Published in final edited form as:

Contemp Clin Trials. 2009 March ; 30(2): 150–154. doi:10.1016/j.cct.2008.11.001.

Analysis of Time-to-Event and Duration Outcomes in Neonatal Clinical Trials with Twin Births

Michele L. Shaffer^{1,2} and Sasiprapa Hiriote³

¹Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, USA

²Department of Pediatrics, Penn State College of Medicine, Hershey, PA, USA

³Department of Statistics, The Pennsylvania State University, State College, PA, USA

Abstract

When conducting neonatal trials in pre-term and/or low-birth-weight infants, twins may represent 10–20% of the study sample. Frailty models and proportional hazards regression with a robust sandwich variance estimate are common approaches for handling correlated time-to-event data or duration outcomes that are subject to censoring. However, the operating characteristics of these methods for mixes of correlated and independent time-to-event data are not well established. Simulation studies were conducted to compare frailty models and proportional hazards regression models with a robust sandwich variance estimate to standard proportional hazards regression models to estimate the treatment effect in two-armed clinical trials. While overall frailty models showed the best performance, caution must be exercised as the interpretation of the parameters differs from the marginal models. Data from the National Institute of Child Health & Human Development sponsored PROPHET trial are used for illustration.

Keywords

correlated data; frailty models; robust variance

Introduction

Neonatal studies present a unique correlated data problem. Data from singletons and twins whose siblings are not included in the study (unmatched twins) can in most cases be considered independent. Data from complete twin births have a hierarchical structure (infants' outcomes nested within a birth) and in most cases would be considered dependent. In the absence of twin births, classical statistical techniques are valid and appropriate. In the absence of singletons, hierarchical methods that account or adjust for the hierarchical structure could be applied. Failing to account for the hierarchical structure within complete twin births may impact the estimated target sample size and the precision of treatment effect estimates and/or decisions regarding treatment efficacy. Therefore, these effects may need to be quantified and accounted for in neonatal studies. If the proportion of infants from complete twin pairs is modest, e.g., less than 5% of the study sample, there may be minimal

© 2008 Elsevier Inc. All rights reserved.

Reprint requests should be addressed to: Michele L. Shaffer, 600 Centerview Drive, Suite 2200, Hershey, PA 17033, shaffer.michele@psu.edu, Phone: 717-531-7178, FAX: 717-531-5779.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

impact if correlation is ignored. Additionally, methods that account for correlation are computationally more difficult and may fail if too little data are available to adequately model or appropriately adjust for correlation. Thus, under some circumstances, it is possible classical statistical techniques may be sufficient or preferred.

Twenty percent or more of infants are products of multiple gestations, primarily twins, in the very low birth weight group (501–1500 g), due in part to the increasing number of pregnancies resulting from assisted reproductive technology [1, 2]. Whether twin outcomes are more similar than unrelated individuals due to genetic similarity, or more different than unrelated individuals because of increased illness severity of one twin, usually the “B” or second twin, is unknown.

For time-to-event data and duration outcomes subject to censoring, frailty models are one approach to modeling correlated data. For a gamma frailty model, the hierarchical structure within births is modeled by adding a random effect with gamma distribution to a Cox proportional hazards regression model [3, 4].

For a model that includes only treatment, we can write the Cox proportional hazards model as

$$\lambda(t|Z) = \lambda_0(t) \exp(\beta Z) \quad (1)$$

where $\lambda(t|Z)$ is the hazard rate at time t for individuals assigned to treatment Z (an indicator variable), $\lambda_0(t)$ is the arbitrary baseline hazard rate, and β measures the treatment effect. Thus, $\exp(\beta)$ is the marginal hazard ratio.

For the gamma frailty model, equation (1) becomes

$$\lambda_{ki}(t) = \xi_k \lambda_0(t) \exp(\beta Z_{ki}) \quad (2)$$

where $\lambda_{ki}(t)$ is the hazard rate for infant i in birth k , ξ_k is the frailty for birth k , and $\exp(\beta)$ is the hazard ratio for a given birth (i.e., a given value of the frailty). The variance of the gamma distributed frailties can be used to obtain a measure of correlation. If we let θ denote the variance of the gamma distribution, $\theta/(\theta + 2)$ is the nonparametric correlation coefficient Kendall's τ [5].

