

Bab 8: Menganalisis dan Menginterpretasikan data

Richard C. Dicker

PENDAHULUAN

Investigasi lapangan biasanya dilakukan untuk mengidentifikasi faktor-faktor yang meningkatkan risiko seseorang untuk terserang penyakit atau kondisi kesehatan lainnya. Dalam investigasi lapangan tertentu, mengidentifikasi penyebab KLB sudah cukup. Jika penyebabnya dapat dihilangkan, masalahnya terpecahkan. Dalam investigasi lain, tujuannya adalah untuk mengukur hubungan antara *exposure* (atau karakteristik populasi apa pun) dengan kondisi kesehatan yang dihasilkan untuk memandu intervensi atau meningkatkan pengetahuan. Kedua jenis investigasi lapangan ini membutuhkan metode analitik yang sesuai, tetapi tidak harus kompleks. Bab ini menjelaskan strategi untuk merencanakan analisis, metode untuk melakukan analisis, dan pedoman untuk menginterpretasikan hasil.

MERENCANAKAN ANALISIS

Analisis yang direncanakan dengan cermat dan dilaksanakan dengan hati-hati sama pentingnya bagi investigasi lapangan dan bagi penelitian yang didasarkan pada protokol penelitian. Perencanaan diperlukan untuk memastikan bahwa hipotesis yang sesuai akan dipertimbangkan dan data yang relevan akan dikumpulkan, dicatat, dikelola, dianalisis, dan diinterpretasikan untuk menjawab hipotesis tersebut. Oleh karena itu, waktu yang tepat untuk memutuskan data apa yang akan dikumpulkan dan bagaimana data tersebut akan dianalisis adalah sebelum Anda merancang kuesioner, *bukan* setelah Anda mengumpulkan data.

Rencana analisis adalah dokumen yang memandu bagaimana Anda akan menangani data sejak data mentah diperoleh hingga laporan akhir dibuat. Rencana ini menjelaskan dengan apa Anda akan memulai (sumber data dan kumpulan data), bagaimana Anda akan melihat dan menganalisis data, dan di titik mana Anda harus menyelesaikannya (laporan akhir). Rencana analisis menjabarkan komponen kunci analisis dalam urutan logis dan memberikan panduan untuk diikuti selama kegiatan analisis yang sesungguhnya.

Rencana analisis mencakup sebagian atau sebagian besar konten yang tercantum dalam [Kotak 8.1](#). Beberapa komponen yang disebutkan dalam kotak tersebut lebih mungkin muncul dalam rencana analisis untuk penelitian terencana dengan protokol. Namun, investigasi KLB juga harus menyertakan komponen-komponen penting dari komponen-komponen yang tercantum ke dalam rencana analisis yang lebih singkat atau setidaknya dalam serangkaian kerangka tabel

KOTAK 8.1 KOMPONEN RENCANA ANALISIS

- Daftar pertanyaan penelitian atau hipotesis
- Sumber data
 - Deskripsi populasi atau kelompok (kriteria inklusi atau eksklusi)
 - Sumber data atau kumpulan data, terutama untuk analisis data sekunder atau penyebut populasi
 - Jenis studi
- Bagaimana data akan dimanipulasi
 - Kumpulan data yang akan digunakan atau digabungkan
 - Variabel baru yang akan dibuat
- Variabel kunci (lampirkan kamus data untuk semua variabel)
 - Variabel demografi dan *exposure*
 - Variabel hasil atau titik akhir
 - Variabel stratifikasi (misalnya, kemungkinan *confounder* atau *effect modifier*)
 - Bagaimana variabel akan dianalisis (misalnya, sebagai variabel kontinu atau dikelompokkan dalam kategori)
- Bagaimana mengelola nilai-nilai yang hilang
- Urutan analisis (misalnya, distribusi frekuensi, tabel dua arah, analisis bertingkat, dosis-respons, atau analisis kelompok)
- Ukuran kejadian, asosiasi, uji kemaknaan (*significance test*), atau derajat kepercayaan yang akan digunakan
- Kerangka tabel yang akan digunakan dalam analisis
- Kerangka tabel untuk dimasukkan dalam laporan akhir

- *Pertanyaan penelitian atau hipotesis.* Rencana analisis biasanya dimulai dengan pertanyaan penelitian atau hipotesis yang rencananya akan Anda jawab. Pertanyaan penelitian atau hipotesis yang baik mengarah langsung ke variabel yang perlu dianalisis dan metode analisisnya. Misalnya, pertanyaan, “Apa yang menyebabkan terjadinya KLB gastroenteritis?” mungkin menjadi tujuan yang cocok untuk investigasi lapangan, tetapi itu bukan pertanyaan penelitian yang spesifik. Pertanyaan yang lebih spesifik—misalnya, “Makanan mana yang lebih mungkin dikonsumsi oleh pasien kasus daripada oleh kontrol?”—menunjukkan bahwa variabel kuncinya adalah jenis makanan dan status kasus-kontrol, dan bahwa metode analisisnya akan menjadi tabel dua-kali-dua untuk setiap makanan.
- *Strategi analisis.* Berbagai jenis studi (misalnya, kohort, kasus-kontrol, atau *cross-sectional*) dianalisis dengan ukuran dan metode yang berbeda. Oleh karena itu, strategi analisis harus konsisten dalam memaparkan bagaimana data akan dikumpulkan. Misalnya, data dari studi kohort retrospektif sederhana harus dianalisis dengan menghitung dan membandingkan *attack rate* pada kelompok *exposure*. Data dari studi kasus-kontrol harus dianalisis dengan membandingkan *exposure* antara pasien kasus dan kontrol, dan data harus memperhitungkan *matching* dalam analisis jika memang digunakan dalam desain. Data dari studi *cross*

sectional atau survei mungkin perlu memasukkan bobot atau efek desain dalam analisis. Rencana analisis harus menentukan variabel mana yang paling penting seperti *exposure* dan hasil yang diinginkan, faktor risiko lain yang diketahui, faktor desain studi (misalnya, variabel yang cocok), kemungkinan *confounder*, dan *effect modifier*

- *Kamus data*. Kamus data adalah dokumen yang menyediakan informasi kunci setiap variabel. Biasanya, kamus data mencantumkan nama setiap variabel, deskripsi singkat, jenis variabel (misalnya, numerik, teks, atau tanggal), nilai yang diizinkan, dan komentar tambahan. Kamus data dapat diatur dengan cara yang berbeda, tetapi format tabel dengan satu baris per variabel, dan kolom untuk nama, deskripsi, jenis, nilai sebenarnya, dan komentar akan lebih mudah disusun (lihat contoh pada Tabel 8.1 dari investigasi KLB tularemia orofaringeal [1]). Suplemen kamus data mungkin mencakup salinan kuesioner dengan nama variabel yang ditulis di sebelah setiap pertanyaan.
- *Kenali data Anda*. Rencanakan untuk mengenal data Anda dengan mengkaji (1) frekuensi respons dan statistik deskriptif untuk setiap variabel; (2) nilai minimum, maksimum, dan rata-rata untuk setiap variabel; (3) apakah ada variabel yang memiliki respons yang sama untuk setiap catatan dan (4) apakah ada variabel yang memiliki banyak, atau bahkan semua, nilai yang hilang. Pola-pola ini akan mempengaruhi cara Anda menganalisis variabel-variabel ini atau menghapusnya dari analisis sama sekali.
- *Kerangka tabel*. Langkah selanjutnya dalam mengembangkan rencana analisis adalah merancang kerangka tabel. Sebuah kerangka tabel yang kadang-kadang disebut *dummy table* adalah sebuah tabel (misalnya, distribusi frekuensi atau tabel dua-kali-dua) yang diberi judul dan label lengkap tetapi tidak berisi data. Angka baru diisikan pada saat analisis berlangsung. Kerangka tabel memberikan panduan untuk analisis, sehingga urutannya harus berjalan dalam urutan logis dari yang sederhana (misalnya, epidemiologi deskriptif) ke yang lebih kompleks (misalnya, epidemiologi analitik) (Kotak 8.2). Setiap tabel harus menunjukkan ukuran mana (misalnya, *attack rate*, *risk ratio* [RR] atau *odds ratio* [OR], derajat kepercayaan 95% [CI]) dan statistik (misalnya, chi-kuadrat dan nilai *p*) yang harus disertakan dalam tabel. Lihat *Handout* 8.1 untuk contoh kerangka tabel yang dibuat untuk investigasi lapangan tularemia orofaringeal (1).

Tabel 8.1

Kamus data parsial dari investigasi KLB tularemia orofaringeal
(Desa Sancaktepe, Provinsi Bayburt, Turki, Juli – Agustus 2013)

Nama variabel	Keterangan	Jenis variabel	Coding	Catatan
ID	Nomor identitas peserta	numerik		Diberikan

HH_size	Jumlah orang yang tinggal dalam rumah tangga	numerik		
DOB	Tanggal lahir	Tanggal		hh/bb/tttt
Lab_tularemia	Hasil uji mikroaglutinasi	numerik	1 = positif 2 = negatif	
Age	Usia (tahun)	numerik		
Sex	Jenis kelamin	Teks	L= laki-laki P= perempuan	
Fever	Demam	numerik	1 = ya 2 = tidak	
Chills	Panas dingin	numerik	1 = ya 2 = tidak	
Sore throat	Sakit tenggorokan	numerik	1 = ya 2 = tidak	
Node_swollen	Pembengkakan kelenjar getah bening	numerik	1 = ya 2 = tidak	
Node_where	Tempat pembengkakan kelenjar getah bening	Teks		
Case_susp	Memenuhi definisi kasus suspek	numerik	1 = ya 2 = tidak	Variabel yang dibuat: Pembengkakan kelenjar getah bening di sekitar leher atau telinga, sakit tenggorokan, konjungtivitis, atau ≥ 2 hari demam, menggigil, mialgia, sakit kepala
Case_prob	Memenuhi definisi kasus <i>probable</i>	numerik	1 = ya 2 = tidak	Variabel yang dibuat: Pembengkakan kelenjar getah bening dan (sakit tenggorokan atau demam)
Case_confirm	Memenuhi definisi kasus terkonfirmasi	numerik	1 = ya 2 = tidak	Variabel yang dibuat: Uji laboratorium-positif
Case_probconf	Memenuhi definisi kasus <i>probable</i> atau terkonfirmasi	numerik	1 = ya 2 = tidak	Variabel yang dibuat: Case_prob = 1 atau Case_confirm = 1
R_Tap_H2O	Minum air keran selama Ramadhan	numerik	1 = ya 2 = tidak	

Sumber: Diadaptasi dari [Referensi 1](#).

