

COMPETING RISK ANALYSIS BAGI PASIEN PNEUMONIA DI SUATU RUMAH SAKIT

B. N. A. Wiyoto¹, *Ruhayat², dan H. Sumarno³

¹Mahasiswa Program Studi S1 Aktuaria. Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor, Jl. Meranti, Kampus IPB Dramaga Bogor
Benediktus_diki@apps.ipb.ac.id

^{2,3}Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor, Jl. Meranti, Kampus IPB Dramaga Bogor.
ruhiyat-mat@apps.ipb.ac.id*, hadisumarno@apps.ipb.ac.id *corresponding author

Abstrak

Analisis *survival* merupakan salah satu bagian dalam ilmu Statistika yang berkaitan dengan kelangsungan hidup suatu entitas. Analisis *survival* memiliki penerapan yang luas dalam bidang biostatistika dan membantu banyak praktisi kesehatan dalam memajukan berbagai metode kesehatan. Pada penelitian ini dibahas salah satu topik penting dalam analisis *survival*, yaitu situasi *competing risks*. *Competing risks* adalah suatu kondisi di mana terdapat lebih dari satu kejadian yang mungkin terjadi pada seorang individu. Penelitian ini berfokus pada ilustrasi bahwa penggunaan penduga Kaplan-Meier dalam situasi *competing risks* tidaklah cukup. Pendekatan dengan menggunakan nilai *cumulative incidence function* (CIF) dari setiap kejadian adalah salah satu solusi dalam situasi *competing risks*, baik secara nonparametrik maupun secara parametrik. Berdasarkan hasil uji hipotesis, situasi *competing risks* benar terjadi pada pasien pneumonia dalam penelitian ini.

Kata kunci: Analisis *survival*, *competing risks*, *cumulative incidence function*, penduga Kaplan-Meier, uji hipotesis

1 Pendahuluan

Analisis *survival* merupakan salah satu cabang ilmu Statistika yang berkaitan dengan kelangsungan hidup suatu entitas yang dapat berupa manusia maupun berupa benda tertentu atau ilmu yang berfokus pada pemodelan waktu hingga kegagalan dalam konteks tertentu. Model *survival* merupakan sebaran peluang untuk suatu peubah acak waktu hingga kegagalan tertentu. Model *survival* memiliki banyak penerapan dalam kehidupan sehari-hari seperti pada bidang asuransi, medis, dan teknik yang terkait dengan suatu peubah acak waktu hingga kegagalan tertentu [1]. Khusus pada bidang asuransi, model *survival* digunakan untuk membangun tabel mortalitas dan selanjutnya dapat diterapkan oleh para aktuaris untuk menentukan tingkat premi asuransi dan memastikan kecukupan cadangan [2]. Dalam penerapannya, model *survival* sering dinyatakan dalam bentuk fungsi *survival*.

Salah satu metode pada analisis *survival* yang sering digunakan adalah metode Kaplan-Meier. Metode ini dapat memberikan dugaan bagi fungsi *survival* dari suatu

populasi tanpa asumsi sebaran tertentu dan banyak digunakan oleh para akademisi. Namun, penduga Kaplan-Meier memiliki kekurangan apabila dikaitkan dengan peluang kematian di kehidupan nyata. Hal ini diakibatkan sebagai adanya situasi *competing risks* dalam suatu populasi [3].

Competing risks adalah suatu kondisi di mana seorang individu dalam populasi dapat mengalami berbagai macam kejadian dan kemunculan kejadian yang satu dapat mencegah munculnya kejadian yang lain [4]. Penduga Kaplan-Meier cenderung menghasilkan dugaan yang nilainya terlalu tinggi dari yang sebenarnya (*overestimate*). Hal ini dikarenakan metode *censoring* dalam metode Kaplan-Meier mengabaikan semua kejadian penyebab kematian selain dari kejadian yang ingin diketahui [5].

Pada penelitian ini dilakukan pendugaan fungsi *survival* dengan metode yang memperhitungkan kondisi *competing risks*. Metode regresi Fine dan Gray atau yang sering disebut juga *competing risks regression* memodifikasi model *Cox proportional hazard* dengan mendefinisikan *cumulative incidence function* untuk menduga peluang kematian yang kemudian dapat diinterpretasikan sebagai banyaknya individu yang meninggal secara aktual [6]. Dengan begitu, masalah pendugaan yang terlalu tinggi pada penduga Kaplan-Meier dapat diperbaiki dan dapat lebih menggambarkan kejadian yang sebenarnya.

Tujuan dari penelitian ini adalah melakukan analisis dampak penyakit Pneumonia pada kematian pasien di unit perawatan intensif dengan metode nonparametrik, mengestimasi *cumulative incidence function* dari setiap jenis kejadian dengan metode nonparametrik dan parametrik, serta menunjukkan hasil dugaan fungsi kumulatif yang terlalu tinggi dari penduga Kaplan-Meier dalam situasi *competing risks*.

2 Metode

Data pada penelitian ini diambil dari *package mvna* pada aplikasi R dengan nama data *sir.adm* yang berisi catatan 747 pasien di unit perawatan intensif, Rumah Sakit Charité University, Berlin, Jerman. Data tersebut pada mulanya berawal dari penelitian para ahli epidemiologi terkait dengan *nosocomial infection*, yaitu penyakit berinfeksi yang menyerang pasien ketika pasien tersebut menerima perawatan di unit kesehatan [7]. Artinya, mereka belum mengidap penyakit tersebut pada saat penerimaan di unit kesehatan. Daftar peubah yang ada dalam data tersebut disajikan pada Tabel 1. Pada penelitian ini peubah-peubah yang digunakan hanyalah *pneu*, *status*, dan *time*.

Tabel 1. Daftar peubah dalam data *sir.adm*

Peubah	Deskripsi
<i>id</i>	Nomor identitas pasien.
<i>pneu</i>	Status penyakit Pneumonia pada pasien: 0 untuk pasien yang tidak mengidap Pneumonia dan 1 untuk pasien yang mengidap Pneumonia.
<i>status</i>	Status akhir pasien saat observasi: 0 untuk pasien tersensor, 1 untuk pasien yang pulang (<i>discharged</i>), dan 2 untuk pasien yang meninggal.
<i>time</i>	Waktu pasien berada di unit perawatan intensif.
<i>age</i>	Usia pasien saat memasuki periode pengamatan.
<i>sex</i>	Jenis kelamin pasien: F untuk wanita dan M untuk pria.

