

# Pengantar Meta- Analisis

---

GITA SEKAR PRIHANTI

ADAPTASI DARI DAHLAN, S. 2012. PENGANTAR META  
ANALISIS

# Tujuan

---

Mampu menjelaskan apa dan mengapa kita memerlukan meta-analisis serta terminologi penting dalam meta-analisis, yaitu *forest plot*, ukuran efek, *fixed effect model*, *random effect model*, pembobotan, heterogenitas, variasi penelitian, variasi antarpencelitian, variasi total, interpretasi interval kepercayaan, interpretasi nilai  $p$ , dan interpretasi klinis

**Tabel 1.1**  
**Rekapitulasi beberapa hasil penelitian**  
**yang membandingkan efektivitas obat A dan Obat B**

No	Peneliti	Tahun	Obat A		Obat B		RR	IK95%	
			n	Sembuh	n	Sembuh		Min	Mak
1	A	1989	20	19	20	16	1,19	0,93	1,51
2	B	1993	40	36	40	33	1,09	0,91	1,30
3	C	1999	80	60	80	65	0,92	0,78	1,09
4	D	2001	200	178	200	183	0,97	0,91	1,04
5	E	2001	160	141	160	142	0,99	0,92	1,07
6	F	2001	300	231	300	206	1,12	1,02	1,24
7	G	2003	100	87	100	71	1,23	1,06	1,42
8	H	2003	250	224	250	218	1,03	0,96	1,09
9	I	2004	50	43	50	47	0,91	0,80	1,04
10	J	2005	50	39	50	41	0,95	0,78	1,16
11	K	2005	55	49	55	49	1,00	0,88	1,14
12	L	2006	300	252	300	248	1,12	0,95	1,09
13	M	2006	20	16	20	17	0,94	0,71	1,25
14	N	2007	350	313	350	285	1,10	1,03	1,17
15	O	2007	20	19	20	17	1,12	0,91	1,38
16	P	2007	250	187	250	199	0,94	0,85	1,03
17	Q	2007	30	25	30	25	1,00	0,80	1,25
18	R	2007	120	95	120	85	1,07	0,93	1,23
19	S	2007	160	135	160	110	1,23	1,08	1,39
20	T	2011	860	806	860	797	1,01	0,99	1,04

RR=resoki relatif; IK=interval kepercayaan; min=minimal; mak=maksimal

Cobalah amati secara seksama hasil penelitian yang disajikan pada tabel 1.1 Tabel ini menyajikan hasil beberapa penelitian yang membandingkan efektifitas obat A dengan obat B. Setelah membaca dengan teliti tabel tersebut, jawablah pertanyaan berikut.

**Tabel 1.2**

**Pertanyaan pendahuluan untuk membahas mengapa kita memerlukan meta-analisis**

No.	Pertanyaan	Jawaban Anda
1.	Dapatkah Anda menyimpulkan mana di antara obat A dan B yang lebih baik?	
2.	Dapatkah Anda menyimpulkan bahwa resiko relatif (RR) antara A dengan B adalah rerata RR dari 20 penelitian?	
3.	Apabila jawaban atas dua pertanyaan sebelumnya adalah 'tidak' metode apa yang digunakan untuk menyimpulkan hasil dari beberapa penelitian tersebut	

Kita tidak dapat menyimpulkan secara pasti obat mana yang lebih baik karena:

1. Dari dua puluh penelitian tersebut, kita harus memilah penelitian mana saja yang valid.
2. Seandainya semua penelitian penelitian valid, kita masih harus memilah penelitian dan kriteria keluaran yang sesuai dengan yang kita inginkan.
3. Seandainya semua penelitian valid, populasi sudah sesuai dengan yang diinginkan, dan kriteria keluaran sama, kita tidak dapat menyimpulkan obat mana yang lebih baik berdasarkan rerata RR.

