Journal of Biostatistics and Epidemiology

*J Biostat Epidemiol. 2018; 4(2): 47-51*

**Original Article**

# Prediction of time to reflux using accelerated failure time model of Weibull distribution in children with antenatal hydronephrosis

Maryam Nazemipour[1], Abdol-Mohammad Kajbafzadeh[2], Kazem Mohammad[3], Mahmood Mahmoudi[3]*

[1] Department of Epidemiology and Biostatistics, School of Public Health, International Campus, Tehran University of Medical Sciences, Tehran, Iran
[2] Pediatric Urology Research Center, Department of Pediatric Urology, Children's Hospital Medical Center, Tehran University of Medical Sciences, Tehran, Iran
[3] Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Received 27.06.2017
Revised 19.09.2017
Accepted 02.10.2017

**Key words**:
Antenatal hydronephrosis;
Reflux;
Risk factor;
Survival model;
Weibull distribution

ABSTRACT

**Background & Aim:** Prediction of time to reflux can aid healthcare providers and preparation programs. We constructed a risk prediction instrument for occurrence reflux in children with antenatal hydronephrosis.
**Methods & Materials:** Demographic and clinical information was collected retrospectively in children with the antenatal hydronephrosis and mostly with reflux, followed at least 5 years.
**Results:** Accelerated failure time model of data from 333 children was developed to assess the risk of time to reflux. Likelihood ratio tests of statistical significant were used to identify best fitting predictive function. Variables "gender", "Sr", and "severity of ANH (in severe level)" were highly significant ($p<0.05$) in multivariate model, adjusting for some traditional risk factors.
**Conclusion:** This proposed risk probability model allows prediction of time to reflux for children with antenatal hydronephrosis to better inform parents from possible time of occurrence reflux and treatment strategies.

## Introduction

Antenatal hydronephrosis, diagnosed by ultrasonography, is known to be associated with vesicouretral reflux (VUR) (1). The incident of VUR in children antenatally diagnosed hydronephrosis, ranges from 7-38% (1-3). VUR in infants and children can lead to renal damage, causing infection and end stage renal disease (4, 5). Clinicians attemp to predict time to occurrence vesicouretral reflux in children borned with antenatal hydronephrosis. This prediction is important due to many reasons. First of all to inform the patient about the outcome of her/his disease. Second, is used as a guideline for th e clinician for an appropriate terapie. Third, to test the differential therapeutic benefits and finally is a guide for researcher in diagnosing, clinical trials, or effect of a single factor on prognosis in an observational study (6).

Survival analysis technique has been applied in medical research widly to explore the duration of time from a certain time until occurrence of the event or events (7) and it is common to have incomplete event time (8). There are two types of regression models in survival analysis, Cox proportional hazard (PH) model (9) and accelerated failure time (AFT) model (10). Proportional hazard model directly refers to the effect of some covariates on the hazard function while the accelerated failure time model describes the failure time affected by the covariates and explains how they accelerate or decelerate the life course of the disease. Cox models have been widely used in medical researches although AFT models are applied just in prediction of time to a certain event accelerated by different factors (11). Theme of this paper is, application of an AFT parametric

* Corresponding Author: Mahmood Mahmoudi, Postal Address: Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran.
Email: mahmoodim@tums.ac.ir

model to predict time to reflux for children with antenatal hydronephrosis.

## Methods

A total of 333 children with ANH, admitted to the Pediatric Urology Research Center of Children's Hospital Medical Center affiliated to Tehran University of Medical Sciences between 2003 and 2005 and nearly most of whom had VUR, were enrolled in this study. They were followed up for at least 5 years for measuring and evaluating time of occurrence their reflux. Information on their demographic and clinical as well as the time were extracted from their medical records and by phoning their family. Data analysis was done using R software version 2.14.1 and p<0.05 was considered statistically significant. This information can prepare a guide for helping a pregnant woman on the postnatal effect of ANH.

We applied an accelerated failure time survival model to analysis the number of days from diagnosis of ANH until the reflux event. Time to reflux was in agreement with the Weibull distribution. The stepwise AFT regression models show five factors are in association with the outcome. The variables included in the stepwise models analysis were "Sr", "gender", "consanguinity marriage", "severity of ANH" and "Direction of ANH".

