

Perbandingan Metode *Propensity Score Matching-Support Vector Machine* dan *Propensity Score Matching-Regresi Logistik Biner* Pada Kasus HIV/AIDS

Silviatul Hasanah^{1*}, Bambang Widjarnarko Otok², Adeni³
*e-mail: silviatulh@walisongo.ac.id

¹*Program Studi Manajemen Pendidikan Islam, Fakultas Ilmu Tarbiyah dan Keguruan
Universitas Islam Negeri Walisongo Semarang*

²*Departemen Statistika, Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember Surabaya*

³*Program Studi Komunikasi Penyiaran Islam, Fakultas Dakwah dan Komunikasi
Universitas Islam Negeri Walisongo Semarang*

ABSTRACT

In the health field, Randomized Controlled Trials (RCT) cannot be done because it relates to human life. The covariates variable as well as the presence of confounding variable in non-experimental research are usually not balanced between the treatment and control groups that cause the estimated treatment effect to be biased. One of the appropriate methods to overcome the biased treatment effect is the Propensity Score Matching (PSM). This study will compare the performance of the Propensity Score Matching-Support Vector Machine (PSM-SVM) method and the Propensity Score Matching-Binary Logistic Regression (PSM-RLB) method in cases of HIV/AIDS opportunistic infection. The confounding variable used was the opportunistic infection variable. The data used were data of HIV/AIDS patients treated at the Grati Public Health Center, Pasuruan Regency in 2016. The results showed that the variable of ARV therapy had a significant effect on opportunistic infections in HIV/AIDS patients. If we look at how much bias can be reduced, PSM-SVM is able to reduce bias more than the PSM-RLB method, which is 60.25%. However, the PSM-SVM method can produce a bias value (after matching) that is smaller than the PSM-RLB method, which is 0.044.

Keywords: Bias, Confounding, HIV/AIDS, PSM-SVM, PSM-RLB.

ABSTRAK

Dalam bidang kesehatan, *Randomized Controlled Trials* (RCT) tidak dapat dilakukan karena berhubungan dengan nyawa manusia. Kovariat pada penelitian *non-experimental* biasanya tidak seimbang antara kelompok *treatment* dan control yang menyebabkan estimasi efek perlakuan menjadi bias. Selain itu, adanya variabel *confounding* juga mengakibatkan estimasi efek perlakuan bias. Metode yang sesuai untuk mengatasi efek perlakuan yang bias adalah metode *Propensity Score* (PS). Salah satu metode yang dikembangkan dari *propensity score* adalah *Propensity Score Matching* (PSM). Pada penelitian ini, peneliti akan membandingkan performa metode *Propensity Score Matching-Support Vector Machine* (PSM-SVM) dan metode *Propensity Score Matching-Regresi Logistik Biner* (PSM-RLB) pada kasus infeksi oportunistik HIV/AIDS. Variabel *confounding* yang digunakan dalam penelitian ini adalah variabel infeksi oportunistik. Data yang digunakan dalam penelitian ini adalah data pasien HIV/AIDS yang di rawat di Puskesmas Grati, Kabupaten Pasuruan pada tahun 2016. Hasil analisis PSM-SVM dan metode PSM-RLB menunjukkan bahwa

variabel pemberian terapi ARV berpengaruh secara signifikan terhadap infeksi oportunistik pada pasien HIV/AIDS. Jika dilihat seberapa bias yang dapat tereduksi, PSM-SVM mampu mereduksi bias lebih besar dibanding metode PSM-RLB yaitu sebesar 60,25%. Namun, metode PSM-SVM dapat menghasilkan nilai bias (setelah *matching*) yang lebih kecil dibanding metode PSM-RLB yaitu sebesar 0,044.

Kata Kunci: Bias, Confounding, HIV/AIDS, PSM-SVM, PSM-RLB

PENDAHULUAN

Penelitian merupakan suatu proses pengumpulan informasi yang dilakukan secara sistematis untuk menyelesaikan suatu masalah yang diteliti. Secara umum penelitian dibagi menjadi dua yaitu penelitian *experimental* dan *non-experimental* (observasi). Penelitian *experimental* merupakan penelitian yang dilakukan secara acak terhadap kelompok *treatment* atau kelompok kontrol, dengan tujuan untuk mengetahui akibat yang ditimbulkan dari suatu perlakuan yang diberikan secara sengaja oleh peneliti, sedangkan penelitian *non-experimental* merupakan penelitian empiris dari efek perlakuan ketika kelompok *treatment* dan kontrol tidak layak dilakukan pengacakan. Penelitian *non-experimental* disusun menyerupai percobaan acak sederhana, sehingga pemahaman mengenai peran pengacakan dalam percobaan ini dianggap penting (Rosenbaum & Rubin, 1985).

Analisis *propensity score* merupakan metode statistika yang dikembangkan untuk mengestimasi efek perlakuan dengan data *non-experimental*. Penelitian *non-experimental* dapat dilakukan diberbagai bidang salah satunya bidang kesehatan, misal melihat pengaruh pemberian terapi ARV dan pendampingan terhadap infeksi oportunistik pada pasien HIV/AIDS. Pada penelitian *non-experimental*, *Randomized Controlled Trials* (RCT) tidak dapat dilakukan karena berhubungan dengan nyawa manusia. Data yang biasa digunakan dalam penelitian *non-experimental* ialah data survei, sensus, administratif, rekam medis atau data lainnya yang tidak mungkin dilakukan pengacakan (Guo & Fraser, 2009).

Kovariat pada penelitian *non-experimental* biasanya tidak seimbang antara kelompok *treatment* dan kontrol. Ketidakseimbangan ini menyebabkan estimasi efek perlakuan menjadi bias. Selain itu, adanya variabel *confounding* juga mengakibatkan estimasi efek perlakuan bias. Metode *propensity score* merupakan metode yang cocok untuk mengatasi efek perlakuan yang bias (Rosenbaum & Rubin, 1983).

Salah satu pengembangan dari *propensity score* adalah metode *Propensity Score Matching* (PSM). PSM merupakan metode yang digunakan untuk menyeimbangkan estimasi kelompok perlakuan dan kontrol dengan cara mencocokkan kelompok perlakuan dan kontrol berdasarkan nilai estimasi *propensity score* yang paling mirip, kemudian membuang unit yang tidak cocok. Algoritma *matching* yang digunakan dalam penelitian ini adalah *Nearest neighbor matching* (Becker & Ichino, 2002). Penelitian sebelumnya tentang metode *Propensity Score Matching* (PSM) telah dilakukan oleh Otok, dkk., (2017) yaitu mengestimasi (PSM) menggunakan metode regresi logistik.

