

PENERAPAN ANALISIS *PROPENSITY SCORE MATCHING* REGRESI LOGISTIK BINER PADA KASUS KANKER PAYUDARA

ALIFA SHAKILA



**DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**

@Hak cipta milik IPB University

IPB University



IPB University
Bogor Indonesia

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penerapan Analisis *Propensity Score Matching* Regresi Logistik Biner pada Kasus Kanker Payudara” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2024

Alifa Shakila
G1401201087

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

ALIFA SHAKILA. Penerapan Analisis *Propensity Score Matching* Regresi Logistik Biner pada Kasus Kanker Payudara. Dibimbing oleh RAHMA ANISA dan AAM ALAMUDI.

Kanker payudara wanita merupakan penyebab utama kematian akibat kanker di Indonesia pada tahun 2022. Kanker payudara tergolong penyakit tidak menular yang belum diketahui penyebab pastinya. Penelitian ini merupakan studi observasi yang dilakukan terhadap wanita Indonesia yang menderita dan tidak menderita kanker payudara. Tujuan penelitian ini adalah mengidentifikasi faktor-faktor yang berpengaruh terhadap kanker payudara setelah mengatasi efek pembaur yang dapat membiaskan hubungan sebenarnya antara faktor risiko dan status kanker payudara. Penelitian ini menggunakan metode *Propensity Score Matching* yang mampu mereduksi efek pembaur khususnya dalam penelitian observasi kesehatan. Peubah pembaur dalam penelitian ini adalah diet tinggi lemak (HFD). Model regresi logistik biner diaplikasikan untuk menduga *propensity score* yang akan digunakan untuk memadankan kelompok yang melakukan HFD >100% AKG (kasus) dengan kelompok yang melakukan HFD 100% AKG (kontrol). Regresi logistik biner tidak menciptakan asumsi sebaran peubah penjelas pada hasil yang biner. Pemadanan dengan algoritma *Nearest Neighbor Matching caliper* berhasil mencapai keseimbangan pembaur antara kedua kelompok dan mereduksi bias sebesar 69,69%. Pemodelan akhir menghasilkan bahwa usia menopause 50 tahun ke atas ($p=0,007$) dengan *odds ratio* (OR) = 2,336 ($p=0,007$) dan riwayat menyusui kurang dari 12 bulan (OR=0,054; $p=0,004$) berpengaruh signifikan terhadap peningkatan risiko kanker payudara.

Kata kunci: kanker payudara, *nearest neighbor*, pembaur, *propensity score matching*, regresi logistik biner

ABSTRACT

ALIFA SHAKILA. Application of Propensity Score Matching Binary Logistic Regression Analysis in Breast Cancer Cases. Supervised by RAHMA ANISA and AAM ALAMUDI.

Female breast cancer is the leading cause of cancer-related deaths in Indonesia by 2022. Breast cancer is classified as a non-communicable disease with unknown exact cause. This observational study conducted on Indonesian women who do and do not suffer from breast cancer. This study aims to identify factors that influence breast cancer after overcoming the confounder that can deflect true relationship between risk factors and breast cancer status. This study utilized Propensity Score Matching method that capable of reducing confounder effect especially in health observational research. The confounder in this study is High-Fat Diet (HFD). Binary logistic regression model was applied to estimate propensity score that will be used to match a group on HFD >100% RDA (case) with a group on HFD 100% RDA (control). Binary logistic regression was chosen because it did not create distribution assumptions of explanatory variables on binary outcomes. Matching with Nearest Neighbor Matching algorithm with caliper achieved confounder balance between the two groups and reduced bias by 69,69%. Final modeling results that menopausal age 50 years and over ($p=0,007$) with odds ratio (OR) = 2,336 ($p=0,007$) and history of breastfeeding less than 12 months (OR=0,054; $p=0,004$) significantly influences the increased risk of breast cancer.

Keywords: binary logistic regression, breast cancer, confounder, nearest neighbor, propensity score matching



© Hak cipta milik IPB University

IPB University

© Hak Cipta milik IPB, tahun 2024
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PENERAPAN ANALISIS *PROPENSITY SCORE MATCHING* REGRESI LOGISTIK BINER PADA KASUS KANKER PAYUDARA

ALIFA SHAKILA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

**Tim Penguji pada Ujian Skripsi:
Dr. Yenni Angraini, S.Si., M.Si.**

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Skripsi : Penerapan Analisis *Propensity Score Matching* Regresi Logistik
Biner pada Kasus Kanker Payudara

Nama : Alifa Shakila
NIM : G1401201087

Disetujui oleh

Pembimbing 1:
Rahma Anisa, S.Stat., M.Si.