Ethics Committee Approval: All the data were collected after Tehran University of Medical Sciences Institutional Review Board (TUMS IRB) approval

Weibull distribution model: The probability distribution of the Weibull distribution is defined by $f(t) = \lambda p t^{p-1} \exp(-\lambda t^{p})$

Where $p > 0$ and $\lambda > 0$. In this formula $p$ is a shape parameter. The survival and hazard function of Weibull are $S(t) = \exp(-\lambda t^{p})$ and $(t) = \lambda p t^{p-1}$ respectively, so that when $p > 1$, hazard increases, when $p < 1$, hazard decreases and if $p = 1$, the hazard will be constant over time (exponential model). For the PH model we reparametrize $\lambda = \exp(\beta_0 + \beta X)$ and for the AFT form is assumed $\frac{1}{\lambda^{\frac{1}{p}}} = \exp(\alpha_0 + \alpha X)$. Therefore, survival function can be written as follows:

$$S(t) = exp\{-[t * exp\{-(\alpha_0 + \alpha X)\}]^{p}\}$$

and finally the risk of an event (here reflux is the event) by time $t$ is computed with $1 - S(t)$.

In another analysis using the Cox proportional hazard model we found that the parameter estimates (hazard ratio) were almost similar between the two methods (analysis available on request). The method of AFT model was used to be consistent with the work of Anderson et al (1991) (12) to estimate the predictive impact of variables from a single observation. Values from the first time of diagnosis ANH were used and the values of predictor variables were not changed during the follow up study.

To assess relative goodness of fit and the discriminatory ability of the models, log-likelihood statistic also C-Statistic (13, 14) was applied. All the HRs, calculated from the variable coefficients in the AFT model, have been presented as the effect of individual predictor variables.

## Results

Clinical and demographical characteristics of 333 children identified with ANH are listed in Table 1. Mean diagnosis of ANH was 29.86

**Table 1.** Characteristics of 333 children with ANH

| Variables | Status | No. | Proportion (%) |
|---|---|---|---|
| Gender | Boy | 278 | 83.5 |
| Consanguinity marriage | YES | 80 | 24 |
| Kidney disease background in parent | YES | 19 | 5.7 |
| UI | YES | 159 | 47.7 |
| Constipation | YES | 145 | 43.5 |
| UC | YES | 130 | 39 |
| UA | YES | 130 | 39 |
| HUN | YES | 255 | 76.6 |
| HN | YES | 328 | 98.5 |
| BWT | Abnormal | 90 | 27 |
| PVR | Abnormal | 5 | 1.5 |
| APD_Right | Abnormal | 67 | 20.1 |
| APD_Left | Abnormal | 82 | 24.6 |
| CA19_9baby | Abnormal | 227 | 68.2 |
| Sr | Abnormal | 22 | 6.6 |
| Cr | Abnormal | 4 | 1.2 |
| Urethra | Abnormal | 254 | 76.3 |
| Bladder | Abnormal | 289 | 86.8 |
| Direction of ANH | Bilateral | 180 | 54.1 |
| Baseline (Right) | Unilateral (Left) | 90 | 27 |
| Grade of ANH | Moderate | 43 | 12.9 |
| Baseline (Mild) | Severe | 68 | 20.4 |
| Week of diagnosis ANH* | | 29.86 | 6.006 |

*Mean (SD), UI or UTI: Urinary tract infection, UC: Urine culture, UA: Urine analysis, HUN: Hydroureteronephrosis, HN: Hydronephrosis, BWT: Bladder wall thickness, PVR: Post-void residual, APD: Anterior-posterior diameter, CA19_9baby: Carbohydrate antigen19-9, Sr: Sacral ratio, Cr: Urine creatinine, ANH: Antenatal hydronephrosis