Tahapan analisis data pada penelitian ini adalah sebagai berikut:

1. menampilkan statistika deskriptif dari data,
2. melakukan pendugaan *cumulative incidence function* dengan metode nonparametrik dari setiap kejadian beserta ilustrasinya berdampingan dengan penduga fungsi sebaran kumulatif Kaplan-Meier,
3. melakukan pendugaan *cumulative incidence function* dengan metode *competing risks regression* beserta ilustrasinya, dan
4. melakukan pengujian hipotesis untuk menguji apakah situasi *competing risks* benar terjadi.

2.1 Kovariat dan Data Tersensor

Sama seperti halnya sebuah peubah bebas, sebuah kovariat melengkapi peubah takbebas atau peubah respon. Suatu peubah dikatakan sebagai sebuah kovariat jika peubah tersebut berhubungan dengan peubah takbebas. Menurut definisi ini, setiap peubah yang dapat diukur dan dianggap memiliki hubungan statistik dengan peubah takbebas akan memenuhi syarat sebagai sebuah kovariat yang potensial. Dengan demikian, sebuah kovariat merupakan suatu peubah prediktif atau suatu peubah penjelas yang mungkin dari peubah takbebas [8].

Data tersensor terjadi ketika waktu saat kejadian tidak teramati untuk sekelompok individu. Ada dua jenis data tersensor, yaitu data tersensor kanan (*right censored*) yang berarti bahwa kejadian yang hendak diamati (*event of interest*) tidak terjadi selama periode pengamatan dan data tersensor kiri (*left censored*) yang berarti bahwa waktu terjadinya suatu kejadian diketahui kurang dari atau sama dengan suatu titik waktu tertentu. Kemunculan kejadian lainnya dalam *competing risks* juga dapat menyebabkan penyensoran. Hal tersebut bergantung pada tujuan studi. Misalkan suatu studi bertujuan untuk menduga waktu kematian karena serangan jantung, maka kematian karena sebab lain dapat mengalami penyensoran [3].

2.2 Competing Risks

Situasi *competing risks* muncul ketika seorang individu dapat mengalami lebih dari satu kejadian, di mana satu kejadian dapat mencegah terjadinya kejadian yang lain [3]. Istilah *competing risks* pertama kali muncul di bidang Onkologi, yaitu cabang ilmu kedokteran yang berkaitan dengan deteksi dan penanganan penyakit kanker. Karena penanganan kanker dapat memperpanjang waktu bertahan hidup, maka penting untuk mempertimbangkan tidak hanya penanganan untuk bertahan hidup dari kanker, namun juga *competing risks*, seperti kematian karena penyebab lain yang mungkin memengaruhi keputusan penanganan [6].

2.3 Cumulative Incidence Function

Cumulative incidence function (CIF) untuk kejadian tipe ke- i ($i = 1, 2, \dots, p$) didefinisikan sebagai peluang bersama [3] yang dirumuskan oleh Persamaan (1) berikut:

$$F_i(t) = \Pr(T \leq t, C = i). \quad (1)$$

Dengan kata lain, CIF adalah peluang terjadinya kejadian tipe ke- i pada atau sebelum waktu t . Penjumlahan seluruh CIF dari semua tipe kejadian sama dengan fungsi sebaran kumulatif dari terjadinya kejadian tipe manapun. Dengan demikian, diperoleh Persamaan (2) berikut:

$$F(t) = \Pr(T \leq t) = \sum_{i=1}^p \Pr(T \leq t, C = i) = \sum_{i=1}^p F_i(t). \quad (2)$$

2.4 Penduga Kaplan-Meier

Misalkan t_1, t_2, \dots, t_r adalah barisan waktu kegagalan sedemikian sehingga $t_1 < t_2 < \dots < t_r$. Misalkan pula d_j menyatakan banyaknya kejadian yang terjadi pada waktu t_j dan n_j menyatakan banyaknya individu yang terpapar risiko sesaat sebelum t_j . [9] mendefinisikan penduga Kaplan-Meier atau *product-limit estimator* untuk fungsi *survival* seperti pada Persamaan (3) berikut:

$$\hat{S}_{KM}(t) = \prod_{t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right). \quad (3)$$

2.5 Pendugaan *Cumulative Incidence Function*

Berikut disajikan pendugaan CIF untuk data tersensor atau data taklengkap. Misalkan $t_1 < t_2 < \dots < t_r$ adalah titik-titik waktu tak tersensor yang unik dan berurutan. Misalkan pula $\hat{S}(t)$ adalah penduga fungsi *survival* seperti halnya penduga Kaplan-Meier tetapi meliputi semua penyebab kejadian. Secara intuitif, penduga CIF merupakan peluang bersama di mana individu bertahan hingga waktu t_{j-1} dengan peluang $\hat{S}(t_{j-1})$ dan peluang individu mengalami kejadian tipe ke- i pada waktu t_j dengan peluang $\frac{d_{ij}}{n_j}$. Definisi secara umum diberikan oleh Persamaan (4) berikut:

$$\hat{F}_i(t) = \sum_{\forall j, t_j \leq t} \frac{d_{ij}}{n_j} \hat{S}(t_{j-1}), \quad (4)$$

dengan $\hat{S}(t_{j-1}) = \prod_{t_j \leq t} \left(\frac{n_j - d_j^*}{n_j} \right)$ di mana d_j^* adalah banyaknya semua kejadian yang terjadi pada titik waktu t_j [3].

Penduga fungsi sebaran kumulatif Kaplan-Meier selalu lebih besar atau sama dengan daripada penduga CIF untuk setiap titik waktu t_k . Secara matematis, hal tersebut disajikan oleh Persamaan (5) berikut:

$$1 - \hat{S}_{KM}(t_k) \geq \hat{F}_i(t). \quad (5)$$

Jika tidak ada situasi *competing risks*, maka kedua nilai tersebut sama [3].

2.6 Fungsi *Partial Likelihood*

Diberikan suatu peubah acak T dan parameter θ dengan $\theta = \{\varphi, \lambda\}$ yang tidak diketahui. Misalkan terdapat peubah acak T_1, T_2, \dots, T_n yang bebas stokastik identik. Untuk suatu model yang kompleks, fungsi *likelihood* dari model tersebut dapat dituliskan secara terpisah seperti pada Persamaan (6) berikut:

$$L(\theta; T) = L_1(\varphi; T) L_2(\varphi, \lambda; T) \quad (6)$$

dengan $L_1(\varphi; T)$ dan $L_2(\varphi, \lambda; T)$ merupakan fungsi-fungsi *partial likelihood* dan λ adalah *nuisance parameter*, yaitu parameter yang bukan merupakan fokus dari analisis namun harus dimasukkan untuk menganalisis parameter yang menjadi perhatian utama, yaitu φ [10].