KOTAK 8.2 URUTAN ANALISIS DATA YANG DISARANKAN

Handout 8.2: Waktu, berdasarkan tanggal timbulnya (*onset*) penyakit (dapat dimasukkan dalam Tabel 1, tetapi untuk KLB, lebih baik ditampilkan sebagai kurva epidemi).

Tabel 1. Gambaran klinis (misalnya, tanda dan gejala, persentase kasus yang dikonfirmasi laboratorium, persentase pasien rawat inap, dan persentase pasien yang meninggal).

Tabel 2. Demografi (misalnya, usia dan jenis kelamin) dan karakteristik kunci lainnya dari peserta studi berdasarkan status kasus-kontrol jika studi berbentuk kasus-kontrol.

Tempat (wilayah geografis tempat tinggal atau kejadian pada Tabel 2 atau di tempat atau peta yang diarsir).

Tabel 3. Tabel utama asosiasi *exposure-outcome*.

Tabel 4. Stratifikasi (Tabel 3 dengan efek terpisah dan penilaian *confounding* dan *effect modification*).

Tabel 5. Penyempurnaan (Tabel 3 dengan, misalnya, dosis respons, masa laten, dan penggunaan definisi kasus yang lebih sensitif atau lebih spesifik).

Tabel 6. Analisis kelompok tertentu.

HANDOUT 8.1

Kerangka Tabel: Hubungan antara air minum dari berbagai sumber dan tularemia orofaringeal (Desa Sancaktepe, Provinsi Bayburt, Turki, Juli – Agustus 2013)

<i>Exposure</i>	Jumlah orang ter-expose				Jumlah orang tidak ter-expose				<i>Risk ratio (95% CI)</i>
	Sakit	Sehat	Total	<i>Attack rate (%)</i>	Sakit	Sehat	Total	<i>Attack rate (%)</i>	
Keran	—	—	—	—	—	—	—	—	— (—)
Sumur	—	—	—	—	—	—	—	—	— (—)
Mata air	—	—	—	—	—	—	—	—	— (—)
Botol	—	—	—	—	—	—	—	—	— (—)
Lainya	—	—	—	—	—	—	—	—	— (—)

Singkatan: CI, *Confidence Interval*.

Sumber: Diadaptasi dari [Referensi 1](#).

MENGANALISIS DATA YANG DIPEROLEH DARI INVESTIGASI LAPANGAN

Dua tabel pertama yang biasanya dihasilkan sebagai bagian dari analisis data investigasi lapangan adalah tabel yang menggambarkan gambaran klinis pasien kasus dan menyajikan epidemiologi deskriptif. Karena epidemiologi deskriptif dibahas dalam [Bab 6](#), bab ini hanya akan membahas epidemiologi analitik yang paling sering digunakan dalam investigasi lapangan.

Handout 8.2 menggambarkan keluaran atau *output* modul Analisis Epi Info 7 (CDC, Atlanta, GA) (2). Ini menunjukkan *output* dari perintah “TABLES” untuk data dari jenis investigasi lapangan. Perhatikan elemen kunci dari *output*: (1) tabel tabulasi silang yang merangkum hasil, (2) estimasi poin dari ukuran asosiasi, (3) 95% CI untuk setiap estimasi poin, dan (4) hasil uji statistik. Masing-masing elemen ini dibahas dalam bagian berikut.

HANDOUT 8.2

Output yang diperoleh dari modul analisis klasik, Epi Info versi 7,
Menggunakan perintah “TABLES”

Tabel Vanilla Sakit

Es krim vanila	Sakit		Total
	Ya	Tidak	
Ya	43	11	54
Baris%	79,63%	20,37%	100,00%
Kolom%	93,48%	37,93%	72,00%
Tidak	3	18	21
Baris%	14,29%	85,71%	100,00%
Kolom%	6,52%	62,07%	28,00%
Total	46	29	75
Baris%	61,33%	38,67%	100,00%
Kolom%	100,00%	100,00%	100,00%

Analisis Tabel Tunggal

	<i>Point estimate</i>	<i>95% Confidence interval</i>	
		Batas bawah	Batas atas
PARAMETER: <i>Odds-based</i>			
<i>Odds ratio (cross product)</i>	23.4545	5.8410	94.1811 (T)
<i>Odds ratio (MLE)</i>	22.1490	5.9280	109.1473 (L)
		5.2153	138.3935 (P)
PARAMETER: Berbasis risiko			
<i>Risk Ratio (RR)</i>	5.5741	1.9383	16,0296 (T)
<i>Risk Difference (RD%)</i>	65.3439	46.9212	83,7666 (T)

(T=Taylor series; MLE= *Maximum Likelihood Estimate*; M=Mid- P; F=Fisher Exact)

Uji Statistik	Chi-kuadrat	1-tailed p	2-tailed p
Chi-kuadrat – tidak dikoreksi	27.2225		0,0000013505
Chi-kuadrat – Mantel-Haenszel	26.8596		0,0000013880
Chi-kuadrat – dikoreksi (Yates)	24.5370		0,0000018982
Mid-p exact		0,0000001349	
Fisher exact		0,0000002597	0,0000002597

Sumber: [Referensi 2](#).

Tabel Dua-kali-Dua

Tabel *dua-kali-dua* dinamakan demikian karena merupakan tabulasi silang dari dua variabel, yaitu *exposure* dan status kesehatan (*outcome*), yang masing-masing memiliki dua kategori, biasanya “ya” dan “tidak” ([Handout 8.3](#)). Tabel dua-kali-dua adalah cara terbaik untuk meringkas data yang mencerminkan hubungan antara *exposure* tertentu (misalnya, konsumsi makanan tertentu) dan status kesehatan yang terkait (misalnya, gastroenteritis). Asosiasi biasanya diukur dengan menghitung ukuran asosiasi (misalnya, *risk ratio* [RR] atau OR) dari data dalam tabel dua-kali-dua (lihat bagian berikut).

HANDOUT 8.3

Kerangka Tabel: Hubungan antara air minum dari berbagai sumber dan tularemia orofaringeal (Desa Sancaktepe, Provinsi Bayburt, Turki, Juli – Agustus 2013)

<i>Exposure</i>	Jumlah orang ter-expose				Jumlah orang tidak ter-expose				<i>Risk ratio (95% CI)</i>
	Sakit	Sehat	Total	<i>Attack rate (%)</i>	Sakit	Sehat	Total	<i>Attack rate (%)</i>	
Keran	—	—	—	—	—	—	—	—	— (_)
Sumur	—	—	—	—	—	—	—	—	— (_)
Mata air	—	—	—	—	—	—	—	—	— (_)
Botol	—	—	—	—	—	—	—	—	— (_)
Lainya	—	—	—	—	—	—	—	—	— (_)

Sumber: Diadaptasi dari [Referensi 1](#). Singkatan: CI, *Confidence Interval*

- Dalam tabel dua-kali-dua yang digunakan untuk epidemiologi lapangan, status penyakit (misalnya, sakit atau sehat, kasus atau kontrol) ditampilkan di bagian atas tabel, dan status *exposure* (misalnya, ter-expose atau tidak) di sepanjang sisinya.
- Bergantung pada *exposure* yang sedang dipelajari, baris dapat diberi label seperti yang ditunjukkan pada [Tabel 8.3](#), atau misalnya dilabeli sebagai ter-

expose dan *tidak ter-expose* atau *pernah* dan *tidak pernah*. Berdasarkan kesepakatan, kelompok yang *ter-expose* ditempatkan di baris atas.

- Bergantung pada penyakit atau status kesehatan yang sedang dipelajari, kolom dapat diberi label seperti yang ditunjukkan pada *Handout 8.3*, atau misalnya, *sakit* dan *sehat*, *kasus* dan *kontrol*, atau *meninggal* dan *hidup*. Berdasarkan kesepakatan, kelompok sakit atau kasus ditempatkan di kolom kiri.
- Perpotongan antara baris dan kolom tempat penghitungan dicatat disebut *sel*. Huruf *a*, *b*, *c*, dan *d* dalam empat sel mengacu pada jumlah orang dengan status penyakit yang ditunjukkan pada judul kolom di atas dan status *exposure* ditunjukkan pada label baris di sebelah kiri. Misalnya, sel *c* berisi jumlah orang yang sakit tetapi tidak terpapar. Total baris diberi label H_1 dan H_0 (atau H_2 [H untuk *horizontal*]) dan kolom diberi label V_1 dan V_0 (atau V_2 [V untuk *vertikal*]). Jumlah orang yang tercakup dalam tabel dua-kali-dua ditulis di sudut kanan bawah dan dilambangkan dengan huruf *T* atau *N*.
- Jika data berasal dari studi kohort, *attack rate* (yaitu, proporsi orang yang jatuh sakit selama periode waktu terkait) terkadang dicantumkan di sebelah kanan total baris. Nilai RR atau OR, CI, atau *p* sering dicantumkan di sebelah kanan atau di bawah tabel.

Contoh tabulasi silang konsumsi air keran (*exposure*) dan status penyakit (*outcome*) dari pemeriksaan tularemia orofaringeal ditampilkan pada [Tabel 8.2 \(1\)](#).

Tabel 8.2

Konsumsi air keran dan risiko tertular tularemia orofaringeal
(Desa Sancaktepe, Turki, Juli–Agustus 2013)

Minum air keran	Sakit	Sehat	Total	<i>Attack rate</i> (%)
Ya	46	127	173	26.6
Tidak	9	76	85	10.6
Total	55	203	258	21.3

Risk ratio = $26,59 / 10,59 = 2,5$; *confidence interval* 95% = (1.3–4.9); chi-kuadrat (tidak dikoreksi) = 8,7 ($p = 0,003$).

Sumber: Diadaptasi dari Referensi 1.

Ukuran Asosiasi

Ukuran *asosiasi* mengukur kekuatan atau besarnya hubungan statistik antara *exposure* dan *outcome*. Ukuran asosiasi kadang-kadang disebut *ukuran efek* karena jika *exposure* berhubungan sebab akibat dengan *outcome* kesehatan, ukuran asosiasi mengukur efek *exposure* terhadap *probabilitas* suatu *outcome* kesehatan akan terjadi.

Ukuran asosiasi yang paling umum digunakan dalam epidemiologi lapangan adalah semua rasio—RR, OR, *prevalence ratio* (PR), dan *prevalence OR* (POR). Rasio ini membandingkan apa yang diamati dengan apa yang diperkirakan—yaitu, jumlah penyakit yang diamati pada orang-orang yang ter-expose dibandingkan dengan yang diperkirakan (*baseline*) di antara orang-orang yang tidak ter-expose. Ukuran-ukuran asosiasi dengan jelas menunjukkan apakah jumlah penyakit di antara kelompok yang ter-expose sama dengan, lebih tinggi dari, atau lebih rendah dari (dan seberapa besar) jumlah penyakit pada kelompok *baseline*.

