

# **PENERAPAN MODEL REGRESI LOGISTIK UNTUK MENGIDENTIFIKASI POTENSI FAKTOR RISIKO MALARIA DI SUMATERA UTARA**

**Riana Sekar Sari<sup>1\*)</sup>, Zahedi<sup>2)</sup>, Pasukat Sembiring<sup>3)</sup> & Asima  
Manurung<sup>4)</sup>**

*Program Studi Matematika, FMIPA, Universitas Sumatera Utara*

*\*email: [rianasekarsari@gmail.com](mailto:rianasekarsari@gmail.com)*

**Abstrak:** Pada penelitian ini memiliki tujuan untuk membentuk model matematika yang didasarkan pada faktor-faktor risiko yang mempunyai pengaruh terhadap terjangkitnya penyakit malaria di Provinsi Sumatera Utara. Data yang digunakan dalam penelitian adalah data sekunder yang diperoleh dari Dinas Kesehatan Provinsi Sumatera Utara. Sampel yang digunakan dalam penelitian ini sebanyak 368 data pasien. Pemodelan dari faktor-faktor risiko yang mempengaruhi terjangkitnya penyakit malaria menggunakan metode regresi logistik biner. Variabel respon yang digunakan bersifat dua kategori. Variabel prediktor yang akan diteliti sebanyak 5 variabel, meliputi umur, jenis kelamin, pekerjaan, klasifikasi penularan, dan jenis parasit. Estimasi parameter menggunakan metode maximum likelihood estimation. Dari 5 variabel prediktor yang diteliti, diperoleh hanya 2 variabel yang berpengaruh secara signifikan terhadap variabel respon, sehingga model regresi logistik biner yang diperoleh yaitu  $g(X)=-3,828+0,792X_1+1,450X_5$ . Dalam penelitian ini ketepatan klasifikasi diperoleh sebesar 92,4%.

**Kata Kunci:** Faktor Risiko, Malaria, Regresi Logistik Biner.

**Abstract:** In this study the aim was to form a mathematical model based on risk factors that have an influence on the spread of malaria in North Sumatra Province. The data used in this research is secondary data obtained from the North Sumatra Provincial Health Office. The sample used in this study was 368 patient data. Modeling of the risk factors that influence the spread of malaria using the binary logistic regression method. The response variable used is of two categories. The predictor variables that will be studied are 5 variables, including age, gender, occupation, classification of transmission, and type of parasite. Parameter estimation uses the maximum likelihood estimation method. Of the 5 predictor variables studied, only 2 variables were found to have a significant effect on the response variable, so that the binary logistic regression model

obtained was  $g(x) = -3,828 + 0,792X_1 + 1,450X_5$ . In this study the accuracy of classification was obtained at 92,4%.

**Keywords:** Risk factors, Malaria, Binary logistic regression.

## PENDAHULUAN

Informasi Dinas Kesehatan (2019) diketahui bahwa, penyakit menular yang diakibatkan karena parasit plasmodium yang berkembang biak dan hidup di dalam eritrosit manusia merupakan pengertian dari malaria. Sumber penyakit malaria adalah nyamuk malaria betina (*Anopheles*). Nyamuk ini bisa menggigit siapa saja, baik laki-laki dan perempuan maupun orang dari segala usia.

Indonesia dan seluruh dunia masih perlu mengatasi tingginya angka kesakitan malaria. Berdasarkan World Malaria Report yang di terbitkan tahun 2017, terdapat 219 juta kasus mengenai malaria yang menyumbang sebesar 435.000 jumlah kematian di seluruh dunia (WHO, 2018). Angka kematian malaria di Indonesia dilaporkan hingga 38.000 kematian diakibatkan malaria tingkat berat karena infeksi Plasmodium falciparum setiap tahunnya (Ika, 2015). Terdapat 415.140 kasus malaria di Indonesia pada tahun 2022. Data tersebut melonjak hingga 36.29% pada tahun sebelumnya yang sebanyak 304.607 kasus malaria (Darnius et al, 2023).

Provinsi Sumatera Utara masih terhitung kedalam kategori penyumbang malaria. Tahun 2013, terdapat 15.131 kasus malaria dengan kasus kematian sebanyak 14 orang (Hannum, 2020; Manumpa, 2016; Siahaan, 2019). Tercatat pada tahun 2020 berjumlah 998 kasus malaria dan meningkat di tahun 2021 yang berjumlah 2.529 kasus malaria di Sumatera Utara. Dari data pasien yang terjangkit malaria di Provinsi Sumatera Utara terdapat faktor-faktor yang berhubungan dengan terjangkitnya malaria dan akan dilihat faktor mana yang paling mempengaruhi dengan tujuan untuk mengkonsentrasikan pencegahan dini terhadap faktor yang paling besar pengaruhnya. Dalam penelitian ini akan digunakan ilmu statistika terapan untuk mengidentifikasi variabel yang mempunyai dampak terbesar dalam model analisis regresi logistik biner.

Metode analisis statistik yang dikenal sebagai regresi logistik biner dipakai untuk menguji relasi antara satu atau lebih variabel prediktor dan respons biner atau dikotomis. Pendekatan dasar untuk mengestimasi parameter pada model regresi logistik disebut metode *Maximum likelihood* (Tampil et al, 2017; Zeleke, 2015). Dalam penelitian ini akan tentukan model yang paling efektif untuk variabel respon kategori dalam model regresi logistik biner untuk kasus malaria.