Instead of a frailty model, a sandwich estimator for variance can be used to adjust for the potential within-birth correlation, which yields a robust estimate of variance for use in confidence interval construction or hypothesis testing [6]. The marginal hazard ratio remains unchanged from the Cox proportional hazards regression model.

While these existing methods may be used to analyze correlated time-to-event data and duration outcomes subject to censoring, the optimal methods for studies that may have as little as 10% of the individuals related are unknown. Also, the interpretation of the regression coefficients for the frailty model differs from the marginal model approaches. Specifically, for the frailty model the regression coefficient measures the effect of the treatment for a given birth. For the marginal models, the regression coefficient measures the effect of treatment averaged over all infants at risk. When there is no treatment effect ($\beta = 0$) or when there is no correlation within births ($\xi_k \equiv 1$), the marginal hazard ratio is equal to the conditional hazard ratio. Otherwise, the ratio of marginal hazards for the frailty model varies with time [7]. Thus, one potential check of whether the marginal or conditional approach is more appropriate would be to examine the proportional hazards assumption for

treatment when fitting the marginal model. Additionally, one can conduct a likelihood ratio test for significant frailty when using the conditional approach [5, 8].

Henderson and Oman reviewed the effects of ignoring frailty in analysis by fitting misspecified Cox proportional hazards models with a robust estimate of variance and showed regression coefficients are biased toward 0 [9]. The extent of the bias depends upon the variance of the frailty and the frailty distribution selected. When censoring is present, the bias is reduced. Glidden and Vittinghoff studied the use of marginal and conditional Cox regression models for analyzing clustered survival data from multicenter clinical trials and found that the gamma frailty model provided estimates with lower mean squared error than fixed effects or stratified methods [7]. In addition, regression coefficient estimators performed adequately even when the frailty distribution was misspecified, indicating the gamma frailty approach may be reasonable even when the random effects interpretation is not of interest.

The purpose of this paper was to determine if, and when, frailty models or a Cox proportional hazards model with a robust estimate of variance were operationally superior to the classic Cox proportional hazards regression model. Specifically, we look at two-armed, neonatal clinical trials with time-to-event data or duration outcomes subject to censoring.

The PROPHET Trial

Our motivating example is the PROPHET trial. The PROPHET trial was a multicenter, randomized clinical trial comparing placebo (n=180) to low-dose hydrocortisone therapy (n=180) in the first two weeks of life in extremely low birth weight babies (500–999 grams) to prevent chronic lung disease sponsored by the National Institute of Child Health & Human Development [10]. The unit of randomization for the study was birth. The targeted enrollment for the study was 712 births, but the enrollment was stopped at 360 babies because of an increase in spontaneous gastrointestinal perforation in the hydrocortisone-treated group. Pregnancies of three or more fetuses were excluded from the study. Of the infants enrolled in the study, 17% were twins whose siblings were also enrolled. While the primary benefit outcome for the study was survival without supplemental oxygen at 36 weeks postmenstrual age, secondary outcomes included length of hospital stay (time to discharge), duration of ventilator support, duration of oxygen, and duration of greater than 40% oxygen.

Simulation Studies

Simulation studies were conducted using R Versions 2.5.1 and 2.7.2 (R Foundation for Statistical Computing, Vienna, Austria) to operationally compare the Cox proportional hazards regression model to frailty models and a Cox proportional hazards model with robust variance estimate for time-to-event data and duration outcomes subject to censoring in two-armed clinical trials. A variety of scenarios were considered, varying the randomization scheme for complete twin pairs, sample size, proportion of infants from complete twin pairs, correlation within twin births, hazard ratio, and censoring rate. Values of parameters used in the simulation studies are shown in Table 1. The largest proportion of twins, 0.6, is unrealistic for low birth weight or preterm birth studies, but this value is included to generalize the results to other areas of application, for example studies that include singles and couples or ophthalmic studies.

Ten thousand data sets were generated for each simulation case, and 95% confidence intervals were computed for the log hazard ratio. Operating characteristics collected included the mean estimate of the log hazard ratio and the mean and median width and coverage rate of the confidence interval.

Data were generated using two different methods. The first method used a gamma frailty model, as described in (2), where the frailties are shared between twins. The baseline hazard rate $\lambda_0(t)$ followed an exponential distribution with rate parameter one-half. We considered gamma-distributed frailties with variances 0 and 2/3, corresponding to a Kendall's τ of 0 and 0.25, respectively. The second method used Gumbel's bivariate exponential model to generate paired times with marginal hazard rates $\exp(\beta Z_{ki})$ for infant i in birth k [11]. We considered correlation coefficients of 0 and 0.25. The Gumbel model has been used previously to test the operating characteristics of the Cox proportional hazards model with a robust variance estimate [12]. For both methods of generation, a binomial random number generator was used to independently assign censor status to the times, including independent assignment for the paired times.