Metode konvensional untuk memperoleh hasil keseluruhan adalah dengan membuat rata-rata (Jumlah risiko relatif dari setiap penelitian)

---

$$\text{rerata RR} = \frac{R_1 + \dots + R_{20}}{20}$$

20

Akan tetapi, kita akan segera menyadari bahwa kita tidak dapat menghitung rerata RR dengan cara tersebut karena jumlah subjek masing-masing penelitian berbeda-beda. Dengan demikian, kita memerlukan metode untuk menyimpulkan hasil dari beberapa penelitian. Metode ini dinamakan meta-analisis

# Meta – Analisis

Perlu dibedakan dari tinjauan pustaka dan telaah sitematis (table 1.3).

Kita dapat membedakan ketiga hal tersebut berdasarkan metode dan teknik statistic.

1. Yang dimaksudkan metode adalah berbagai komponen yang biasanya kita temukan dalam usulan penelitian, misalnya latar belakang, pertanyaan, tujuan, dan criteria.
2. Teknik statistic adalah jenis analisis yang digunakan untuk menghitung efek gabungan beberapa penelitian secara formal.

Pada tinjauan pustaka umumnya kita tidak menemukan adanya metode,. Sementara itu pada meta-analisis kita menemukan bukan saja metode namun juga teknik statistic. Oleh karena itu mudah dipahami kalau meta-analisis didefinisikan sebagai telaah sistematik yang disertai teknik statistic untuk menghitung kesimpulan beberapa hasil penelitian.

Table 1.3

Perbandingan tinjauan pustaka, telaah sistematis, dan meta-analisis

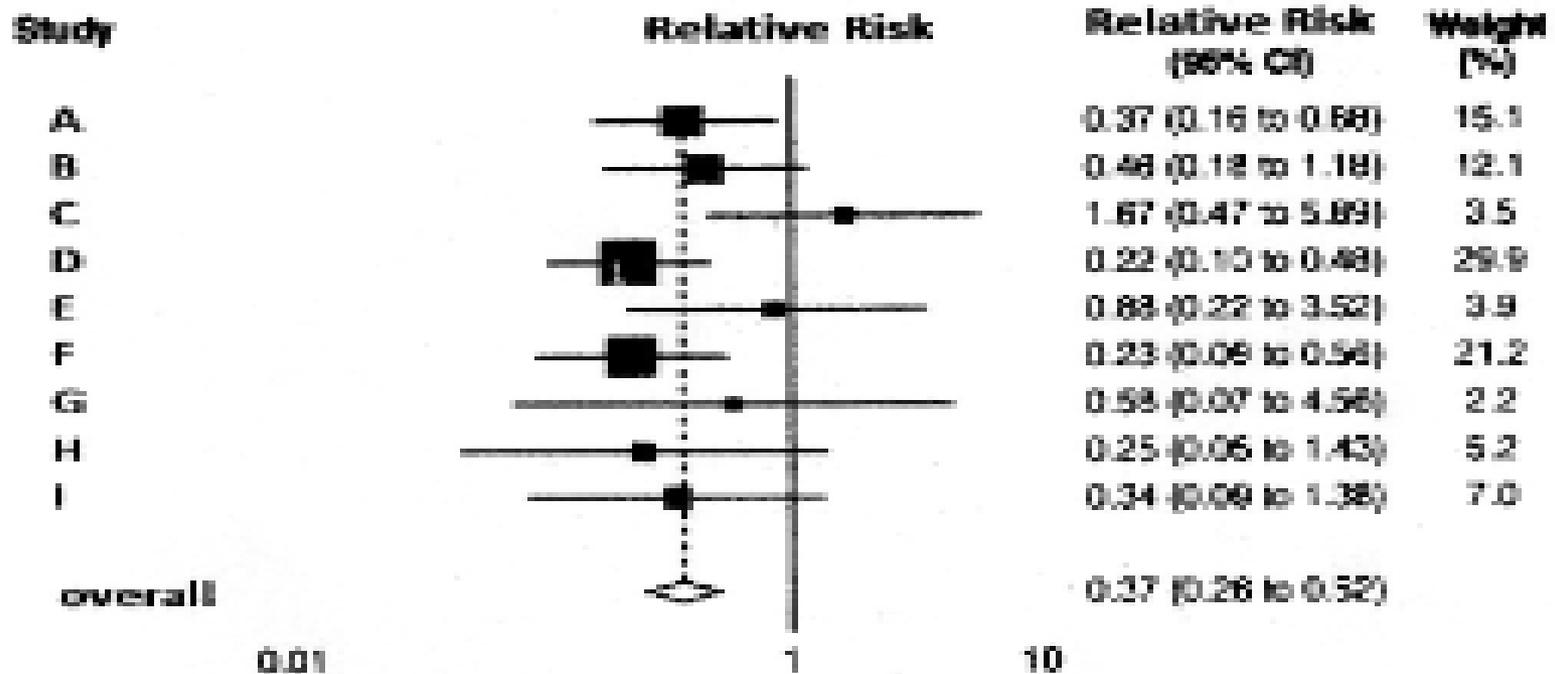
	Metode	Teknik statistik
Tinjauan pustaka	X	x
Telaah sistematis	✓	x
Meta - analisis	✓	✓