Regresi logistik biner sering dikenal sebagai regresi logistik merupakan jenis regresi yang membantu dalam memodelkan hubungan antara variabel respon dan prediktor dengan syarat variabel respon yang bersifat biner atau dikotomis (misal: positif atau negatif, suka atau tidak suka) sementara untuk variabel prediktor bisa berbentuk jenis data nominal, interval, ordinal, atau rasio (Talombo et al, 2018; Yamin, Rachmach, & Kurniawan, 2011; Zaen, 2019). Berikut model dari regresi logistik biner lebih dari satu variabel prediktor adalah:

$$\pi(x_i) = \frac{e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}}}{1 + e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}}}$$

Keterangan:

$\pi(x_i)$  = Peluang kejadian dengan probabilitas  $0 \leq \pi(x) \leq 1$

$\beta_0$  = Konstanta

$x_{ki}$  = Variabel bebas ke-i ( $i = 1, 2, 3, \dots$ )

$\beta_k$  = Koefisien dari variabel bebas ke-i ( $i = 1, 2, 3, \dots$ )

Fungsi  $\pi(x_i)$  adalah fungsi non linear sehingga untuk mempermudah estimasi parameter regresi  $\pi(x_i)$  dari persamaan di atas, maka akan ditransformasikan untuk mendapatkan fungsi yang linear agar bisa dilihat hubungan antara variabel respon dan predictor (Algifari, 2000; Chin, 1998). Berikut model regresi logistik yang terbentuk yang dinyatakan sebagai  $g(x_i)$ :

$$g(x_i) = \ln \left[ \frac{\pi(x_i)}{1 - \pi(x_i)} \right] = (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})$$

Pada Metode *Maximum Likelihood* akan diperoleh hasil yang dapat memaksimalkan fungsi *likelihood*, sehingga dapat digunakan untuk mengestimasi parameter (Rizki, et al, 2015). Secara sistematis, fungsi *likelihood* dari model regresi logistik biner adalah sebagai berikut (David et al, 2000):

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}$$

Keterangan:

$y_i$  = Pengamatan saat variabel ke-i

$\pi(x_i)$  = Peluang saat variabel prediktor ke-i

Untuk memperoleh perhitungan yang mudah digunakan metode log *likelihood* seperti berikut:

$$L(\beta) = \sum_{j=0}^p \left[ \sum_{i=1}^n y_i x_{ij} \right] \beta_j - \sum_{i=1}^n n_i \ln \left[ 1 + e^{\sum_{j=0}^p \beta_j x_{ij}} \right]$$

Turunan pertama  $L(\beta)$  terhadap  $\beta$  disamakan dengan 0 untuk memberikan nilai interpretasi koefisien regresi logistik ( $\hat{\beta}$ ), tetapi hasil implisit masih dapat dihasilkan. Oleh sebab itu dibutuhkan suatu metode numerik agar dapat mencapai nilai estimasi parameter tersebut, yaitu menggunakan metode Newton Raphson. Untuk solusi iteratif persamaan nonlinear digunakan metode Newton-Raphson (Agresti, 2007).

*Cox and Snell's R Square* adalah pengukuran yang berfungsi untuk mencoba menirukan ukuran  $R^2$  pada *Multiple regression* berdasarkan cara estimasi *likelihood* serta nilai maksimal kurang dari 1, maka sulit untuk diinterpretasikan. Koefisien determinasi sering disebut dengan nilai  $R^2$  dalam *Multiple regression* dapat dihitung menggunakan *Nigelkerke R Square* (Ghozali, 2011). Nilai Koefisien Determinasi berguna untuk menggambarkan seberapa besar variabel prediktor bisa menerangkan

variabel respon. Biasa nilai *Nagelkerke R<sup>2</sup>* disajikan dalam bentuk persentase agar dapat memastikan seberapa baik variabel prediktor menjelaskan variabel respon.

## METODE PENELITIAN

Data diperoleh dari Dinas Kesehatan Provinsi Sumatera Utara. Data tersebut adalah data demografi pasien malaria yang ada di wilayah Sumatera Utara dalam periode bulan Januari sampai Desember 2022.

Penelitian ini mengambil 5 kabupaten yang memiliki jumlah orang yang terjangkit malaria cukup banyak dan memiliki data yang cukup lengkap yaitu kabupaten Labuhan Batu Utara, Batu Bara, Asahan, Serdang Bedagai, dan Kota Medan. Jadi total populasi dalam penelitian sebanyak 4,662 orang. Untuk mengambil sampel yang akan diteliti, penelitian ini menggunakan rumus Slovin dengan populasi 5 kabupaten, maka sampel diambil menggunakan rumus berikut:

$$n = \frac{N}{1+N.e^2}$$

Keterangan:

$N$  = Ukuran Populasi

$n$  = Ukuran sampel

$e$  = Persentase atau toleransi kesalahan (5%)

Berdasarkan rumus diatas jumlah sampel yang diambil yaitu:

$$\begin{aligned} n &= \frac{4,662}{1+4,662.(0,05)^2} \\ &= \frac{4,662}{12,655} \\ &= 368 \text{ sampel} \end{aligned}$$

Dari hasil perhitungan diatas sampel yang akan diambil yaitu 368 sampel.

