Published in final edited form as:

Epidemiology. 2010 July ; 21(4): 557–565. doi:10.1097/EDE.0b013e3181e090ed.

Inference for mutually exclusive competing events through a mixture of generalized gamma distributions

William Checkley¹, Roy G Brower¹, and Alvaro Muñoz² for the NIH Acute Respiratory Distress Syndrome Network Investigators

¹Division of Pulmonary and Critical Care, School of Medicine, Johns Hopkins University, Baltimore, MD

²Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

Abstract

Background—Time-to-event data with two or more types of endpoints are found in many epidemiologic settings. Instead of treating the times for one of the endpoints as censored observations for the other, we present an alternative approach where we treat competing events as distinct outcomes in a mixture. Our objective was to determine if and how the mixture was modified in response to an intervention.

Methods—We used a mixture of generalized gamma distributions to concatenate the overall frequency and distribution of the times of two competing events commonly observed in critical care trials, namely unassisted breathing followed by hospital discharge alive and in-hospital death. We applied our proposed methods to data from two randomized clinical trials of critically ill patients.

Results—Mechanical ventilation with lower tidal volumes modified the mixture ($P = 0.103$) when compared with traditional tidal volumes by lowering the overall frequency of death ($P = 0.005$), rather than through affecting either the distributions of times to unassisted breathing ($P = 0.477$) or times to death ($P = 0.718$). Likewise, use of a conservative versus a liberal fluid management modified the mixture ($P < 0.001$) by achieving earlier times to unassisted breathing ($P < 0.001$) and not through affecting the overall frequency of death ($P = 0.202$) or the distribution of times to death ($P = 0.693$).

Conclusions—A mixture approach to competing risks provides a means to determine the overall effect of an intervention and insights into how intervention modifies the components of the mixture.

Time-to-event data with two or more types of endpoints are found in many epidemiological settings and are commonly analyzed using methods for competing risks.^{1–3} A competing risk is defined as an endpoint that precludes the observation of another event.¹ An example of data with competing risks is found in studies of critically ill patients who require mechanical ventilation, where observation of important clinical events such as recovery of the ability to breathe without assistance is precluded by the competing event of death. The most common approach for the analysis of data with competing risks is to treat the times for one of the endpoints as censored observations for the other. This approach is closely linked to the use of semi-parametric models for the analysis of cause-specific hazards under a proportionality assumption.^{1–3}

The act of censoring the competing event is also an attempt to be linked to the causal question and the counterfactual outcome of “what would be the distribution of times to unassisted breathing had nobody died?” Questions of this sort are problematic not only in the context of our example but also in other epidemiologic settings where it is unrealistic to expect that any of the competing events, such as death in the intensive care unit, would never occur or that the competing event could be completely prevented by an intervention. Interventions in such settings are designed to modify the mixture of the two competing events towards a more favorable health outcome, such as reducing overall mortality, achieving unassisted breathing at earlier times, or both. Analytic methods that measure how the mixture is modified according to an intervention are important because they provide the basis for a flexible predictive model in the setting of competing risks such that several different measures of association, including those based on cause-specific hazards, can be derived and their departure from null results can be assessed. Such methods should simultaneously incorporate all the competing events as bona fide outcomes and not simply as one event censoring the other.

The purpose of this paper is to describe an application of a mixture of parametric survival distributions for the analysis of time-to-event data with competing risks, which is an approach previously used in several areas of epidemiologic research.^{4–9} Specifically, we treat competing events as distinct outcomes in a mixture of parametric survival distributions and use the three-parameter generalized gamma distribution¹⁰ to summarize the times to events. We demonstrate the value of our approach using data from two clinical trials in mechanically ventilated, critically ill patients with acute lung injury^{11,12} in which there were two competing events, namely unassisted breathing and in-hospital death. Our inferential objective was to determine how the mixture of unassisted breathing and death changed as a result of two distinct protocol-driven interventions which were assigned randomly in the setting of a clinical trial. We did not attempt to make inferences about the causal effects of treatments, for which there are well-developed methods based on the concept of principal stratification and the estimation of survivor-average causal effects.^{13–17} Instead, on the basis of an appropriate goodness-of-fit of our proposed models, we provide a full description of all components in the mixture and we characterize how interventions modify the mixture. Of particular interest is to jointly contrast the cumulative percentages of patients who achieved unassisted breathing and the cumulative percentages of patients who died at different times after randomization according to treatment arms.

METHODS

Study sample

We used data from two clinical trials conducted by the National Institutes of Health (NIH) Acute Respiratory Distress Syndrome Network investigators.^{11,12} The first clinical trial randomized 861 patients to either a traditional tidal volume strategy (mechanical ventilation with tidal volumes of 12 ml per kilogram of predicted body weight and plateau pressure < 50 cm H₂O) or a lower tidal volume strategy (mechanical ventilation with tidal volumes of 6 ml per kilogram of predicted body weight and plateau pressure < 30 cm water). The second clinical trial randomized 1000 patients to either a fluid liberal strategy (central venous pressure goal of 10–14 mm Hg or pulmonary artery occlusion pressure of 14–18 mm Hg) or a fluid-conservative strategy (central venous pressure goal < 4 mmHg or pulmonary artery occlusion pressure < 8 mm Hg).