- Nilai setiap ukuran asosiasi sama dengan 1,0 bila jumlah penyakit antara kelompok yang terpapar dan tidak terpapar sama
- Ukuran memiliki nilai lebih besar dari 1,0 ketika jumlah penyakit lebih besar pada kelompok yang terpapar daripada pada kelompok yang tidak terpapar, konsisten dengan efek berbahaya.
- Ukuran tersebut memiliki nilai kurang dari 1,0 bila jumlah penyakit pada kelompok yang ter-expose lebih sedikit daripada pada kelompok yang tidak ter-expose, seperti ketika *exposure* melindungi terhadap terjadinya suatu penyakit (misalnya, vaksinasi).

Ukuran asosiasi yang berbeda digunakan untuk berbagai macam studi. Ukuran yang paling umum digunakan dalam studi kohort retrospektif investigasi KLB adalah RR, yang merupakan rasio dari *attack rate*. Untuk sebagian besar studi kasus-kontrol, karena *attack rate* tidak dapat dihitung, ukuran pilihannya adalah OR.

Studi atau survei *cross-sectional* biasanya mengukur *prevalence* (kasus yang ada) dan bukan kasus baru dari suatu kondisi kesehatan. Ukuran *prevalence* asosiasi yang memiliki analogi dengan RR dan OR adalah PR dan POR yang digunakan secara luas.

Risk Ratio (Relative Risk)

RR yang merupakan ukuran yang lebih disukai untuk studi kohort, dihitung sebagai *attack rate* (risiko) pada kelompok yang terpapar dibagi dengan *attack rate* (risiko) pada kelompok yang tidak terpapar. Dengan menggunakan rumus pada *Handout 8.3*,

$$RR = \text{risiko}_{\text{ter-expose}} / \text{risiko}_{\text{tidak ter-expose}} = (a/H_1) / (c/H_0)$$

Dari [Tabel 8.2](#), *attack rate* (yaitu, risiko) untuk tertular tularemia orofaringeal pada orang-orang yang minum air keran di jamuan makan adalah 26,6%. *Attack rate* (yaitu, risiko) pada mereka yang tidak minum air keran adalah 10,6%. Jadi, RR dihitung sebagai $0,266 / 0,106 = 2,5$. Artinya, orang yang minum air keran 2,5 kali lebih mungkin untuk jatuh sakit dibandingkan mereka yang tidak minum air keran ([1](#)).

Odds ratio

OR adalah ukuran asosiasi yang digunakan untuk data kasus-kontrol. Secara konseptual, ukuran ini dihitung sebagai *odd* dari *exposure* pada pasien-kasus dibagi dengan peluang *exposure* pada kontrol. Namun, dalam praktiknya, rasio ini dihitung sebagai *cross product ratio*. Menggunakan notasi di *Handout 8.3*,

$$OR = ad/bc$$

Contoh data pada *Handout 8.4* berasal dari studi kasus-kontrol gagal ginjal akut di Panama pada tahun 2006 (3). Karena data berasal dari studi kasus-kontrol, baik *attack rate* (risiko) maupun RR tidak dapat dihitung. OR, yang dihitung sebagai $37 \times 110 / (29 \times 4) = 35,1$, sangat tinggi sehingga menunjukkan hubungan yang kuat antara konsumsi sirup obat batuk cair dan gagal ginjal akut.

HANDOUT 8.4

Konsumsi sirup batuk cair resep berdasarkan jawaban terhadap pertanyaan langsung: Studi Kasus-Kontrol Gagal Ginjal Akut (Panama, 2006)

Pernah mengonsumsi obat batuk cair?	Kasus	Kontrol	Total
Ya	37	29	81
Tidak	4	110	35
Total	41	139	116

Odds ratio = 35,1; selang kepercayaan 95% = (11,6–106,4); chi-kuadrat (tidak dikoreksi)=65,6 (p<0,001).

Sumber: Diadaptasi dari [Referensi 3](#).

Prevalence Ratio (PR) dan Prevalence Odds Ratio (POR)

Studi atau survei *cross sectional survey* biasanya lebih mengukur *prevalence* dibandingkan *incidence* status kesehatan (misalnya, status vaksinasi) atau kondisi (misalnya, hipertensi) di suatu populasi. Ukuran asosiasi *prevalence* memiliki analogi dengan RR dan OR adalah PR untuk RR dan POR untuk OR

PR dihitung sebagai *prevalence* pada kelompok indeks dibagi dengan *prevalence* pada kelompok pembanding. Dengan menggunakan rumus pada [Handout 8.3](#),

$$PR = prevalence_{\text{indeks}} / prevalence_{\text{pembanding}} = (a/H_1)/(c/H_0)$$

POR dihitung seperti menghitung OR.

$$POR = ad/bc$$

Dalam sebuah penelitian tentang seroprevalensi HIV di kalangan pengguna *cocaine* dibanding yang tidak pernah menggunakan *cocaine crack*, 165 dari 780 pengguna saat

ini berstatus HIV-positif (*prevalence* = 21,2%), dibandingkan dengan 40 dari 464 yang tidak pernah menggunakan (*prevalence* = 8,6%) (4). PR dan POR berdekatan (masing-masing adalah 2,5 dan 2,8, secara berurutan), tetapi PR lebih mudah dijelaskan.

Ukuran Dampak Kesehatan-Masyarakat

Ukuran dampak kesehatan-masyarakat menempatkan hubungan *exposure*-penyakit dalam perspektif kesehatan-masyarakat. Ukuran dampak mencerminkan kontribusi nyata dari *exposure* terhadap status kesehatan di suatu populasi. Misalnya, untuk *exposure* yang terkait dengan peningkatan risiko penyakit (misalnya, merokok dan kanker paru-paru), persentase risiko yang dapat *attributable risk percent* mewakili jumlah kanker paru-paru di kalangan perokok yang dianggap berasal dari merokok, yang juga dapat dianggap sebagai pengurangan beban penyakit yang diharapkan jika *exposure* dapat dihilangkan atau tidak pernah ada.

Untuk *exposure* yang terkait dengan penurunan risiko penyakit (misalnya, vaksinasi), *prevented fraction* (yang dapat dicegah) mewakili pengurangan beban penyakit yang diamati yang disebabkan oleh tingkat *exposure* saat ini di populasi. Perhatikan bahwa istilah *attributable risk* tidak lebih dari sekadar asosiasi statistik. Kedua istilah tersebut menyiratkan hubungan sebab-akibat langsung antara *exposure* dan penyakit. Oleh karena itu, langkah-langkah ini harus disajikan hanya setelah kesimpulan kausalitas yang tepat berhasil dibuat.

Attributable Risk Percent

Attributable Risk Percent adalah proporsi kasus pada kelompok yang terexpose yang mungkin disebabkan oleh *exposure*. Ukuran ini mengasumsikan bahwa tingkat risiko pada kelompok yang tidak ter-expose (yang dianggap memiliki risiko dasar atau latar belakang penyakit) juga berlaku untuk kelompok yang ter-expose, sehingga hanya risiko berlebih yang harus dikaitkan dengan *exposure*. *Attributable risk percent* dapat dihitung dengan menggunakan salah satu dari rumus berikut yang ekuivalen secara aljabar:

$$\textit{Attributable Risk Percent} = (\text{risiko}_{\text{ter-expose}} / \text{risiko}_{\text{tidak ter-expose}}) / \text{risiko}_{\text{ter-expose}} = \\ (RR-1)/RR$$

Dalam studi kasus-kontrol, jika OR merupakan estimasi sesuai untuk RR, *attributable risk percent* dapat dihitung dari OR.

$$\textit{Attributable Risk Percent} = \frac{(OR - 1)}{OR}$$

Dalam situasi KLB, *Attributable Risk Percent* dapat digunakan untuk mengukur seberapa banyak beban penyakit yang berasal dari *exposure* tertentu.

Prevented Fraction Pada Kelompok Ter-expose (Efikasi Vaksin)

Prevented fraction pada kelompok yang terpapar dapat dihitung ketika RR atau OR kurang dari 1,0. Ukuran ini mencerminkan kemungkinan proporsi kasus yang dicegah oleh *exposure* yang menguntungkan (misalnya kelambu yang mencegah gigitan nyamuk malam hari dan malaria sebagai akibatnya). Ukuran ini juga dapat dianggap sebagai proporsi kasus baru yang akan terjadi tanpa adanya *exposure* yang menguntungkan. Secara aljabar, *Prevented fraction* di populasi yang ter-expose identik dengan efikasi vaksin.

$$\begin{aligned} \text{Prevented fraction pada kelompok ter-expose} &= \text{efikasi vaksin} \\ &= (\text{risiko}_{\text{tidak ter-expose}} - \text{risiko}_{\text{ter-expose}}) / \text{risiko}_{\text{tidak ter-expose}} = 1 - \text{RR} \end{aligned}$$

[Handout 8.5](#) menampilkan data dari KLB varicella (cacar air) di sebuah sekolah dasar di Nebraska pada tahun 2004 (5). Risiko varicella adalah 13,0% pada anak-anak yang divaksinasi dan 66,7% pada anak-anak yang tidak divaksinasi. Efikasi vaksin berdasarkan data ini dihitung sebagai $(0,667 - 0,130) / 0,667 = 0,805$, atau 80,5%. Efikasi vaksin sebesar 80,5% ini menunjukkan bahwa vaksinasi mencegah sekitar 80% kasus yang seharusnya terjadi pada anak-anak yang divaksinasi jika mereka tidak divaksinasi.

HANDOUT 8.5

Status vaksinasi dan kejadian Varicella: KLB di Sekolah Dasar (Nebraska, 2004)

	Sakit	Sehat	Total	Risiko Varicella
Divaksinasi	15	100	115	13,0%
Tidak divaksinasi	18	9	27	66,7%
Total	33	109	142	23,2%

Risk ratio = $13,0/66,7 = 0,195$; efikasi vaksin = $(66,7-13,0) / 66,7 = 80,5\%$.

Sumber: Diadaptasi dari [Referensi 5](#).

Uji Kemaknaan Statistik (Statistical Significance)

Uji kemaknaan statistik digunakan untuk menentukan seberapa besar kemungkinan hasil yang diamati akan terjadi secara kebetulan saja jika *exposure* tidak terkait dengan status kesehatan. Bagian ini menjelaskan faktor-faktor utama yang perlu dipertimbangkan saat menerapkan uji statistik pada data dari tabel dua-kali-dua.