X= tidak ada, ✓ = ada

Terminologi  
penting  
Meta – analisis

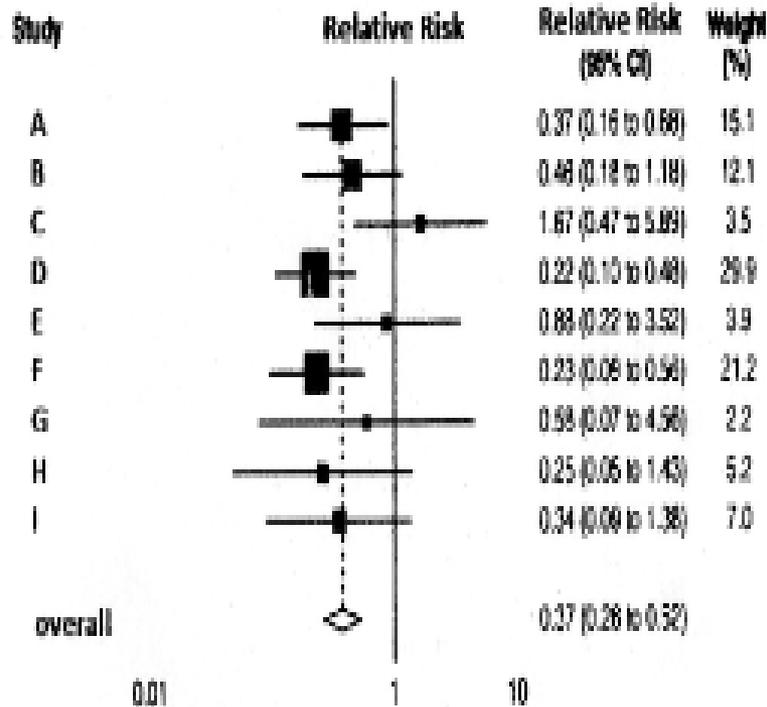
---

# Hasil meta-analisis yang membandingkan efektivitas obat A dengan obat B



Test for heterogeneity  $\chi^2 = 9,7$  (df=8,  $p=0,287$ )  $I^2 = 25,5\%$   
Test for overall effect  $Z=2,55$  ( $p=0,005$ ; fixed effect model)

1. Ukuran efek yang disajikan pada gambar 1.1 adalah resiko relatif (RR). Selain RR, apalagi ukuran efek yang Anda Ketahui?