**Tabel 1.** Variabel Penelitian

Kode	Nama Variabel	Kategori/Keterangan	Skala
$Y$	Malaria	0 : Malaria tanpa komplikasi 1 : Malaria komplikasi	Kategorik
$X_1$	Umur	0 : 12-25 tahun 1 : 26-45 tahun 2 : $\geq 46$	Kategorik
$X_2$	Jenis Kelamin	0 : Laki-laki 1 : Perempuan	Kategorik
$X_3$	Pekerjaan	0 : Tidak Bekerja 1 : Bekerja 2 : Pelajar	Kategorik
$X_4$	Klasifikasi Penularan	0 : Impor 1 : Indigenous	Kategorik
$X_5$	Jenis Parasit	0 : Vivax	Kategorik

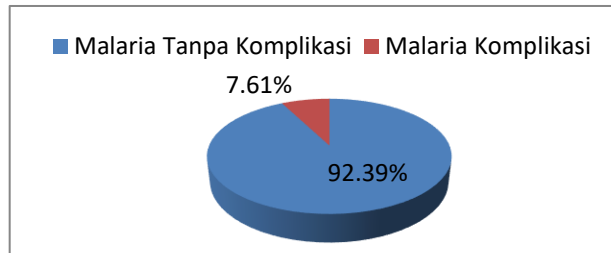
Adapun langkah-langkah penelitian yang akan dilakukan yaitu sebagai berikut:

1. Melakukan studi literatur dari beberapa buku dan jurnal mengenai analisis regresi logistik biner.
2. Mengidentifikasi masalah dengan membuat perumusan masalah hingga batasan masalah.
3. Pengumpulan data yang diambil dari Dinas Kesehatan Provinsi Sumatera Utara yang dimulai dari bulan Januari hingga Desember 2022.
4. Menginput data pasien malaria sebagai variabel (Y) dan faktor-faktor yang mempengaruhinya sebagai variabel (X).
5. Deskripsi data yang bertujuan untuk menganalisis data yang akan dipakai dengan menganalisis jumlah sampel data, jumlah yang terjangkit malaria dengan komplikasi, dan jumlah yang terjangkit malaria tanpa komplikasi.
6. Estimasi parameter yang bertujuan untuk menghilangkan variabel yang nilai signifikansi lebih besar dari 0,05. Proses tersebut diulang hingga tidak ada lagi variabel yang tidak signifikan yang tersisa dalam model yang telah dibuat.
7. Uji signifikansi  
Uji signifikansi berfungsi untuk menguji variabel independen (X) yang mempengaruhi variabel dependen (Y) dengan menjalankan dua pengujian yaitu (Suharjo, 2008):
  - a. Uji Serentak  
Uji serentak berguna untuk melihat hubungan diantara variabel independen dengan variabel dependen berpengaruh signifikan secara serentak. Uji serentak dilakukan dengan uji G menggunakan kriteria penolakan (tolak  $H_0$ ) jika  $G \geq X_{(ab,\alpha)}^2$  atau  $P_{value} \leq \alpha$ .
  - b. Uji Parsial  
Uji parsial berguna dalam melihat hubungan diantara variabel independen dengan variabel dependen berpengaruh signifikan secara parsial. Uji parsial dilakukan dengan uji *Wald* menggunakan kriteria penolakan (tolak  $H_0$ ) jika  $W^2 \geq x_{(df,\alpha)}^2$  atau  $P_{value} \leq \alpha$ .
8. Menentukan model terbaik.
9. Melakukan uji kesesuaian model.
10. Menentukan nilai *Odds Ratio*.
11. Melakukan interpretasi koefisien determinasi
12. Melakukan interpretasi ketetapan klasifikasi
13. Penarikan kesimpulan.

## **HASIL DAN PEMBAHASAN**

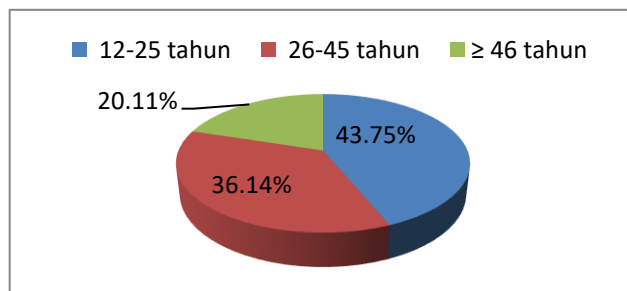
### **Deskripsi Data Malaria**

Berdasarkan data yang diperoleh peubah dependen yang diamati dalam penelitian ini adalah malaria tanpa komplikasi ( $Y = 0$ ) dan malaria komplikasi ( $Y = 1$ ) sedangkan untuk peubah independennya yaitu umur ( $X_1$ ), jenis kelamin ( $X_2$ ), pekerjaan ( $X_3$ ), klasifikasi penularan ( $X_4$ ), dan jenis parasit ( $X_5$ ).



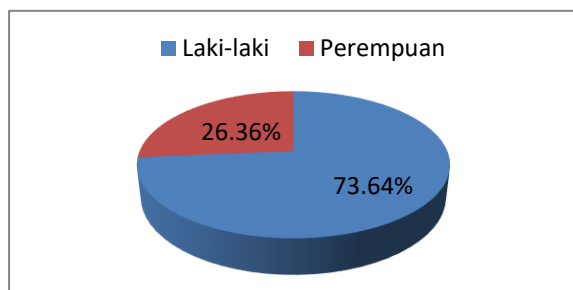
Gambar 1. Diagram Lingkaran Penyakit Malaria

Dapat dilihat berdasarkan jumlah sampel pasien malaria di Provinsi Sumatera Utara yaitu sebanyak 368 orang, terdapat sebesar 7,61% atau sebanyak 28 orang terjangkit penyakit malaria komplikasi dan terdapat sebesar 92,39% atau sebanyak 340 orang terjangkit penyakit malaria tanpa komplikasi.