- Pengujian statistik dimulai dengan asumsi bahwa, pada populasi sumber, *exposure* tidak berhubungan dengan penyakit. Asumsi ini dikenal sebagai hipotesis nol. Hipotesis alternatif, yang akan diadopsi jika hipotesis nol terbukti tidak masuk akal, adalah bahwa *exposure* berhubungan dengan penyakit.
- Selanjutnya, hitung ukuran asosiasi (misalnya, RR atau OR).

- Kemudian, pilih dan hitung uji kemaknaan statistik (misalnya, chi-kuadrat). Epi Info dan program komputer lainnya melakukan pengujian ini secara otomatis. Pengujian ini memberikan kemungkinan menemukan hubungan kuat, atau lebih kuat dari, yang diamati jika hipotesis nol itu benar. Probabilitas ini disebut nilai p .
 - Nilai p kecil berarti tidak akan mungkin mengamati hubungan seperti itu jika hipotesis nol benar. Dengan kata lain, nilai p *kecil* menunjukkan bahwa hipotesis nol tidak masuk akal, mengingat data yang tersedia.
 - Jika nilai p ini lebih kecil dari batas yang ditentukan sebelumnya, yang disebut *alfa* (biasanya 0,05 atau 5%), Anda menolak hipotesis nol dan lebih memilih hipotesis alternatif. Asosiasi tersebut kemudian dikatakan *bermakna secara statistik*.
 - Jika nilai p lebih besar dari *cutoff* (misalnya, nilai $p > 0,06$), jangan tolak hipotesis nol; asosiasi yang terlihat mungkin adalah temuan yang diperoleh secara kebetulan.
- Dalam mencapai keputusan tentang hipotesis nol, Anda mungkin membuat salah satu dari dua jenis kesalahan.
 - Dalam *kesalahan tipe I* (juga disebut *kesalahan alfa*), hipotesis nol ditolak padahal sebenarnya hipotesis tersebut benar.
 - Dalam *kesalahan tipe II* (juga disebut *kesalahan beta*), hipotesis nol tidak ditolak padahal sebenarnya hipotesis itu salah.

Menguji dan Menginterpretasikan Data dalam Tabel Dua-kali-Dua

Untuk data dalam tabel dua-kali-dua, Epi Info melaporkan hasil dari dua pengujian berbeda, yaitu uji *chi-kuadrat* dan uji *Fisher Exact*, yang masing-masing memiliki variasi ([Handout 8.2](#)). Kedua pengujian atau uji ini tidak spesifik untuk ukuran asosiasi tertentu. Uji yang sama dapat digunakan selain pada RR, OR, atau *attributable risk percent*

- *Uji mana yang harus digunakan?*
 - *Jika expected value dalam sel mana pun kurang dari 5. Fisher Exact test* adalah standar yang diterima secara umum ketika *expected value* dalam salah satu sel kurang dari 5. (Ingat: *expected value* untuk setiap sel dapat ditentukan dengan mengalikan total baris dengan total kolom dan membaginya dengan total tabel.)
 - *Jika semua expected value dalam tabel dua-kali-dua adalah 5 atau lebih besar.* Pilih salah satu uji chi-kuadrat. Untungnya, untuk sebagian besar analisis, ketiga rumus chi-kuadrat memberikan nilai p yang cukup mirip untuk membuat keputusan yang sama mengenai hipotesis nol berdasarkan ketiganya. Namun, ketika rumus yang berbeda menunjukkan keputusan yang berbeda (biasanya ketika ketiga nilai p kira-kira 0,05), penilaian epidemiologi diperlukan. Beberapa ahli epidemiologi lapangan lebih memilih rumus *Yates-corrected* karena rumus ini memiliki kemungkinan paling kecil untuk membuat kesalahan tipe I (tetapi memiliki kemungkinan besar membuat kesalahan tipe II).

Ahli-ahli lain mengakui bahwa *Yates-corrected* sering kali berlebihan. Oleh karena itu, mereka lebih memilih formula yang tidak dikoreksi. Ahli epidemiologi yang sering melakukan analisis dengan stratifikasi terbiasa menggunakan rumus Mantel-Haenszel sehingga mereka cenderung menggunakan rumus ini, bahkan untuk tabel sederhana dua-kali-dua.

- *Ukuran asosiasi dibanding uji kemaknaan.*
 - *Ukuran asosiasi.* Ukuran asosiasi (misalnya, RR dan OR) menunjukkan kekuatan hubungan antara *exposure* dan penyakit. Ukuran-ukuran ini biasanya tidak bergantung pada ukuran penelitian dan dapat dianggap sebagai perkiraan terbaik dari tingkat asosiasi yang sebenarnya pada populasi sumber. Namun, ukuran tersebut tidak memberikan indikasi reliabilitasnya (yaitu, seberapa besar keyakinan yang harus dimasukkan ke dalamnya).
 - *Uji kemaknaan.* Sebaliknya, uji kemaknaan memberikan indikasi seberapa besar kemungkinan asosiasi yang diamati adalah hasil kebetulan. Meskipun statistik uji *chi-kuadrat* dipengaruhi baik oleh besaran asosiasi maupun ukuran studi, kontribusi masing-masing tidak menjadi dibedakan. Jadi, ukuran asosiasi dan uji kemaknaan (atau CI; lihat *Confidence Intervals for Measures of Association*) memberikan informasi pelengkap.
- *Menafsirkan hasil uji statistik. Tidak bermakna* bukan berarti tidak ada asosiasi. Ukuran asosiasi (RR atau OR) menunjukkan arah dan kekuatan asosiasi. Uji statistik menunjukkan seberapa besar kemungkinan bahwa asosiasi yang diamati mungkin terjadi secara kebetulan saja. Ketidakmaknaan mungkin mencerminkan tidak ada hubungan di antara sumber populasi, tetapi mungkin juga mencerminkan ukuran studi yang terlalu kecil untuk mendeteksi hubungan yang sebenarnya di sumber populasi.
 - *Peran kemaknaan statistik.* Kemaknaan statistik tidak dengan sendirinya menunjukkan hubungan sebab-akibat. Sebuah asosiasi yang diamati mungkin memang mewakili hubungan sebab akibat, tetapi mungkin juga merupakan hasil dari kebetulan, bias seleksi, bias informasi, *confounding*, atau sumber kesalahan lain dalam desain, pelaksanaan, atau analisis studi. Pengujian statistik hanya berhubungan dengan peran peluang dalam menjelaskan asosiasi yang diamati, dan kemaknaan statistik hanya menunjukkan bahwa peluang adalah penjelasan yang tidak mungkin, meskipun bukan tidak mungkin, dari asosiasi tersebut. Penilaian epidemiologi diperlukan ketika mempertimbangkan kriteria ini dan kriteria lain untuk menyimpulkan penyebab (misalnya, konsistensi temuan dengan penelitian lain, hubungan temporal antara *exposure* dan penyakit, atau kemungkinan biologis).
 - *Implikasi kemaknaan statistik pada kesehatan-masyarakat.* Akhirnya, kemaknaan statistik tidak selalu berarti kemaknaan kesehatan-masyarakat. Pada penelitian besar, hubungan yang lemah dengan sedikit

relevansi untuk kesehatan-masyarakat atau relevansi klinis mungkin tetap bermakna secara statistik. Lebih umum lagi bahwa dalam sebuah penelitian kecil, asosiasi kesehatan-masyarakat atau kepentingan klinis mungkin gagal mencapai kemaknaan statistik.

Confidence Interval (CI) untuk Ukuran Asosiasi

Banyak jurnal kedokteran dan kesehatan-masyarakat saat ini mengharuskan asosiasi dijelaskan dengan ukuran asosiasi dan CI, bukan dengan nilai p atau uji statistik lainnya. Ukuran asosiasi seperti RR atau OR memberikan nilai *point estimate* yang mengukur hubungan antara *exposure* dan status kesehatan. CI memberikan estimasi interval atau rentang nilai yang mengakui ketidakpastian estimasi titik bilangan tunggal, terutama yang didasarkan pada sampel populasi.

Confidence Interval 95%

Ahli statistik mendefinisikan 95% CI sebagai interval yang, dengan pengambilan sampel berulang dari populasi sumber, akan mencakup nilai asosiasi sebenarnya pada 95% sepanjang waktu. Konsep epidemiologi CI 95% adalah bahwa CI 95% mencakup rentang nilai yang konsisten dengan data dalam penelitian (6).

Hubungan Antara Uji Chi-kuadrat dan Confidence Interval

Uji chi-kuadrat dan CI terkait erat. Uji chi-kuadrat menggunakan data yang diamati untuk menentukan probabilitas (nilai p) di bawah hipotesis nol, dan hipotesis nol ditolak jika probabilitasnya lebih kecil dari alfa (misalnya, 0,05). CI menggunakan nilai probabilitas yang telah dipilih sebelumnya, yaitu alfa (misalnya, 0,05), untuk menentukan batas interval ($1 - \alpha = 0,95$), dan hipotesis nol ditolak jika interval tidak menyertakan nilai asosiasi nol. Keduanya menunjukkan ketepatan asosiasi yang diamati dan keduanya dipengaruhi oleh besarnya asosiasi dan besar kecilnya kelompok studi. Meskipun keduanya mengukur presisi, tidak satu pun di antaranya yang membahas validitas (kurangnya bias).

Menafsirkan Confidence Interval

- *Arti Confidence Interval.* CI dapat dianggap sebagai rentang nilai yang konsisten dengan data dalam suatu penelitian. Misalkan sebuah penelitian yang dilakukan secara lokal menghasilkan RR 4,0 untuk hubungan antara penggunaan obat intravena dan penyakit X; 95% CI berkisar antara 3,0 hingga 5,3. Dari penelitian tersebut, estimasi terbaik dari hubungan antara penggunaan obat intravena dan penyakit X di populasi umum adalah 4,0, tetapi datanya konsisten dengan nilai dari 3,0 hingga 5,3. Sebuah studi mengenai asosiasi yang sama dilakukan di tempat lain yang kemudian menghasilkan RR 3,2 atau 5,2 akan dianggap kompatibel. Namun, studi yang menghasilkan RR 1,2 atau 6,2 tidak akan dianggap kompatibel. Sekarang, pertimbangkan studi berbeda yang menghasilkan RR 1,0, CI dari 0,9 hingga 1,1, dan nilai $p = 0,9$. Daripada menafsirkan hasil ini sebagai tidak bermakna dan tidak

informatif, Anda dapat menyimpulkan bahwa *exposure* tidak meningkatkan atau mengurangi risiko penyakit. Pesan itu bisa meyakinkan jika *exposure* itu menjadi perhatian publik yang sedang merasa khawatir. Dengan demikian, nilai-nilai yang termasuk dalam CI dan nilai-nilai yang dikecualikan oleh CI sama-sama memberikan informasi penting.

