SPPH 501 TERM PROJECT: UNDERSTANDING JOINT FRAILTY MODELS FOR RECURRENT EVENTS AND A TERMINAL EVENT WITH AN APPLICATION TO SEVERE COPD EXACERBATIONS

Tae Yoon Lee
Department of Pharmaceutical Sciences
University of British Columbia
Vancouver, BC
taeyoon.lee@alumni.ubc.ca

December 4, 2019

Contents

1 Introduction

Sudden episodes of progressive worsening of respiratory symptoms are referred to as exacerbations or flare-ups. Exacerbations constitute an important component of the natural history of chronic respiratory diseases, such as asthma

and chronic obstructive pulmonary disease (COPD). Moreover, they are a prominent source of economic burden for these diseases. For instance, there are higher healthcare-related costs for patients with asthma exacerbations than those with none (Ivanova et al., 2012), and COPD exacerbations are one of the main causes for hospitalization in Canada (Benady, 2010). Hence, understanding risk factors that associated with exacerbations and predicting the occurrence of exacerbations are imperative research goals.

For the former research goal, it is deemed that the occurrence of an exacerbation is a crucial risk factor for a subsequent exacerbation. This topic has been addressed for COPD in (Suissa et al., 2012) with an appropriate method (Sadatsafavi et al., 2018). Sadatsafavi et al. (2018) shows that there is a positive association between each severe exacerbation and the risk of its subsequent exacerbations in the British Columbia administrative data. On the other hand, this question has not been investigated for asthma. The aim of this report is to understand the methodology used in Sadatsafavi et al. (2018) and apply it to pooled data of two randomized controlled trials for COPD (unfortunately, not asthma data because I did not have access to it at the time).

The rest of this report is organized as follows. The clinical trial data is described along with summary statistics and explanatory plots in Chapter 2. In Chapter 3, motivation for joint modeling of recurrent events and a terminal event is provided, and a joint frailty model is described with a couple of methods, including the one used in Sadatsafavi et al. (2018), for fitting it. In Chapter 4, analysis results are provided and interpreted. Finally, concluding remarks are stated in Chapter 5.

2 Data

The data is pooled from two randomized controlled trials: MACRO (Albert et al., 2011) and OPTIMAL (Aaron et al., 2007) studies; the pooling is supported by Adibi et al. (2019). The purpose of these studies is to investigate the effects of drugs on COPD exacerbations during the study period of 12 months. The study selection criteria are quite similar for these two studies (Table 1). One notable difference is that all OPTIMAL patients were required to have at least one year of COPD exacerbations in the past year prior to entering the study, whereas MACRO patients were not.

In MACRO, 1106 patients were randomized to two groups. The placebo and drug treatment groups were given a sugar pill capsule and a 250 mg capsule of Azithromycin, respectively, to take once a day for 12 months. In OPTIMAL, 417 patients were randomized to three groups: placebo, Triotropium + Salmeterol, and Triotropium + Salmeterol + Fluticasone. These groups were given 2 puffs per day from a drug-free inhaler, a Salmeterol inhaler containing 25 μg/puff of salmeterol, and an Advair inhaler containing 25 μg/puff of salmeterol and 250 μg/puff of fluticasone, respectively, for 12 months. In the pooled data, the placebo groups from both studies are merged together to form the placebo group for the pooled data. The treatment group of MACRO and the Triotropium + Salmeterol treatment group of OPTIMAL are merged together and referred to as the first treatment group or Sal group. Finally, the Triotropium + Salmeterol + Fluticasone treatment group of OPTIMAL is referred to as the second treatment group or Flu in the pooled data. This implies that the sample size for the second treatment group is much smaller than the other two groups (Table 2).

RCT Name	age (years)	Has COPD diagnosis	Has asthma?	pack years of smoking	# of COPD exacerbations in the past year
MACRO	>40	Yes	No	>10	NA
OPTIMAL	>35	Yes	No	>10	>= 1

Table 1: The selection criteria for MACRO and OPTIMAL.

2.1 Exploratory data analysis

The baseline (time-independent) variables are provided in Table 2. There are more male patients than female patients. All the patients were actually more than 40 years old even though the OPTIMAL study age requirement is lower. Looking at BMI, half of the patients are overweight (BMI > 25), and about 25% of them are obese (BMI > 30). Note that patients entered the study at different time points, and based on the recording of those dates, the season (the four quarters of year) of the randomization dates can be used for modeling. The patients were recruited throughout the year fairly evenly, as there is little variability in the sample size across the four quarters.

The lung-related variables are also only measured at the baseline (hence time-independent) and summarized in Table 3. The variable, nowsmk, is the indicator variable for whether a patient smoked in the past year prior to entering the study. The majority of the patients (77%) did not smoke in the past year and possibly quit given the nature of smoking

OPTIMAL	treatment group	gender	age	BMI	randomization season
0:1106	Placebo : 685	Female : 634	Min. :40.66	Min. :13.91	Q1: 448
1: 417	Sal : 698	Male : 889	1st Qu.:60.00	1st Qu.:23.51	Q2: 446
	Flu : 140		Median :66.00	Median :26.77	Q3: 430
			Mean :65.92	Mean :27.70	Q4: 500
			3rd Qu.:72.00	3rd Qu.:31.31	
			Max. :89.00	Max. :54.56	

Table 2: Summary statistics of the baseline variables. OPTIMAL is the indicator variable for whether a patient is in the OPTIMAL clinical study. Randomization season indicates which quarter of the year that a patient entered the study.

addiction. The variable, oxygen, is the indicator variable for whether a patient used oxygen support in the past year. About half of the patients used oxygen support. The variable, packyears, is the number of years that a patient smoked for. Notice that the summary statistics of packyears has nonsensical values; any number in the summary statistics of packyears cannot exceed its corresponding number in that of age (in Table 2). Under a hypothesis that some of the values are measured in months instead of years, a simple correction was applied: packyears was divided by 12 if packyears exceeded age or if the difference between age and packyears is less than 18 (the legal age of smoking).

nowsmk	oxygen	fev1pp	fev1	fvc	packyears	packyears_corrected
0:1167	0:818	Min. : 9.00	Min. :0.150	Min. :0.690	Min. : 4.00	Min. : 2.50
1: 356	1:705	1st Qu.:28.15	1st Qu.:0.720	1st Qu.:1.965	1st Qu.: 36.00	1st Qu.: 5.83
		Median :38.78	Median :0.97	Median :2.45	Median : 50.00	Median :10.25
		Mean :40.09	Mean :1.08	Mean :2.59	Mean : 56.31	Mean :19.54
		3rd Qu.:51.00	3rd Qu.:1.33	3rd Qu.:3.10	3rd Qu.: 74.00	3rd Qu.:35.00
		Max. :85.41	Max. :3.20	Max. :6.45	Max. :230.00	Max. :64.50

Table 3: Summary statistics of the baseline variables related to lung function.

A severe COPD exacerbation is the recurrent event of interest, and death is the terminal event of interest. Only 458 out of 1981 patients had at least one severe COPD exacerbation, and only 70 patients died during the study period. Small event sample sizes are an issue for analysis and interpretation of results due to low statistical power. As there are only 78 patients that experienced more than 2 severe COPD exacerbations (Table 4), two or a higher number of severe COPD exacerbations are grouped together as one level, 3 hereinafter. As the date of severe COPD exacerbations and the death date were recorded, the season (the four quarters of the year) of those events can be extracted and used for modeling (Table 5). The season could serve as a proxy variable for temperature, air pollution, and other time-changing factors that could affect the risk of death or severe COPD exacerbations.

Figure 1 shows severe COPD exacerbation (green circles) and death (red triangle) events across time (measured from randomization) for each patient who experienced any of those events. The plot can be stratified by categorical factors such as, gender (Figure 2) and by treatment group (Figure 3). The stratified plots are useful for examining patterns in each group but are not useful for comparison unless the sample sizes are similar across the levels of the group variable of interest. There seems to be more severe COPD exacerbations in the Sal group than in the placebo group while the number of deaths seems equal, hinting that the exacerbation rate is higher for the Sal group.

