



**EPIDEMIOLOGI ANALITIK:
PENDEKATAN KUANTITATIF DALAM
PENELITIAN KESEHATAN**

Sanksi Pelanggaran Pasal 113

Undang-Undang No. 28 Tahun 2014 Tentang Hak Cipta

- I. Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp 100.000.000 (seratus juta rupiah).
- II. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp 500.000.000,00 (lima ratus juta rupiah).
- III. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp1.000.000.000,00 (satu miliar rupiah).
- IV. Setiap Orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan/atau pidana denda paling banyak Rp 4.000.000.000,00 (empat miliar rupiah).

EPIDEMIOLOGI ANALITIK: PENDEKATAN KUANTITATIF DALAM PENELITIAN KESEHATAN

Dewi Ariyani Wulandari | Dian Rezki Wijaya |
Rimawati Aulia Insani Sadarang | Vitasari Indriani |
Dian Ihwana Ansyar | Rojali | Fitri Rachmillah Fadmi |
Bs. Titi Haerana | Debi Susanti

Editor
Ronald



EPIDEMIOLOGI ANALITIK : PENDEKATAN KUANTITATIF DALAM PENELITIAN KESEHATAN

Penulis:

**Dewi Ariyani Wulandari | Dian Rezki Wijaya |
Rimawati Aulia Insani Sadarang | Vitasari Indriani |
Dian Ihwana Ansyar | Rojali | Fitri Rachmillah Fadmi |
Bs. Titi Haerana | Debi Susanti**

Editor:

Ronald

Tata Letak:

Bhaskara B.B. Barung

Desainer:

Tim SMI-Kesehatan

Sumber Gambar Cover:

www.freepik.com

Ukuran:

vii, 193 hlm., 15,5 cm x 23 cm

ISBN:

Cetakan Pertama:

Mei 2026

Hak Cipta Dilindungi oleh Undang-undang. Dilarang menerjemahkan, memfotokopi, atau memperbanyak sebagian atau seluruh isi buku ini tanpa izin tertulis dari Penerbit.

PT. KAKINAAN PUSTAKA UTAMA

Kantor Cabang:

Sleman, Daerah Istimewa Yogyakarta, 55581

Kontak: 0813-5604-7006

Website: <https://kakinaan.com>

E-mail: admin@kakinaan.com

KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga buku Epidemiologi Analitik: Pendekatan Kuantitatif dalam Penelitian Kesehatan ini dapat diselesaikan dan hadir ke tangan para pembaca.

Epidemiologi analitik merupakan salah satu pilar utama dalam ilmu kesehatan masyarakat. Kemampuan untuk merancang studi yang tepat, mengukur frekuensi penyakit secara akurat, serta menganalisis data dengan metode yang sah adalah kompetensi yang mutlak diperlukan oleh setiap peneliti dan praktisi di bidang kesehatan. Namun demikian, tidak sedikit akademisi maupun mahasiswa yang masih menghadapi tantangan dalam memahami pendekatan kuantitatif secara menyeluruh dan terpadu. Buku ini lahir dari kebutuhan nyata tersebut.

Buku ini disusun secara sistematis untuk memandu pembaca mulai dari fondasi konseptual hingga penerapan metode analisis yang mutakhir. Diawali dengan pemahaman dasar epidemiologi analitik, pembahasan berlanjut pada ukuran frekuensi penyakit, berbagai desain studi utama — kohort, case-control, cross-sectional, dan ekologi — serta uji klinis dan studi eksperimental. Pada bagian akhir, disajikan metode analisis statistik lanjutan yang sangat relevan dalam riset epidemiologi modern, yakni regresi logistik, analisis survival, model Cox Proportional Hazard, serta inferensi kausal dengan pendekatan kontemporer.

Buku ini merupakan buah dari kolaborasi sembilan penulis yang memiliki latar belakang keilmuan dan pengalaman penelitian di bidang epidemiologi dan kesehatan masyarakat. Setiap bab ditulis dengan semangat untuk menyajikan materi yang tidak hanya teoritis, tetapi juga aplikatif dan mudah dipahami. Kami berharap buku ini dapat menjadi sumber belajar yang andal bagi mahasiswa strata satu maupun pascasarjana, peneliti, dosen, serta tenaga kesehatan yang aktif dalam kegiatan riset.

Kami menyadari bahwa buku ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun dari para

pembaca sangat kami harapkan demi penyempurnaan pada edisi-edisi berikutnya. Semoga kehadiran buku ini memberikan kontribusi nyata bagi pengembangan ilmu epidemiologi di Indonesia dan pada akhirnya berdampak pada peningkatan derajat kesehatan masyarakat.

Akhir kata, kami menyampaikan terima kasih yang sebesar-besarnya kepada seluruh pihak yang telah mendukung penyelesaian buku ini — rekan-rekan penulis, keluarga, serta institusi yang telah memberikan ruang dan kesempatan untuk berkarya. Semoga karya ini bermanfaat dan menjadi amal ilmu yang tiada putus pahalanya..

Tim Penulis

DAFTAR ISI

Kata Pengantar (v)

Daftar Isi..... (vi)

01	Fondasi Epidemiologi Analitik <i>Dewi Ariyani Wulandari</i>	(01)
02	Ukuran Frekuensi Penyakit: Insidensi, Prevalensi, dan Mortalitas <i>Dian Rezki Wijaya</i>	(18)
03	Desain Studi Kohort: Prospektif dan Retrospektif <i>Rimawati Aulia Insani Sadarang</i>	(42)
04	Desain Studi Case-Control: Prinsip dan Aplikasi <i>Vitasari Indriani</i>	(64)
05	Desain Studi Cross-Sectional dan Ekologi <i>Dian Ihwana Ansyar</i>	(81)
06	Uji Klinis dan Studi Eksperimental dalam Epidemiologi <i>Rojali</i>	(98)
07	Regresi Logistik dalam Epidemiologi <i>Fitri Rachmillah Fadmi</i>	(125)
08	Analisis Survival dan Model Cox Proportional Hazard <i>Bs. Titi Haerana</i>	(148)
09	Inferensi Kausal: Teori dan Metode Kontemporer <i>Debi Susanti</i>	(177)

1



FONDASI EPIDEMIOLOGI ANALITIK

DEWI ARIYANI WULANDARI

E-mail: dewiwulan02@gmail.com

PENDAHULUAN

Epidemiologi merupakan ilmu yang mempelajari distribusi dan determinan kejadian penyakit dalam populasi serta penerapannya dalam upaya pencegahan dan pengendalian masalah kesehatan. Dalam perkembangannya, epidemiologi tidak hanya berfokus pada penggambaran pola penyakit, tetapi juga pada analisis hubungan antara berbagai faktor yang mempengaruhi terjadinya penyakit. Pendekatan ini dikenal sebagai epidemiologi analitik, yang berperan penting dalam memahami faktor-faktor yang berkontribusi terhadap masalah kesehatan di masyarakat. Seiring dengan meningkatnya kompleksitas masalah kesehatan, baik penyakit menular maupun tidak menular, diperlukan pendekatan yang mampu menjelaskan hubungan antara berbagai determinan kesehatan secara lebih mendalam.

Penyakit tidak lagi dipandang sebagai akibat dari satu faktor tunggal, melainkan sebagai hasil interaksi antara faktor biologis, lingkungan, perilaku, dan sosial. Oleh karena itu, pemahaman mengenai konsep sebab-akibat, serta hubungan antara paparan dan outcome, menjadi sangat penting dalam mengkaji kejadian penyakit secara komprehensif. Dalam konteks penelitian kesehatan,

epidemiologi analitik berkontribusi dalam mengidentifikasi faktor risiko, mengevaluasi hubungan antar variabel, serta mendukung pengambilan keputusan berbasis bukti. Pendekatan ini memungkinkan dihasilkannya informasi yang lebih akurat untuk perencanaan intervensi kesehatan, sehingga menjadi landasan penting dalam upaya peningkatan derajat kesehatan masyarakat secara berkelanjutan.

DEFINISI DAN RUANG LINGKUP EPIDEMIOLOGI ANALITIK

Epidemiologi merupakan ilmu yang mempelajari distribusi dan determinan kejadian penyakit atau masalah kesehatan dalam populasi serta penerapannya untuk pengendalian masalah kesehatan. Dalam perkembangan disiplin ini, epidemiologi terbagi menjadi dua cabang utama, yaitu epidemiologi deskriptif dan epidemiologi analitik. Epidemiologi analitik secara khusus berfokus pada identifikasi hubungan sebab-akibat antara paparan (exposure) dan kejadian penyakit atau outcome kesehatan.

Epidemiologi analitik dapat didefinisikan sebagai cabang epidemiologi yang bertujuan untuk menguji hipotesis mengenai hubungan antara faktor risiko dengan

kejadian penyakit melalui pendekatan ilmiah dan metode penelitian yang sistematis. Berbeda dengan epidemiologi deskriptif yang hanya menggambarkan distribusi penyakit berdasarkan orang, tempat, dan waktu, epidemiologi analitik berusaha menjawab pertanyaan “mengapa” dan “bagaimana” suatu penyakit terjadi.

Ruang lingkup epidemiologi analitik mencakup berbagai aspek penting, antara lain (Sastroasmoro, 2018):

1. Identifikasi faktor risiko penyakit
2. Pengujian hubungan antara paparan dan outcome
3. Penilaian kekuatan asosiasi (misalnya melalui *Odds Ratio* dan *Relative Risk*)
4. Pengendalian faktor perancu (*confounding*)
5. Pengujian hipotesis kausal
6. Evaluasi efektivitas intervensi kesehatan

Dalam praktiknya, epidemiologi analitik menggunakan berbagai desain penelitian seperti studi kohort, studi kasus-kontrol, dan uji klinis (eksperimental). Desain-desain ini memungkinkan peneliti untuk menilai apakah suatu paparan benar-benar berkontribusi terhadap terjadinya penyakit.

Sebagai contoh, dalam konteks tuberkulosis (TB), epidemiologi analitik dapat digunakan untuk menguji apakah paparan asap rokok dalam rumah tangga

meningkatkan risiko infeksi TB pada anak dan hasilnya adalah asap rokok melemahkan sistem kekebalan tubuh dalam melawan infeksi TB (Altet et al., 2022). Melalui pendekatan ini, peneliti tidak hanya mengetahui bahwa TB lebih banyak terjadi pada kelompok tertentu, tetapi juga dapat menjelaskan faktor penyebab yang mendasarinya.

PERBEDAAN EPIDEMIOLOGI DESKRIPTIF DAN ANALITIK

Epidemiologi deskriptif dan analitik memiliki tujuan dan pendekatan yang berbeda, meskipun keduanya saling melengkapi dalam memahami masalah kesehatan masyarakat. Epidemiologi deskriptif berfokus pada penggambaran distribusi penyakit berdasarkan karakteristik populasi, seperti usia, jenis kelamin, lokasi geografis, dan waktu kejadian. Tujuan utamanya adalah untuk mengidentifikasi pola dan tren penyakit serta menghasilkan hipotesis awal. Misalnya, peningkatan kasus TB di suatu wilayah tertentu dalam kurun waktu tertentu dapat diidentifikasi melalui pendekatan deskriptif. Sebaliknya, epidemiologi analitik bertujuan untuk menguji hipotesis yang dihasilkan dari studi deskriptif. Pendekatan ini melibatkan perbandingan antara kelompok yang

terpapar dan tidak terpapar terhadap suatu faktor risiko untuk menentukan apakah terdapat hubungan yang bermakna secara statistik.

Perbedaan utama antara keduanya dapat dijelaskan sebagai berikut :

Tabel 1.1

Perbedaan Epidemiologi Deskriptif dan Analitik

No.	Aspek	Epidemiologi Deskriptif	Epidemiologi Analitik
1.	Tujuan	menggambarkan distribusi penyakit	menjelaskan hubungan sebab-akibat
2.	Pendekatan	menggunakan data agregat dan observasi umum	menggunakan desain penelitian komparatif.
3.	Pertanyaan utama	menjawab “siapa, di mana, dan kapan”	analitik menjawab “mengapa dan bagaimana”
4.	Output	menghasilkan hipotesis	menguji hipotesis

Sumber : (Murti, 2018)

Sebagai ilustrasi, jika ditemukan bahwa prevalensi TB lebih tinggi pada kelompok masyarakat dengan status sosial ekonomi rendah (Pandey et al., 2025), maka epidemiologi deskriptif hanya mencatat fenomena tersebut. Namun, epidemiologi analitik akan meneliti lebih lanjut apakah

faktor seperti kepadatan hunian, malnutrisi, atau akses layanan kesehatan menjadi penyebab utama peningkatan risiko tersebut. Dengan demikian, epidemiologi deskriptif dan analitik merupakan dua tahapan penting dalam siklus penelitian epidemiologi yang tidak dapat dipisahkan.

KONSEP SEBAB - AKIBAT DALAM EPIDEMIOLOGI

Konsep sebab-akibat (causality) merupakan inti dari epidemiologi analitik. Tidak semua hubungan antara paparan dan outcome bersifat kausal. Oleh karena itu, diperlukan pendekatan ilmiah untuk menentukan apakah suatu hubungan benar-benar bersifat sebab-akibat atau hanya kebetulan (*spurious*).

Dalam epidemiologi, penyebab penyakit umumnya bersifat multifaktorial, artinya satu penyakit dapat disebabkan oleh kombinasi berbagai faktor, baik faktor biologis, lingkungan, perilaku, maupun sosial. Model yang sering digunakan untuk menjelaskan hal ini adalah model “*web of causation*”, yang menggambarkan bahwa penyakit merupakan hasil interaksi kompleks dari berbagai faktor (Najmah, 2019).

Untuk menilai hubungan kausal, epidemiologi menggunakan kriteria tertentu, salah satunya adalah

kriteria Bradford Hill, yang meliputi (Shimonovich et al., 2021) :

1. **Kekuatan asosiasi (*strength*)** yaitu Kekuatan asosiasi menunjukkan besarnya hubungan antara paparan dan outcome, di mana hubungan yang lebih kuat lebih mendukung adanya kausalitas.
2. **Konsistensi (*consistency*)** yaitu kondisi di mana hubungan antara paparan dan outcome ditemukan secara berulang pada berbagai penelitian, populasi, dan situasi.
3. **Spesifisitas (*specificity*)** yaitu kondisi di mana satu paparan berhubungan dengan satu outcome tertentu secara khusus.
4. **Temporalitas (*temporality*)** yaitu prinsip yang menyatakan bahwa paparan harus terjadi sebelum timbulnya outcome.
5. **Gradien biologis (*dose-response*)** yaitu peningkatan tingkat paparan diikuti dengan peningkatan risiko terjadinya outcome.
6. **Plausibilitas biologis (*biological plausibility*)** yaitu prinsip adanya penjelasan ilmiah yang masuk akal mengenai bagaimana suatu paparan dapat menyebabkan suatu outcome.

7. **Koherensi (*coherence*)** yaitu prinsip yang menyatakan bahwa hubungan antara paparan dan outcome harus konsisten dengan pengetahuan ilmiah, fakta, dan teori yang sudah ada mengenai penyakit tersebut.
8. **Eksperimen (*experiment*)** yaitu prinsip yang menyatakan bahwa hubungan antara paparan dan outcome harus konsisten dengan pengetahuan ilmiah, fakta, dan teori yang sudah ada mengenai penyakit tersebut.
9. **Analogi (*analogy*)** prinsip yang menyatakan bahwa hubungan kausal dapat didukung jika terdapat hubungan serupa yang telah terbukti pada kondisi atau paparan yang mirip.

Dari semua kriteria tersebut, temporalitas merupakan syarat mutlak dalam menentukan hubungan kausal. Tanpa adanya urutan waktu yang jelas antara paparan dan outcome, hubungan sebab-akibat tidak dapat ditegakkan. Sebagai contoh, dalam penelitian TB, paparan terhadap penderita TB aktif harus terjadi sebelum seseorang mengalami infeksi TB. Jika tidak, maka hubungan tersebut tidak dapat dikatakan kausal. Selain itu, penting untuk membedakan antara asosiasi dan kausalitas. Asosiasi hanya menunjukkan adanya

hubungan statistik antara dua variabel, sedangkan kausalitas menunjukkan bahwa satu variabel menyebabkan perubahan pada variabel lain.

VARIABEL PAPARAN DAN OUTCOME

Dalam epidemiologi analitik, konsep variabel sangat penting karena menjadi dasar dalam analisis hubungan antara faktor risiko dan penyakit (Murti, 2018; Najmah, 2019; Ye, 2023).

1. Variabel Paparan (Exposure)

Variabel paparan adalah faktor yang diduga mempengaruhi terjadinya suatu penyakit atau outcome.

Paparan dapat berupa:

- a) Faktor lingkungan (misalnya polusi udara)
- b) Faktor perilaku (misalnya merokok)
- c) Faktor biologis (misalnya status imun)
- d) Faktor sosial (misalnya tingkat pendidikan)

Paparan dapat bersifat:

- a) Dikotomik (ya/tidak)
- b) Kategorik (rendah, sedang, tinggi)
- c) Kontinu (misalnya kadar zat tertentu dalam darah)

Dalam penelitian epidemiologi, pengukuran paparan harus dilakukan secara valid dan reliabel untuk menghindari bias.

2. Variabel Outcome

Outcome adalah hasil atau kejadian yang diteliti, biasanya berupa penyakit atau kondisi kesehatan tertentu. Outcome dapat berupa:

- Kejadian penyakit
- Kematian
- Status kesehatan tertentu
- Perubahan perilaku

Outcome juga dapat diklasifikasikan sebagai:

- Biner (sakit/tidak sakit)
- Ordinal (ringan, sedang, berat)
- Kontinu (misalnya tekanan darah)

3. Hubungan antara Paparan dan Outcome

Tujuan utama epidemiologi analitik adalah menilai hubungan antara paparan dan outcome. Hubungan ini diukur menggunakan ukuran asosiasi seperti:

- *Relative Risk* (RR) pada studi kohort
- *Odds Ratio* (OR) pada studi kasus-kontrol

Sebagai contoh, jika anak yang terpapar asap rokok memiliki risiko TB dua kali lebih besar dibandingkan yang tidak terpapar, maka $RR = 2$. Hal ini

menunjukkan adanya asosiasi yang signifikan antara paparan dan outcome.

4. Variabel Perancu dan Efek Modifikasi

Selain paparan dan outcome, terdapat variabel lain yang dapat mempengaruhi hubungan tersebut, yaitu:

- ***Confounding* (perancu)**: variabel yang berhubungan dengan paparan dan outcome
- ***Effect modifier* (modifikator efek)**: variabel yang mengubah kekuatan hubungan

Misalnya, status gizi dapat menjadi variabel perancu dalam hubungan antara kemiskinan dan TB.

PERAN EPIDEMIOLOGI ANALITIK DALAM PENELITIAN KESEHATAN

Epidemiologi analitik memiliki peran yang sangat penting dalam penelitian kesehatan masyarakat, terutama dalam menghasilkan bukti ilmiah yang dapat digunakan untuk pengambilan keputusan. Beberapa peran utama epidemiologi analitik antara lain (Aschengrau & Seage III, 2020; Murti, 2018; Najmah, 2019):

1. Identifikasi Faktor Risiko

Epidemiologi analitik membantu mengidentifikasi faktor-faktor yang berkontribusi terhadap terjadinya

penyakit. Informasi ini penting untuk perencanaan program pencegahan.

2. Pengembangan Intervensi

Dengan mengetahui faktor penyebab penyakit, intervensi dapat dirancang secara lebih tepat sasaran. Misalnya, jika ditemukan bahwa kepadatan hunian berkontribusi terhadap TB, maka intervensi dapat difokuskan pada perbaikan lingkungan rumah.

3. Evaluasi Program Kesehatan

Epidemiologi analitik digunakan untuk menilai efektivitas program kesehatan, seperti program skrining atau vaksinasi.

4. Mendukung Kebijakan Kesehatan

Hasil penelitian epidemiologi analitik menjadi dasar dalam penyusunan kebijakan berbasis bukti (*evidence-based policy*).

5. Pengendalian Penyakit

Melalui pemahaman faktor risiko, strategi pengendalian penyakit dapat dilakukan secara lebih efektif.

Dalam konteks TB, epidemiologi analitik sangat berperan dalam memahami rendahnya partisipasi masyarakat dalam skrining, mengidentifikasi hambatan

seperti stigma, serta merancang intervensi berbasis komunitas.

PENUTUP

Epidemiologi analitik merupakan fondasi penting dalam penelitian kesehatan masyarakat yang berfokus pada pengujian hubungan sebab-akibat antara paparan dan outcome. Dengan memahami konsep-konsep dasar seperti variabel, kausalitas, dan desain penelitian, peneliti dapat menghasilkan bukti ilmiah yang kuat untuk mendukung upaya pencegahan dan pengendalian penyakit.

Kombinasi antara epidemiologi deskriptif dan analitik memungkinkan pendekatan yang komprehensif dalam memahami masalah kesehatan, mulai dari identifikasi pola hingga penentuan faktor penyebab. Oleh karena itu, penguasaan epidemiologi analitik menjadi keterampilan esensial bagi peneliti dan praktisi kesehatan masyarakat

DAFTAR PUSTAKA

- Altet, N., Latorre, I., Jiménez-Fuentes, M. Á., Soriano-Arandes, A., Villar-Hernández, R., Milà, C., Rodríguez-Fernández, P., Muriel-Moreno, B., Comella-del-Barrio, P., Godoy, P., Millet, J.-P., de Souza-Galvão, M. L., Jiménez-Ruiz, C. A., Domínguez, J., & on behalf of PII Smoking SEPAR Working Group. (2022). Tobacco Smoking and Second-Hand Smoke Exposure Impact on Tuberculosis in Children. *Journal of Clinical Medicine*, 11(7), 2000. <https://doi.org/10.3390/jcm11072000>
- Aschengrau, A., & Seage III, G. R. (2020). *Essential of Epidemiology in Public Health* (4 th Edition). Jones & Bartlett Learning, LLC, an Ascend Learning Company.
- Murti, B. (2018). *Prinsip dan metode riset epidemiologi (Epidemiological research principles and methods)* (Vol. 5). Surakarta: Program Studi Ilmu Kesehatan Masyarakat, Program Pascasarjana, Universitas Sebelas Maret.
- Najmah, N. (2019). *Epidemiologi: Untuk Mahasiswa Kesehatan Masyarakat*. Depok : Rajawali Press.
- Pandey, A. K., Chaturvedi, H. K., & Choraria, P. (2025). Relationship between Household Tuberculosis and Socioeconomic and Bioenvironmental Factors: A Statistical Model Approach Using NFHS-5 Data. *Indian Journal of Community Medicine*, 50(4). https://journals.lww.com/ijcm/fulltext/2025/07000/relationship_between_household_tuberculosis_and.24.aspx
- Sastroasmoro, S. (2018). *Dasar-Dasar Metodologi Penelitian Klinis* (Edisi ke-4). Jakarta : Sagung Seto.
- Shimonovich, M., Pearce, A., Thomson, H., Keyes, K., & Katikireddi, S. V. (2021). Assessing causality in epidemiology: Revisiting Bradford Hill to incorporate

developments in causal thinking. *European Journal of Epidemiology*, 36(9), 873–887.
<https://doi.org/10.1007/s10654-020-00703-7>

Ye, L. (2023). Cause of Disease and Causal Inference. In C. Wang & F. Liu (Eds.), *Textbook of Clinical Epidemiology: For Medical Students* (pp. 153–171). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-3622-9_9

BIODATA PENULIS



Dewi Ariyani Wulandari, S.K.M.,M.P.H Penulis lahir di Palembang pada 9 Juni 1984. Ia menyelesaikan pendidikan Sarjana Kesehatan Masyarakat di Universitas Sriwijaya (2006) dan meraih gelar Master of Public Health dari Universitas Gadjah Mada (2012), serta saat ini tengah menempuh studi doktoral di bidang Kesehatan Masyarakat di Universitas Diponegoro. Penulis merupakan dosen tetap pada Program Studi S1 Kesehatan Masyarakat STIKES Wira Husada Yogyakarta dan aktif dalam organisasi profesi sebagai Pengurus Harian IAKMI Daerah D.I. Yogyakarta. Fokus keilmuan penulis meliputi epidemiologi, khususnya pendekatan berbasis komunitas, pemberdayaan masyarakat, analisis perilaku kesehatan, serta manajemen dan pembiayaan kesehatan. Sepanjang karier akademiknya, penulis telah memperoleh berbagai hibah penelitian dan pengabdian kepada masyarakat, serta aktif sebagai narasumber dalam pelatihan di Puskesmas dan rumah sakit.

2



UKURAN FREKUENSI PENYAKIT : INSIDENSI, PREVALENSI DAN MORTALITAS

DIAN REZKI WIJAYA

E-mail: dianrezky.wijaya@uin-alauddin.ac.id

PENDAHULUAN

Pemahaman mengenai pola dan distribusi penyakit dalam populasi merupakan dasar penting dalam perencanaan intervensi kesehatan masyarakat yang efektif. Salah satu cara untuk menggambarkan besarnya masalah kesehatan tersebut adalah melalui ukuran frekuensi penyakit, yaitu ukuran yang menunjukkan seberapa sering suatu penyakit, kejadian kesehatan, atau kematian terjadi pada suatu populasi dalam kurung waktu tertentu. Pada kajian epidemiologi, ukuran frekuensi penyakit umumnya dinyatakan dalam bentuk rasio, proporsi, dan rate yang menjadi dasar dalam perhitungan berbagai indikator kesehatan serta memungkinkan perbandingan antar kelompok populasi, wilayah, maupun periode waktu yang berbeda.

Ukuran frekuensi penyakit juga dinyatakan melalui indikator epidemiologi yang lebih spesifik, yaitu insidensi, prevalensi dan mortalitas. Insidensi merupakan ukuran yang menggambarkan banyaknya kasus baru suatu penyakit yang terjadi dalam populasi selama periode waktu tertentu. Oleh karena itu, ukuran ini sering dimanfaatkan untuk menilai tingkat risiko terjadinya penyakit serta membantu peneliti dalam mengidentifikasi

faktor risiko yang berperan dan menetapkan prioritas intervensi kesehatan masyarakat.

Sementara itu, prevalensi menggambarkan keseluruhan jumlah kasus penyakit yang ada dalam suatu populasi pada waktu tertentu, baik yang merupakan kasus baru maupun kasus yang telah berlangsung sebelumnya. Ukuran ini memberikan gambaran mengenai besarnya beban penyakit dalam masyarakat dan sering digunakan sebagai dasar dalam perencanaan kebutuhan pelayanan kesehatan.

Mortalitas adalah ukuran yang digunakan untuk menunjukkan frekuensi kematian yang terjadi dalam populasi pada periode waktu tertentu. Analisis mortalitas menggambarkan tingkat keparahan suatu penyakit serta dapat digunakan untuk mengevaluasi efektivitas sistem pelayanan dan kebijakan kesehatan. Dengan memahami ukuran frekuensi penyakit peneliti, tenaga kesehatan, dan pembuat kebijakan dapat memantau tren penyakit, menilai tingkat keparahan masalah kesehatan, serta merumuskan strategi pencegahan dan pengendalian penyakit secara lebih tepat.

KONSEP DASAR UKURAN FREKUENSI

Ukuran frekuensi penyakit merupakan konsep fundamental yang digunakan untuk menunjukkan seberapa sering suatu kejadian kesehatan muncul dalam suatu populasi. Konsep ini menjadi landasan dalam proses analisis epidemiologi karena melalui pengukuran frekuensi, peneliti dapat memahami pola kejadian penyakit, mengidentifikasi kelompok populasi yang berisiko, serta menilai besarnya masalah kesehatan dalam masyarakat. Tanpa adanya ukuran frekuensi yang jelas, upaya analisis distribusi penyakit dan penentuan prioritas intervensi kesehatan akan sulit dilakukan secara objektif dan terukur (Friis & Sellers, 2021).

Secara umum, ukuran frekuensi penyakit menggambarkan hubungan antara jumlah kejadian penyakit dengan populasi yang memiliki risiko terhadap suatu kejadian dalam kurun waktu tertentu. Dalam konteks ini, epidemiologi tidak hanya menghitung jumlah kasus secara absolut, tetapi juga mempertimbangkan besarnya populasi yang menjadi dasar perhitungan. Oleh karena itu, ukuran frekuensi biasanya dinyatakan dalam bentuk rasio, proporsi, atau rate (laju). Pendekatan ini memungkinkan perbandingan kejadian penyakit

antarwilayah, antarperiode waktu, maupun antarkelompok populasi yang berbeda.

Rasio adalah ukuran yang menunjukkan perbandingan antara dua nilai yang tidak memiliki hubungan bagian terhadap keseluruhan.

$$\text{Rasio} = \frac{\text{Kejadian/Kelompok A}}{\text{Kejadian /Kelompok B}}$$

Rasio sering digunakan untuk membandingkan dua kelompok populasi atau dua jenis kejadian kesehatan. Misalnya, jika suatu penyakit ditemukan pada 200 laki-laki dan 100 perempuan, hasil perbandingan menunjukkan rasio **2 : 1** antara laki-laki dan perempuan, yang mengindikasikan bahwa jumlah kasus pada laki-laki dua kali lipat dibandingkan perempuan..