- **Lebar Confidence Interval.** Lebar CI (yaitu, nilai yang disertakan) mencerminkan presisi yang dapat digunakan oleh penelitian untuk menentukan hubungan. CI yang lebar mencerminkan sejumlah besar variabilitas atau ketidaktepatan. CI yang sempit mencerminkan lebih sedikit variabilitas dan presisi yang lebih tinggi. Biasanya, semakin besar jumlah subjek atau pengamatan dalam suatu penelitian, semakin besar presisi dan semakin sempit CI.
- **Hubungan Confidence Interval dengan hipotesis nol.** Karena CI mencerminkan rentang nilai yang konsisten dengan data dalam penelitian, CI dapat digunakan sebagai pengganti pengujian statistik (yaitu, untuk menentukan apakah data konsisten dengan hipotesis nol). Ingat: hipotesis nol menetapkan bahwa RR atau OR sama dengan 1,0. Oleh karena itu, CI yang mencakup 1,0 kompatibel dengan hipotesis nol. Ini sama dengan menyimpulkan bahwa hipotesis nol tidak dapat ditolak. Sebaliknya, CI yang tidak menyertakan 1,0 menunjukkan bahwa hipotesis nol harus ditolak karena tidak sesuai dengan hasil penelitian. Dengan demikian, CI dapat digunakan sebagai pengganti uji kemaknaan statistik.

Confidence Interval dalam Peristiwa KLB Keracunan Pangan

Dalam peristiwa KLB keracunan pangan, tujuan investigasi adalah untuk mengidentifikasi makanan atau perantara lain yang menyebabkan penyakit. Dalam peristiwa semacam ini, ukuran asosiasi (misalnya, RR atau OR) dihitung untuk mengidentifikasi makanan atau bahan habis pakai lainnya dengan nilai tinggi yang mungkin menyebabkan KLB. Peneliti biasanya tidak peduli jika RR untuk jenis makanan tertentu adalah 5,7 atau 9,3. Hanya saja, RR tinggi dan tidak mungkin disebabkan oleh kebetulan dan, oleh karena itu, item tersebut harus dievaluasi lebih lanjut. Untuk tujuan itu, *point estimate* (RR atau OR) ditambah nilai *p* cukup memadai dan CI tidak diperlukan.

Ringkasan Tabel Exposure

Untuk investigasi lapangan yang ditujukan untuk mengidentifikasi satu atau lebih perantara atau faktor risiko penyakit, pertimbangkan untuk membuat tabel tunggal yang dapat merangkum asosiasi untuk beberapa *exposure* yang menjadi fokus. Untuk investigasi KLB keracunan makanan, tabel biasanya mencakup satu baris untuk setiap jenis makanan dan kolom untuk nama makanan; jumlah orang sakit dan sehat, berdasarkan riwayat konsumsi makanan; *attack rate* khusus makanan (jika studi kohort dilakukan); RR atau OR; nilai chi-kuadrat atau *p*; dan, terkadang, 95% CI. Makanan yang paling mungkin menyebabkan penyakit biasanya memiliki dua karakteristik berikut:

1. Peningkatan RR, OR, atau chi-kuadrat (nilai *p* kecil), mencerminkan perbedaan besar dalam *attack rate* di kalangan mereka yang mengonsumsi makanan tersebut dan mereka yang tidak.
2. Mayoritas orang sakit telah mengonsumsi makanan itu. Oleh karena itu, *exposure* dapat menjelaskan atau menjelaskan sebagian besar, jika tidak semua, kasus yang ada.

Dalam contoh rangkuman [Tabel 8.3](#), air keran memiliki RR tertinggi (dan satu-satunya nilai $p < 0,05$, berdasarkan CI 95% tidak termasuk 1,0) dan mungkin mencakup 46 dari 55 kasus.

Tabel 8.3 *Attack rate* tularemia orofaringeal berdasarkan sumber air (Desa Sancaktepe, Turki, Juli – Agustus 2013).

<i>Exposure</i>	Jumlah orang yang ter- <i>expose</i>			Jumlah orang yang tidak ter- <i>expose</i>			<i>Attack rate</i> (%)	<i>Attack rate</i> (%)	Risk ratio (95% CI)
	Sakit	Sehat	Total	Sakit	Sehat	Total			
Keran	46	127	173	27	9	76	85	11	2.5 (1.3–4.9)
Sumur	2	6	8	25	53	198	250	21	1.2 (0.4–4.0)
Mata air	25	111	136	18	30	92	122	25	0,7 (0,5–1,2)
Botol	5	26	31	16	50	177	227	22	0,7 (0,3–1,7)
Lainnya	2	6	8	25	53	198	250	21	1.2 (0.4–4.0)

Singkatan: CI, *Confidence Interval*/derajat kepercayaan.

Sumber: Diadaptasi dari [Referensi 1](#).

Analisis Bertingkat (Stratifikasi)

Stratifikasi adalah analisis hubungan *exposure* -penyakit dalam dua atau lebih kategori (*strata*) dari variabel ketiga (misalnya, usia). Analisis ini merupakan alat yang berguna untuk menilai apakah ada *confounding* dan, jika ya, mengendalikannya. *Stratifikasi* juga merupakan metode terbaik untuk mengidentifikasi *effect modification*. Baik *confounding* dan *effect modification* dibahas di bagian berikut.

Stratifikasi juga merupakan metode yang efektif untuk memeriksa efek dari dua *exposure* yang berbeda pada suatu penyakit. Misalnya, dalam KLB keracunan makanan, dua makanan mungkin tampak terkait dengan penyakit berdasarkan peningkatan RR atau OR. Mungkin kedua makanan tersebut terkontaminasi atau mengandung bahan terkontaminasi yang sama. Alternatifnya, kedua makanan tersebut mungkin telah dimakan bersama-sama (misalnya, selai kacang dan jeli atau donat dan susu), dengan hanya salah satu yang terkontaminasi dan yang lainnya dianggap bersalah karena asosiasi. *Stratifikasi* adalah salah satu cara untuk memisahkan efek dari kedua makanan tersebut.

1. Peningkatan RR, OR, atau chi-kuadrat (nilai *p* kecil), mencerminkan perbedaan substansial dalam *attack rate* pada mereka yang mengonsumsi makanan itu dan mereka yang tidak.
2. Mayoritas orang sakit telah mengonsumsi makanan itu. Oleh karena itu, *exposure* dapat menjelaskan atau menjelaskan sebagian besar, jika tidak semua, kasus yang ditemui.

Dalam contoh rangkuman [Tabel 8.3](#), air keran memiliki RR tertinggi (dan satu-satunya nilai $p < 0,05$, berdasarkan CI 95% tidak termasuk 1,0) dan mungkin mencakup 46 dari 55 kasus.

Membuat Strata Tabel Dua-kali-Dua

- Pertama, pertimbangkan kategori variabel ketiga.
 - Untuk mengelompokkan berdasarkan jenis kelamin, buat tabel dua-kali-dua untuk laki-laki dan tabel lain untuk perempuan.
 - Untuk membuat stratifikasi berdasarkan usia, tentukan pengelompokan usia, pastikan untuk tidak memiliki usia yang tumpang tindih; lalu buat tabel dua-kali-dua terpisah untuk setiap kelompok usia.
- Sebagai contoh, data pada [Tabel 8.2](#) dikelompokkan menurut jenis kelamin pada [Handout 8.6](#) dan [8.7](#). RR untuk minum air keran dan mengalami tularemia orofaringeal adalah 2,3 pada perempuan dan 3,6 pada laki-laki, tetapi stratifikasi juga memungkinkan Anda untuk melihat bahwa perempuan memiliki risiko lebih tinggi daripada laki-laki, terlepas dari konsumsi air keran.

HANDOUT 8.6 Konsumsi air keran dan risiko terkena tularemia orofaring pada perempuan (Desa Sancaktepe, Turki, Juli – Agustus 2013)

Minum air keran?	Sakit	Sehat	Total	<i>Attack Rate</i> (%)	<i>Risk Ratio</i>
Ya	30	60	90	33.3	2.3
Tidak	7	41	48	14.6	
Total	37	101	138	37.9	

Sumber: Diadaptasi dari [Referensi 1](#).

HANDOUT 8.7 Konsumsi air keran dan risiko terkena tularemia orofaring pada laki-laki (Desa Sancaktepe, Turki, Juli-Agustus 2013)

Minum air keran?	Sakit	Sehat	Total	<i>Attack Rate</i> (%)	<i>Risk Ratio</i>
Ya	16	67	83	19.3	3.6
Tidak	2	35	37	5.4	
Total	18	102	120	15.0	

Sumber: Diadaptasi dari Referensi 1.

Tabel Dua-kali-Empat

Tabel bertingkat/terstratifikasi (misalnya, [Handout 8.6](#) dan [8.7](#)) berguna bila variabel stratifikasi tidak menjadi perhatian utama (yaitu, tidak diperiksa sebagai penyebab KLB). Namun, ketika masing-masing dari dua *exposure* mungkin menjadi penyebabnya, tabel dua kali empat lebih baik untuk menguraikan efek dari dua variabel. Pertimbangkan studi kasus-kontrol untuk KLB hepatitis A hipotetis yang menghasilkan OR yang tinggi pada donat (OR = 6.0) dan susu (OR = 3.9). Data yang disusun dalam tabel dua kali empat ([Handout 8.8](#)) menguraikan efek dari kedua makanan tersebut, yaitu hanya *exposure* donat saja yang sangat terkait dengan penyakit (OR = 6.0), tetapi *exposure* susu saja tidak terkait (OR = 1,0).

HANDOUT 8.8 Tampilan tabel dua kali empat hubungan antara hepatitis a dan konsumsi donat dan susu: studi kasus-kontrol dari hipotetis KLB

Donat	Susu	Kasus	Kontrol	Odds Ratio
Ya	Ya	36	18	6.0
Tidak	Ya	1	3	1.0
Ya	Tidak	4	2	6.0
Tidak	Tidak	9	27	1.0 (Ref.)
Total		50	50	

Crude odds ratio untuk donat = 6,0; *Crude odds ratio* untuk susu = 3,9.