Test for heterogeneity  $\chi^2=9.7$  (df=8, p=0.287) ( $I^2=25.5\%$ )  
 Test for overall effect  $Z=7.55$  (p=0.005; fixed effect model)

Ukuran efek yang disajikan pada gambar 1.1 adalah resiko relatif (RR).

Risiko Relatif adalah salah satu ukuran efek.

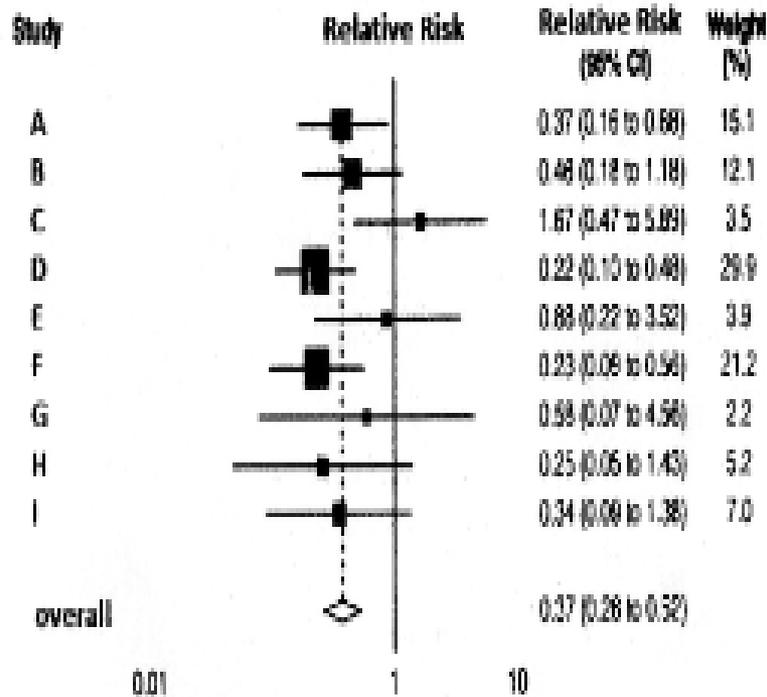
Ukuran efek lainnya antara lain odds rasio (OR), hazard rasio (HR), prevalenrasio (PR), selisih proporsi, selisih rerata, dan korelasi.

Berbagai ukuran efek tersebut dapat diklasifikasikan kedalam ukuran efek yang mempunyai konsep pembagian (RR, OR, HR, PR), konsep pengurangan (selisih rerata, selisih proporsi), dan korelasi.

Berbagai ukuran efek tersebut dapat dijadikan sebagai keluaran pada penelitian meta-analisis.

Ukuran statistic ada yang menganut konsep pembagian (RR, OR, HR) dan ada yang menganut konsep selisih (selisih rerata, selisih proporsi)

## 2. Apa artinya gambar berbentuk segiempat, garis horizontal pada segiempat, ukuran segiempat, gambar berbentuk wajik, dan garis vertical pada *forest plot*?



Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 7.55$  ( $p = 0.005$ ; fixed effect model)

Gambar berbentuk segiempat dan garis horizontalnya adalah visualisasi hasil masing-masing penelitian.

Kotak segiempat adalah nilai RR pada subjek penelitian sementara garis horizontal menunjukkan interval kepercayaan.

Ujung kiri adalah nilai minimal sedangkan ujung kanan adalah nilai maksimal.