Competing events

Our analyses had two competing events, unassisted breathing followed by discharge home alive and in-hospital death. Uncensored observations correspond to either the time when a

patient achieved unassisted breathing and was subsequently discharged home alive with unassisted breathing or to the time when a patient died in the hospital.

There were some patients for whom the date of unassisted breathing could not be precisely determined but the only information available was that it occurred between the last date a person was known to be on mechanical ventilation and the date the patient was discharged home alive with unassisted breathing. We handled these observations as interval censored data.

In addition, few patients were right-censored because they were neither discharged nor had died by the end of the study, but the day of their last follow-up was recorded. The time at which these patients were right-censored varied according to the type of competing event and by clinical trial. Specifically, in the tidal volume trial, times to unassisted breathing were collected only up to day 28. Thus, we censored putative days to unassisted breathing at 28 days, but we censored days to in-hospital death at the last day of follow-up. In the fluid management trial, times to unassisted breathing and times to death were collected for the duration of follow-up and hence we censored both the times to unassisted breathing and death at the last day of follow-up.

Biostatistical methods

In our analysis, the event of interest was unassisted breathing followed by discharge home alive and the competing event was in-hospital death. If π is the proportion of the total population of patients who achieve unassisted breathing and $(1 - \pi)$ is the complementary proportion of patients who died in the hospital, we used a mixture according to π and $(1 - \pi)$ of two generalized gamma (GG) distributions to model times to unassisted breathing and times to death.¹⁰ We chose the three-parameter GG distribution for this application because this parametric distribution can flexibly accommodate various shapes of hazard patterns. We modeled the times to unassisted breathing with a GG distribution with density $f(t)$, and the times to death with another GG distribution with density $g(t)$. Hence, if T denotes the time to either unassisted breathing or death, the proportion with $T < t$ is given by:

$$\begin{aligned}\Pr(T < t) &= \Pr(T < t, \text{event} = \text{UAB}) + \Pr(T < t, \text{event} = \text{death}) \\ &= \pi[1 - F(t)] + (1 - \pi)[1 - G(t)]\end{aligned}$$

where F and G were the survival functions corresponding to the f and g densities, respectively; $\pi[1 - F(t)]$ is the cumulative incidence of unassisted breathing and $(1 - \pi)[1 - G(t)]$ is the cumulative incidence of death. In Table 1, we describe the types of observed data and corresponding expressions for the contributions to the likelihood function in terms of π , the densities $f(t)$, and $g(t)$ (for uncensored data), and the survival functions $F(t)$ and $G(t)$ (for censored data). The mixture approach appropriately incorporated different types of incomplete data due to study design or conduct.

For the tidal volume trial, study group “0” was the 12 ml/kg strategy (reference) and study group “1” was the 6 ml/kg strategy. For the fluid management trial, study group “0” was the liberal strategy (reference) and study group “1” was the conservative strategy. Hence, for each study group $i = \{0, 1\}$, we described the distribution of times to unassisted breathing with three parameters $\varphi_i = \{\beta_i, \sigma_i, \lambda_i\}$, the distribution of times to death with another three parameters $\varphi_i^* = \{\beta_i^*, \sigma_i^*, \lambda_i^*\}$, and the overall probability of unassisted breathing (π_i) with one parameter, for a total of seven parameters for each study group. The full model for the two study groups had 14 parameters.

We used maximum likelihood for parameter estimation and the likelihood ratio tests to compare nested models. The first two tests were fully supported by the randomization assignment. First, we performed an omnibus test to determine if there were any differences in either the overall frequency or the distribution of times in any of the competing events between study groups ($H_0: \pi_0 = \pi_1$ and $\varphi_0 = \varphi_1$ and $\varphi_0^* = \varphi_1^*$ versus $H_A: \pi_0 \neq \pi_1$ or $\varphi_0 \neq \varphi_1$ or $\varphi_0^* \neq \varphi_1^*$). Second, we tested for differences only in the overall frequency of the competing events between study groups $H_0: \pi_0 = \pi_1$ and $\varphi_0 \neq \varphi_1$ and $\varphi_0^* \neq \varphi_1^*$ versus $H_A: \pi_0 \neq \pi_1$ and $\varphi_0 \neq \varphi_1$ and $\varphi_0^* \neq \varphi_1^*$). We defined the relative risk (RR) of unassisted breathing of study group

