

Penerapan Analisis Survival untuk Menaksir Waktu Bertahan Hidup bagi Penderita Penyakit Jantung

Oleh :

Yani Hendrajaya (me_yen2@yahoo.co.id),
Adi Setiawan dan Hanna A. Parhusip
Program Studi Matematika, Fakultas Sains dan Matematika
Universitas Kristen Satya Wacana
Jl. Diponegoro 52-60 Salatiga 50711

Abstract

Applying survival analysis in survival data from coroner heart (*acute miocard infark*) patients is discussed in this short paper. The goal of this research is to determine the treatment that gives a longer survival time. The treatments are the ring treatment, the bypass treatment and the medicine treatment. Data was collected from medical record patients who regularly going to control and check their healthy and conditions to a heart specialist doctor. The sample consists of 90 patients; 30 patients used ring treatment, 30 patients used bypass treatment, and 30 patients used medicine treatment.

Survival analysis by using parametric and non parametric estimation are used to estimate the survival time for coroner heart patients using three treatments. The result of this research shows that by using both estimations there is a difference survival time among ring treatment, bypass treatment and medicine treatment. A better medical treatment that gives longer survival time for coroner heart patients is the ring treatment.

Key words : survival analysis, parametric estimation, non parametric estimation, survival time.

1. Pendahuluan

Saat ini, kemungkinan atau peluang seseorang terkena penyakit semakin besar karena banyak jenis penyakit berbahaya yang disebabkan pola hidup masyarakat yang kurang sehat. Salah satu jenis penyakit yang berbahaya dan mematikan tersebut adalah penyakit jantung. Walaupun penyakit jantung adalah jenis penyakit yang sulit atau tidak mungkin disembuhkan, paling tidak pengidap penyakit jantung akan berusaha meminimalkan resiko kematian dengan melakukan tindakan atau usaha pengobatan tertentu. Untuk pengobatan penyakit jantung, ada 3 macam teknik yang dapat dilakukan, yaitu teknik pengobatan dengan menggunakan ring, teknik pengobatan dengan by pass, dan teknik pengobatan yang menggunakan obat.

Tabel 1. Data waktu bertahan hidup penderita penyakit jantung

No	Waktu Bertahan Hidup (bulan)	Status	Kelompok
1	6	0	ring
2	7	0	ring
3	7	0	ring
4	8	0	ring
5	11	0	ring
6	17	0	ring
7	20	0	ring
8	21	0	ring
9	21	0	ring
10	25	0	ring
11	26	0	ring
12	26	0	ring
13	38	0	ring
14	51	0	ring
15	52	0	ring
16	56	0	ring
17	57	0	ring
18	61	1	ring
19	62	0	ring
20	62	0	ring
21	66	0	ring
22	71	1	ring
23	71	0	ring
24	75	0	ring
25	83	0	ring
26	106	0	ring
27	123	0	ring
28	128	0	ring
29	156	0	ring
30	183	0	ring
31	6	0	By pass
32	6	0	By pass
33	7	0	By pass
34	12	0	By pass
35	12	0	By pass
36	16	0	By pass
37	17	0	By pass
38	17	0	By pass
39	21	0	By pass
40	26	0	By pass
41	32	0	By pass
42	33	0	By pass
43	42	0	By pass
44	42	0	By pass
45	56	0	By pass
46	56	0	By pass
47	60	1	By pass
48	65	0	By pass
49	78	0	By pass
50	87	0	By pass
51	87	0	By pass
52	93	0	By pass
53	102	0	By pass
54	116	0	By pass
55	116	1	By pass
56	146	1	By pass
57	161	0	By pass
58	173	1	By pass
59	178	1	By pass
60	182	1	By pass
61	6	0	Obat
62	8	0	Obat
63	8	0	Obat
64	11	0	Obat
65	12	0	Obat
66	12	0	Obat
67	16	0	Obat
68	30	0	Obat
69	33	0	Obat
70	33	0	Obat
71	35	0	Obat
72	38	0	Obat
73	47	1	Obat
74	48	0	Obat
75	62	1	Obat
76	62	0	Obat
77	88	0	Obat
78	88	0	Obat
79	92	1	Obat
80	97	0	Obat
81	98	1	Obat
82	101	1	Obat
83	102	1	Obat
84	116	0	Obat
85	132	1	Obat
86	137	1	Obat
87	141	1	Obat
88	142	1	Obat
89	151	1	Obat
90	178	1	Obat

Dengan menganalisis data waktu bertahan hidup penderita penyakit jantung koroner yang melakukan pemeriksaan secara teratur ke salah seorang dokter spesialis jantung di Solo, akan ditentukan teknik pengobatan yang lebih baik bagi penderita penyakit jantung. Analisis ini akan dilakukan dengan analisis *survival* estimasi

parametrik dan non parametrik, serta menggunakan program S-PLUS 2000. Data waktu bertahan hidup pasien penderita penyakit jantung disajikan pada Tabel 1.