Salah satu metode klasik yang sering digunakan untuk mengestimasi *Propensity score* adalah regresi logistik (Hosmer & Lemeshow, 2000). Namun, metode tersebut memiliki beberapa kelemahan yaitu hanya dapat digunakan pada data linier dan tidak dapat digunakan pada data berdimensi tinggi sehingga dalam penelitian ini diusulkan menggunakan metode *machine learning* untuk mengestimasi *propensity score*. Metode

machine learning yang digunakan adalah *Support Vector Machine* (SVM). SVM memiliki beberapa kelebihan diantaranya adalah SVM memiliki sifat global optima, fungsi kernel pada SVM dapat digunakan untuk data tidak liner karena secara implisit memuat transformasi *non-linear* dan SVM bisa digunakan untuk mengatasi masalah data berdimensi tinggi dan sampel *training* yang sedikit (Pal & Mather, 2005),(Hardle, Prastyo, & Hafner, 2014). SVM bisa digunakan untuk data berdimensi tinggi dan untuk data sampling kecil. Yao, *et al.*,(2012) memprediksi *B-cell epitope* menggunakan metode kombinasi *Tri-peptide similarity and Propensity score* (SVMTriP), AAP dan BCPred. Didapatkan hasil bahwa *propensity score* dengan SVM merupakan metode terbaik dengan hasil sensitivitas sebesar 80,1%. Dalam penelitian ini peneliti ingin meneliti apakah metode *machine learning* lebih baik daripada metode klasik. Metode *machine learning* yang digunakan adalah metode *propensity score matching-support vector machine* dan metode klasik yang digunakan adalah metode *propensity score matching -regresi logistik biner*.

Human Immunodeficiency Virus (HIV) merupakan virus yang menyebar melalui cairan tubuh tertentu yang menyerang system kekebalan tubuh, sehingga seseorang yang terkena virus tersebut akan mengalami penurunan kekebalan tubuh. Seiring waktu, penderita HIV menjadi rentan terhadap berbagai infeksi dan penyakit. HIV akan berkembang menjadi *Acquired Immuno Defisiensi Syndrome* (AIDS) ketika system kekebalan tubuh telah rusak parah (Center for Disease Control and Prevention, 2018). Dampak paling fatal dari HIV/AIDS adalah kematian. Kematian yang paling banyak terjadi akibat adanya Infeksi Oportunistik (IO) pada pasien penderita HIV/AIDS. Adapun infeksi oportunistik yang berhubungan dengan HIV diantaranya *tuberculosis* (TB), *cryptococcal meningitis*

(CM), and *Pneumocystis jirovecii pneumonia* (PCP) (Fenner et al., 2013).

Pada penelitian ini, peneliti menggunakan data sekunder yaitu data pasien HIV/AIDS yang di rawat di Puskesmas Grati, Kabupaten Pasuruan, Provinsi Jawa Timur pada tahun 2016. Tujuan dari penelitian ini adalah membandingkan performa metode *propensity score matching -support vector machine* dan metode *propensity score matching-regresi logistik biner* untuk kasus infeksi oportunistik pada pasien HIV/AIDS.

BAHAN DAN METODE

Data yang digunakan dalam penelitian ini merupakan data sekunder yang berasal dari data survei dan rekam medis pasien HIV/AIDS pada tahun 2016 di Puskesmas Grati, Kabupaten Pasuruan, Jawa timur. Jumlah responden yang digunakan sebanyak 150 pasien. Variabel penelitian yang digunakan dalam penelitian ini disajikan pada Tabel 1.

Tabel 1. Variabel dan Skala Data

Variabel	Deskripsi	Keterangan
Y	Infeksi Oportunistik	1= Tidak ada IO
		0= Ada IO
X ₁	Usia	Tahun
X ₂	Pengetahuan	1= Baik
		0= Kurang Baik
X ₃	Sikap terhadap HIV/AIDS	1= Positif
		0= Negatif
X ₄	Konsep Diri	1= Positif
		0= Negatif
X ₅	Dukungan Keluarga	1= Mendukung
		0= Tidak Mendukung
X ₆	Lama Menderita	Bulan
X ₇	Perlakuan (Pemberian terapi ARV)	1= Mendapat ARV dan Pendampingan
		0= Hanya Mendapat ARV

Langkah analisis dalam penelitian ini adalah sebagai berikut (Yanovitzky, Zanutto, & Hornik, 2005).

1. Mendeskripsikan karakteristik pasien HIV/AIDS.
2. Langkah-langkah *propensity score-support vector machine*.
 - a) Menentukan variabel *confounding*, dimana variabel tersebut dinotasikan dengan z (Porta, Greenland, Hernan, Silva, & Last, 2000).
 - b) Mengestimasi nilai *propensity score* dengan model *support vector machine*. Kernel yang digunakan pada metode *propensity score-support vector machine* adalah kernel *radial basis function*. Nilai *cost* dan *gamma* yang digunakan secara berturut-turut adalah (1, 10, 50, 100) dan (0,1; 1; 3; 5; 10; 12). Metode optimasi parameter yang digunakan adalah optimasi *taguchi*.
 - c) Melakukan analisis *matching* dengan metode *nearest neighbor matching*.
 - d) Menghitung nilai *Percent Bias Reduction* (PBR) (Pan & Bai, 2015).
 - e) Mendapatkan signifikansi hasil *Average Treatment of Treated* (ATT) dengan menggunakan metode *propensity score matching-support vector machine*.
3. Langkah-langkah *propensity score-regresi logistik biner*.
 - a) Menentukan variabel *confounding*, dimana variabel tersebut dinotasikan dengan z.
 - b) Mengestimasi nilai *propensity score-regresi logistik biner*.
 - c) Melakukan analisis *matching* dengan metode *nearest neighbor matching*.
 - d) Menghitung nilai *Percent Bias Reduction* (PBR).
 - e) Mendapatkan signifikansi hasil *Average Treatment of Treated* (ATT) dengan menggunakan metode *propensity score matching-regresi logistik biner*.

4. Membandingkan metode *propensity score-support vector machine* dan *propensity score-regresi logistik biner*.
5. Membuat kesimpulan.

HASIL DAN PEMBAHASAN

A. Karakteristik Pasien HIV/AIDS

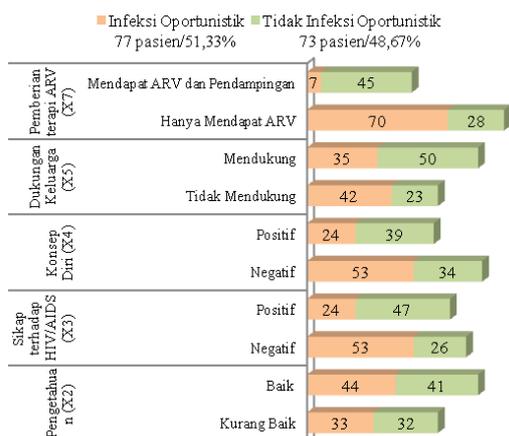
Deskripsi awal untuk mendapatkan gambaran umum atau karakteristik dari suatu data penelitian dapat dilakukan dengan analisis deskriptif. Analisis deskriptif tersebut dilakukan pada masing-masing variabel. Variabel yang digunakan dalam penelitian ini, terbagi menjadi dua yaitu variabel data kontinu dan variabel data kategorik. Variabel data kontinu dideskripsikan berdasarkan nilai rata-rata, standar deviasi, minimum, dan maksimum, sedangkan variabel kategorik dideskripsikan berdasarkan variabel infeksi oportunistik (Y) menggunakan diagram batang. Karakteristik pasien HIV/AIDS untuk data kontinu disajikan pada Tabel 2, sedangkan untuk data kategorik disajikan pada Gambar 1.