“1” to study group “0” as the ratio of probabilities of unassisted breathing $\left(\frac{\hat{\pi}_1}{\hat{\pi}_0}\right)$, and the RR of death of study group “1” to study group “0” as the ratio of probabilities of death $\left(\frac{1 - \hat{\pi}_1}{1 - \hat{\pi}_0}\right)$. The third set of comparisons tested if the relative cumulative incidence (RCI) of unassisted

breathing defined as $\frac{\hat{\pi}_1[1 - F_1(t; \hat{\varphi}_1)]}{\hat{\pi}_0[1 - F_0(t; \hat{\varphi}_0)]}$ and, simultaneously, if the RCI of death defined as $\frac{(1 - \hat{\pi}_1)[1 - G_1(t; \hat{\varphi}_1^*)]}{(1 - \hat{\pi}_0)[1 - G_0(t; \hat{\varphi}_0^*)]}$ were different from 1 at any time after randomization. Since the cumulative incidence of one event excludes the participants who previously experienced the other event, the ratio of cumulative incidences of unassisted breathing and the ratio of cumulative incidences of death between study groups should be examined jointly (e.g., simultaneously depicting them in a graph) to provide meaningful inferences. The fourth set of comparisons characterized differences in the conditional distributions of times of the two competing events between study groups. Specifically, we compared $\hat{\varphi}_1$ of the $\hat{\pi}_1$ percent of participants who achieved unassisted breathing to $\hat{\varphi}_0$ of the $\hat{\pi}_0$ percent who achieved unassisted breathing; and, simultaneously, we compared $\hat{\varphi}_1^*$ of the $(1 - \hat{\pi}_1)$ percent of participants who died to $\hat{\varphi}_0^*$ of the $(1 - \hat{\pi}_0)$ percent of participants who died. A limitation of this last set of comparisons is they are performed on subsets of survivors and non-survivors and do not follow the original randomization assignment.

For the calculations of confidence intervals, we used 1,000 bootstrap replicates to obtain percentile-based 95% percent confidence intervals (CIs).⁽¹⁸⁾ The 95% bootstrap CIs closely corresponded to estimates based on the delta method, which can be implemented for the mixture model using PROC NLMIXED in SAS (SAS Institute Inc., Cary, NC).

All analyses were conducted in R (R Foundation for Statistical Computing, www.r-project.org). Software to fit these types of mixture models is not commercially available; however, publicly-available algorithms at the Johns Hopkins STATEPI website (www.statepi.jhsph.edu) greatly facilitate the development of maximum likelihood procedures in R, STATA (StataCorp, College Station, TX) or SAS (SAS Institute, Cary, NC).

RESULTS

We summarized baseline characteristics for both clinical trials in Table 2. There were no differences in age, sex, APACHE III, baseline plateau pressure and $\text{PaO}_2/\text{FiO}_2$ (ratio of partial pressure of oxygen in arterial blood to fraction of inspired oxygen) between the 2 study groups in either clinical trial.^{11,12}

Tidal volume trial

The mixture of generalized gamma distributions summarized the cumulative incidence of unassisted breathing in the 12 ml/kg strategy with $\hat{\phi}_0 = \{2.19, 1.13, 0.09\}$ and $\hat{\pi}_0 = 0.60$, and in the 6 ml/kg strategy with $\hat{\phi}_1 = \{2.20, 1.07, 0.34\}$ and $\hat{\pi}_1 = 0.69$; and, it summarized the cumulative incidence of death in the 12 ml/kg strategy with $\hat{\varphi}_0^* = \{2.40, 1.03, 0.18\}$ and $(1 - \hat{\pi}_0) = 0.40$, and in the 6 ml/kg strategy with $\hat{\varphi}_1^* = \{2.61, 1.02, 0.45\}$ and $(1 - \hat{\pi}_1) = 0.31$. To assess goodness-of-fit of the parametric models, we compared parametric and non-parametric estimates of the cumulative percentages of individuals in the tidal volume trial who either achieved unassisted breathing or died at various days after randomization (Figure 1). There was good agreement between the non-parametric and parametric estimates. The non-parametric estimates of the cumulative incidences of unassisted breathing stopped at day 28 because data on the times to unassisted breathing were not collected beyond that time period in this clinical trial. We present descriptive statistics of the observed outcomes in Table 3.

The omnibus test resulted in a marginally-significant difference between study groups (P -value = 0.103; likelihood ratio tests). The overall estimated probability of death was 9 percent higher (top panel of Figure 2, P -value = 0.005; likelihood ratio tests) in the 12 ml/kg strategy than in the 6 ml/kg strategy (40 percent versus 31 percent, respectively). We show the three components of the mixture model in Figure 2: the overall probability of unassisted breathing and its complement, the overall probability of death; the conditional times of unassisted breathing among survivors; and, the conditional times of death. The conditional times to unassisted breathing and the conditional times to death were similar between study groups (center and bottom panels of Figure 2, respectively). The estimated median time-to-unassisted breathing among survivors was 8.6 days (95% bootstrap CI = 7.3 – 10.3 days) in the 12 ml/kg strategy and 8.0 days (6.9 – 9.3 days) in the 6 ml/kg strategy.