2.7 Himpunan Risiko

Himpunan risiko merupakan semua subjek yang memiliki risiko untuk mengalami suatu kejadian [11]. Pada awal studi, himpunan risiko terdiri atas semua subjek. Sepanjang berjalannya studi, subjek dapat mengalami kejadian yang hendak diamati atau menjadi amatan tersensor sehingga menyebabkan subjek tersebut keluar dari himpunan risiko [12].

2.8 Uji Hipotesis *k-sample* Gray

Uji hipotesis *k-sample* Gray merupakan suatu uji hipotesis yang dapat digunakan untuk mengetahui apakah beberapa CIF dari kejadian tertentu cukup berbeda secara statistik dan mencerminkan kondisi *competing risks*. Misalkan CIF untuk suatu sebab j dari grup k pada waktu ke- t adalah $F_{jk}(t)$, maka hipotesis dari uji ini untuk suatu sebab (dinotasikan dengan 1) adalah sebagai berikut:

$H_0 : F_{11}(t) = \dots = F_{1K}(t) = F_1^0(t)$ untuk semua $t \leq \tau$.

H_a : Setidaknya ada satu $F_{1K}(t)$ yang berbeda untuk beberapa $t \leq \tau$.

$F_1^0(t)$ adalah fungsi subdistribusi yang tak ditentukan, τ waktu pengamatan terlama dalam studi dan K adalah peubah acak banyaknya grup dalam pengujian. Ada dua langkah yang harus dilakukan untuk mendapatkan statistik uji dari pengujian ini, yaitu menghitung *score* pengujian sebagai *numerator* dan mengestimasi ragam dari CIF sebagai *denominator*. Bentuk umum dari *score* untuk grup k diberikan oleh Persamaan (7) berikut:

$$z_k = \int_0^{\tau} W_k(t) (\gamma_k(t) - \gamma_0(t)) dt \quad (7)$$

dengan $W_k(t)$ adalah fungsi bobot, $\gamma_k(t)$ adalah fungsi *hazard* dari subdistribusi untuk grup k , dan $\gamma_0(t)$ adalah fungsi *hazard* dari subdistribusi untuk semua grup. Bentuk umum dari fungsi bobot adalah $W_k(t) = L(t)R_k(t)$. Menurut [13], untuk alasan kesederhanaan dan kejelasan, dipilih $L(t) = 1$ dan $R_k(t)$ seperti pada Persamaan (8) berikut:

$$R_k(t) = n_k(t) \frac{1 - \hat{F}_k(t-)}{\hat{S}_k(t-)} \quad (8)$$

dengan $n_k(t)$ adalah banyaknya individu yang terpapar risiko untuk grup k pada waktu ke- t , $F(t-)$ adalah penduga *left-hand limit* dari CIF kejadian yang utama (*event of interest*), dan $S(t-)$ adalah *left-hand limit* dari peluang bertahan dari kejadian apapun yang diduga dengan metode Kaplan-Meier. Dengan demikian, $R_k(t)$ dapat dipahami sebagai himpunan risiko yang disesuaikan.

Misalkan t_1, t_2, \dots, t_n adalah titik-titik waktu terurut ketika kejadian apapun dapat terjadi di grup manapun, d_{1j} adalah banyaknya kejadian utama di grup 1 pada waktu t_j ,

n_{1j} adalah banyaknya individu yang terpapar risiko di grup 1 pada waktu t_j dan ada $k = 2$ samples, maka berlaku Persamaan (9) berikut:

$$R_{1j} = n_{1j}(t) \frac{1 - \hat{F}_1(t_{j-1})}{\hat{S}_1(t_{j-1})} \quad (9)$$

dan *score* untuk pengujian kejadian ini bagi kasus diskret [13] diberikan oleh Persamaan (10) berikut:

$$Z_1 = \sum_{\text{semua } t_j} R_{1j} \left(\frac{d_{1j}}{R_{1j}} - \frac{d_{1j} + d_{2j}}{R_{1j} + R_{2j}} \right). \quad (10)$$

Langkah selanjutnya adalah menduga ragam dari CIF sebagai *denominator* dari statistik uji. Pada penelitian ini, digunakan pendekatan dengan sebaran multinomial. Misalkan untuk setiap waktu t ada $\bar{Y}(t)$ individu yang terpapar risiko untuk setiap kejadian J , dan $\bar{N}_j(t)$ individu yang telah mengalami kejadian j . Dengan demikian, $\bar{N}(t) = \{\bar{N}_1(t), \bar{N}_2(t), \dots, \bar{N}_j(t)\}$ menyebar multinomial dengan parameter $\bar{Y}(t)$ dan $h(t) = \{h_1(t), h_2(t), \dots, h_j(t)\}$ sebagai fungsi *hazard* subdistribusi dari setiap kejadian.

Misalkan ada N waktu hingga kejadian $\{T_1, T_2, \dots, T_N\}$ dan waktu-waktu tersebut diurutkan untuk $M \leq N$ menjadi $\{Z_1, Z_2, \dots, Z_M\}$, maka diperoleh Persamaan (11) berikut:

$$\hat{F}_j(t) = \sum_{Z_i \leq t} \hat{h}_j(Z_i) \left\{ \prod_{Z_i < t} \left[1 - \sum_j \hat{h}_j(Z_i) \right] \right\} = \sum_{Z_i \leq t} \hat{h}_j(Z_i) \hat{S}(Z_i) \quad (11)$$

sedemikian sehingga berlaku Persamaan (12) berikut:

$$\begin{aligned} \text{Var}\{\hat{F}_j(t)\} &= \sum_{Z_i \leq t} \text{Var}\{\hat{h}_j(Z_i) \hat{S}(Z_i)\} \\ &+ 2 \sum_{Z_i \leq t} \sum_{Z_i < Z_i' \leq t} \text{Cov}\{\hat{h}_j(Z_i) \hat{S}(Z_i), \hat{h}_j(Z_i') \hat{S}(Z_i')\}. \end{aligned} \quad (12)$$