To mimic the balance that would occur by using permuted blocks in practice, fifty percent of the randomizations were to the treatment arm and fifty percent to the placebo arm. Randomized treatment assignments for complete twin pairs were dependent upon the twin randomization scheme. When twins were randomized independently, the process was the same as for singletons and unmatched twins. When complete twin pairs always were assigned to the same treatment, or opposite treatments, a treatment assignment was generated for a single twin in the pair, which automatically determined the second twin's assignment.

Results

Results when generating the data under a frailty model are presented in Tables 2–3. When the correlation within twin births was 0, disregarding the correlation by using the Cox proportional hazards regression model performed best as expected; however, using a robust estimator of variance or a frailty model still resulted in good coverage values. When there was no treatment effect, but a positive correlation, using a robust estimator of variance showed the best coverage values, but using a frailty model still produced acceptable results. When there was a treatment effect present together with a positive correlation, the frailty model had much higher coverage values, but the coverage fell as the correlation increased and the treatment effect increased, while the coverage increased as the proportion of twins increased. The coverage for the frailty model fell slightly as the censoring rate increased, but increased modestly when correlation was disregarded or when a robust estimator of variance was used (results not shown). Looking at randomization scheme, only in the most extreme cases did there appear to be a small advantage to randomizing twins independently or to opposite treatment arms rather than by birth, which is the most commonly used method (results not shown). However, in these extreme cases no method performed very well.

Results when the data were generated under Gumbel's bivariate exponential model are presented in Tables 4–5. When the correlation within twin births was 0, again disregarding the correlation by using the Cox proportional hazards regression model performed best as expected, but using a robust estimator of variance or a frailty model still resulted in good coverage values. When there was no treatment effect, but a positive correlation, all methods performed similarly, but disregarding the correlation resulted in the best coverage values. When there was a treatment effect present together with a positive correlation, again disregarding correlation produced the best coverage values, but no method performed poorly. In these cases, for the frailty model the coverage increased as the proportion of twins increased. Censoring had little impact on the coverage, and there was no apparent advantage to randomizing twins independently or to opposite treatment arms rather than by birth (results not shown).

Revisiting the PROPHET Trial

Analyses of length of hospital stay in days and days of greater than 40% oxygen are shown in Tables 6 and 7, respectively. Infants' values were censored if they died during the period of data collection for the outcome. Note for these outcomes that a smaller value is a better value. The hazard ratios shown use placebo as the reference group. Thus, a hazard ratio of greater than 1 indicates that hydrocortisone yields a better outcome, while a hazard ratio of less than 1 indicates that placebo yields a better outcome.

For length of stay, there was no evidence of non-proportionality observed from plotting Kaplan-Meier curves, testing for an interaction with logged length of stay ($p=0.35$), or plotting the scaled Schoenfeld residuals. Thus, we examine the marginal results. From the simulation findings, the estimates disregarding the correlation should be reported. However, we note in this case that all analyses lead to the same conclusion of no difference in treatment groups.

For days of greater than 40% oxygen, crossing Kaplan-Meier curves showed evidence of non-proportionality, although the test of interaction with logged days was not significant (0.63). The test of significant frailty ($p=0.05$) indicated the presence of positive correlation between twins. Thus, the findings based on the frailty model should be reported, although again, as for length of stay, all analyses lead to the same conclusion of no difference in treatment groups.

Discussion

While frailty models and proportional hazards regression with a robust sandwich variance estimate are common approaches for handling correlated time-to-event data and duration outcomes that are subject to censoring, the operating characteristics of these methods for mixes of correlated and independent time-to-event data, such as the data encountered in neonatal trials that include multiple gestations, are not well understood. For moderately sized treatment effects and correlations within births, gamma frailty models perform adequately, even when no correlation within births is present or the data are generated from a model where marginal proportional hazards holds. If the underlying model is truly a frailty model, there is a large penalty imposed for incorrectly fitting a marginal model when there is correlation present together with a positive treatment effect. While disregarding the correlation results in the best coverage values when a marginal model holds, all methods studied showed good performance. None of the methods studied performed well when large treatment effects and positive correlation were present, and the data were generated under a frailty model. While these large effects are rare in the authors' experience, the question of how to best analyze such outcomes is an area of future research. One possible approach is to use within-birth resampling introduced by Hoffman, Sen, and Weinberg for binary outcomes [13].