The ratio of cumulative incidences of unassisted breathing favored the 6 ml/kg strategy over the 12 ml/kg strategy at any time after randomization (P -value = 0.03; likelihood ratio tests). On average, the cumulative incidence of unassisted breathing was about 20 percent greater in the 6 ml/kg strategy than the 12 ml/kg (Figure 3). The ratio of cumulative incidences of unassisted breathing of the 6 ml/kg to 12 ml/kg strategy was statistically significant by day 8 after randomization, at which time 21 percent more patients had achieved unassisted breathing in the 6 ml/kg than in the 12 ml/kg strategy (ratio of cumulative incidences of unassisted breathing of the 6 ml/kg to the 12 ml/kg strategy = 1.21 (95% bootstrap CI = 1.00 – 1.49). The relative risk of achieving unassisted breathing of the 6 ml/kg to the 12 ml/kg strategy was 1.16 (1.04 – 1.29). The ratio of cumulative incidences of unassisted breathing of the 6 ml/kg to 12 ml/kg strategy was fairly constant and was not statistically different from the RR of unassisted breathing (P -value = 0.477; LRT). Similarly, the ratio of cumulative incidences of death of the 6 ml/kg to 12 ml/kg strategy was fairly constant and not statistically different from the RR of death (Figure 3; P -value = 0.718; likelihood ratio tests). These findings are congruent with those shown in Figure 2 whereby differences in the ratio of cumulative incidences of unassisted breathing were attributed to differences in the overall frequency of the competing events.

Fluid management trial

The mixture of generalized gamma distributions summarized the cumulative incidence of unassisted breathing in the liberal strategy with $\hat{\phi}_0 = \{2.22, 0.81, -0.01\}$ and $\hat{\pi}_0 = 0.68$, and in the conservative strategy with $\hat{\phi}_1 = \{1.80, 0.74, -0.44\}$ and $\hat{\pi}_1 = 0.72$; and, it summarized the cumulative incidence of death in the liberal strategy with $\hat{\varphi}_0^* = \{2.71, 1.23, 0.35\}$ and $(1 - \hat{\pi}_0) = 0.32$, and in the conservative strategy with $\hat{\varphi}_1^* = \{2.47, 1.34, 0.16\}$ and $(1 - \hat{\pi}_1) = 0.28$. We show good agreement between parametric and non-parametric estimates of the cumulative

percentages of individuals in the fluid-management trial who either achieved unassisted breathing or died at various days after randomization (Figure 4). We present descriptive statistics of the observed outcomes for this trial in Table 4.

The omnibus test detected a significant difference in the overall frequency or timing in at least one of the competing events between study groups (p -value < 0.001 ; likelihood ratio tests). The estimated overall probability of death was 4 percent higher in the liberal strategy than in the conservative strategy (32% versus 28%, respectively); however, this difference was not statistically significant (top panel of Figure 5, P -value = 0.202; likelihood ratio tests). We show the three components of our modeling approach in Figure 5. The conditional times to unassisted breathing were significantly longer in the liberal strategy than in the conservative strategy whereas the conditional times to death were not (center and bottom panels of Figure 5, respectively). The estimated median time-to-unassisted breathing among survivors was 9.2 days (95% bootstrap CI = 8.4 – 10.1 days) for the liberal strategy and 6.7 days (6.2 – 7.4 days) for the conservative strategy.

The ratio of cumulative incidences of unassisted breathing favored the conservative strategy over the liberal strategy at any time after randomization (p -value < 0.001 ; likelihood ratio tests). Shortly after randomization, the cumulative incidence of unassisted breathing was about 50% greater in the conservative strategy than in the liberal strategy (Figure 6). The ratio of cumulative incidences of unassisted breathing was statistically significant by day 3, and favored the conservative strategy over the liberal strategy (ratio of cumulative incidences of unassisted breathing of the conservative to the liberal strategy = 1.55 [95% bootstrap CI = 1.13 – 2.64]). That is, 55% more patients had achieved unassisted breathing by day 3 in the conservative strategy than in the liberal strategy. The ratio of cumulative incidences of unassisted breathing decreased thereafter but remained statistically significant through day 28 after randomization (1.10 [1.01 – 1.20]). The RR of achieving unassisted breathing of the conservative to the liberal strategy was 1.05 (95% bootstrap CI = 0.97 – 1.14). The ratio of cumulative incidences of unassisted breathing of the conservative to the liberal strategy was statistically greater than the RR of unassisted breathing at some time after randomization (P -value < 0.001 ; likelihood ratio tests). Indeed, the ratio of cumulative incidences of unassisted breathing of the conservative to liberal strategy was greater than the RR of unassisted breathing in the first 12 days after randomization. In contrast, the ratio of cumulative incidences of death of the conservative to liberal strategy was not statistically different from the RR of death (P -value = 0.693; likelihood ratio tests). These findings are congruent with those shown in Figure 5 whereby conservative strategy achieved unassisted breathing earlier than patients in the liberal strategy but that overall mortality was not affected between study groups.