Berdasarkan sifat sebaran multinomial, diperoleh Persamaan (13), (14), dan (15) berikut:

$$E[\hat{h}_j(t)] = h_j(t) \quad (13)$$

$$\text{Var}[\hat{h}_j(t)] = \frac{h_j(t)[1 - h_j(t)]}{\bar{Y}(t)} \quad (14)$$

$$\text{Cov}[\hat{h}_j(t), \hat{h}_k(t)] = -\frac{h_j(t)h_k(t)}{\bar{Y}(t)}, j \neq k \quad (15)$$

sehingga $\hat{F}_j(t)$ dapat diduga ketika $\hat{h}(t)$ telah diduga [14]. Statistik uji untuk uji hipotesis *k-sample* Gray pada grup 1 dapat didefinisikan dengan Persamaan (16) berikut:

$$T = \frac{Z_1}{\text{Var}\{\hat{F}_j(t)\}} \quad (16)$$

2.9 Model Cox Proportional Hazard

Misalkan *hazard function* untuk individu i dengan himpunan kovariat $x_{1i}, x_{2i}, \dots, x_{mi}$ dapat dibagi ke dalam dua komponen. Komponen pertama melibatkan waktu namun tidak melibatkan kovariat dan komponen kedua melibatkan kovariat namun tidak melibatkan waktu, sehingga diperoleh Persamaan (17) berikut:

$$h_i(t) = \exp\{\beta_1 x_{1i} + \dots + \beta_m x_{mi}\} h_0(t) \quad (17)$$

dengan $h_0(t)$ adalah *baseline hazard function* dan $\beta_1, \beta_2, \dots, \beta_m$ adalah beberapa koefisien yang perlu diduga. Untuk titik-titik waktu $t_1 < t_2 < \dots < t_r$ yang terurut dan unik, nilai dugaan koefisien $\beta_1, \beta_2, \dots, \beta_m$ dapat diperoleh dengan memaksimumkan fungsi *partial likelihood* yang diberikan oleh Persamaan (18) berikut:

$$L(\beta_1, \beta_2, \dots, \beta_m) = \prod_{j=1}^r \left(\frac{\exp\{\beta_1 x_{1j} + \dots + \beta_m x_{mj}\}}{\sum_{i \in R_j} \exp\{\beta_1 x_{1i} + \dots + \beta_m x_{mi}\}} \right) \quad (18)$$

di mana perkalian diambil untuk tiap titik waktu hingga t_j dan R_j merupakan himpunan individu yang terpapar risiko hingga usia t_j [15].

2.10 Competing Risks Regression (CRR)

Salah satu metode untuk mendapatkan nilai dugaan dari CIF dikemukakan oleh [16]. Pada awalnya, metode ini digunakan untuk data lengkap. Namun, dengan sedikit modifikasi, metode ini dapat pula diterapkan untuk data taklengkap. Model tersebut didasarkan pada Persamaan (19) berikut:

$$\gamma(t, \mathbf{x}) = \gamma_0(t) \exp\{\boldsymbol{\beta}^T \mathbf{x}\} \quad (19)$$

dengan γ adalah fungsi *hazard* dari subdistribusi, γ_0 adalah *baseline hazard function* dari subdistribusi, \mathbf{x} adalah vektor kovariat, dan $\boldsymbol{\beta}$ adalah vektor koefisien.

Bentuk *partial likelihood function* yang digunakan untuk menduga vektor $\boldsymbol{\beta}$ pada metode ini identik dengan model *Cox proportional hazard*. Apabila dituliskan untuk satu kovariat, maka *partial likelihood function* dari metode ini adalah sebagai berikut:

$$L(\boldsymbol{\beta}) = \prod_{j=1}^r \left(\frac{\exp\{\boldsymbol{\beta} x_j\}}{\sum_{i \in R_j} w_{ji} \exp\{\boldsymbol{\beta} x_j\}} \right) \quad (20)$$

di mana perbedaannya terdapat pada penambahan bobot w_{ji} dan himpunan risiko R_j yang dibentuk oleh himpunan individu yang belum pernah mengalami kejadian apapun hingga waktu t dan yang sudah mengalami kejadian *competing risks* pada waktu t . Dengan kata lain, individu yang telah mengalami kejadian lain tetap berada di himpunan risiko sepanjang waktu. Bobot w_{ji} didefinisikan dengan Persamaan (21) berikut:

$$w_{ji} = \frac{\hat{S}_{KM}(t_j)}{\hat{S}_{KM}(\min(t_j, t_i))} \quad (21)$$

dengan \hat{S}_{KM} merupakan dugaan fungsi *survival* Kaplan-Meier pada sebaran yang tersensor. Dengan menurunkan fungsi log *partial likelihood* pada Persamaan (20), diperoleh fungsi statistik berikut:

$$U(\beta) = \sum_{j=1}^r x_j - \left\{ \frac{\sum_{i \in R_j} w_{ji} x_i \exp(x_i \beta)}{\sum_{i \in R_j} w_{ji} \exp(x_i \beta)} \right\}. \quad (22)$$

Nilai dugaan β dilambangkan dengan $\hat{\beta}$ dan dapat diperoleh dengan memaksimumkan fungsi $U(\beta)$ pada Persamaan (22). Selanjutnya, untuk memperoleh nilai CIF dari tiap kejadian, fungsi kumulatif *hazard* diduga terlebih dahulu dengan bantuan penduga *Breslow-type* seperti pada Persamaan (23) berikut:

$$\hat{\Lambda}(t; x_0, \hat{\beta}) = \sum_{t_j \leq t} \left\{ \frac{\exp\{x_0 \hat{\beta}\}}{\sum_{i \in R_j} w_{ji} \exp\{x_i \hat{\beta}\}} \right\}. \quad (23)$$

Dengan begitu, nilai CIF dari tiap kejadian dapat diduga dengan menggunakan Persamaan (24) berikut [16]:

$$\hat{F}(t; x_0, \hat{\beta}) = 1 - \exp\{-\hat{\Lambda}(t; x_0, \hat{\beta})\}. \quad (24)$$

3 Hasil dan Pembahasan

3.1 Statistika Deskriptif Data

Pada penelitian ini digunakan peubah-peubah *pneu*, *status*, dan *time*. Untuk kejelasan mengenai status penyakit pasien, peubah *pneu* diubah namanya menjadi *dis* yang diambil dari kata *disease*. Ringkasan dari ketiga peubah tersebut dapat dilihat pada Tabel 2. Informasi mengenai pasien untuk setiap status penyakit pada setiap kejadian dapat dilihat pada Tabel 3.

Tabel 2. Peubah-peubah dalam data *sir.adm* yang digunakan dalam penelitian

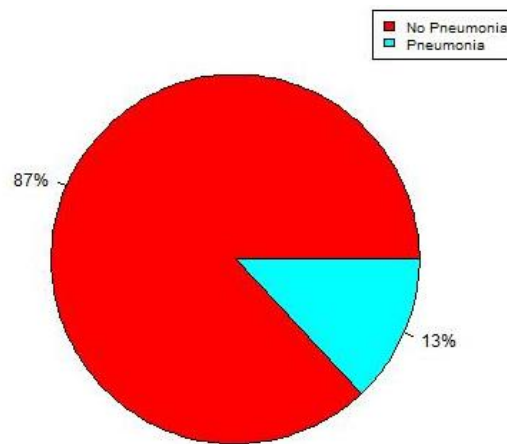
Peubah	Ringkasan
<i>dis</i>	Pneumonia: 97 pasien
	Tidak Pneumonia: 650 pasien
<i>status</i>	<i>Censored</i> : 14 pasien
	<i>Discharged</i> : 657 pasien
	<i>Dead</i> : 76 pasien
<i>time</i>	Minimum: 2 hari
	Kuartil ke-1: 5 hari
	Kuartil ke-2 (Median): 9 hari
	Rata-rata: 14.52 hari
	Kuartil ke-3: 18 hari
	Maksimum: 183 hari