Gambar 2. Diagram Lingkaran Umur

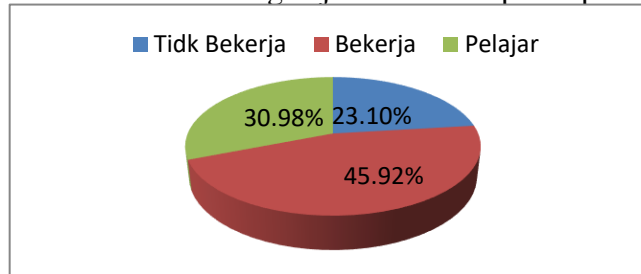
Dapat dilihat berdasarkan jumlah sampel pasien malaria di Provinsi Sumatera Utara yaitu sebanyak 368 orang, dalam rentang umur 12 – 25 tahun terdapat sebanyak 161 orang atau sebesar 43,75% terjangkit penyakit malaria. Dalam rentang umur 26 – 45 tahun terdapat sebanyak 133 orang atau sebesar 36,14% terjangkit penyakit malaria dan rentang umur  $\geq 46$  tahun terdapat sebanyak 74 orang atau sebesar 20,11% terjangkit penyakit malaria.



Gambar 3. Diagram Lingkaran Jenis Kelamin

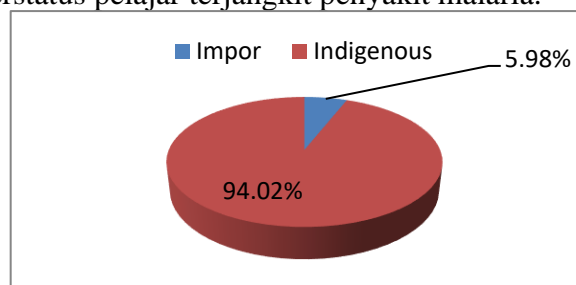
Dapat dilihat berdasarkan jumlah sampel pasien malaria di Provinsi Sumatera Utara yaitu sebanyak 368 orang, terdapat sebesar 73,64% atau sebanyak 271 orang

berjenis kelamin laki-laki terjangkit penyakit malaria sedangkan terdapat sebesar 26,36% atau sebesar 97 orang bejenis kelamin perempuan terjangkit penyakit malaria.



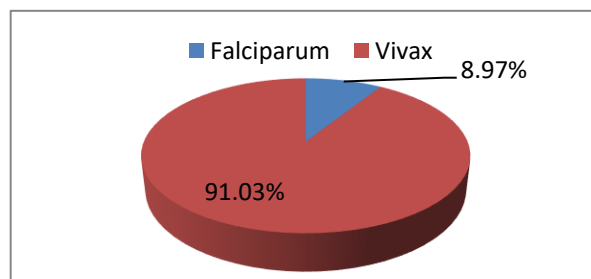
**Gambar 4.** Diagram Lingkaran Pekerjaan

Dapat dilihat berdasarkan jumlah sampel pasien malaria di Provinsi Sumatera Utara yaitu sebanyak 368 orang, terdapat sebesar 23,10% atau sebanyak 85 orang berstatus tidak bekerja terjangkit penyakit malaria. Terdapat sebesar 45,92% atau 169 orang berstatus bekerja terjangkit penyakit malaria dan terdapat sebesar 30,98% atau 114 orang berstatus pelajar terjangkit penyakit malaria.



**Gambar 5.** Diagram Lingkaran Klasifikasi Penularan

Dapat dilihat berdasarkan jumlah sampel pasien malaria di Provinsi Sumatera Utara yaitu sebanyak 368 orang, terdapat sebesar 5,98% atau sebanyak 22 orang tertular malaria dari luar wilayah (impor). Terdapat sebesar 94,02% atau 346 orang tertular malaria yang sumber penularannya berasal dari wilayah setempat (indigenous).



**Gambar 6.** Diagram Lingkaran Jenis Parasit

Dapat dilihat bahwa berdasarkan jumlah sampel pasien malaria di Provinsi Sumatera Utara yaitu sebanyak 368 orang, terdapat sebesar 8,97% atau sebanyak 33 orang terkena jenis parasit plasmodium falciparum. Terdapat sebesar 91,03% atau 335 orang terkena jenis parasit plasmodium vivax.

## Pembahasan

### Estimasi Parameter

Untuk mendapatkan model regresi logistik biner akan di cari nilai estimasi parameter dari  $\beta_1, \beta_2, \beta_3, \beta_4$ , dan  $\beta_5$ . Dengan menghilangkan faktor-faktor yang memiliki nilai signifikansi tinggi, variabel-variabel tersebut akan dievaluasi lebih dulu untuk mengidentifikasi variabel-variabel mana yang memiliki pengaruh. Apabila nilai signifikansi suatu variabel lebih rendah dari tingkat signifikansi yang sudah ditetapkan, yaitu 5% atau 0,05 maka proses eliminasi berakhir. Untuk mencari penaksiran tersebut akan digunakan bantuan *Software SPSS 24*.