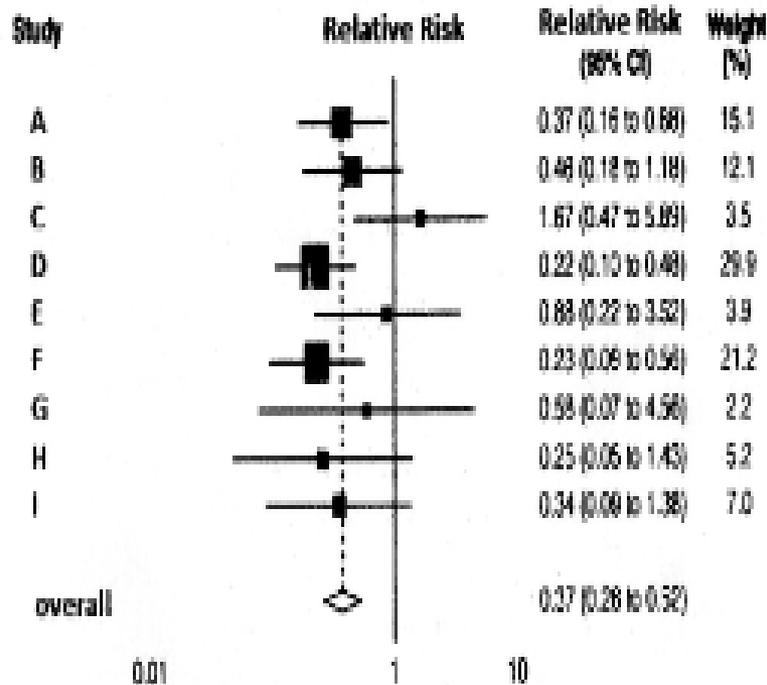
Ukuran segiempat disesuaikan dengan bobot (*weighting*) masing-masing penelitian.

Semakin besar bobot, semakin besar ukuran segiempat.

Gambar berbentuk wajik adalah visualisasi dari hasil keseluruhan. Nilai di tengah wajik adalah nilai pada subjek sedangkan ujung kiri sampai ujung kanan wajik adalah nilai interval kepercayaannya. Ujung kiri adalah nilai minimal sedangkan ujung kanan adalah nilai maksimal.

Garis vertical adalah garis yang menunjukkan tidak ada perbedaan efek. Pada gambar 1.1,

.....2. Apa artinya gambar berbentuk segiempat, garis horizontal pada segiempat, ukuran segiempat, gambar berbentuk wajik, dan garis vertical pada *forest plot*?



Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 2.55$  ( $p = 0.005$ ; fixed effect model)

Garis vertical terdapat pada angka satu  
 Karena ukuran efek yang dilihat adalah RR.

Risiko relative adalah Ukuran statistik yang menganut konsep pembagian.

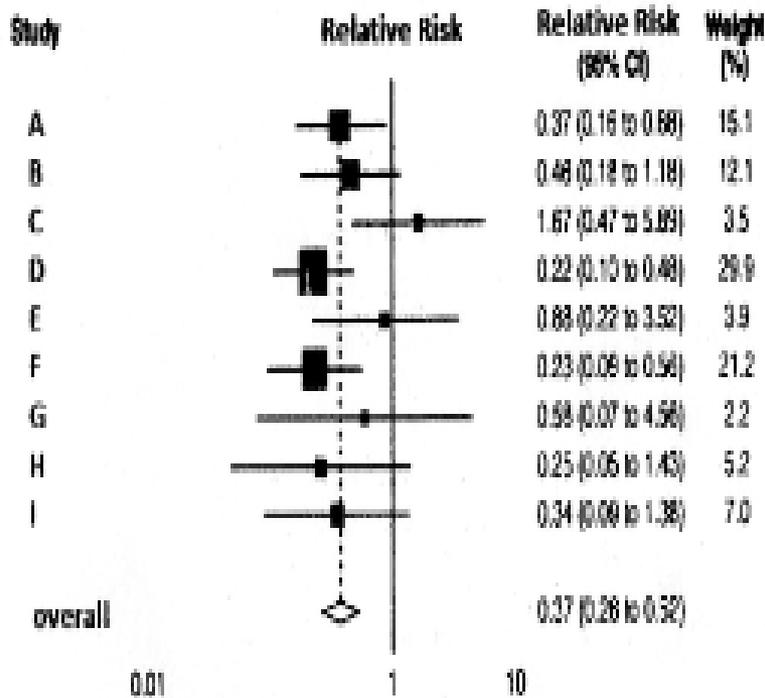
Pada semua ukuran statistik yang menganut konsep pembagian (RR, OR, HR, PR), garis vertikal berada pada angka satu.

