

Faktor-Faktor Risiko yang Mempengaruhi Kanker Paru-Paru dengan Menggunakan Analisis Regresi Logistik

Juwita^{#1}, Nonong Amalita^{#2}, Meira Parma Dewi^{#3}

[#]*Student of Mathematic Departement Universitas Negeri Padang, Indonesia*

[#]*Lecture of Mathematical Department Universitas Negeri Padang, Indonesia*

¹whiwyt_whita@yahoo.com

²nongaditya@gmail.com

³meiraparma@fmipa.unp.ac.id

Abstract – Lung cancer is a disease with characteristic presence of uncontrolled cell growth in the lung tissue. If not treated, the growth of these cells can spread beyond the lung. So if it is not taken seriously, will be ensured with lung cancer longer will increase. The factors that affect the risk of lung cancer are age, gender, cigarette consumption and lung disease history, family history, and type of work. To determine the factors that most influence the risk of lung cancer will be established a model which can help the application of a causal relationship between two or more variables by using logistic regression analysis. Which aims to determine the factors that influence and determine the chances of each of the factors that affect the risk of lung cancer. Based on the results of the study, the factors that affect the risk of lung cancer at the Hospital Dr. M. Djamil Padang is age cigarette consumption and lung disease history. If a person older than 40 years, smokers and have a history of lung disease, the chances that a person will suffer from lung cancer increases.

Keywords—lung cancer, risk factor, logistic regression analysis

Abstrak– Kanker paru-paru merupakan penyakit dengan ciri khas adanya pertumbuhan sel yang tidak terkontrol pada jaringan paru-paru. Bila tidak dirawat, pertumbuhan sel ini dapat menyebar ke luar dari paru-paru. Jika tidak ditangani dengan serius, akan dipastikan penderita kanker paru semakin lama akan semakin bertambah. Faktor-faktor risiko yang mempengaruhi kanker paru-paru adalah umur, jenis kelamin, konsumsi rokok, riwayat penyakit paru, riwayat keluarga, dan jenis pekerjaan. Untuk mengetahui faktor-faktor risiko yang paling mempengaruhi kanker paru maka dibentuklah sebuah model yang dapat membantu penerapan hubungan kausal (sebab-akibat) antara dua atau lebih dua peubah yaitu dengan menggunakan analisis regresi logistik. Yang bertujuan untuk menentukan faktor-faktor yang mempengaruhi dan mengetahui besar peluang masing-masing faktor risiko yang mempengaruhi kanker paru-paru. Berdasarkan hasil penelitian, faktor-faktor risiko yang berpengaruh terhadap kanker paru-paru di RSUP Dr. M. Djamil Padang umur, konsumsi rokok, dan riwayat penyakit paru. Jika seseorang berumur lebih dari 40 tahun, perokok dan memiliki riwayat penyakit paru-paru maka peluang seseorang tersebut akan menderita kanker paru-paru semakin besar.

Kata Kunci—kanker paru-paru, faktor risiko, analisis regresi logistik

PENDAHULUAN

Kanker merupakan penyakit akibat pertumbuhan tidak normal dari sel-sel jaringan tubuh yang berubah menjadi sel kanker. Dalam perkembangannya, sel-sel kanker ini dapat menyebar ke bagian tubuh lainnya sehingga dapat menyebabkan kematian. Kanker sering dikenal oleh masyarakat sebagai tumor, padahal tidak semua tumor adalah kanker[1]. Tumor adalah segala benjolan tidak normal atau abnormal. Kanker paru-paru merupakan penyakit dengan ciri khas adanya pertumbuhan sel yang tidak terkontrol pada jaringan paru-paru. Bila tidak dirawat, pertumbuhan sel ini dapat menyebar ke luar dari paru-paru

melalui suatu proses yang disebut metastasis ke jaringan yang terdekat atau bagian tubuh yang lainnya. Kanker paru jenis penyakit keganasan yang menjadi penyebab kematian utama pada kelompok kematian akibat keganasan, bukan hanya pada laki laki tetapi juga pada perempuan. Buruknya prognosis penyakit ini mungkin berkaitan erat dengan jaranganya penderita datang ke dokter ketika penyakitnya masih berada dalam stadium awal[4].