**Tabel 2.** Penaksiran Model Regresi Logistik Biner

Parameter	Nilai Taksiran $\beta$	S.E.	Df	Sig.
<b>Umur</b>	0,792	0,315	1	0,012
<b>Jenis Kelamin</b>	0,396	0,487	1	0,417
<b>Pekerjaan</b>	0,356	0,383	1	0,353
<b>Kalsifikasi Penularan</b>	-0,104	0,784	1	0,895
<b>Jenis Parasit</b>	1,450	0,509	1	0,004

Berdasarkan tabel diatas dapat dilihat untuk variabel jenis kelamin, pekerjaan, dan klasifikasi penularan memiliki nilai signifikansi yang lebih besar dari taraf signifikansi yang sudah ditetapkan yaitu 0,05 artinya ketiga variabel tersebut tidak mempunyai pengaruh yang signifikan terhadap terjangkitnya penyakit malaria, sehingga variabel tersebut dieliminasi.

**Tabel 3.** Penaksiran Model Regresi Logistik Biner Tanpa Variabel  $X_4, X_2$ , dan  $X_3$

Parameter	Nilai Taksiran $\beta$	S.E.	Df	Sig.
<b>Umur</b>	0,792	0,315	1	0,012
<b>Jenis Parasit</b>	1,450	0,509	1	0,004

Dapat dilihat pada tabel diatas hanya tersisa variabel umur dan jenis parasit yang memiliki nilai signifikansi lebih rendah dari taraf signifikansi yang sudah ditetapkan yaitu 0,05 artinya kedua variabel tersebut mempunyai pengaruh yang signifikan terhadap terjangkitnya penyakit malaria, maka variabel tersebut akan dimasukkan ke dalam model regresi logistik biner.

### Uji Signifikansi Parameter

Uji serentak maupun parsial adalah cara untuk uji signifikansi parameter. Uji ini berguna untuk melihat signifikansi dari semua variabel prediktor dan variabel respon

#### Uji Serentak

Hipotesis yang digunakan:

$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$  (tidak ada variabel prediktor yang berpengaruh)

$H_1$ : paling sedikit koefisien  $\beta_j \neq 0$  (ada variabel prediktor yang berpengaruh)

Taraf signifikansi:  $\alpha = 0,05$



Kriteria penolakan: (tolak  $H_0$ ) jika nilai  $G \geq X^2_{(db,\alpha)}$  atau  $P_{value} \leq \alpha$ .

**Tabel 4.** Uji Serentak

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
<b>44,044</b>	0,088	0,449

Hasil uji serentak yang memakai uji *Likelihood Ratio Test* yaitu -2 Log *likelihood* atau uji G sebesar 44,044 lebih besar dari nilai *chi square* tabel (4 ; 0,05) = 9,488 dengan kriteria penolakan (tolak  $H_0$ ) jika nilai  $G \geq X^2_{(db,\alpha)}$  atau  $P_{value} \leq \alpha$  maka diperoleh keputusan menolak  $H_0$ . Ditarik kesimpulan bahwasanya variabel prediktor mempengaruhi secara serentak pada variabel respon.

• **Uji Parsial**

Hipotesis yang digunakan:

$H_0: \beta_j = 0$  (variabel prediktor ke- $j$  tidak mempengaruhi variabel respon)

$H_1: \beta_j \neq 0$  (variabel prediktor ke- $j$  ada yang mempengaruhi variabel respon)

Taraf signifikansi:  $\alpha = 0,05$

Kriteria penolakan: tolak  $H_0$  apabila nilai  $W^2 \geq x^2_{(df,\alpha)}$  atau  $P_{value} \leq \alpha$ ,  $df = 1$ .

**Tabel 5.** Uji Parsial

	B	S.E.	Wald	Df	Sig.	Exp (B)
<b>Umur</b>	0,792	0,315	6,302	1	0,012	2,208
<b>Jenis Kelamin</b>	0,396	0,487	0,660	1	0,417	1,486
<b>Pekerjaan</b>	0,356	0,383	0,862	1	0,353	1,427
<b>Klasifikasi Penularan</b>	-0,104	0,784	0,017	1	0,895	0,902
<b>Jenis Parasit</b>	1,450	0,509	8,132	1	0,004	4,025
<b>Constant</b>	-3,828	1,021	14,046	1	0,000	0,022

Statistik uji parsial yang menggunakan uji *wald* memiliki taraf signifikansi 0,05 dan mengikuti distribusi *chi square* dengan derajat kebebasan 1 memberikan nilai 3,841 untuk *chi Square* tabel.

1. Variabel Umur ( $X_1$ )

Pada tabel diperoleh hasil uji *wald* variabel umur ( $X_1$ ) adalah 6,302 lebih besar dari nilai *chi square* yaitu 3,841. Pada nilai signifikansi diperoleh 0,012 yang lebih rendah dari taraf signifikansi 0,05 artinya keputusan menolak  $H_0$  sebab variabel umur berpengaruh terhadap terjangkitnya penyakit malaria.

2. Variabel Jenis Kelamin ( $X_2$ )

Pada tabel diperoleh hasil uji *wald* variabel jenis kelamin ( $X_2$ ) adalah 0,660 lebih rendah dari nilai *chi square* yaitu 3,841. Pada nilai signifikansi diperoleh 0,417 yang lebih besar dari taraf signifikansi 0,05 artinya keputusan terima  $H_0$  sebab

variabel jenis kelamin tidak memiliki pengaruh terhadap terjangkitnya penyakit malaria.

3. Variabel Pekerjaan ( $X_3$ )

Pada tabel diperoleh hasil uji *wald* variabel pekerjaan ( $X_3$ ) adalah 0,862 yang lebih rendah dari nilai *chi square* yaitu 3,841. Pada nilai signifikansi diperoleh 0,353 yang lebih besar dari taraf signifikansi 0,05 artinya keputusan terima  $H_0$  sebab variabel pekerjaan tidak memiliki pengaruh terhadap terjangkitnya penyakit malaria.