Tabel 2. Karakteristik Pasien HIV/AIDS berdasarkan Usia (X_1) dan Lama Menderita (X_6)

Kovariat (X)	Mean	Standar Deviasi	minimum	Maksimum
X_1	33,84 ^a	8,62 ^a	4 ^a	61 ^a
X_6	39,03 ^b	26,74 ^b	3 ^b	147 ^b

^atahun; ^bbulan

Tabel 2 menunjukkan bahwa rentang usia pasien yang terkena HIV/AIDS antara usia 4 sampai 61 tahun, dengan rata-rata usia pasien sekitar 33,84 tahun dan standar deviasi 8,62. Rentang lama pasien menderita HIV/AIDS antara 3 sampai 147 bulan, dengan rata-rata lama pasien menderita HIV/AIDS sekitar 39,03 bulan dan standar deviasi 26,74.



Gambar 1. Karakteristik Pasien HIV/AIDS Berdasarkan Variabel Infeksi Oportunistik

Gambar 1 memberikan informasi bahwa pasien HIV/AIDS yang memiliki infeksi oportunistik sebesar 51,33% atau bisa dikatakan dari 150 pasien HIV/AIDS terdapat 77 Pasien HIV/AIDS yang terkena infeksi oportunistik. Jumlah ini cukup banyak sehingga perlu adanya penanganan agar pasien HIV/AIDS tidak terkena infeksi oportunistik karena kematian paling banyak disebabkan adanya infeksi oportunistik.

Berdasarkan Gambar 1 juga dapat dilihat bahwa pasien HIV/AIDS yang terkena infeksi oportunistik dan memiliki pengetahuan yang kurang baik lebih sedikit dibanding yang memiliki pengetahuan baik yaitu 33 pasien (22,00%). Sedangkan pasien HIV/AIDS yang terkena infeksi oportunistik dan memiliki sikap negatif terhadap HIV/AIDS lebih banyak dibanding yang memiliki sikap positif yaitu 53 pasien (35,33%). Pasien HIV/AIDS yang terkena infeksi oportunistik dan mempunyai konsep diri yang negatif lebih banyak dibanding yang mempunyai konsep diri positif yaitu 53 pasien (35,33%). Pasien HIV/AIDS yang terkena infeksi oportunistik dan tidak mendapat dukungan dari keluarga lebih banyak dibanding yang mendapat dukungan dari keluarga yaitu 42 pasien (28%), serta pasien HIV/AIDS yang terkena infeksi oportunistik dan hanya mendapat terapi ARV lebih banyak dibanding yang mendapat terapi ARV beserta pendampingan yaitu 70 pasien (46,67%).

B. Pemilihan Variabel Confounding

Memilih kovariat sebagai variabel *confounding* merupakan langkah awal sebelum melakukan estimasi nilai *Propensity Score* menggunakan *Support Vector Machine* dan *Propensity Score* menggunakan Regresi Logistik Biner. Pemilihan variabel *confounding* dapat dilakukan berdasarkan teori dan kemudian di buktikan secara empiris dengan melihat hubungan antar variabel. Melihat hubungan antar variabel dapat menggunakan statistik uji *Chi-square*.

Berdasarkan teori yang telah dijelaskan pada tinjauan pustaka, faktor-faktor yang memengaruhi pasien HIV/AIDS terkena infeksi oportunistik adalah faktor usia, pengetahuan, sikap terhadap HIV/AIDS, konsep diri, dukungan keluarga, lama menderita dan pemberian terapi ARV. Secara empiris pengujian hubungan variabel yang diduga sebagai variabel *confounding* dengan kovariat (X) lainnya disajikan pada Tabel 3.

Tabel 3. Pengujian Hubungan Variabel yang Diduga Sebagai Variabel *Confounding* dengan Kovariat (X) Lainnya

Variabel	χ^2	P-value	Keputusan
$X_3 * X_2$	28,6993	0,0000	*
$X_4 * X_2$	10,4864	0,0012	*
$X_5 * X_2$	0,5190	0,4713	**
$X_7 * X_2$	0,2579	0,6116	**
$X_2 * X_3$	28,6993	0,0000	*
$X_4 * X_3$	1,1102	0,2920	**
$X_5 * X_3$	1,5451	0,2139	**
$X_7 * X_3$	4,8162	0,0282	*
$X_2 * X_4$	10,4864	0,0012	*
$X_3 * X_4$	1,1102	0,2920	**
$X_5 * X_4$	0,3221	0,5704	**
$X_7 * X_4$	6,1946	0,0128	*
$X_2 * X_5$	0,5190	0,4713	**
$X_3 * X_5$	1,5451	0,2139	**
$X_4 * X_5$	0,3221	0,5704	**
$X_7 * X_5$	6,8028	0,0091	*
$X_2 * X_7$	0,2579	0,6116	**
$X_3 * X_7$	4,8162	0,0282	*
$X_4 * X_7$	6,1946	0,0128	*
$X_5 * X_7$	6,8028	0,0091	*

*p-value < $\alpha = 0.05$ (Tolak H_0), **p-value > $\alpha = 0.05$ (Gagal Tolak H_0)

Tabel 3 menyajikan tentang uji dependensi antara variabel yang diduga sebagai *confounding* dengan kovariat (X) lainnya. Jika variabel (X₂) yang diduga sebagai variabel *confounding*, memiliki dua hubungan yang signifikan dengan variabel kovariat lainnya yaitu variabel (X₃) dan variabel (X₄). Begitupula variabel (X₃) dan variabel (X₄) jika diduga sebagai variabel *confounding* juga memiliki dua hubungan yang signifikan dengan variabel kovariat lainnya yaitu variabel (X₂) dan variabel (X₇). Sedangkan jika variabel (X₅) yang diduga sebagai variabel *confounding*, hanya memiliki satu hubungan yang signifikan dengan variabel kovariat lainnya yaitu variabel (X₇) dan jika variabel (X₇) yang diduga sebagai variabel *confounding* terdapat tiga hubungan yang signifikan dengan variabel kovariat lainnya yaitu variabel (X₃), variabel (X₄) dan variabel (X₅).

Pengujian hubungan kovariat dengan variabel yang diduga sebagai *confounding* telah dilakukan, selanjutnya pengujian hubungan variabel yang diduga sebagai *confounding* dengan variabel infeksi oportunistik (Y) yang disajikan pada Tabel 4.