DISCUSSION

There are several approaches that an epidemiologist can use to analyze time-to-event data when there are competing risks. The choice of approach depends on the inferential objective of the analysis. One approach is to use a composite event or an index, such as the ventilator-free days score.⁽¹⁹⁾ The attractive feature of such an approach is that one can use standard statistical methods for the analysis of data; however, it does not provide any information on how an intervention or exposure differentially modifies the distribution of times and the overall occurrence of each competing event.

A second approach is to use methods for competing risks,^{1–3} in which the times of the competing event are treated as censored observations of the event of interest. This approach is used to compare cause-specific hazards for one of competing events between interventions; however, it cannot be used to determine whether a difference is due only to a

change in overall frequency or due only to a change in the distribution of times for this event. Furthermore, the assumption of proportionality of the cause-specific hazards may preclude the ability to determine if there are any true differences in the distribution of times between interventions of either endpoint.

A third approach is that of Fine and Gray, which has the attractive feature that it compares the sub-hazards associated with the cumulative incidence functions but it is also subjected to proportionality of sub-hazards functions.²⁰ The Fine and Gray method makes the descriptive analysis using cumulative incidences congruent with the relative hazards from regression models of the sub-hazards functions, and it is widely available in commercial statistical packages, eg, STATA version 11.

A fourth approach is to use methods based on principal stratification for the estimation of survivor-average causal effects.^{13–17} This approach is used to estimate causally-interpretable differences between treatments, i.e., differences in the distribution of times of unassisted breathing between study groups in the subgroup of patients who would have survived regardless of assignment to either study group. A drawback of this approach is that it relies on additional assumptions as some of the potential outcomes are de facto unobservable. Moreover, in the context of the survivor average causal effects, the assumptions required are often implausible. Thus, a sensitivity analysis is typically conducted to assess how large of a violation of the assumptions is necessary to completely eliminate the effect.¹⁷

Our approach was to concatenate the two competing events as distinct outcomes in a mixture.^{4,6,9} The inferential aim of this approach was to determine how the mixture of unassisted breathing and of death was jointly affected by an intervention. By incorporating parameters for the overall frequency and for the distribution of times of each competing event in the mixture, a full range of contrasts can be described under a unified regression approach based on different measures including: an omnibus test to determine if there are differences in either the overall frequencies or in the distribution of times of the competing events between study groups; differences in only the overall frequencies between study groups; differences in the sub-distribution cumulative incidences of either competing event between study groups; and, differences in the conditional hazards of either competing event between study groups. Both the omnibus test and the test of the overall frequencies follow the original randomization assignment and are therefore causally interpretable. Tests of the ratios of cumulative incidences of the competing events between study groups in settings in which the events can be viewed as mutually exclusive are causally interpretable because these comparisons are being made over the entire study sample; however, we stress that these need to be evaluated jointly to provide meaningful inferences. On the other hand, tests of the conditional hazards are not causally interpretable because they are performed on subsets of survivors and non-survivors and do not follow the original randomization assignment. Our goal was to provide a descriptive framework to identify what components in a mixture were differentially affected by interventions in the setting of a randomized clinical trial. As we have presented here (Figures 1 – 3 for the tidal volume trial and Figures 4 – 6 for the fluid management trial), it is important to jointly interpret the multiple tests as it is possible that the omnibus test is not significant because only a component is different (eg, tidal volume trial); and conversely, that the omnibus test may significant but only one component is different (eg, fluid management trial). Other inferential challenges may be present, such as a non-significant omnibus test because the components cancel each other. Causal inferences on potential outcomes are more appropriately achieved by principal stratification approaches.

The use of mixtures with two exponential distributions for the analysis of multiple failure types was originally proposed by Cox.^{4,5} Subsequent applications of mixtures have been

used in epidemiological studies of cancer and HIV/AIDS.^{6,7-9,21} Our proposed approach offers additional methodological advantages over traditional methods for competing risks. First, it utilizes a flexible yet succinct, fully parametric approach that can accommodate different forms of censoring to model both the overall frequency and timing of two endpoints. In particular, our methods can incorporate partial information on patients for whom we do not have a date of discontinuation of mechanical ventilation but who were discharged alive without unassisted breathing sometime between 28 days and the day of discharge. Not only can we incorporate different types of censoring with this method, but as evident from the contributions of the likelihood function, our approach does not handle an observed time of a competing event as a censored time for the event of interest, but rather as a distinct outcome in the mixture. Second, relative to non-parametric or semi-parametric approaches, a parametric model affords opportunities for increased efficiency and allows for the complete characterization of the variability of competing events between interventions. Third, in contrast to classical competing risks methods,¹⁻³ our model is more flexible because we do not need to assume proportionality of hazards. Indeed, in our analysis of the fluid management trial, the cause-specific hazards between the study groups did not fulfill the assumption of proportionality. A limitation of our approach is that the generalized gamma distribution may not appropriately summarize the times to events in all cases thus requiring more complex models.^{6,21} Hence, a graphical assessment of goodness-of-fit is imperative. Another limitation is that in the presence of substantial amount of censored observations, the estimation of the mixture probabilities could be unstable; in such cases, the Fine and Gray approach may be preferable.