Pembimbing 2:
Ir. Aam Alamudi, M.S.



Diketahui oleh

Ketua Departemen Statistika:
Dr. Bagus Sartono, S.Si., M.Si.
NIP 197804112005011002



Tanggal Ujian:
8 Agustus 2024

Tanggal Lulus:
(tanggal penandatanganan oleh Dekan
Fakultas/Sekolah ...)



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Penelitian ini dilaksanakan sejak bulan Januari 2024 sampai bulan Juli 2024 dengan judul "Penerapan Analisis *Propensity Score Matching* Regresi Logistik Biner pada Kasus Kanker Payudara".

Terima kasih penulis ucapkan kepada seluruh pihak yang turut serta dalam membantu penulis menyelesaikan karya ilmiah ini, di antaranya:

1. Ibu Rahma Anisa, S.Stat., M.Si. dan Bapak Aam Alamudi, M.S. selaku komisi pembimbing yang telah memberikan saran, semangat, dan ilmu baru kepada penulis selama penyusunan karya ilmiah ini.
2. Bapak/Ibu selaku dosen penguji yang telah memberikan saran dan masukan dalam penulisan karya ilmiah ini.
3. Seluruh dosen dan staf Departemen Statistika FMIPA IPB yang telah membantu penulis selama masa kuliah sehingga berhasil menyelesaikan pendidikan sampai lulus.
4. Ayah, Bunda, dan keempat adik penulis, serta seluruh keluarga besar yang selalu memberikan perhatian dan dukungan tak terkira baik berupa materi maupun do'a.
5. Teman-teman Statistika 57 yang selalu bersedia memberikan masukan, do'a, dan motivasi kepada penulis selama berkuliah di IPB.
6. Seluruh orang kreatif yang telah memotivasi penulis dengan karya dan semangatnya, serta seluruh pihak yang tidak dapat penulis sebutkan satu persatu.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2024

Alifa Shakila

DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	x
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
II TINJAUAN PUSTAKA	3
2.1 Kanker Payudara	3
2.2 Regresi Logistik Biner	3
2.3 <i>Propensity Score Matching</i> (PSM)	5
2.4 <i>Nearest Neighbor Matching</i> (NNM)	6
2.5 <i>Percent Bias Reduction</i> (PBR)	8
2.6 Uji Keseimbangan Kovariat	8
2.7 <i>Average Treatment Effect on the Treated</i> (ATT)	8
III METODE	10
3.1 Data	10
3.2 Prosedur Analisis	10
IV HASIL DAN PEMBAHASAN	12
4.1 Eksplorasi Data	12
4.2 Pemeriksaan Kesesuaian <i>Confounder</i>	13
4.3 Pendugaan <i>Propensity Score</i> dengan Regresi Logistik Biner	13
4.4 <i>Matching</i> Data Menggunakan Algoritma NNM	14
4.5 Pemeriksaan Kualitas <i>Matching</i>	15
4.6 Uji Signifikansi Nilai Dugaan ATT	17
4.7 Pemodelan Setelah Penanganan Efek <i>Confounder</i>	18
V SIMPULAN DAN SARAN	20
5.1 Simpulan	20
5.2 Saran	20
DAFTAR PUSTAKA	21
LAMPIRAN	24
RIWAYAT HIDUP	30

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR TABEL

1	Ilustrasi <i>matching</i>	7
2	Peubah penelitian	10
3	Hubungan antara <i>confounder</i> potensial (Z) dan peubah lainnya	13
4	Pendugaan parameter model regresi logistik biner antara peubah penjelas (X) dan <i>confounder</i> (Z)	14
5	Hasil pendugaan <i>propensity score largest to smallest</i>	14
6	Jumlah unit berpasangan dan tidak berpasangan	14
7	Hasil <i>matching</i> kelompok kasus dan kelompok kontrol	15
8	Hasil perhitungan bias sebelum dan setelah <i>matching</i>	16
9	Hasil uji keseimbangan kovariat	16
10	Hasil uji signifikansi nilai dugaan ATT	17
11	Hasil uji signifikansi parameter secara simultan	18
12	Hasil uji signifikansi parameter secara parsial	18

DAFTAR GAMBAR

1	Karakteristik responden berdasarkan status kanker payudara	12
2	Keseimbangan kovariat <i>confounder</i> sebelum <i>matching</i>	17
3	Keseimbangan kovariat <i>confounder</i> setelah <i>matching</i>	17

DAFTAR LAMPIRAN

1	Hasil pendugaan <i>propensity score largest to smallest</i>	25
2	Hasil <i>matching</i> kelompok kasus dan kelompok kontrol	27