Sementara itu, semua ukuran statistik yang menganut konsep pengurangan (selisih proporsi dan selisih rerata) serta korelasi, garis vertikal berada pada angka nol.

Hasil tiap penelitian disajikan dalam bentuk gambar segi empat yang ukurannya sesuai dengan bobot disertai interval kepercayaan berupa garis horizontal.

Efek gabungan disajikan dalam bentuk wajik di mana ujung kanan dan kirinya menggambarkan interval kepercayaan.

### 3. Apakah yang dimaksud dengan *test of heterogeneity*? Bagaimana cara mengetahui dan interpretasi heterogenitas?



Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 7.55$  ( $p = 0.005$ ; fixed effect model)

Heterogenitas adalah variasi.

Variasi total bersumber dari variasi dalam penelitian (misalnya kita sebut X) dan variasi antar penelitian (misalnya kita sebut Y).

Jadi, variasi total adalah variasi penelitian ditambah variasi antar penelitian =  $x+y$ .

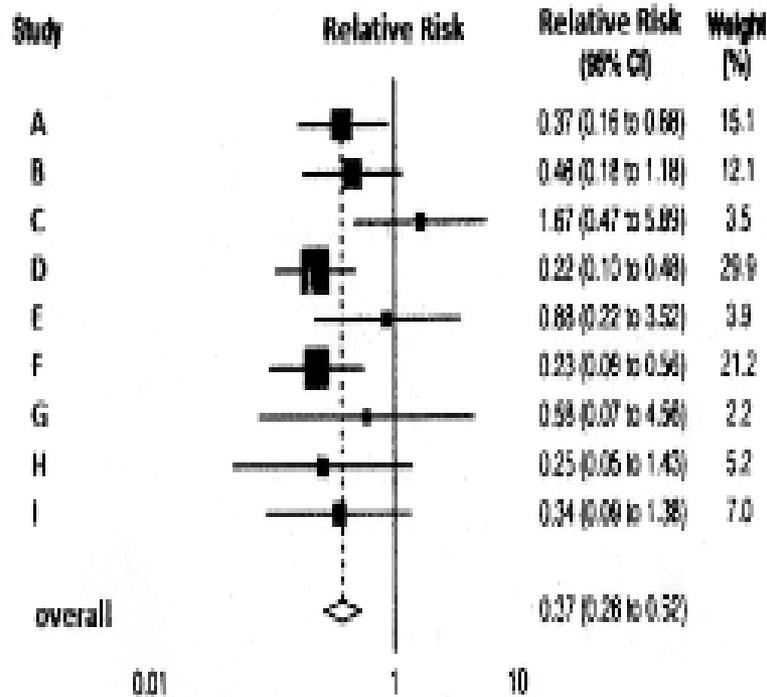
Variasi total = variasi penelitian + variasi antar penelitian

Pada uji heterogenitas, kita menghitung peran variasi antarpemeriksaan (dengan simbol  $I^2$ ) terhadap variasi total.

Secara matematis, peran variasi antarpemeriksaan adalah variasi antarpemeriksaan (y) dibagi variasi total (x+y) dikalikan 100%.

$$I^2 = \frac{y}{x+y} \times 100\%$$

### .....3. Apakah yang dimaksud dengan *test of heterogeneity*? Bagaimana cara mengetahui dan interpretasi heterogenitas?



Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 7.55$  ( $p = 0.005$ ; fixed effect model)

Apabila peran  $\gamma$  kecil, kita bisa mengabaikannya ( $\gamma$  dianggap nol) sehingga nilai  $I^2$  sama dengan nol.

Bila  $\gamma = 0$ , variasi total bisa dianggap sama dengan  $x$ .