Acknowledgments

Partial support for this research was provided under Grant No. R03 HD049373 from the National Institute of Child Health & Human Development. The authors would like to thank Kristi L. Watterberg for permission to use the PROPHET trial data and the National Institute of Child Health & Human Development Neonatal Research Network for permission to use the IVIG trial database. The authors also thank two reviewers for their valuable comments and suggestions.

References

1. Schieve LA, Meikle SF, Ferre C, et al. Low and very low birth weight in infants conceived with use of assisted reproductive technology. *N Engl J Med*. 2002; 346:731–737. [PubMed: 11882728]

2. Martin JA, Hamilton BE, Sutton PD, et al. Births: final data for 2002. National Vital Statistics Reports. 2003; 52:1–113. [PubMed: 14717305]
3. Cox DR. The analysis of multivariate binary data. Appl Stat. 1972; 21:113–120.
4. Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika. 1978; 65:141–152.
5. Therneau, TM.; Grambsch, PM. Modeling Survival Data: Extending the Cox Model. New York: Springer; 2000.
6. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. J Am Stat Assoc. 1989; 84:1074–1078.
7. Glidden DV, Vittinghoff E. Modelling clustered survival data from multicentre clinical trials. Stat Med. 2004; 23:369–388. [PubMed: 14748034]
8. Nielsen GG, Gill RD, Andersen PK, Sorensen TI. A counting process approach to maximum likelihood estimation in frailty models. Scandinavian J Stat. 1992; 19:25–43.
9. Henderson R, Oman P. Effect of frailty on marginal regression estimates in survival analysis. J R Statist Soc B. 1999; 61:367–379.
10. Watterberg KL, Gerdes JS, Cole CH, et al. Prophylaxis of early adrenal insufficiency to prevent bronchopulmonary dysplasia: a multicenter trial. Pediatrics. 2004; 114:1649–1657. [PubMed: 15574629]
11. Gumbel EJ. Bivariate Exponential Distributions. J Am Stat Assoc. 1960; 55:698–707.
12. Lin DY. Cox regression analysis of multivariate failure time data: the marginal approach. 1994; 13:2233–2247.
13. Hoffman EB, Sen PK, Weinberg CR. Within-cluster resampling. Biometrika. 2001; 88:1121–1134.

Table 1

Parameter values for simulation studies

Simulation parameter	Values
Randomization of twin pairs	Same arm, independently, opposite arms
Total sample size	250, 500
Proportion of infants from twin pairs	0.1, 0.2, 0.6
Log hazard ratio for Cox model	$\ln(1)=0$, $\ln(1.5)=0.405$, $\ln(2)=0.693$
Correlation within twin births	
Variance of gamma distribution (frailty data generation)	0, 2/3
Correlation coefficient (Gumbel data generation)	0, 0.25
Censoring rate	0.05, 0.10

Table 2

Findings under gamma frailty model data generation when twins are randomized to the same arm; log hazard ratio=0; censoring rate=0.10

Total Sample Size	Proportion Of twins	Kenall's τ	Method	Mean β	Median Width	Coverage
250	0.1	0	Cox	1.95E-3	0.527	0.949
			Frailty	2.03E-3	0.528	0.948
			Robust	1.95E-3	0.524	0.949
		0.25	Cox	1.80E-4	0.527	0.946
			Frailty	7.75E-4	0.535	0.940
			Robust	1.80E-4	0.533	0.948
	0.2	0	Cox	-1.14E-3	0.527	0.948
			Frailty	-1.40E-3	0.528	0.946
			Robust	-1.14E-3	0.524	0.946
		0.25	Cox	3.13E-3	0.527	0.940
			Frailty	2.98E-3	0.672	0.942
			Robust	3.13E03	0.544	0.948
500	0.6	0	Cox	4.60E-4	0.527	0.950
			Frailty	4.49E-4	0.529	0.950
			Robust	4.60E-4	0.523	0.948
		0.25	Cox	-3.42E-4	0.528	0.914
			Frailty	-1.43E-3	0.785	0.944
			Robust	-3.42E-4	0.586	0.945
	0.1	0	Cox	5.01E-4	0.371	0.953
			Frailty	3.53E-4	0.372	0.952
			Robust	5.01E-4	0.370	0.952
		0.25	Cox	-2.48E-4	0.371	0.944
			Frailty	-6.86E-4	0.456	0.943
			Robust	-2.48E-4	0.378	0.950
500	0.2	0	Cox	-7.04E-4	0.371	0.949
			Frailty	-6.09E-4	0.372	0.948
			Robust	-7.04E-4	0.370	0.948
		0.25	Cox	-3.90E-4	0.371	0.941