To compare the Flu group with others, a cumulative sample mean (CSM) plot is a better alternative:

$$\text{CSM}(t) = \frac{1}{n}\sum_{i=1} N_i(t), \tag{1}$$

where $N_i(t)$ is the number of events of interest over the time interval $[0, t]$ for patient i (Cook and Lawless, 2007). It is another way of visualizing the event data by looking at the proportion of events occurred at any time during the study period. In Figure 4, the slope for the Sal group, is higher than those of the placebo and the Flu group, indicating that the exacerbation rate is the highest for the Sal group. There seems to be no difference in the exacerbate rate between the placebo and Flu groups. However, this could be due to survival bias. Suppose the drug is indeed effective. Then patients

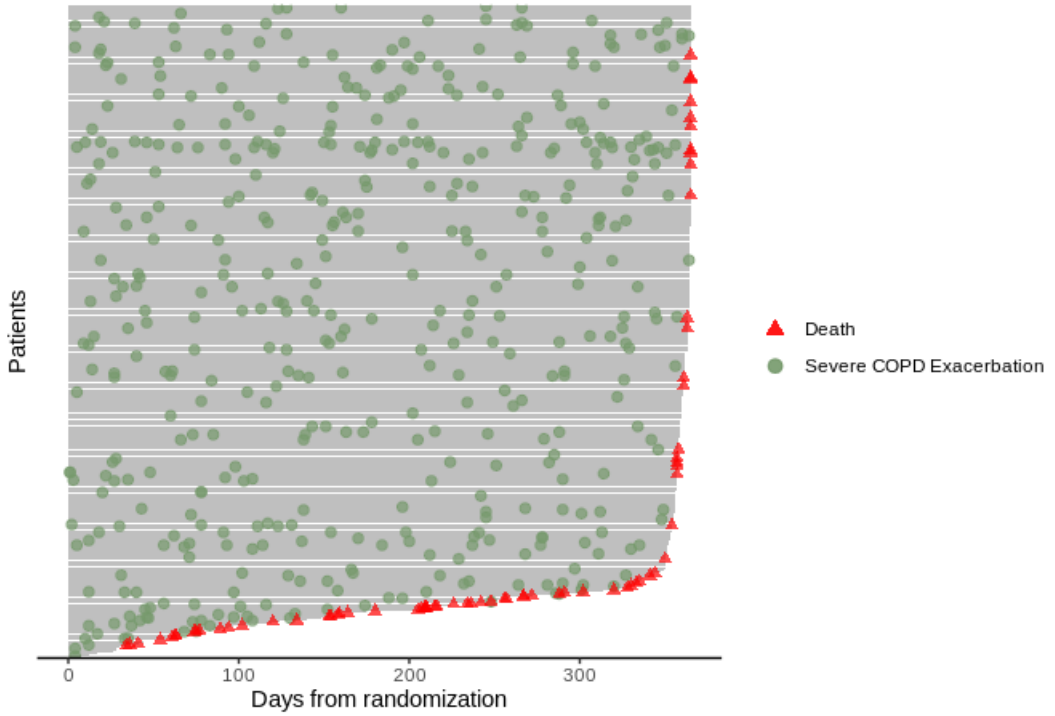

Figure 1: A plot of severe COPD exacerbations (green circle) and deaths (red circle)

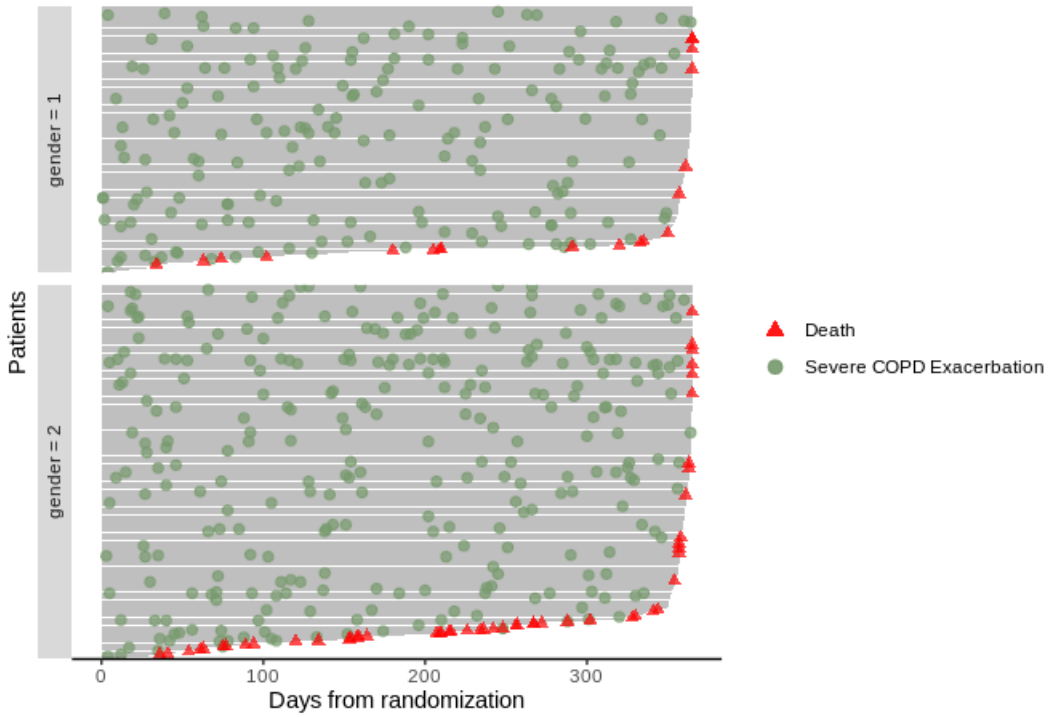

Figure 2: A plot of severe COPD exacerbations (green circle) and deaths (red circle) by gender (1: female; 2: male)

Number of severe COPD exacerbations	frequency
0	1523
1	311
2	83
3	37
4	13
5	7
6	3
7	3
8	1

Table 4: Frequency table of the severe COPD exacerbations experienced by the patients in the study during the 12 months.

severe COPD exacerbation season	severe COPD exacerbation seasonality	death season	death seasonality
1:438	0:842	1:15	0:39
2:432	1:925	2:29	1:31
3:414		3:20	
4:483		4:16	

Table 5: Summary statistics of the season (1: Jan-March; 2: April-June; 3: July-Sept; 4: Oct-Dec) and seasonality (1: Jan-June; 2: July-Dec) of the randomization date and severe COPD exacerbation and death dates.

in the Sal group had a higher chance of being alive, did not die during the trial, and hence could have experienced more severe COPD exacerbations. On the other hand, patients in the placebo group could have died earlier and experienced a smaller number of severe COPD exacerbations as the group.

An important time-dependent variable is the number of severe COPD exacerbations that a patient experienced up to time t during the study period. Its association with the time it takes to get a subsequent severe COPD exacerbation is illustrated in Figure 5. The orange, green, blue, and purple curves are the empirical distributions of the time it took to get the first, second, third, and fourth or higher number of severe COPD exacerbations, respectively, from the zeroth, first, second, and third or higher number of severe COPD exacerbations, respectively. Having at least one severe COPD exacerbation leads to a major change in the distribution of gap times from a somewhat flat distribution to a right skewed distribution. In addition, the gap time appears to become shorter as the number of severe COPD exacerbations increases. Figure 6 shows the boxplot of time to death in days from randomization for each of the total number of severe COPD exacerbations that patients experienced during the study. There seems to be a decrease in time to death as the number of severe COPD exacerbations increases. It is difficult to make any conclusion assertively given much variability due to the small sample size.