Apabila peran  $\gamma$  besar, kita tidak dapat mengabaikannya sehingga variasi total adalah penjumlahan  $x$  dan  $\gamma$ .

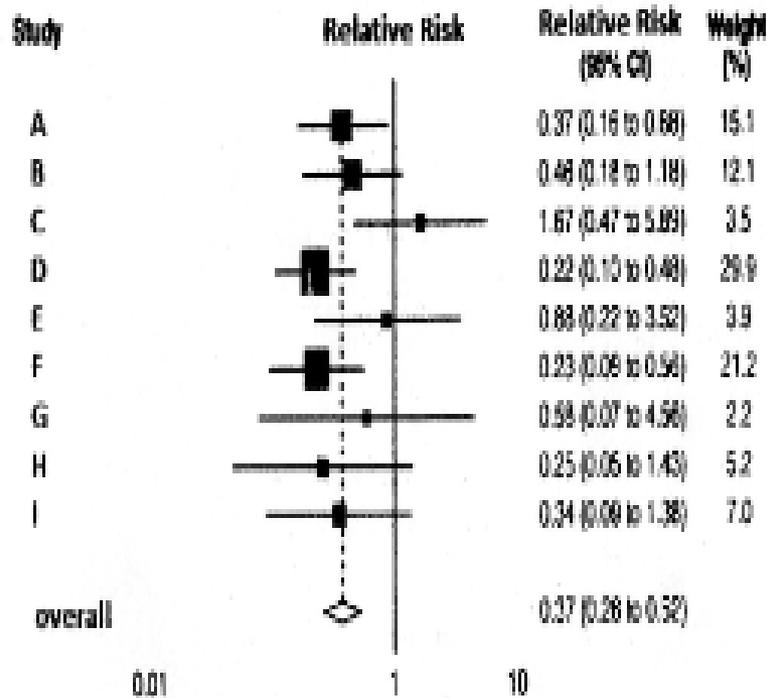
Penentuan besar kecilnya peranan variasi antarpenelitian bergantung pada penelitian masing-masing peneliti.

Masalahnya, berapa persentase  $\gamma$  minimal yang dikatakan besar sangat subjektif.

Pada gambar 1.1 misalnya, peran variasi antarpenelitian ( $I^2$ ) adalah 25,5%.

Kita mungkin kesulitan untuk menentukan apakah angka 25,5% besar atau kecil.

### .....3. Apakah yang dimaksud dengan *test of heterogeneity*? Bagaimana cara mengetahui dan interpretasi heterogenitas?



Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 7.55$  ( $p = 0.005$ ; fixed effect model)

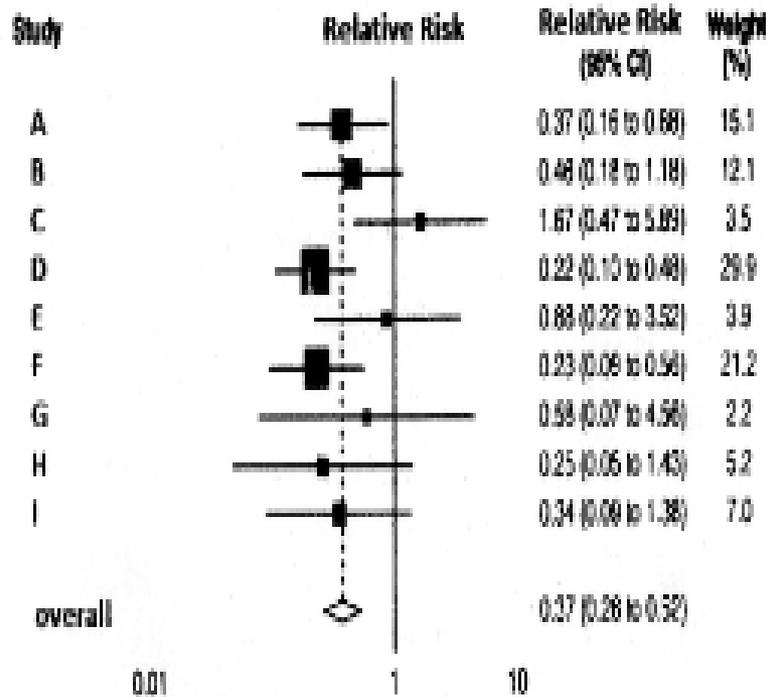
Supaya lebih objektif, penentuan peranan variasi antarpemeriksaan diuji dengan uji heterogenitas.