Seperti umumnya kanker yang lain, penyebab yang pasti dari kanker paru belum diketahui, sebagian besar kanker paru-paru terjadi tanpa gejala. Seseorang dapat hidup dengan kanker paru-paru selama bertahun-tahun tanpa

mengetahuinya. Kanker paru-paru baru diketahui jika mereka melakukan CT scan atau sinar-X dada. Sehingga perlu diketahui faktor-faktor apa saja yang mempengaruhi terjadinya kanker paru-paru. Karena jika tidak ditanggapi dengan serius faktor resiko penyebab terjadinya akan dipastikan penderita kanker paru-paru semakin lama akan semakin bertambah. Faktor yang mempengaruhi kanker paru-paru yang akan diteliti sebanyak enam faktor, faktor tersebut yaitu: umur, jenis kelamin, konsumsi rokok, riwayat keluarga, riwayat penyakit paru-paru, dan jenis pekerjaan.

Berdasarkan hal diatas, tujuan dari penelitian ini adalah menentukan bentuk model regresi logistik yang sesuai untuk menjelaskan faktor-faktor risiko dan mengetahui besar peluang masing-masing faktor risiko yang paling mempengaruhi kanker paru-paru. Untuk mengetahui faktor-faktor resiko yang sangat mempengaruhi kanker paru-paru maka dibentuklah sebuah model yang dapat membantu penerapan hubungan kausal (sebab-akibat) antara dua atau lebih peubah yaitu menggunakan model regresi. Karena variabel respons Y memiliki dua kategori dan bertipe non metrik, maka model yang digunakan adalah model regresi logistik biner. Bentuk model regresi logistik adalah

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)} \quad (1)$$

Model regresi logistik binomial dengan variabel respon bernilai 1 dan 0, dimana antar pengamatan diasumsikan saling bebas. Maka untuk pendugaan parameter β , salah satunya dapat dilakukan dengan metode *Maximum Likelihood Estimation* (MLE)[3]. Setelah menaksir parameter maka langkah selanjutnya menguji signifikansi parameter tersebut. Untuk itu digunakan uji serentak atau disebut juga uji model *chi-Square*, yaitu:

$$G^2 = -2 \ln \left[\frac{L_1}{L_0} \right]; G^2 = -2 \ln \left[\frac{\binom{n_1}{n} \binom{n_0}{n}^{n_0}}{\prod_{i=1}^n \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{1 - y_i}} \right] \quad (2)$$

Hipotesis:

$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$ (variabel prediktor tidak berpengaruh terhadap variabel respon)

$H_1 : \text{paling sedikit ada satu } \beta_k \neq 0$ (variabel prediktor berpengaruh terhadap variabel respon)

Statistik uji G^2 ini mengikuti distribusi χ^2 dengan derajat prediktornya adalah k (banyaknya variabel prediktor). Dengan kriteria pengujian, $G^2 > \chi^2_{\alpha, k}$ maka tolak H_0 yang berarti pada model regresi terdapat sekurang-kurangnya satu penduga parameter yang tidak sama dengan nol.

Untuk menguji pengaruh setiap β secara individu, hasil pengujian secara parsial atau individual akan

menunjukkan apakah suatu variabel prediktor layak untuk masuk dalam model atau tidak, maka dilakukan uji parsial.

Hipotesis:

$H_0: \beta_j = 0$, untuk $j=1,2,\dots,k$ (variabel X_j tidak berpengaruh nyata)

$H_1: \beta_j \neq 0$, (variabel X_j berpengaruh nyata)

Statistik Uji:

$$Wald(W) = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \quad (3)$$

Rasio yang dihasilkan dari statistik uji, dibawah hipotesis H_0 akan mengikuti sebaran baku normal. Kriteria penolakan (tolak H_0) jika $W > Z_{\alpha/2}$, uji hipotesis dapat juga dibandingkan dengan *p-value*, jika nilai *p-value* kurang dari α , maka tolak H_0 . Parameter dalam model regresi logistik dapat diinterpretasikan menggunakan *odd ratio*. Nilai *odd ratio* didefinisikan sebagai berikut:

$$OR = \frac{e^{\beta_0 + \beta_1}}{e^{\beta_0}} = e^{\beta_1} \quad (4)$$

Jadi, nilai *odd ratio* untuk model regresi logistik adalah e^{β_1} .

METODE

Jenis data dalam penelitian ini adalah data sekunder. Data diperoleh dari bagian Rekam Medis (Medical Record) di RSUP Dr. M.Djamil Padang dari bulan Januari 2014 sampai dengan Desember 2014. Populasi pada penelitian ini adalah Penderita kanker paru-paru pada stadium 2 dan stadium 3 yang dirawat inap di RSUP Dr. M. Djamil Padang dari Januari 2014 sampai dengan Desember 2014 berjumlah 126 orang. Dimana jumlah sampel dalam penelitian ini adalah 96 orang. Yang ditentukan dengan menggunakan rumus slovin dengan taraf kepercayaan 95%.