4. Variabel Klasifikasi Penularan ( $X_4$ )

Pada tabel diperoleh hasil uji *wald* variabel klasifikasi penularan ( $X_4$ ) adalah 0,017 yang lebih rendah dari nilai *chi square* yaitu 3,841. Pada nilai signifikansi diperoleh 0,895 yang lebih besar dari taraf signifikansi 0,05 artinya keputusan terima  $H_0$  sebab variabel klasifikasi penularan tidak memiliki pengaruh terhadap terjangkitnya penyakit malaria.

5. Variabel Jenis Parasit ( $X_5$ )

Pada tabel diperoleh hasil uji *wald* variabel jenis parasit ( $X_5$ ) adalah 8,132 yang lebih besar dari nilai *chi square* yaitu 3,841. Pada nilai signifikansi diperoleh 0,004 yang lebih rendah dari taraf signifikansi 0,05 artinya keputusan menolak  $H_0$  sebab variabel jenis parasit memiliki pengaruh terhadap terjangkitnya penyakit malaria.

**Model Regresi Logistik Biner**

Setelah melakukan proses eliminasi terhadap variabel prediktor yang memiliki nilai signifikansi tinggi atau lebih besar dari taraf signifikansi yang sudah ditentukan yaitu 0,05 maka variabel tidak baik untuk dimasukkan ke dalam model, maka didapat hasil penaksiran parameter sebagai berikut:

**Tabel 6.** Penaksiran Model Regresi Logistik Biner Tanpa Variabel  $X_4$ ,  $X_2$ , dan  $X_3$

Parameter	Nilai Taksiran $\beta$	S.E.	Df	Sig.
<b>Umur</b>	0,792	0,315	1	0,012
<b>Jenis Parasit</b>	1,450	0,509	1	0,004
<b>Constant</b>	-3,828	1,021	1	0,000

Model yang akan terbentuk berdasarkan hasil penaksiran parameter diatas adalah:

$$\pi(x_i) = \frac{e^{-3,828+0,792X_1+1,450X_5}}{1+e^{-3,828+0,792X_1+1,450X_5}}$$

$$\{\pi(x_i)\}\{1 + e^{-3,828+0,792X_1+1,450X_5}\} = e^{-3,828+0,792X_1+1,450X_5}$$

$$\{\pi(x_i)\}\{\pi(x_i) + e^{-3,828+0,792X_1+1,450X_5}\} = e^{-3,828+0,792X_1+1,450X_5}$$

$$\pi(x_i) = e^{-3,828+0,792X_1+1,450X_5} - \pi(x_i)e^{-3,828+0,792X_1+1,450X_5}$$

$$\pi(x_i) = \{1 - \pi(x_i)\}e^{-3,828+0,792X_1+1,450X_5}$$

$$\frac{\pi(x_i)}{1-\pi(x_i)} = e^{-3,828+0,792X_1+1,450X_5}$$

$$\ln \frac{\pi(x_i)}{1-\pi(x_i)} = \ln e^{-3,828+0,792X_1+1,450X_5}$$

$$\ln \frac{\pi(x_i)}{1-\pi(x_i)} = -3,828 + 0,792X_1 + 1,450X_5$$

Sehingga diperoleh model regresi logistik biner sebagai berikut:

$$g(x) = -3,828 + 0,792X_1 + 1,450X_5$$

Berikut interpretasi dari model yang dihasilkan yaitu:

1. Dari model diperoleh bahwa nilai intersep adalah -3,828 yang memiliki arti apabila saat semua variabel bernilai 0, maka probabilitas terjangkitnya penyakit malaria mengalami penurunan sebesar -3,828.
2. Pada variabel umur ( $X_1$ ) diperoleh nilai 0,792 yang memiliki arti apabila variabel umur meningkat, maka probabilitas terjangkitnya penyakit malaria mengalami kenaikan sebesar 0,792.
3. Pada variabel jenis parasit ( $X_5$ ) diperoleh nilai 1,450 yang memiliki arti apabila variabel jenis parasit meningkat, maka probabilitas terjangkitnya penyakit malaria mengalami kenaikan sebesar 1,450.

### Uji Kesesuaian Model

Uji kesesuaian model dilakukan dengan uji Hosmer dan Lemeshow menggunakan hipotesis seperti berikut:

$H_0$  : Model sudah sesuai dengan nilai observasinya

$H_1$ : Ada perbedaan signifikan antara model dengan nilai observasinya.

Taraf signifikansi:  $\alpha = 0,05$

Kriteria penolakan: Tolak  $H_0$  apabila  $\hat{C} \geq \chi_{df,\alpha}^2$  atau  $P_{value} \leq \alpha$

**Tabel 7. Hosmer and Lemeshow Test**

Chi Square	Df	Sig.
3,122	7	0,874

Berdasarkan tabel didapat nilai dari *chi square* hitung adalah 3,122 sementara nilai dari *chi square* tabel dengan derajat kebebasan sebesar 7 dan taraf signifikansi senilai 0,05 dihasilkan nilai dari *chi square* tabel adalah 14,017 maka  $3,122 < 14,017$ . Dari hasil tersebut dapat diambil keputusan bahwa terima  $H_0$  yang artinya model telah sesuai berdasarkan nilai observasinya atau tidak ada beda signifikansi antara model dengan nilai observasinya.