Total Sample Size	Proportion Of twins	Kenall's τ	Method	Mean β	Median Width	Coverage
0.6	0.6	0	Frailty	1.67E-4	0.513	0.944
			Robust	-3.90E-4	0.385	0.950
			Cox	-1.24E-3	0.371	0.954
			Frailty	-1.22E-3	0.372	0.954
			Robust	-1.24E-3	0.370	0.953
			Cox	-1.63E-4	0.372	0.918
			Frailty	7.80E-4	0.569	0.948
			Robust	-1.63E-4	0.415	0.949

Table 3

Findings under gamma frailty model data generation when twins are randomized to the same arm; log hazard ratio=0.405; censoring rate=0.10

Total Sample Size	Proportion of twins	Kenall's τ	Method	Mean β	Median Width	Coverage
250	0.1	0	Cox	0.409	0.536	0.947
			Frailty	0.418	0.538	0.941
			Robust	0.409	0.532	0.945
	0.2	0.25	Cox	0.252	0.529	0.785
			Frailty	0.318	0.581	0.849
			Robust	0.252	0.539	0.794
	0.6	0	Cox	0.409	0.536	0.953
			Frailty	0.418	0.538	0.949
			Robust	0.409	0.532	0.950
500	0.1	0.25	Cox	0.250	0.529	0.772
			Frailty	0.348	0.692	0.891
			Robust	0.250	0.549	0.790
	0.2	0	Cox	0.409	0.536	0.951
			Frailty	0.416	0.538	0.949
			Robust	0.409	0.531	0.949
	0.6	0.25	Cox	0.248	0.530	0.755
			Frailty	0.376	0.793	0.937
			Robust	0.248	0.591	0.817
	0.1	0	Cox	0.407	0.378	0.952
			Frailty	0.414	0.379	0.944
			Robust	0.407	0.376	0.950
	0.2	0.25	Cox	0.246	0.373	0.610
			Frailty	0.342	0.483	0.845
			Robust	0.246	0.381	0.626
	0.6	0	Cox	0.407	0.378	0.945
			Frailty	0.413	0.379	0.941
			Robust	0.407	0.376	0.942
		0.25	Cox	0.246	0.373	0.602

Total Sample Size	Proportion of twins	Kenall's τ	Method	Mean β	Median Width	Coverage
			Frailty	0.364	0.521	0.905
			Robust	0.246	0.389	0.638
	0.6	0	Cox	0.407	0.378	0.950
			Frailty	0.411	0.379	0.949
			Robust	0.407	0.376	0.948
		0.25	Cox	0.246	0.373	0.592
			Frailty	0.385	0.574	0.942
			Robust	0.246	0.420	0.673

Table 4

Findings under Gumbel's bivariate exponential model data generation when twins are randomized to the same arm; log hazard ratio=0; censoring rate=0.10

Total Sample Size	Proportion of twins	Correlation ρ	Method	Mean $\hat{\beta}$	Median Width	Coverage
250	0.1	0	Cox	-8.85E-4	0.527	0.950
			Frailty	-1.06E-3	0.528	0.949
			Robust	-8.85E-4	0.524	0.948
	0.25		Cox	2.11E-3	0.527	0.950
			Frailty	1.78E-3	0.528	0.948
			Robust	2.11E-3	0.525	0.949
	0.2	0	Cox	2.00E-3	0.527	0.949
			Frailty	1.82E-3	0.528	0.949
			Robust	2.00E-3	0.524	0.947
500	0.1	0.25	Cox	-3.21E-4	0.527	0.948
			Frailty	-3.31E-4	0.528	0.947
			Robust	-3.21E-4	0.524	0.946
		0	Cox	-1.23E-3	0.527	0.949
			Frailty	-9.21E-4	0.529	0.948
			Robust	-1.23E-3	0.523	0.947
	0.25	0.25	Cox	-1.08E-3	0.527	0.950
			Frailty	-1.21E-3	0.529	0.950
			Robust	-1.08E-3	0.523	0.948
		0	Cox	4.95E-4	0.371	0.950
			Frailty	3.89E-4	0.372	0.948
			Robust	4.95E-4	0.370	0.949
	0.2	0.25	Cox	6.13E-4	0.371	0.948
			Frailty	6.47E-4	0.372	0.946
			Robust	6.13E-4	0.370	0.947
		0	Cox	1.32E-3	0.372	0.951
			Frailty	1.48E-3	0.372	0.950
			Robust	1.32E-3	0.370	0.951
	0.25	0.25	Cox	-1.24E-3	0.371	0.949