Ketika dua makanan menyebabkan penyakit, misalnya ketika keduanya terkontaminasi atau memiliki bahan yang sama, tabel dua kali empat adalah cara terbaik untuk melihat efek individu dan efek gabungan.

Confounding

Confounding adalah distorsi asosiasi *exposure-outcome* oleh efek dari faktor ketiga (*confounder*). Faktor ketiga mungkin menjadi *confounder* jika faktor tersebut:

- Terkait dengan hasil yang tidak bergantung pada *exposure*, yaitu harus merupakan faktor risiko yang independen; dan,
- Terkait dengan *exposure* tetapi bukan merupakan konsekuensinya.

Pertimbangkan studi kohort retrospektif hipotetis tentang kematian di kalangan pekerja manufaktur yang menemukan bahwa pekerja yang terlibat dalam proses manufaktur lebih mungkin meninggal dunia selama periode *follow up* daripada pekerja kantor dan wiraniaga di industri yang sama.

- Peningkatan jumlah kematian secara otomatis dapat dikaitkan dengan satu atau lebih *exposure* selama proses manufaktur.
- Namun, jika usia rata-rata pekerja manufaktur ternyata 15 tahun lebih tua dari pekerja lain, kematian secara wajar dapat diharapkan atau diperkirakan lebih tinggi pada pekerja yang lebih tua.
- Dalam situasi itu, usia kemungkinan merupakan faktor *confounder* yang dapat menjelaskan setidaknya sebagian dari peningkatan kematian. (Perhatikan

bahwa usia memenuhi dua kriteria yang dijelaskan sebelumnya: bertambahnya usia dikaitkan dengan peningkatan kematian, terlepas dari pekerjaan orang tersebut, dan, dalam industri tersebut, usia dikaitkan dengan pekerjaan, atau secara khusus ditemukan bahwa pekerja manufaktur berusia lebih tua daripada pekerja kantoran).

Sayangnya, keberadaan *confounder* merupakan hal umum terjadi. Langkah pertama untuk menangani *confounding* adalah dengan mencarinya. Jika *confounding* telah diidentifikasi, langkah kedua adalah mengontrol atau menyesuaikan efek distorsinya dengan menggunakan metode statistik yang tersedia.

Mencari Confounding

Metode yang paling umum digunakan untuk mencari *confounding* adalah dengan melakukan stratifikasi hubungan *exposure-outcome* yang menjadi fokus studi dengan variabel ketiga yang diduga sebagai *confounder*.

- Untuk membuat stratifikasi (lihat bagian sebelumnya), buat tabel dua-kali-dua terpisah untuk setiap kategori atau strata yang diduga merupakan *confounder* dan pertimbangkan hal-hal berikut saat menilai *confounder* yang dicurigai:
 - Karena salah satu dari dua kriteria untuk variabel *confounding* adalah bahwa variabel tersebut harus dikaitkan dengan hasil akhir (*outcome*), daftar *confounder* potensial harus mencakup faktor risiko penyakit yang diketahui. Daftar juga harus menyertakan variabel yang cocok. Karena usia sering menjadi faktor *confounder*, usia harus dianggap memiliki potensi *confounder* dalam kumpulan data apa pun.
- Untuk setiap strata, hitung ukuran asosiasi khusus strata. Jika variabel stratifikasi adalah jenis kelamin, hanya perempuan yang akan berada di satu strata dan hanya laki-laki di strata yang lain. Asosiasi *exposure-outcome* dihitung secara terpisah untuk perempuan dan laki-laki. Jenis kelamin tidak bisa lagi menjadi *confounder* dalam strata ini karena perempuan dibandingkan dengan perempuan dan laki-laki dibandingkan dengan laki-laki.
- Untuk mencari *confounding*, pertama-tama periksa nilai terkecil dan terbesar dari ukuran asosiasi khusus strata dan bandingkan dengan nilai tabel gabungan (disebut *crude value*). *Confounding* hadir jika nilai kasar berada di luar kisaran antara nilai spesifik strata terkecil dan terbesar.
- Sering kali, *confounding* tidak begitu jelas. Metode yang lebih tepat untuk menilai *confounding* adalah dengan menghitung rangkuman ukuran asosiasi yang disesuaikan sebagai rata-rata tertimbang dari nilai-nilai spesifik strata (lihat bagian berikut, Pengendalian untuk *Confounding*). Setelah menghitung nilai rangkuman, bandingkan nilai rangkuman dengan *crude value* untuk melihat apakah keduanya cukup berbeda. Sayangnya, tidak ada aturan universal atau uji statistik untuk menentukan apa yang merupakan "sangat berbeda." Dalam praktiknya, asumsikan bahwa nilai rangkuman yang disesuaikan lebih akurat. Pertanyaannya kemudian menjadi, "Apakah *crude*

value-nya cukup mendekati nilai yang disesuaikan, atau apakah *crude value*-nya akan menyesatkan pembaca?" Jika *crude value* dan nilai yang disesuaikan mendekati, gunakan *crude value* karena tidak menyesatkan dan lebih mudah dijelaskan. Jika kedua nilai tersebut cukup berbeda (beberapa ahli epidemiologi menggunakan perbedaan 10%, yang lain menggunakan 20%), gunakan nilai yang disesuaikan ([Kotak 8.3](#)).

KOTAK 8.3
METODE UNTUK MENENTUKAN APAKAH ADA CONFOUNDING

- Jika *crude risk ratio* atau *odds ratio* berada di luar kisaran strata-spesifik.
- Jika *crude risk ratio* atau *odds ratio* berbeda >10% atau >20% dari rasio yang disesuaikan Mantel-Haenszel.

Mengendalikan Confounding

- Salah satu metode untuk mengendalikan *confounding* adalah dengan menghitung rangkuman RR atau OR berdasarkan rata-rata tertimbang dari data spesifik strata. Teknik Mantel-Haenszel ([6](#)) adalah metode yang populer untuk melakukannya.
- Metode kedua adalah dengan menggunakan model regresi logistik yang mencakup *exposure* yang diteliti dan satu atau lebih variabel *confounding*. Model menghasilkan perkiraan OR yang mengontrol pengaruh variabel *confounding*.

Effect Modification

Effect modification atau *effect measure modification* berarti bahwa tingkat hubungan antara *exposure* dan *outcome* berbeda pada kelompok populasi yang berbeda. Misalnya, vaksin campak biasanya sangat efektif untuk mencegah penyakit jika diberikan pada anak berusia 12 bulan atau lebih, tetapi kurang efektif jika diberikan sebelum usia 12 bulan. Demikian pula, tetrasiklin dapat menyebabkan bintik-bintik pada gigi pada anak-anak, tetapi tidak pada orang dewasa. Dalam kedua contoh tersebut, asosiasi (atau efek) *exposure* (vaksin campak atau tetrasiklin) adalah fungsi dari, atau dimodifikasi oleh, variabel ketiga (usia dalam kedua contoh tersebut).

Karena *effect modification* berarti efek yang berbeda pada kelompok yang berbeda, langkah pertama dalam mencari *effect modification* adalah membuat stratifikasi asosiasi *exposure-outcome* yang menjadi fokus oleh variabel ketiga yang diduga sebagai *effect modifier*. Selanjutnya, hitung ukuran asosiasi (misalnya, RR atau OR) untuk setiap strata. Terakhir, nilai apakah ukuran asosiasi khusus strata sangat berbeda atau tidak dengan menggunakan salah satu dari dua metode berikut.

1. Periksa ukuran asosiasi khusus strata. Apakah ukuran tersebut cukup berbeda untuk kesehatan-masyarakat atau kepentingan ilmiah?
2. Tentukan apakah variasi dalam besaran asosiasi bermakna secara statistik dengan menggunakan Uji *Breslow-Day* untuk homogenitas *odds ratio* atau dengan menguji istilah interaksi dalam regresi logistik.

Jika ada *effect modification*, sajikan setiap hasil spesifik strata secara terpisah.

Dosis-Respons

Dalam epidemiologi, dosis-respons berarti peningkatan risiko untuk status kesehatan yang mengalami peningkatan (atau, untuk *exposure* pelindung, penurunan) jumlah *exposure*. Jumlah *exposure* mencerminkan kuantitas *exposure* (misalnya, miligram asam folat atau jumlah sendok es krim yang dikonsumsi), atau durasi *exposure* (misalnya, jumlah bulan atau tahun *exposure*), atau keduanya.

Adanya efek dosis-respons adalah salah satu kriteria yang telah diakui untuk menyimpulkan sebab-akibat. Oleh karena itu, ketika hubungan antara *exposure* dan status kesehatan telah diidentifikasi berdasarkan peningkatan RR atau OR, pertimbangkan untuk menilai efek dosis-respons.

Seperti biasa, langkah pertama adalah mengatur data. Salah satu format yang mudah adalah tabel 2-kali-H, dengan H mewakili kategori atau dosis *exposure*. RR untuk studi kohort atau OR untuk studi kasus-kontrol dapat dihitung untuk setiap dosis relatif terhadap dosis terendah atau kelompok yang tidak ter-*expose* ([Handout 8.9](#)). CI dapat dihitung untuk setiap dosis. Pengkajian data dan ukuran asosiasi dalam format ini dan ditampilkannya ukuran secara grafis dapat memberikan gambaran apakah ada asosiasi dosis-respons. Selain itu, teknik statistik dapat digunakan untuk menilai asosiasi tersebut, bahkan ketika *confounder* harus dipertimbangkan.

HANDOUT 8.9 Tataletak data dan notasi untuk tabel dosis respons

TATA LETAK DATA DAN NOTASI UNTUK TABEL DOSIS-RESPONS

Dosis	Sakit atau kasus	Sehat atau kontrol	Total	Risiko	<i>Risk ratio</i>	<i>Odds ratio</i>
Dosis 3	a_3	b_3	H_3	a_3/H_3	$\text{Risiko}_3/\text{Risiko}_0$	a_3d/b_3c
Dosis 2	a_2	b_2	H_2	a_2/H_2	$\text{Risiko}_2/\text{Risiko}_0$	a_2d/b_2c
Dosis 1	a_1	b_1	H_1	a_1/H_1	$\text{Risiko}_1/\text{Risiko}_0$	a_1d/b_1c
Dosis 0	c	d	H_0	c/H_0	1.0 (Ref.)	1.0 (Ref.)
Total	V_1	V_0				

Matching

Tata letak data dasar untuk analisis *matched-pair* adalah tabel dua-kali-dua yang tampak menyerupai tabel dua-kali-dua *unmatched* yang disajikan sebelumnya dalam bab ini, tetapi sebenarnya berbeda ([Handout 8.10](#)). Dalam tabel dua-kali-dua *matched pair*, setiap sel mewakili jumlah pasangan yang cocok yang memenuhi kriteria baris dan kolom. Dalam tabel dua-kali-dua *unmatched*, setiap sel mewakili jumlah orang yang memenuhi kriteria.