Tabel 4. Pengujian Hubungan Variabel yang Diduga Sebagai Variabel *Confounding* dengan Variabel Komplikasi Penyakit (Y)

Variabel	χ^2	P-value	Keputusan
X ₂ *Y	0,0146	0,9038	**
X ₃ *Y	16,5837	0,0000	*
X ₄ *Y	7,6196	0,0058	*
X ₅ *Y	8,1000	0,0044	*
X ₇ *Y	45,6951	0,0000	*

*p - value < $\alpha = 0.05$ (Tolak H₀)

**p - value > $\alpha = 0.05$ (Gagal Tolak H₀)

Berdasarkan Tabel 4 dapat dilihat bahwa variabel X₃, X₄, X₅ dan X₇ memiliki hubungan yang signifikan dengan variabel infeksi oportunistik (Y) karena memiliki nilai p - value < $\alpha = 0.05$. Dalam hal ini yang dipilih sebagai variabel *confounding* adalah variabel pemberian terapi ARV (X₇) karena

variabel tersebut selain memiliki hubungan yang signifikan dengan variabel infeksi oportunistik (Y) juga memiliki tiga hubungan yang signifikan dengan variabel kovariat lainnya. Hal ini sesuai dengan teori yang dilakukan oleh B-Lajoie, Drouin, Bartlett, Nguyen, Low, Gavriilidis, ... & Muhe yaitu pasien yang melakukan terapi ARV sebagian besar mengalami penurunan risiko terkena infeksi oportunistik (B-Lajoie et al., 2016).

C. Propensity Score Matching-Support Vector Machine

Pemilihan variabel *confounding* telah dilakukan, maka tahap selanjutnya yang perlu dilakukan pada metode *Propensity Score Matching-Support Vector Machine* adalah (1) mengestimasi nilai *Propensity Score* (PS) menggunakan model *Support Vector Machine* (SVM), (2) melakukan analisis *Matching*, (3) menghitung *Percent Bias Reduction* (PBR), dan yang terakhir (4) menghitung signifikansi hasil *Average Treatment of Treated* (ATT).

1. Estimasi Nilai Propensity Score dengan SVM

Variabel respon yang digunakan pada saat mengestimasi nilai *propensity score-support vector machine* adalah variabel *confounding* yang telah ditentukan di langkah sebelumnya yaitu variabel pemberian terapi ARV (X₇) yang selanjutnya dinotasikan (Z) dan variabel prediktor yang digunakan adalah variabel usia (X₁), pengetahuan (X₂), sikap terhadap HIV/AIDS (X₃), konsep diri (X₄), dukungan keluarga (X₅), lama menderita (X₆) dan variabel pemberian terapi ARV (X₇). Berdasarkan persamaan (6) dapat diperoleh nilai *propensity score* dengan model SVM, namun untuk mendapatkan nilai tersebut perlu didapatkan dahulu fungsi klasifikasi (*score*) dari model SVM yang disimbolkan dengan $f(x)$. Nilai *cost* dan *gamma* yang optimum diperlukan untuk mendapatkan nilai $f(x)$. Nilai *cost*

dan γ yang optimum dengan menggunakan metode *Taguchi* adalah 1 dan 0,1. Berikut ini *score* dari fungsi klasifikasi model SVM.

Tabel 5. *Score* Fungsi Klasifikasi Model SVM ($f(x)$)

Data ke-	$f(x)$					
1-6	0,202	0,054	0,162	0,453	0,148	0,746
7-12	0,401	0,082	0,092	0,160	0,922	0,880
⋮	⋮	⋮	⋮	⋮	⋮	⋮
139-144	0,048	0,035	0,063	0,048	0,024	0,023
145-150	0,006	-	0,081	0,484	0,175	0,341

Score dari fungsi klasifikasi model SVM telah diperoleh, maka tahap selanjutnya adalah mendapatkan nilai *propensity score* dengan menggunakan model *support vector machine*, dan hasilnya disajikan pada Tabel 6.

Tabel 6 Estimasi Nilai *Propensity Score* dengan Model SVM

Data ke-	$\hat{e}(x_i)$					
1-6	0,550	0,513	0,540	0,611	0,537	0,678
7-12	0,599	0,520	0,523	0,540	0,715	0,707
⋮	⋮	⋮	⋮	⋮	⋮	⋮
139-144	0,512	0,509	0,516	0,512	0,494	0,506
145-150	0,502	0,491	0,520	0,619	0,544	0,584

2. Analisis Matching

Langkah selanjutnya setelah mendapatkan nilai estimasi *propensity score-support vector machine* adalah melakukan analisis *matching*. Metode yang digunakan dalam analisis *matching* adalah metode *nearest neighbor matching*. Konsep dari analisis *matching* tersebut adalah mencocokkan data pada kelompok *treatment* dengan data pada kelompok kontrol berdasarkan nilai *propensity score* yang terdekat. Pada saat pencocokan, kelompok kontrol yang tidak memiliki pasangan akan dikeluarkan dari analisis selanjutnya.

Berdasarkan hasil analisis *matching*, dari 150 pasien HIV/AIDS terdapat 52 pasien yang mendapat terapi ARV dan pendampingan (kelompok *treatment*) yang

dipasangkan dengan pasien yang hanya mendapat terapi ARV (kelompok kontrol). Pasien yang hanya mendapat terapi ARV (kelompok kontrol) dan tidak memiliki pasangan dikeluarkan dari himpunan pasien yang hanya mendapat terapi ARV (kelompok kontrol) dan tidak diikuti pada analisis selanjutnya. Hasil analisis *matching* tersebut dapat ditunjukkan pada Tabel 7.

Tabel 7 Jumlah Pasangan yang Terbentuk

No	Kategori	Jumlah
1	Kelompok <i>treatment</i>	52
2	Pasangan kelompok <i>treatment</i>	52
3	Kelompok <i>non-treatment</i>	98
4	Pasangan kelompok <i>non-treatment</i>	52
5	Kelompok yang <i>matched</i>	104
6	Kelompok yang <i>not-matched</i>	46

Jumlah pasangan yang terbentuk dari hasil analisis *matching* telah didapat. Selanjutnya data kelompok *treatment* dan kontrol yang sudah *matching* dapat disajikan pada Tabel 8.

Tabel 8 Anggota *Treatment* dan Kontrol yang *Matching*

No	Data yang dipasangkan	
	<i>Treatment</i>	Kontrol
1	135	118
2	104	46
⋮	⋮	⋮
51	15	55
52	82	101

Tabel 8 menunjukkan bahwa pasien HIV/AIDS yang mendapat terapi ARV dan pendampingan merupakan pasien dengan nomor urut ke-135, 104, 109, 96, 18, 88, 2, 73, 78, 8, 20, 40, 62, 58, 100, 30, 123, 10, 111, 81, 22, 33, 42, 116, 54, 110, 115, 36, 74, 103, 16, 134, 7, 136, 19, 28, 14, 37, 35, 44, 13, 132, 6, 29, 12, 23, 21, 11, 24, 34, 15, dan 82. Karakteristik pasangan pasien yang telah dicocokkan (*matching*), jika dilihat berdasarkan pasien yang terkena infeksi oportunistik terdapat 41 pasien dan apabila dilihat berdasarkan pasien yang

memiliki pengetahuan kurang baik terdapat 49 pasien sedangkan jika dilihat berdasarkan pasien yang memiliki sikap negatif terhadap HIV/AIDS terdapat 47 pasien dan pasien yang memiliki konsep diri negatif terdapat 49 pasien, serta apabila dilihat berdasarkan pasien yang tidak mendapat dukungan keluarga terdapat 33 pasien.