Tabel 3. Informasi pasien untuk setiap status penyakit pada setiap kejadian

Status penyakit	<i>Censored</i>	<i>Discharged</i>	<i>Dead</i>	Total
Pneumonia	8	68	21	97
Tidak Pneumonia	6	589	55	650
Total	14	657	76	747

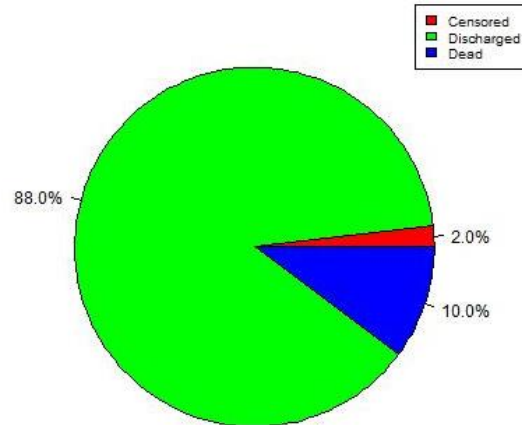
Berdasarkan Tabel 2, dapat diketahui bahwa ada 97 pasien di unit perawatan intensif yang mengidap pneumonia, 657 pasien berhasil pulang pada saat studi berlangsung, 76 pasien meninggal, dan sisanya tetap berada di rumah sakit hingga studi berakhir. Tabel 2 juga memberikan informasi mengenai beberapa statistika deskriptif dari waktu terjadinya suatu kejadian (dalam satuan hari). Secara lebih spesifik, berdasarkan Tabel 3, dapat diketahui bahwa dari semua pasien yang mengidap pneumonia, ada 8 pasien tersensor, 68 pasien pulang ke rumah, dan 21 pasien meninggal. Sementara itu, dari semua pasien yang tidak mengidap pneumonia, ada 6 pasien tersensor, 589 pasien pulang ke rumah, dan 55 pasien meninggal.

Diagram pai proporsi pasien yang mengidap pneumonia dan yang tidak mengidap pneumonia pada awal studi disajikan pada Gambar 1. Dari Gambar 1, dapat diketahui bahwa ada 87% pasien yang tidak mengidap pneumonia dan ada 13% pasien yang mengidap pneumonia. Proporsi pasien yang tidak mengidap pneumonia sekitar enam kali lebih besar daripada proporsi pasien yang mengidap pneumonia.



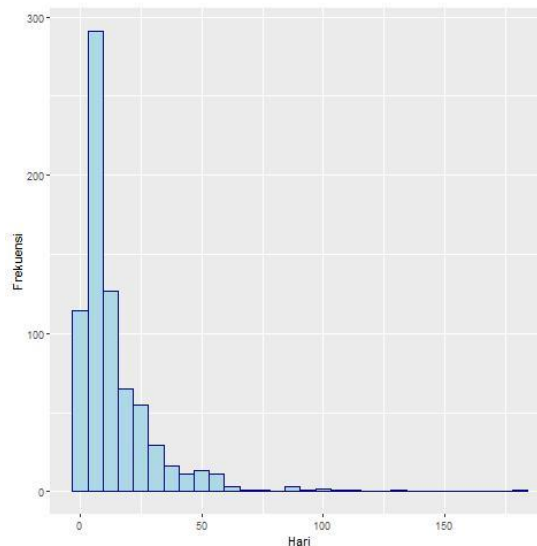
Gambar 1. Proporsi pasien yang mengidap dan tidak mengidap pneumonia pada awal studi

Diagram pai proporsi status pasien pada akhir studi disajikan pada Gambar 2. Berdasarkan Gambar 2, pasien dengan status tersensor memiliki proporsi paling kecil, yakni hanya sebesar 2%, pasien dengan status meninggal dunia memiliki proporsi sebesar 10%. Pasien dengan status pulang dari rumah sakit memiliki proporsi yang paling besar, yakni sebesar 88%.



Gambar 2. Persentase pasien menurut status pada akhir studi

Histogram dari peubah *time* (dalam satuan hari) disajikan pada Gambar 3. Berdasarkan Gambar 3, dapat diketahui bahwa mayoritas pasien berada di unit perawatan intensif selama 2-50 hari. Hanya sedikit saja pasien yang berada di unit perawatan intensif selama lebih dari 50 hari.



Gambar 3. Histogram peubah *time* (dalam satuan hari) untuk semua jenis kejadian

3.2 Pendugaan CIF dengan Metode Nonparametrik beserta Ilustrasinya

Pertama, nilai dugaan CIF dari setiap kejadian dihitung dengan metode nonparametrik. Hasilnya dapat dilihat pada Tabel 4. Selanjutnya nilai dugaan CIF yang telah dihitung ditinjau berdasarkan status penyakit pasien. Terakhir, nilai dugaan CIF yang telah dihitung secara nonparametrik dibandingkan dengan fungsi kumulatif yang dihitung dari penduga Kaplan-Meier.

Nilai dugaan CIF dari setiap kejadian memiliki interpretasi secara langsung. Sebagai contoh, pada hari ke-5 studi, ada 30.12% pasien yang sudah pulang ke rumah, ada 1.47% pasien yang meninggal dunia, dan sisanya masih berlanjut ke waktu selanjutnya. Pada akhir studi, jumlah CIF dari setiap kejadian akan sama dengan $1 - \hat{S}(t)$ untuk semua kejadian dan umumnya akan mendekati satu. Hasil dugaan CIF dua kejadian

secara nonparametrik pada Tabel 4 dapat digambarkan dalam bentuk grafik fungsi sebaran kumulatif seperti yang dapat dilihat pada Gambar 4.