3 Method

Two event processes are of interest: a recurrent event process of severe COPD exacerbations and a terminal event process of death. While one can model the two processes independently of each other, Figures 5 and 6 indicate that these two processes might be dependent. More specifically, the follow-up of potential severe COPD exacerbations was stopped by death for some patients. In addition, the recurrence of severe COPD exacerbations appeared to increase the risk of death. There could be unobserved factors that are related to the recurrent event process and affect the risk of death. Hence, this dependence must be taken into consideration to account for potential survival bias caused by death and biased estimation of the risk factors for death due to recurrence of severe COPD exacerbations.

To model the potential dependence, Sadatsafavi et al. (2018) used a joint or shared frailty model (Liu et al., 2004; Cook and Lawless, 2007), in which the two processes are jointly connected through a common frailty term. As the model is complex, the rest of this section is devoted to explaining the model. After formally setting up the problem, gap times

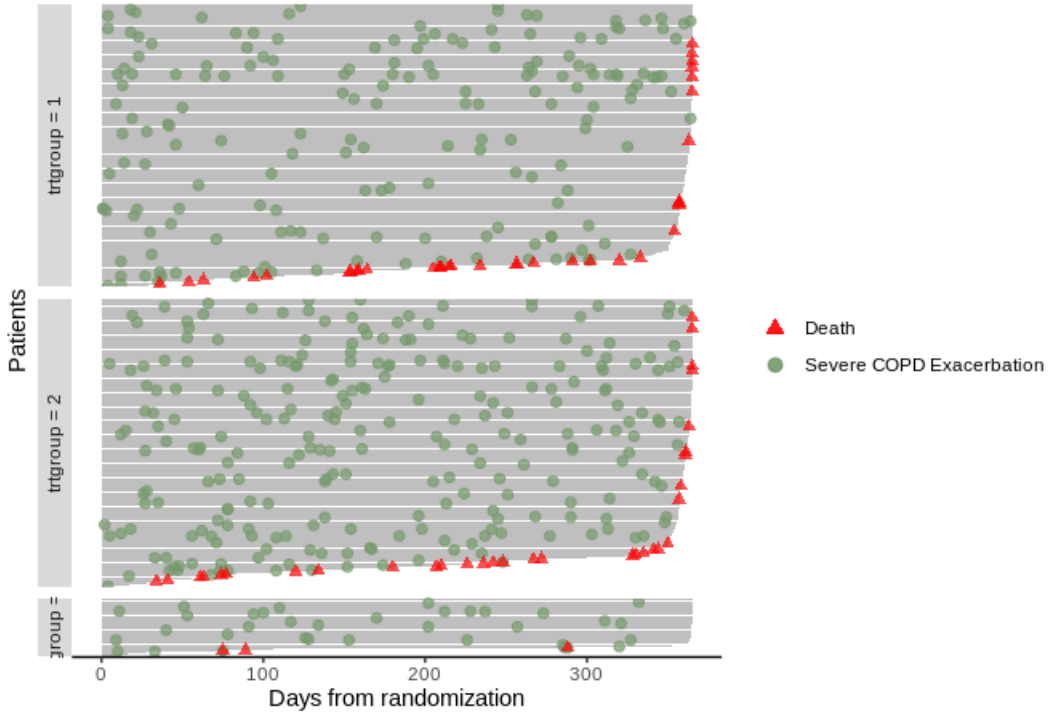

Figure 3: A plot of severe COPD exacerbations (green circle) and deaths (red circle) by treatment group (1: placebo; 2: Sal; 3:Sal+Flu)

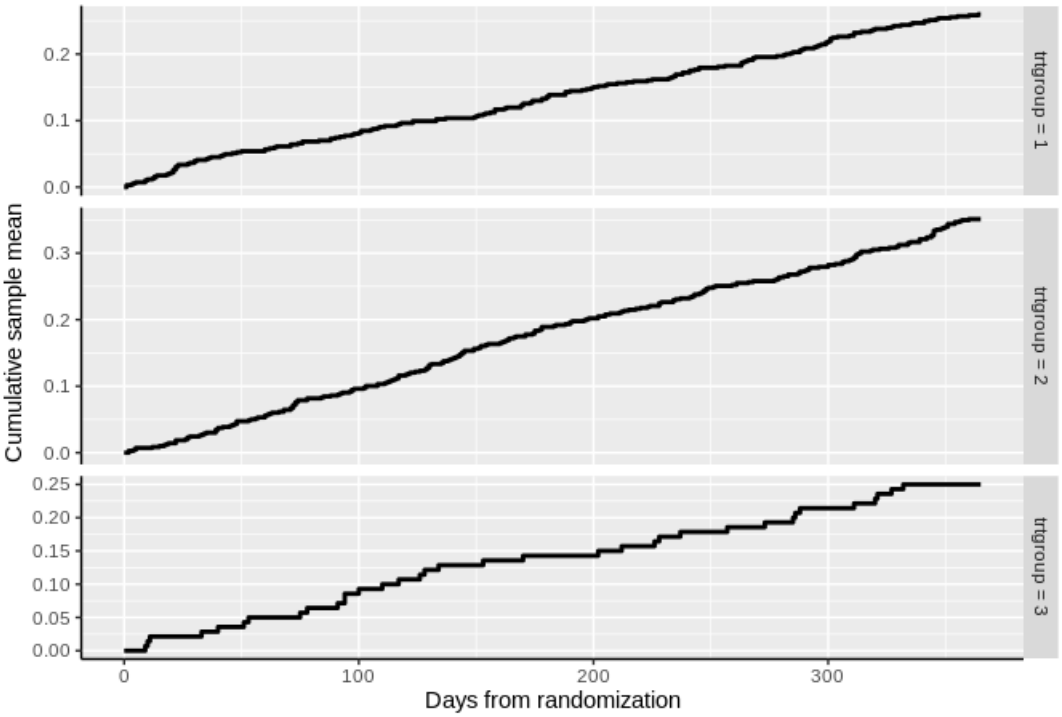

Figure 4: A CSM plot of severe COPD exacerbations by the treatment group (1: placebo; 2: Sal; 3: Sal+Flu).

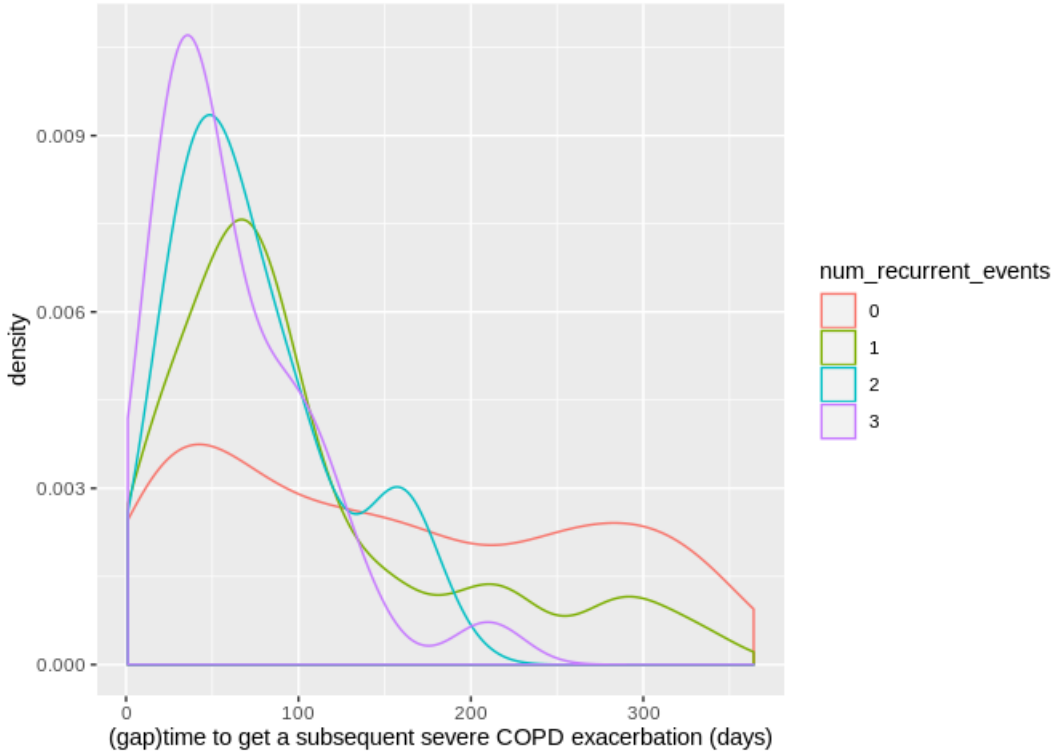

Figure 5:

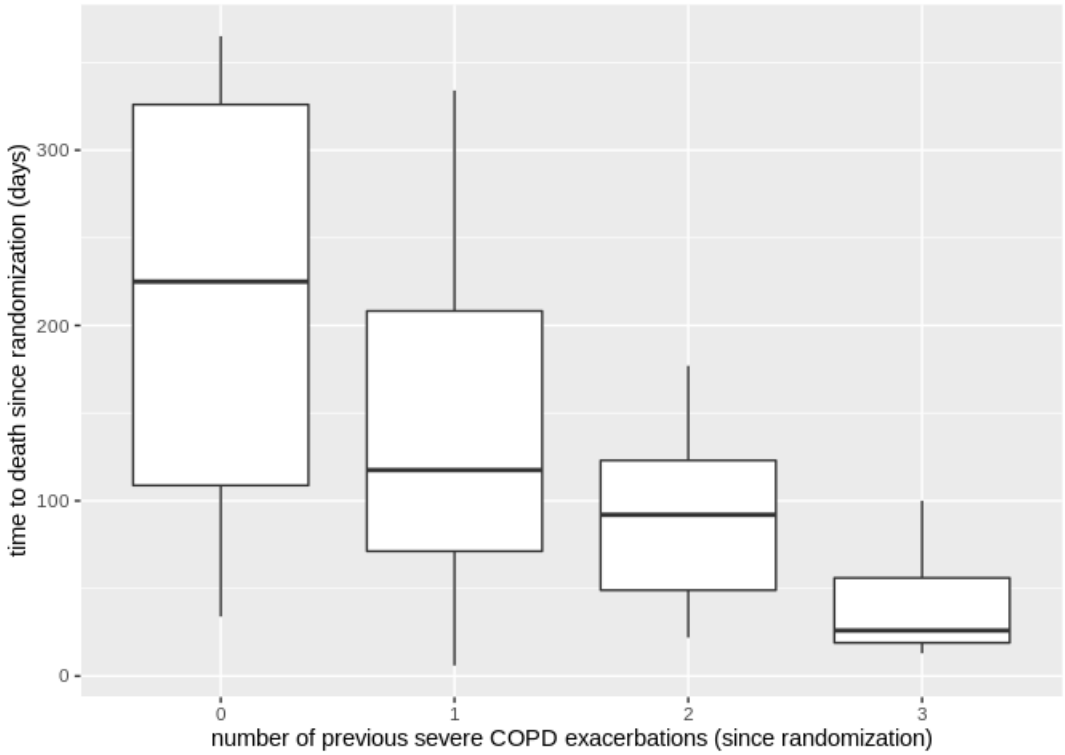

Figure 6:

and common models for them are briefly described. Then frailty models are introduced along with several methods for fitting them. Finally, joint (or shared) frailty models are presented to extend univariate frailty models and model the two processes jointly.

3.1 Set-up

The joint frailty model proposed by Liu et al. (2004) is designed for left-truncated and/or right-censored data, which is setting for the data used in this report and for other datasets, such as the administrative data. For each patient $i \in \{1, \ldots, n\}$, set the left-truncating time (i.e. the start date of the trials) as the starting or zero time. Let T_{ij} be the jth recurrent time or time-to-event of a severe COPD exacerbation for $j = 1, \ldots, n_i$, C_i be the (right-)censoring time, and D_i be the death time (if observed; otherwise set $D_i = \infty$). Furthermore, let $Y_{ij} = \min(T_{ij}, C_i, D_i)$ be the observation time, $\delta_{ij} = 1_{\{Y_{ij}=T_{ij}\}}$ be the recurrent event indicator, $T_i^* = \min(C_i, D_i)$ be the last follow-up time, and $\delta_i^* = 1_{\{T_i^*=D_i\}}$ be the death indicator.

3.2 Gap times

There are two fundamental ways to model recurrent event data: time-to-event and gap (or waiting) times (Cook and Lawless, 2007). For this report, the gap time, $W_{ij} = T_{ij} - T_{i(j-1)}$, is chosen and used to determine the association of a severe COPD exacerbation with the risk of a subsequent severe COPD exacerbation. For time-dependent or independent covariates $z = z(t)$, two commonly used families to model gap times are the proportional hazards model, whose hazard function of W_{ij} given z_i is:

$$h(w_{ij}|z_i) = h_0(w_{ij}) \exp(z_i^T \beta), \tag{2}$$

and the accelerated failure time (AFT) model, whose hazard function is:

$$h(w_{ij}|z_i) = h_0(w_{ij} \exp(z_i^T \beta)) \exp(z_i^T \beta). \tag{3}$$

In each of (2) and (3), $h_0(\cdot)$ is the baseline hazard function. For the Weibull AFT model, the relationship between (2) and (3) can be seen through the interpretation of the parameters. For the proportional hazards model, the anti-log of β is the hazard ratio. For the Weibull AFT model, the anti-log of the multiplication of the corresponding $-\beta$ and the shape parameter of the Weibull distribution corresponds to the hazard ratio of the proportional hazards model. This property is referred to as the accelerated failure time property (Collett, 2015). In general, AFT models offer very different interpretations from proportional hazards models. AFT models assume that a covariate affects the rate or speed of a disease progression, whereas proportional hazards models assume that a covariates affects the hazard or risk multiplicatively. The key assumption for these models is that the gap times are independent given observed covariates within each patient.

The gap time analysis is not straightforward for left-truncated data. Left-truncated data implies that full information on the patient history of events prior to the study is not available. One needs to carefully determine how the initial conditions (i.e. the history of events prior to the study) affect the event process of interest and model accordingly. This issue is discussed in depth in Chapters 4 and 5 of Cook and Lawless (2007) with several ways of coping with it, and it is not addressed in this report.

3.3 Frailty models

To account for individual-specific effect or heterogeneity that are not captured by observed covariates, random effects can be introduced to Cox proportional hazards or AFT models. As a side note, in survival analysis, a random effect is often referred to as a frailty Collett (2015). It originates from the fact that an individual with a larger random effect has a larger risk (of death) and therefore is deemed to be more frail. A frailty model can be defined as:

$$\lambda_{ij}(w|v_i, z_i) = v_i h(w|z_i), \tag{4}$$

where $h(\cdot)$ can be (2) or (3) and v_i, which is the frailty for patient i, is assumed to be independent and identically distributed (Liu et al., 2004; Cook and Lawless, 2007). When v_i has a distribution belonging to a specific family of Gamma distributions, $\{\Gamma(1/\theta, 1/\theta), \theta > 0\}$, it leads to an analytical form of likelihood, for which standard optimization techniques, such as the Newton-Raphson method, can be used to obtain MLE estimates of the parameters (Rondeau et al., 2012). This frailty is referred to as the gamma frailty, and note that $E(v_i) = 1$ and $\text{var}(v_i) = \theta$. If a more flexible Gamma distribution, where the mean is no longer unity, one will run into the non-identifiablity problem Liu et al. (2004). Equivalent to (4), one can introduce the frailty as an additive term to the linear component of the model:

$$\lambda_{ij}(w|v_i, z_i) = h_0(w \exp(z_i^T \beta + v_i)) \exp(z_i^T \beta + v_i), \tag{5}$$

Package	**Distribution of** v_i	$h_0(\cdot)$	$h(\cdot)$
frailtypack	Gamma	Splines	(2)
parfm	Gamma	Weibull	(2)
NLMIXED	Normal or Gamma	Weibull	(5)

Table 6: The list of packages used in this report is provided along with the distributions used for the frailty term, the type of the baseline hazard function, and the type of the hazard model.

where v_i follows a log-Gamma or Normal distribution (Liu and Huang, 2008). Again, the key assumption is that the gap times are independent given the frailty and observed covariates. As a side note, including random effects does not necessarily resolve the issue with the initial conditions for gap time analysis.