3. Percent Bias Reduction (PBR)

Analisis *matching* telah dilakukan, langkah selanjutnya adalah menghitung *percent bias reduction*. *Percent bias reduction* digunakan untuk menghitung berapa persen bias dapat tereduksi dari sebelum dilakukan *matching* sampai setelah dilakukan *matching*. *Percent bias reduction* pada metode *propensity score-support vector machine* dapat dilihat pada Tabel 9.

Tabel 9 *Percent Bias Reduction*

Bias	PSM dengan <i>support vector machine</i>
Sebelum <i>Matching</i>	0,062
Setelah <i>Matching</i>	0,044
PBR	28,57%

Tabel 9 menunjukkan bahwa bias yang dihasilkan setelah *matching* lebih kecil dibandingkan bias sebelum *matching* yaitu sebesar 0,044. Sebelum dilakukan *matching* sampai setelah dilakukan *matching*, bias dapat tereduksi sebesar 28,57%.

4. Signifikansi Hasil *Average Treatment of Treated (ATT)*

Pada tahap selanjutnya adalah menghitung signifikansi hasil *Average Treatment of Treated (ATT)*, namun sebelum menghitung signifikansi hasil ATT perlu di uji keseimbangan (*balance*) kovariatnya terlebih dahulu.

a. Pengujian *Balance Kovariat*

Pengujian *balance* kovariat dalam *propensity score* digunakan untuk mengecek keseimbangan kovariat pada

variabel *confounding* antara kelompok *treatment* dan kontrol. Pada penelitian ini untuk menguji perbedaan antara kelompok *treatment* dan kontrol menggunakan uji kesamaan dua rata-rata (uji t) untuk data kontinu dan menggunakan uji kesamaan dua proporsi (uji z) untuk data kategori. Pengujian *balance* kovariat terbagi menjadi dua tahap, pada tahap pertama pengujian dilakukan sebelum *matching* dan pada tahap kedua pengujian dilakukan sesudah *matching*. Hasil pengujian *balance* kovariat ditunjukkan pada Tabel 10.

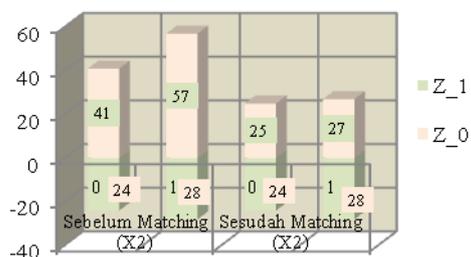
Tabel 10 Hasil Pengujian *Balance Kovariat*

Variabel	<i>Balance</i>	<i>P-value</i>
X ₁	Sebelum	0,774
	Sesudah	0,530
X ₂	Sebelum	0,616
	Sesudah	0,846
X ₃	Sebelum	0,029
	Sesudah	0,329
X ₄	Sebelum	0,014
	Sesudah	0,560
X ₅	Sebelum	0,007
	Sesudah	0,532
X ₆	Sebelum	0,896
	Sesudah	0,899

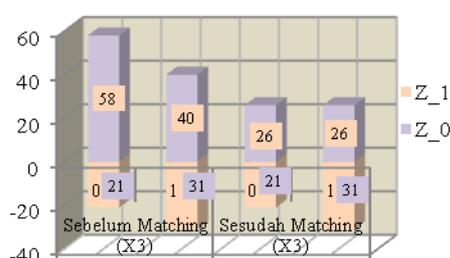
p - value > $\alpha = 0.05$ (Gagal Tolak H₀)

Berdasarkan Tabel 10 dapat dilihat bahwa sebelum dilakukan analisis PSM dengan *support vector machine*, variabel (X₃), variabel (X₄) dan variabel (X₅) tidak *balance* yang artinya variabel tersebut memiliki perbedaan proporsi antara kelompok *treatment* dan kelompok kontrol. Sedangkan setelah dilakukan analisis *propensity score matching-support vector machine*, variabel (X₁), variabel (X₂), variabel (X₃), variabel (X₄), variabel (X₅) dan variabel (X₆) sudah *balance* yang artinya variabel tersebut tidak memiliki perbedaan proporsi antara kelompok *treatment* dan kelompok kontrol atau dapat diartikan pula variabel (X₁), variabel (X₂), variabel (X₃), variabel (X₄), variabel (X₅) dan variabel (X₆) merupakan variabel yang berpengaruh secara langsung terhadap pemberian terapi ARV sehingga kesimpulan dari studi observasi menjadi

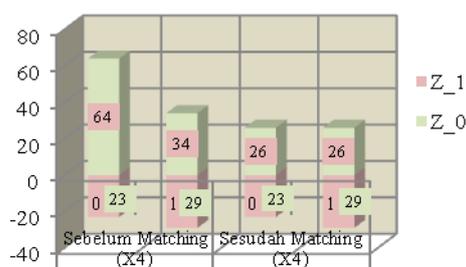
tidak bias. *Balance* kovariat untuk data kategori secara visual dapat dilihat pada Gambar 2.



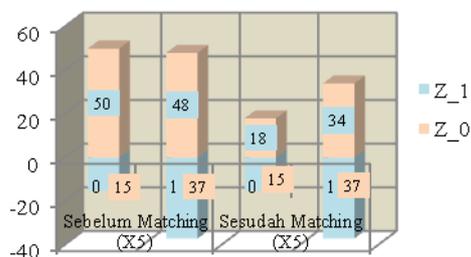
(a)



(b)



(c)



(d)

Gambar 2 *Balance* Kovariat Secara Visual, (a) variabel (X₂); (b) Variabel (X₃); (c) Variabel (X₄); (d) Variabel (X₅)

Gambar 6 menunjukkan bahwa jumlah kelompok kontrol dan *treatment* setelah dilakukan *matching* lebih *balance* daripada sebelum dilakukan *matching*. Hal

ini dapat dilihat pada variabel X₂ untuk kelompok kontrol, sebelum dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 24 dan jumlah *treatment* pada variabel Z sebanyak 41. Sedangkan setelah dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 24 dan jumlah *treatment* pada variabel Z sebanyak 57. Pada variabel X₃ untuk kelompok kontrol, sebelum dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 21 dan jumlah *treatment* pada variabel Z sebanyak 58. Sedangkan setelah dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 21 dan jumlah *treatment* pada variabel Z sebanyak 40. Pada variabel X₄ untuk kelompok kontrol, sebelum dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 23 dan jumlah *treatment* pada variabel Z sebanyak 64. Sedangkan setelah dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 23 dan jumlah *treatment* pada variabel Z sebanyak 34. Pada variabel X₅ untuk kelompok kontrol, sebelum dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 15 dan jumlah *treatment* pada variabel Z sebanyak 50. Sedangkan setelah dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 15 dan jumlah *treatment* pada variabel Z sebanyak 48.

b. *Signifikansi Hasil Average Treatment of Treated*

Signifikansi hasil ATT dilakukan untuk mengetahui seberapa besar pengaruh pemberian terapi ARV (Z) terhadap kejadian infeksi oportunistik (Y) pada saat pengaruh dari variabel lain (kovariat) sudah direduksi. Hasil signifikansi hasil ATT dapat dilihat pada Tabel 11.