HANDOUT 8.10 Tataletak data dan notasi untuk tabel dua-kali-dua berpasangan (*matched pair*)

Kasus	Kontrol	Kontrol	Total
	Ter-expose	Tidak Ter-expose	
Ter-expose	e	f	e + f
Tidak ter-expose	g	h	g + h
Total	e + g	f + h	pasangan e + f + g + h

Odds ratio = f/g .

Dalam [Handout 8.10](#), sel *e* berisi jumlah pasangan di mana kasus-pasien ter-expose dan kontrol ter-expose. Sel *f* berisi jumlah pasangan dengan pasien kasus yang ter-expose dan kontrol yang tidak ter-expose sedangkan sel *g* berisi jumlah pasangan dengan pasien kasus yang tidak ter-expose dan kontrol yang ter-expose. Sel *h* berisi jumlah pasangan kasus-pasien maupun kontrol yang cocok tidak ter-expose. Sel *e* dan *h* disebut *pasangan konkordan* karena pasien kasus dan kontrol berada dalam kategori *exposure* yang sama. Sel *f* dan *g* disebut *pasangan diskordan*.

Dalam analisis pasangan yang cocok, hanya pasangan diskordan yang digunakan untuk menghitung OR. OR dihitung sebagai rasio pasangan diskordan.

OR = f/g

Uji kemaknaan untuk analisis pasangan yang cocok adalah uji McNemar chi-kuadrat.

[Handout 8.11](#) menampilkan data dari studi kasus-kontrol berpasangan yang dilakukan pada tahun 1980 untuk menilai hubungan antara penggunaan tampon dan sindrom syok toksik (7).

HANDOUT 8.11 Penggunaan tampon berkelanjutan selama indeks periode menstruasi: Pusat Pengendalian Penyakit *Toxic Shock Syndrome (Matched-pair)* studi kasus-kontrol, 1980

Kasus	Kontrol Ter-expose	Kontrol Tidak Ter-expose	Total
Ter-expose	33	9	42

Tidak ter-expose	1	1	2
Total	34	10	44 pasang

Odds ratio = $9/1 = 9,0$; uji chi-kuadrat McNemar yang tidak dikoreksi = 6,40 ($p = 0,01$).

Sumber: Diadaptasi dari [Referensi 7](#).

- *Rangkaian pencocokan yang lebih besar dan pencocokan bervariasi*. Dalam studi tertentu, dua, tiga, empat, atau sejumlah variabel kontrol dicocokkan dengan pasien kasus. Cara terbaik untuk menganalisis rangkaian pencocokan yang bervariasi dan yang lebih besar ini adalah dengan mempertimbangkan setiap rangkaian (misalnya, triplet atau kuadruplet) sebagai strata unik dan kemudian menganalisis data dengan menggunakan metode Mantel-Haenszel atau regresi logistik untuk merangkum strata (lihat Mengendalikan *confounding*).
- *Apakah matched design memerlukan matched analysis?* Biasanya, memang diperlukan *matched analysis*. Dalam studi berpasangan, jika pasangan itu unik (misalnya, saudara kandung atau teman), analisis berpasangan diperlukan. Jika pasangan didasarkan pada karakteristik yang tidak unik (misalnya, jenis kelamin atau kelas di sekolah), semua pasien kasus dan semua kontrol dari strata yang sama (jenis kelamin atau kelas) dapat dikelompokkan bersama, dan analisis bertingkat dapat dilakukan.

Pada praktiknya, beberapa ahli epidemiologi melakukan *matched analysis* tetapi kemudian melakukan *unmatched analysis* pada data yang sama. Jika hasilnya serupa, mereka mungkin memilih untuk menyajikan data dengan cara *unmatched*. Dalam kebanyakan kasus, OR yang *unmatched* akan lebih dekat ke 1,0 daripada OR yang *matched* (bias terhadap nol). Bias yang terkait dengan *confounding* ini mungkin sepele atau substansial. Hasil uji chi-kuadrat dari data yang *unmatched* bisa sangat menyesatkan karena biasanya lebih besar daripada hasil uji McNemar untuk data *matched*. Keputusan untuk menggunakan analisis *matched* atau analisis *unmatched* dianalogikan dengan keputusan untuk menyajikan hasil kasar (*crude*) atau *adjusted*. Penilaian epidemiologi harus digunakan untuk menghindari penyajian hasil yang tidak sesuai yang menyesatkan.

Regresi logistik

Dalam beberapa tahun terakhir, regresi logistik telah menjadi alat standar dalam perangkat epidemiologi lapangan. Hal ini disebabkan karena tersedianya perangkat lunak yang mudah digunakan untuk analisis ini telah tersedia secara luas, serta telah terbukti kemampuannya untuk menilai efek dari banyak variabel. *Regresi logistik* adalah metode pemodelan statistik yang analog dengan regresi linier, tetapi digunakan untuk hasil yang bersifat biner (misalnya, sakit/sehat atau kasus/kontrol). Seperti jenis regresi lainnya, *outcome* (variabel terikat) dalam regresi logistik

dimodelkan sebagai fungsi dari satu atau lebih variabel bebas. Variabel bebas tersebut meliputi *exposure* yang menjadi fokus dan sering kali, *confounding* dan variabel interaksi

- Paket perangkat lunak menyesuaikan data dengan model logistik dan memberikan *output* dengan koefisien beta untuk setiap variabel bebas
 - Eksponensial dari koefisien beta yang diberikan (e^β) sama dengan OR untuk variabel tersebut sambil mengontrol efek dari semua variabel lain dalam model.
 - Jika model hanya mencakup variabel *outcome* dan variabel *exposure* utama yang diberi kode (0,1), e^β seharusnya sama dengan OR yang dapat Anda hitung dari tabel dua-kali-dua.
 - Sebagai contoh, model regresi logistik dari data tularemia orofaringeal dengan air keran sebagai satu-satunya variabel bebas menghasilkan OR 3,06, yaitu nilai yang sama persis dengan dua desimal pada *crude* OR. Demikian pula, model yang memasukkan air keran dan jenis kelamin sebagai variabel bebas menghasilkan OR untuk air keran sebesar 3,24 yang hampir identik dengan OR Mantel-Haenszel untuk air keran yang dikontrol berdasarkan jenis kelamin sebesar 3,26. (Perhatikan bahwa regresi logistik memberikan OR dan bukan RR, sehingga tidak ideal untuk studi kohort epidemiologi lapangan.)
- Regresi logistik juga dapat digunakan untuk menilai asosiasi dosis-respons, *effect modification*, dan asosiasi yang lebih kompleks. Varian regresi logistik yang disebut regresi logistik bersyarat (*conditional logistic regression*) sangat sesuai untuk data berpasangan (*paired*).

Teknik analisis yang kompleks tidak dapat mengoreksi data yang tidak akurat! Teknik analisis seperti yang dijelaskan dalam bab ini hanya akan menghasilkan analisis sebaik data yang digunakan. Teknik analisis, baik yang sederhana, bertingkat, maupun pemodelan, menggunakan informasi yang ada. Teknik-teknik tersebut tidak mengetahui atau menilai apakah kelompok pembanding yang benar telah dipilih, tingkat respons memadai, *exposure* dan *outcome* telah ditentukan secara akurat, atau pengkodean dan entri data telah bebas dari kesalahan. Teknik analisis hanyalah alat. Peneliti yang melakukan analisis bertanggung jawab untuk mengetahui kualitas data dan menafsirkan hasilnya dengan tepat.

MENGINTERPRETASIKAN DATA DARI INVESTIGASI LAPANGAN

Komputer dapat menghitung angka lebih cepat dan akurat daripada yang dapat dilakukan oleh peneliti dengan tangan, tetapi komputer tidak dapat menginterpretasikan hasilnya. Untuk tabel dua-kali-dua, Epi Info menyediakan RR dan OR, tetapi peneliti harus memilih mana yang terbaik berdasarkan jenis penelitian yang dilakukan. Untuk tabel tersebut, RR dan OR mungkin meningkat; nilai p mungkin kurang dari 0,05; dan CI 95% mungkin tidak menyertakan 1.0. Namun, apakah hasil

statistik tersebut menjamin bahwa *exposure* tersebut merupakan penyebab penyakit yang sebenarnya? Belum tentu. Meskipun asosiasi mungkin bersifat sebab akibat (kausal), kekurangan dalam desain studi, pelaksanaan, dan analisis dapat menghasilkan asosiasi yang sebenarnya hanya merupakan *artifact*. Peluang bias seleksi, bias informasi, *confounding*, dan kesalahan peneliti harus dievaluasi sebagai penjelasan yang mungkin untuk hubungan yang diamati. Langkah pertama dalam mengevaluasi apakah hubungan yang tampak nyata dan bersifat sebab akibat tersebut memang benar adalah mengkaji daftar faktor-faktor yang dapat menyebabkan asosiasi palsu, seperti yang tercantum dalam Daftar Periksa Interpretasi Epidemiologi 1 ([Kotak 8.4](#)).

KOTAK 8.4
DAFTAR PERIKSA INTERPRETASI EPIDEMIOLOGI 1

- Peluang (*chance*)
- Bias seleksi
- Bias informasi
- *Confounding*
- Kesalahan peneliti
- Asosiasi yang benar

Daftar tilik Interpretasi Epidemiologi 1

Peluang (chance) adalah salah satu penjelasan yang mungkin untuk hubungan yang diamati antara *exposure* dan *outcome*. Dalam hipotesis nol, Anda berasumsi bahwa populasi penelitian Anda adalah sampel dari populasi sumber di mana *exposure* tersebut tidak terkait dengan penyakit; yaitu, RR dan OR sama dengan 1. Bisakah peningkatan (atau penurunan) OR hanya disebabkan oleh variasi yang disebabkan oleh kebetulan? Peran kesempatan dinilai dengan menggunakan uji kemaknaan (atau, seperti disebutkan sebelumnya, dengan menafsirkan CI). Peluang (*chance*) adalah penjelasan yang *tidak mungkin jika*:

- Nilai p kurang dari alfa (biasanya ditetapkan pada 0,05), atau
- CI untuk RR atau OR tidak mencakup 1.0.