As a result of recent methodology development in this field, one can use other distributions for v_i and model the baseline hazard either parametrically or nonparametrically; see the list of available R (Team, 2019) packages and their comparisons in Balan and Putter (2017). To fit a proportional hazards frailty model, two R packages are used for this report: `parfm` (Munda et al., 2012) and `frailtypack` (Rondeau et al., 2012). With `parfm`, a parametric baseline hazard function can be used with the gamma frailty. With `frailtypack`, a nonparametric baseline hazard function (based on penalized splines) can be used with the gamma frailty. To fit a Weibull AFT frailty model with the gamma or Normal frailty, where the Normal frailty is a Normal distribution with zero mean, `PROC NLMIXED` (McMahon et al., 2006) in SAS (Institute, 2017) can be used (Liu and Huang, 2008). The comparison of the three packages are summarized in Table 6. Further details on `frailtypack` and `PROC NLMIXED` are given in Subsections 3.4.1 and 3.4.2, respectively.

3.4 Joint/shared frailty models

The joint frailty model for recurrent events (RE) and a terminal event is characterized by two frailty models linked by a common frailty term (Liu et al., 2004):

$$\begin{aligned} \text{(RE)}: \ & \lambda_{ij}(w|v_i, z_{ij}) = v_i h_{re}(w|z_{ij}) \\ \text{(Terminal)}: \ & d_i(t|v_i, z_i) = v_i^{\alpha} h_d(t|z_i), \end{aligned} \tag{6}$$

where $\alpha \geq 0$ and v_i is the gamma frailty (or log-Normal$(0, \sigma^2)$). In this case, the dependence between the two processes is determined by (α, θ). Suppose that θ is null (otherwise, α is meaningless, and it suggests that there is no need for v_i. In other words, there is no unobserved covariate that accounts for heterogeneity not explained by observed covariates). If $\alpha = 0$, then it implies the two processes are independent given the covariates z_i. If $\alpha > 0$, then it implies that the two processes are dependent through unobserved covariates, whose effects are captured by v_i and α.

Maximum likelihood estimation of the joint frailty model is not straightforward. While using the gamma frailty results in a closed form of the likelihood for frailty models, this is not the case for the joint frailty model (Rondeau et al., 2007). To deal with this issue, Liu et al. (2004) proposed a Monte-Carlo (MC) Expectation-Maximization (EM) algorithm. The basic idea is to treat the frailty as missing data and apply the framework of the EM algorithm. There are two disadvantages with this method (Liu and Huang, 2008). First, it is too computationally intensive. Second, it is difficult to implement the MC-EM method. Two alternatives are considered in this report: Rondeau et al. (2007) and Liu and Huang (2008). Both of them rely on less computationally intensive methods at the cost of accuracy.

3.4.1 Rondeau et al. (2007)

Rondeau et al. (2007) apply penalized maximum likelihood estimation for (6), for which the baseline hazard functions, $h_0^{re}(\cdot)$ and $h_0^{re}(\cdot)$, are estimated nonparametrically using splines. To be exact, M-splines (Ramsay et al., 1988) of order 3 or 4 are used for a pre-specified number of knots. Rondeau et al. (2007) recommends to start with a small number of knots, say 6, and increase the number of knots until there is no substantial change in the shape of the baseline function. As splines can easily overfit the data, penalization is imposed on the roughness of the spline functions through the likelihood. The resulting penalized log-likelihood is defined as:

$$\sum_i^n l_i(h_0^{re}(\cdot), h_0^d(\cdot), \beta_{re}, \beta_d, \alpha, \theta | z_i, t, w) - \left(\kappa_1 \int_0^{\infty} ((h_0^{re})'')^2(t)dt + \kappa_2 \int_0^{\infty} ((h_0^d)'')^2(t)dt \right), \tag{7}$$

where $l_i(\cdot)$ is the individual log-likelihood, $\kappa_j \geq 0$ are positive smoothing parameters for the splines, and $\left(\kappa_1 \int_0^{\infty} ((h_0^{re})'')^2(t)dt + \kappa_2 \int_0^{\infty} ((h_0^d)'')^2(t)dt \right)$ is the penalty term. The parameter estimates that maximize (7)

are referred to as the penalized maximum likelihood estimates (PMLE). To approximate the log-likelihood term, which is the first term in (7), Gaussian quadrature (Weisstein, 2008) is used. To approximate the integrals in the penalty term, Languerre polynomials (Weisstein, 2002) is used. Finally, to optimize the parameters and obtain the PMLE, a Levenberg-Marquardt (LM) algorithm (Strutz, 2010) is used.

There could be an issue with the above method. The LM algorithm is a heuristic method in the sense that it does not guarantee an optimal solution. In a nutshell, it uses the gradient descent method to approach the neighborhood of the global maximum/minimum quickly and then applies the Gauss-Newton method to exactly pin the global maximum down. One pitfall is that it highly depends on the initial values of the parameters of interest. In addition, the algorithm can easily get stuck on a local maximum or saddle point if the surface of an objective function is not smooth.

Next, Rondeau et al. (2007) has implemented this method in R and made it available in `frailtypack`. The package is missing a critical component; there is no method to optimize the smoothing parameters,κ_j, which control the bias-variance trade-off. If the smoothing parameters are too big, then fitted splines (baseline hazard functions) are forced to be smooth and hence would have low variance but potentially have high bias (in other words, they do not fit the data well). On the other hand, if they are too low, then fitted splines can fit the data too closely ,resulting in low bias but potentially high variance. Currently, Rondeau et al. (2007) suggest to use the smoothing parameters that are optimized by cross-validation from a frailty model for each event process. These optimal smoothing estimates could be very different from the optimal values of the smoothing parameters of a joint frailty model.

With those two problems combined, one would need to implement a strategy to fit a joint frailty model if the suggestion provided in Rondeau et al. (2012) does not work. It could be a simple grid search or a pseudo-Bayesian search. However, as the package is mostly written in FORTRAN, a low-level scripting language like C, it is quite time intensive to examine the code and implement a modification. It poses a great practical barrier to applied scientists, who may not have necessary computing skills.

3.4.2 Liu and Huang (2008)

Under the framework of incorporating the frailty as the additive term, Liu and Huang (2008) apply Gaussian quadrature (Pinheiro and Bates, 1995) to estimate the parameters in the joint frailty model with piece-wise constant baseline hazard functions. To see how the Gaussian quadrature method can be used, first re-write the individual likelihood, $l_i(\cdot)$ as:

$$\int \prod_{j=1}^{n_i} l_{ij}(\cdot|v_i) f_\theta(v_i) dv_i, \tag{8}$$

where $l_{ij}(\cdot|v_i)$ is the conditional likelihood given the frailty and $f_\theta(\cdot)$ has a Normal distribution with zero mean. This is not exactly in the form, say (2.11) in Pinheiro and Bates (1995), for Gaussian quadrature. (8) can be re-written as:

$$\int \prod_{j=1}^{n_i} \exp(\tilde{l}_{ij}(\cdot|v_i)) f_\theta(v_i) dv_i, \tag{9}$$

where $\tilde{l}_{ij}(\cdot|v_i)) = \log(l_{ij}(\cdot|v_i))$. (9) can be approximated as a weighted sum of the integrand evaluated at Q number of predetermined quadrature abscissas $\{u_q\}_{q=1}^{Q}$ over the frailty v_i (Pinheiro and Bates, 1995). Explicitly, the full likelihood (9) is approximated as:

$$\sum_{q=}^{Q} \prod_{j=1}^{n_i} \exp(\tilde{l}_{ij}(\cdot|u_q)) f_\theta(u_q) w_q, \tag{10}$$

where u_q and w_q can be determined using an algorithm Pinheiro and Bates (1995). To use another distribution for the frailty, one can invoke the probability integral transformation (Nelson et al., 2006).