**Table 2.** Result of accelerated failure time for Weibull regression

| Variables | Status | Weibull | | |
|---|---|---|---|---|
| | | Coef | Std.Err | HR (CI.95%) |
| Week of diagnosis ANH | | 0.016 | 0.011 | 0.985 (0.966 , 1.006) |
| Gender* | Girl | 0 | - | 1 |
| | Boy | -0.446 | 0.190 | 1.501 (1.070 , 2.107) |
| UI | No | 0 | - | 1 |
| | Yes | -0.123 | 0.136 | 1.119 (0.878 , 1.426) |
| Constipation | No | 0 | - | 1 |
| | Yes | 0.120 | 0.145 | 0.897 (0.693 , 1.161) |
| UC | No | 0 | - | 1 |
| | Yes | 0.067 | 0.150 | 0.941 (0.719 , 1.230) |
| UA | No | 0 | - | 1 |
| | Yes | -0.214 | 0.148 | 1.215 (0.932 , 1.584) |
| HUN | No | 0 | - | 1 |
| | Yes | -0.042 | 0.164 | 1.040 (0.775 , 1.392) |
| BWT | Normal | 0 | - | 1 |
| | Abnormal | -0.068 | 0.153 | 1.064 (0.810 , 1.398) |
| APD_Right | Normal | 0 | - | 1 |
| | Abnormal | 0.216 | 0.244 | 0.821 (0.531 , 1.270) |
| APD_Left | Normal | 0 | - | 1 |
| | Abnormal | 0.245 | 0.222 | 0.800 (0.539 , 1.188) |
| Direction of ANH | Unilateral (Right) | 0 | - | 1 |
| | Bilateral | -0.178 | 0.189 | 1.176 (0.840 , 1.647) |
| | Unilateral (Left) | 0.153 | 0.233 | 0.870 (0.575 , 1.318) |
| Severity of ANH* | Mild | 0 | - | 1 |
| | moderate | -0.414 | 0.254 | 1.457 (0.926 , 2.294) |
| | Severe* | -0.683 | 0.235 | 1.862 (1.223 , 2.837) |
| CA19-9-baby | Normal | 0 | - | 1 |
| | Abnormal | -0.179 | 0.153 | 1.177 (0.896 , 1.546) |
| Sr* | Normal | 0 | - | 1 |
| | Abnormal | -1.132 | 0.253 | 2.803 (1.767 , 4.447) |
| Urethra | Normal | 0 | - | 1 |
| | Abnormal | -0.106 | 0.167 | 1.101 (0.818 , 1.484) |
| Bladder | Normal | 0 | - | 1 |
| | Abnormal | -0.105 | 0.204 | 1.101 (0.764 , 1.585) |
| Kidney disease background in parent | No | 0 | - | 1 |
| | Yes | -0.060 | 0.287 | 1.056 (0.633 , 1.763) |
| Consanguinity marriage | No | 0 | - | 1 |
| | Yes | -0.256 | 0.159 | 1.262 (0.951 , 1.676) |
| Cons | | 8.280 | 0.586 | |
| Ancillary parameters | | P=0.911 | | |

*Significant variable

weeks with the standard deviation of 6.006. Of these patients 83.5% were boy and 24% had consanguinity marriage. 20.4% of all had ANH in severe level also 54.1% were bilateral. The highest abnormality was for bladder with 86.8% and the lowest was in PVR and Cr. Table 2 presents result of weibull accelerated failure time model.

The highest HR is seen for the “Sr” with 2.803 and the lowest for the “APD_Left” with 0.8 so one of them is very risky and the other one is protective. Table 3 presents result of the stepwise accelerated failure time models. Variety of AFT models for the outcome were used. Variables “Sr”, “gender”, “consanguinity marriage”, “severity of ANH” and “Direction of ANH” come to stepwise models step by step from model A through E in Table 3 along with the HR for the all five predictor variables of reflux. The effect estimates for all these variables vary little across the models. The magnitude of the log-likelihood statistic similar to the C-statistics was the greatest one (the smallest one without considering the negative) in the full multivariate analysis (model E), although adding variable to model always increase the log-likelihood. The values of C-statistics for the five reflux models vary little from 3.727 to 4.885 and its p-value from 0.444 to 0.299. The p-value of 0.80 is usually reguarded as boundary between acceptable and excellent discrimination (14). The discriminatory power of fit as measured by the log-likelihood and C-statistics are virtually equivalent for these models.

**Table 3.** Result from AFT models for the time of reflux

| Predictor | Status | HR (CI .95) | | | | |
|---|---|---|---|---|---|---|
| | | A | B | C | D | E |
| Sr | Normal | 1 | 1 | 1 | 1 | 1 |
| | Abnormal | 2.526 (1.624,3.928) | 2.560 (1.646,3.983) | 2.639 (1.694, 4.112) | 2.651 (1.701,4.131) | 2.676 (1.716,4.171) |
| gender | Girl | - | 1 | 1 | 1 | 1 |
| | Boy | - | 1.542 (1.115,2.135) | 1.563 (1.129,2.164) | 1.546 (1.117,2.141) | 1.486 (1.070,2.063) |
| Consanguinity marriage | No | - | - | 1 | 1 | 1 |
| | Yes | - | - | 1.398 (1.067,1.832) | 1.381 (1.054,1.810) | 1.350 (1.029,1.770) |
| Severity of ANH (Severe level) | No | - | - | - | 1 | 1 |
| | Yes | - | - | - | 1.380 (1.038,1.833) | 1.347 (1.013,1.792) |
| Direction of ANH (Bilateral)* | No | - | - | - | - | 1 |
| | Yes | - | - | - | - | 1.249 (0.983,1.587) |
| log likelihood | | -556.813 | -553.045 | -550.248 | -547.934 | -546.268 |
| C-statistic | | 3.727 | 4.784 | 4.996 | 4.885 | 4.885 |
| P-value | | 0.444 | 0.310 | 0.288 | 0.299 | 0.299 |