An important difference between our approach and one based on the analysis of cause-specific hazards is that our approach provides a practical strategy to separate the effects of interventions on overall mortality from that on the timing of unassisted breathing under a single analytical framework. In this analysis we report that the use of lower tidal volumes in patients with acute lung injury increased the overall chance of achieving unassisted breathing and decreased overall mortality but did not affect the times to unassisted breathing when compared with traditional tidal volumes. In contrast, a conservative protocol for fluid management in patients with acute lung injury resulted in earlier times to unassisted breathing but it did not increase the overall chance of achieving unassisted breathing or decrease overall mortality when compared with a liberal protocol for fluid management. Such determinations cannot be accomplished with an analysis based on cause-specific hazards under a proportionality assumption.

In summary, we used a mixture of generalized gamma distributions to concatenate the times to events of multiple endpoints. In contrast to methods for competing risks, our approach does not treat one event as a censored observation of the other but instead as bona fide outcomes in a mixture. Unlike methods for competing risks, our approach is free of the assumption of proportionality and can be used to fully characterize differences in the overall frequency and timing of the competing events.

Acknowledgments

Sources of support: Supported in part by a post-doctoral National Research Service Award of the National Heart, Lung, and Blood Institute (F32-HL090179), a Clinician Scientist Award from the Johns Hopkins University, and a Pathway to Independence Award (K99/R00) of the National Heart, Lung, and Blood Institute (K99-HL096955) awarded to William Checkley. The clinical trials were supported by the National Heart, Lung, and Blood Institute (contracts NO1-HR-46046-64 and NO1-HR-16146-54). Development of statistical methods by Alvaro Muñoz was supported by the National Institutes of National Diabetes and Digestive and Kidney Diseases (U01-DK-66116).

We thank members of the Acute Respiratory Distress Syndrome Network Publications Committee for helpful comments in an earlier version of this manuscript.

References

1. Marubini, E.; Valsecchi, MG. Analysing survival data from clinical trials and observational studies. John Wiley and Sons; Chichester: 2004.
2. Pintilie, M. Competing risks: a practical perspective. John Wiley and Sons; Chichester: 2006.
3. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med.* 2007; 26:2389–430. [PubMed: 17031868]
4. Cox DR. The analysis of exponentially distributed lifetimes with 2 types of failure. *J R Stat Soc Series B Stat Methodol.* 1959; 21:411–421.
5. Cox, DR.; Oakes, D. Analysis of Survival Data. Chapman & Hall; London: 1996.
6. Larson MG, Dinse GE. A Mixture Model for the Regression Analysis of Competing Risks Data. *Applied Stat.* 1985; 34:201–11.
7. Cole SR, Li R, Anastos K, Detels R, et al. Accounting for leadtime in cohort studies: evaluating when to initiate HIV therapies. *Stat Med.* 2004; 23:3351–63. [PubMed: 15493031]
8. Lau B, Cole SR, Moore RD, et al. Evaluating competing adverse and beneficial outcomes using a mixture model. *Stat Med.* 2008; 27:4313–27. [PubMed: 18416435]
9. Lau B, Cole SR, Gange SJ. Competing Risk Regression Models for Epidemiologic Data. *Am J Epidemiol.* 2009; 170:244–56. [PubMed: 19494242]
10. Cox C, Chu H, Schneider MF, et al. Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Stat Med.* 2007; 26:4352–74. [PubMed: 17342754]
11. The ARDS Network. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the ARDS. *N Engl J Med.* 2000; 342:1301–8. [PubMed: 10793162]
12. Wiedemann HP, Wheeler AP, Bernard GR, Thompson BT, Hayden D, deBoisblanc B, Connors AF Jr, Hite RD, Harabin AL. National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Comparison of two fluid-management strategies in acute lung injury. *N Engl J Med.* 2006; 354:2564–75. [PubMed: 16714767]
13. Robins JM. A new approach to causal inference in mortality studies with sustained exposure period – application to control of the healthy worker survivor effect. *Mathematical Modelling.* 1986; 7:1393–1512.
14. Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu Rev Public Health.* 2000; 21:121–45. [PubMed: 10884949]
15. Frangakis CE, Rubin DB, An MW, MacKenzie E. Principal stratification designs to estimate input data missing due to death. *Biometrics.* 2007; 63:641–9. [PubMed: 17824995]
16. Hayden D, Pauler DK, Schoenfeld D. An Estimator for Treatment Comparisons among Survivors in Randomized Trials. *Biometrics.* 2005; 61:305–310. [PubMed: 15737107]
17. Egleston BL, Scharfstein DO, Freeman EE, West SK. Causal inference for non-mortality outcomes in the presence of death. *Biostatistics.* 2007; 8:526–45. [PubMed: 16980696]
18. Efron, B.; Tibshirani, RJ. An Introduction to the Bootstrap. Chapman & Hall; London: 1989.
19. Schoenfeld DA, Bernard GR. ARDS Network. Statistical evaluation of ventilator-free days as an efficacy measure in clinical trials of treatments for acute respiratory distress syndrome. *Crit Care Med.* 2002; 30:1772–7. [PubMed: 12163791]
20. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc.* 1999; 94:496–509.
21. Peng Y, Dear KB, Denham JW. A generalized F mixture model for cure rate estimation. *Statist Med.* 1998; 17:813–30.
22. Li X, Buechner JM, Tarwater PM, et al. A diamond-shaped equiponderant graphical display of the effects of two categorical predictors on continuous outcomes. *The American Statistician.* 2003; 57:1–7.