Liu and Huang (2008) use a commercially available library is SAS called `PROC NLMIXED` to estimate the parameters. `PROC NLMIXED` allows users to write down any likelihood. Moreover, the Gaussian quadrature method and its variants are built in it along with various optimization methods to find the MLE of the parameters. Hence, one can easily implement their proposed method using`PROC NLMIXED`. Sadatsafavi et al. (2018) implement their method in this way using the Weibull AFT model (i.e. Weibull baseline hazard functions with covariates) for both of the event processes. Their code is adopted for the analysis performed in this report.

3.4.3 Practicality

Using the method proposed by Liu and Huang (2008) by implementing it through `PROC NLMIXED` in SAS was easier and more reliable for analysis of the data for this report. Even though it took time to implement the method and took

more time to run the algorithm, it was easier to debug with help of excellent documentation and fit a joint frailty model. Moreover, as mentioned above, it is easy to add more features to frailty models and write down the corresponding likelihood.

In contrast, it was not even possible to fit the model using `frailtypack`. Because the package is mostly written in FORTRAN, it was not practically feasible to debug random crashing and uninformative error messages (which simply said "there is some problem with the likelihood"). As the package is maintained by the leading author of Rondeau et al. (2012), there is virtually no support available (I did contact the author for help and have not heard back after one reply that she would get back to me *soon*). The errors are likely due to insufficient data or due to bad initial values of the parameters and smoothing parameters.

4 Result

To cross-check the three packages/methods, frailty models were fit using them for each of the recurrent event process of severe COPD exacerbations and the terminal event of death. Estimations of the variance of the gamma frailty, $v_i \sim^d \Gamma(1/\theta, 1/\theta)$, and the hazard ratio of the parameters are compared. As discussed in Section 3.4.3, only `PROC NLMIXED` was used for fitting a joint frailty model. The results are interpreted with caution due to low statistical power.

4.1 Frailty models

To begin with, there is one advantage of using `frailtypack`; it is extremely fast regardless of whether it succeeds in fitting the model. With `frailtypack`, it only took less than 5 seconds to fit a frailty model for both of the recurrent event process and terminal event process with the default initial value (which is 0.1) for each parameter. With `parfm`, it took about 5 minutes to fit the same model with the default initial values. The syntax for `parfm` is almost identical to that for `frailtypack`. With `PROC NLMIXED`, it took more than 30 minutes, and some of the covariates had to be manually removed due to numerical convergence issues. For `PROC NLMIXED`, a strategy used in this report was to set a maximum running time (15 min), update the parameter values manually, and repeat until convergence.

The estimated parameter values from the three methods were converted to the corresponding hazard ratios (HR). As a reminder, for the AFT-Weibull model, the hazard ratio of a parameter can be calculated by invoking the accelerated failure time property (Section 3.2). The estimated HRs with 95% confidence interval (CI) for the three methods are summarized in Figure 7 for the recurrent event process and in Figure 8 for the terminal event process. The parameter names are listed on the y-axis, and the black dots are the point estimates of the HRs with the bars being the 95% CIs. The dashed vertical line is the value of 1. The estimates for the variance, θ, of the gamma frailty are provided in Tables 7 and 8 for the recurrent event and terminal event processes, respectively.

For the recurrent event process, the three models produce almost identical parameter estimates and CIs (Figure 7). Most of the parameter CIs contain the value of 1, and hence most of the parameters are insignificant. In particular, the estimated treatment HRs agree with the observation made in Fig. 4, where the Sal group had a higher rate of the COPD severe COPD exacerbations than the placebo group, and the Flu group had about the same rate as the placebo group. The HR estimates show that 1) there is no significant difference in the risk of getting a severe COPD exacerbation between the placebo group and the Flu group and that 2) the risk is marginally higher with the Sal group. It implies the drug treatments are ineffective or actually worse. However, this result could be due to the survival bias as mentioned in Section 2.1 or low statistical power. For the variance estimation, the three models produce similar estimates, which were significant (Table 4).

For the terminal event process, the three models produce similar estimates, but the CIs by `PROC NLMIXED` are generally smaller than the others. The treatments decrease the risk of death, as the CIs are tightly centered around 1. The risk of death strongly decreases with a healthier lung, denoted by a very low HR of the predicted lung health, fev1pp. The risk of death generally increases as patients experience more severe COPD exacerbations. As for the variance estimation, the three models all produce insignificant estimates of θ of the gamma frailty; the standard error is at least twice as big as the estimate for each model (Table 8). This implies the unobserved individual-level effects or characteristics are not significant given the observed covariates. Once again, this could be due to low statistical power. In other words, the frailty term might be significant with a richer dataset, such as an administrative dataset, in which the study period is longer and the sample size is much bigger.

4.2 Joint frailty models

Joint frailty models are used to capture potential dependence between the two processes. A priori, as the frailty term is estimated to be insignificant for the terminal event process of death, the dependence parameter, α, might be also

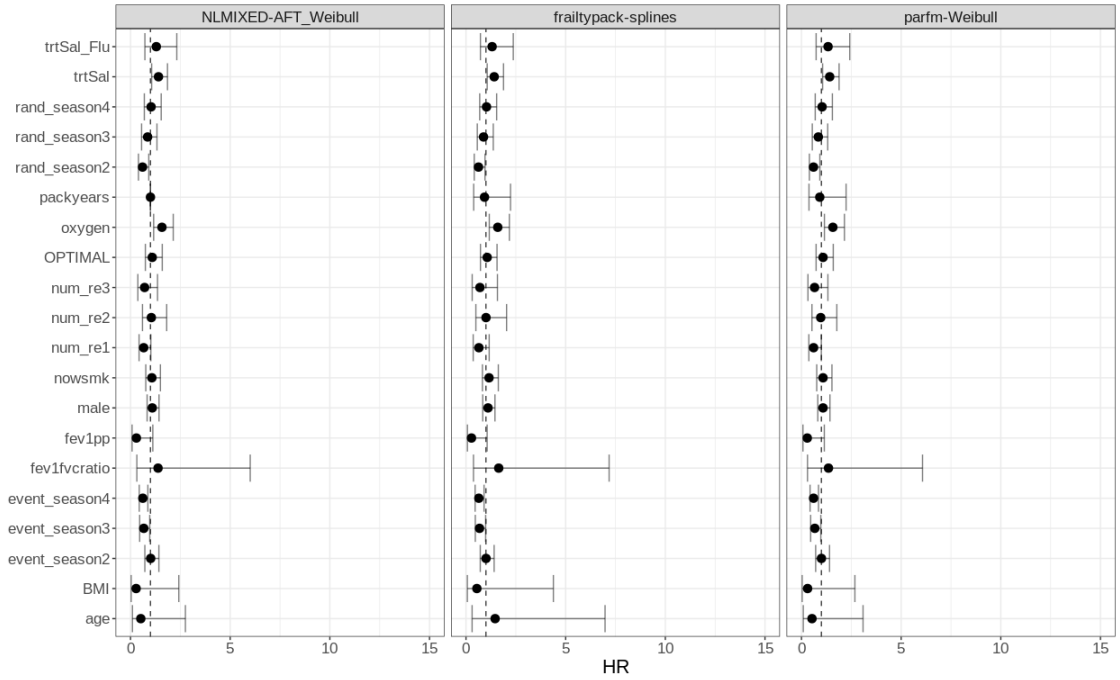

Figure 7: Estimated hazard ratios of the parameters by the three different methods for the the recurrent event process.