Penarikan sampel dilakukan dengan menggunakan metode *probability sampling* yaitu *systematic sampling*. Penerapan metode ini dimungkinkan untuk mengetahui dengan jelas sumber kesalahan yang ada. Penarikan sampel acak sistematis merupakan metode pengambilan sampel dimana hanya unsur pertama saja dari sampel dipilih secara acak, sedangkan unsur-unsur selanjutnya dipilih secara sistematis menurut suatu pola tertentu[6].

Langkah-langkah dalam penelitian adalah sebagai berikut:

1. Mendeskripsikan data pengamatan pada pasien penderita kanker paru-paru.
2. Melakukan pendugaan parameter model dengan metode maksimum.

3. Membentuk model dugaan regresi logistik antara variabel respons dengan variabel bebas
4. Melakukan uji signifikansi model regresi logistik dengan menggunakan uji rasio Likelihood.
5. Melakukan uji signifikansi parameter untuk setiap model regresi logistik individu untuk mengetahui variabel-variabel prediktor yang berpengaruh dan mereduksi variabel bebas yang tidak berpengaruh terhadap variabel respons dengan menggunakan uji Wald.
6. Mendapatkan model terbaik dengan variabel prediktor yang berpengaruh secara signifikan
7. Mencari nilai *odd ratio* model terbaik untuk masing-masing variabel prediktor.
8. Melakukan interpretasi terhadap hasil yang diperoleh.

HASIL DAN PEMBAHASAN

A. Deskripsi data

Penelitian ini menggunakan data sekunder yang diperoleh dari Rumah Sakit Umum Pusat M. Djamil Padang. Jumlah pasien kanker paru-paru stadium 2 sebanyak 64 orang atau sebesar 66,67% dan pasien kanker paru-paru stadium 3 sebanyak 32 orang atau sebesar 33,33%. Berikut deskripsi data pasien kanker paru-paru: penderita kanker paru-paru lebih banyak diderita oleh pasien yang berusia ≥ 40 tahun yang berjumlah 72 orang atau sebesar 75% dari pada pasien yang berusia < 40 tahun yang berjumlah 25 orang. Pasien penderita kanker paru-paru lebih banyak diderita oleh laki-laki yang berjumlah 76 orang atau 79,16% dibandingkan perempuan yang hanya 20 orang atau 20,84%. Pasien penderita kanker paru-paru lebih banyak dialami oleh pasien perokok yang berjumlah 66 orang, dimana pasien perokok berat berjumlah 25 orang, perokok sedang berjumlah 33 orang, dan perokok ringan berjumlah 8 orang. Sedangkan pasien yang bukan perokok berjumlah 30 orang atau 31,26%. Pasien yang memiliki riwayat penyakit paru-paru yang berjumlah 69 orang atau sebesar 71,87% daripada pasien yang tidak memiliki riwayat penyakit paru-paru yang berjumlah 27 orang. Pasien yang tidak memiliki riwayat keluarga yang berjumlah 86 orang atau sebesar 89,58% daripada pasien yang memiliki riwayat penyakit paru-paru yang berjumlah 10 orang atau sebesar 10,42%. Pasien dengan jenis pekerjaan yang terpapar zat karsinogenik berjumlah 31 orang atau sebesar 32,29% dan pasien yang jenis pekerjaannya tidak terpapar zat karsinogenik berjumlah 65 orang atau sebesar 67,71%.

B. Analisis data

Adapun langkah-langkah yang dilakukan dalam analisis regresi logistik dalam penelitian ini adalah sebagai berikut:

1). Penaksiran parameter model regresi logistik

Penaksiran parameter model dilakukan dengan menggunakan MLE (*Maximum Likelihood Estimator*), nilai fungsi likelihood dengan mengikutsertakan keenam variabel prediktor hasil dugaan parameter model regresi logistik dengan kemungkinan maksimum likelihood yang terdapat pada tabel berikut.