### Odds Ratio

*Odds ratio* berguna untuk menggambarkan koefisien variabel prediktor serta menentukan kecenderungan antara variabel prediktor dengan variabel respon.

**Tabel 8.** Odds Ratio Variabel  $Y$  pada  $X_1$

Umur ( $X_1$ )	Derajat Malaria ( $Y$ )		Total
	Malaria Tanpa Komplikasi	Malaria Komplikasi	
12 – 25 tahun	154	7	161
26 – 45 tahun	122	11	133
$\geq 46$ tahun	64	10	74
Total	340	28	368

Berdasarkan tabel diatas menunjukkan bahwa tabel yang digunakan adalah tabel 3x2 maka *Odds Ratio* dihitung secara manual seperti berikut:

$$\begin{aligned}
 \text{Odds ratio} &= \frac{\text{odds ratio (umur 12-25 tahun)}}{\text{odds ratio (umur 26-45 tahun)}} \\
 \psi &= \frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]} \\
 &= \frac{154/122}{7/11} \\
 &= \frac{1,262}{0,636} \\
 &= 1,983 \\
 \text{Odds ratio} &= \frac{\text{odds ratio (umur 12-25 tahun)}}{\text{odds ratio (umur } \geq 46 \text{ tahun)}} \\
 \psi &= \frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]} \\
 &= \frac{154/64}{7/10} \\
 &= \frac{2,406}{0,7} \\
 &= 3,437
 \end{aligned}$$

Dari hasil perhitungan *odds ratio* tersebut diperoleh bahwa pada rentang umur 12 – 25 tahun mempunyai kecenderungan terjangkau penyakit malaria tanpa komplikasi 1,983 lebih besar daripada rentang umur 26 - 45 tahun. Sedangkan diperoleh juga bahwa pada rentang umur 12 – 25 tahun mempunyai kecenderungan terjangkau penyakit malaria tanpa komplikasi 3,437 lebih besar daripada rentang umur  $\geq 46$  tahun.

**Tabel 9.** Odds Ratio Variabel  $Y$  pada  $X_5$

Jenis Parasit ( $X_5$ )	Derajat Malaria ( $Y$ )		Total
	Malaria Tanpa Komplikasi	Malaria Komplikasi	
Vivax	314	21	335
Falciparum	26	7	33
Total	340	28	368

$$\begin{aligned}
 \text{Odd ratio} &= \frac{\text{odds ratio (vivax)}}{\text{odds ratio (falciparum)}} \\
 \psi &= \frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]} \\
 &= \frac{314/26}{21/7} \\
 &= \frac{12,076}{3} \\
 &= 4,025
 \end{aligned}$$

Dari hasil perhitungan *odds ratio* tersebut diperoleh pada jenis parasit plasmodium vivax mempunyai kecenderungan menjangkitkan penyakit malaria tanpa komplikasi 4,025 lebih besar daripada jenis parasit *plasmodium falciparum*.

### Koefisien Determinasi

Nilai *Cox and Snell* dan *Nagelkerke* yang ditampilkan pada tabel berikut dapat digunakan untuk menentukan koefisien determinasi:

**Tabel 10.** Koefisien Determinasi

<i>Cox and Snell R Square</i>	<i>Nagelkerke R Square</i>
<b>0,088</b>	0,449

Pada tabel diatas terlihat nilai dari *Cox and Snell* yaitu 0,088 dengan nilai *Nagelkerke R Square* yaitu 0,449 atau 44,9% artinya bahwa kemampuan variabel prediktor untuk menjelaskan variabel respon moderat (cukup kuat) dan terdapat 55,1% lainnya dijelaskan oleh faktor-faktor yang tidak di uji dalam penelitian ini.

### Ketepatan Klasifikasi

Berikut ini merupakan tabel ketepatan klasifikasi dari model faktor-faktor yang mempengaruhi terjangkitnya penyakit malaria tanpa komplikasi berdasarkan hasil observasi dan prediksi di Provinsi Sumatera Utara.

**Tabel 11.** Ketepatan Klasifikasi

Observasi	Prediksi		Presentase Benar
	Derajat Malaria Tanpa Komplikasi	Derajat Malaria Komplikasi	
Derajat Malaria Tanpa Komplikasi	340	0	100%
Derajat Malaria Komplikasi	28	0	7,60%
<b>Total Presentase</b>			<b>92,4%</b>

Berdasarkan tabel diatas data observasi dari 368 sampel orang yang terjangkit penyakit malaria yang dinyatakan terkena malaria komplikasi sebanyak 28 orang dan yang dinyatakan terkena malaria tanpa komplikasi sebanyak 340 orang. Namun hasil prediksi menyatakan yang terkena malaria komplikasi sebanyak 0 orang dan yang

dinyatakan terkena malaria tanpa komplikasi sebanyak 340 orang. Terdapat kesalahan klasifikasi sebanyak 28 orang yang seharusnya terkena malaria komplikasi, maka akan digunakan rumus *Apparent Error Rate* (APER) untuk menghitung besar kesalahan yang diprediksi yaitu:

$$\begin{aligned} APER &= \frac{n_{1M}+n_{2M}}{n_{1C}+n_{1M}+n_{2M}+n_{2C}} \times 100\% \\ &= \frac{0+28}{340+0+28+0} \times 100 \\ &= \frac{28}{368} \times 100 \\ &= 7,60\% \end{aligned}$$

sehingga diperoleh:  $100\% - 7,60\% = 92,4\%$ .