*This variable was in the borderline ($0.05 < p < 0.1$) and important. So it was included in the model

Finally it should be mention that the data was fitted with an AFT survival model based on the weibull distribution. This model yeilds a linear function of the predictor variables X and coefficients $\alpha$ with an intercept $\alpha_0$ as follows

$$L(X) = \alpha_0 + \alpha_1 X_1 + \cdots + \alpha_5 X_5$$

(the model for the time to reflux event has five predictor variable)

Along with the shape parameter $p$, for a person with the particular predictor variables, the predicted survival at time $t$ is given by

$$S(t) = \exp(-(t * \exp(-L(X)))^p)$$

So the risk of the event by time $t$ is obtained as follows

$$1 - S(t)$$

## Discussion

From five predictor variables in the Table 3, all of them were negetively related to the long-term incidence of reflux. All the five predictor variables were statistically significant and the difference between the value of the log-likelihood statistics among the models were modest. The variables, Sr, gender of child, consanguinity marriage, severity of ANH and Direction of ANH were hazardious. This study has some limitation in which has been reported only the experience of one hospital. Obtaining the result in other geographic areas, for managing the risk factors, require considering race, ethnicity of population as well as calender year of the experience.

## Conclusion

Being interested to analyzing children detected with ANH and avaluating factors affecting on the time to occurrence VUR, this study analysed children diagnosed with ANH most of whom have VUR. It showed gender of child, severity of ANH and Sr in addition to some other clinical factors were related to the risk of occurence reflux. These information can prepare a guideline for helping the pregnant women on the postnatal effect of ANH.

## Acknowledgments

The authors are grateful to the staff of the Pediatric Urology Research Center of Children's Hospital Medical Center, affiliated to Tehran University of Medical Sciences for providing data. Research leading to this paper was supported by the International Campus of Tehran University of Medical Sciences.

## Conflicts of interests

The authors declare that there is no conflict of interest regarding the publication of this article.

## References

1. Vesicoureteric reflux: all in the genes? Report of a meeting of physicians at the Hospital for Sick Children, Great Ormond Street, London. The Lancet. 1996;348(9029):725-8.
2. Kramer S. Vesico-Ureteral reflux. In Clinical

Pediatric Urology. 4th ed. Belman AK, Kramer SA, editors. London: Martin Dunitz; 2002. 749 p.
3. Williams G, Fletcher JT, Alexander SI, Craig JC. Vesicoureteral reflux. J Am Soc Nephrol. 2008;19(5):847-62.
4. Mathews R, Carpenter M, Chesney R, Hoberman A, Keren R, Mattoo T, et al. Controversies in the management of vesicoureteral reflux: the rationale for the RIVUR study. J Pediatr Surg. 2009;5(5):336-41.
5. Sharbaf FG, Fallahzadeh MH, Modarresi AR, Esmaeili M. Primary vesicoureteral reflux in Iranian children. Indian Pediatr. 2007;44(2):128-30.
6. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med. 2007;26(11):2389-430.
7. Hosmer DW, Lemeshow S, May S. Applied survival analysis: regression modeling of time-to-event data. 2nd ed. Hoboken, N.J: Wiley-Interscience; 2008;13:392.
8. Southern DA, Faris PD, Brant R, Galbraith PD, Norris CM, Knudtson ML, et al. Kaplan-Meier methods yielded misleading results in competing risk scenarios. J Clin Epidemiol. 2006;59(10): 1110-4.
9. Cox DR. Regression models and life tables. J Royal Stat Soc. 1972;34(B):187-202.
10. Collett D. Modelling survival data in medical research. 2nd ed. Boca Raton, Fla.: Chapman & Hall/CRC; 2003:391.
11. Lawless JF. Statistical models and methods for lifetime data. 2nd ed. Hoboken, N.J.: Wiley-Interscience; 2003;2:630.
12. Anderson KM, Wilson PW, Odell PM, Kannel WB. An updated coronary risk profile. A statement for health professionals. Circulation. 1991;83(1):356-62.
13. Harrell JFE. Regression modeling strategies : with applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag; 2001;12:568.
14. Hosmer DW, Lemeshow S. Applied logistic regression. 2nd ed. New York: Wiley; 2000;12:375.