Proporsi adalah ukuran yang menggambarkan bagian dari suatu keseluruhan, sehingga nilai pembilang selalu merupakan bagian dari penyebut.

$$\text{Proporsi} = \frac{\text{Jumlah kejadian dalam kelompok tertentu}}{\text{Jumlah seluruh kejadian yang diamati}} \times k$$

Proporsi sering digunakan untuk menggambarkan distribusi kasus penyakit dalam suatu populasi. Misalnya, jika dalam suatu populasi berjumlah 1.000 orang ditemukan 100 orang yang menderita hipertensi, maka proporsi penderita hipertensi adalah 10%.

Rate menunjukkan besarnya kejadian kesehatan yang terjadi dalam populasi tertentu selama periode waktu tertentu

$$\text{Rate} = \frac{\text{Jumlah kejadian dalam periode tertentu}}{\text{Jumlah populasi berisiko pada periode tersebut}} \times k$$

Konsep dasar ukuran frekuensi penyakit juga berkaitan erat dengan tiga komponen utama, yaitu populasi yang diamati, jumlah kejadian penyakit, dan dimensi waktu. Populasi yang diamati merujuk pada kelompok individu yang memiliki kemungkinan mengalami suatu kejadian kesehatan. Jumlah kejadian penyakit menggambarkan banyaknya kasus yang terjadi dalam populasi tersebut, baik berupa kejadian kasus baru maupun total kasus yang telah terdapat dalam populasi. Sementara itu, dimensi waktu menunjukkan periode terjadinya kejadian penyakit, yang dapat berupa titik

waktu tertentu atau rentang waktu tertentu (Rothman et al., 2014; Somerville, 2012).

Dalam praktik epidemiologi, ukuran frekuensi penyakit memiliki beberapa fungsi antara lain untuk menggambarkan besarnya masalah kesehatan dalam suatu populasi, memantau perubahan pola penyakit dari waktu ke waktu, membandingkan kejadian penyakit antarkelompok populasi yang berbeda, serta mendukung proses perencanaan dan evaluasi program kesehatan masyarakat. Melalui pemanfaatan ukuran frekuensi penyakit secara tepat, para peneliti dan pengambil kebijakan dapat memperoleh gambaran yang lebih komprehensif mengenai kondisi kesehatan masyarakat serta menentukan strategi intervensi yang lebih efektif.

Konsep dasar ukuran frekuensi tidak hanya memberikan kerangka analitis untuk memahami distribusi penyakit dalam populasi, tetapi juga menjadi landasan dalam penggunaan indikator epidemiologi yang lebih spesifik, seperti insidensi, prevalensi, dan mortalitas, dalam analisis kesehatan masyarakat.

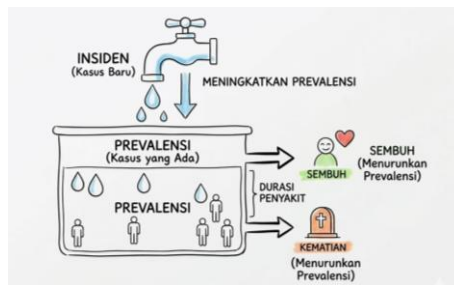
Prevalensi dan insidensi merupakan dua indikator utama yang digunakan untuk menggambarkan frekuensi terjadinya penyakit dalam suatu populasi. Kedua ukuran ini saling berkaitan, namun masing-masing

merepresentasikan dimensi yang berbeda dari kejadian penyakit. Insidensi menunjukkan jumlah kasus baru yang muncul dalam populasi selama periode waktu tertentu. Sebaliknya, prevalensi menggambarkan total kasus penyakit yang terdapat dalam populasi pada waktu atau periode tertentu, yang mencakup baik kasus yang sudah ada sebelumnya maupun kasus yang baru terjadi (Friis & Sellers, 2021; Somerville, 2012).

Secara konseptual, hubungan antara prevalensi dan insidensi dapat digambarkan melalui persamaan epidemiologi sederhana yaitu

$$P = I \times D$$

Dimana P (prevalensi) dipengaruhi oleh I (insidens) dan D (durasi penyakit).



Gambar 2.1 Hubungan Prevalensi dan Insidensi

Kondisi tersebut menunjukkan bahwa prevalensi suatu penyakit dapat meningkat apabila jumlah kasus baru

bertambah atau apabila durasi penyakit berlangsung lebih lama dalam populasi. Sebaliknya, prevalensi akan menurun apabila jumlah kasus baru berkurang atau ketika durasi penyakit menjadi lebih singkat, misalnya akibat proses penyembuhan yang lebih cepat ataupun meningkatnya angka kematian yang disebabkan oleh penyakit tersebut.

INSIDENSI KUMULATIF DAN INSIDENSI RATE

Insidensi digunakan sebagai ukuran untuk menunjukkan jumlah kejadian kasus baru penyakit yang timbul pada populasi yang memiliki risiko selama periode pengamatan tertentu. Ukuran ini memberikan gambaran mengenai tingkat risiko individu dalam suatu populasi untuk mengalami suatu penyakit. Insidensi sering digunakan dalam penelitian epidemiologi analitik untuk menilai hubungan antara faktor risiko dengan kejadian penyakit, serta untuk mengevaluasi efektivitas intervensi pencegahan yang dilakukan pada populasi tertentu (Friis & Sellers, 2021). Secara umum, insidensi dinyatakan menjadi dua yaitu insidensi kumulatif (*Cumulative Incidence*) dan laju insidensi (*Incidence Rate*).

Insidensi kumulatif (CI) adalah proporsi individu dalam populasi yang awalnya bebas dari penyakit tetapi kemudian mengalami penyakit tersebut selama periode waktu tertentu. Ukuran ini menggambarkan probabilitas atau risiko terjadinya penyakit pada suatu populasi selama periode pengamatan yang nilainya berada antara 0 sampai 1. Perhitungan insidensi kumulatif dapat dinyatakan sebagai berikut:

$$CI = \frac{\text{Jumlah yang menderita penyakit selama periode tertentu}}{\text{jumlah orang tanpa penyakit dalam populasi berisiko di awal periode}}$$

Ukuran ini biasanya digunakan dalam penelitian kohort dengan periode pengamatan yang jelas dan populasi yang relatif stabil. Misalnya, dalam studi mengenai kejadian penyakit tertentu pada kelompok pekerja selama satu tahun, insidensi kumulatif dapat menggambarkan proporsi pekerja yang mengalami penyakit tersebut selama periode pengamatan.

Sementara itu, laju insidensi (IR) merupakan ukuran yang memperhitungkan waktu individu berada dalam kondisi berisiko terhadap penyakit. Ukuran ini digunakan ketika populasi yang diamati bersifat dinamis, di mana individu dapat masuk atau keluar dari populasi

selama periode pengamatan. Dalam perhitungan insidensi rate, penyebutnya bukan hanya jumlah individu, tetapi total waktu paparan atau *person-time* yang dialami oleh seluruh individu dalam populasi berisiko.

Konsep *person-time* merujuk pada akumulasi waktu yang dihabiskan oleh setiap individu dalam keadaan berisiko terhadap penyakit. Misalnya, jika sepuluh orang diamati selama satu tahun, maka total *person-time* adalah sepuluh orang-tahun (*person-years*). Namun, apabila sebagian individu keluar dari pengamatan sebelum periode berakhir, maka waktu yang dihitung hanya selama individu tersebut berada dalam penelitian (Carneiro & Howard, 2011; Rothman et al., 2014). Insidensi rate dapat dirumuskan sebagai berikut:

$$IR = \frac{\text{Jumlah kasus baru dalam periode tertentu}}{\text{Jumlah orang yang terpapar risiko selama periode tersebut}} \times k$$

PREVALENSI TITIK DAN PERIODE

Prevalensi merupakan salah satu indikator frekuensi penyakit yang digunakan untuk menunjukkan total jumlah kasus penyakit yang terdapat dalam suatu populasi pada waktu tertentu, yang mencakup baik kasus

baru maupun kasus yang telah ada sebelumnya. Oleh karena itu, prevalensi sering dimanfaatkan untuk menggambarkan besarnya beban penyakit dalam populasi serta menjadi dasar dalam perencanaan pelayanan kesehatan dan pengalokasian sumber daya Kesehatan (Friis & Sellers, 2021).

Prevalensi dipengaruhi oleh dua faktor utama, yaitu insidensi penyakit dan lamanya penyakit berlangsung. Penyakit dengan angka insidensi tinggi dan durasi penyakit yang panjang cenderung menghasilkan prevalensi yang tinggi pula. Sebaliknya, penyakit dengan durasi singkat atau tingkat kesembuhan yang tinggi biasanya memiliki prevalensi yang lebih rendah meskipun angka insidensinya tinggi. Hubungan ini menunjukkan bahwa prevalensi tidak hanya menggambarkan frekuensi kejadian penyakit, tetapi juga mencerminkan dinamika perjalanan penyakit dalam populasi (Gordis, 2020). Prevalensi dibedakan menjadi dua jenis yaitu prevalensi titik (*point prevalence*) dan prevalensi periode (*period prevalence*).

Prevalensi titik adalah proporsi individu dalam suatu populasi yang mengalami suatu penyakit pada satu titik waktu tertentu. Pengukuran ini menggambarkan situasi penyakit pada saat tertentu, misalnya pada tanggal

tertentu atau pada saat survei dilakukan. Rumus prevalensi titik dapat dinyatakan sebagai berikut:

$$\text{Prevalensi titik} = \frac{\text{Jumlah penderita pada satu titik waktu tertentu}}{\text{Jumlah populasi pada waktu tersebut}} \times k$$

Prevalensi titik memberikan gambaran cepat mengenai distribusi penyakit dalam masyarakat dan sering digunakan untuk mengidentifikasi kebutuhan pelayanan kesehatan pada saat tertentu.

Sementara itu, prevalensi periode adalah proporsi individu dalam populasi yang mengalami suatu penyakit selama rentang waktu tertentu, misalnya selama satu bulan, enam bulan, atau satu tahun. Dalam pengukuran ini, kasus yang dihitung mencakup seluruh individu yang mengalami penyakit selama periode tersebut, baik yang sudah sakit pada awal periode maupun yang mengalami penyakit selama periode pengamatan. Prevalensi periode dapat dirumuskan sebagai berikut:

$$\text{Prevalensi periode} = \frac{\text{Jumlah kasus penyakit dalam satuan waktu tertentu}}{\text{Jumlah populasi selama periode tersebut}} \times k$$

Prevalensi periode memberikan gambaran yang lebih luas mengenai beban penyakit dalam populasi sehingga sering digunakan untuk menilai dampak

penyakit kronis atau penyakit yang memiliki durasi cukup lama dalam populasi (Carneiro & Howard, 2011; Somerville, 2012).

Perubahan Angka Prevalensi Penyakit (APP) dalam suatu populasi dapat dipengaruhi oleh berbagai faktor yang menyebabkan peningkatan maupun penurunan jumlah kasus yang ada.

Faktor yang dapat meningkatkan APP antara lain masuknya penderita ke dalam suatu populasi, keluarnya individu yang sehat dari populasi, masuknya individu yang memiliki potensi mengalami penyakit, peningkatan lama masa sakit, bertambahnya jumlah kasus baru yang terjadi dalam populasi.

Sebaliknya, penurunan APP dapat terjadi apabila individu yang sehat masuk ke dalam populasi, penderita keluar dari populasi, meningkatnya angka kesembuhan, meningkatnya angka kematian, serta masa sakit yang lebih singkat (Nasry N, 2022).

ANGKA MORTALITAS DAN CASE FATALITY RATE

Angka mortalitas merupakan salah satu indikator yang digunakan untuk menggambarkan tingkat kematian dalam suatu populasi selama periode waktu tertentu.

Ukuran mortalitas tidak hanya memberikan gambaran mengenai tingkat keparahan suatu penyakit, tetapi juga mencerminkan kondisi kesehatan masyarakat secara keseluruhan, termasuk kualitas pelayanan kesehatan, kondisi sosial ekonomi, serta efektivitas program kesehatan yang telah diterapkan. Oleh karena itu, analisis mortalitas sering digunakan untuk menilai dampak penyakit terhadap populasi dan menentukan prioritas intervensi kesehatan masyarakat (Friis & Sellers, 2021).

Angka mortalitas biasanya dinyatakan dalam bentuk *rate* atau laju, yaitu perbandingan antara jumlah kematian yang terjadi dalam suatu populasi selama periode tertentu dengan jumlah populasi yang berisiko pada periode yang sama. Dalam epidemiologi, terdapat berbagai jenis angka mortalitas yang digunakan untuk menggambarkan tingkat kematian berdasarkan karakteristik tertentu, seperti angka kematian kasar (CDR), angka kematian bayi (IMR), angka kematian anak (CMR), angka kematian ibu (MMR) dan *case fatality rate* (CFR).

Angka kematian kasar (CDR) merupakan indikator demografi dan kesehatan masyarakat yang digunakan untuk menggambarkan tingkat kematian secara umum dalam suatu populasi selama periode waktu tertentu,

biasanya dalam satu tahun serta sebagai dasar dalam evaluasi pembangunan kesehatan suatu negara atau daerah.

$$\text{CDR} = \frac{\text{jumlah seluruh kematian dalam satu tahun tertentu}}{\text{Jumlah penduduk pada pertengahan tahun tersebut}} \times k$$

Angka kematian bayi (IMR) adalah indikator kesehatan masyarakat yang menunjukkan jumlah kematian bayi berusia kurang dari satu tahun dalam suatu populasi pada periode waktu tertentu. IMR merupakan salah satu indikator paling sensitif dalam menilai derajat kesehatan masyarakat karena sangat dipengaruhi oleh berbagai faktor seperti status gizi ibu dan bayi, kualitas pelayanan kesehatan ibu dan anak, kondisi sosial ekonomi, serta lingkungan tempat tinggal

$$\text{IMR} = \frac{\text{Jumlah kematian bayi (< 1 tahun) dalam satu tahun}}{\text{Jumlah kelahiran hidup tahun tersebut}} \times k$$

Angka kematian anak (CMR) menggambarkan tingkat kematian pada anak dalam kelompok usia tertentu dalam suatu populasi selama periode waktu tertentu. Indikator ini sangat penting karena mencerminkan kondisi kesehatan anak secara keseluruhan, termasuk kualitas

gizi, cakupan imunisasi, akses terhadap pelayanan kesehatan dasar, serta kondisi lingkungan seperti sanitasi dan air bersih.

$$\text{CMR} = \frac{\text{Jumlah kematian anak usia } 1- < 5 \text{ tahun dalam satu tahun}}{\text{Jumlah populasi anak usia } 1- < 5 \text{ tahun pada pertengahan tahun}} \times k$$

Maternal Mortality Ratio merupakan indikator yang digunakan untuk menggambarkan risiko kematian yang dialami perempuan selama masa kehamilan, persalinan, atau dalam periode tertentu setelah persalinan yang berkaitan dengan proses reproduksi. Indikator ini menilai kualitas pelayanan kesehatan maternal serta efektivitas sistem kesehatan dalam memberikan perawatan selama kehamilan dan persalinan.

$$\text{MMR} = \frac{\text{Jumlah kematian ibu karena proses reproduksi dalam satu tahun}}{\text{Jumlah kelahiran hidup pada tahun tersebut}} \times k$$

CFR merupakan ukuran yang digunakan untuk menggambarkan proporsi individu yang meninggal akibat suatu penyakit di antara seluruh individu yang didiagnosis menderita penyakit tersebut dalam periode waktu tertentu. Nilai CFR yang tinggi menunjukkan bahwa penyakit tersebut memiliki tingkat keparahan yang tinggi

atau memiliki peluang kematian yang besar bagi penderita. Sebaliknya, nilai CFR yang rendah menunjukkan bahwa sebagian besar penderita penyakit tersebut memiliki peluang untuk bertahan hidup atau sembuh (Rothman et al., 2014).

Selama terjadinya kejadian luar biasa (*outbreak*) atau pandemi, CFR menjadi indikator penting untuk memantau perkembangan situasi epidemiologis dan menilai efektivitas penanganan kasus oleh sistem pelayanan Kesehatan.

$$CFR = \frac{\text{Jumlah kematian akibat penyakit tertentu dalam periode tertentu}}{\text{Jumlah seluruh kasus penyakit tersebut dalam periode yang sama}} \times k$$

INTERPRETASI UKURAN FREKUENSI DALAM KEBIJAKAN KESEHATAN

Dalam epidemiologi kesehatan masyarakat, ukuran frekuensi penyakit seperti insidensi, prevalensi, dan mortalitas tidak hanya berfungsi sebagai indikator statistik semata, tetapi juga memiliki peran strategis dalam proses pengambilan keputusan dan perumusan kebijakan kesehatan. Data mengenai frekuensi penyakit memberikan gambaran objektif mengenai besarnya

masalah kesehatan dalam suatu populasi, sehingga dapat digunakan sebagai dasar dalam menentukan prioritas program kesehatan, alokasi sumber daya, serta perencanaan intervensi kesehatan yang tepat sasaran (Friis & Sellers, 2021).

Interpretasi ukuran frekuensi penyakit dalam kebijakan kesehatan harus dilakukan secara hati-hati dengan mempertimbangkan konteks epidemiologis, karakteristik populasi, serta faktor sosial dan lingkungan yang memengaruhi distribusi penyakit. Misalnya, angka insidensi yang tinggi pada suatu penyakit menunjukkan bahwa penularan atau kejadian kasus baru masih berlangsung secara aktif dalam populasi. Kondisi ini mengindikasikan perlunya penguatan upaya pencegahan primer, seperti promosi kesehatan, peningkatan cakupan imunisasi, maupun pengendalian faktor risiko.

Sementara itu, prevalensi memberikan informasi mengenai besarnya beban penyakit yang sedang ditanggung oleh suatu populasi. Prevalensi yang tinggi, terutama pada penyakit kronis, menunjukkan bahwa terdapat banyak individu yang hidup dengan kondisi penyakit tersebut dan membutuhkan pelayanan kesehatan dalam jangka panjang. Informasi ini sangat penting dalam perencanaan sistem pelayanan kesehatan, termasuk

penyediaan fasilitas kesehatan, tenaga medis, serta ketersediaan obat dan layanan rehabilitasi (WHO, 2023)

Di sisi lain, angka mortalitas memberikan gambaran mengenai dampak fatal suatu penyakit terhadap populasi. Mortalitas yang tinggi pada penyakit tertentu dapat menunjukkan keterlambatan diagnosis, keterbatasan akses pelayanan kesehatan, atau kurang efektifnya sistem penanganan penyakit. Oleh karena itu, analisis mortalitas sering digunakan untuk mengevaluasi kinerja sistem kesehatan serta mengidentifikasi penyakit yang memerlukan perhatian khusus dalam kebijakan kesehatan Masyarakat (Rothman et al., 2014).

Interpretasi ukuran frekuensi penyakit juga sangat penting dalam proses pemantauan dan evaluasi program kesehatan. Perubahan nilai insidensi, prevalensi, atau mortalitas dari waktu ke waktu dapat digunakan sebagai indikator keberhasilan atau kegagalan suatu program kesehatan. Misalnya, penurunan angka insidensi suatu penyakit setelah pelaksanaan program imunisasi dapat menunjukkan bahwa intervensi tersebut efektif dalam mengurangi penularan penyakit

Dengan demikian, interpretasi yang tepat terhadap ukuran frekuensi penyakit merupakan langkah penting dalam menjembatani antara data epidemiologi dan

kebijakan kesehatan. Informasi yang dihasilkan dari analisis frekuensi penyakit dapat membantu pembuat kebijakan dalam memahami situasi kesehatan masyarakat secara lebih komprehensif, menentukan prioritas intervensi, serta merancang kebijakan kesehatan yang lebih efektif dan berkelanjutan dalam upaya meningkatkan derajat kesehatan masyarakat.

SIMPULAN

Ukuran frekuensi penyakit merupakan konsep fundamental dalam epidemiologi yang digunakan untuk menganalisis distribusi dan besaran masalah kesehatan dalam suatu populasi. Melalui pengukuran yang sistematis terhadap kejadian penyakit, para peneliti dan praktisi kesehatan masyarakat dapat memberikan gambaran yang lebih komprehensif tentang perubahan dan perkembangan penyakit, tingkat risiko populasi, serta dampak kesehatan yang ditimbulkan. Dalam konteks ini, indikator seperti insidensi, prevalensi, dan mortalitas menjadi alat analisis yang penting untuk menggambarkan kondisi kesehatan masyarakat secara kuantitatif dan terukur.

Pemahaman yang komprehensif mengenai berbagai ukuran frekuensi penyakit memiliki peran strategis dalam

penelitian epidemiologi, perencanaan, pelaksanaan, dan evaluasi program kesehatan masyarakat. Data yang dihasilkan dari analisis ukuran frekuensi penyakit dapat menjadi dasar dalam penentuan prioritas masalah kesehatan, pengalokasian sumber daya kesehatan, serta perumusan kebijakan Kesehatan.

DAFTAR PUSTAKA

- Carneiro, I., & Howard, N. (2011). Understanding Public Health: Introduction to Epidemiology. In *Clinical and Translational Science*.
- Friis, R. H., & Sellers, T. A. (2021). *Epidemiology for Public Health Practice SIXTH EDITION*.
www.jblearning.com.
- Nasry N, N. (2022). Epidemiologi Dasar Disiplin Ilmu dalam Kesehatan Masyarakat. In *Unhas Prees* (Vol. 3, Number 3).
- Rothman, K. J., Greenland, S., & Associate, T. L. L. (2014). Modern Epidemiology, 3rd Edition. *The Hastings Center Report, 44 Suppl 2*.
- Somerville, M. (2012). Public Health and Epidemiology at a Glance. In *John Wiley & Sons* (Vol. 51, Number 2).
- WHO. (2023). World health statistics 2023: monitoring health for the sdgs, sustainable development goals. In *The Milbank Memorial Fund quarterly* (Vol. 27, Number 2)

BIODATA PENULIS



Dian Rezki Wijaya, SKM., M.Kes, lahir di Soppeng pada 8 Oktober 1994. Penulis menempuh pendidikan tinggi pada Program Studi Kesehatan Masyarakat di Universitas Hasanuddin, dengan menyelesaikan jenjang Sarjana (S1) dan Magister (S2) pada bidang yang sama. Latar belakang pendidikan tersebut membentuk minat akademik penulis dalam kajian kesehatan masyarakat, khususnya bidang epidemiologi. Sejak tahun 2019, penulis berkiprah sebagai dosen pada Fakultas Kedokteran dan Ilmu Kesehatan Universitas Islam Negeri Alauddin Makassar. Dalam menjalankan tugas akademik, penulis aktif melaksanakan kegiatan tridharma perguruan tinggi yang meliputi pendidikan, penelitian, serta pengabdian kepada masyarakat. Berbagai kegiatan penelitian yang dilakukan berfokus pada isu-isu kesehatan masyarakat dan telah dipublikasikan dalam jurnal ilmiah terakreditasi nasional maupun jurnal internasional.



DESAIN STUDI KOHORT: PROSPEKTIF DAN RETROSPEKTIF

RIMAWATI AULIA INSANI SADARANG

E-mail: rimawati.aulia@uin-alauddin.ac.id

PENDAHULUAN

Banyak fenomena, baik dalam konteks klinis, perilaku, sosial, maupun kebijakan, tidak terjadi secara instan, melainkan berkembang melalui proses yang bertahap dan sering kali berlangsung dalam rentang waktu yang panjang. Perubahan kondisi kesehatan dan pembentukan faktor risiko terjadi melalui proses yang kompleks yang terakumulasi dari waktu ke waktu. Ketika tujuan suatu penelitian adalah memahami bagaimana suatu kondisi atau paparan awal memengaruhi kemungkinan terjadinya outcome di masa mendatang, maka diperlukan pendekatan yang mampu merekam dinamika tersebut secara longitudinal.

Dalam konteks epidemiologi analitik, penggunaan desain studi kohort menjadi sangat relevan karena menyediakan kerangka metodologis yang memastikan bahwa paparan diidentifikasi sebelum outcome terjadi. Struktur ini memungkinkan peneliti menegakkan urutan temporal secara jelas, yang merupakan prasyarat penting dalam evaluasi hubungan sebab-akibat. Melalui pengamatan berkelanjutan terhadap kelompok yang berbeda status paparannya, studi kohort memungkinkan estimasi insidens secara langsung serta perhitungan ukuran asosiasi seperti relative risk atau hazard ratio.

Pendekatan ini memberikan kedalaman analitik yang tidak dapat dicapai oleh desain potong lintang atau cross sectional, yang hanya menggambarkan kondisi pada satu titik waktu tanpa mampu memastikan arah hubungan antara paparan dan outcome.

PRINSIP DASAR STUDI KOHORT

1. Prinsip temporalitas

Studi kohort mengamati sekelompok individu dalam periode waktu tertentu untuk menilai hubungan paparan (*exposure*) dan kejadian (*outcome*). Pengamatan dimulai sejak *baseline* dan dilanjutkan selama masa tindak lanjut sehingga hubungan temporal dapat diidentifikasi secara jelas.

2. Prinsip populasi berisiko

Paparan harus terjadi sebelum munculnya outcome. Penegakan urutan waktu ini memungkinkan interpretasi hubungan sebab-akibat menjadi lebih kuat dalam analisis epidemiologi.

3. Prinsip populasi berisiko (*population at risk*)

Semua individu dalam kohort harus bebas dari outcome pada saat baseline. Dengan demikian, setiap kasus yang

muncul selama tidak lanjut dapat dikategorikan sebagai kasus insidens baru.

4. Prinsip perbandingan kelompok

Analisis dilakukan dengan membandingkan kelompok terpapar dan tidak terpapar untuk menilai perbedaan risiko atau laju kejadian outcome. Perbandingan ini menghasilkan ukuran asosiasi seperti relative risk, incidence rate ratio, dan hazard ratio.

5. Prinsip tindak lanjut (*follow up*)

Keberhasilan studi kohort bergantung pada sistem tindak lanjut yang baik. Kehilangan peserta selama observasi dapat menimbulkan bias seleksi apabila terjadi secara tidak seimbang antara kelompok paparan.

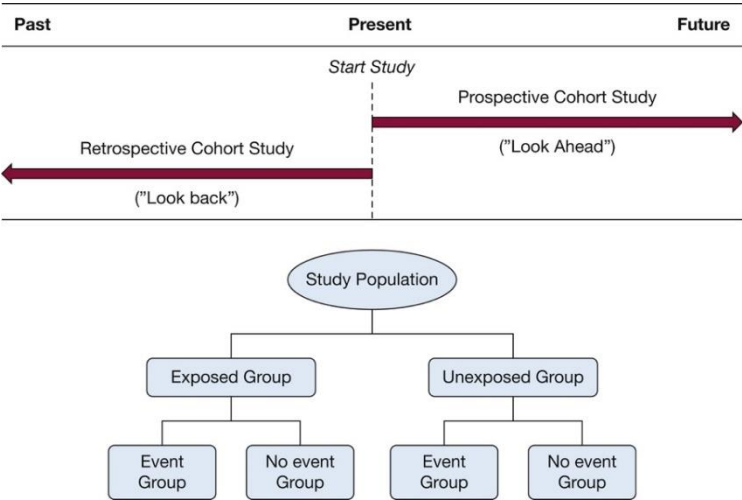
6. Prinsip pengukuran insidens dan ukuran efek

Studi kohort memungkinkan perhitungan langsung incidence proportion dan incidence rate, serta estimasi ukuran asosiasi untuk menggambarkan kekuatan hubungan antara paparan dan outcome.

KOHORT PROSPEKTIF DAN RETROSPEKTIF

Dalam praktik epidemiologi analitik, desain kohort diklasifikasikan terutama berdasarkan orientasi temporal pengumpulan datanya, yakni arah hubungan antara

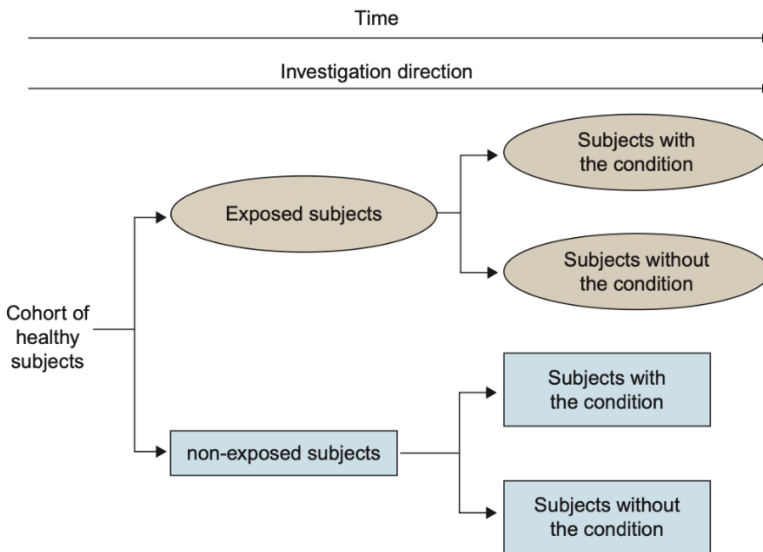
waktu identifikasi paparan dan terjadinya outcome. Pada pendekatan prospektif, peneliti menetapkan status paparan terlebih dahulu, kemudian mengikuti individu hingga outcome muncul di masa mendatang. Sebaliknya, pada pendekatan retrospektif, kohort disusun dengan memanfaatkan data historis sehingga paparan dan outcome telah terjadi ketika penelitian dirancang. Perbedaan orientasi waktu ini tidak semata-mata bersifat teknis, tetapi berimplikasi langsung terhadap strategi pengukuran variabel, potensi bias informasi, efisiensi pelaksanaan penelitian, serta kekuatan inferensi yang dapat ditarik dari hasil analisis.