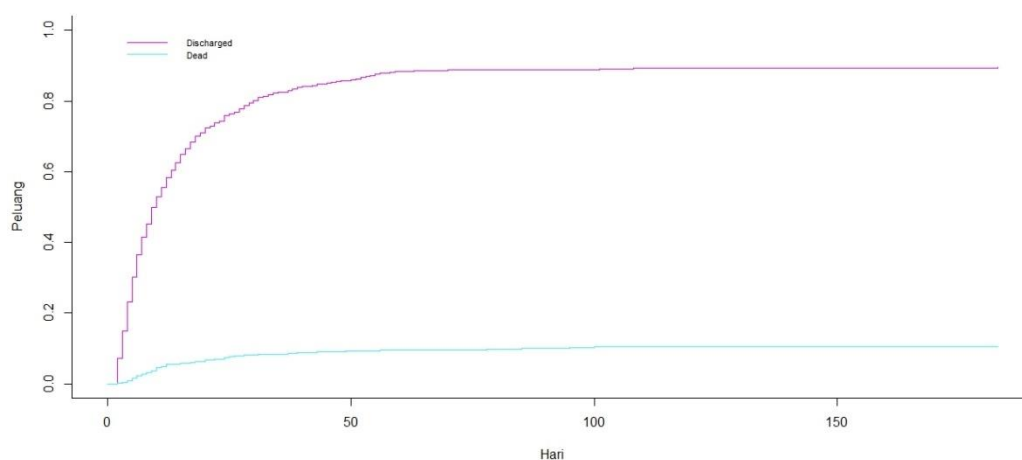
Tabel 4. Hasil dugaan CIF dua kejadian secara nonparametrik

Hari	<i>Discharged</i>	<i>Dead</i>
0	0	0
5	0.30120	0.01473
10	0.52975	0.04569
50	0.85872	0.09269
100	0.88876	0.10351
140	0.89185	0.10506
180	0.89185	0.10506

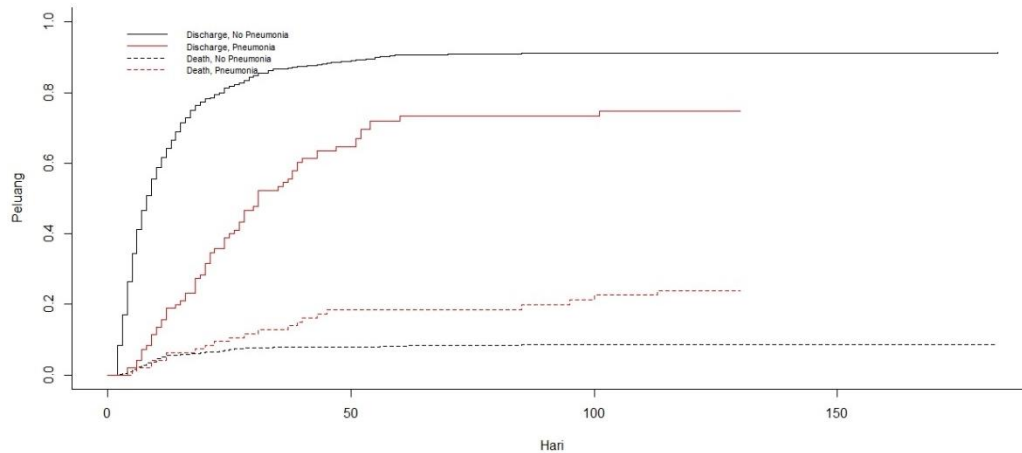
Berdasarkan Gambar 4, dapat dilihat bahwa pada hari ke-0 nilai fungsi kumulatif untuk setiap kejadian bernilai 0. Pada hari ke-100 nilai dugaan CIF untuk kejadian *discharged* ada di atas 0.8 dan nilai dugaan CIF untuk kejadian *dead* ada di kisaran 0.1. Nilai dugaan CIF untuk setiap kejadian menggambarkan sifat fungsi takturun hingga akhir studi. Nilai dugaan CIF untuk setiap kejadian juga dapat ditinjau berdasarkan status penyakit pasien seperti yang ada pada Gambar 5.

Berdasarkan Gambar 5, pasien yang tidak mengidap pneumonia cenderung memiliki lebih banyak kemungkinan untuk pulang dan pasien yang mengidap pneumonia lebih mungkin untuk meninggal dunia. Hal ini masuk akal secara biologis bahwa pasien yang sejak awal mengidap pneumonia akan lebih mungkin untuk meninggal dunia. Selanjutnya, kedua kejadian tersebut dibandingkan dengan dugaan fungsi sebaran kumulatif Kaplan-Meier (1 dikurangi dengan dugaan fungsi *survival* Kaplan-Meier). Ilustrasi perbandingannya dapat dilihat pada Gambar 6.

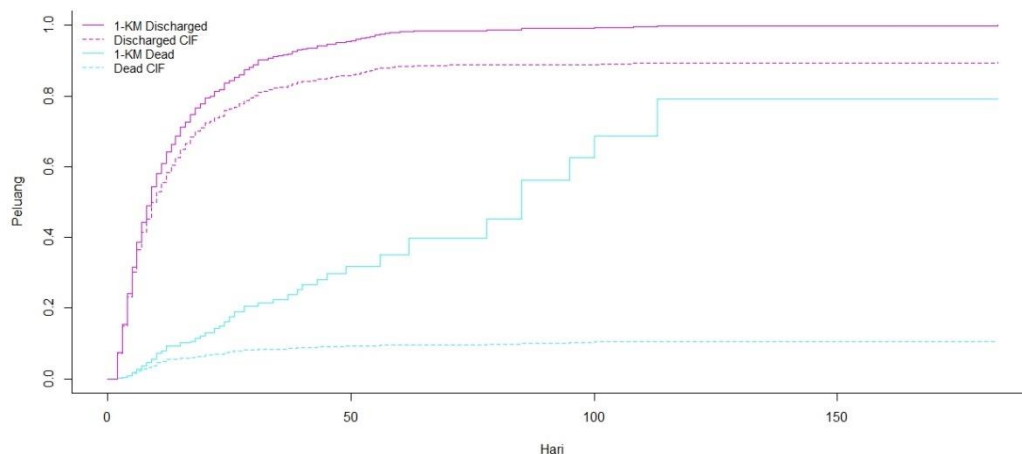
Berdasarkan Gambar 6, dapat dilihat bahwa nilai dugaan fungsi sebaran kumulatif Kaplan-Meier cenderung *overestimate* apabila dibandingkan dengan pendekatan dengan CIF. Hal ini menandakan adanya indikasi situasi *competing risks* dan perlunya penyesuaian untuk mempelajari peluang kejadian yang lebih mendekati kenyataan. Nilai CIF dari setiap kejadian tidak hanya dapat diperoleh melalui metode nonparametrik namun dapat juga diperoleh dengan metode parametrik, yaitu *competing risks regression*.