Tabel 11 Signifikansi Hasil ATT

Variabel	ATT (θ)	SE (ATT)	T	P-value
X ₁ ,X ₂ ,X ₃ ,X ₄ ,X ₅ ,X ₆	0,566	0,106	5,322	0,000*

*p - value < $\alpha = 0.05$ (Tolak H₀)

Tabel 11 menunjukkan bahwa variabel pemberian terapi ARV (Z) berpengaruh secara signifikan terhadap infeksi oportunistik pada pasien HIV/AIDS (Y) karena memiliki nilai $p\text{-value}$ $(0,0000) < \alpha (0,05)$. Berdasarkan Tabel 11 diketahui pula bahwa efek hanya mendapat terapi ARV terhadap infeksi oportunistik adalah sebesar 0,566 atau bisa dikatakan peluang pasien HIV/AIDS yang hanya mendapat terapi ARV untuk terindikasi infeksi oportunistik adalah 0,566 kali lebih besar dibandingkan pasien yang mendapat terapi ARV dan pendampingan.

D. Propensity Score Matching- Regresi Logistik Biner

Tahap awal yang perlu dilakukan pada metode *Propensity Score Matching - Regresi Logistik Biner* adalah (1) mengestimasi nilai *Propensity Score* (PS) menggunakan model Regresi Logistik Biner, (2) melakukan analisis *Matching*, (3) menghitung *Percent Bias Reduction* (PBR), dan yang terakhir (4) menghitung signifikansi hasil *Average Treatment of Treated* (ATT).

1. Estimasi Nilai Propensity Score dengan Regresi Logistik Biner

Pada tahap mengestimasi nilai *propensity score* menggunakan model regresi logistik biner, variabel respon yang digunakan adalah variabel *confounding* yang telah ditentukan di langkah sebelumnya yaitu variabel pemberian terapi ARV (X_7) yang selanjutnya dinotasikan (Z) dan variabel prediktor yang digunakan adalah variabel usia (X_1), pengetahuan (X_2), sikap terhadap HIV/AIDS (X_3), konsep diri (X_4), dukungan keluarga (X_5), lama menderita (X_6) dan variabel pemberian terapi ARV (X_6). Nilai *propensity score* sering dikenal dengan model regresi logistik yang disimbolkan dengan $\pi(x_i)$ seperti dalam persamaan (5). Nilai *propensity score* akan diperoleh jika parameter model regresi logistik biner telah didapatkan. Hasil estimasi parameter nilai

propensity score dengan regresi logistik biner dapat dilihat pada Tabel 12.

Tabel 12 Estimasi Parameter Model Regresi Logistik Biner antara Kovariat (X) dengan Pemberian terapi ARV dan Pendampingan (Z)

Kovariat	Parameter (β)	SE	P-value	OR
Intercept	-2,294	0,961	0,017	0,101
X_1	-0,001	0,022	0,941	0,999
$X_{2,1}$	0,512	0,430	0,234	1,669
$X_{3,1}$	0,854	0,410	0,037	2,349
$X_{4,1}$	1,056	0,394	0,007	2,875
$X_{5,1}$	1,007	0,386	0,009	2,737
X_6	-0,002	0,007	0,782	0,998

Tabel 12 menunjukkan bahwa variabel yang berpengaruh secara signifikan terhadap infeksi oportunistik (Z) adalah variabel pasien yang memiliki sikap positif terhadap penyakit HIV/AIDS ($X_{3,1}$), variabel pasien yang memiliki konsep diri positif ($X_{4,1}$) dan variabel pasien yang mendapat dukungan keluarga ($X_{5,1}$). Berdasarkan tabel 12 fungsi $f(x)$ dapat dituliskan sebagai berikut.

$$\hat{f}(x_i) = -2,294 - 0,001X_1 + 0,512X_{2,1} + 0,854X_{3,1} + 1,056X_{4,1} + 1,007X_{5,1} - 0,002X_6$$

Berdasarkan persamaan di atas dapat diperoleh nilai *propensity score* sebagai berikut.

Tabel 13 Estimasi Nilai *Propensity Score* dengan Model Regresi Logistik Biner

Data ke-	$\hat{e}(x_i)$					
1-6	0,288	0,173	0,494	0,614	0,289	0,727
7-12	0,387	0,190	0,391	0,183	0,633	0,562
⋮	⋮	⋮	⋮	⋮	⋮	⋮
139-144	0,315	0,253	0,473	0,179	0,128	0,134
145-150	0,296	0,131	0,396	0,630	0,411	0,542

2. Analisis Matching

Setelah mendapatkan nilai estimasi *propensity score* menggunakan model Regresi Logistik Biner langkah selanjutnya adalah melakukan analisis *matching*. Metode *nearest neighbor matching* merupakan metode yang digunakan dalam analisis *matching*. Algoritma dari analisis *matching* adalah mencari selisih nilai *propensity score* terdekat antara data pada

kelompok *treatment* dengan data pada kelompok kontrol. Pada saat analisis *matching* kelompok kontrol yang tidak memiliki pasangan dapat dikeluarkan dari himpunan kelompok kontrol dan tidak digunakan dalam analisis selanjutnya.

Berdasarkan hasil analisis *matching*, dari 150 pasien HIV/AIDS terdapat 52 pasien yang mendapat terapi ARV dan pendampingan (kelompok *treatment*) yang dipasangkan dengan pasien yang hanya mendapat terapi ARV (kelompok kontrol). Pasien yang hanya mendapat terapi ARV dan tidak memiliki pasangan dikeluarkan dari himpunan pasien yang hanya mendapat terapi ARV (kelompok kontrol) dan tidak diikuti pada analisis selanjutnya. Hasil analisis *matching* dapat ditunjukkan pada Tabel 14.

Tabel 14 Jumlah Pasangan yang Terbentuk

No	Kategori	Jumlah
1	Kelompok <i>treatment</i>	52
2	Pasangan kelompok <i>treatment</i>	52
3	Kelompok <i>non-treatment</i>	98
4	Pasangan kelompok <i>non-treatment</i>	52
5	Kelompok yang <i>matched</i>	104
6	Kelompok yang <i>not-matched</i>	46

Data pasien HIV/AIDS dari kelompok *treatment* dan *propensity score* yang terdekat. Berikut ini pembagian data pada kelompok kontrol dipasangkan berdasarkan nilai *pok treatment* dan kontrol yang dipasangkan.