Gambar 3.1. Skema orientasi waktu pada desain kohort prospektif dan retrospektif
 Sumber: Wang, X., & Kattan, M. W. (2020)

1. Kohort Prospektif

Studi kohort prospektif dimulai dengan identifikasi populasi pada saat *baseline* sebelum kejadian outcome terjadi, kemudian subjek diikuti secara sistematis selama periode waktu yang telah ditentukan untuk mengamati kemunculan outcome yang berkaitan dengan paparan yang diukur di awal. Dalam pendekatan ini, peneliti mengumpulkan data paparan dan variabel lain secara prospektif, sehingga secara metodologis desain ini kuat dalam menegakkan urutan temporal antara paparan dan outcome.



Sumber: Suchmacher, M., & Geller, M. (2012)
Gambar 3.2. Skema desain kohort prospektif

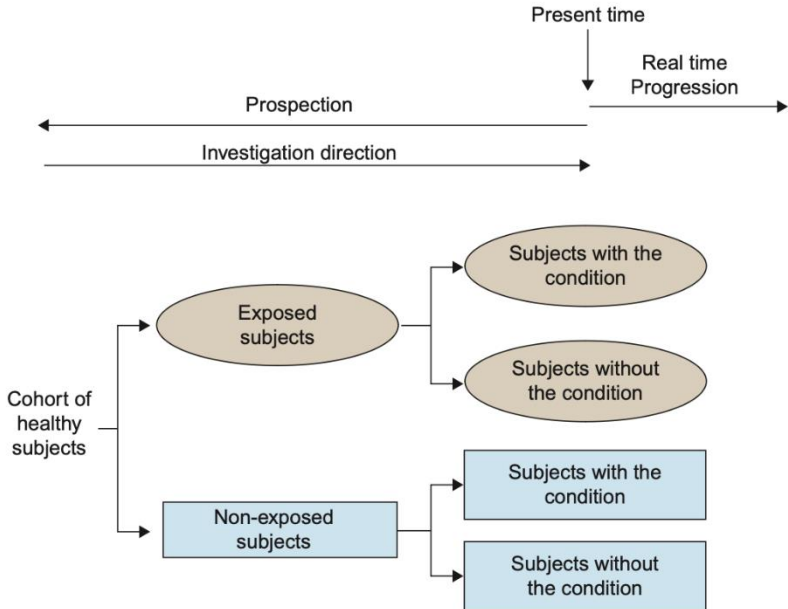
Kohort prospektif memungkinkan peneliti mempelajari progresi alami penyakit dan mengukur insidens outcome secara lebih akurat karena paparan dicatat sebelum outcome terjadi. Selain itu, desain ini dapat mengevaluasi berbagai paparan dan outcome dalam satu kerangka longitudinal yang sama, serta menghasilkan informasi yang merefleksikan dinamika kondisi dunia nyata, meskipun tetap bersifat observasional dan tidak dapat membuktikan kausalitas secara mutlak (Wang & Kattan, 2020; Zouhair, 2023).

Meskipun memiliki kekuatan dalam menegaskan urutan temporal dan pengukuran insidens secara langsung, desain ini sering menghadapi tantangan operasional yang signifikan. Salah satu isu utama adalah kebutuhan waktu tindak lanjut yang panjang untuk memungkinkan outcome berkembang secara alami. Sebagai ilustrasi klasik, apabila paparan seperti kebiasaan merokok diidentifikasi pada remaja, sementara penyakit seperti kanker paru memiliki periode laten rata-rata 20 tahun, maka penelitian harus berlangsung selama beberapa dekade untuk memperoleh outcome yang relevan (Celentano & Szklo, 2019). Konsekuensinya, studi kohort prospektif dapat memerlukan durasi sangat panjang, menghadapi

keterbatasan pendanaan, risiko kehilangan tindak lanjut (*loss to follow up*), serta tantangan keberlanjutan tim peneliti. Oleh karena itu, desain ini menuntut perencanaan jangka panjang yang matang serta sistem pemantauan yang kuat.

2. Kohort Retrospektif

Pendekatan kohort retrospektif atau *historical cohort study*, secara struktur tetap membandingkan kelompok terpapar dan tidak terpapar, tetapi memanfaatkan data paparan dari catatan masa lalu sebagai alternatif untuk mengatasi kendala waktu yang panjang pada kohort rospektif. Secara operasional, penelitian dimulai dengan mengidentifikasi populasi pada periode waktu tertentu di masa lalu, kemudian mengklasifikasikan subjek berdasarkan status paparan yang tercatat dalam dokumen atau basis data yang tersedia. Selanjutnya, peneliti menelusuri kejadian outcome yang telah terjadi hingga titik waktu tertentu. Meskipun disebut “retrospektif”, pendekatan ini tetap mempertahankan logika kohort, yakni bergerak dari paparan menuju outcome, hanya saja seluruh proses tersebut direkonstruksi melalui data yang sudah ada



Sumber: Suchmacher, M., & Geller, M. (2012)
 Gambar 3.3. Skema Desain Kohort Retrospektif

Walaupun lebih efisien dan relatif lebih murah, pendekatan retrospektif memiliki keterbatasan dalam pengendalian kualitas pengukuran paparan, kovariat, dan faktor perancu karena data tidak dikumpulkan secara khusus untuk tujuan penelitian. Risiko bias informasi dan misklasifikasi paparan lebih tinggi apabila pencatatan historis tidak lengkap atau tidak seragam. Oleh karena itu, validitas internal penelitian sangat bergantung pada kualitas sistem pencatatan

yang digunakan sebagai sumber data (Celentano & Szklo, 2019; Suchmacher & Geller, 2012).

PENGUKURAN PAPANAN DAN OUTCOME

Pengukuran paparan dan outcome merupakan komponen metodologis yang sangat penting karena menentukan validitas hubungan antara faktor risiko dan kejadian penyakit. Pengukuran paparan dilakukan pada awal penelitian (*baseline*) dan dapat diperbaharui selama masa tindak lanjut apabila paparan bersifat dinamis. Semua individu dalam kohort harus bebas dari outcome pada saat *baseline* sehingga setiap kejadian yang muncul selama periode observasi dapat dihitung sebagai kasus insidens baru.

Dalam studi kohort, analisis hubungan antara paparan dan outcome biasanya disajikan menggunakan tabel kontingensi 2 x 2. Tabel ini menjadi dasar untuk menghitung berbagai ukuran asosiasi seperti relative risk (RR), risk difference (RD), dan hazard ratio (HR).

Tabel 3.1 Tabel Kontingensi (2x2) pada Studi Kohort

Status Paparan	Outcome (+)	Outcome (-)	Total
Terpapar (E+)	a	b	a + b
Tidak terpapar (E-)	c	d	c + d
Total	a + b	c + d	N

Sumber: Roberts et al. (2019)

1. Relative Risk (RR)

Ukuran ini menunjukkan perbandingan risiko terjadinya outcome pada kelompok terpapar ($a/a+b$) dibandingkan kelompok tidak terpapar ($c/c+d$). Interpretasi nilai dari RR ini terbagi menjadi tiga.

Jika nilai $RR = 1$ maka tidak ada hubungan antara paparan dan outcome. Jika nilai $RR > 1$, maka paparan meningkatkan risiko outcome. Jika $RR < 1$, paparan bersifat protektif terhadap outcome

2. Risk Difference (RD)

Ukuran ini sering digunakan untuk menghitung attributable risk dalam epidemiologi. Ukuran ini juga menunjukkan selisih risiko antara kelompok terpapar ($a/a+b$) dan tidak terpapar ($c/c+d$). Interpretasi nilai RD

berdasarkan nilai 0. Jika nilai $RD > 0$ maka paparan meningkatkan risiko penyakit, nilai $RD = 0$ maka tidak ada perbedaan risiko, dan nilai $RD < 0$ maka paparan bersifat protektif.

3. Hazard Ratio (HR)

Dalam studi kohort yang menggunakan analisis survival (*time to event*), ukuran asosiasi yang sering digunakan adalah hazard ratio (HR). Ukuran ini merupakan perbandingan laju kejadian outcome antara kelompok terpapar dan tidak terpapar pada suatu waktu tertentu selama periode pengamatan.

Berbeda dengan relative risk yang menggambarkan risiko kumulatif, HR merepresentatifkan risiko instan terjadinya kejadian serta mempertimbangkan waktu hingga outcome muncul. Analisis ini juga memperhitungkan *censoring* terutama *right censoring*, yaitu ketika peserta belum mengalami kejadian hingga akhir pengamatan atau hilang selama *follow up*.

Estimasi HR umumnya dilakukan menggunakan model Cox proportional hazards. Nilai $HR=1$ menunjukkan tidak ada perbedaan laju kejadian, $HR>1$ menunjukkan risiko lebih tinggi pada kelompok terpapar, dan $HR<1$ menunjukkan risiko lebih rendah

dibandingkan kelompok tidak terpapar / kelompok referensi.

KELEBIHAN DAN KETERBATASAN STUDI KOHORT

Studi kohort, seperti semua desain penelitian pada umumnya, memiliki keunggulan yang spesifik sekaligus keterbatasan yang perlu dipahami dalam konteks perancangan, pelaksanaan, dan interpretasi hasil penelitian.

Kelebihan studi kohort

1. Menegakkan urutan temporal

Paparan diidentifikasi sebelum terjadinya outcome sehingga hubungan sebab akibat lebih jelas dibandingkan desain potong lintang (Song & Chung, 2010; Wang & Kattan, 2020).

2. Dapat menghitung insidens dan ukuran risiko secara langsung

Karena peserta bebas dari outcome pada awal penelitian, studi kohort memungkinkan perhitungan cumulative incidence, incidence rate, serta ukuran asosiasi seperti Relative Risk (RR), Risk Difference (RD) (Andrade, 2022).

3. Dapat mengamati banyak outcome dari satu paparan

Satu populasi kohort dapat digunakan untuk menilai berbagai dampak kesehatan dari suatu faktor risiko tertentu.

4. Efektif untuk meneliti paparan jarang (*rare exposure*)
Pemilihan kelompok berdasarkan status paparan memungkinkan peneliti mempelajari efek paparan yang tidak umum di populasi.

Keterbatasan studi kohort

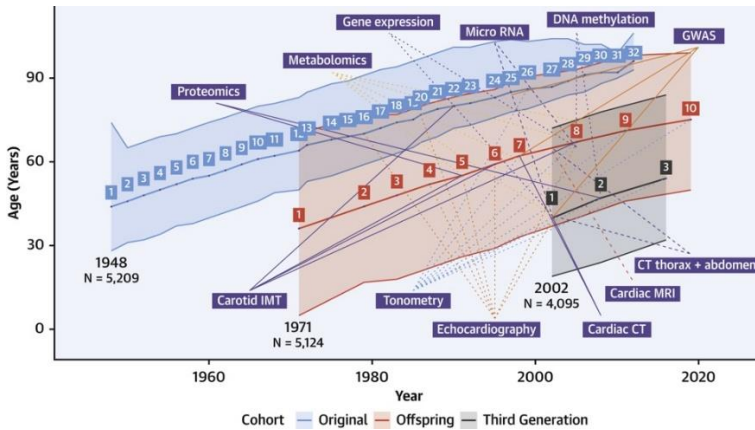
1. Membutuhkan biaya dan sumber daya besar
Studi kohort, terutama yang prospektif, memerlukan sistem pelacakan partisipan, tenaga penelitian, dan infrastruktur yang memadai (Wang & Kattan, 2020).
2. Memerlukan periode tindak lanjut yang panjang
Outcome yang berkembang lambat, seperti penyakit kronis, menuntut *follow up* jangka panjang sehingga penelitian menjadi lebih lama.
3. Risiko kehilangan tindak lanjut (*loss to follow up*)
Kehilangan peserta selama observasi dapat menimbulkan bias seleksi apabila terjadi secara tidak proporsional antar kelompok terpapar dan tidak terpapar (Song & Chung, 2010).
4. Potensi kekacauan variabel (*confounding*)
Sebagai desain observasional, perbedaan karakteristik awal antar kelompok dapat memengaruhi hasil sehingga

diperlukan pengendalian melalui desain penelitian atau analisis statistik multivariat (Andrade, 2022; Song & Chung, 2010).

CONTOH APLIKASI DALAM PENELITIAN KESEHATAN

1. Kohort Prospektif

Salah satu contoh paling terkenal dari studi kohort prospektif dalam bidang kesehatan adalah Framingham Heart Study, yang dimulai pada tahun 1948 di Amerika Serikat. Penelitian ini merekrut 5.209 orang dewasa yang pada awal penelitian belum menderita penyakit kardiovaskular, kemudian mengikuti mereka secara berkala untuk mencatat kejadian baru seperti penyakit jantung dan stroke. Dalam studi ini, berbagai faktor risiko seperti tekanan darah, kolesterol, kebiasaan merokok, serta status diabetes diukur terlebih dahulu pada saat *baseline* sebelum timbulnya penyakit. Pendekatan tersebut memungkinkan peneliti mengamati secara langsung hubungan temporal antara paparan dan outcome, yang merupakan karakteristik utama desain kohort prospektif.



Sumber: Andersson et al. (2021)

Gambar 3.3. Rentang Usia Tiga Generasi Kohort Framingham Menurut Periode dan Siklus Pemeriksaan

Seiring perkembangannya, penelitian ini diperluas menjadi tiga generasi kohort, yaitu *original cohort*, *offspring cohort*, dan *third-generation cohort*). Struktur multidimensi tersebut memungkinkan analisis yang lebih mendalam mengenai perubahan faktor risiko penyakit kardiovaskular antar generasi, serta interaksi antara faktor genetik dan lingkungan. Melalui tindak lanjut jangka panjang dan pemeriksaan kesehatan yang dilakukan secara berulang, studi ini berhasil mengidentifikasi berbagai determinan utama penyakit jantung dan menjadi dasar pengembangan model prediksi risiko kardiovaskular yang digunakan secara

luas dalam praktik klinis dan kesehatan masyarakat (Andersson et al., 2021).

2. Kohort Retrospektif

Contoh penerapan desain kohort retrospektif dapat dilihat pada penelitian yang meneliti hubungan antara viral load dan Human Herpesvirus 8 (HHV-8) dan prognosis kaposi sarcoma pada orang yang hidup dengan HIV/AIDS (ODHA) di Jepang. Penelitian ini menggunakan pendekatan kohort retrospektif dengan menelusuri rekam medis pasien yang pernah didiagnosis dengan kaposi sarcoma di Tokyo Medical University selama periode 2000 hingga 2023. Kohort awal diidentifikasi berdasarkan diagnosis kaposi sarcoma dan viral load HHV-8 dalam catatan laboratorium. Selanjutnya, pasien diklasifikasikan ke dalam kelompok paparan berdasarkan tingkat viral load HHV-8 yang tercatat dalam sampel darah, misalnya kelompok dengan viral load tinggi (≥ 615 copies/ 10^6 sel darah putih) dan kelompok dengan viral load lebih rendah (Ishikawa et al., 2026).

Outcome utama yang dianalisis adalah mortalitas semua penyebab selama periode tindak lanjut yang ditelusuri melalui rekam medis rumah sakit. Melalui pendekatan ini, hubungan temporal antara paparan

(tingkat viral load HHV-8) dan outcome (kelangsungan hidup atau kematian) tetap dapat diidentifikasi meskipun seluruh data diperoleh dari catatan historis. Hasil penelitian menunjukkan bahwa viral load HHV-8 yang tinggi berhubungan dengan prognosis yang lebih buruk pada pasien dengan kaposi sarcoma, terutama pada pasien yang mengalami *kaposi sarcoma inflammatory cytokine syndrome* (KICS). Pada kelompok dengan kondisi tersebut, seluruh pasien meninggal selama periode tindak lanjut. Temuan ini menunjukkan bahwa desain kohort retrospektif tetap dapat membandingkan kelompok terpapar dan tidak terpapar berdasarkan data historis untuk menilai risiko terjadinya outcome kesehatan.

SIMPULAN

Studi kohort adalah desain observasional longitudinal yang menilai hubungan paparan dan outcome dengan mengikuti individu yang bebas dari penyakit pada *baseline* sehingga urutan temporal dapat ditegakkan dan insidens serta ukuran asosiasi seperti relative risk, risk difference, dan hazard ratio dapat dihitung. Berdasarkan orientasi waktunya, kohort dibedakan menjadi prospektif

yang mengikuti subjek ke masa depan dan retrospektif yang menggunakan data historis. Meskipun kuat untuk analisis risiko, desain ini memerlukan sumber daya besar, waktu tindak lanjut panjang, serta berpotensi mengalami *loss to follow up* dan *confounding*.

DAFTAR PUSTAKA

- Andersson, C., Naylor, M., Tsao, C. W., Levy, D., & Vasan, R. S. (2021). Framingham Heart Study: JACC Focus Seminar, 1/8. *Journal of the American College of Cardiology*, 77(21), 2680–2692. <https://doi.org/10.1016/J.JACC.2021.01.059>
- Andrade, C. (2022). Research Design: Cohort Studies. *Indian Journal of Psychological Medicine*, 44(2), 191. <https://doi.org/10.1177/025371762111073764>
- Celentano, D. D., & Szklo, M. (2019). *Gordis Epidemiology, Sixth Edition*.
- Ishikawa, K., Muramatsu, T., Kaneko, S., Harada, Y., Miyashita, R., Kamikubo, Y., Yamaguchi, T., Ichiki, A., Chikasawa, Y., Bingo, M., Sekiya, R., Yotsumoto, M., Hagiwara, T., Amano, K., & Kinai, E. (2026). A Retrospective Cohort Study on HHV-8 Viral Load and Prognosis in HIV-Associated Kaposi Sarcoma Among People Living with HIV in Japan. *Viruses* 2026, Vol. 18, 18(2). <https://doi.org/10.3390/V18020161>
- Song, J. W., & Chung, K. C. (2010). Observational Studies: Cohort and Case-Control Studies. *Plastic and Reconstructive Surgery*, 126(6), 2242. <https://doi.org/10.1097/PRS.0B013E3181F44ABC>
- Suchmacher, M., & Geller, M. (2012). Study Type Determination. In *Practical Biostatistics* (pp. 3–15). Elsevier. <https://doi.org/10.1016/B978-0-12-415794-1.00001-X>
- Wang, X., & Kattan, M. W. (2020). Cohort Studies: Design, Analysis, and Reporting. *Chest*, 158(1), S72–S78. <https://doi.org/10.1016/j.chest.2020.03.014>

Zouhair, J. (2023). Longitudinal Prospective Cohort Studies: A Research Method Primer. *Undergraduate Research in Natural and Clinical Sciences and Technology Journal*, 7(1), 1–5.
<https://doi.org/10.26685/URNCST.399->

BIODATA PENULIS



Rimawati Aulia Insani Sadarang, M.P.H. lahir di Ujung Pandang pada September 1993. Ia merupakan dosen pada Fakultas Kedokteran dan Ilmu Kesehatan, UIN Alauddin Makassar. Aktivitas akademiknya berfokus kajian kesehatan masyarakat, khususnya epidemiologi penyakit menular dan analisis data kesehatan. Ia aktif meneliti dan mempublikasikan karya ilmiah di jurnal nasional maupun internasional serta berkontribusi dalam penulisan buku ajar dan buku referensi di bidang kesehatan masyarakat.

4



DESAIN STUDI CASE-CONTROL: PRINSIP DAN APLIKASI

VITASARI INDRIANI

E-mail: vasari.indriani@unsoed.ac.id

PENDAHULUAN

Studi *case-control* merupakan salah satu desain penelitian observasional analitik yang penting dalam epidemiologi untuk mengidentifikasi hubungan antara faktor paparan (*exposure*) dan kejadian penyakit (*outcome*). Desain ini dilakukan dengan membandingkan individu yang telah mengalami penyakit tertentu (kasus) dengan individu yang tidak mengalami penyakit tersebut (kontrol), kemudian menelusuri riwayat paparan terhadap faktor risiko yang diduga berperan dalam terjadinya penyakit. Melalui perbandingan distribusi paparan pada kedua kelompok tersebut, peneliti dapat menilai apakah suatu paparan berhubungan dengan peningkatan atau penurunan risiko penyakit (Szklo & Nieto, 2019; Lash *et al.*, 2021).

Berbeda dengan studi kohort, studi *case-control* dimulai dari outcome yang telah terjadi, sehingga sering dipandang sebagai pendekatan retrospektif (Pottegård, 2022; Mann, 2020). Desain ini sangat efisien untuk mempelajari penyakit langka atau penyakit dengan periode laten panjang, karena tidak memerlukan pengamatan jangka panjang maupun sampel yang sangat besar seperti pada studi kohort (Pearce *et al.*, 2020). Selain itu, studi *case-control* memungkinkan peneliti

mengevaluasi beberapa faktor risiko sekaligus dalam satu penelitian, sehingga bermanfaat untuk penyakit multifaktorial.

Meskipun demikian, desain ini tetap memiliki keterbatasan, terutama potensi *selection bias*, *recall bias*, dan kesulitan dalam memastikan hubungan temporal antara paparan dan penyakit. Oleh karena itu, perancangan penelitian harus dilakukan secara cermat agar hasilnya valid dan bermakna (Setia, 2016; Vandenbroucke *et al.*, 2021).

KONSEP DASAR STUDI CASE-CONTROL

Studi *case-control* merupakan desain penelitian observasional analitik yang dimulai dengan identifikasi individu berdasarkan status penyakitnya. Kelompok kasus terdiri dari individu yang mengalami penyakit tertentu, sedangkan kelompok kontrol terdiri dari individu yang tidak mengalami penyakit tersebut tetapi berasal dari populasi sumber yang sama (Lash *et al.*, 2021; Szklo & Nieto, 2019).

Setelah kelompok kasus dan kontrol ditentukan, peneliti menelusuri riwayat paparan terhadap faktor risiko tertentu yang diduga berhubungan dengan kejadian

penyakit. Dengan membandingkan frekuensi paparan pada kedua kelompok tersebut, peneliti dapat menilai apakah paparan tertentu berhubungan dengan peningkatan atau penurunan risiko penyakit (Pearce *et al.*, 2020; Lash *et al.*, 2021).

Berbeda dengan studi kohort yang dimulai dari paparan dan diikuti hingga munculnya penyakit, studi *case-control* bergerak dari penyakit menuju penelusuran paparan sebelumnya (*retrospective investigation*) (Szklo & Nieto, 2019).

Pendekatan ini sangat bermanfaat untuk mempelajari penyakit dengan insidensi rendah seperti beberapa jenis kanker atau penyakit genetik. Dalam kondisi tersebut, studi kohort akan memerlukan jumlah sampel yang besar dan waktu penelitian yang lama, sedangkan studi *case-control* dapat memberikan hasil yang lebih efisien (Pearce *et al.*, 2020; Szklo & Nieto, 2019).

PEMILIHAN KASUS DAN KONTROL

Pemilihan kasus dan kontrol merupakan tahap krusial dalam studi *case-control* karena sangat menentukan validitas internal penelitian. Kasus adalah individu yang mengalami penyakit atau kondisi kesehatan

yang menjadi fokus penelitian, sedangkan kontrol adalah individu yang tidak mengalami penyakit tersebut tetapi berasal dari populasi sumber yang sama dengan kasus (Vandenbroucke & Pearce, 2021). Oleh karena itu, kelompok kontrol harus merepresentasikan populasi yang berisiko menjadi kasus apabila mereka mengalami penyakit yang diteliti.

Kasus perlu ditetapkan berdasarkan kriteria diagnostik yang jelas, konsisten, dan dapat dipertanggungjawabkan, baik melalui standar klinis, hasil laboratorium, maupun kriteria epidemiologis tertentu. Peneliti juga perlu menetapkan kriteria inklusi dan eksklusi agar kelompok kasus tetap homogen dan sesuai tujuan penelitian (Setia, 2016). Sementara itu, kontrol harus dipilih secara hati-hati agar distribusi paparannya mencerminkan populasi sumber. Jika tidak, hasil penelitian dapat mengalami *selection bias*. Dengan demikian, proses seleksi kasus dan kontrol harus dilakukan secara sistematis, transparan, dan konsisten agar interpretasi *Odds Ratio* tetap valid.

Tabel 4.1. Perbandingan Jenis Kontrol dalam Studi Case-Control

Jenis kontrol	Definisi	Kelebihan	Kelemahan	Potensi bias
Hospital control	Individu yang dirawat di rumah sakit tetapi tidak memiliki penyakit yang sedang diteliti	Mudah direkrut, data klinis biasanya lengkap, biaya lebih rendah	Penyakit pada kontrol dapat berhubungan dengan paparan yang diteliti, sehingga kurang merepresentasikan populasi sumber	<i>Selection bias</i> akibat paparan pada kontrol tidak mencerminkan populasi sumber
Community control	Individu dari komunitas yang sama dengan kasus tetapi tidak mengalami penyakit yang diteliti	Lebih representatif terhadap lingkungan sosial dan karakteristik populasi kasus	Rekrutmen lebih sulit, waktu dan sumber daya lebih besar	Bias nonrespons dan kemungkinan perbedaan karakteristik dengan kasus
Population-based control	Individu yang dipilih secara acak dari populasi umum melalui metode sampling	Representatif terhadap populasi sumber, validitas eksternal lebih baik	Biaya lebih besar, proses sampling lebih kompleks, partisipasi responden bisa rendah	Bias seleksi akibat nonrespons atau kerangka sampling yang tidak memadai

Sumber: Szklo & Nieto (2019); Setia (2016); Vandenbroucke & Pearce (2021).

Pemilihan jenis kontrol harus disesuaikan dengan tujuan penelitian dan sumber data yang tersedia. *Hospital control* lebih praktis, *community control* lebih dekat dengan konteks sosial kasus, sedangkan *population-based control* umumnya memberikan representasi terbaik terhadap populasi sumber. Namun, apa pun jenis kontrol yang dipilih, prinsip utamanya tetap sama: kontrol harus berasal dari populasi yang sama dengan kasus dan memiliki peluang yang sebanding untuk mengalami paparan yang diteliti (Setia, 2016; Vandenbroucke & Pearce, 2021).

MATCHING DALAM STUDI CASE-CONTROL

Matching merupakan teknik dalam studi *case-control* yang digunakan untuk mengendalikan variabel perancu (*confounding variables*), yaitu variabel yang berhubungan dengan paparan dan penyakit sehingga dapat memengaruhi interpretasi hubungan keduanya (Pearce *et al.*, 2020). Melalui *matching*, peneliti berupaya membuat karakteristik tertentu pada kelompok kasus dan kontrol menjadi sebanding agar validitas internal penelitian meningkat (Szklo & Nieto, 2019).

Matching umumnya diterapkan ketika terdapat variabel perancu yang kuat, seperti usia, jenis kelamin, atau status sosial ekonomi. Misalnya, dalam penelitian tentang hubungan merokok dengan kanker paru, seorang kasus berupa laki-laki berusia 55 tahun dapat dipasangkan dengan kontrol berupa laki-laki berusia 55 tahun. Dalam praktiknya, terdapat dua bentuk *matching* yang paling sering digunakan, yaitu *individual matching* dan *frequency matching*. *Individual matching* memasangkan setiap kasus dengan satu atau lebih kontrol yang memiliki karakteristik sama, sedangkan *frequency matching* menyamakan distribusi karakteristik tertentu antara kelompok kasus dan kontrol (Lash *et al.*, 2021).

Meskipun bermanfaat, *matching* tidak selalu dapat menghilangkan *confounding* sepenuhnya. Oleh karena itu, teknik ini sering dikombinasikan dengan analisis statistik seperti stratifikasi atau regresi logistik. Penggunaannya juga harus hati-hati karena dapat menimbulkan *overmatching*, yaitu pencocokan yang justru mengaburkan hubungan antara paparan dan penyakit (Pearce *et al.*, 2020).