Gambar 4. Ilustrasi nilai CIF dari setiap kejadian berdasarkan metode nonparametrik



Gambar 5. Ilustrasi nilai CIF dari status penyakit pneumonia pasien



Gambar 6. Ilustrasi nilai dugaan CIF dan dugaan fungsi sebaran kumulatif Kaplan-Meier

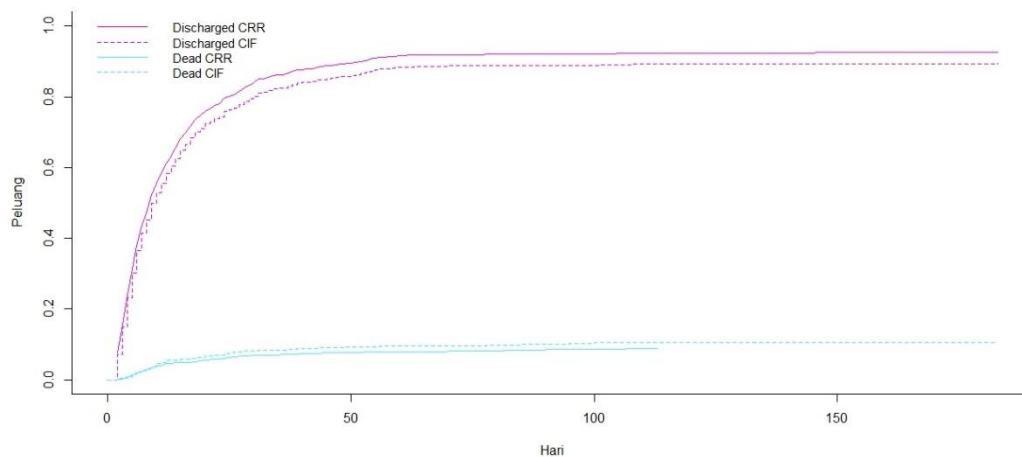
3.3 Pendugaan CIF dengan CRR beserta Ilustrasinya

Competing risks regression merupakan salah satu metode parametrik yang digunakan untuk menduga nilai CIF dari setiap kejadian pada situasi *competing risks*. Konsep dari metode ini mirip dengan regresi linear di mana terdapat peubah-peubah penjelas dan peubah respon. Pada metode ini, peubah-peubah penjelas disebut dengan kovariat. Pada penelitian ini, digunakan status penyakit pasien sebagai masukan. Keluaran metode ini berupa dugaan CIF dari setiap kejadian. Peubah-peubah lain dalam penelitian ini seperti *age* dan *sex* tidak dimasukkan dalam pendugaan karena kedua peubah tersebut tidak signifikan secara statistik. Hasil dugaan dengan *competing risks regression* dapat dilihat pada Tabel 5.

Tabel 5. Hasil dugaan CIF dengan CRR untuk kejadian *dead* dan *discharged*

Kejadian	Karakteristik	Koefisien	<i>p-value</i>
<i>Dead</i>	<i>Dis</i> (pneumonia)	0.975	9.4×10^{-5}
<i>Discharged</i>	<i>Dis</i> (pneumonia)	-0.907	0.000

Berdasarkan Tabel 5, hasil dugaan untuk vektor koefisien karakteristik status pneumonia pasien yang diperoleh adalah $\beta_{pneumonia} = 0.975$ pada kejadian kematian dan $\beta_{pneumonia} = -0.907$ pada kejadian kepulangan pasien. Perlu dipahami bahwa vektor koefisien tersebut hanya mengandung sebuah bilangan dikarenakan kovariat yang digunakan hanya satu peubah. Dengan vektor koefisien tersebut, hasil dugaan untuk CIF setiap kejadian dapat diperoleh dengan ilustrasi seperti pada Gambar 7.



Gambar 7. Ilustrasi nilai dugaan CIF dengan metode nonparametrik dan dengan CRR

Berdasarkan Gambar 7, dapat dilihat bahwa kedua metode menghasilkan nilai dugaan CIF yang berupa fungsi tak turun. Pada hari ke-0, nilai dugaan CIF untuk semua kejadian dengan semua metode sama dengan 0. Pada akhir studi, penjumlahan nilai dugaan CIF untuk semua kejadian mendekati nilai 1 karena ada data yang tersensor. Hasil dugaan CIF dengan CRR tidak jauh berbeda dengan hasil dugaan CIF dengan metode nonparametrik. Untuk dugaan CIF kejadian *dead* dengan CRR hanya sampai dengan 113 hari dikarenakan 113 hari merupakan waktu maksimum untuk kejadian *dead*.

3.4 Uji Hipotesis *k-sample* Gray

Selanjutnya dilakukan uji hipotesis untuk mengetahui apakah benar data pada penelitian ini mencerminkan situasi *competing risks* atau tidak. Misalkan kejadian pasien pulang dilambangkan dengan 1 dan kejadian kematian dilambangkan dengan 2. Misalkan pula kelompok pasien yang mengidap pneumonia dilambangkan dengan 1 dan pasien yang tidak mengidap pneumonia dilambangkan dengan 0. Hipotesis ujinya adalah sebagai berikut:

$H_0 : F_{10}(t) = F_{11}(t) = F_{20}(t) = F_{21}(t) = F_1^0(t)$ untuk semua $t \leq \tau$.

H_a : Setidaknya ada satu $F_{1K}(t)$ yang berbeda untuk beberapa $t \leq \tau$.

$F_1^0(t)$ adalah fungsi subdistribusi yang tak ditentukan dan τ adalah waktu pengamatan terlama dalam studi. Statistik uji yang digunakan untuk uji hipotesis ini seperti pada Persamaan (16). Taraf nyata yang digunakan adalah sebesar 5%. Hasil pengujian dapat dilihat pada Tabel 6.

Tabel 6. Hasil uji hipotesis k-sample Gray

Kejadian	Statistik uji	<i>p-value</i>	Derajat bebas
<i>Discharged</i>	42.07480	8.784784×10^{-11}	1
<i>Dead</i>	16.16525	5.804939×10^{-5}	1

Berdasarkan Tabel 6, diperoleh nilai statistik uji yang menyebar khi-kuadrat dengan derajat bebas 1 atau dinotasikan $\chi_{2-1}^2 = \chi_1^2$. Untuk kejadian *discharged*, karena uji hipotesis Gray menunjukkan *p-value* yang kurang dari taraf nyata, maka terdapat cukup bukti untuk menolak H_0 . Dengan kata lain, nilai dugaan CIF kejadian *discharged* antara kelompok pasien yang mengidap pneumonia dan kelompok pasien yang tidak mengidap pneumonia berbeda. Hasil yang sama untuk kejadian *dead*.

Nilai dugaan CIF untuk setiap kejadian berdasarkan status penyakit dapat dilihat pada Tabel 7. Tabel 7 dapat membantu interpretasi dari uji hipotesis *k-sample* Gray. Hal tersebut dikarenakan uji hipotesis hanya memberikan informasi apakah nilai dugaan CIF cukup berbeda secara statistik. Informasi nilai CIF untuk setiap status penyakit pasien pada setiap kejadian dapat membantu dalam menganalisis kejadian yang lebih mungkin dialami pasien selama studi.