Jadi, ketepatan model regresi logistik biner cukup baik sebab mampu memprediksi 92,4% kondisi yang terjadi.

## SIMPULAN

Berdasarkan hasil dan pembahasan yang dilakukan, maka dapat diperoleh kesimpulan sebagai berikut:

1. Terdapat 5 variabel penelitian sebagai faktor yang mempunyai pengaruh pada penyakit malaria yaitu umur ( $X_1$ ), jenis kelamin ( $X_2$ ), pekerjaan ( $X_3$ ), klasifikasi penularan ( $X_4$ ), dan jenis parasit ( $X_5$ ). Dari kelima variabel yang diteliti terdapat hanya ada 2 variabel yang berpengaruh terhadap terjangkitnya penyakit malaria yaitu variabel umur ( $X_1$ ) dan jenis parasit ( $X_5$ ). Variabel tersebut dianggap sebagai pemicu terjangkitnya penyakit malaria.
2. Model regresi logistik biner yang diperoleh dari 2 variabel yang berpengaruh yaitu:  $g(X) = -3,828 + 0,792X_1 + 1,450X_5$ .

## SARAN

Hasil penelitian ini dapat digunakan sebagai informasi bagi Pemerintah Provinsi Sumatera Utara khususnya terkait dengan instansi kesehatan untuk menekan variabel yang berpengaruh terhadap terjangkitnya penyakit malaria. Bagi peneliti selanjutnya agar dapat memasukkan lebih banyak lagi variabel risiko yang menyebabkan penyebaran penyakit malaria atau menggunakan kasus yang berbeda serta diharapkan menggunakan metode pemodelan lain agar menjadi bahan perbandingan dengan metode lainnya.

## DAFTAR PUSTAKA

- Agresti, A. (2007). *An Introduction To Categorical Data Analysis*. Canada: John Wiley and Sons, Inc., Hoboken, New Jersey.
- Algifari. (2000). *Analisis Regresi Teori, Kasus, dan Solusi*. Yogyakarta: BPFY-YOGYAKARTA.

- Chin, W. W. (1998). *The Partial Least Squares Approach to Structural Equation Modeling. Modern Methods for Business Research*. London: Mahwah, New Jersey.
- Darnius, O., & Siahaan, D. (2023). Modelling of risk factors that influence malaria infection using Binary Logistic Regression. *2nd TALENTA-International Conference on Science Technology*, 1-7.
- David W Hosmer, J., & Lemeshow, S. (2000). *Applied Logistic Regression*. Canada: John Wiley dan Sons, Inc.
- Dinas Kesehatan. (2019). *Profil Kesehatan Provinsi Sumatera Utara*. [www.dinkes.sumutprov.go.id](http://www.dinkes.sumutprov.go.id).
- Ghozali, I. (2011). *Aplikasi Analisis Multivariate dengan Program SPSS*. Semarang: Badan Penerbit Universitas Diponegoro.
- Hannum, L. (2020). *Faktor Penyebab Malaria Pada Ibu Hamil di Wilayah Kerja Puskesmas Panyambungan Jae Mandailing Natal*. [Skripsi].
- Ika. (2015, Desember 31). *Kasus Malaria di Indonesia Masih Tinggi*. Dipetik Mei 14, 2023, dari [www.ugm.ac.id](http://www.ugm.ac.id): <https://www.ugm.ac.id/id/berita/10980-kasus-malaria-di-indonesia-masih-tinggi>
- Manumpa, S. (2016). Pengaruh Faktor Demografi dan Riwayat Malaria Terhadap Kejadian Malaria. *Jurnal Berkala Epidemiologi*, 4(3), 338-348.
- Rizki, F., Widodo, D. A., & Wulandari, S. P. (2015). Faktor Risiko Penyakit Anemia Gizi Besi pada Ibu Hamil di Jawa Timur Menggunakan Analisis Regresi Logistik. *Jurnal Sains dan Seni ITS*, 4(2), 2337-3520.
- Siahaan, D. R. (2019). *Pemodelan Faktor-Faktor Risiko yang Memengaruhi Infeksi Malaria dengan Regresi Logistik Biner*. [Skripsi]
- Suharjo, B. (2008). *Analisis Regresi Terapan dengan SPSS*. Yogyakarta: Graha Ilmu.
- Talombo, U. B., Munir, M. A., & Lintin, G. (2018). Analisis Faktor Risiko Utama Terhadap Kejadian Malaria Di Wilayah Puskesmas Kampung Baru Luwuk Tahun 2013-2015.
- Tampil, Y. A., Komalig, H., & Langi, Y. (2017). Analisis Regresi Logistik Untuk Menentukan Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa FMIPA Universitas Sam Ratulangi Manado. *JdC*, 6(2), 56-62.
- WHO. (2018). *The World Malaria Report 2018*. Geneva: World Health Organization.
- Yamin, S., Rachmach, L. A., & Kurniawan, H. (2011). *Regresi dan Korelasi Dalam Genggaman Anda*. Jakarta: Salemba Empat.
- Zaen, N. J. (2019). *Diagnosis Penyakit Stroke dengan Metode Regresi Logistik Biner*. [Skripsi].
- Zelege, G. T., Egide, H., & Francois, T. (2015). Application of Regression Model to Identify Potential Risk Factors of Malaria in Rwanda using 2010 Demographic and Health Survey. *International Journal of Applied Science and Mathematics*, 2(3), 2894-2394.