2. Dasar Teori

Analisis *survival* adalah salah satu cabang statistika yang mempelajari teknik analisis data *survival*. Data *survival* adalah data waktu bertahan sampai munculnya kejadian tertentu. Data *survival* dikumpulkan dalam suatu periode waktu yang terbatas, dan sebagai konsekuensinya bisa saja data yang diperoleh tidak mencakup total waktu bertahan seseorang. Hal inilah yang kemudian dalam analisis *survival* disebut dengan data tersensor.

2.1 Estimasi Parametrik

Misalkan Y adalah waktu bertahan hidup sampai munculnya kejadian tertentu. Fungsi *survival*, $S(y)$, mendefinisikan probabilitas dari suatu individu untuk bertahan setelah waktu yang ditetapkan, namakan y , $S(y) = P(Y > y)$. Grafik fungsi *survival* adalah grafik fungsi yang tidak naik. Nilai fungsi $S(0) = 1$ dan $S(\infty) = 0$, artinya dapat dipastikan bahwa semua orang yang diamati pasti akan mengalami kejadian tertentu.

Fungsi *survival* dapat pula diperoleh dengan cara mengintegalkan fungsi kepadatan probabilitas (*probability density function*) dari Y yaitu $f(y)$,

$$S(y) = P(Y > y) = \int_y^{\infty} f(y) dy .$$

Fungsi *hazard*, $h(y)$, mendefinisikan laju kegagalan dari suatu individu untuk mampu bertahan setelah melewati waktu yang ditetapkan yaitu y (Klein dan Moeschberger, 1997).

$$h(y) = \frac{f(y)}{S(y)} = - \frac{d}{dy} [\ln S(y)] .$$

Sedangkan fungsi *hazard* kumulatif (*cumulative hazard function*) didefinisikan oleh :

$$H(y) = \int_0^y h(u) du = - \int_0^y \frac{d}{du} [\ln S(u)] du = - \ln S(y) .$$

2.1.1 Distribusi Weibull

Suatu variabel acak Y dikatakan berdistribusi Weibull dengan parameter β dan θ jika memiliki fungsi kerapatan probabilitas sebagai berikut :

$$f(y; \beta, \theta) = \frac{\beta}{\theta^\beta} y^{\beta-1} \exp \left[-\left(\frac{y}{\theta}\right)^\beta \right].$$

Fungsi *survival* dan fungsi *hazard* untuk distribusi Weibull, yaitu :

$$S(y) = \exp \left[-\left(\frac{y}{\theta}\right)^\beta \right] \quad ; \quad h(y) = \frac{\beta}{\theta} \left(\frac{y}{\theta}\right)^{\beta-1} = \frac{\beta}{\theta^\beta} y^{\beta-1}.$$

Notasi yang menunjukkan bahwa Y berdistribusi Weibull adalah : $Y \sim \text{WEI}(\beta, \theta)$.

Jika Y_1, Y_2, \dots, Y_n adalah suatu sampel acak dari distribusi $f(y; \beta, \theta)$, maka untuk mencari nilai dari parameter β dan θ dapat digunakan sistem persamaan :

$$E(Y) = \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad , \quad E(Y^2) = \overline{Y^2} = \frac{1}{n} \sum_{i=1}^n Y_i^2 .$$

Nilai rata-rata dari $Y \sim \text{WEI}(\beta, \theta)$ adalah : $E(Y) = \theta \Gamma\left(1 + \frac{1}{\beta}\right)$. Sedangkan nilai rata-rata dari Y^2 adalah $E(Y^2) = \theta^2 \Gamma\left(1 + \frac{2}{\beta}\right)$.