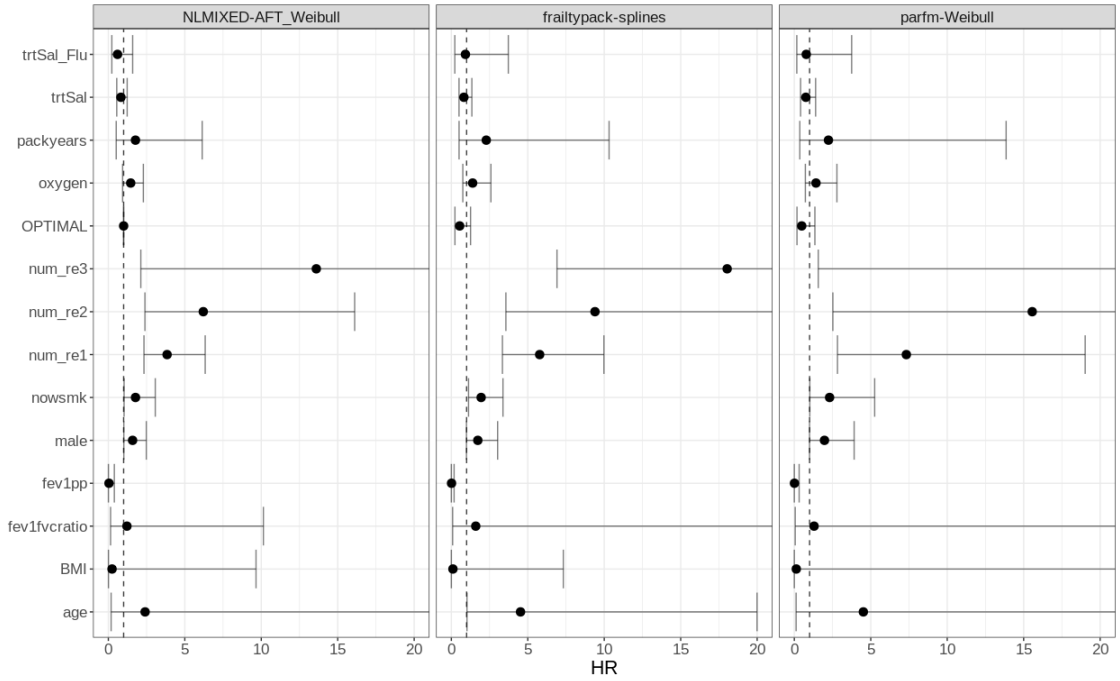

Figure 8: Estimated hazard ratios of the parameters by the three different methods for the terminal event process.

Model	Estimate	Standard error
NLMIXED-AFT_Weibull	2.60	0.55
frailtypack-splines	3.10	1.20
parfm-Weibull	3.28	0.95

Table 7: Estimated variance of the gamma frailty by the three different methods for the recurrent event process.

Model	Estimate	Standard error
NLMIXED-AFT_Weibull	1.16	2.41
frailtypack-splines	0.00	0.03
parfm-Weibull	2.96	4.59

Table 8: Estimated variance of the gamma frailty by the three methods of the terminal event process.

insignificant. The following results come from fitting a joint frailty model, for which the standard AFT-Weibull hazards model (where the scale parameter is set to 1) is used for both of the event processes with the Normal frailty using the method described in Section 3.4.2 (I ran into convergence issues using the gamma frailty).

As suspected, the standard error of α is about three times bigger than the estimate of α (Table 9), indicating that there is no strong evidence to support that the two processes are independent. Again, the large standard error might be due to low statistical. Given that the two processes are not significantly dependent, most of the HR estimates are similar to those of the non-joint ones (Figures 7 and 8). There are a couple of notable differences, which might be caused by using the Normal frailty instead of the gamma frailty. The HR estimate of num_re3, which is the indicator of having three or more severe COPD exacerbations, is now significantly associated with a decreased risk for the recurrent event process, whereas the HR estimate is much higher for the terminal event process.

While Sadatsafavi et al. (2018) found a positive, diminishing association between each severe COPD exacerbation and the risk of future exacerbations, no such pattern is found with the data. The first exacerbation was significantly associated with a decreased risk of 36% (HR=0.64, 95% 0.42-0.99), and the second exacerbation was weakly associated with an increased risk of 18% (HR=1.18, 95% 0.42-0.99), and the third or higher was significantly associated with a decreased risk of 50% (HR=0.50, 95% 0.29-0.89). Retrospectively, an interaction effect with the drug treatments and the number of severe COPD exacerbations should have been included to cope with the confounding effect.

As for the risk of death, only the first one was significantly associated with an increased risk of 420% (HR=5.22, 95% 2.91-9.40). The second and third and higher are insignificant but associated with an increased risk. To report this finding assertively, one needs to check the assumption that the patient history prior to the study is similar or well captured by the frailty term and observed covariates (so that the unobserved history that is not accounted for by the model does not affect the event process) in addition to other assumptions.

Parameter	Estimate	Standard error
α	0.02	0.06
σ^2	1.96	0.36

Table 9: Estimated dependence parameters of the joint frailty model.

5 Conclusion

Joint shared frailty models for left-truncated and right-censored data were described with motivation for joint modeling of a recurrent event process of severe COPD exacerbations and a terminal event process of death. Using the combined RCT data, two methods, `frailtypack` and `PROC NLMIXED` implemented in R and SAS, respectively, for fitting a joint frailty model were briefly explained and compared. There was no significant difference in the estimated parameters between these two methods for a frailty model for each of the two processes. However, the joint frailty model was only successfully fitted using `PROC NLMIXED`. It was found that there was insufficient data to support dependence of the two processes, and thus the HR estimates under joint modeling were similar to those under non-joint modeling. The

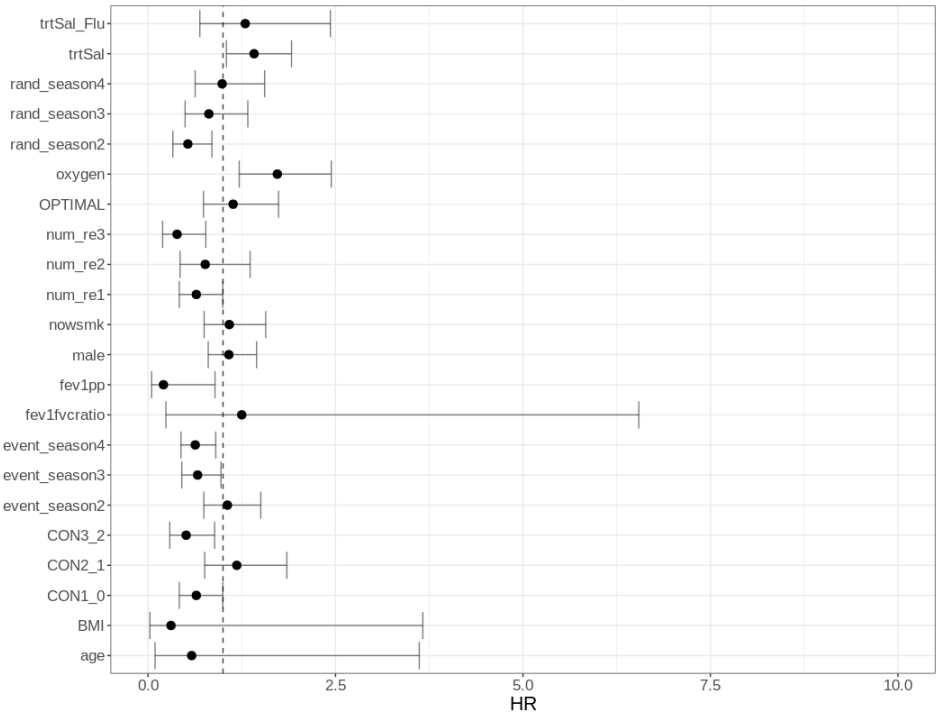

Figure 9: Estimated hazard ratios of the parameters of the joint frailty model of the recurrent event process.