ANALISIS DAN INTERPRETASI ODDS RATIO

Odds Ratio (OR) merupakan ukuran asosiasi yang paling umum digunakan dalam studi *case-control* untuk menilai hubungan antara paparan dan kejadian penyakit. OR menggambarkan perbandingan peluang paparan pada kelompok kasus dibandingkan kelompok kontrol, sehingga dapat menunjukkan kekuatan hubungan antara faktor risiko dan *outcome* kesehatan (Pearce *et al.*, 2020). Karena studi *case-control* tidak menghitung insidensi penyakit secara langsung, OR digunakan sebagai alternatif *Relative Risk*. Perhitungan OR biasanya menggunakan tabel kontingensi 2×2 dan juga banyak diterapkan dalam regresi logistik untuk menghasilkan *adjusted odds ratio* (Lash *et al.*, 2021; Szklo & Nieto, 2019).

Tabel 4.2. Tabel Kontingensi 2×2 dalam Studi Case-Control

Status paparan	Kasus	Kontrol
Terpapar	a	b
Tidak terpapar	c	d

Sumber: Lash *et al.* (2021) dan Szklo & Nieto (2019).

Berdasarkan tabel tersebut, rumus *Odds Ratio* adalah:

$$OR = (a \times d) / (b \times c)$$

Untuk mempermudah pemahaman, berikut contoh perhitungan dengan angka nyata. Misalnya, suatu penelitian ingin menilai hubungan antara kebiasaan merokok dan kanker paru. Hasil pengumpulan data menunjukkan bahwa dari 80 responden, terdapat 30 kasus yang merokok, 10 kasus yang tidak merokok, 15 kontrol yang merokok, dan 25 kontrol yang tidak merokok.

Tabel 3. Contoh Data Perhitungan Odds Ratio

Status paparan	Kasus	Kontrol
Merokok	30	15
Tidak merokok	10	25

Sumber: Pearce *et al.*, (2020); Lash *et al.*, (2021)

Berdasarkan tabel di atas, maka:

- a = 30
- b = 15
- c = 10
- d = 25

Langkah perhitungan OR dilakukan sebagai berikut:

1. Kalikan nilai a dengan d
 $30 \times 25 = 750$
2. Kalikan nilai b dengan c
 $15 \times 10 = 150$
3. Bagi hasil langkah pertama dengan langkah kedua
 $OR = 750 / 150 = 5$

Nilai OR sebesar 5 menunjukkan bahwa individu yang merokok memiliki peluang lima kali lebih besar untuk mengalami kanker paru dibandingkan individu yang tidak merokok. Hasil ini mengindikasikan bahwa merokok merupakan faktor risiko yang kuat terhadap kanker paru dalam penelitian tersebut (Lash *et al.*, 2021). Secara umum, interpretasi OR dapat dibagi menjadi tiga. Jika $OR > 1$, maka paparan berhubungan dengan peningkatan peluang terjadinya penyakit dan dapat dianggap sebagai faktor risiko. Jika $OR = 1$, maka tidak terdapat hubungan antara paparan dan penyakit. Jika $OR < 1$, maka paparan berhubungan dengan penurunan peluang terjadinya penyakit sehingga dapat dianggap memiliki efek protektif (Pearce *et al.*, 2020).

Sebagai ilustrasi, $OR = 0,5$ pada hubungan aktivitas fisik dan penyakit jantung menunjukkan bahwa individu yang rutin beraktivitas fisik memiliki peluang setengah kali lebih kecil mengalami penyakit jantung dibandingkan yang tidak aktif (Szklo & Nieto, 2019). Namun, interpretasi OR harus selalu mempertimbangkan *confidence interval* (CI). Jika CI 95% tidak mencakup angka 1, maka hubungan dianggap bermakna secara statistik. Selain itu, pada penyakit yang jarang terjadi, OR dapat mendekati *Relative Risk*, sehingga interpretasinya menjadi hampir

serupa. Oleh karena itu, pemahaman OR sangat penting dalam analisis studi *case-control*.

KEKUATAN DAN KELEMAHAN DESAIN CASE-CONTROL

Studi *case-control* merupakan desain penelitian observasional analitik yang banyak digunakan dalam epidemiologi karena efisien untuk menilai hubungan antara paparan dan penyakit, terutama pada penyakit yang jarang terjadi atau memiliki periode laten panjang (Lash *et al.*, 2021). Dibandingkan studi kohort, desain ini lebih hemat waktu dan biaya karena penelitian dimulai dari *outcome* yang telah terjadi, sehingga peneliti tidak perlu menunggu munculnya kasus baru dalam jangka waktu lama (Szklo & Nieto, 2019). Selain itu, studi *case-control* memungkinkan peneliti mengevaluasi beberapa faktor risiko sekaligus dalam satu penelitian, sehingga sangat berguna untuk penyakit multifaktorial seperti kanker, penyakit kardiovaskular, dan penyakit metabolik (Pearce *et al.*, 2020).

Meskipun demikian, desain ini juga memiliki keterbatasan. Studi *case-control* rentan terhadap *selection bias*, *recall bias*, dan *information bias* yang dapat

memengaruhi validitas hasil penelitian (Lash *et al.*, 2021). Penelitian ini juga sulit memastikan hubungan temporal antara paparan dan kejadian penyakit karena data paparan umumnya diperoleh secara retrospektif (Szklo & Nieto, 2019). Selain itu, desain ini tidak dapat menghitung insidensi penyakit secara langsung. Oleh karena itu, studi *case-control* paling tepat digunakan untuk penyakit langka dengan sumber daya terbatas, asalkan dirancang dan diinterpretasikan secara cermat (Pearce *et al.*, 2020).

SIMPULAN

Studi *case-control* merupakan desain penelitian observasional analitik yang penting dalam epidemiologi untuk mengidentifikasi hubungan antara faktor paparan dan kejadian penyakit. Desain ini dilakukan dengan membandingkan individu yang telah mengalami penyakit (kasus) dengan individu yang tidak mengalami penyakit (kontrol), kemudian menelusuri riwayat paparan pada kedua kelompok tersebut. Pendekatan ini sangat bermanfaat untuk mempelajari penyakit yang jarang terjadi atau memiliki periode laten yang panjang.

Keunggulan utama studi *case-control* terletak pada efisiensi waktu dan biaya dibandingkan dengan studi kohort, serta kemampuannya untuk menilai beberapa

faktor risiko dalam satu penelitian. Validitas hasil penelitian sangat ditentukan oleh ketepatan pemilihan kasus dan kontrol, penggunaan *matching* untuk mengendalikan variabel perancu, serta analisis data menggunakan *Odds Ratio* sebagai ukuran asosiasi antara paparan dan penyakit.

Meskipun demikian, desain ini tetap memiliki keterbatasan, terutama karena rentan terhadap *selection bias*, *recall bias*, dan kesulitan dalam memastikan urutan waktu antara paparan dan kejadian penyakit. Oleh sebab itu, perencanaan penelitian harus dilakukan secara cermat.

DAFTAR PUSTAKA

- Andrade, C. (2022). Research design: Case-control studies. *Indian Journal of Psychological Medicine*, 44(3), 307–308. <https://doi.org/10.1177/02537176221090104>
- Lash, T. L., VanderWeele, T. J., Haneuse, S., & Rothman, K. J. (2021). *Modern epidemiology* (4th ed.). Wolters Kluwer.
- Mann, C. J. (2020). Observational research methods: Research design II—Cohort, cross sectional, and case-control studies. *Emergency Medicine Journal*, 37(1), 38–42. <https://doi.org/10.1136/emered-2019-208843>
- Pearce, N., Vandenbroucke, J. P., & Lawlor, D. A. (2020). Causal inference in environmental epidemiology. *International Journal of Epidemiology*, 49(1), 378–387. <https://doi.org/10.1093/ije/dyz259>
- Pottegård, A. (2022). Core concepts in pharmacoepidemiology: Fundamentals of the cohort and case-control study designs. *Pharmacoepidemiology and Drug Safety*, 31(8), 817–826. <https://doi.org/10.1002/pds.5482>
- Setia, M. S. (2016). Methodology series module 2: Case-control studies. *Indian Journal of Dermatology*, 61(2), 146–151. <https://doi.org/10.4103/0019-5154.177773>
- Song, J. W., & Chung, K. C. (2020). Observational studies: Cohort and case-control studies. *Plastic and*

Reconstructive Surgery, 126(6), 2234–2242.
<https://doi.org/10.1097/PRS.0b013e3181f44abc>

Szklo, M., & Nieto, F. J. (2019). *Epidemiology: Beyond the basics* (4th ed.). Jones & Bartlett Learning

Vandenbroucke, J. P., Pearce, N., & Case-control study group. (2021). Case-control studies: Basic concepts. *International Journal of Epidemiology*, 50(3), 971–975. <https://doi.org/10.1093/ije/dyab031>.

BIODATA PENULIS



Dr. dr. Vitasari Indriani, MM, SpPK, M.Si.Med lahir di Jember, pada 3 April 1980. Ia tercatat sebagai lulusan S1 dan S2 Fakultas Kedokteran Universitas Diponegoro Semarang. S3 dari Universitas Gadjah Mada. Saat ini bekerja sebagai dosen Departemen Patologi Klinik Fakultas Kedokteran Universitas Jenderal Soedirman Purwokerto



DESAIN STUDI CROSS-SECTIONAL DAN EKOLOGI

DIAN IHWANA ANSYAR

E-mail: dian.ihwana@uin-alauddin.ac.id

PENDAHULUAN

Pendekatan penelitian dalam epidemiologi modern tidak hanya bertujuan menggambarkan distribusi penyakit dalam populasi, tetapi juga menganalisis hubungan antara faktor paparan dengan terjadinya penyakit. Pendekatan ini dikenal sebagai epidemiologi analitik, yaitu suatu pendekatan penelitian yang digunakan untuk menguji hipotesis mengenai hubungan kausal antara paparan tertentu dan hasil kesehatan pada suatu populasi (Kirollos, 2024).

Dalam epidemiologi analitik digunakan berbagai desain penelitian, baik observasional maupun eksperimental, untuk mengkaji asosiasi antara faktor risiko dan kejadian penyakit. Dalam penelitian kesehatan masyarakat, salah satu desain yang paling umum digunakan adalah studi potong lintang (*cross-sectional study*). Desain ini banyak digunakan karena relatif sederhana, efisien, dan mampu memberikan gambaran kondisi kesehatan suatu populasi pada waktu tertentu. Pada penelitian ini, variabel paparan dan outcome diukur secara bersamaan pada satu titik waktu sehingga dapat menggambarkan situasi kesehatan populasi saat pengamatan dilakukan (Bonita Beaglehole, 2006).

Selain studi potong lintang, desain penelitian lain yang cukup sering digunakan dalam epidemiologi adalah **studi ekologi (ecological study)**. Pada desain ini, unit analisis yang digunakan bukan individu, melainkan kelompok atau populasi, seperti wilayah geografis, komunitas, maupun negara. Studi ekologi menggunakan data agregat untuk menganalisis asosiasi antara paparan dan outcome pada tingkat populasi, sehingga sering digunakan untuk menganalisis pengaruh faktor lingkungan, sosial, maupun kebijakan kesehatan terhadap kejadian penyakit pada masyarakat.

Baik studi cross-sectional maupun studi ekologi memiliki peranan penting dalam penelitian epidemiologi, khususnya dalam menggambarkan distribusi penyakit, mengidentifikasi kemungkinan faktor risiko, serta menyediakan informasi awal bagi perumusan kebijakan kesehatan. Kedua desain tersebut juga banyak digunakan dalam kegiatan surveilans kesehatan masyarakat dan penelitian berbasis populasi untuk memahami pola permasalahan kesehatan yang terjadi di masyarakat.

KONSEP STUDI POTONG LINTANG (*CROSS-SECTIONAL*)

Studi potong lintang (cross-sectional study) merupakan desain penelitian epidemiologi yang digunakan untuk mengukur suatu kaitan antara paparan dan kejadian penyakit dalam suatu populasi pada waktu tertentu. Pada desain ini, pengukuran variabel paparan dan variabel penyakit dilakukan secara simultan dalam satu waktu pengamatan sehingga dapat menggambarkan distribusi penyakit serta faktor risiko pada populasi yang diteliti (Gordis, 2014)

Dalam penelitian kesehatan masyarakat, studi potong lintang sering digunakan karena relatif mudah dilaksanakan, waktu yang dibutuhkan relative singkat, serta memerlukan biaya yang lebih minim dibandingkan dengan desain studi longitudinal seperti studi kohort. Melalui pendekatan ini, peneliti dapat memperoleh gambaran mengenai kondisi kesehatan masyarakat pada periode tertentu sekaligus mengidentifikasi adanya probabilitas hubungan antara faktor risiko terhadap kejadian penyakit (Friis & Sellers, 2021).

Secara konseptual, studi potong lintang berfokus pada pengukuran prevalensi penyakit dalam populasi. Prevalensi merupakan ukuran yang menunjukkan

proporsi suatu individu dalam populasi yang memiliki penyakit maupun kondisi kesehatan tertentu pada waktu pengamatan tertentu (WHO, 2020).

Oleh karena itu, desain penelitian ini sangat berguna untuk menggambarkan besarnya masalah kesehatan dalam masyarakat sekaligus mendukung perencanaan program kesehatan.

Unit analisis yang digunakan dalam studi potong lintang umumnya adalah individu dalam populasi tertentu. Peneliti mengumpulkan berbagai informasi mengenai responden, seperti karakteristik demografi, perilaku kesehatan, paparan lingkungan, serta status penyakit. Selanjutnya, data yang terkumpul dianalisis untuk menilai hubungan antara paparan dan kondisi kesehatan yang diamati.

Salah satu keunggulan utama dari studi potong lintang adalah kemampuannya memberikan gambaran distribusi penyakit secara relatif cepat dalam populasi. Informasi ini sangat penting bagi pengambil kebijakan untuk menentukan prioritas intervensi kesehatan serta merancang program pencegahan penyakit. Selain itu, desain penelitian ini juga sering digunakan dalam survei kesehatan berskala besar baik pada tingkat nasional maupun internasional.

Walaupun demikian, studi potong lintang memiliki keterbatasan dalam menentukan hubungan kausal antara paparan dan penyakit. Hal ini dikarenakan pengukuran kedua variabel tersebut dilakukan secara simultan sehingga tidak dapat dipastikan apakah paparan mendahului terjadinya penyakit atau sebaliknya. Oleh sebab itu, hasil penelitian dari studi potong lintang sering digunakan sebagai dasar dalam pembentukan hipotesis yang selanjutnya dapat diuji dengan desain penelitian yang lebih kuat, seperti studi kohort atau studi kasus-kontrol (Rothman et al., 2021).

Dalam praktik kesehatan masyarakat, studi potong lintang sering digunakan dalam berbagai survei kesehatan untuk menilai kondisi kesehatan populasi. Salah satu contohnya adalah survei nasional yang dilakukan oleh Kementerian Kesehatan Republik Indonesia melalui Riset Kesehatan Dasar atau Riskesdas yang menggunakan pendekatan potong lintang untuk mengumpulkan data mengenai berbagai indikator kesehatan masyarakat, seperti prevalensi penyakit tidak menular, status gizi, perilaku merokok, serta akses terhadap pelayanan kesehatan.

Data yang dihasilkan dari survei tersebut menjadi dasar penting dalam perencanaan kebijakan kesehatan

nasional. Dengan demikian, studi potong lintang memiliki peranan penting dalam epidemiologi, khususnya dalam menggambarkan distribusi penyakit dalam populasi serta mengidentifikasi faktor-faktor yang berkaitan dengan kondisi kesehatan masyarakat. Walaupun memiliki keterbatasan dalam menentukan hubungan kausal, desain penelitian ini tetap menjadi salah satu metode yang paling banyak digunakan dalam penelitian kesehatan masyarakat.

Keunggulan dan keterbatasan studi *cross-sectional* dapat terlihat jelas pada tabel berikut :

Tabel 5.1. Keunggulan dan Keterbatasan Studi Cross-Sectional

Aspek	Keunggulan	Keterbatasan
Waktu penelitian	Dilakukan dalam waktu relatif singkat	Tidak dapat menggambarkan perubahan penyakit dari waktu ke waktu
Biaya	Relatif lebih murah dibandingkan studi kohort	Analisis lanjutan sering memerlukan penelitian tambahan
Pengukuran epidemiologi	Dapat mengukur prevalensi penyakit	Tidak dapat menghitung insidensi penyakit
Analisis hubungan	Dapat menilai asosiasi antara paparan dan penyakit	Tidak dapat memastikan hubungan sebab-akibat

Aspek	Keunggulan	Keterbatasan
Kegunaan kesehatan masyarakat	Berguna dalam survei kesehatan populasi	Rentan terhadap bias seperti survival bias

Sumber : diadaptasi dari Gordis (2014) dan Rothman et al. (2021).

STUDI EKOLOGI DAN UNIT ANALISIS POPULASI

Studi ekologi merupakan salah satu desain penelitian epidemiologi yang menggunakan kelompok atau populasi sebagai unit analisis. Berbeda dengan studi potong lintang yang menggunakan individu sebagai unit analisis, studi ekologi menganalisis data pada tingkat kelompok atau populasi tertentu (Morgenstern H., 2008).

Data yang digunakan dalam penelitian ini biasanya berupa data agregat yang menggambarkan karakteristik suatu populasi, misalnya wilayah geografis, kelompok sosial, atau periode waktu tertentu. Pendekatan ini memungkinkan peneliti untuk menilai hubungan antara paparan dan kejadian penyakit pada tingkat populasi (Gordis, 2014).

Sumber informasi yang digunakan dalam studi ekologi umumnya berasal dari data sekunder seperti data sensus penduduk, registrasi penyakit, laporan surveilans

kesehatan, maupun statistik kesehatan nasional. Sebagai contoh, peneliti dapat membandingkan rata-rata konsumsi rokok di berbagai negara dengan angka kejadian kanker paru pada populasi negara tersebut. Analisis tersebut dilakukan pada tingkat kelompok, bukan pada individu yang merokok.

Keunggulan utama dari studi ekologi adalah efisiensi dalam pemanfaatan data. Penelitian ini relatif lebih cepat dan ekonomis karena sering memanfaatkan data yang sudah tersedia. Selain itu, studi ekologi juga bermanfaat untuk mengidentifikasi pola distribusi penyakit dalam skala yang luas serta menghasilkan hipotesis awal mengenai hubungan antara faktor risiko dan penyakit (Friis & Sellers, 2021).

Namun demikian, penggunaan unit analisis populasi juga memiliki keterbatasan. Data agregat tidak mampu menggambarkan variasi individu dalam kelompok tersebut. Oleh karena itu, hasil studi ekologi umumnya digunakan sebagai dasar eksplorasi atau pembentukan hipotesis yang selanjutnya perlu diuji dengan desain penelitian yang menggunakan unit analisis individu.

PENGUKURAN PREVALENSI DAN ASOSIASI

Salah satu tujuan utama dari studi potong lintang adalah mengukur prevalensi penyakit atau kondisi kesehatan dalam suatu populasi. Prevalensi merupakan proporsi individu dalam populasi yang mengalami penyakit atau kondisi kesehatan tertentu pada waktu pengamatan tertentu (WHO, 2020).

Prevalensi dihitung dengan membandingkan jumlah kasus yang ditemukan pada saat pengamatan dengan jumlah total populasi yang diteliti. Ukuran ini memberikan gambaran mengenai beban penyakit (disease burden) dalam masyarakat. Informasi mengenai prevalensi sangat penting dalam perencanaan program kesehatan karena dapat membantu menentukan prioritas intervensi kesehatan masyarakat.

Sebagai contoh, peneliti dapat mengkaji hubungan antara kebiasaan merokok dan kejadian hipertensi pada populasi dewasa. Apabila Hipertensi memiliki prevalensi lebih tinggi pada kelompok perokok dibandingkan kelompok bukan perokok, maka hal tersebut menunjukkan adanya asosiasi antara perilaku merokok dan hipertensi.

Pada studi potong lintang, pengukuran paparan dan outcome dilakukan secara simultan dalam satu waktu pengamatan. Selain itu, desain ini dapat dimanfaatkan

untuk menganalisis hubungan antara faktor paparan dan kondisi kesehatan yang terjadi pada periode yang sama. Hubungan tersebut umumnya diukur menggunakan indikator asosiasi seperti *prevalence ratio* atau *prevalence odds ratio* (Szklo, 2019; Kenneth J. Rothman, 2021)

Meskipun demikian, karena pengukuran paparan dan penyakit dilakukan secara bersamaan, studi potong lintang memiliki keterbatasan dalam menentukan urutan waktu antara paparan dan penyakit. Oleh sebab itu, desain penelitian ini tidak dapat memastikan hubungan sebab-akibat secara pasti (Gordis, 2014).

Contoh Perhitungan *Prevalence Ratio* : Sebagai ilustrasi, dilakukan survei terhadap 200 orang dewasa untuk menilai hubungan antara kebiasaan merokok dan kejadian hipertensi.

Status Merokok	Hipertensi	Tidak Hipertensi	Total
Perokok	40	60	100
Tidak Perokok	20	80	100
Total	60	140	200

Prevalensi hipertensi pada perokok:

$$40 / 100 = 0,40$$

Prevalensi hipertensi pada non-perokok:

$$20 / 100 = 0,20$$

Prevalence Ratio (PR):

$$PR = 0,40 / 0,20 = 2,0$$

Interpretasi:

Individu yang merokok memiliki prevalensi hipertensi dua kali lebih tinggi dibandingkan individu yang tidak merokok.

RISIKO *ECOLOGICAL FALLACY*

Salah satu keterbatasan penting dalam studi ekologi adalah adanya potensi *ecological fallacy*. *Ecological fallacy* merupakan kesalahan interpretasi yang terjadi ketika hubungan yang ditemukan pada tingkat kelompok dianggap berlaku juga pada tingkat individu.

Kesalahan ini dapat terjadi karena data yang digunakan dalam studi ekologi bersifat agregat. Hubungan yang terlihat pada tingkat populasi belum tentu mencerminkan hubungan yang sama pada individu dalam populasi tersebut (Rothman et al., 2021).

Sebagai contoh, suatu negara mungkin memiliki tingkat konsumsi alkohol yang tinggi dan pada saat yang sama juga memiliki angka kejadian penyakit jantung yang tinggi. Namun demikian, dari informasi tersebut tidak dapat disimpulkan bahwa individu yang mengonsumsi alkohol memiliki risiko penyakit jantung lebih tinggi dibandingkan individu yang tidak mengonsumsi alkohol.

Fenomena *ecological fallacy* menunjukkan pentingnya kehati-hatian dalam menginterpretasikan hasil penelitian ekologi. Hasil studi ekologi sebaiknya digunakan sebagai dasar untuk merancang penelitian lanjutan dengan desain yang lebih kuat, seperti studi kohort atau studi kasus kontrol. Oleh karena itu, hasil penelitian ekologi perlu ditafsirkan secara hati-hati dan sebaiknya diikuti oleh penelitian dengan unit analisis individu (Morgenstern H., 2008; Last John M, 2014).

APLIKASI DALAM SURVEI KESEHATAN MASYARAKAT

Studi potong lintang merupakan salah satu desain penelitian yang paling sering digunakan dalam survei kesehatan masyarakat. Desain ini memungkinkan peneliti mengumpulkan informasi mengenai status kesehatan, perilaku kesehatan, serta berbagai faktor risiko dalam suatu populasi pada waktu tertentu.

Dalam praktik kesehatan masyarakat, survei potong lintang digunakan untuk mengukur berbagai indikator kesehatan, seperti prevalensi penyakit tidak menular, status gizi masyarakat, perilaku merokok, tingkat aktivitas fisik, serta akses terhadap pelayanan kesehatan.

Di Indonesia, pendekatan ini digunakan dalam berbagai survei nasional, salah satunya adalah Riset Kesehatan Dasar (Riskesdas) yang diselenggarakan oleh Kementerian Kesehatan Republik Indonesia (Kemenkes, 2018). Survei tersebut memberikan gambaran mengenai kondisi kesehatan masyarakat Indonesia, termasuk prevalensi penyakit tidak menular, status gizi, serta berbagai faktor risiko kesehatan lainnya.

Data yang dihasilkan dari survei tersebut menjadi dasar bagi pemerintah dalam merancang kebijakan kesehatan nasional serta menentukan prioritas program kesehatan masyarakat. Oleh karena itu, studi potong lintang memiliki peran penting dalam perencanaan, pemantauan, dan evaluasi program kesehatan masyarakat.

SIMPULAN

Studi potong lintang merupakan salah satu desain penelitian dalam epidemiologi yang banyak dimanfaatkan dalam penelitian kesehatan masyarakat. Desain ini memungkinkan peneliti untuk melakukan pengukuran paparan dan kejadian penyakit pada waktu yang sama dalam suatu populasi pada periode tertentu sehingga dapat

menggambarkan prevalensi penyakit serta penyebaran faktor risiko di masyarakat.

Selain studi potong lintang, studi ekologi juga termasuk desain penelitian yang penting dalam epidemiologi karena memungkinkan analisis keterkaitan antara faktor lingkungan dan terjadinya penyakit pada tingkat populasi.. Namun demikian, studi ekologi memiliki keterbatasan karena menggunakan data agregat sehingga berpotensi menimbulkan kesalahan dalam penafsiran yang dikenal sebagai ecological fallacy.

Walaupun memiliki keterbatasan dalam menentukan hubungan sebab-akibat, studi potong lintang dan studi ekologi tetap memiliki peranan penting dalam survei kesehatan masyarakat. Kedua desain penelitian tersebut memberikan informasi yang bermanfaat untuk memahami distribusi penyakit, mengidentifikasi faktor risiko, serta mendukung proses perencanaan dan evaluasi program kesehatan masyarakat.

DAFTAR PUSTAKA

- Bonita Beaglehole, R. R. (2006). *Basic epidemiology 2nd edition*.
- Friis, R. H. , & S. T. A. (2021). *Epidemiology for Public Health Practice (6th ed.)*. Jones & Bartlett Learning.
- Kementerian Kesehatan Republik Indonesia. (2018). *Laporan Nasional Riskesdas 2018*.
- Kenneth J. Rothman, G. S. , & L. T. L. (2021). *Modern Epidemiology (4th ed.)*. Wolters Kluwer.
- Kirollos, M. , & G. S. (2024). *Epidemiology*. In *Handbook for Designing and Conducting Clinical and Translational Research*. Elsevier.
- Last John M. (2014). *A dictionary of epidemiology, 6th edn*. . Oxford University Press.
- Leon Gordis. (2014). *Epidemiology (5th ed.)*. Elsevier.
- Morgenstern H. (2008). *Ecological studies*. In: *Modern Epidemiology. 3rd Edition*.
- Szklo, M. , & N. F. J. (2019). *Epidemiology: Beyond the Basics*.
- World Health Organization. (2020). *Basic Epidemiology (2nd ed.)*. WHO Press

BIODATA PENULIS



Dian Ihwana Ansyar lahir pada 26 Desember 1993. Ia merupakan alumni Universitas Hasanuddin dan saat ini berkarier sebagai dosen di bidang Kesehatan Masyarakat dengan minat utama pada epidemiologi. Selain menjalankan kegiatan pengajaran, ia juga aktif terlibat dalam berbagai kegiatan penelitian serta pengabdian kepada masyarakat dengan beragam topik. Dengan latar belakang akademik dan pengalaman di bidang kesehatan masyarakat, ia berkomitmen untuk mengintegrasikan pelaksanaan Tri Dharma Perguruan Tinggi dalam upaya meningkatkan kualitas kesehatan masyarakat.



**UJI KLINIS DAN STUDI
EKSPERIMENTAL DALAM
EPIDEMIOLOGI**

ROJALI

E-mail: rojali@untirta.ac.id

PENDAHULUAN

Uji klinis merupakan tulang punggung ilmu pengetahuan biomedis modern. Sebagai desain penelitian dengan tingkat bukti tertinggi dalam hierarki epidemiologi, uji klinis terkontrol acak (Randomized Controlled Trial/RCT) memberikan landasan yang paling kokoh untuk menyimpulkan hubungan kausal antara intervensi dan luaran kesehatan (Schulz & Grimes, 2002). Bab ini membahas secara mendalam konsep, desain, pelaksanaan, dan analisis uji klinis serta studi eksperimental dalam konteks epidemiologi analitik.

DEFINISI DAN KARAKTERISTIK UJI KLINIS

Uji klinis (clinical trial) didefinisikan sebagai penelitian prospektif yang memberikan intervensi kepada manusia untuk mengevaluasi efeknya terhadap luaran kesehatan tertentu. Menurut World Health Organization (WHO, 2020), uji klinis merupakan setiap penelitian yang secara prospektif menugaskan peserta manusia kepada satu atau lebih intervensi kesehatan terkait untuk mengevaluasi efeknya terhadap luaran kesehatan.

Ciri utama yang membedakan uji klinis dari desain observasional adalah adanya penugasan aktif (active

assignment) intervensi oleh peneliti, bukan sekadar pengamatan pasif terhadap paparan yang sudah terjadi secara alami. Penugasan inilah yang memungkinkan peneliti mengendalikan berbagai faktor perancu (confounders) yang dapat mengacaukan interpretasi hasil (Friedman et al., 2015).