Parameter β dan θ dapat ditentukan dengan menyelesaikan sistem persamaan :

$$\frac{1}{n} \sum_{i=1}^n Y_i = \theta \Gamma\left(1 + \frac{1}{\beta}\right) \quad , \quad \frac{1}{n} \sum_{i=1}^n Y_i^2 = \theta^2 \Gamma\left(1 + \frac{2}{\beta}\right)$$

2.1.2 Uji Kecocokan Distribusi Probabilitas

Uji probabilitas digunakan untuk menguji kecocokan dari distribusi probabilitas pada data waktu bertahan hidup penderita kanker payudara. Uji kecocokan distribusi ini menggunakan uji *Kolmogorov-Smirnov*, atau yang lebih dikenal dengan *Goodness-Of-Fit Test*. Jika fungsi distribusi $F(y)$ akan diduga dengan $F_0(y)$, maka akan ditetapkan hipotesis nol dan pengujian yang sesuai, yaitu :

$$H_0 : F(y) = F_0(y) \quad , \quad \text{untuk semua nilai } y.$$

$$H_1 : F(y) \neq F_0(y) \quad , \quad \text{untuk paling sedikit satu nilai } y.$$

Statistik uji *Kolmogorov-Smirnov* (D_n) untuk pengujian dua sisi diberikan sebagai berikut :

$$D_n = \text{Maks}_{1 \leq i \leq r} \left\{ \text{Maks} \left[\left| F(y_i) - \frac{i}{n} \right|, \left| F(y_i) - \frac{i-1}{n} \right| \right] \right\} ,$$

untuk $i = 1, 2, \dots, r + 1$, dengan r adalah banyaknya nilai y yang berbeda dan $F(y_i)$ adalah fungsi distribusi yang diduga. Kriteria penolakan hipotesis nol (H_0) adalah sebagai berikut :

1. Hipotesis nol (H_0) diterima jika nilai- $p \geq$ taraf signifikansi (*level of significance*) α yang dipilih,
2. Sebaliknya jika nilai- $p < \alpha$ maka hipotesis nol (H_0) akan ditolak, yang artinya distribusi probabilitas yang diduga tidak sama dengan distribusi tertentu.

2.2 Estimasi Non Parametrik

Misalkan suatu peristiwa terjadi pada waktu D dan $t_1 < t_2 < \dots < t_D$, pada saat t_i terdapat d_i peristiwa, dan Y_i adalah banyaknya individu yang memiliki resiko setelah waktu t_i . Setelah mengubah notasi, maka untuk menaksir peluang suatu individu bertahan pada waktu t_i adalah :

$$\hat{S}(t) = \prod_{t_i \leq t} \frac{Y_i - d_i}{Y_i} \quad \dots (3)$$

yang selanjutnya dikenal sebagai estimasi Kaplan-Meier. Nilai variansi untuk $\hat{S}(t)$ dirumuskan oleh :

$$\hat{V}[\hat{S}(t)] = \hat{S}(t)^2 \sum_{t_i \leq t} \frac{d_i}{Y_i(Y_i - d_i)}$$

dengan standar error diberikan oleh $[\hat{V}[\hat{S}(t)]]^{1/2}$. Selang kepercayaan $(1-\alpha) \times 100\%$ untuk $\hat{S}(t)$ pada titik t_0 ,

$$\hat{S}(t_0) - Z_{1-\frac{\alpha}{2}} [\hat{V}[\hat{S}(t_0)]]^{1/2} \leq \hat{S}(t_0) \leq \hat{S}(t_0) + Z_{1-\frac{\alpha}{2}} [\hat{V}[\hat{S}(t_0)]]^{1/2} .$$

2. Metodologi Penelitian

Langkah-langkah yang dilakukan untuk menyelesaikan permasalahan menentukan teknik pengobatan yang lebih baik bagi penderita penyakit jantung adalah :

Estimasi Parametrik

1. Mengestimasi nilai dari parameter β dan θ dengan menyelesaikan persamaan (2) dan menguji kecocokan distribusi probabilitas menggunakan nilai- p untuk statistika *Kolmogorov-Smirnov* yang diperoleh dari hasil keluaran perintah *ks.gof* pada SPLUS,
2. Memodelkan data waktu bertahan hidup penderita penyakit jantung dengan menggunakan perintah *cursorReg*,
3. Menguji ada tidaknya perbedaan lama waktu bertahan hidup penderita penyakit jantung dengan tiga teknik pengobatan. Pengujian ini menggunakan nilai- p untuk statistika *Likelihood Ratio* yang diperoleh dari hasil keluaran perintah *anova*,
4. Memperkirakan lama waktu bertahan hidup dan peluang kegagalan bertahan hidup dengan perintah *predict*.

Estimasi Non Parametrik

1. Memodelkan data waktu bertahan hidup penderita penyakit jantung dengan menggunakan perintah *survfit*,
2. Membuat grafik perbandingan antara tiga teknik pengobatan dengan perintah *plot*.

