS Institute Inc., Cary, NC, USA, 1995.

[3] J.D. Kalbfleisch, R.L. Prentice. The Statistical Analysis of Failure Time Data. John Wiley & Sons Inc., New York, 1980.

[4] L.J. Wei, D.Y. Lin, L. Weissfeld, "Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions," Journal of the American Statistical Association, vol.84, pp. 1065-1073, 1989.

[5] A. Agresti, *Introduction to Categorical Data Analysis*. Wiley Series in Probability and Statistics, 1996.

[6] D. W. Hosmer Jr., S. Lemeshow, *Applied Logistic Regression*. John Wiley & Sons, Inc., 1989.

[7] P. J. Diggle, K. Y. Liang, S. L. Zeger, *Analysis of Longitudinal Data.* Oxford Univeristy Press, 1994.

[8] J. A. K. Suykens, J. Vandewalle, "Least squares support vector machines classifiers," *Neural Processing Letters*, vol. 9(3), pp. 293-300, 1999.

[9] T. Van Gestel, J. A. K. Suykens, B. Baessens, S. Viaene, J. Vanthienen, G. Dedene, B. De Moor, J. Vandewalle, " *Machine Learning,* vol. 54(1), pp. 5-32, 2004.

[10] V. Vapnik, *The nature of statistical learning theory*. Springer-Verlag, 1995.

[11] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle. *Least Squares Support Vector Machines*. World Scientific Publishing Co., Pte, Ltd. (Singapore), 2002.

[12] K. Pelckmans, J. A. K. Suykens, T. Van Gestel, J. De Brabanter, L. Lukas, B. Hamers, B. De Moor, J. Vamdewalle, "LS-SVMlab: a Matlab/C toolbox for least squares support vector machines," unpublished
(downloadable from http://www.esat.kuleuven.ac.be/sista/lssvmlab)

[13] A. J. Hanley, B. J. McNeil, "The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve," *Radiology*, vol. 143(1), pp. 29-36, 1982.

[14] E. R. De Long, D. M. De Long, D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," *Biometrics*, vol. 44, pp. 837-845, 1988.

[15] F. E. Harrell Jr., *Regression Modeling Strategies*. Springer Series in Statistics, Springer – Verlag, New York, 2002.G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.