PRINSIP FUNDAMENTAL: RANDOMISASI

Randomisasi adalah proses penugasan peserta ke kelompok intervensi atau kontrol berdasarkan mekanisme peluang (chance) yang telah ditentukan sebelumnya. Tujuan utama randomisasi adalah menghasilkan komparabilitas antara kelompok dalam hal variabel yang diukur maupun yang tidak diukur, sehingga perbedaan luaran yang terdeteksi dapat diatribusikan kepada intervensi yang diberikan (Schulz & Grimes, 2002).

Terdapat beberapa metode randomisasi yang umum digunakan dalam praktik:

- Randomisasi sederhana (simple randomization): setiap peserta memiliki probabilitas yang sama untuk masuk ke kelompok mana pun, seperti pelemparan koin.

- Randomisasi blok (block randomization): memastikan keseimbangan jumlah peserta antar kelompok pada setiap interval waktu perekrutan.
- Randomisasi stratifikasi (stratified randomization): dilakukan terpisah dalam setiap strata variabel prognostik penting (misalnya usia, jenis kelamin, keparahan penyakit) untuk memastikan distribusi yang seimbang.
- Randomisasi adaptif (adaptive randomization): probabilitas penugasan disesuaikan secara dinamis berdasarkan respons peserta sebelumnya.

PRINSIP PENYAMARAN (BLINDING/MASKING)

Penyamaran adalah prosedur untuk mencegah peserta, pemberi intervensi, dan/atau penilai luaran mengetahui penugasan kelompok. Tujuannya adalah meminimalkan bias yang mungkin timbul akibat harapan atau persepsi yang berbeda antara kelompok (Jadad & Enkin, 2007). Klasifikasi penyamaran meliputi:

- Uji terbuka (open-label trial): baik peserta maupun peneliti mengetahui penugasan kelompok; cocok untuk kondisi di mana penyamaran tidak praktis.
- Penyamaran tunggal (single-blind): hanya peserta yang tidak mengetahui penugasan kelompoknya.

- Penyamaran ganda (double-blind): baik peserta maupun pemberi intervensi tidak mengetahui penugasan kelompok; dianggap standar emas.
- Penyamaran tiga (triple-blind): peserta, pemberi intervensi, dan penilai luaran semuanya tidak mengetahui penugasan kelompok.

KERANGKA ETIKA UJI KLINIS

Setiap uji klinis wajib memenuhi prinsip etika yang ditetapkan dalam Deklarasi Helsinki (WMA, 2013) dan pedoman CIOMS (2016). Prinsip-prinsip utama meliputi: (a) *equipe* atau *clinical equipoise*, yakni ketidakpastian tulus tentang superioritas salah satu intervensi; (b) *informed consent* yang komprehensif dan bebas paksaan; (c) perlindungan kelompok rentan; serta (d) pengawasan independen oleh Data Safety Monitoring Board (DSMB) yang memiliki wewenang menghentikan uji jika ditemukan bukti bahaya signifikan.

Tabel 6.1 Prinsip-Prinsip Desain Uji Klinis dan Implikasi Metodologis

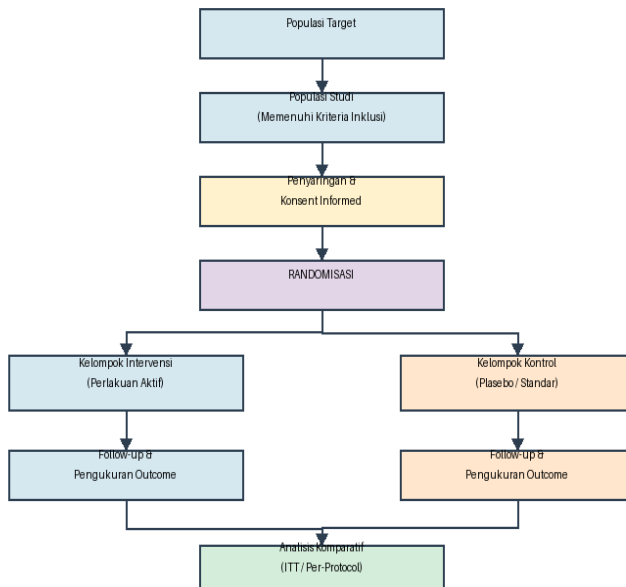
Prinsip	Definisi Operasional	Tujuan Utama	Risiko Jika Diabaikan
Randomisasi	Penugasan peserta secara acak ke kelompok	Menyeimbangkan faktor perancu	Confounding bias
Penyamaran (Blinding)	Menyembunyikan penugasan kelompok	Mencegah performance & detection bias	Placebo effect & ascertainment bias
Kontrol	Kelompok pembanding yang tidak menerima intervensi aktif	Menyediakan baseline komparatif	Kurangnya referensi untuk estimasi efek
Alokasi tersembunyi (Concealment)	Menyembunyikan urutan randomisasi dari perekrut	Mencegah selection bias sebelum randomisasi	Selection bias dalam perekrutan
Analisis Intention-to-Treat (ITT)	Analisis berdasarkan alokasi awal	Mempertahankan keuntungan randomisasi	Overestimasi efek intervensi
Informed Consent	Persetujuan sukarela berbasis informasi lengkap	Melindungi otonomi peserta	Pelanggaran etika & validitas hukum

Sumber: Diadaptasi dari Friedman et al. (2015); Schulz & Grimes (2002).

DESAIN UJI KLINIS: JENIS DAN KEUNGGULAN

Uji Klinis Terkontrol Acak (RCT)

RCT merupakan desain penelitian yang dianggap sebagai standar emas (gold standard) dalam evaluasi intervensi kesehatan. Ciri pembeda RCT yang paling fundamental adalah penugasan random peserta ke dalam dua atau lebih kelompok komparasi. Alur sistematis pelaksanaan RCT klasik ditampilkan pada Gambar 6.1 di bawah ini.



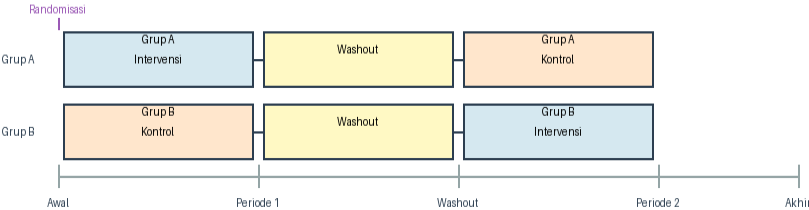
Gambar 6.1 Alur Pelaksanaan Uji Klinis Terkontrol Acak (RCT)

Berdasarkan Gambar 6.1, peserta yang memenuhi kriteria inklusi dan eksklusi direkrut dari populasi target.

Setelah melalui proses penyaringan dan pemberian informed consent, peserta dirandomisasi ke kelompok intervensi atau kontrol. Kedua kelompok kemudian di-follow up selama periode yang telah ditentukan, dan luaran diukur secara terstandar untuk selanjutnya dianalisis secara komparatif (Pocock, 2013).

Desain Crossover

Dalam desain crossover, setiap peserta menerima semua intervensi yang dibandingkan, namun dalam urutan yang berbeda-beda. Setiap peserta berfungsi sebagai kontrol bagi dirinya sendiri (within-subject comparison), sehingga variabilitas antar-individu dapat dieliminasi secara efektif dari estimasi efek. Ilustrasi desain crossover dua-periode ditampilkan pada Gambar 6.2.



Gambar 6.2 Desain Crossover Dua Periode dengan Periode Washout

Persyaratan kritis dalam desain crossover adalah adanya periode washout (pencucian) yang memadai antara dua periode perlakuan untuk memastikan efek intervensi periode pertama tidak terbawa ke periode kedua (carry-over effect). Desain ini tidak sesuai untuk kondisi yang dapat disembuhkan permanen (permanent cure) oleh salah satu intervensi atau kondisi dengan perjalanan penyakit yang tidak stabil (Senn, 2002).

Desain Faktorial

Desain faktorial memungkinkan evaluasi dua atau lebih intervensi secara simultan dalam satu uji klinis. Dalam desain faktorial 2×2 , peserta dirandomisasi ke salah satu dari empat kelompok: intervensi A saja, intervensi B saja, kombinasi A+B, atau kontrol (tanpa keduanya). Keuntungan utama adalah efisiensi sumber daya karena dua pertanyaan penelitian dapat dijawab dengan satu sampel yang sama, serta kemampuan untuk mendeteksi interaksi antara dua intervensi (Montgomery, 2017).

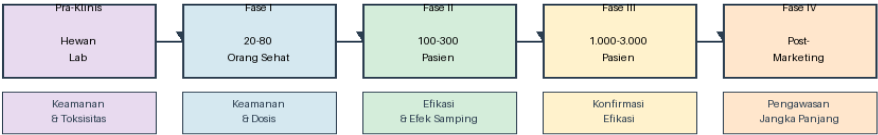
Uji Klinis Klaster (Cluster Randomized Trial)

Pada cluster randomized trial, unit randomisasi bukan individu melainkan kelompok atau klaster seperti sekolah, puskesmas, desa, atau klinik. Desain ini tepat

digunakan ketika intervensi diberikan pada level kelompok (misalnya pelatihan tenaga kesehatan, program promosi kesehatan komunitas) atau ketika kontaminasi antar-kelompok sangat sulit dihindari jika individu yang dirandomisasi. Analisis data dari desain kluster harus memperhitungkan efek desain (design effect) akibat korelasi intra-kluster (intraclass correlation coefficient/ICC) (Hayes & Moulton, 2017).

Fase Uji Klinis

Pengembangan intervensi farmakologis melewati serangkaian fase uji klinis yang terstruktur sebagaimana digambarkan dalam Gambar 6.3 berikut ini.



Gambar 6.3 Fase-Fase Uji Klinis dalam Pengembangan Obat

Tabel 6.2 Perbandingan Jenis-Jenis Desain Uji Klinis

Desain	Unit Randomisasi	Keunggulan Utama	Keterbatasan Utama	Contoh Penerapan
RCT Paralel	Individu	Standar emas; kontrol optimal	Biaya & waktu tinggi	Uji efikasi vaksin
Crossover	Individu (urutan)	Efisien; kontrol variabilitas antar-individu	Risiko carry-over effect	Uji obat kronik stabil
Faktorial	Individu	Evaluasi ≥ 2 intervensi sekaligus	Asumsi tidak ada interaksi	Uji kombinasi terapi
Cluster RCT	Kelompok/klaster	Tepat untuk intervensi berbasis komunitas	Membutuhkan lebih banyak klaster	Program kesehatan sekolah
Stepped Wedge	Klaster (berurutan)	Semua klaster akhirnya menerima intervensi	Analisis kompleks	Implementasi kebijakan
Adaptive Design	Individu	Fleksibel; efisien secara etis	Kompleksitas statistik tinggi	Uji onkologi fase III

Sumber: Diadaptasi dari Pocock (2013); Hayes & Moulton (2017).

METODOLOGI PELAKSANAAN UJI KLINIS

Perencanaan dan Protokol

Protokol uji klinis adalah dokumen komprehensif yang menjabarkan semua aspek pelaksanaan penelitian secara rinci dan harus diselesaikan sebelum pengumpulan data dimulai. Komponen wajib protokol mencakup: pertanyaan penelitian dan hipotesis yang akan diuji; defisiensi populasi target beserta kriteria inklusi dan eksklusi; deskripsi intervensi dan kontrol; definisi luaran primer dan sekunder; perhitungan besar sampel; rencana randomisasi dan penyamaran; prosedur follow-up; serta rencana analisis statistik. Registrasi prospektif protokol di platform seperti ClinicalTrials.gov atau ISRCTN adalah standar internasional yang diterapkan oleh sebagian besar jurnal ilmiah bereputasi (DeAngelis et al., 2004).

Kriteria Kelayakan dan Rekrutmen

Kriteria inklusi menentukan karakteristik yang harus dimiliki peserta agar dapat diikutsertakan, sedangkan kriteria eksklusi menentukan kondisi yang mendiskualifikasi calon peserta meskipun memenuhi kriteria inklusi. Definisi kriteria kelayakan yang terlalu ketat akan meningkatkan homogenitas sampel dan validitas internal, namun sekaligus mengurangi

generalisabilitas (validitas eksternal) hasil penelitian (Rothman et al., 2008).

Strategi rekrutmen yang efektif meliputi kerja sama dengan fasilitas kesehatan, penggunaan registri penyakit, iklan yang disetujui komite etik, serta keterlibatan komunitas (community engagement). Tingkat respons rekrutmen yang rendah dapat mengancam kecukupan kekuatan statistik (statistical power) dan menimbulkan bias seleksi.

Perhitungan Besar Sampel

Besar sampel yang memadai merupakan prasyarat fundamental agar uji klinis memiliki kekuatan statistik yang cukup untuk mendeteksi perbedaan klinis yang bermakna antara kelompok. Formula umum untuk uji klinis dua-kelompok dengan luaran biner menggunakan pendekatan Fleiss et al. (2003):

$$n = [(z_{1-\alpha/2} + z_{1-\beta})^2 \times (p_1(1-p_1) + p_2(1-p_2))] / (p_1 - p_2)^2$$

di mana n adalah jumlah sampel per kelompok; $z_{1-\alpha/2}$ adalah nilai- z untuk tingkat signifikansi α (umumnya 1,96 untuk $\alpha=0,05$); $z_{1-\beta}$ adalah nilai- z untuk kekuatan statistik $1-\beta$ (umumnya 0,84 untuk kekuatan 80%); p_1 dan p_2 adalah proporsi kejadian luaran pada masing-masing kelompok. Faktor tambahan yang harus diperhitungkan adalah

ekspektasi drop-out, efek desain (untuk cluster RCT), dan koreksi multiplisitas jika ada lebih dari satu luaran primer.

Manajemen Data dan Kualitas

Manajemen data yang ketat merupakan tulang punggung integritas hasil uji klinis. Prinsip-prinsip utama meliputi:

- Data entry ganda (double data entry) dengan rekonsiliasi otomatis untuk meminimalkan kesalahan input.
- Penerapan sistem manajemen data elektronik (Electronic Data Capture/EDC) dengan audit trail yang tidak dapat dimanipulasi.
- Validasi data secara berkala dengan pemeriksaan rentang nilai (range checks) dan konsistensi logis (logical consistency checks).
- Pemantauan situs penelitian (site monitoring) oleh Clinical Research Associate (CRA) yang independen.
- Pengarsipan dokumen sesuai regulasi Good Clinical Practice (GCP) untuk memungkinkan rekonstruksi dan audit.

Peran Data Safety Monitoring Board (DSMB)

DSMB atau Independent Data Monitoring Committee (IDMC) adalah komite independen yang memiliki akses terblinding terhadap data interim (tidak tersembunyi) dan bertanggung jawab memberikan rekomendasi kepada sponsor untuk melanjutkan, memodifikasi, atau menghentikan uji klinis. DSMB biasanya menggunakan aturan penghentian (stopping rules) berbasis batas pengeluaran alpha (alpha-spending boundaries) seperti prosedur O'Brien-Fleming atau Pocock untuk mengendalikan tingkat kesalahan tipe I secara keseluruhan dalam analisis interim berulang (DeMets & Lan, 1994).

Tabel 6.3 Jenis-Jenis Bias dalam Uji Klinis dan Strategi Pengendalian

Jenis Bias	Definisi	Fase yang Rentan	Strategi Pengendalian
Selection bias	Perbedaan sistematis antara karakteristik peserta antar kelompok	Rekrutmen & randomisasi	Randomisasi + alokasi tersembunyi
Performance bias	Perbedaan perawatan yang diterima	Pelaksanaan intervensi	Penyamaran ganda (double-blind)

Jenis Bias	Definisi	Fase yang Rentan	Strategi Pengendalian
	kelompok selain intervensi yang diteliti		
Detection bias	Perbedaan dalam pengukuran luaran antar kelompok	Pengukuran outcome	Penyamaran penilai (blinded outcome assessment)
Attrition bias	Kehilangan follow-up yang berbeda antar kelompok	Follow-up	Analisis ITT; strategi retensi aktif
Reporting bias	Pelaporan selektif luaran berdasarkan hasil signifikansi	Analisis & pelaporan	Registrasi protokol prospektif; CONSORT
Contamination bias	Peserta kontrol mendapat akses ke intervensi	Pelaksanaan	Cluster randomization; pemisahan geografis

Sumber: Diadaptasi dari Higgins & Green (2011); Cochrane Handbook.

ANALISIS STATISTIK DAN PELAPORAN UJI KLINIS

Pendekatan Analisis: ITT vs Per-Protocol

Dua pendekatan analisis utama dalam uji klinis adalah analisis intention-to-treat (ITT) dan analisis per-protocol (PP). Dalam analisis ITT, semua peserta yang telah dirandomisasi dianalisis sesuai kelompok yang dialokasikan, terlepas dari kepatuhan atau penghentian dini. ITT mempertahankan keuntungan komparabilitas yang diperoleh dari randomisasi dan memberikan estimasi efektivitas dalam kondisi praktis (effectiveness). Sebaliknya, analisis PP hanya menyertakan peserta yang patuh terhadap protokol, sehingga memberikan estimasi efikasi murni (efficacy) intervensi dalam kondisi ideal, namun rentan terhadap bias seleksi pasca-randomisasi (Lachin, 2000).

Konsensus metodologi saat ini merekomendasikan analisis ITT sebagai analisis primer, dengan analisis PP sebagai analisis sensitivitas untuk mendukung interpretasi. Modifikasi modified ITT (mITT) yang mengeksklusikan peserta yang tidak pernah menerima setidaknya satu dosis intervensi juga lazim digunakan, terutama dalam uji klinis fase III (Gillings & Koch, 1991).

Ukuran Efek dan Estimasi Presisi

Pilihan ukuran efek bergantung pada jenis luaran yang diukur. Untuk luaran dikotomi (biner), ukuran efek yang relevan meliputi risiko relatif (relative risk/RR), odds ratio (OR), perbedaan risiko (risk difference/RD), dan number needed to treat (NNT). Untuk luaran kontinu, perbedaan rata-rata (mean difference/MD) atau perbedaan rata-rata terstandarisasi (standardized mean difference/SMD) digunakan. Untuk luaran berbasis waktu (time-to-event), hazard ratio (HR) dari model regresi Cox merupakan ukuran yang paling tepat (Altman & Bland, 2011).

Semua estimasi efek harus dilengkapi dengan interval kepercayaan 95% (95% CI) yang mencerminkan ketidakpastian estimasi. Nilai p tunggal tanpa CI dianggap tidak memadai untuk pelaporan hasil uji klinis yang komprehensif (Rothman, 2012).

Analisis Subkelompok dan Sensitivitas

Analisis subkelompok mengevaluasi apakah efek intervensi berbeda pada subpopulasi tertentu yang telah ditetapkan sebelumnya (misalnya berdasarkan usia, jenis kelamin, keparahan penyakit). Interpretasi analisis subkelompok harus sangat hati-hati karena: (a) kekuatan

statistik yang jauh lebih rendah; (b) peningkatan risiko kesalahan tipe I akibat komparasi multipel; dan (c) kemungkinan temuan spurious akibat variasi acak. Panduan umum mengharuskan analisis subkelompok ditetapkan apriori dalam protokol, menggunakan uji interaksi formal (interaction test) sebagai landasan, dan menginterpretasikan hasil secara konservatif (Sun et al., 2010).

Analisis sensitivitas bertujuan menguji robustness temuan primer terhadap berbagai asumsi yang dibuat dalam analisis utama, seperti metode imputasi data hilang, definisi alternatif luaran, atau pengecualian situs dengan kualitas data rendah.

Penanganan Data Hilang (Missing Data)

Data hilang merupakan tantangan universal dalam uji klinis. Mekanisme data hilang yang relevan meliputi: Missing Completely at Random (MCAR), di mana kehilangan data tidak berkaitan dengan nilai yang hilang maupun variabel lain; Missing at Random (MAR), di mana kehilangan data berkaitan dengan variabel yang diamati namun tidak dengan nilai yang hilang itu sendiri; dan Missing Not at Random (MNAR), di mana kehilangan data berkaitan dengan nilai yang hilang itu sendiri. Metode

imputasi berganda (multiple imputation) berdasarkan asumsi MAR saat ini merupakan pendekatan yang paling direkomendasikan untuk data hilang yang substansial (Little & Rubin, 2002).

Pelaporan Standar: Pedoman CONSORT

Consolidated Standards of Reporting Trials (CONSORT) menyediakan pedoman pelaporan uji klinis yang telah diadopsi oleh lebih dari 600 jurnal biomedis di seluruh dunia. Dua elemen pelaporan inti dalam pedoman CONSORT adalah: (1) diagram alur CONSORT yang mendokumentasikan jumlah peserta di setiap tahap penelitian dari perekrutan hingga analisis akhir; dan (2) daftar periksa (checklist) 25 item yang mencakup semua aspek metodologi dan hasil yang harus dilaporkan. Kepatuhan terhadap CONSORT secara signifikan meningkatkan transparansi, reproduisibilitas, dan utilisasi temuan uji klinis dalam praktik klinis berbasis bukti (Schulz et al., 2010).

Tabel 6.4 Ukuran Efek dalam Uji Klinis: Definisi, Formula, dan Interpretasi

Ukuran Efek	Jenis Luaran	Formula	Interpretasi	Keunggulan
Risk Ratio (RR)	Dikotomi	$RR = R_1 / R_0$	RR>1: risiko meningkat; <1: menurun; =1: tidak ada efek	Intuitif; mudah dikomunikasikan
Odds Ratio (OR)	Dikotomi	$OR = (a/b) / (c/d)$	Mendekati RR bila luaran jarang (<10%)	Standar dalam regresi logistik
Risk Difference (RD)	Dikotomi	$RD = R_1 - R_0$	Efek absolut; mudah untuk NNT	Relevan secara klinis & kebijakan
NNT	Dikotomi	$NNT = 1 / RD $	Jumlah yang perlu ditangani untuk 1 outcome	Langsung relevan secara klinis
Mean Difference (MD)	Kontinu	$MD = X_1 - X_0$	Perbedaan rata-rata dalam satuan asli	Mudah diinterpretasi bila skala sama
Hazard Ratio (HR)	Time-to-event	Model Cox regresi	HR<1: menurunkan laju kejadian	Memanfaatkan semua data termasuk censored

Sumber: Altman & Bland (2011); Rothman et al. (2008).

PENUTUP

Uji klinis, khususnya RCT, menempati posisi tertinggi dalam hierarki bukti ilmiah epidemiologi karena kemampuannya mengendalikan confounding melalui randomisasi. Bab ini telah menguraikan empat pilar utama topik tersebut.

- Konsep dasar uji klinis: Randomisasi, penyamaran, dan prinsip etika (clinical equipoise dan informed consent) merupakan fondasi yang membedakan uji klinis dari desain observasional dan menjamin validitas internal penelitian.
- Ragam desain uji klinis: RCT paralel, crossover, faktorial, cluster RCT, dan adaptive design masing-masing memiliki keunggulan dan keterbatasan yang menentukan kesesuaiannya dengan pertanyaan penelitian tertentu. Pemilihan desain harus mempertimbangkan sifat intervensi, unit analisis, dan kendala praktis.
- Metodologi pelaksanaan: Perencanaan protokol yang komprehensif, perhitungan besar sampel yang memadai, manajemen data yang ketat, dan pengawasan independen oleh DSMB merupakan komponen operasional yang menentukan kualitas dan integritas uji klinis.

Analisis dan pelaporan: Pendekatan ITT sebagai analisis primer, pemilihan ukuran efek yang sesuai dengan jenis luaran, penanganan data hilang yang tepat, dan pelaporan sesuai pedoman CONSORT adalah standar minimal yang harus dipenuhi untuk memastikan transparansi dan reproduibilitas temuan uji klinis,

DAFTAR PUSTAKA

Altman, D. G., & Bland, J. M. (2011). How to obtain the confidence interval from a P value. *BMJ*, 343, d2090. <https://doi.org/10.1136/bmj.d2090>

Council for International Organizations of Medical Sciences (CIOMS). (2016). *International ethical guidelines for health-related research involving humans* (4th ed.). CIOMS.

DeAngelis, C. D., Drazen, J. M., Frizelle, F. A., Haug, C., Hoey, J., Horton, R., Kotzin, S., Laine, C., Marusic, A., Overbeke, A. J. P. M., Schroeder, T. V., Sox, H. C., & Van Der Weyden, M. B. (2004). Clinical trial registration: A statement from the International Committee of Medical Journal Editors. *JAMA*, 292(11), 1363–1364. <https://doi.org/10.1001/jama.292.11.1363>

DeMets, D. L., & Lan, K. K. G. (1994). Interim analysis: The alpha spending function approach. *Statistics in Medicine*, 13(13–14), 1341–1352. <https://doi.org/10.1002/sim.4780131308>

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). John Wiley & Sons. <https://doi.org/10.1002/0471445428>

Friedman, L. M., Furberg, C. D., DeMets, D. L., Reboussin, D. M., & Granger, C. B. (2015). *Fundamentals of clinical trials* (5th ed.). Springer. <https://doi.org/10.1007/978-3-319-18539-2>

Gillings, D., & Koch, G. (1991). The application of the principle of intention-to-treat to the analysis of clinical trials. *Drug*

Information Journal, 25(3), 411–424.
<https://doi.org/10.1177/009286159102500316>

Hayes, R. J., & Moulton, L. H. (2017). Cluster randomised trials (2nd ed.). CRC Press/Chapman & Hall.
<https://doi.org/10.1201/9781315370286>

Higgins, J. P. T., & Green, S. (Eds.). (2011). Cochrane handbook for systematic reviews of interventions (Version 5.1.0). The Cochrane Collaboration.
<https://training.cochrane.org/handbook>

Jadad, A. R., & Enkin, M. W. (2007). Randomised controlled trials: Questions, answers, and musings (2nd ed.). Blackwell Publishing/BMJ Books.
<https://doi.org/10.1002/9780470691823>

Lachin, J. M. (2000). Statistical considerations in the intent-to-treat principle. *Controlled Clinical Trials*, 21(3), 167–189.
[https://doi.org/10.1016/s0197-2456\(00\)00046-5](https://doi.org/10.1016/s0197-2456(00)00046-5)

Little, R. J. A., & Rubin, D. B. (2002). Statistical analysis with missing data (2nd ed.). John Wiley & Sons.
<https://doi.org/10.1002/9781119013563>

Montgomery, D. C. (2017). Design and analysis of experiments (9th ed.). John Wiley & Sons.

Pocock, S. J. (2013). Clinical trials: A practical approach. John Wiley & Sons. <https://doi.org/10.1002/9781118793916>

Rothman, K. J. (2012). Epidemiology: An introduction (2nd ed.). Oxford University Press.

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). Modern epidemiology (3rd ed.). Lippincott Williams & Wilkins.

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel

group randomised trials. *BMJ*, 340, c332.
<https://doi.org/10.1136/bmj.c332>

Schulz, K. F., & Grimes, D. A. (2002). Allocation concealment in randomised trials: Defending against deciphering. *The Lancet*, 359(9306), 614–618.
[https://doi.org/10.1016/S0140-6736\(02\)07750-4](https://doi.org/10.1016/S0140-6736(02)07750-4)

Senn, S. (2002). *Cross-over trials in clinical research* (2nd ed.). John Wiley & Sons. <https://doi.org/10.1002/0470854596>

Sun, X., Briel, M., Walter, S. D., & Guyatt, G. H. (2010). Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ*, 340, c117.
<https://doi.org/10.1136/bmj.c117>

World Health Organization (WHO). (2020). International clinical trials registry platform (ICTRP). <https://www.who.int/clinical-trials-registry-platform>

World Medical Association (WMA). (2013). Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, 310(20), 2191–2194.
<https://doi.org/10.1001/jama.2013.281053>

BIODATA PENULIS



Dr. Rojali, S.KM., M.Epid. adalah akademisi dan praktisi kesehatan masyarakat dengan keahlian epidemiologi. Lulusan FKM Universitas Indonesia dan UIN Syarif Hidayatullah Jakarta. Pernah bekerja di Kementerian Kesehatan RI, BNPB, dan Kemenko Maritim. Saat ini dosen Fakultas Kedokteran dan Ilmu Kesehatan Universitas Sultan Ageng Tirtayasa serta aktif di Asosiasi Dosen Indonesia, HAKLI, penelitian, dan pengabdian masyarakat.



REGRESI LOGISTIK DALAM EPIDEMIOLOGI

FITRI RACHMILLAH FADMI

E-mail: fitri.rachmillahfadmi@gmail.com

PENDAHULUAN

Epidemiologi sebagai disiplin ilmu yang mempelajari distribusi dan determinan masalah kesehatan dalam populasi membutuhkan pendekatan analisis statistik yang tepat untuk memahami hubungan antara berbagai faktor risiko dan kejadian penyakit. Dalam praktik penelitian kesehatan masyarakat, peneliti sering kali berhadapan dengan variabel hasil (outcome) yang bersifat kategorik, khususnya dalam bentuk dikotomi atau biner, seperti sakit dan tidak sakit, stunting dan tidak stunting, anemia dan tidak anemia, serta hidup dan meninggal. Karakteristik data semacam ini memerlukan metode analisis statistik yang mampu menggambarkan hubungan antara faktor determinan dengan kemungkinan terjadinya suatu kejadian kesehatan secara lebih tepat (Schober & Vetter, 2021).