3. Analisis dan Pembahasan

Estimasi Parametrik

Uji kecocokan distribusi probabilitas dilakukan dengan metode *trial and error*. Beberapa nilai parameter β dan θ disajikan pada Tabel 2.

Tabel 2. Estimasi Parameter β dan θ

Nilai β yang dipilih	θ
0,5	59,25
1	118,5
2,5	133,5567

Hipotesis yang disusun untuk melakukan uji ini adalah :

H_0 : Data waktu bertahan hidup penderita penyakit jantung berdistribusi Weibull untuk parameter yang ditetapkan

H_1 : Data waktu bertahan hidup penderita penyakit jantung tidak berdistribusi Weibull untuk parameter yang ditetapkan

Jika dipilih nilai $\beta = 0,5$ dan $\theta = 59,25$ maka dari hasil uji *Kolmogorov-Smirnov* diperoleh nilai- p sebesar 0. Dengan mengambil taraf signifikansi $\alpha = 0,05$, maka nilai- $p < \alpha$ yang artinya hipotesis nol (H_0) ditolak. Ini berarti pemilihan nilai $\beta = 0,5$ dan $\theta = 59,25$ belum tepat. Untuk itu perlu dilakukan uji *Kolmogorov-Smirnov* lagi dengan mengambil nilai yang berbeda untuk masing-masing parameter β dan parameter θ . Dari hasil uji *Kolmogorov-Smirnov* untuk nilai $\beta = 2,5$ dan nilai $\theta = 133,5567$, diperoleh nilai- p sebesar 0,8961. Karena nilai- $p > \alpha = 0,05$ maka hipotesis nol (H_0) diterima. Jadi data waktu bertahan hidup penderita penyakit jantung berdistribusi Weibull dengan nilai parameter $\beta = 2,5$ dan nilai parameter $\theta = 133,5567$.

Dari Tabel 3. terlihat bahwa peluang kegagalan bertahan hidup pasien penderita penyakit jantung yang melakukan pengobatan hanya dengan menggunakan obat lebih tinggi bila dibandingkan dengan penderita penyakit jantung yang melakukan pengobatan dengan menggunakan ring ataupun by pass. Pada kelompok pasien penderita penyakit jantung yang diobati dengan menggunakan ring, peluang kegagalan penderita penyakit jantung untuk mampu bertahan hidup selama 100 bulan (± 8 tahun) sebesar 8 % atau dengan kata lain peluang keberhasilan pasien penderita penyakit jantung yang diobati dengan menggunakan ring adalah 92% untuk bertahan hidup selama 8 tahun.

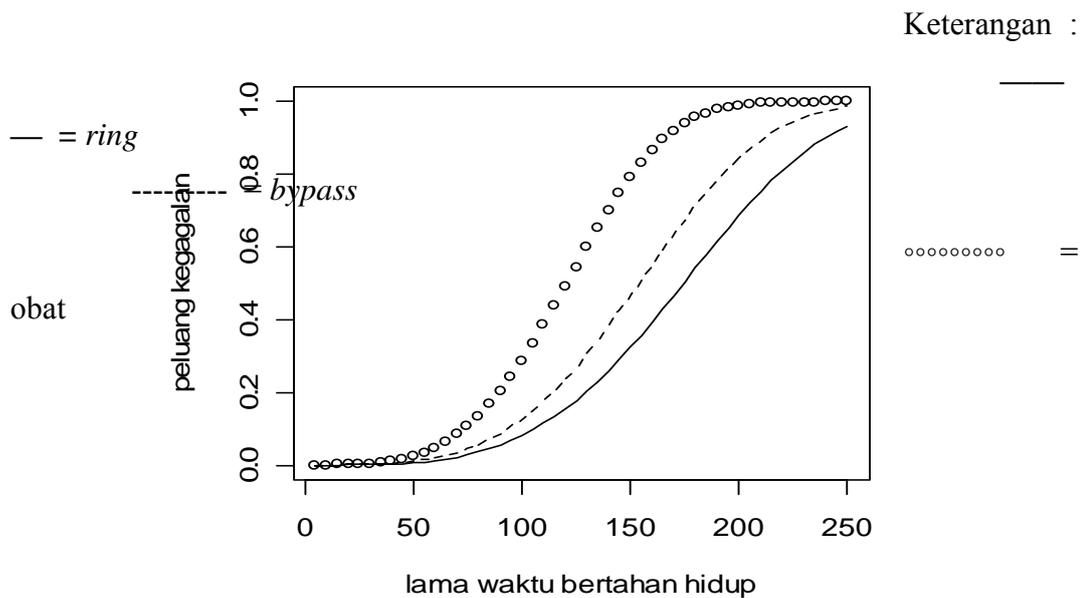
Tabel 3. Peluang kegagalan bertahan hidup pasien penderita penyakit jantung