TABEL I
HASIL DUGAAN PARAMETER REGRESI LOGISTIK DENGAN SEMUA VARIABEL PREDIKTOR

Variabel Prediktor	Nilai Penduga Parameter
Umur (X_1)	-1,235
Jenis Kelamin (X_2)	0,434
Konsumsi rokok (X_3)	1,526
Riwayat Penyakit Paru (X_4)	1,565
Riwayat Keluarga (X_5)	0,692
Jenis Pekerjaan (X_6)	-0,737
Konstanta	-1,696

Berdasarkan Tabel di atas diperoleh model regresi logistik dengan mengikutsertakan semua variabel prediktor adalah sebagai berikut:

$$\pi(x) = \frac{e^{(-1,235+0,434X_1+1,526X_2+1,565X_3+0,692X_4-0,737X_5-1,696X_6)}}{1+e^{(-1,235+0,434X_1+1,526X_2+1,565X_3+0,692X_4-0,737X_5-1,696X_6)}} \quad (5)$$

Dengan melakukan transformasi logit terhadap $\pi(x)$ maka sifat liner dapat terpenuhi, sehingga model tersebut dapat menghasilkan bentuk logit seperti dibawah ini.

$$g(x) = -1,235 + 0,434X_1 + 1,526X_2 + 1,565X_3 + 0,692X_4 - 0,737X_5 - 1,696X_6 \quad (6)$$

Pengujian signifikansi model dengan mengikutsertakan semua variabel prediktor dapat menggunakan uji G yang disebut juga uji *Chi-Square*. Hasil uji signifikansi model dengan menggunakan uji G dapat dilihat pada tabel berikut.

TABEL II
UJI KEBAIKAN MODEL PENUH

	Chi-Square	Sig.
Model Regresi Logistik	17,600	0,007

Pengujian pada hipotesis ini menggunakan taraf nyata 0,1 dan derajat bebas 6. Nilai *Chi-Square* yang diperoleh adalah sebesar 17,600 sedangkan pada tabel *Chi-Square* nilai $\chi^2_{0,1,6} = 10,64$. Dapat dilihat nilai $G > \chi^2_{a,k}$ sehingga keputusan tolak H_0 , yang berarti sekurang-kurangnya ada satu penduga parameter yang tidak sama dengan nol. Dan terlihat juga bahwa nilai signifikansi model regresi logistik lebih kecil dari taraf nyata 0,1. Jadi, variabel prediktor secara bersama-sama mempunyai pengaruh yang signifikan

terhadap variabel respon, dengan kata lain model ini boleh disarankan tapi model tersebut bukanlah model terbaik. Karena dengan menggunakan semua variabel prediktor dalam model tidak akan mengetahui variabel mana yang lebih berpengaruh terhadap variabel respon. Untuk menentukan variabel mana saja yang harus direduksi dari model, selanjutnya analisis dapat dilakukan dengan menggunakan uji Wald untuk melihat kesignifikansian variabel terhadap model. Untuk memperoleh nilai Wald dan nilai signifikansi variabel prediktor dapat dilihat pada Tabel berikut.

TABEL III
SIGNIFIKANSI PARAMETER REGRESI LOGISTIK DENGAN SEMUA VARIABEL PREDIKTOR

Variabel Prediktor	β	SE(β)	Wald	Signifikansi
Umur (X_1)	-1,235	0,629	0,852	0,050
Jenis Kelamin (X_2)	0,434	0,662	0,429	0,512
Konsumsi rokok (X_3)	1,526	0,555	7,578	0,006
Riwayat Penyakit Paru (X_4)	1,565	0,652	5,762	0,016
Riwayat Keluarga (X_5)	0,692	0,756	0,838	0,360
Jenis Pekerjaan (X_6)	-0,737	0,552	1,785	0,182
Konstanta	-1,696	0,805	4,438	0,035

Berdasarkan Tabel di atas diperoleh nilai wald dan nilai signifikansi variabel prediktor. Dari tabel tersebut nilai signifikansi variabel umur (X_1) adalah 0,050 dengan nilai wald sebesar 3,852, variabel konsumsi rokok (X_3) adalah 0,006 dengan nilai wald sebesar 7,578 dan variabel riwayat penyakit paru (X_4) adalah 0,016 dengan nilai wald sebesar 5,762. Karena nilai signifikansinya lebih kecil dari 0,1, hal ini menunjukkan variabel umur (X_1), variabel konsumsi rokok (X_3) dan riwayat penyakit paru (X_4) mempunyai peranan yang signifikan terhadap variabel respon. Sedangkan variabel prediktor yang lain tidak signifikan terhadap variabel respon pada taraf nyata 0,1. Sehingga model regresi logistik dengan seluruh variabel prediktor harus direduksi untuk mendapatkan variabel yang benar-benar memiliki nilai signifikansi yang lebih kecil dari 0,1.