Salah satu pendekatan analisis yang paling banyak digunakan untuk tujuan tersebut adalah regresi logistik. Metode ini dirancang untuk memodelkan hubungan antara satu atau lebih variabel independen dengan variabel dependen yang bersifat kategorik, khususnya variabel biner. Dalam konteks epidemiologi, regresi logistik memungkinkan peneliti untuk mengestimasi peluang terjadinya suatu kejadian kesehatan berdasarkan

keberadaan faktor-faktor tertentu serta menilai kekuatan hubungan antara faktor risiko dan outcome melalui ukuran asosiasi seperti odds ratio (Dey, 2025; Panda, 2022).

Berbeda dengan regresi linear yang digunakan untuk variabel dependen kontinu, regresi logistik dikembangkan untuk mengatasi keterbatasan analisis ketika variabel hasil bersifat kategorik. Penggunaan regresi linear pada data biner dapat menghasilkan nilai prediksi yang berada di luar rentang probabilitas yang logis, yaitu antara 0 dan 1. Oleh karena itu, regresi logistik menggunakan transformasi fungsi logit untuk menghubungkan variabel independen dengan probabilitas kejadian suatu outcome, sehingga model yang dihasilkan tetap konsisten dengan konsep probabilitas dalam analisis statistik (Hua, Stead, George, & Ganti, 2025).

Dalam perkembangan penelitian epidemiologi modern, regresi logistik tidak hanya digunakan untuk mengidentifikasi faktor risiko suatu penyakit, tetapi juga untuk mengendalikan pengaruh variabel perancu (confounding) serta mengevaluasi kemungkinan adanya interaksi antarvariabel dalam suatu model analisis. Dengan memasukkan beberapa variabel independen secara simultan, peneliti dapat memperoleh gambaran

yang lebih komprehensif mengenai faktor-faktor yang berperan dalam kejadian suatu masalah kesehatan (Dey, 2025).

Metode regresi logistik telah banyak diterapkan dalam berbagai penelitian kesehatan masyarakat, mulai dari analisis determinan stunting pada balita, faktor risiko penyakit tidak menular seperti diabetes dan hipertensi, hingga studi mengenai perilaku kesehatan masyarakat. Selain itu, perkembangan metodologi statistik dalam beberapa tahun terakhir juga menunjukkan adanya perluasan penggunaan regresi logistik, seperti pada analisis data spasial, analisis longitudinal, serta integrasi dengan pendekatan analitik yang lebih kompleks dalam epidemiologi modern (Das, 2023; Lee, Zito, Sang, & Dunson, 2024).

Dengan demikian, pemahaman yang baik mengenai konsep dasar, model matematis, serta interpretasi hasil regresi logistik menjadi hal yang sangat penting bagi peneliti di bidang epidemiologi dan kesehatan masyarakat. Pemahaman tersebut tidak hanya membantu dalam melakukan analisis data secara tepat, tetapi juga mendukung proses penarikan kesimpulan yang lebih akurat mengenai faktor-faktor yang memengaruhi terjadinya suatu masalah kesehatan. Oleh karena itu, bab

ini akan membahas secara sistematis mengenai konsep dasar regresi logistik, model analisis yang digunakan, tahapan penerapannya dalam penelitian epidemiologi, serta cara menginterpretasikan hasil analisis regresi logistik dalam konteks kesehatan masyarakat.

KONSEP DASAR REGRESI LOGISTIK

Dalam penelitian epidemiologi, analisis statistik berperan penting dalam membantu menjelaskan hubungan antara paparan (exposure) dan kejadian penyakit (outcome). Banyak permasalahan kesehatan masyarakat dinyatakan dalam bentuk outcome yang bersifat kategorik, terutama biner, seperti sakit dan tidak sakit, stunting dan tidak stunting, atau terpapar dan tidak terpapar suatu faktor risiko. Kondisi ini menuntut penggunaan metode analisis yang mampu menggambarkan hubungan antara berbagai faktor penentu kesehatan dengan kemungkinan terjadinya suatu kejadian penyakit secara tepat. Salah satu pendekatan analisis yang banyak digunakan untuk tujuan tersebut adalah regresi logistik, karena metode ini memungkinkan peneliti untuk memperkirakan peluang terjadinya suatu

outcome berdasarkan faktor-faktor yang memengaruhinya (Dabrowska-Szalotka et al. 2024; Rothman et al, 2024).

Secara konseptual, regresi logistik memodelkan hubungan antara variabel independen dan probabilitas kejadian melalui fungsi logit, sehingga nilai probabilitas yang dihasilkan tetap berada dalam rentang antara 0 hingga 1. Dalam penelitian epidemiologi, hasil analisis regresi logistik umumnya diinterpretasikan menggunakan ukuran asosiasi berupa Odds Ratio (OR), yang menggambarkan perbandingan peluang terjadinya suatu kejadian pada kelompok yang terpapar faktor risiko dibandingkan dengan kelompok yang tidak terpapar. Selain itu, kemampuan regresi logistik untuk menganalisis beberapa variabel secara simultan memungkinkan peneliti mengendalikan pengaruh faktor perancu dan memperoleh gambaran yang lebih komprehensif mengenai determinan suatu masalah kesehatan dalam populasi (Hartmann, Kopp, & Lois, 2023).

MODEL REGRESI LOGISTIK

Dalam epidemiologi, model regresi logistik digunakan untuk menjelaskan hubungan antara faktor paparan dengan kemungkinan terjadinya suatu kejadian

kesehatan. Model ini banyak diterapkan ketika variabel hasil bersifat biner, seperti adanya penyakit atau tidak adanya penyakit pada individu atau kelompok tertentu. Melalui model regresi logistik, peneliti dapat menilai seberapa besar kontribusi suatu faktor risiko terhadap peluang terjadinya suatu outcome dengan tetap mempertimbangkan faktor lain yang mungkin ikut berperan. Dengan demikian, pendekatan ini membantu peneliti memperoleh pemahaman yang lebih jelas mengenai hubungan antara paparan dan kejadian penyakit dalam suatu populasi (Rothman et al., 2024).

Secara matematis, regresi logistik memodelkan hubungan antara variabel independen dan probabilitas kejadian melalui fungsi logit. Transformasi ini digunakan untuk mengubah probabilitas kejadian menjadi logaritma dari perbandingan peluang (odds), sehingga hubungan antara variabel independen dan outcome dapat dinyatakan dalam bentuk linear. Model regresi logistik secara umum dapat dituliskan sebagai berikut:

$$\text{logit}(p) = \ln(p/(1-p)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

Keterangan :

p : probabilitas terjadinya suatu kejadian atau outcome yang diteliti (misalnya kejadian penyakit).

$1 - p$: probabilitas tidak terjadinya kejadian tersebut.

$\frac{p}{1-p}$: *odds*, yaitu perbandingan antara probabilitas terjadinya kejadian dan tidak terjadinya kejadian.

$\ln\left(\frac{p}{1-p}\right)$: fungsi **logit**, yaitu logaritma natural dari nilai odds.

Transformasi ini digunakan untuk mengubah hubungan yang bersifat nonlinier menjadi hubungan linier.

β_0 : konstanta atau *intercept*, yaitu nilai log odds ketika seluruh variabel independen bernilai nol.

$\beta_1, \beta_2, \dots, \beta_k$: koefisien regresi yang menunjukkan besarnya perubahan log odds pada outcome untuk setiap perubahan satu unit pada variabel independen.

X_1, X_2, \dots, X_k : variabel independen atau faktor risiko yang dianalisis dalam model regresi logistik.

Dalam persamaan tersebut, p menunjukkan probabilitas terjadinya suatu kejadian kesehatan, sedangkan β_0 merupakan konstanta model. Koefisien β_1, β_2 , hingga β_k menggambarkan perubahan nilai logit dari probabilitas kejadian yang berkaitan dengan masing-masing variabel independen (x_1, x_2, \dots, x_k). Dalam praktik analisis epidemiologi, nilai koefisien tersebut biasanya ditransformasikan dalam bentuk eksponensial untuk memperoleh odds ratio, yang digunakan untuk menafsirkan kekuatan hubungan antara faktor paparan dan outcome kesehatan (Kleinbaum & Klein, 2023).

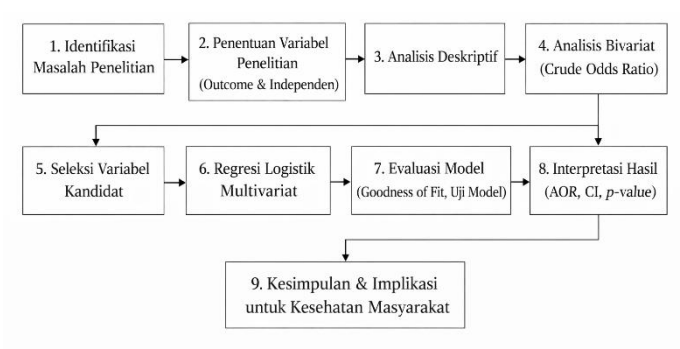
Dalam penerapannya, model regresi logistik dapat digunakan untuk analisis sederhana maupun analisis multivariat. Analisis sederhana biasanya digunakan untuk menilai hubungan antara satu faktor risiko dengan outcome, sedangkan analisis multivariat memungkinkan peneliti memasukkan beberapa variabel sekaligus dalam satu model. Pendekatan multivariat sangat penting dalam epidemiologi karena kejadian penyakit umumnya dipengaruhi oleh berbagai faktor yang saling berkaitan. Dengan mempertimbangkan berbagai variabel secara bersama-sama, model regresi logistik dapat memberikan gambaran yang lebih komprehensif mengenai faktor-faktor yang berkontribusi terhadap terjadinya suatu masalah kesehatan dalam populasi (Celentano, Szklo, & Farag, 2023; Pampel, 2020).

TAHAPAN ANALISIS REGRESI LOGISTIK DALAM PENELITIAN EPIDEMIOLOGI

Dalam penelitian epidemiologi, analisis regresi logistik dilakukan melalui beberapa tahapan yang terstruktur agar model yang dihasilkan mampu menggambarkan hubungan antara faktor paparan dan kejadian penyakit secara tepat. Proses analisis ini tidak

hanya berfokus pada pembentukan model statistik, tetapi juga mempertimbangkan relevansi epidemiologis dari setiap variabel yang dianalisis. Dengan mengikuti tahapan yang sistematis, peneliti dapat memperoleh hasil analisis yang lebih akurat serta mudah diinterpretasikan dalam konteks kesehatan masyarakat (hosmer, lemeshow, & sturdivant, 2013).

Tahapan analisis regresi logistik dalam penelitian epidemiologi secara umum dapat dijelaskan sebagai berikut:



1. Analisis deskriptif

Tahap awal dilakukan dengan menggambarkan karakteristik responden dan distribusi variabel penelitian. Analisis ini bertujuan untuk memberikan gambaran umum mengenai kondisi

data, termasuk proporsi outcome serta distribusi faktor risiko dalam populasi penelitian.

2. Analisis bivariat

Analisis bivariat dilakukan untuk menilai hubungan awal antara masing-masing variabel independen dengan variabel dependen. Hasil analisis ini membantu peneliti mengidentifikasi variabel yang berpotensi memiliki hubungan dengan outcome yang diteliti.

3. Penentuan variabel kandidat

Variabel yang menunjukkan hubungan pada tahap analisis bivariat, serta variabel yang secara teoritis relevan berdasarkan kajian epidemiologi, dipertimbangkan sebagai variabel kandidat yang akan dimasukkan ke dalam model regresi logistik.

4. Penyusunan model regresi logistik multivariat

Pada tahap ini beberapa variabel independen dimasukkan secara bersamaan dalam satu model analisis. Pendekatan ini memungkinkan peneliti menilai pengaruh masing-masing faktor risiko terhadap outcome setelah mempertimbangkan pengaruh variabel lain dalam model.

5. Evaluasi kesesuaian model

Model yang telah dibangun kemudian dievaluasi untuk menilai sejauh mana model tersebut mampu menjelaskan data penelitian. Evaluasi ini penting untuk memastikan bahwa model yang dihasilkan memiliki kemampuan yang baik dalam menggambarkan hubungan antara variabel yang dianalisis.

6. Interpretasi hasil analisis

Tahap akhir adalah menafsirkan hasil regresi logistik, terutama nilai odds ratio, interval kepercayaan, serta tingkat signifikansi statistik. Interpretasi ini kemudian dikaitkan dengan konsep epidemiologi untuk menjelaskan bagaimana faktor-faktor tertentu berperan dalam terjadinya suatu masalah kesehatan dalam populasi (rothman, greenland, & lash, 2021; szklo & nieto, 2019).

INTERPRETASI HASIL REGRESI LOGISTIK

Interpretasi hasil regresi logistik merupakan tahap penting dalam analisis epidemiologi karena melalui proses ini peneliti memahami makna hubungan antara faktor paparan dan kejadian penyakit berdasarkan model

statistik yang telah dibangun. Hasil analisis regresi logistik umumnya disajikan dalam beberapa komponen utama, seperti koefisien regresi, nilai odds ratio, interval kepercayaan (confidence interval), serta nilai signifikansi statistik. Komponen-komponen tersebut memberikan gambaran mengenai arah hubungan, kekuatan asosiasi, serta tingkat ketepatan estimasi yang diperoleh dari model analisis (Hosmer, Lemeshow, & Sturdivant, 2013).

Dalam praktik penelitian epidemiologi, interpretasi hasil regresi logistik biasanya dilakukan secara bertahap mengikuti proses analisis yang telah dilakukan sebelumnya.

1. Interpretasi Analisis Deskriptif

Pada tahap ini interpretasi difokuskan pada gambaran umum data penelitian, seperti karakteristik responden dan distribusi kejadian outcome dalam populasi penelitian. Analisis deskriptif membantu menjelaskan proporsi atau prevalensi kejadian penyakit serta distribusi faktor-faktor yang diduga berperan dalam populasi yang diteliti.

2. Interpretasi Analisis Bivariat

Analisis bivariat memberikan gambaran awal mengenai hubungan antara masing-masing variabel

independen dan outcome. Pada tahap ini hasil analisis biasanya dinyatakan dalam bentuk Odds Ratio (OR) atau Crude Odds Ratio (COR), yaitu ukuran asosiasi yang menggambarkan peluang terjadinya suatu kejadian pada kelompok yang terpapar dibandingkan dengan kelompok yang tidak terpapar tanpa mempertimbangkan pengaruh variabel lain dalam model.

3. Interpretasi Model Regresi Logistik Multivariat

Pada tahap multivariat, interpretasi difokuskan pada nilai $\text{Exp}(B)$ yang dikenal sebagai Adjusted Odds Ratio (AOR). Nilai ini menunjukkan perbandingan peluang terjadinya suatu kejadian pada kelompok yang terpapar suatu faktor setelah mempertimbangkan pengaruh variabel lain yang terdapat dalam model. Nilai AOR yang lebih besar dari satu menunjukkan bahwa faktor tersebut berkaitan dengan peningkatan peluang kejadian, sedangkan nilai yang lebih kecil dari satu menunjukkan kemungkinan adanya efek perlindungan dari faktor yang dianalisis. Dalam interpretasinya, nilai AOR biasanya disertai dengan interval kepercayaan dan nilai signifikansi statistik untuk menilai kekuatan serta ketepatan estimasi

hubungan yang diperoleh (Rothman, Greenland, & Lash, 2021).

4. Interpretasi Kesesuaian Model

Selain menafsirkan hubungan antarvariabel, peneliti juga perlu menilai kesesuaian model regresi logistik yang digunakan. Evaluasi ini bertujuan untuk memastikan bahwa model yang dibangun mampu menggambarkan hubungan antara variabel independen dan outcome secara memadai serta memberikan estimasi yang stabil.

5. Interpretasi dalam Konteks Epidemiologi

Tahap akhir interpretasi adalah mengaitkan hasil analisis statistik dengan kerangka teori epidemiologi serta temuan penelitian sebelumnya. Peneliti perlu mempertimbangkan plausibilitas biologis, karakteristik populasi yang diteliti, serta faktor lingkungan dan sosial yang mungkin memengaruhi kejadian penyakit. Dengan pendekatan ini, hasil analisis regresi logistik tidak hanya memberikan makna statistik, tetapi juga membantu menjelaskan faktor-faktor yang berperan dalam terjadinya suatu masalah kesehatan serta implikasinya bagi upaya pencegahan dan intervensi

kesehatan masyarakat (Jewell, 2020; Szklo & Nieto, 2019).

APLIKASI REGRESI LOGISTIK DALAM EPIDEMIOLOGI

Regresi logistik merupakan salah satu metode analisis yang banyak digunakan dalam penelitian epidemiologi untuk mempelajari hubungan antara berbagai faktor paparan dengan kejadian penyakit yang dinyatakan dalam bentuk outcome kategorik, khususnya biner. Metode ini memungkinkan peneliti menilai pengaruh beberapa faktor risiko secara bersamaan serta mengidentifikasi faktor mana yang tetap berhubungan dengan kejadian penyakit setelah mempertimbangkan pengaruh variabel lain. Oleh karena itu, regresi logistik sering digunakan dalam studi epidemiologi analitik untuk memahami determinan suatu masalah kesehatan dalam populasi serta mendukung pengambilan keputusan berbasis bukti dalam bidang kesehatan masyarakat ((Pampel, 2020). Penerapan regresi logistik dalam penelitian epidemiologi umumnya dilakukan melalui tahapan analisis yang sistematis, dimulai dari analisis deskriptif, analisis bivariat, hingga analisis regresi logistik

multivariat. Setiap tahapan memberikan informasi yang berbeda mengenai hubungan antara variabel yang diteliti dan kejadian penyakit. Contoh berikut menggambarkan penerapan regresi logistik dalam penelitian mengenai faktor-faktor yang berhubungan dengan kejadian stunting pada balita.

1. Analisis Deskriptif

Tahap awal analisis dilakukan dengan menggambarkan karakteristik responden dan distribusi variabel penelitian. Analisis deskriptif bertujuan untuk memberikan gambaran umum mengenai kondisi populasi penelitian, termasuk distribusi faktor risiko dan proporsi kejadian outcome.

Sebagai contoh, dari 400 balita yang diteliti diperoleh prevalensi stunting sebesar 28%. Selain itu, sebagian besar ibu memiliki tingkat pendidikan menengah (52%), sedangkan sekitar 30% ibu memiliki pendidikan rendah. Dari sisi kondisi sosial ekonomi, sekitar 35% keluarga berada dalam kategori ekonomi rendah. Sementara itu, sekitar 32% anak dilaporkan pernah mengalami penyakit infeksi dalam enam bulan terakhir (Fadmi, Mulyani, Justin, & Riza, 2025).

Informasi ini memberikan gambaran awal mengenai kondisi populasi penelitian serta membantu peneliti memahami distribusi faktor-faktor yang diduga berperan dalam kejadian stunting.

2. Analisis Bivariat

Tahap berikutnya adalah analisis hubungan antara masing-masing variabel independen dan kejadian stunting. Analisis ini bertujuan untuk mengidentifikasi variabel yang memiliki hubungan awal dengan outcome yang diteliti. Hasil analisis biasanya dinyatakan dalam bentuk Odds Ratio (OR) atau Crude Odds Ratio (COR), yaitu ukuran asosiasi yang menggambarkan peluang terjadinya suatu kejadian pada kelompok yang terpapar dibandingkan dengan kelompok yang tidak terpapar tanpa mempertimbangkan pengaruh variabel lain.

Tabel 7.1. Analisis Bivariat Faktor Risiko Stunting

Variabel	Kategori	Stunting n (%)	Tidak Stunting n (%)	OR	95% CI	p- value
Pendidikan Ibu	Rendah	60 (40,0)	90 (60,0)	2,1	1,3– 3,4	0,002
	Tinggi	35 (22,6)	120 (77,4)			
Status Ekonomi	Rendah	70 (43,8)	90 (56,2)	2,4	1,5– 3,8	0,001
	Tinggi	25 (18,5)	110 (81,5)			
	Ya	50 (41,7)	70 (58,3)	1,9		0,006

Variabel	Kategori	Stunting n (%)	Tidak Stunting n (%)	OR	95% CI	p-value
Riwayat Infeksi	Tidak	45 (23,7)	145 (76,3)		1,2–3,0	

Hasil analisis bivariat menunjukkan bahwa pendidikan ibu, status ekonomi keluarga, dan riwayat infeksi pada anak memiliki hubungan awal dengan kejadian stunting. Misalnya, anak yang berasal dari keluarga dengan status ekonomi rendah memiliki peluang sekitar 2,4 kali lebih besar mengalami stunting dibandingkan dengan anak dari keluarga dengan status ekonomi lebih baik.

3. Analisis Regresi Logistik Multivariat

Variabel yang memiliki hubungan pada analisis bivariat kemudian dimasukkan ke dalam model regresi logistik multivariat. Analisis ini bertujuan untuk menilai hubungan antara faktor risiko dan outcome setelah mempertimbangkan pengaruh variabel lain dalam model.

Pada tahap ini interpretasi difokuskan pada nilai Exp(B) atau Adjusted Odds Ratio (AOR), yaitu ukuran asosiasi yang menunjukkan peluang kejadian setelah dilakukan penyesuaian terhadap variabel lain dalam model.

Tabel 2. Hasil Analisis Regresi Logistik Multivariat Faktor risiko Stunting

Variabel	Kategori	Exp(B) / AOR	95% CI	p-value
Pendidikan Ibu	Rendah	1,7	1,1–2,9	0,021

Variabel	Kategori	Exp(B) / AOR	95% CI	p-value
Status Ekonomi	Rendah	1,9	1,2–3,1	0,008
Riwayat Infeksi	Ya	1,6	1,0–2,6	0,037

Hasil analisis multivariat menunjukkan bahwa status ekonomi keluarga tetap berhubungan dengan kejadian stunting setelah mempertimbangkan variabel lain dalam model. Hal ini menunjukkan bahwa faktor sosial ekonomi memiliki peran penting dalam menentukan status gizi anak.

4. Interpretasi dalam Konteks Epidemiologi

Berdasarkan hasil analisis regresi logistik multivariat, anak yang berasal dari keluarga dengan status ekonomi rendah memiliki peluang sekitar **1,9 kali lebih besar** mengalami stunting dibandingkan dengan anak dari keluarga dengan kondisi ekonomi yang lebih baik (**AOR = 1,9; 95% CI: 1,2–3,1**). Selain itu, pendidikan ibu yang rendah serta adanya riwayat infeksi pada anak juga menunjukkan hubungan yang bermakna dengan kejadian stunting.

Temuan ini menunjukkan bahwa kejadian stunting tidak hanya dipengaruhi oleh faktor gizi semata, tetapi juga oleh kondisi sosial ekonomi keluarga, tingkat pendidikan ibu, serta kondisi kesehatan anak. Oleh karena itu, upaya pencegahan stunting perlu dilakukan secara komprehensif dengan mempertimbangkan faktor-faktor sosial, ekonomi, dan kesehatan yang memengaruhi pertumbuhan dan perkembangan anak (Szklo & Nieto, 2019; Jewell, 2020).

DAFTAR PUSTAKA

- Celentano, D. D., Szklo, M., & Farag, Y. (2023). *Gordis epidemiology e-book: Gordis epidemiology e-book*. Elsevier health sciences.
- Dabrowska-Szalotka, A., Stanczyk-Szalotka, K., & Stanczyk-Szalotka, P. (2024). *Consumer and Innovations in e-Health Services: Examples of European Countries*. Routledge.
- Das, A. (2023). Logistic Regression. In F. Maggino (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 3985–3986). Cham: Springer International Publishing.
https://doi.org/10.1007/978-3-031-17299-1_1689
- Dey, D. (2025). The Proper Application of Logistic Regression Model in Complex Survey Data: A Systematic Review. *Journal of Applied Statistics*.
- Fadmi, F. R., Mulyani, S., Justin, W. O. S., & Riza, Y. (2025). Logistic Regression Analysis of Risk Factors for Stunting Among Toddlers Aged 24-59 Months in Southeast Sulawesi, Indonesia. *Health Dynamics*, 2(2), 85–91.
- Hartmann, F. G., Kopp, J., & Lois, D. (2023). Logistic Regression. In F. G. Hartmann, J. Kopp, & D. Lois, *Social Science Data Analysis* (pp. 141–163). Wiesbaden: Springer Fachmedien Wiesbaden.
https://doi.org/10.1007/978-3-658-41230-2_8
- Hua, Y., Stead, T. S., George, A., & Ganti, L. (2025). Logistic Regression as a Tool for Clinical Risk Prediction and Decision Making. *Academic Medicine and Surgery*.

- Kleinbaum, D. G., & Klein, M. (2023). *Logistic Regression: A Self-Learning Text* (4th ed.). New York: Springer.
- Lee, C. J., Zito, A., Sang, H., & Dunson, D. B. (2024). Logistic-beta processes for modeling dependent random probabilities with beta marginals. *arXiv Preprint arXiv:2402.07048*. processes-for-dependent-random-probabilities-with-beta-marginals.pdf
- Pampel, F. C. (2020). *Logistic regression: A primer*. Sage publications.
- Panda, N. R. (2022). A review on logistic regression in medical research. *National Journal of Community Medicine, 13*(04), 265–270.
- Rothman, K. J., Huybrechts, K. F., & Murray, E. J. (2024). *Epidemiology: An introduction*. Oxford university press

BIODATA PENULIS



Dr. Fitri Rachmillah Fadmi, S.KM., M.Kes lahir di Kolaka, Sulawesi Tenggara, pada 26 Oktober 1986. Penulis merupakan dosen di Universitas Mandala Waluya sejak tahun 2009. Pendidikan Magister Kesehatan Masyarakat dengan peminatan Biostatistik dan Kependudukan ditempuh di Fakultas Kesehatan Masyarakat Universitas Airlangga (2012–2014), dan melanjutkan Program Doktor (S3) Kesehatan Masyarakat sejak tahun 2021 dengan dukungan Beasiswa Pendidikan Indonesia (BPI), hingga selesai pada tahun 2024. Selain aktif dalam pengajaran, penulis juga konsisten dalam penelitian dan pengabdian kepada masyarakat, serta telah memperoleh berbagai hibah dari Kementerian sejak tahun 2016 hingga sekarang. Pengalaman tersebut menjadi landasan dalam menghasilkan karya ilmiah untuk mendukung pembangunan kesehatan dan peningkatan kualitas sumber daya manusia di Indonesia.



ANALISIS SURVIVAL DAN MODEL COX PROPORTIONAL HAZARDS

BS. TITI HAERANA

E-mail: bs.titihaerana@uin-alauddin.ac.id

PENDAHULUAN

Pada penelitian kesehatan, banyak pertanyaan ilmiah tidak hanya menanyakan apakah suatu kejadian terjadi, tetapi juga kapan kejadian itu muncul selama masa pengamatan. Waktu hingga kematian, progresi penyakit, kekambuhan, remisi, atau putus berobat merupakan luaran yang memiliki dimensi temporal, sehingga tidak memadai jika hanya direduksi menjadi status ya atau tidak pada akhir observasi. Dua pasien dapat sama-sama meninggal, tetapi kematian pada bulan ke-2 dan bulan ke-24 jelas memiliki makna klinis dan epidemiologis yang berbeda.

Analisis survival dikembangkan untuk menganalisis data waktu hingga kejadian (*time-to-event data*), terutama ketika sebagian subjek belum mengalami kejadian hingga akhir follow-up atau datanya tersensor. Berbeda dari regresi linear atau regresi logistik, analisis survival mempertahankan informasi waktu kejadian sekaligus mengakomodasi penyensoran, sehingga lebih sesuai untuk data klinis dan epidemiologis.

Dalam praktik, dua metode utama yang paling sering digunakan adalah estimator Kaplan–Meier dan model Cox proportional hazards. Kaplan–Meier digunakan untuk menggambarkan peluang bertahan secara deskriptif

sepanjang waktu, sedangkan model Cox digunakan untuk menilai pengaruh kovariat terhadap hazard kejadian. Bab ini membahas konsep dasar analisis survival, kurva Kaplan–Meier, hazard ratio, model Cox proportional hazards, serta contoh aplikasi pada pasien *diffuse histiocytic lymphoma*.

Waktu hingga kejadian (*time-to-event*)

Objek utama analisis survival adalah waktu sejak titik awal yang didefinisikan secara konsisten hingga terjadinya suatu kejadian tertentu. Titik awal dapat berupa waktu diagnosis, awal terapi, atau masuk studi, sedangkan kejadian dapat berupa kematian, kekambuhan, progresi penyakit, remisi, atau luaran lain yang dirumuskan secara operasional. Karena memuat informasi tentang waktu observasi dan status kejadian sekaligus, data *time-to-event* berbeda dari luaran biner biasa dan lebih tepat digunakan ketika fokus penelitian terletak pada dinamika kejadian sepanjang waktu.