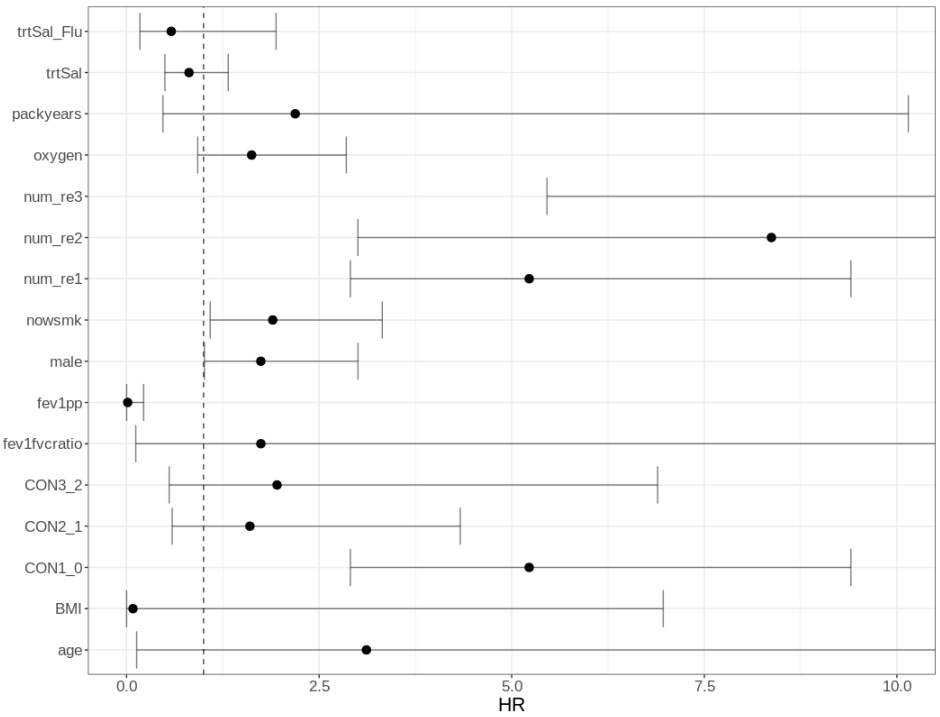

Figure 10: Estimated hazard ratios of the parameters of the joint frailty model of the death event process .

association between severe exacerbations was inconclusive due to a potential confounding effect with the drugs. One interesting finding was that the treatments were not effective for reducing the risk for severe COPD exacerbations when the papers related to the trials show that they are effective for reducing the risk for mild and severe COPD exacerbations; however, this could be due to low statistical power.

Several important model diagnostics and validation of model assumptions are not performed in this report. First, the assumption of individual-level conditional independence of gap times was not checked. It appears there is no method or way to check or validate this assumption. Second, potential issues with the initial conditions for gap times were not explored. For the asthma administrative data, this problem can be avoided by defining an incidence cohort (i.e. people who have no hospitalization or drugs related to asthma for five years prior to their first asthma exacerbations reported in the database) given that the resulting sample size is sufficient. Third, the assumptions for either Cox proportional hazards or AFT models were not checked. These can be checked by existing diagnostic plots based on residuals. Fourth, other distributions for parametric hazards functions and for frailty terms were not considered. Different distributions can be used and compared based on some goodness-of-fit criteria, such as AIC. Fifth, the assumption on the distribution of the frailty term was not checked. To check this assumption, a histogram of the predicted random effects can be drawn and compared with the fitted random effects distribution.

References

Aaron, S. D., K. L. Vandemheen, D. Fergusson, F. Maltais, J. Bourbeau, R. Goldstein, M. Balter, D. O'Donnell, A. McIvor, and S. Sharma (2007). Tiotropium in combination with placebo, salmeterol, or fluticasone–salmeterol for treatment of chronic obstructive pulmonary disease: a randomized trial. *Annals of internal medicine 146*(8), 545–555.

Adibi, A., D. D. Sin, A. Safari, K. M. Johnson, S. Aaron, J. M. FitzGerald, and M. Sadatsafavi (2019). Development and external validation of the acute copd exacerbation prediction tool (accept). *bioRxiv*, 651901.

Albert, R. K., J. Connett, W. C. Bailey, R. Casaburi, J. A. D. Cooper, G. J. Criner, J. L. Curtis, M. T. Dransfield, M. K. Han, S. C. Lazarus, B. Make, N. Marchetti, F. J. Martinez, N. E. Madinger, C. McEvoy, D. E. Niewoehner, J. Porsasz, C. S. Price, J. Reilly, P. D. Scanlon, F. C. Sciurba, S. M. Scharf, G. R. Washko, P. G. Woodruff, and N. R. Anthonisen (2011, August). Azithromycin for Prevention of Exacerbations of COPD. *The New England journal of medicine 365*(8), 689–698.

Balan, T. A. and H. Putter (2017). frailtyem: An r package for estimating semiparametric shared frailty models.

Benady, S. (2010). *The human and economic burden of COPD: a leading cause of hospital admission in Canada*. Canadian Thoracic Society.

Collett, D. (2015). *Modelling survival data in medical research*. Chapman and Hall/CRC.

Cook, R. J. and J. Lawless (2007). *The statistical analysis of recurrent events*. Springer Science & Business Media.

Institute, S. (2017). *Base SAS 9.4 procedures guide: Statistical procedures*. SAS Institute.

Ivanova, J. I., R. Bergman, H. G. Birnbaum, G. L. Colice, R. A. Silverman, and K. McLaurin (2012). Effect of asthma exacerbations on health care costs among asthmatic patients with moderate and severe persistent asthma. *Journal of Allergy and Clinical Immunology 129*(5), 1229–1235.

Liu, L. and X. Huang (2008). The use of Gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in medicine 27*(14), 2665–2683.

Liu, L., R. A. Wolfe, and X. Huang (2004). Shared Frailty Models for Recurrent Events and a Terminal Event. *Biometrics 60*(3), 747–756.

McMahon, J. M., E. R. Pouget, and S. Tortu (2006). A guide for multilevel modeling of dyadic data with binary outcomes using sas proc nlmixed. *Computational statistics & data analysis 50*(12), 3663–3680.

Munda, M., F. Rotolo, C. Legrand, et al. (2012). Parfm: parametric frailty models in r. *Journal of Statistical Software 51*(11), 1–20.

Nelson, K. P., S. R. Lipsitz, G. M. Fitzmaurice, J. Ibrahim, M. Parzen, and R. Strawderman (2006). Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. *Journal of Computational and Graphical Statistics 15*(1), 39–57.

Pinheiro, J. C. and D. M. Bates (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of computational and Graphical Statistics 4*(1), 12–35.

Ramsay, J. O. et al. (1988). Monotone regression splines in action. *Statistical science 3*(4), 425–441.

Rondeau, V., S. Mathoulin-Pelissier, H. Jacqmin-Gadda, V. Brouste, and P. Soubeyran (2007, January). Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events. *Biostatistics 8*(4), 708–721.

Rondeau, V., Y. Mazroui, and J. R. Gonzalez (2012). frailtypack: an R package for the analysis of correlated survival data with frailty models using penalized likelihood estimation or parametrical estimation. *J Stat Softw 47*(4), 1–28.

Sadatsafavi, M., H. Xie, M. Etminan, K. Johnson, J. M. FitzGerald, and Canadian Respiratory Research Network (2018). The association between previous and future severe exacerbations of chronic obstructive pulmonary disease: Updating the literature using robust statistical methodology. *PloS One 13*(1), e0191243.

Strutz, T. (2010). *Data fitting and uncertainty: A practical introduction to weighted least squares and beyond*. Vieweg and Teubner.

Suissa, S., S. Dell'Aniello, and P. Ernst (2012, November). Long-term natural history of chronic obstructive pulmonary disease: severe exacerbations and mortality. *Thorax 67*(11), 957–963.

Team, R. C. (2019). R: A language and environment for statistical computing, version 3.3. 1. vienna, austria: R foundation for statistical computing; 2016.

Weisstein, E. W. (2002). Polynomial.

Weisstein, E. W. (2008). Legendre-gauss quadrature. *From MathWorld–A Wolfram Web Resource, http://mathworld. wolfram. com/Legendre-GaussQuadrature. html, Accessed 30*.