2) Pemilihan Model Regresi Logistik Terbaik

Pemilihan model regresi logistik dilakukan dengan metode langkah mundur (*backward method*) yaitu penyederhanaan model dengan mengeluarkan satu per satu variabel prediktor yang memiliki signifikansi yang lebih besar dari taraf nyata 0,05. Untuk melihat pengaruh model

terbaik diatas, dapat dilihat nilai statistik uji G berdasarkan Lampiran 2 diperoleh sebagai berikut.

TABEL IV
UJI KEBAIKAN MODEL REDUKSI

	Chi-Square	Sig.
Model Regresi Logistik	14,053	0,003

Nilai signifikansi model regresi logistik lebih kecil dari taraf nyata 0,1 yaitu sebesar 0,003 yang berarti model reduksi yang diperoleh dapat menggambarkan data. Sehingga dilihat bahwa nilai wald dan signifikansi variabel prediktor pada model yang telah direduksi pada tabel berikut.

TABEL V
HASIL ANALISIS REGRESI LOGISTIK REDUKSI

Variabel Prediktor	β	SE (β)	Wald	Sig.
Umur (X_1)	-1,120	0,591	3,593	0,058
Konsumsi rokok (X_3)	1,512	0,525	8,289	0,004
Riwayat Penyakit Paru (X_4)	1,575	0,624	6,363	0,012
Konstanta	-1,576	0,606	6,757	0,009

Berdasarkan Tabel diatas diperoleh model regresi logistik:

$$\pi(x) = \frac{e^{(-1,576 - 1,120X_1 + 1,512X_3 + 1,575X_4)}}{1 + e^{(-1,576 - 1,120X_1 + 1,512X_3 + 1,575X_4)}} \quad (7)$$

dengan nilai logit $g(x)$, yaitu:

$$g(x) = -1,576 - 1,120X_1 + 1,512X_3 + 1,575X_4 \quad (8)$$

Model ini memiliki signifikansi yang lebih kecil dari taraf nyata 0,1. Sehingga variabel prediktor tersebut berpengaruh secara nyata terhadap terjadinya kanker paru-paru dan dapat diartikan juga bahwa faktor-faktor yang sangat mempengaruhi terjadinya kanker paru-paru adalah umur, konsumsi rokok dan riwayat penyakit paru-paru.

3) Interpretasi Koefisien

Untuk mengetahui seberapa besar pengaruh yang diberikan oleh model terbaik dapat dilihat dari nilai *odd ratio*. Nilai *odds ratio* dapat dilihat dari tabel adalah sebagai berikut.

TABEL VI
NILAI ODDS RATIO MODEL REGRESI LOGISTIK

Variabel Prediktor	Exp (β)
Konsumsi Rokok (X_3)	3,357
Riwayat Penyakit Paru (X_4)	3,402
Konstanta	0,130

Dalam menginterpretasikan seberapa besar pengaruh variabel prediktor terhadap terjadinya kanker paru-paru dapat dilihat dari nilai *odds ratio*. Berdasarkan tabel diatas diketahui bahwa variabel perokok berat memiliki nilai *odds ratio* sebesar 3,357 dan variabel riwayat penyakit paru dengan nilai *odds ratio* sebesar 3,402.

C. Pembahasan

Setelah dilakukan penerapan analisis regresi logistik binomial, didapatkan model yang menggambarkan faktor-faktor yang lebih berpengaruh terhadap pasien rawat inap penderita kanker paru-paru di RSUP Dr. M. Djamil Padang berdasarkan stadium kanker paru-paru. Dimana faktor-faktor yang mempengaruhinya adalah konsumsi rokok dan riwayat penyakit paru-paru. Ini dapat dibuktikan dengan *p-value* yang lebih kecil dari 0,1. Dari kedua faktor tersebut, faktor konsumsi rokok merupakan faktor yang paling berpengaruh terhadap pasien kanker paru-paru, karena faktor konsumsi rokok memiliki *p-value* yang paling kecil diantara kedua faktor berpengaruh tersebut yaitu 0,003. Artinya pasien yang perokok akan lebih rentan terkena kanker paru-paru.