Lama waktu bertahan hidup (dalam bulan)	Peluang kegagalan bertahan hidup		
	ring	by pass	Obat
50	0,0061	0,0098	0,0242
100	0,0808	0,1265	0,2873
150	0,3237	0,4663	0,7926
200	0,6874	0,8454	0,9907

Pada Tabel 4, jika diberikan peluang kegagalan bertahan hidup sebesar 0,1 atau dengan kata lain peluang keberhasilan hidup sebesar 90%, maka kelompok penderita yang diobati dengan ring dapat bertahan hidup selama 106 bulan (± 9 tahun), kelompok penderita yang diobati dengan by pass dapat bertahan hidup selama 93 bulan (± 8 tahun), dan kelompok penderita yang minum obat dapat bertahan hidup selama 73 bulan (± 6 tahun).

Tabel 4. Lama waktu bertahan hidup penderita penyakit jantung

Peluang kegagalan bertahan hidup	Lama waktu bertahan hidup (dalam bulan)		
	ring	by pass	Obat
0,1	106,0835	93,6169	73,4612
0,3	146,3868	129,1839	101,3706
0,5	174,4626	153,9603	120,8126
0,7	201,8486	178,1280	139,7770
0,9	239,5465	211,3958	165,8822



Gambar 1. Perbandingan peluang kegagalan bertahan hidup antara kelompok ring, bypass dan obat

Untuk mengetahui ada tidaknya perbedaan lama waktu bertahan hidup penderita penyakit jantung yang diobati dengan tiga teknik pengobatan, maka akan dilakukan uji *Likelihood-Ratio*. Hipotesis yang disusun untuk melakukan uji ini adalah :

H_0 : lama waktu bertahan hidup penderita kanker payudara dengan dua teknik pengobatan adalah sama.

H_1 : terdapat perbedaan lama waktu bertahan hidup penderita kanker payudara dengan tiga teknik pengobatan.

Hipotesis nol (H_0) ditolak jika nilai- $p < \alpha$ untuk tingkat signifikansi α yang dipilih. Pada kasus ini penulis mengambil nilai $\alpha = 0,1$. Dengan perintah *anova* pada S-PLUS diperoleh nilai- p sebesar 0,0546, yang berarti nilai- $p < \alpha = 0,1$ sehingga H_0 ditolak. Jadi terdapat perbedaan lama waktu bertahan hidup dengan tiga teknik pengobatan tersebut sehingga lama waktu bertahan hidup penderita penyakit jantung tergantung pada teknik pengobatan yang diterima oleh penderita.

Estimasi Non Parametrik

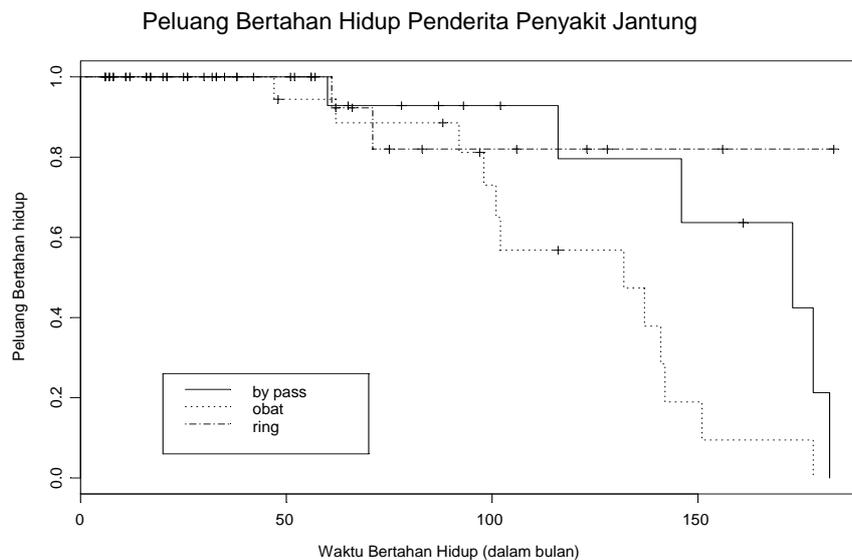
Hasil olahan data waktu bertahan hidup penderita penyakit jantung dengan estimasi non parametrik disajikan pada Tabel 5. Pada kelompok pengobatan ring, terdapat 2 penderita penyakit jantung yang telah meninggal dunia. Pada kelompok pengobatan by pass, terdapat 6 penderita penyakit jantung yang telah meninggal dunia. Sedangkan pada kelompok pengobatan dengan menggunakan obat, penderita penyakit jantung yang telah meninggal dunia sebanyak 12 penderita.