Tabel 15 Anggota *Treatment* dan Kontrol yang *Matching*

No	Data yang dipasangkan	
	<i>Treatment</i>	Kontrol
1	2	63
2	6	117
⋮	⋮	⋮
51	135	76
52	136	57

Tabel 15 menunjukkan bahwa pasien HIV/AIDS yang mendapat terapi

ARV dan pendampingan merupakan pasien dengan nomor urut ke-2, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 18, 20, 21, 22, 23, 24, 28, 29, 30, 34, 35, 36, 37, 40, 42, 44, 54, 58, 62, 73, 74, 78, 81, 82, 88, 96, 100, 103, 104, 109, 110, 111, 115, 116, 123, 132, 134, 135, dan 136. Karakteristik pasangan pasien yang telah dicocokkan, jika dilihat berdasarkan pasien yang terkena infeksi oportunistik terdapat 40 pasien dan apabila dilihat berdasarkan pasien yang memiliki pengetahuan kurang baik terdapat 55 pasien sedangkan jika dilihat berdasarkan pasien yang memiliki sikap negatif terhadap HIV/AIDS terdapat 43 pasien dan pasien yang memiliki konsep diri negatif terdapat 50 pasien, serta apabila dilihat berdasarkan pasien yang tidak mendapat dukungan keluarga terdapat 32 pasien.

3. Percent Bias Reduction (PBR)

Tahap selanjutnya setelah menganalisis *matching* adalah menghitung *percent bias reduction*. *Percent bias reduction* digunakan untuk menghitung berapa persen bias dapat tereduksi dari sebelum dilakukan *matching* sampai setelah dilakukan *matching*. *Percent bias reduction* dapat dilihat pada Tabel 16.

Tabel 16 *Percent Bias Reduction*

Bias	PSM dengan <i>support vector machine</i>
Sebelum <i>Matching</i>	0,124
Setelah <i>Matching</i>	0,050
PBR	60.25%

Tabel 16 menunjukkan bahwa bias yang dihasilkan setelah *matching* lebih kecil dibandingkan bias sebelum *matching* yaitu sebesar 0,05. Sebelum dilakukan *matching* sampai setelah dilakukan *matching*, bias dapat tereduksi sebesar 60,25%.

4. Signifikansi Hasil *Average Treatment of Treated (ATT)*

Pada tahap selanjutnya adalah menghitung signifikansi hasil *Average*

Treatment of Treated (ATT), dimana sebelum menghitung signifikansi hasil ATT perlu di uji keseimbangan (*balance*) kovariatnya terlebih dahulu.

a. Pengujian Balance Kovariat

Pada bagian ini, pengujian *balance* kovariat digunakan untuk mengecek keseimbangan kovariat pada variabel *confounding* antara kelompok *treatment* dan kontrol. Pada penelitian ini untuk menguji perbedaan antara kelompok *treatment* dan kontrol menggunakan uji kesamaan dua rata-rata (uji t) untuk data kontinu dan menggunakan uji kesamaan dua proporsi (uji z) untuk data kategori. Pengujian *balance* kovariat dibagi menjadi dua tahapan yaitu sebelum dilakukan *matching* dan sesudah dilakukan *matching*. Hasil pengujian *balance* kovariat ditunjukkan pada Tabel 17.

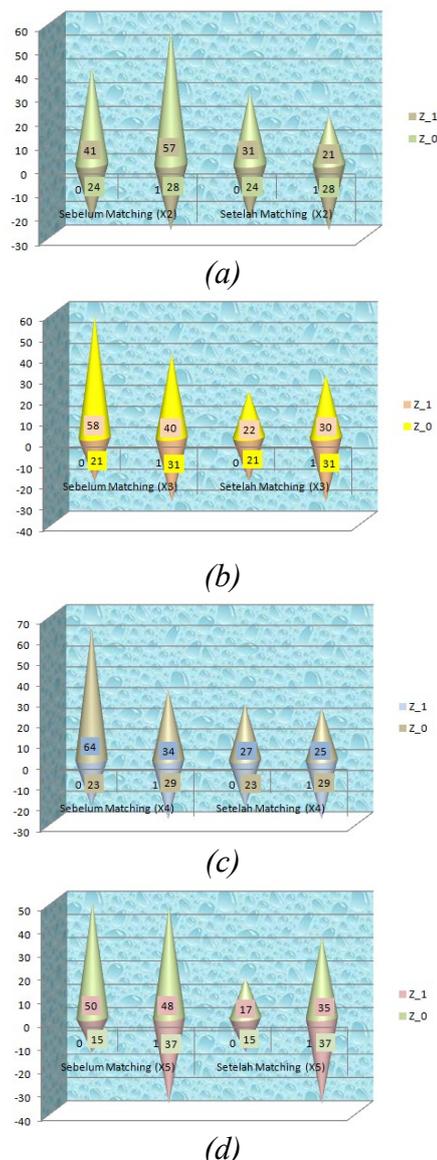
Tabel 17 Hasil Pengujian Balance Kovariat

Variabel	Balance	P-value
X ₁	Sebelum	0,774
	Sesudah	0,700
X ₂	Sebelum	0,616
	Sesudah	0,172
X ₃	Sebelum	0,029
	Sesudah	0,844
X ₄	Sebelum	0,014
	Sesudah	0,437
X ₅	Sebelum	0,007
	Sesudah	0,675
X ₆	Sebelum	0,896
	Sesudah	0,077

p – value > α = 0.05 (Gagal Tolak H₀)

Berdasarkan Tabel 17 dapat dilihat bahwa sebelum dilakukan analisis PSM dengan model regresi logistik biner, variabel (X₃), variabel (X₄) dan variabel (X₅) tidak *balance* yang artinya variabel tersebut memiliki perbedaan proporsi antara kelompok *treatment* dan kelompok kontrol. Sedangkan setelah dilakukan analisis *propensity score matching-regresi logistik biner*, variabel (X₁), variabel (X₂), variabel (X₃), variabel (X₄), variabel (X₅) dan

variabel (X₆) sudah *balance* yang artinya variabel tersebut tidak memiliki perbedaan proporsi antara kelompok *treatment* dan kelompok kontrol atau dapat diartikan pula variabel (X₁), variabel (X₂), variabel (X₃), variabel (X₄), variabel (X₅) dan variabel (X₆) merupakan variabel yang berpengaruh secara langsung terhadap pemberian terapi ARV sehingga kesimpulan dari studi observasi menjadi tidak bias. *Balance* kovariat untuk data kategori secara visual dapat dilihat pada Gambar 3.



Gambar 3 Balance Kovariat Secara Visual, (a) variabel (X₂); (b) Variabel (X₃); (c) Variabel (X₄); (d) Variabel (X₅)

Gambar 3 menunjukkan bahwa jumlah kelompok kontrol dan *treatment*

setelah dilakukan *matching* lebih *balance* daripada sebelum dilakukan *matching*. Hal ini dapat dilihat pada variabel X_2 untuk kelompok kontrol, sebelum dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 24 dan jumlah *treatment* pada variabel Z sebanyak 41. Sedangkan setelah dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 24 dan jumlah *treatment* pada variabel Z sebanyak 31. Pada variabel X_3 untuk kelompok kontrol, sebelum dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 21 dan jumlah *treatment* pada variabel Z sebanyak 58. Sedangkan setelah dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 21 dan jumlah *treatment* pada variabel Z sebanyak 22. Pada variabel X_4 untuk kelompok kontrol, sebelum dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 23 dan jumlah *treatment* pada variabel Z sebanyak 64. Sedangkan setelah dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 23 dan jumlah *treatment* pada variabel Z sebanyak 27. Pada variabel X_5 untuk kelompok kontrol, sebelum dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 15 dan jumlah *treatment* pada variabel Z sebanyak 50. Sedangkan setelah dilakukan *matching* jumlah kontrol pada variabel Z sebanyak 15 dan jumlah *treatment* pada variabel Z sebanyak 17.

b. Signifikansi Hasil Average Treatment of Treated (ATT)

Signifikansi Hasil ATT dilakukan untuk mengetahui seberapa besar pengaruh pemberian terapi ARV (Z) terhadap kejadian infeksi oportunistik (Y) pada saat pengaruh dari variabel lain (kovariat) sudah direduksi. Signifikansi hasil ATT dapat dilihat pada Tabel 18.