Namun, peluang tidak pernah bisa dikesampingkan sepenuhnya. Bahkan jika nilai p sekecil 0,01, studi yang dilakukan mungkin satu-satunya studi dari 100 studi yang menyatakan bahwa hipotesis nol *benar* dan peluang adalah penjelasannya. Perhatikan bahwa uji kemaknaan hanya mengevaluasi peran peluang (*chance*). Uji tersebut tidak membahas adanya bias seleksi, bias informasi, *confounding*, atau kesalahan peneliti.

Bias seleksi adalah kesalahan sistematis dalam penetapan kelompok studi atau dalam perekrutan peserta penelitian yang menghasilkan perkiraan yang salah tentang efek

exposure terhadap risiko penyakit. Bias seleksi dapat dianggap sebagai masalah yang dihasilkan dari siapa yang masuk ke dalam penelitian atau bagaimana mereka dimasukkan ke dalam penelitian. Bias seleksi dapat timbul dari desain yang salah dari studi kasus-kontrol melalui, misalnya, penggunaan definisi kasus yang terlalu luas (sehingga beberapa orang dalam kelompok kasus tidak benar-benar mengalami penyakit yang sedang dipelajari) atau kelompok kontrol yang tidak tepat, atau ketika kasus tanpa gejala tidak terdeteksi pada kontrol. Pada fase pelaksanaan, bias seleksi dapat terjadi jika orang yang memenuhi syarat dengan *exposure* dan karakteristik penyakit tertentu memilih untuk tidak berpartisipasi atau tidak dapat ditemukan. Misalnya, jika orang sakit dengan *exposure* yang diteliti mengetahui hipotesis penelitian dan lebih bersedia untuk berpartisipasi daripada orang dengan penyakit lainnya, sel *a* dalam tabel dua-kali-dua akan meningkat secara artifisial dibandingkan dengan sel *c*, dan OR juga akan meningkat. Evaluasi terhadap kemungkinan peran bias seleksi memerlukan pemeriksaan terkait bagaimana pasien kasus dan kontrol ditentukan dan direkrut.

Bias informasi adalah kesalahan sistematis dalam pengumpulan data dari atau tentang partisipan studi yang menghasilkan perkiraan yang salah tentang efek *exposure* terhadap risiko penyakit. Bias informasi mungkin timbul dengan memasukkan kata-kata yang kurang tepat atau pemahaman terhadap pertanyaan pada kuesioner yang kurang tepat; ingatan yang buruk; teknik wawancara yang tidak konsisten; atau jika seseorang dengan sengaja memberikan informasi palsu, baik untuk menyembunyikan kebenaran atau, seperti yang umum terjadi pada budaya tertentu, sebagai upaya untuk menyenangkan pewawancara.

Confounding adalah distorsi hubungan *exposure-outcome* oleh efek faktor ketiga, seperti yang dibahas sebelumnya dalam bab ini. Untuk mengevaluasi peran *confounding*, pastikan bahwa *confounder* potensial telah diidentifikasi, dievaluasi, dan dikendalikan sesuai dengan kebutuhan.

Kesalahan investigator dapat terjadi pada setiap langkah investigasi lapangan, termasuk dalam desain, pelaksanaan, analisis, dan interpretasi. Dalam analisis, titik koma yang salah tempat dalam program komputer, transkripsi nilai yang salah, penggunaan rumus yang salah, atau kesalahan dalam membaca hasil dapat menghasilkan asosiasi sesuai dengan kenyataan. Mencegah jenis kesalahan ini membutuhkan pemeriksaan pekerjaan yang ketat dan meminta rekan kerja untuk meninjau ulang pekerjaan dan kesimpulan dengan cermat.

Sebagai penekanan, **sebelum mempertimbangkan apakah suatu asosiasi benar-benar bersifat sebab akibat, pertimbangkan apakah asosiasi tersebut dapat dijelaskan sebagai peluang, bias seleksi, bias informasi, *confounding*, atau kesalahan peneliti.** Sekarang anggaplah bahwa RR atau OR yang meningkat memiliki nilai *p kecil* dan CI sempit yang tidak mencakup 1,0. Maka, kebetulan adalah penjelasan yang tidak

mungkin. Spesifikasi pasien kasus dan kontrol yang masuk akal dan partisipasi yang juga baik akan membuat bias seleksi menjadi penjelasan yang tidak mungkin. Informasi dikumpulkan dengan menggunakan kuesioner standar oleh pewawancara yang berpengalaman dan terlatih. *Confounding* yang berasal dari faktor risiko lain dinilai dan ditentukan untuk tidak ada atau telah dikendalikan. Entri data dan perhitungan diverifikasi. Namun, sebelum menyimpulkan bahwa asosiasi tersebut bersifat sebab akibat, kekuatan asosiasi, hubungan biologis yang masuk akal, konsistensi dengan hasil dari penelitian lain, urutan temporal, dan asosiasi dosis-respons, jika ada, perlu dipertimbangkan ([Kotak 8.5](#)).

KOTAK 8.5
CHEKLIST INTERPRETASI EPIDEMIOLOGI 2

- Kekuatan asosiasi
- Masuk akal secara biologis (*biological plausibility*)
- Konsistensi dengan penelitian-penelitian lain
- *exposure* mendahului penyakit
- Efek dosis-respons

Daftar tilik Interpretasi Epidemiologi 2

Kekuatan asosiasi berarti bahwa asosiasi yang lebih kuat memiliki kredibilitas kausal yang lebih besar dibandingkan dengan asosiasi yang lemah. Jika RR yang sebenarnya adalah 1,0; bias seleksi yang lemah, bias informasi, atau *confounding* dapat menghasilkan RR 1,5, tetapi bias harus sangat dramatis dan diharapkan tampak jelas bagi peneliti untuk menjelaskan RR sebesar 9,0.

Biological plausibility berarti suatu asosiasi memiliki kredibilitas sebab akibat jika asosiasi tersebut konsisten dengan patofisiologi yang diketahui, perantara yang diketahui, riwayat alami dari status kesehatan, model hewan, dan faktor biologis lain yang relevan. Untuk perantara makanan yang terlibat dalam KLB penyakit menular, apakah makanan terlibat dalam KLB sebelumnya, atau, lebih baik lagi, apakah agen telah diidentifikasi dalam makanan? Meskipun beberapa KLB disebabkan oleh patogen, perantara, atau faktor risiko baru atau yang sebelumnya tidak diketahui, sebagian besar disebabkan oleh hal-hal yang telah diketahui sebelumnya. Pertimbangkan konsistensi dengan penelitian lain. Apakah hasilnya konsisten dengan penelitian sebelumnya? Sebuah temuan lebih dapat diterima jika telah direplikasi oleh peneliti berbeda dan menggunakan metode berbeda untuk populasi yang berbeda.

Exposure mendahului penyakit tampak jelas, tetapi dalam studi kohort retrospektif, mendokumentasikan bahwa *exposure* mendahului penyakit bisa jadi sulit. Misalkan, bahwa orang dengan jenis leukemia tertentu lebih mungkin memiliki antibodi terhadap virus tertentu daripada kontrol. Mungkin peneliti akan tergoda untuk

menyimpulkan bahwa virus menyebabkan leukemia, tetapi kehati-hatian diperlukan karena infeksi virus mungkin terjadi setelah timbulnya perubahan leukemia.

Bukti efek dosis-respons menambah bobot pada bukti penyebab. Efek dosis-respons bukanlah fitur yang diperlukan untuk membuat suatu asosiasi menjadi kausal. Beberapa asosiasi kausal mungkin menunjukkan efek ambang, misalnya. Meskipun demikian, efek dosis respons biasanya dianggap menambah kredibilitas asosiasi.

Dalam banyak investigasi lapangan, terdapat kemungkinan bahwa faktor yang bertanggung jawab mungkin tidak memenuhi semua kriteria yang dibahas dalam bab ini. Mungkin tingkat respons kurang dari ideal, agen etiologi tidak dapat diisolasi dari makanan yang terlibat, atau tidak ada dosis respons yang diidentifikasi. Meskipun demikian, jika kesehatan-masyarakat terancam, kegagalan untuk memenuhi setiap kriteria tidak boleh digunakan sebagai alasan untuk tidak bertindak. Seperti yang dikatakan George Comstock, "Seni penalaran epidemiologi adalah menarik kesimpulan yang masuk akal dari data yang tidak sempurna" (8). Bagaimanapun, epidemiologi lapangan adalah alat untuk mengarahkan tindakan kesehatan-masyarakat untuk mempromosikan dan melindungi kesehatan-masyarakat atas dasar ilmu pengetahuan (metode epidemiologi yang baik), penalaran sebab akibat (kausal), dan dosis yang sehat dari akal sehat praktis.

Semua karya ilmiah tidak lengkap, baik itu bersifat observasional atau eksperimental. Semua karya ilmiah dapat diubah atau dimodifikasi sejalan dengan kemajuan ilmu pengetahuan. Namun, kondisi ini tidak kemudian memberi kita kebebasan untuk mengabaikan pengetahuan yang sudah kita miliki, atau menunda tindakan yang tampaknya harus dilakukan pada waktu tertentu (9).

— Sir Austin Bradford Hill (1897–1991),
Ahli Epidemiologi dan Statistik Inggris

REFERENSI

1. Aktas D, Celebi B, Isik ME, dkk. Oropharyngeal tularemia outbreak associated with drinking contaminated tap water, Turkey, July–September 2013. *Emerg Infect Dis.* 2015;21:2194–6.
2. Centers for Disease Control and Prevention. Epi Info. <https://www.cdc.gov/epiinfo/index.html>
3. Rentz ED, Lewis L, Mujica OJ, dkk. Outbreak of acute renal failure in Panama in 2006: a case–control study. *Bull World Health Organ.* 2008;86:749–56.
4. Edlin BR, Irwin KL, Faruque S, dkk. Intersecting epidemics—crack cocaine use and HIV infection among inner-city young adults. *N Eng J Med.* 1994;331:1422–7.
5. Centers for Disease Control and Prevention. Varicella outbreak among vaccinated children—Nebraska, 2004. *MMWR.* 2006;55:749–52.
6. Rothman KJ. *Epidemiology: an introduction.* New York: Oxford University Press; 2002: p. 113–29.

7. Shands KN, Schmid GP, Dan BB, dkk. Toxic-shock syndrome in menstruating women: association with tampon use and *Staphylococcus aureus* and clinical features in 52 women. *N Engl J Med*. 1980;303:1436–42.
8. Comstock GW. Vaccine evaluation by case–control or prospective studies. *Am J Epidemiol*. 1990;131:205–7.
9. Hill AB. The environment and disease: association or causation? *Proc R Soc Med*. 1965;58:295–300.