Jika hasil olahan data waktu bertahan hidup penderita kanker payudara dengan estimasi non parametrik dinyatakan dalam bentuk grafik, maka akan muncul gambar grafik seperti tersaji pada Gambar 2. Bentuk grafik pada Gambar 2. menyerupai anak tangga (*stepwise*), sehingga nilai fungsi *survival* $\hat{S}(y)$ akan sama untuk suatu interval waktu. Agar lebih jelasnya dapat dilihat pada Tabel 6.

Tabel 5. Hasil olahan dengan perintah survfit untuk data yang tidak tersensor

No	Ring		by pass		Obat	
	Waktu bertahan hidup	Peluang bertahan hidup	Waktu bertahan hidup	Peluang bertahan hidup	Waktu bertahan hidup	Peluang bertahan hidup
1	61	0,923	60	0,929	47	0,944
2	71	0,821	116	0,796	62	0,885
3			146	0,637	92	0,812
4			173	0,424	98	0,731
5			178	0,212	101	0,649
6			182	0,000	102	0,568
7					132	0,474
8					137	0,379
9					141	0,284
10					142	0,189
11					151	0,095
12					178	0,000

Gambar 2. Grafik Perbandingan Peluang Bertahan Hidup



Tabel 6. Nilai $\hat{S}(y)$ untuk beberapa interval waktu

No	Ring		By Pass		Obat	
	Interval waktu bertahan hidup	Peluang bertahan hidup	Interval waktu bertahan hidup	Peluang bertahan hidup	Interval waktu bertahan hidup	Peluang bertahan hidup
1	$0 \leq y < 61$	1	$0 \leq y < 60$	1	$0 \leq y < 47$	1
2	$61 \leq y < 71$	0,923	$60 \leq y < 116$	0,929	$47 \leq y < 62$	0,944
3	$71 \leq y < 183$	0,821	$116 \leq y < 146$	0,796	$62 \leq y < 92$	0,885
4			$146 \leq y < 173$	0,637	$92 \leq y < 98$	0,812
5			$173 \leq y < 178$	0,424	$98 \leq y < 101$	0,731
6			$178 \leq y < 182$	0,212	$101 \leq y < 102$	0,649
7			$182 \leq y < \infty$	0,000	$102 \leq y < 132$	0,568
8					$132 \leq y < 137$	0,474
9					$137 \leq y < 141$	0,379
10					$141 \leq y < 142$	0,284
11					$142 \leq y < 151$	0,189
12					$151 \leq y < 178$	0,095
13					$178 \leq y < \infty$	0,000

Berdasarkan hasil yang telah diperoleh dapat diambil beberapa kesimpulan, yang pertama adalah bahwa analisis *survival* dapat digunakan untuk menentukan teknik pengobatan yang lebih baik bagi pasien penderita penyakit jantung. Kesimpulan yang kedua adalah lama waktu bertahan hidup yang mampu dicapai pasien penderita penyakit jantung berhubungan dengan teknik pengobatan yang diterima oleh pasien penderita penyakit jantung, dan kesimpulan yang ketiga adalah bahwa teknik pengobatan yang memberikan waktu bertahan hidup lebih lama bagi pasien penderita penyakit jantung adalah teknik pengobatan dengan menggunakan ring.

DAFTAR PUSTAKA

- Dobson, J.A., 2002, *An Introduction to Generalized Linear Models*, Chapman & Hall, USA.
- Harinaldi, 2005, *Prinsip-Prinsip Statistik untuk Teknik dan Sains*, Erlangga, Jakarta.
- Klein, J.P and Moeschberger, M.L., 1997, *Survival Analysis : Techniques for Censored and Truncated Data*, Springer-Verlag New York Inc, New York.
- Venables, W.N and Ripley, B.D., 1994, *Modern Applied Statistics with S-PLUS*, Springer-Verlag New York Inc, New York.