Tabel 18 Signifikansi Hasil ATT

Variabel	ATT (θ)	SE (ATT)	T	P-value
$X_1, X_2, X_3, X_4, X_5, X_6$	0,474	0,097	4,911	0,000

*p - value < $\alpha = 0.05$ (Tolak H_0)

Tabel 18 menunjukkan bahwa variabel pemberian terapi ARV (Z) berpengaruh secara signifikan terhadap infeksi oportunistik pada pasien HIV/AIDS (Y) karena memiliki nilai *p-value* (0,0000) < alpha (0,05). Berdasarkan Tabel 11 diketahui pula bahwa efek hanya mendapat terapi ARV terhadap infeksi oportunistik adalah sebesar 0,474 atau bisa dikatakan peluang pasien HIV/AIDS yang hanya mendapat terapi ARV untuk terindikasi infeksi oportunistik adalah 0,474 kali lebih besar dibandingkan pasien yang mendapat terapi ARV dan pendampingan.

E. Perbandingan Metode Propensity Score Matching-Support Vector Machine dan Propensity Score Matching-Regresi Logistik Biner

Berdasarkan ulasan diatas dapat dilihat bahwa pemberian terapi ARV (Z) berpengaruh secara signifikan terhadap infeksi oportunistik pada pasien HIV/AIDS (Y) baik menggunakan metode *propensity score matching-support vector machine* maupun metode *propensity score matching-regresi logistik Biner*. Jika menggunakan metode *propensity score matching-support vector machine* nilai bias dapat tereduksi sebesar 28,57% dan jika menggunakan metode *propensity score matching-regresi logistik Biner* nilai bias dapat tereduksi sebesar 60,25%. Namun jika dibandingkan antara metode *propensity score matching-support vector machine* dan *propensity score matching-regresi logistik Biner*, metode *propensity score matching-support vector machine* menghasilkan nilai bias (setelah *matching*) yang lebih kecil yaitu sebesar 0,044.

KESIMPULAN

Signifikansi hasil *Average Treatment of Treated* (ATT) pada metode *propensity score matching-support vector machine* dan *propensity score matching-regresi logistik biner* menunjukkan bahwa variabel pemberian terapi ARV (Z)

berpengaruh secara signifikan terhadap infeksi oportunistik pada pasien HIV/AIDS (Y). Pada metode *propensity score matching-support vector machine* nilai bias yang dapat tereduksi sebesar 28,57% sedangkan pada metode *propensity score matching-regresi logistik Biner* nilai bias yang dapat tereduksi sebesar 60,25%. Meskipun, jumlah bias yang tereduksi pada metode regresi logistik biner lebih besar tetapi, nilai bias setelah *matching* pada metode *propensity score matching-support vector machine* lebih kecil dibanding metode *propensity score matching-regresi logistik Biner* yaitu sebesar 0,044.

Pada penelitian ini peneliti hanya menggunakan algoritma *Nearest Neighbor Matching*. Saran yang dapat diberikan untuk penelitian selanjutnya, peneliti dapat juga menggunakan algoritma *matching* yang lainnya, seperti metode *Stratified Matching*, *N:N Matching*, *Mahalanobis Metric Matching*, dan *Kernel Matching*.

DAFTAR PUSTAKA

- Becker, S. O., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal*, 2(4), 358–377. doi: <https://doi.org/10.1177/1536867X0200200403>
- B-Lajoie, M.-R., Drouin, O., Bartlett, G., Nguyen, Q., Low, A., Gavriilidis, G., ... Muhe, L. (2016). Incidence and Prevalence of Opportunistic and Other Infections and the Impact of Antiretroviral Therapy Among HIV-infected Children in Low- and Middle-income Countries: A Systematic Review and Meta-analysis. *Clinical Infectious Diseases*, 62(12), 1586–1594. doi: 10.1093/cid/ciw139
- Center for Disease Control and Prevention, C. for D. C. and P. (2018). *About HIV/AIDS*.
- Fenner, L., Reid, S., Fox, M. P., Garone, D., Wellington, M., Prozesky, H., ... Egger, M. (2013). Tuberculosis and the risk of opportunistic infections and cancers in HIV-infected patients starting ART in Southern Africa. *Tropical Medicine & International Health*, 18(2), 194–198. doi: 10.1111/tmi.12026
- Guo, S. Y., & Fraser, M. W. (2009). *Propensity score analysis: Statistical methods and applications*. SAGE Publications, Inc.
- Hardle, W. K., Prastyo, D. D., & Hafner, C. M. (2014). Support Vector Machines with Evolutionary Model Selection for Default Prediction. In *The Oxford Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics*. Oxford University Press. doi: 10.1093/oxfordhb/9780199857944.013.011
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression, Second Edition*. John Wiley & Sons, Inc. doi: 10.1002/0471722146
- Otok, B. W., Aisyah, A., Puhadi, & Andari, S. (2017). Propensity score matching of the gymnastics for diabetes mellitus using logistic /regression. *AIP Conference Proceedings*, 1913(1), 1–8. doi: <https://doi.org/10.1063/1.5016668>
- Pal, M., & Mather, P. . (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26(5), 1007–1011. doi: <https://doi.org/10.1080/01431160512331314083>
- Pan, W., & Bai, H. (2015). *Propensity Score Analysis Fundamentals and Developments*. New York: The Guilford Press.
- Porta, M., Greenland, S., Hernan, M., Silva, I. D. S., & Last, J. M. (2000). A Dictionary of Epidemiology 6 th edition. In *A Handbook Sponsored By The IEA* (p. 343). Oxford University Press.
- Rosenbaum, P. R., & Rubin, D. B. (1983).

- The Central Role of the Propensity Score in Observational Studies For Causal Effects. *Biometrika*, 70(1), 41–55. doi: 10.1093/biomet/70.1.41
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician*, 39(1), 33–38. doi: <https://doi.org/10.2307/2683903>
- Yanovitzky, I., Zanutto, E. L., & Hornik, R. (2005). Estimating Causal Effects of Public Health Education Campaigns Using Propensity Score Methodology. *Evaluation and Program Planning*, 28(2), 209–220. doi: 10.1016/j.evalprogplan.2005.01.004
- Yao, B., Zhang, L., Liang, S., & Zhang, C. (2012). SVMTriP: A Method to Predict Antigenic Epitopes Using Support Vector Machine to Integrate Tri-Peptide Similarity and Propensity. *Plos One*, 7(9), 1–5. doi: <https://doi.org/10.1371/journal.pone.0045152>