Event dan censoring

Dalam analisis survival, *event* adalah kejadian utama yang diteliti dan harus didefinisikan secara spesifik serta konsisten. Ciri penting data survival adalah adanya *censoring*, terutama *right censoring*, yaitu ketika subjek belum mengalami *event* hingga studi berakhir, hilang dari

tindak lanjut, atau keluar dari penelitian sebelum kejadian terjadi. Dalam kondisi ini, waktu kejadian tidak diketahui secara lengkap, tetapi informasi hingga waktu terakhir pengamatan tetap bernilai sehingga subjek tersensor tetap berkontribusi dalam analisis.

Risk set dan person-time

Pada setiap titik waktu, hanya subjek yang belum mengalami *event* dan belum tersensor yang termasuk dalam *risk set*. Analisis survival juga menggunakan konsep *person-time*, yaitu akumulasi waktu berisiko yang disumbangkan setiap subjek selama masa pengamatan. Karena lama observasi antarindividu dapat berbeda, kontribusi informasi tiap subjek tidak hanya ditentukan oleh status kejadian, tetapi juga oleh lamanya waktu ia berada dalam kondisi berisiko.

Fungsi survival dan fungsi hazard

Dua fungsi utama dalam analisis survival adalah fungsi survival dan fungsi hazard. Fungsi survival, $S(t)$, menyatakan probabilitas bahwa subjek tetap bebas dari *event* hingga melewati waktu t , sedangkan fungsi hazard, $h(t)$, menggambarkan intensitas sesaat terjadinya kejadian pada waktu t di antara subjek yang masih berada dalam *risk set*. Dengan demikian, fungsi survival menunjukkan peluang bertahan secara

kumulatif, sedangkan fungsi hazard menunjukkan dinamika risiko sesaat.

KAPLAN–MEIER SURVIVAL CURVE

Estimator Kaplan–Meier adalah metode nonparametrik untuk menaksir fungsi survival dari data time-to-event. Metode ini tidak mensyaratkan bentuk distribusi waktu-kejadian tertentu, sehingga sangat berguna untuk menggambarkan peluang bertahan secara empiris. Kurva Kaplan–Meier berbentuk tangga karena probabilitas survival hanya berubah saat event terjadi, dengan tetap mempertahankan kontribusi subjek tersensor hingga waktu terakhir pengamatannya.

Kurva Kaplan–Meier dibaca melalui waktu pada sumbu horizontal dan probabilitas survival pada sumbu vertikal. Semakin tinggi kurva, semakin baik pengalaman survival kelompok tersebut. Kurva yang turun tajam menunjukkan event terjadi lebih cepat atau lebih sering, sedangkan kurva yang lebih landai menunjukkan subjek bertahan lebih lama. Titik sensor menandai subjek yang keluar dari pengamatan atau belum mengalami event hingga waktu tertentu.

Salah satu ringkasan penting dari kurva Kaplan–

Meier adalah median survival time (MST), yaitu waktu ketika probabilitas survival mencapai 0,50. Ukuran ini sering lebih informatif daripada rerata karena data survival umumnya tidak simetris dan mengandung penyensoran. Namun, median survival tidak selalu dapat dihitung apabila hingga akhir follow-up kurva tetap berada di atas 0,50.

Uji log-rank digunakan untuk menilai apakah perbedaan kurva Kaplan–Meier antar kelompok didukung secara statistik. Uji ini membandingkan jumlah event yang teramati dan yang diharapkan pada setiap waktu kejadian. Jika nilai-p kecil, maka fungsi survival antar kelompok dianggap berbeda. Namun, hasilnya perlu dibaca hati-hati, terutama bila kurva saling berpotongan atau efek kelompok berubah sepanjang follow-up.

MODEL COX PROPORTIONAL HAZARDS

Kaplan–Meier dan log-rank bersifat deskriptif dan belum memadai untuk menilai beberapa faktor secara simultan. Karena waktu kejadian dalam penelitian kesehatan biasanya dipengaruhi oleh banyak variabel, diperlukan model yang dapat menilai pengaruh kovariat terhadap hazard secara bersamaan. Model Cox

proportional hazards merupakan pendekatan yang paling sering digunakan untuk tujuan ini.

Model Cox adalah model semiparametrik yang menyatakan hazard individu sebagai hasil kali antara *baseline hazard* dan komponen eksponensial dari kombinasi kovariat. Keunggulannya adalah peneliti dapat menaksir efek relatif kovariat tanpa harus menetapkan bentuk distribusi hazard dasar. Dalam model Cox, eksponensial koefisien regresi menghasilkan *hazard ratio*. Jika model memuat beberapa kovariat, maka HR yang diperoleh merupakan *adjusted hazard ratio*, yaitu rasio hazard yang telah disesuaikan terhadap kovariat lain dalam model. Nilai ini lebih informatif daripada HR kasar, tetapi tidak otomatis menunjukkan hubungan kausal karena tetap bergantung pada kualitas data, variabel penyesuaian, dan kecocokan model.

Model Cox bertumpu pada asumsi *proportional hazards*, yaitu bahwa rasio hazard antar kelompok relatif tetap sepanjang waktu. Jika asumsi ini terpenuhi, satu nilai HR dapat digunakan sebagai ringkasan efek. Namun, jika efek kovariat berubah sepanjang *follow-up*, maka satu HR tunggal dapat menyederhanakan data secara berlebihan.

Jika asumsi *proportional hazards* tidak terpenuhi, model Cox dapat dimodifikasi, misalnya dengan memasukkan kovariat yang berubah menurut waktu, interaksi dengan waktu, stratifikasi, atau ukuran alternatif seperti *restricted mean survival time* (RMST). Intinya, satu HR tunggal tidak selalu cukup untuk menggambarkan dinamika kejadian bila efek kovariat berubah sepanjang pengamatan.

Tahapan umum analisis survival

Analisis survival yang baik tidak dimulai dari menjalankan perintah perangkat lunak, tetapi dari penataan logika analitik yang benar. Secara umum, tahapan analisis survival dapat diringkas sebagai berikut.

Pertama, peneliti harus mendefinisikan dengan jelas titik awal waktu, event, dan status sensor. Ketidakjelasan pada tahap ini akan menghasilkan waktu kejadian yang tidak sebanding antar subjek dan pada akhirnya menurunkan validitas analisis.

Kedua, data perlu dieksplorasi secara deskriptif. Peneliti perlu mengetahui jumlah subjek, jumlah event, jumlah data tersensor, lama pengamatan minimum dan maksimum, serta distribusi karakteristik dasar yang relevan.

Ketiga, data perlu ditetapkan sebagai data survival. Dalam perangkat lunak seperti Stata, tahap ini dilakukan melalui perintah `stset`, yang memberi tahu perangkat lunak tentang variabel waktu dan variabel status event.

Keempat, peneliti membangun deskripsi pola survival menggunakan kurva Kaplan–Meier. Pada tahap ini, fokus utamanya adalah menggambarkan peluang bertahan sepanjang waktu, baik secara keseluruhan maupun menurut kelompok tertentu.

Kelima, jika ada dua atau lebih kelompok yang dibandingkan, uji log-rank digunakan untuk menilai apakah perbedaan kurva survival didukung secara statistik.

Keenam, analisis dapat dilanjutkan dengan model Cox proportional hazards untuk menaksir *hazard ratio* dan menilai pengaruh kovariat secara simultan. Model ini sangat berguna ketika peneliti ingin mengontrol faktor lain seperti usia, jenis kelamin, stadium penyakit, atau komorbiditas.

Ketujuh, peneliti perlu memeriksa asumsi proportional hazards. Langkah ini tidak boleh diabaikan karena validitas interpretasi *hazard ratio* dalam model Cox bergantung pada masuk akal nya asumsi tersebut.

Kedelapan, hasil harus disusun menjadi interpretasi substantif yang hati-hati. Pembacaan hasil survival tidak cukup berhenti pada nilai-*p*. Peneliti juga harus memperhatikan bentuk kurva, arah efek, interval kepercayaan, dan konteks klinis atau epidemiologis dari kejadian yang dipelajari.

Contoh aplikasi: survival pasien *diffuse histiocytic lymphoma*

Subbab ini menyajikan contoh aplikasi analisis survival menggunakan data 80 pasien *diffuse histiocytic lymphoma* dengan stadium 3 dan stadium 4. Tujuan contoh ini adalah memperlihatkan bagaimana konsep dasar survival analysis diterapkan dalam analisis deskriptif dan inferensial.

Deskripsi data dan variabel (akses data set <https://acesse.one/ztkrlq8>)

No	Nama Variabel	Deskripsi	Kode
1	Stadium	Stadium lymphoma	3 Stadium 3 4 Stadium 4
2	Survival	Waktu pengamatan	
3	Censor	Penyebab kematian	0 Censor/masih hidup 1 Event/meninggal
4	Age_years	Usia responden	

Analisis menggunakan *software Stata*. Tampilan data seperti berikut

The screenshot shows the Stata Data Editor interface for a file named [berry.dta]. The main window displays a table with three columns: stadium, survival, and censor. The 'Snapshots' panel on the left indicates that rows 14 through 27 are currently visible. The data is as follows:

	stadium	survival	censor
14	3	255	0
15	3	270	0
16	3	310	0
17	3	316	0
18	3	335	0
19	3	346	0
20	4	4	1
21	4	6	1
22	4	10	1
23	4	11	1
24	4	11	1
25	4	11	1
26	4	13	1
27	4	17	1

```
. tabu censor
```

censor	Freq.	Percent	Cum.
0	26	32.50	32.50
1	54	67.50	100.00
Total	80	100.00	

Dari 80 subjek, terdapat 54 event kematian dan 26 observasi tersensor. Dengan demikian, proporsi data tersensor adalah 32,5%, sedangkan proporsi event adalah 67,5%.

```
. summ if stadium==3
Variable |      Obs      Mean      Std. Dev.      Min      Max
-----+-----
stadium |        19         3         0         3         3
survival |        19    173.8421    118.9305         6        346
censor  |        19    .4210526    .5072573         0         1

. summ if stadium==4
Variable |      Obs      Mean      Std. Dev.      Min      Max
-----+-----
stadium |        61         4         0         4         4
survival |        61    105.1639    99.53696         4        345
censor  |        61    .7540984    .4341942         0         1
```

Secara deskriptif, jumlah pasien stadium 3 adalah 19 orang, sedangkan pasien stadium 4 berjumlah 61 orang. Rata-rata waktu survival pada stadium 3 adalah sekitar 173,84 hari, sedangkan pada stadium 4 sekitar 105,16 hari. Gambaran awal ini mengisyaratkan bahwa pasien stadium 3 cenderung bertahan lebih lama daripada pasien stadium 4. Namun, perbedaan rerata seperti ini belum cukup untuk menarik kesimpulan inferensial karena data survival harus dibaca dengan mempertimbangkan censoring dan distribusi waktu kejadian.

Menyiapkan data survival di Stata

Dalam Stata, data survival harus terlebih dahulu dideklarasikan menggunakan perintah `stset`. Pada contoh ini, variabel waktu adalah `survival`, sedangkan variabel event adalah `cancel`. Setelah deklarasi dilakukan, diperoleh bahwa jumlah subjek adalah 80, jumlah kegagalan atau event adalah 54, dan total waktu risiko adalah 9.718 hari. Waktu keluar terakhir dari pengamatan adalah 346 hari.

```
. stset survival, failure (cancel)
      failure event:  cancel != 0 & cancel < .
obs. time interval:  (0, survival]
exit on or before:   failure
-----
      80 total obs.
       0 exclusions
-----
      80 obs. remaining, representing
      54 failures in single record/single failure data
9718 total analysis time at risk, at risk from t =      0
           earliest observed entry t =      0
           last observed exit t =      346
```

Secara substantif, hasil ini menunjukkan bahwa seluruh analisis selanjutnya akan dibangun di atas akumulasi waktu berisiko selama 9.718 hari. Angka total waktu risiko ini penting karena menegaskan bahwa survival analysis tidak hanya menghitung berapa banyak event yang terjadi, tetapi juga menghitung berapa lama subjek berada dalam keadaan berisiko sebelum event atau censoring terjadi.

. sts list, by {stadium}

Time	Day	Full	Leaf	Survivor Function	EG. Means	[95% Conf. Int.]

stadium=1						
8	18	1	0	0.8974	0.0211	0.8828 0.8998
18	18	1	0	0.8947	0.0704	0.8408 0.8726
22	17	1	0	0.8921	0.0827	0.8283 0.8948
42	16	2	0	0.7968	0.1920	0.4789 0.8810
43	14	2	1	0.7968	0.1920	0.4789 0.8810
84	13	1	0	0.8902	0.1980	0.4214 0.8401
126	11	2	1	0.8902	0.1980	0.4214 0.8401
168	11	2	1	0.8902	0.1980	0.4214 0.8401
207	10	1	0	0.8121	0.1167	0.2487 0.7921
211	8	2	1	0.8121	0.1167	0.2487 0.7921
227	8	2	1	0.8121	0.1167	0.2487 0.7921
232	7	1	0	0.2247	0.1287	0.2270 0.7583
232	6	2	1	0.2247	0.1287	0.2270 0.7583
270	2	2	1	0.2247	0.1287	0.2270 0.7583
210	4	2	1	0.2247	0.1287	0.2270 0.7583
216	2	2	1	0.2247	0.1287	0.2270 0.7583
220	2	2	1	0.2247	0.1287	0.2270 0.7583
246	1	2	1	0.2247	0.1287	0.2270 0.7583

stadium=2						
6	61	1	0	0.8826	0.0162	0.8882 0.8977
8	61	1	0	0.8871	0.0228	0.8728 0.8917
10	58	1	0	0.8908	0.0377	0.8528 0.8928
11	58	2	0	0.8916	0.0381	0.7961 0.8908
13	53	1	0	0.8821	0.0408	0.7743 0.8426
17	54	1	0	0.8889	0.0421	0.7548 0.8921
20	52	2	0	0.8921	0.0474	0.7167 0.8928
21	51	1	0	0.8187	0.0481	0.6881 0.8821
22	50	1	0	0.8923	0.0509	0.6787 0.8821
24	49	2	0	0.7702	0.0228	0.6422 0.8071
28	47	1	0	0.7941	0.0225	0.6237 0.8428
30	46	2	0	0.7822	0.0274	0.5896 0.8186
31	44	1	0	0.7908	0.0284	0.5722 0.8037
32	42	1	0	0.8823	0.0282	0.5282 0.7887
34	42	1	0	0.8721	0.0401	0.5282 0.7743
35	41	1	0	0.8227	0.0408	0.5282 0.7601
38	40	1	0	0.8282	0.0413	0.5207 0.7426
40	38	1	0	0.8220	0.0421	0.4882 0.7210
41	38	2	1	0.8220	0.0421	0.4882 0.7210
43	37	2	1	0.8220	0.0421	0.4882 0.7210
43	36	1	0	0.8226	0.0427	0.4714 0.7126
48	33	1	0	0.2882	0.0421	0.4242 0.7001
50	34	1	0	0.2710	0.0427	0.4270 0.6846
58	32	1	0	0.2227	0.0441	0.4200 0.6882
61	32	2	1	0.2227	0.0441	0.4200 0.6882
62	30	1	0	0.2223	0.0448	0.4014 0.6217
68	29	1	0	0.2168	0.0449	0.3822 0.6247
82	28	1	0	0.6884	0.0421	0.2626 0.6273
82	27	1	0	0.6789	0.0425	0.2478 0.6028
88	25	1	0	0.8824	0.0424	0.5226 0.8887
88	23	1	0	0.8820	0.0422	0.5122 0.8820
90	24	1	0	0.8242	0.0421	0.3881 0.8471
92	22	1	0	0.8821	0.0448	0.3782 0.8280
104	22	1	0	0.2878	0.0442	0.2626 0.5126
110	21	1	0	0.2882	0.0440	0.2442 0.4822
124	20	1	0	0.2227	0.0424	0.2200 0.4727
127	19	1	0	0.2221	0.0427	0.2141 0.4588
160	18	2	1	0.2221	0.0427	0.2141 0.4588
168	17	1	0	0.2127	0.0420	0.1871 0.4221
171	16	1	0	0.2821	0.0411	0.1802 0.4120
173	13	1	0	0.2728	0.0401	0.1662 0.3866
173	14	1	0	0.2241	0.0289	0.1482 0.2740
184	12	1	0	0.2242	0.0272	0.1287 0.2221
201	12	1	0	0.2120	0.0229	0.1176 0.2216
222	11	1	0	0.1824	0.0242	0.1026 0.2120
222	10	2	1	0.1824	0.0242	0.1026 0.2120
247	8	2	1	0.1824	0.0242	0.1026 0.2120
260	8	2	1	0.1824	0.0242	0.1026 0.2120
284	7	2	1	0.1824	0.0242	0.1026 0.2120
280	6	2	1	0.1824	0.0242	0.1026 0.2120
281	5	2	1	0.1824	0.0242	0.1026 0.2120
282	4	2	1	0.1824	0.0242	0.1026 0.2120
284	3	2	1	0.1824	0.0242	0.1026 0.2120
241	2	2	1	0.1824	0.0242	0.1026 0.2120
242	1	2	1	0.1824	0.0242	0.1026 0.2120

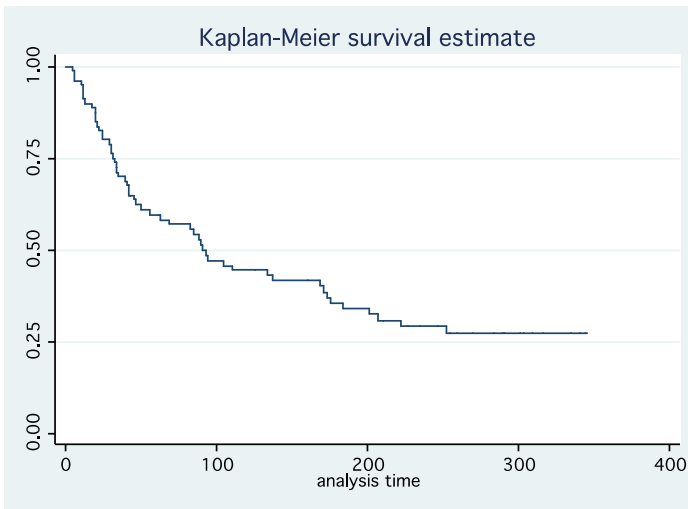
Berdasarkan hasil tersebut, terlihat bahwa kontribusi waktu pengamatan (*person-time*) antarindividu bervariasi,

dengan lama observasi minimum 4 hari dan maksimum 346 hari. Pada akhir periode pengamatan, probabilitas survival pasien stadium 3 masih sebesar 52%, sedangkan pada pasien stadium 4 hanya sekitar 19%. Temuan ini menunjukkan bahwa ketahanan hidup pasien stadium 3 relatif lebih baik dibandingkan pasien stadium 4 sepanjang masa follow-up.

Kaplan–Meier menurut stadium

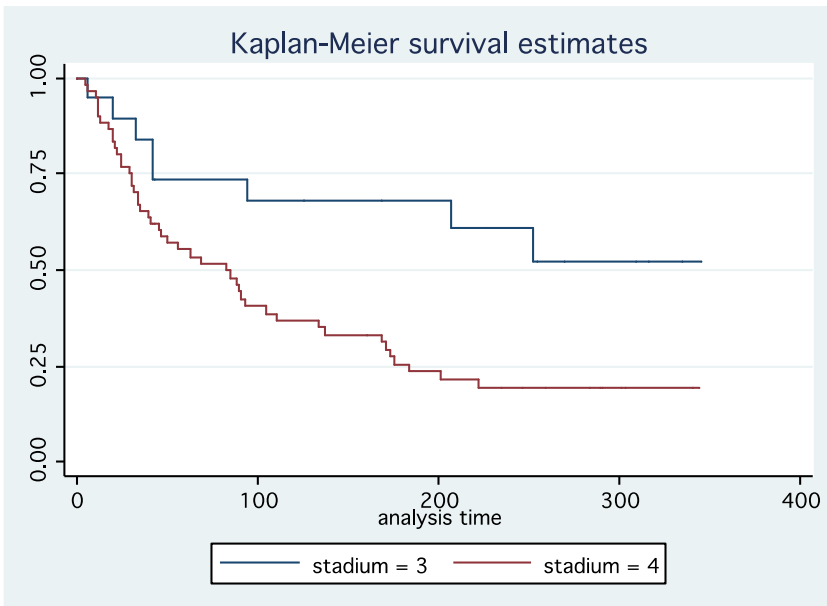
Setelah data dideklarasikan sebagai data survival, langkah berikutnya adalah membangun kurva Kaplan–Meier.

sts graph



Secara umum, kurva survival menunjukkan bahwa semakin lama pengamatan, semakin besar proporsi pasien yang mengalami event kematian. Secara keseluruhan, median survival pada data ini berada di sekitar 90 hari, yaitu saat probabilitas survival turun mendekati 0,50. Untuk melihat detailnya masukkan command dalam stata `.sts list`

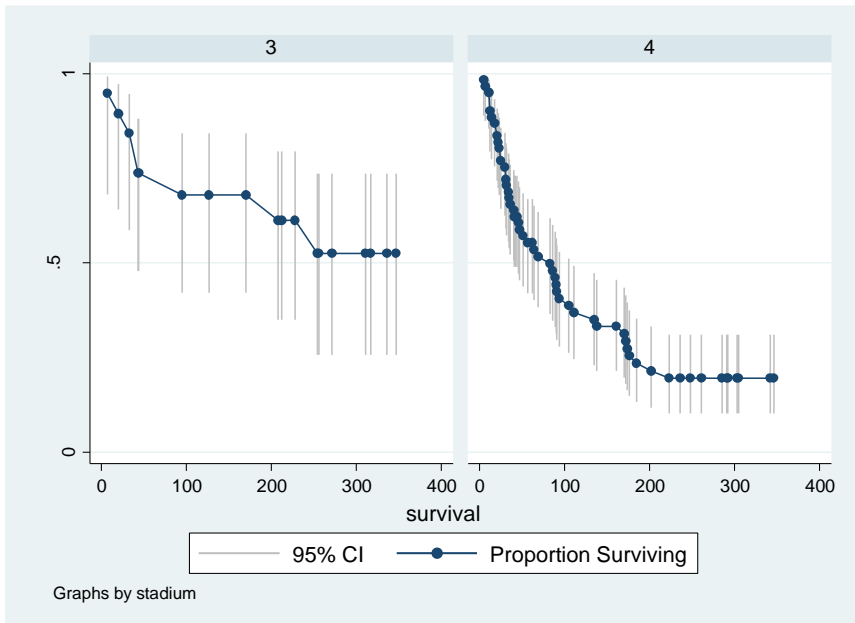
```
.sts graph, by (stadium)  
failure _d: censor  
analysis time _t: survival
```



Ketika kurva Kaplan–Meier ditampilkan menurut stadium, tampak bahwa pasien stadium 3 memiliki pola survival yang lebih baik daripada pasien stadium 4. Kurva stadium 3 berada lebih tinggi hampir sepanjang masa

pengamatan, yang berarti peluang bertahan pasien stadium 3 relatif lebih besar.

. ltable survival censor, by (stadium) graph ci



Daftar estimasi Kaplan–Meier menurut stadium juga memperlihatkan kontras ini dengan jelas. Pada stadium 3, probabilitas survival pada akhir pengamatan masih sekitar 52,47%, sedangkan pada stadium 4 hanya sekitar 19,54%. Artinya, pada akhir periode observasi, sekitar separuh pasien stadium 3 masih hidup, sementara pada stadium 4 hanya sekitar seperlima pasien yang masih bertahan.

```

.   sts list, by (stadium) compare
      failure _d: censor
      analysis time _t: survival
              Survivor Function
stadium              3              4
-----
time          4          1.0000          0.9836
              46          0.7368          0.5883
              88          0.7368          0.4614
             130          0.6802          0.3692
             172          0.6802          0.2932
             214          0.6121          0.2150
             256          0.5247          0.1954
             298          0.5247          0.1954
             340          0.5247          0.1954
             382          .              .
-----

```

Sebagai ilustrasi, pada hari ke-88 probabilitas survival pasien stadium 3 masih sekitar 73,68%, sedangkan stadium 4 sudah turun menjadi 46,14%. Pada hari ke-340, stadium 3 masih menunjukkan probabilitas survival sekitar 52,47%, sedangkan stadium 4 tetap berada pada kisaran 19,54%. Pola ini menunjukkan bahwa perbedaan survival antarkedua stadium bukan hanya terjadi pada akhir pengamatan, tetapi sudah tampak sejak fase yang relatif dini.

Uji log-rank

Untuk menilai apakah perbedaan kurva Kaplan–Meier menurut stadium didukung secara statistik, dilakukan uji

log-rank. Hasil uji menunjukkan nilai chi-kuadrat sebesar 6,71 dengan $p = 0,0096$.

```
.sts test stadium
      failure _d:  censor
      analysis time _t:  survival
Log-rank test for equality of survivor functions
-----+-----
stadium |      Events      Events
         | observed      expected
-----+-----
3       |          8          16.69
4       |         46          37.31
-----+-----
Total   |         54          54.00
         |      chi2(1) =         6.71
         |      Pr>chi2 =         0.0096
```

Karena nilai p lebih kecil dari 0,05, maka terdapat bukti yang cukup untuk menolak hipotesis nol bahwa fungsi survival pasien stadium 3 dan stadium 4 adalah sama. Dengan demikian, perbedaan kurva survival antara kedua kelompok didukung secara statistik.

Secara substantif, hasil ini menunjukkan bahwa stadium penyakit berhubungan dengan perbedaan pengalaman survival pada pasien *diffuse histiocytic lymphoma*. Pasien stadium 4 cenderung mengalami kematian lebih cepat dibandingkan pasien stadium 3. Namun, perlu diingat bahwa uji log-rank masih bersifat komparatif univariat. Uji ini belum menyesuaikan faktor lain seperti usia.

Model Cox bivariat

Langkah berikutnya adalah membangun model Cox bivariat dengan stadium sebagai satu-satunya kovariat. Hasil model menunjukkan bahwa *hazard ratio* untuk stadium 4 dibandingkan stadium 3 adalah 2,61 dengan interval kepercayaan 95% sebesar 1,23 sampai 5,57 dan nilai $p = 0,013$.

```
. stcox i.stadium

      failure _d:  censor
analysis time _t:  survival

Iteration 0:  log likelihood = -207.5548
Iteration 1:  log likelihood = -203.86666
Iteration 2:  log likelihood = -203.73805
Iteration 3:  log likelihood = -203.73761
Refining estimates:
Iteration 0:  log likelihood = -203.73761

Cox regression -- Breslow method for ties

No. of subjects =          80          Number of obs =          80
No. of failures =          54
Time at risk   =         9718

Log likelihood = -203.73761          LR chi2(1) =          7.63
                                      Prob > chi2 =          0.0057

-----+-----
      _t | Haz. Ratio  Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
    4.stadium |  2.614362   1.008191    2.49  0.013    1.227756   5.566976
-----+-----
```

Interpretasi yang tepat adalah bahwa hazard kematian pada pasien stadium 4 diperkirakan sekitar 2,61 kali hazard kematian pada pasien stadium 3. Karena interval kepercayaan seluruhnya berada di atas 1, bukti statistik untuk adanya perbedaan hazard dapat dianggap cukup kuat.

Hasil ini konsisten dengan kurva Kaplan–Meier dan uji log-rank. Secara umum, stadium 4 berhubungan dengan laju kematian yang lebih tinggi dibandingkan

stadium 3. Namun, model ini masih belum memperhitungkan faktor lain yang mungkin berperan sebagai perancu.

Model Cox multivariat

Analisis kemudian dilanjutkan dengan model Cox multivariat yang memasukkan stadium dan usia (*age_years*). Hasil model menunjukkan bahwa HR untuk stadium 4 dibandingkan stadium 3 adalah 1,91 dengan interval kepercayaan 95% sebesar 0,87 sampai 4,20 dan nilai $p = 0,108$. Sementara itu, usia memiliki HR sebesar 1,06 dengan interval kepercayaan 95% sebesar 1,03 sampai 1,09 dan nilai $p < 0,001$.

```
. stcox i.stadium age_years
      Failure _d: event
      Analysis time _t: survival_time

Iteration 0:  log likelihood = -207.5548
Iteration 1:  log likelihood = -196.86375
Iteration 2:  log likelihood = -196.75419
Iteration 3:  log likelihood = -196.75409
Refining estimates:
Iteration 0:  log likelihood = -196.75409

Cox regression with Breslow method for ties

No. of subjects =      80                Number of obs =      80
No. of failures =      54
Time at risk   = 9,718

Log likelihood = -196.75409                LR chi2(2)    = 21.60
                                           Prob > chi2   = 0.0000
```

	_t	Haz. ratio	Std. err.	z	P> z	[95% conf. interval]
4.stadium		1.908886	.7682037	1.61	0.108	.8674137 4.200818
age_years		1.05974	.0166812	3.69	0.000	1.027544 1.092944

Setelah dikontrol terhadap usia, stadium 4 masih menunjukkan arah efek yang sama, yaitu hazard kematian

yang lebih tinggi dibandingkan stadium 3. Namun, bukti statistiknya tidak lagi cukup kuat karena interval kepercayaan melintasi angka 1 dan nilai p lebih besar dari 0,05.