Tabel 7. Nilai dugaan CIF untuk setiap kejadian berdasarkan status penyakit

Waktu (hari)	<i>Discharged</i>		<i>Dead</i>	
	Mengidap pneumonia	Tidak mengidap pneumonia	Mengidap pneumonia	Tidak mengidap pneumonia
0	0	0	0	0
5	0.02602	0.34308	0.01030	0.01538
10	0.13499	0.58850	0.04147	0.04628
30	0.47660	0.84811	0.11604	0.07607
50	0.64633	0.88999	0.18393	0.07930
100	0.73293	0.91100	0.22550	0.08577
120	0.74680	0.91262	0.23936	0.08577

Berdasarkan Tabel 7, dapat dilihat bahwa nilai dugaan CIF untuk kejadian *discharged* pada pasien yang tidak mengidap pneumonia lebih besar dibandingkan dengan pasien yang mengidap pneumonia sehingga dapat dikatakan bahwa kelompok pasien yang tidak mengidap pneumonia lebih mungkin untuk pulang ke rumah. Mulai dari hari ke-30, nilai dugaan CIF untuk kejadian *dead* pada pasien yang mengidap pneumonia lebih besar dibandingkan dengan pasien yang tidak mengidap pneumonia sehingga dapat dikatakan bahwa kelompok pasien yang mengidap pneumonia lebih mungkin untuk meninggal dunia. [17] memberikan contoh cara membaca *output software R* untuk uji hipotesis *k-sample* Gray.

Berdasarkan pendugaan nilai CIF baik secara nonparametrik maupun secara parametrik serta berdasarkan hasil uji hipotesis *k-sample* Gray, dapat diketahui bahwa

situasi *competing risks* benar terjadi pada pasien pneumonia yang digunakan dalam penelitian ini. Penduga Kaplan-Meier perlu penyesuaian lebih lanjut dalam menduga fungsi *survival* dalam situasi *competing risks* agar tidak *overestimate*. Pada situasi seperti ini, kedua metode pendugaan nilai CIF dapat digunakan dan penggunaannya bergantung pada tujuan penelitian dan ketersediaan data.

4 Simpulan

Penduga Kaplan-Meier memberikan nilai dugaan fungsi sebaran kumulatif kejadian yang *overestimate* sehingga situasi *competing risks* harus ditangani dengan lebih hati-hati. Situasi *competing risks* benar terjadi pada pasien pengidap pneumonia di rumah sakit Charité University. Penyakit pneumonia yang menyerang pasien juga memiliki pengaruh yang signifikan. Pasien dengan penyakit pneumonia lebih mungkin untuk meninggal dunia selama periode studi berlangsung. Nilai dugaan CIF dengan metode nonparametrik dan metode parametrik mendekati satu sama lain. Dengan mempertimbangkan situasi *competing risks*, juga dapat diketahui pengaruh dari suatu penyakit atau suatu pengobatan tertentu yang memengaruhi kelompok studi.

Daftar Pustaka

- [1] Dickson DCM, Hardy MR, Waters HR. 2020. *Actuarial Mathematics for Life Contingent Risks*. 3rd ed. Cambridge: Cambridge University Press.
- [2] Pavlov V, Mihova V. 2018. An application of survival model in insurance. Di dalam: Todorov MD. *Applications of Mathematics in Technical and Natural Sciences*. AIP Conference Proceedings; 2018 Jun 20-25; Albena, Bulgaria. Albena: hlm 1-12; [diakses 2022 Jun 08]. <https://aip.scitation.org/doi/abs/10.1063/1.5064883>.
- [3] Pintilie M. 2006. *Competing Risks: A Practical Perspective*. West Sussex: John Wiley & Sons Ltd.
- [4] Kalbfleisch JD, Prentice RL. 2022. *The Statistical Analysis of Failure Time Data*. 2nd ed. Hoboken (NJ): John Wiley & Sons, Inc.
- [5] Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. 2004. A note on competing risks in survival data analysis. *British Journal of Cancer*. 91(7):1229-1235. <https://doi.org/10.1038/sj.bjc.6602102>.
- [6] Berry SD, Ngo L, Samelson EJ, Kiel DP. 2010. Competing risk of death: An important consideration in studies of older adults. *Journal of the American Geriatrics Society*. 58(4):783-787. <https://doi.org/10.1111/j.1532-5415.2010.02767.x>.
- [7] Beyersmann J, Gastmeier P, Grundmann H, Bärwolff S, Geffers C, Behnke M, Rüdén H, Schumacher M. 2006. Use of multistate models to assess prolongation of intensive care unit stay due to nosocomial infection. *Infection Control & Hospital Epidemiology*. 27(5):493-499. <https://doi.org/10.1086/503375>.
- [8] Frey, BB. 2022. *The SAGE Encyclopedia of Research Design*. 2nd Ed. Thousand Oaks (CA): SAGE Publications, Inc.
- [9] London D. 1997. *Survival Models and Their Estimation*. 3rd Ed. Winsted (CT): ACTEX Publications.
- [10] Reid N. 2000. Likelihood. *Journal of the American Statistical Association*. 95(452):1335-1340. <https://doi.org/10.1080/01621459.2000.10474343>.
- [11] Noordzij M, Leffondré K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ. 2013. When do we need competing risks methods for survival analysis in nephrology?

- Nephrology Dialysis Transplantation*. 28(11):2670-2677.
<https://doi.org/10.1093/ndt/gft355>.
- [12] Hsu JY, Roy JA, Xie D, Yang W, Shou H, Anderson AH, Landis JR, Jepson C, Wolf M, Isakova T, Rahman M, Feldman HI. 2017. Statistical methods for cohort studies of CKD: Survival analysis in the setting of competing risks. *Clinical Journal of the American Society of Nephrology*. 12(7):1181-1189.
<https://doi.org/10.2215/CJN.10301016>.
- [13] Bajorunaite R, Klein JP. 2008. Comparison of failure probabilities in the presence of competing risks. *Journal of Statistical Computation and Simulation*. 78(10):951-966.
<https://doi.org/10.1080/00949650701473791>.
- [14] Braun TM, Yuan Z. 2007. Comparing the small sample performance of several variance estimators under competing risks. *Statistics in Medicine*. 26(5):1170-1180.
<https://doi.org/10.1002/sim.2661>.
- [15] Cox DR. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*. 34(2):187-202. <https://doi.org/10.1111/j.2517-6161.1972.tb00899.x>.
- [16] Fine JP, Gray RJ. 1999. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. 94(446):496-509.
<https://doi.org/10.2307/2670170>.
- [17] Zhang Z. 2017. Survival analysis in the presence of competing risks. *Annals of Translational Medicine*. 5(3):1-9. <https://doi.org/10.21037/atm.2016.08.62>.