Figure 1. Parametric and non-parametric estimates of the cumulative percentage of ventilated patients in the tidal volume trial who either achieved unassisted breathing or died at various days after randomization, stratified by study group. Parametric estimates are depicted with smooth lines and non-parametric estimates with steps.

Figure 2.

Estimated values for the overall probability of unassisted breathing, the conditional times to unassisted breathing, and the conditional times to death in the tidal volume trial, stratified by study group. Top panel: Percentages above and below the filled circle correspond to the overall probabilities of death ($1 - \hat{\pi}$) and of unassisted breathing ($\hat{\pi}$), respectively. The whisker-and-brackets correspond to the 95% CIs. Center panel: Median and interquartile range of the conditional times to unassisted breathing in a diamond-and-whiskers plot. The diamond corresponds to the median, and the whiskers correspond to the interquartile range. The amount of filling inside the diamond is proportional to the overall probability of unassisted breathing and the percentage in the center corresponds to the overall probability of unassisted breathing.²² Bottom panel: Similar to the center panel but for the times to death.

Figure 3.

Ratio of cumulative incidences of unassisted breathing and death of the 6 ml/kg to the 12 ml/kg study groups in the NIH Acute Respiratory Distress Syndrome tidal volume trial for the first 28 days after randomization. The continuous, blue line depicts the ratio of cumulative incidences of unassisted breathing of the 6 ml/kg to 12 ml/kg study group. The dashed, blue lines correspond to a 95% bootstrap CI. A ratio of cumulative incidences of unassisted breathing of the 6 ml/kg to the 12 ml/kg study groups above 1 favors the 6 ml/kg strategy and below 1 favors the 12 ml/kg strategy. The blue circle corresponds to the RR of unassisted breathing of the 6 ml/kg to the 12 ml/kg strategy, and vertical blue line corresponds to its 95% bootstrap CI. The continuous, red line depicts the ratio of cumulative incidences of death of the 6 ml/kg to 12 ml/kg study group. The dashed, red lines correspond to a 95% bootstrap CI. A ratio of cumulative incidences of death of the 6 ml/kg to the 12 ml/kg study groups below 1 favors the 6 ml/kg strategy and above 1 favors the 12 ml/kg strategy. The red circle corresponds to the RR of death of the 6 ml/kg to the 12 ml/kg study groups, and the vertical red line corresponds to its 95% bootstrap CI.

Figure 4. Parametric and non-parametric estimates of the cumulative percentage of ventilated patients in the fluid management trial who either achieved unassisted breathing or died at various days after randomization, stratified by study group. Parametric estimates are represented with lines and non-parametric estimates are represented with steps.

Figure 5.

Estimated values for the overall probability of unassisted breathing (and its complement, the overall probability of death), the conditional times to unassisted breathing, and the conditional times to death in the fluid management trial, stratified by study group. Top panel: Percentages above and below the filled circle correspond to the overall probabilities of death ($1 - \hat{\pi}$) and of unassisted breathing ($\hat{\pi}$), respectively. The whisker-and-brackets correspond to the 95% CIs. Center panel: Median and interquartile range of the conditional times to unassisted breathing in a diamond-and-whiskers plot. The diamond corresponds to the median, and the whiskers correspond to the interquartile range. The amount of filling inside the diamond is proportional to the overall probability of unassisted breathing and the percentage in the center corresponds to the overall probability of unassisted breathing.²² Bottom panel: Similar to the center panel but for the times to death.

Figure 6.

Ratio of cumulative incidences of unassisted breathing of the conservative to liberal strategy of fluid management in the NIH Acute Respiratory Distress Syndrome fluid management trial for the first 28 days after randomization. The continuous, blue line depicts the ratio of cumulative incidences of unassisted breathing of the conservative to liberal strategy. The dashed, blue lines correspond to a 95% bootstrap CI. A ratio of cumulative incidences of unassisted breathing of the conservative to liberal strategy above 1 favors the conservative strategy and below 1 favors the liberal strategy. The horizontal, continuous, dark grey line corresponds to the RR of unassisted breathing of the conservative to the liberal strategy, and the light grey shaded area corresponds to its 95% bootstrap CI. The continuous, red line depicts the ratio of cumulative incidences of death of the conservative to liberal strategy. The dashed, red lines correspond to a 95% bootstrap CI. A ratio of cumulative incidences of death of the conservative to liberal strategy below 1 favors the conservative strategy and above 1 favors the liberal strategy. The red circle corresponds to the RR of death of the conservative to liberal strategy, and the vertical red line corresponds to its 95% bootstrap CI.