Sebaliknya, usia menunjukkan hubungan yang lebih stabil. Setiap kenaikan 1 tahun usia berhubungan dengan peningkatan hazard kematian sekitar 6%, dengan bukti statistik yang kuat. Temuan ini menunjukkan bahwa usia merupakan faktor prognosis yang penting dalam data ini.

Interpretasi hasil

Contoh ini memberikan pelajaran metodologis yang penting. Pada analisis bivariat, stadium tampak berhubungan kuat dengan hazard kematian. Namun, setelah usia dimasukkan ke dalam model, besar efek stadium menurun dan tidak lagi signifikan secara statistik. Hal ini menunjukkan bahwa pembacaan hubungan antara stadium dan survival dapat berubah setelah dilakukan penyesuaian terhadap kovariat lain.

Bagi pembaca, perbedaan antara hasil bivariat dan multivariat ini penting untuk dipahami. Hasil bivariat membantu menggambarkan asosiasi kasar, sedangkan hasil multivariat memberikan gambaran yang telah disesuaikan terhadap faktor lain. Karena itu, interpretasi

akhir dalam studi klinis umumnya lebih bertumpu pada model multivariat, selama pemilihan kovariat dan asumsi model dapat dipertanggungjawabkan.

Selain itu, contoh ini juga memperlihatkan bahwa statistik survival tidak boleh dibaca hanya dari satu angka. Kurva Kaplan–Meier, log-rank test, *hazard ratio*, interval kepercayaan, dan konteks klinis semuanya perlu dibaca secara terpadu.

SIMPULAN

Analisis survival digunakan ketika outcome penelitian berupa waktu hingga terjadinya suatu kejadian. Berbeda dari analisis biner biasa, metode ini mempertahankan informasi waktu kejadian sekaligus mengakomodasi penyensoran.

Konsep inti dalam survival analysis meliputi *time-to-event*, event, censoring, *risk set*, *person-time*, fungsi survival, dan fungsi hazard. Kaplan–Meier digunakan untuk menggambarkan peluang bertahan secara deskriptif, sedangkan uji log-rank digunakan untuk membandingkan kurva survival antar kelompok.

Model Cox proportional hazards digunakan ketika peneliti ingin menilai pengaruh satu atau lebih kovariat

terhadap hazard kejadian. Ukuran efek utama yang dihasilkan adalah *hazard ratio*, yang harus dibaca secara hati-hati karena berbeda dari *risk ratio* dan *odds ratio*.

Validitas interpretasi model Cox bergantung pada asumsi proportional hazards. Oleh sebab itu, pelaporan analisis survival yang baik harus mencakup bukan hanya nilai HR dan nilai-*p*, tetapi juga interval kepercayaan, konteks kejadian, dan pemeriksaan asumsi model.

Contoh aplikasi pada pasien *diffuse histiocytic lymphoma* menunjukkan bahwa pasien stadium 4 memiliki pengalaman survival yang lebih buruk daripada pasien stadium 3. Namun, setelah penyesuaian terhadap usia, efek stadium tidak lagi menunjukkan bukti statistik yang kuat, sedangkan usia tetap tampak sebagai faktor yang berhubungan dengan peningkatan hazard kematian.

DAFTAR PUSTAKA

- Abebe, T. B. et al. (2017) 'Anemia in Severe Heart Failure Patients: Does It Predict Prognosis?', *BMC Cardiovascular Disorders*, 17(1). doi: 10.1186/s12872-017-0680-5.
- Ali, A. K. H. and Mohammed, A. I. A. (2022) 'Log-Rank Test for Comparing Survival Curves of Neonatal Mortality Characteristic Groups in River Nile State-Sudan', *Open Journal of Statistics*, 12(04), pp. 532–547. doi: 10.4236/ojs.2022.124032.
- Aranda-Michel, E. et al. (2021) 'A Machine Learning Approach to Model for End-stage Liver Disease Score in Cardiac Surgery', *Journal of Cardiac Surgery*, 37(1), pp. 29–38. doi: 10.1111/jocs.16076.
- Bewick, V., Cheek, L. and Ball, J. (2004) 'No Title', *Critical Care*, 8(5), p. 389. doi: 10.1186/cc2955.
- Borgne, F. L. et al. (2015) 'Comparisons of the Performance of Different Statistical Tests for Time-to-event Analysis With Confounding Factors: Practical Illustrations in Kidney Transplantation', *Statistics in Medicine*, 35(7), pp. 1103–1116. doi: 10.1002/sim.6777.
- Cekic, S. et al. (2021) 'A Tutorial for Joint Modeling of Longitudinal and Time-to-Event Data in R', *Quantitative and Computational Methods in Behavioral Sciences*, 1. doi: 10.5964/qcmb.2979.
- Cioci, A. C. et al. (2021) 'Advanced Statistics: Multiple Logistic Regression, Cox Proportional Hazards, and Propensity Scores', *Surgical Infections*, 22(6), pp. 604–610. doi: 10.1089/sur.2020.425.

- Graeff, P. d. et al. (2006) 'Factors Influencing P53 Expression in Ovarian Cancer as a Biomarker of Clinical Outcome in Multicentre Studies', *British Journal of Cancer*, 95(5), pp. 627–633. doi: 10.1038/sj.bjc.6603300.
- Hemade, A. et al. (2025) 'Risk Factors for Infectious Disease Mortality in Breast Cancer: A Retrospective Cohort Study and Nomogram Development'. doi: 10.21203/rs.3.rs-6209450/v1.
- Hessol, N. A. et al. (2000) 'Factors Associated With Incident Self-Reported AIDS Among Women Enrolled in the Women's Interagency HIV Study (WIHS)', *Aids Research and Human Retroviruses*, 16(12), pp. 1105–1111. doi: 10.1089/088922200414947.
- Hua, Q. et al. (2021) 'Prognostic Factors of Survival of Advanced Liver Cancer Patients Treated With Palliative Radiotherapy: A Retrospective Study', *Frontiers in Oncology*, 11. doi: 10.3389/fonc.2021.658152.
- Jung, J., Song, H. and Han, S. (2023) 'Propensity Score Methods for Estimating Treatment Delay Effects', *Korean Journal of Applied Statistics*, 36(5), pp. 415–445. doi: 10.5351/kjas.2023.36.5.415.
- Karim, M. E. et al. (2016) 'Comparison of Statistical Approaches Dealing With Time-Dependent Confounding in Drug Effectiveness Studies', *Statistical Methods in Medical Research*, 27(6), pp. 1709–1722. doi: 10.1177/0962280216668554.
- Langholz, B. and Richardson, D. B. (2009) 'Fitting General Relative Risk Models for Survival Time and Matched Case-Control Analysis', *American Journal of Epidemiology*, 171(3), pp. 377–383. doi: 10.1093/aje/kwp403.
- Liem, Y. S. et al. (2010) 'Propensity Scores in the Presence of Effect Modification: A Case Study Using the

- Comparison of Mortality on Hemodialysis Versus Peritoneal Dialysis', *Emerging Themes in Epidemiology*, 7(1). doi: 10.1186/1742-7622-7-1.
- Lucijanić, M. and Petrovečki, M. (2012) 'Analysis of Censored Data', *Biochemia Medica*, pp. 151–155. doi: 10.11613/bm.2012.018.
- Möllenhoff, K. and Tresch, A. (2023) 'Investigating Non-Inferiority or Equivalence in Time-to-Event Data Under Non-Proportional Hazards', *Lifetime Data Analysis*, 29(3), pp. 483–507. doi: 10.1007/s10985-023-09589-5.
- Nguyen, T. et al. (2017) 'Comparison of the Ability of Double-robust Estimators to Correct Bias in Propensity Score Matching Analysis. A Monte Carlo Simulation Study', *Pharmacoepidemiology and Drug Safety*, 26(12), pp. 1513–1519. doi: 10.1002/pds.4325.
- Qiu, X. et al. (2021) 'The Role of Radiotherapy in Soft Tissue Sarcoma on Extremities With Lymph Nodes Metastasis: An IPTW Propensity Score Analysis of the SEER Database', *Frontiers in Oncology*, 11. doi: 10.3389/fonc.2021.751180.
- Rich, J. T. et al. (2010) 'A Practical Guide to Understanding Kaplan-Meier Curves', *Otolaryngology*, 143(3), pp. 331–336. doi: 10.1016/j.otohns.2010.05.007.
- Salazar, A. M. and Huertas, J. (2023) 'A Joint Model of Competing Risks in Discrete Time With Longitudinal Information', *Revista Colombiana de Estadística*, 46(2), pp. 145–161. doi: 10.15446/rce.v46n2.98005.
- Schmid, M. and Berger, M. (2020) 'Competing Risks Analysis for Discrete Time-to-event Data', *Wiley Interdisciplinary Reviews Computational Statistics*, 13(5). doi: 10.1002/wics.1529.
- Suhartini, A., Rahmawati, R. and Suparti, S. (2018) 'Analisis Kurva Survival Kaplan Meier

- Menggunakan Uji Log Rank', *Jurnal Gaussian*, 7(1), pp. 33–42. doi: 10.14710/j.gauss.v7i1.26633.
- Tolosie, K. and Sharma, M. (2014) 'Application of Cox Proportional Hazards Model in Case of Tuberculosis Patients in Selected Addis Ababa Health Centres, Ethiopia', *Tuberculosis Research and Treatment*, 2014, pp. 1–11. doi: 10.1155/2014/536976.
- Voeltz, D. et al. (2024) 'A Parametric Additive Hazard Model for Time-to-Event Analysis', *BMC Medical Research Methodology*, 24(1). doi: 10.1186/s12874-024-02180-y.
- Xiang, L., Ma, X. and Yau, K. K. (2011) 'Mixture Cure Model With Random Effects for Clustered Interval-censored Survival Data', *Statistics in Medicine*, 30(9), pp. 995–1006. doi: 10.1002/sim.4170.
- Xie, J. and Liu, C. (2005) 'Adjusted Kaplan-Meier Estimator and Log-rank Test With Inverse Probability of Treatment Weighting for Survival Data', *Statistics in Medicine*, 24(20), pp. 3089–3110. doi: 10.1002/sim.2174.
- Xue, X. et al. (2013) 'Testing the Proportional Hazards Assumption in Case-Cohort Analysis', *BMC Medical Research Methodology*, 13(1). doi: 10.1186/1471-2288-13-88.
- Yan, L. et al. (2016) 'Revealing Facts and Avoiding Biases: A Review of Several Common Problems in Statistical Analyses of Epidemiological Data', *Frontiers in Public Health*, 4. doi: 10.3389/fpubh.2016.00207.
- Zhou, B., Fine, J. P. and Laird, G. (2013) 'Goodness-of-fit Test for Proportional Subdistribution Hazards Model', *Statistics in Medicine*, 32(22), pp. 3804–3811. doi: 10.1002/sim.5815

BIODATA PENULIS



Dr. Bs. Titi Haerana, SKM., M.Kes lahir di Liu (Kab. Wajo, Sul-Sel). Menyelesaikan jenjang S1-S3 dengan konsentrasi keilmuan Epidemiologi di Universitas Hasanuddin dan Universitas Indonesia. Melakukan kajian dengan pemanfaatan data epidemiologi baik data primer maupun data sekunder menjadi fokus keilmuan

9



INFERENSI KAUSAL: TEORI DAN METODE KONTEMPORER

DEBI SUSANTI

E-mail: debisusantidosen@wdh.ac.id

PENDAHULUAN

Inferensi kausal merupakan salah satu konsep fundamental yang memiliki peranan sangat penting dalam perkembangan ilmu epidemiologi serta penelitian kesehatan masyarakat. Dalam konteks penelitian kesehatan, inferensi kausal digunakan untuk memahami serta menjelaskan hubungan sebab–akibat antara suatu paparan dengan luaran kesehatan yang diamati pada individu maupun populasi. Tujuan utama dari pendekatan ini adalah untuk menentukan apakah suatu faktor benar-benar berperan sebagai penyebab terjadinya suatu penyakit atau kondisi kesehatan tertentu (Hernán & Robins, 2020). Dengan demikian, inferensi kausal tidak hanya berfokus pada hubungan statistik antarvariabel, tetapi juga pada interpretasi ilmiah mengenai mekanisme biologis, sosial, maupun lingkungan yang mendasari hubungan tersebut.

Dalam ilmu kesehatan masyarakat, kemampuan untuk mengidentifikasi hubungan sebab-akibat memiliki implikasi yang sangat luas dan strategis. Hasil penelitian epidemiologi sering kali digunakan sebagai dasar dalam pengambilan keputusan klinis, penyusunan kebijakan kesehatan, serta perancangan berbagai program intervensi kesehatan masyarakat. Sebagai contoh, kebijakan

mengenai pengendalian tembakau, promosi aktivitas fisik, serta program pencegahan penyakit menular sebagian besar didasarkan pada bukti epidemiologis mengenai hubungan antara faktor risiko dan kejadian penyakit (Gordis, 2014). Oleh karena itu, kesimpulan mengenai hubungan sebab-akibat harus didasarkan pada analisis yang valid dan metodologi yang kuat.

Dalam praktik penelitian ilmiah, hubungan antara dua variabel biasanya pertama kali diidentifikasi melalui analisis statistik. Analisis tersebut dapat menunjukkan adanya hubungan atau korelasi antara variabel paparan dan variabel luaran kesehatan. Namun demikian, keberadaan korelasi tidak selalu menunjukkan adanya hubungan sebab-akibat. Dua variabel dapat terlihat saling berkaitan secara statistik karena dipengaruhi oleh faktor lain yang tidak diukur atau tidak diperhitungkan dalam analisis penelitian (Rothman et al., 2021). Fenomena ini menunjukkan bahwa interpretasi terhadap hasil analisis statistik memerlukan kehati-hatian serta pemahaman konseptual mengenai kausalitas.

Perbedaan antara asosiasi dan kausalitas merupakan salah satu isu metodologis yang paling penting dalam epidemiologi. Asosiasi merujuk pada adanya hubungan statistik antara dua variabel yang dapat diamati

dalam data penelitian. Hubungan tersebut dapat bersifat positif atau negatif, kuat maupun lemah. Sebaliknya, kausalitas merujuk pada situasi di mana perubahan pada suatu variabel secara langsung menyebabkan perubahan pada variabel lain melalui mekanisme tertentu (Szklo & Nieto, 2019). Dengan kata lain, kausalitas merujuk pada perubahan pada luaran kesehatan yang terjadi sebagai akibat dari perubahan pada suatu paparan, dengan membandingkan hasil yang akan terjadi pada kondisi paparan yang berbeda

Kesalahan dalam menafsirkan asosiasi sebagai hubungan kausal dapat menimbulkan berbagai konsekuensi yang merugikan. Misalnya, suatu intervensi kesehatan mungkin dikembangkan berdasarkan hubungan statistik yang sebenarnya bersifat semu. Jika intervensi tersebut diterapkan secara luas, maka sumber daya kesehatan masyarakat dapat terbuang tanpa menghasilkan manfaat yang signifikan. Oleh karena itu, penilaian hubungan kausal harus dilakukan secara hati-hati dengan mempertimbangkan berbagai faktor metodologis yang dapat memengaruhi validitas hasil penelitian (Rothman et al., 2021).

Dalam penelitian observasional, tantangan dalam mengidentifikasi hubungan kausal menjadi lebih

kompleks. Hal ini disebabkan karena peneliti tidak memiliki kendali penuh terhadap proses penentuan paparan yang dialami oleh subjek penelitian. Berbeda dengan penelitian eksperimental seperti randomized controlled trial (RCT), dalam studi observasional paparan terjadi secara alami tanpa intervensi langsung dari peneliti (Hernán & Robins, 2020). Kondisi ini menyebabkan hubungan antara paparan dan luaran kesehatan dapat dipengaruhi oleh berbagai faktor lain yang sulit dikendalikan secara langsung.

Salah satu tantangan utama dalam penelitian observasional adalah adanya variabel perancu atau confounding. Variabel perancu merupakan variabel yang berhubungan dengan paparan sekaligus berkaitan dengan luaran kesehatan. Kehadiran variabel ini dapat menciptakan hubungan semu antara paparan dan luaran sehingga menimbulkan kesan bahwa paparan tersebut menyebabkan luaran tertentu, padahal sebenarnya hubungan tersebut dipengaruhi oleh faktor lain (VanderWeele, 2015). Oleh karena itu, identifikasi dan pengendalian variabel perancu menjadi langkah yang sangat penting dalam analisis epidemiologi.

Selain confounding, penelitian observasional juga rentan terhadap bias seleksi dan bias informasi. Bias

seleksi terjadi ketika proses pemilihan subjek penelitian tidak mewakili populasi sasaran secara tepat. Hal ini dapat terjadi, misalnya, apabila partisipasi dalam penelitian dipengaruhi oleh status kesehatan atau karakteristik tertentu yang berkaitan dengan paparan yang diteliti. Sementara itu, bias informasi terjadi akibat kesalahan dalam proses pengukuran atau pencatatan variabel penelitian, seperti kesalahan dalam diagnosis penyakit atau ketidakakuratan dalam pelaporan paparan (Rothman et al., 2021).

Perkembangan metodologi inferensi kausal modern tidak terlepas dari kontribusi sejumlah tokoh penting dalam bidang epidemiologi dan statistika. Salah satu tokoh yang memiliki pengaruh besar adalah Austin Bradford Hill, yang memperkenalkan seperangkat kriteria untuk menilai kemungkinan adanya hubungan sebab-akibat antara paparan dan penyakit (Hill, 1965). Kriteria ini kemudian menjadi salah satu dasar penting dalam evaluasi bukti epidemiologi.

Selain itu, Donald Rubin mengembangkan kerangka konseptual yang dikenal sebagai Rubin Causal Model, yang didasarkan pada konsep potential outcomes. Kerangka ini memberikan landasan teoritis yang kuat dalam memahami efek kausal suatu perlakuan atau intervensi (Rubin, 1974).

Kontribusi penting lainnya berasal dari James Robins yang mengembangkan berbagai metode analisis untuk mengatasi masalah confounding yang berubah terhadap waktu dalam studi longitudinal (Robins et al., 2000).

Tokoh lain yang memberikan kontribusi besar dalam pengembangan inferensi kausal adalah Judea Pearl. Ia memperkenalkan pendekatan grafik kausal melalui penggunaan Directed Acyclic Graph (DAG), yang memungkinkan peneliti memvisualisasikan hubungan antarvariabel serta mengidentifikasi potensi bias dalam analisis kausal (Pearl, 2009). Pendekatan ini menjadi sangat penting dalam epidemiologi modern karena membantu peneliti merumuskan asumsi kausal secara eksplisit sebelum melakukan analisis statistik.

Seiring dengan perkembangan teknologi komputasi serta meningkatnya ketersediaan data berskala besar, metode inferensi kausal juga mengalami perkembangan pesat melalui integrasi metode inferensi kausal dengan teknik pembelajaran mesin untuk meningkatkan akurasi estimasi efek kausal. Pendekatan-pendekatan ini memungkinkan analisis hubungan kausal yang lebih kompleks dan akurat dalam berbagai konteks penelitian kesehatan.

Dengan demikian, inferensi kausal dapat dipahami sebagai proses ilmiah yang bertujuan untuk menilai apakah hubungan yang diamati dalam data benar-benar mencerminkan hubungan sebab-akibat. Proses ini melibatkan kombinasi antara analisis statistik, kerangka konseptual teoritis, serta pemahaman mendalam mengenai mekanisme biologis dan sosial yang mendasari terjadinya penyakit (Hernán & Robins, 2020).

KONSEP KAUSALITAS DALAM EPIDEMIOLOGI

Epidemiologi merupakan cabang ilmu kesehatan yang mempelajari distribusi serta determinan penyakit dalam populasi manusia. Ilmu ini berperan penting dalam memahami bagaimana penyakit terjadi, bagaimana penyakit menyebar dalam populasi, serta faktor-faktor apa saja yang berkontribusi terhadap kejadian penyakit tersebut (Gordis, 2014).

Melalui pendekatan ilmiah yang sistematis, epidemiologi berupaya mengidentifikasi faktor risiko yang berkaitan dengan berbagai kondisi kesehatan. Faktor risiko tersebut dapat berupa faktor biologis, lingkungan, perilaku, maupun sosial ekonomi. Dengan memahami faktor-faktor tersebut, para peneliti dan praktisi kesehatan

dapat merancang strategi pencegahan yang lebih efektif untuk mengurangi beban penyakit dalam masyarakat.

Secara historis, konsep kausalitas dalam epidemiologi mengalami perkembangan yang cukup panjang. Pada awal perkembangan ilmu kedokteran modern, pendekatan yang digunakan untuk menentukan penyebab penyakit terutama didasarkan pada postulat Henle–Koch. Postulat ini menekankan bahwa suatu mikroorganisme dapat dianggap sebagai penyebab penyakit apabila memenuhi serangkaian kriteria tertentu, seperti ditemukan pada individu yang sakit dan mampu menimbulkan penyakit yang sama ketika diinokulasikan pada organisme sehat (Evans, 1976).

Namun demikian, pendekatan tersebut memiliki keterbatasan ketika diterapkan pada penyakit kronis. Penyakit seperti penyakit jantung, diabetes, kanker, dan penyakit paru kronis tidak disebabkan oleh satu faktor tunggal, melainkan oleh kombinasi berbagai faktor yang saling berinteraksi (Szklo & Nieto, 2019). Oleh karena itu, epidemiologi modern mengembangkan pendekatan multikausal untuk menjelaskan proses terjadinya penyakit.

Salah satu model yang banyak digunakan adalah sufficient-component cause model yang dikembangkan oleh

Rothman. Dalam model ini, suatu penyakit dianggap terjadi ketika berbagai komponen penyebab bergabung membentuk kombinasi yang cukup untuk memicu terjadinya penyakit (Rothman et al., 2021). Model ini sering digambarkan dalam bentuk causal pie, di mana setiap irisan pie merepresentasikan faktor penyebab yang berkontribusi terhadap terjadinya penyakit.

Kriteria Bradford Hill

Kriteria Bradford Hill merupakan salah satu kerangka kerja yang paling berpengaruh dalam penilaian hubungan kausal dalam epidemiologi. Hill mengemukakan sembilan pertimbangan (considerations) yang dapat membantu peneliti menilai kemungkinan hubungan kausal, apakah suatu hubungan statistik antara paparan dan penyakit kemungkinan mencerminkan hubungan sebab-akibat yang sebenarnya (Hill, 1965).

Kesembilan kriteria tersebut meliputi:

- 1) Kekuatan asosiasi
- 2) Konsistensi hasil penelitian
- 3) Spesifisitas hubungan
- 4) Urutan temporal
- 5) Gradien biologis (dose-response)
- 6) Plausibilitas biologis

- 7) Koherensi dengan pengetahuan ilmiah
- 8) Bukti eksperimental
- 9) Analogi

Kriteria ini tidak dimaksudkan sebagai aturan yang harus dipenuhi secara mutlak, tetapi sebagai panduan dalam menilai bukti epidemiologi secara keseluruhan (Rothman et al., 2021).

DIRECTED ACYCLIC GRAPH (DAG)

Directed Acyclic Graph (DAG) merupakan alat konseptual yang digunakan untuk merepresentasikan hubungan kausal antarvariabel dalam bentuk grafik berarah tanpa siklus (Pearl, 2009). Dalam grafik tersebut, setiap simpul merepresentasikan variabel, sedangkan panah menunjukkan arah hubungan kausal antarvariabel.

Penggunaan DAG memungkinkan peneliti untuk menyatakan asumsi kausal secara eksplisit sebelum melakukan analisis statistik. Dengan memvisualisasikan hubungan antarvariabel, peneliti dapat mengidentifikasi variabel perancu, mediator, maupun collider yang dapat memengaruhi interpretasi hubungan kausal (Pearl et al., 2016).

COUNTERFACTUAL FRAMEWORK

Kerangka counterfactual merupakan salah satu dasar utama dalam inferensi kausal modern. Pendekatan ini menyatakan bahwa efek kausal dapat dipahami sebagai perbandingan antara dua kemungkinan hasil yang dapat dialami oleh individu yang sama dalam kondisi paparan yang berbeda (Rubin, 1974).

Dalam kerangka ini, setiap individu secara teoritis memiliki dua kemungkinan hasil yang disebut *potential outcomes*. Namun dalam praktik penelitian, hanya satu dari kedua kemungkinan tersebut yang dapat diamati. Kondisi ini dikenal sebagai *fundamental problem of causal inference* (Holland, 1986).

PENDEKATAN MODERN DALAM INFERENSI KAUSAL

Perkembangan metodologi inferensi kausal modern telah menghasilkan berbagai metode analitik yang digunakan untuk meningkatkan validitas estimasi efek kausal. Salah satu metode yang banyak digunakan adalah *propensity score*, yang bertujuan menyeimbangkan karakteristik antara kelompok perlakuan dan kelompok pembanding (Rosenbaum & Rubin, 1983).

Selain itu, *marginal structural models* dikembangkan untuk mengatasi masalah confounding yang berubah terhadap waktu dalam studi longitudinal (Robins et al., 2000). Metode instrumental variable juga digunakan ketika terdapat variabel perancu yang tidak dapat diukur secara langsung (Angrist & Pischke, 2009).

Dalam beberapa tahun terakhir, integrasi metode pembelajaran mesin dengan analisis kausal juga semakin berkembang. Pendekatan seperti *targeted maximum likelihood estimation* memungkinkan analisis data dengan dimensi tinggi serta hubungan antarvariabel yang kompleks (van der Laan & Rose, 2011).

SIMPULAN

Inferensi kausal merupakan komponen yang sangat penting dalam penelitian epidemiologi dan kesehatan masyarakat. Melalui pendekatan ini, peneliti dapat memahami hubungan sebab-akibat antara faktor paparan dan luaran kesehatan secara lebih mendalam.

Berbagai kerangka konseptual seperti kriteria Bradford Hill, Directed Acyclic Graph, serta kerangka counterfactual memberikan dasar teoritis yang kuat dalam analisis kausal. Selain itu, perkembangan metode statistik

modern telah meningkatkan kemampuan peneliti dalam mengestimasi efek kausal secara lebih akurat.

Dengan penerapan metodologi yang tepat, inferensi kausal dapat menghasilkan bukti ilmiah yang lebih valid dan dapat diandalkan untuk mendukung pengambilan keputusan dalam praktik klinis maupun kebijakan kesehatan masyarakat,

DAFTAR PUSTAKA

- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Evans, A. S. (1976). Causation and disease: The Henle–Koch postulates revisited. *Yale Journal of Biology and Medicine*, 49(2), 175–195.
- Gordis, L. (2014). *Epidemiology* (5th ed.). Elsevier Saunders.
- Hernán, M. A., & Robins, J. M. (2020). *Causal inference: What if*. Chapman & Hall/CRC.
- Hill, A. B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*, 58(5), 295–300.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396), 945–960. <https://doi.org/10.1080/01621459.1986.10478354>
- Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge University Press.
- Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal inference in statistics: A primer*. Wiley.
- Robins, J. M., Hernán, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5), 550–560. <https://doi.org/10.1097/00001648-200009000-00011>
- 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.

- <https://doi.org/10.1093/biomet/70.1.41>
- Rothman, K. J., Greenland, S., & Lash, T. L. (2021). *Modern epidemiology* (4th ed.). Wolters Kluwer.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. <https://doi.org/10.1037/h0037350>
- Szklo, M., & Nieto, F. J. (2019). *Epidemiology: Beyond the basics* (4th ed.). Jones & Bartlett Learning.
- van der Laan, M. J., & Rose, S. (2011). *Targeted learning: Causal inference for observational and experimental data*. Springer.
- VanderWeele, T. J. (2015). *Explanation in causal inference: Methods for mediation and interaction*. Oxford University Press

BIODATA PENULIS



Debi Susanti, S.Si., Apt., M.Farm

Penulis lahir di Cianjur pada Bulan Oktober 1982. Saat ini mengajar di Program Studi Farmasi Klinis dan Komunitas, Fakultas Farmasi STIKes Widya Dharma Husada Tangerang. Penulis menyelesaikan Pendidikan S1 di Jenderal Achmad Yani (Unjani), Cimahi dan melanjutkan Pendidikan Profesi Apoteker di Universitas Indonesia, Depok, serta Magister (S2) dengan Peminatan Farmasi Rumah Sakit di Universitas Pancasila, Jakarta. Selain aktif sebagai Dosen, Penulis juga bekerja sebagai Tenaga Kesehatan, Unit Farmasi di salah satu Rumah Sakit Swasta di wilayah Parung, Bogor. Penulis dapat dihubungi melalui
email: debisusantidosen@wdh.ac.id.