Total Sample Size	Proportion of twins	Correlation ρ	Method	Mean $\hat{\beta}$	Median Width	Coverage
0.6	0	0	Frailty	-1.21E-3	0.372	0.948
			Robust	-1.24E-3	0.370	0.949
			Cox	-2.41E-4	0.372	0.947
			Frailty	-2.22E-4	0.372	0.948
	0.25	0.25	Robust	-2.41E-4	0.370	0.947
			Cox	-1.36E-4	0.372	0.946
			Frailty	-1.20E-4	0.372	0.946
			Robust	-1.36E-4	0.370	0.945

Table 5

Findings under Gumbel's bivariate exponential model data generation when twins are randomized to the same arm; log hazard ratio=0.693; censoring rate=0.10

Total Sample Size	Proportion of twins	Correlation ρ	Method	Mean $\hat{\beta}$	Median Width	Coverage
250	0.1	0	Cox	0.698	0.554	0.948
			Frailty	0.713	0.557	0.937
			Robust	0.698	0.547	0.944
		0.25	Cox	0.701	0.554	0.945
			Frailty	0.715	0.557	0.937
			Robust	0.701	0.547	0.943
	0.2	0	Cox	0.696	0.554	0.949
			Frailty	0.709	0.557	0.941
			Robust	0.696	0.547	0.946
500	0.6	0.25	Cox	0.699	0.554	0.947
			Frailty	0.713	0.557	0.941
			Robust	0.699	0.547	0.944
			Cox	0.698	0.553	0.954
		0	Frailty	0.708	0.557	0.951
			Robust	0.698	0.547	0.950
	0.1	0.25	Cox	0.699	0.554	0.950
			Frailty	0.709	0.557	0.947
			Robust	0.699	0.547	0.947
			Cox	0.694	0.390	0.949
		0	Frailty	0.703	0.392	0.941
			Robust	0.694	0.387	0.947
	0.2	0.25	Cox	0.696	0.390	0.949
			Frailty	0.705	0.392	0.937
			Robust	0.696	0.388	0.947
			Cox	0.695	0.390	0.943
		0	Frailty	0.704	0.392	0.935
			Robust	0.695	0.387	0.943
		0.25	Cox	0.695	0.390	0.947

Total Sample Size	Proportion of twins	Correlation ρ	Method	Mean β	Median Width	Coverage
			Frailty	0.704	0.392	0.939
			Robust	0.695	0.387	0.946
	0.6	0	Cox	0.695	0.390	0.950
			Frailty	0.701	0.392	0.946
			Robust	0.695	0.387	0.948
		0.25	Cox	0.696	0.391	0.953
			Frailty	0.702	0.392	0.949
			Robust	0.696	0.388	0.951

Table 6

Analyses of length of hospital stay (in days) for the PROPHET trial

Analysis	Variance of Random Effect	Hazard Ratio ^a	Standard Error (Log Scale)	95% Confidence Interval
Cox	N/A	0.957	0.117	(0.760, 1.20)
Frailty	0.326	0.935	0.151	(0.696, 1.26)
Robust	N/A	0.957	0.119	(0.758, 1.21)

^aPlacebo is the reference group.

Table 7

Analyses of days of greater than 40% oxygen for the PROPHET trial

Analysis	Variance of Random Effect	Hazard Ratio ^a	Standard Error (Log Scale)	95% Confidence Interval
Cox	N/A	1.13	0.116	(0.902, 1.42)
Frailty	0.519	1.25	0.164	(0.908, 1.73)
Robust	N/A	1.13	0.120	(0.894, 1.43)

^aPlacebo is the reference group.