Untuk mengetahui kecenderungan faktor resiko yang berpengaruh terhadap penderita kanker paru-paru berdasarkan stadium digunakan *odds ratio*. *Odds ratio* merupakan ukuran untuk mengetahui kecenderungan untuk mengalami suatu kejadian tertentu antara kategori yang satu dengan yang lain dalam satu variabel. Pada Tabel 14, dapat juga dilihat untuk variabel umur, variabel konsumsi rokok dan variabel riwayat penyakit paru-paru dengan nilai *odds ratio* masing-masingnya adalah 0,326, 4,537 dan 4,832. Dapat dijelaskan sebagai berikut:

1. Untuk variabel umur memiliki nilai *odds ratio* sebesar 0,326. Ini dapat kita artikan bahwa kecenderungan pasien yang berumur ≥ 40 tahun dibandingkan pasien yang berumur < 40 tahun adalah 0,326 kali untuk menderita kanker paru-paru stadium 3 dibandingkan dengan stadium 2.
2. Untuk variabel konsumsi rokok memiliki nilai *odds ratio* sebesar 4,537. Ini dapat kita artikan bahwa kecenderungan pasien yang perokok dibandingkan pasien yang tidak perokok adalah 4,537 kali untuk menderita kanker paru-paru stadium 3 dibandingkan dengan stadium 2.

3. Untuk variabel riwayat penyakit paru-paru memiliki nilai *odds ratio* sebesar 4,832. Ini dapat kita artikan bahwa kecenderungan pasien yang memiliki riwayat penyakit paru-paru dibandingkan yang tidak memiliki riwayat penyakit paru-paru adalah 4,832 kali untuk menderita kanker paru-paru stadium 3 dibandingkan dengan stadium 2.

SIMPULAN

Berdasarkan hasil penelitian dan pembahasan, didapat kesimpulan bahwa:

1. Faktor-faktor risiko yang paling mempengaruhi kanker paru-paru di RSUP Dr. M. Djamil Padang adalah umur, konsumsi rokok dan riwayat penyakit paru-paru.
2. Model regresi logistik untuk menggambarkan faktor-faktor risiko yang berpengaruh terhadap kanker paru-paru di RSUP Dr. M. Djamil Padang pada pasien rawat inap adalah sebagai berikut.

$$\pi(x) = \frac{e^{(-1,576 - 1,120X_1 + 1,512X_3 + 1,575X_4)}}{1 + e^{(-1,576 - 1,120X_1 + 1,512X_3 + 1,575X_4)}}$$

dengan nilai logit $g(x)$, yaitu:

$$g(x) = -1,576 - 1,120X_1 + 1,512X_3 + 1,575X_4$$

Dari model di atas, dapat diartikan bahwa peluang seseorang yang berumur < 40 tahun akan menderita kanker paru sebesar 1,120 kali lebih kecil dibandingkan seseorang yang berumur ≥ 40 tahun jika konsumsi rokok dan riwayat penyakit paru mereka sama. Peluang seseorang yang perokok akan menderita kanker paru-paru sebesar 1,512 kali lebih besar dibandingkan seseorang yang tidak perokok jika umur dan riwayat penyakit paru mereka sama. Dan peluang seseorang yang memiliki riwayat penyakit paru akan menderita kanker paru-paru sebesar 1,575 kali lebih besar dibandingkan seseorang yang tidak memiliki riwayat penyakit paru jika umur dan konsumsi rokok mereka sama.

REFERENSI

- [1] Agresti, A. (2007). *Categorical Data Analysis*. John Willey & Sons. New York.
- [2] Christine. (2011). *Pendeteksian Kanker Paru-Paru Dan Beberapa Faktor Yang Mempengaruhinya Menggunakan Transformasi Wavelet*. Skripsi Mahasiswa. Universitas Sumatera Utara.
- [3] Hosmer, D. W. dan Lemeshow, S., (2000), *Applied Logistic Regression*, John Willey and Sons, Inc. New York.
- [4] Perhimpunan Dokter Paru Indonesia. (2003). *KANKER PARU; Pedoman Diagnosis & Penatalaksanaan Asma di Indonesia*. *Jurnal Kesehatan*.
- [5] Sembiring, R.K. (1995). *Analisis Regresi*. ITB Bandung.
- [6] Sugiarto, dkk. (2003). *Teknik Sampling*. PT Gramedia Pustaka Umum. Jakarta.
- [7] Syahrudji Naseh, dkk. (1992). *Beberapa Model Statistik Dalam Penelitian Kesehatan/ Kedokteran*. *Artikel Media Litbangkes Vol II No.02/1992*.
- [8] Tabrani Rab. (2003). *Ilmu Penyakit Paru*. CV Trans Info Media. DKI Jakarta.