Table 1

Types of observed data and the corresponding expressions for the contributions to the likelihood function in terms of π , the densities $f(t)$ and $g(t)$ (for uncensored data), and the survival functions $F(t)$ and $G(t)$ (for censored data).

Clinical trial	Observation	In general	Likelihood contribution In terms of the mixture model
Tidal volume	Achieved unassisted breathing at day $t < 28$	$\Pr(T = t, \text{event} = \text{UAB})$	$\pi f(t)$
	In-hospital death at t days	$\Pr(T = t, \text{event} = \text{death})$	$(1 - \pi)g(t)$
	On mechanical ventilation at 28 days and discharged on day $c > 28$	$\Pr(28 < T < c, \text{event} = \text{UAB})$	$\pi[F(28) - F(c)]$
	On mechanical ventilation at 28 days and alive in the hospital at day $c > 28$	$\Pr\left(\begin{array}{l} T > 28, \text{event} = \text{UAB} \\ \text{or} \\ T > c, \text{event} = \text{death} \end{array}\right)$	$\pi F(28) + (1 - \pi)G(c)$
Fluid management	Achieved unassisted breathing at day t	$\Pr(T = t, \text{event} = \text{UAB})$	$\pi f(t)$
	In-hospital death at day t	$\Pr(T = t, \text{event} = \text{death})$	$(1 - \pi)g(t)$
	On mechanical ventilation at t^* days but discharged on day t	$\Pr(t^* < T < t, \text{event} = \text{UAB})$	$\pi[F(t^*) - F(t)]$
	On mechanical ventilation and alive in the hospital at c days	$\Pr(T > c, \text{event} = \text{UAB or death})$	$\pi F(c) + (1 - \pi)G(c)$

Baseline characteristics of two clinical trials conducted by the National Institutes of Health Acute Respiratory Distress Syndrome Network investigators.^a

Table 2

	Tidal volume trial		Fluid management trial	
	Group "0" 12 ml/kg	Group "1" 6 ml/kg	Group "0" Liberal	Group "1" Conservative
Age in years; mean (SD)	51.9 (17.7)	50.2 (16.7)	49.5 (15.9)	50.1 (16.2)
Men; no. (%)	252 (59)	261 (60)	271 (55)	263 (52)
APACHE III score; mean (SD)	84.1 (28.2)	81.0 (28.3)	95.2 (30.7)	93.1 (31.1)
Baseline plateau pressure cm water; mean (SD)	30.5 (8.0)	29.9 (7.7)	26.2 (7.2)	26.1 (6.7)
Baseline PaO ₂ /FiO ₂ ; mean (SD)	148 (71)	149 (68)	150 (67)	154 (74)
Mortality at the end of the study period; no. (%)	173 (40)	134 (31)	150 (30)	134 (27)

^a Sample size numbers for age, male sex, and APACHE score were 429, 432, 497, and 503 for the 4 columns; for baseline plateau pressure, 425, 430, 479, and 486; for baseline PaO₂/FiO₂, 332, 338, 335, and 355; and for mortality 398, 401, 475, and 485 APACHE indicates Acute Physiology and Chronic Health Evaluation; PaO₂/FiO₂ the ratio of partial pressure of oxygen in arterial blood to fraction of inspired oxygen.

Table 3

Descriptive statistics of the outcomes observed in the tidal volume trial.

Observation	12 ml/kg strategy (n = 429)		6 ml/kg strategy (n = 432)	
	No.	Day of observation Median (IQR)	No.	Day of observation Median (IQR)
Unassisted breathing	200	6 (3, 10)	250	6.5 (3, 12)
Death	173	12 (6, 18)	134	12 (5, 21)
Interval censored for Unassisted breathing	48	after 28 (28, 28) before 63 (47, 103)	41	after 28 (28, 28) before 52 (38, 85)
Right censored for Unassisted breathing	8 ^a	after 28 (28, 28)	7 ^b	after 28 (28, 28)
Right censored for death	8 ^a	after 184 (128, 194)	7 ^b	after 165 (68, 189.5)

^a, ^b These numbers represent the same individuals but with differentially right-censored data.
IQR indicates interquartile range.

Table 4

Descriptive statistics of the outcomes observed in the fluid management trial.

Observation	Liberal strategy (n = 497)			Conservative strategy (n = 503)		
	No.	Day of observation	Median (IQR)	No.	Day of observation	Median (IQR)
Unassisted breathing	334	9 (7, 15)		356	7 (4, 11)	
Death	150	14 (6, 25)		134	10 (4, 23)	
Interval censored for unassisted breathing	4	after 0.5 (0.5, 1.4) before 41 (22, 58)		5	after 0.5 (0.5, 4) before 48 (48, 70)	
Right censored for unassisted breathing or death	9	after 90 (90, 91)		8	after 91 (90, 91)	

IQR indicates interquartile range.