Hipotesis nol dari uji heterogenitas adalah variasi antarpemeriksaan tidak memberikan peranan pada variasi total.

Dalam bahasa yang lebih populer, hipotesis nol ini adalah antarpemeriksaan tidak bervariasi atau antarpemeriksaan mempunyai variasi yang sama.

Hipotesis nol ditolak apabila nilai  $p$  pada uji heterogenitas kurang dari 0,05..

### .....3. Apakah yang dimaksud dengan *test of heterogeneity*? Bagaimana cara mengetahui dan interpretasi heterogenitas?



Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 7.55$  ( $p = 0.005$ ; fixed effect model)

Pada gambar 1.1, hasil uji heterogenitas adalah 0,287 sehingga hipotesis nol diterima, artinya, variasi antarpelelitian tidak memberikan peranan pada variasi total.

Dengan kata lain, antarpelelitian homogen.

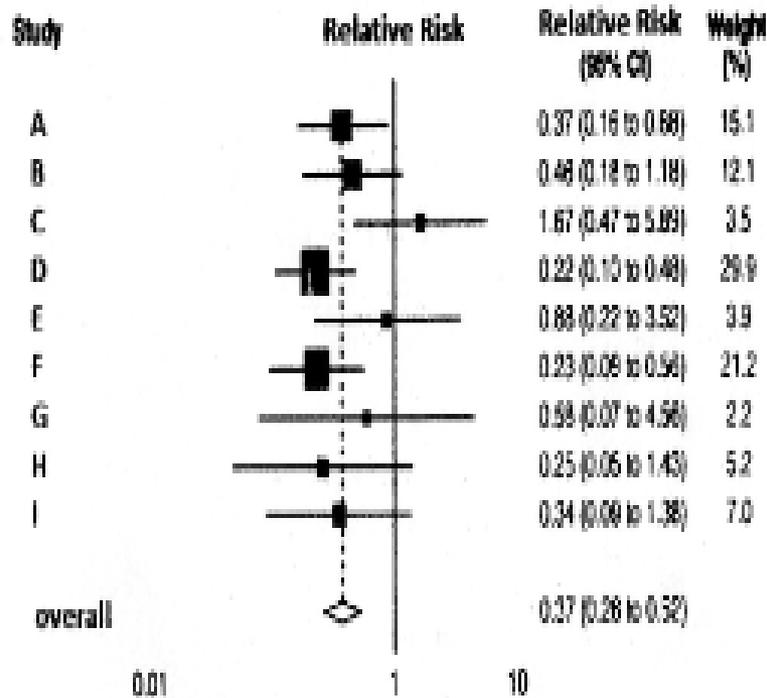
Untuk melihat heterogenitas, kita dapat melihat nilai  $I^2$  atau nilai  $p$  pada uji heterogenitas.

Kesimpulan keduanya pada umumnya sama, apakah homogen atau heterogen.

Dalam kondisi tertentu keputusan berdasarkan  $I^2$  dapat berbeda dengan keputusan uji heterogenitas.

Hal tersebut biasanya terjadi pada meta-analisis dengan jumlah penelitian yang terlalu sedikit atau terlalu banyak.

### .....3. Apakah yang dimaksud dengan *test of heterogeneity*? Bagaimana cara mengetahui dan interpretasi heterogenitas?



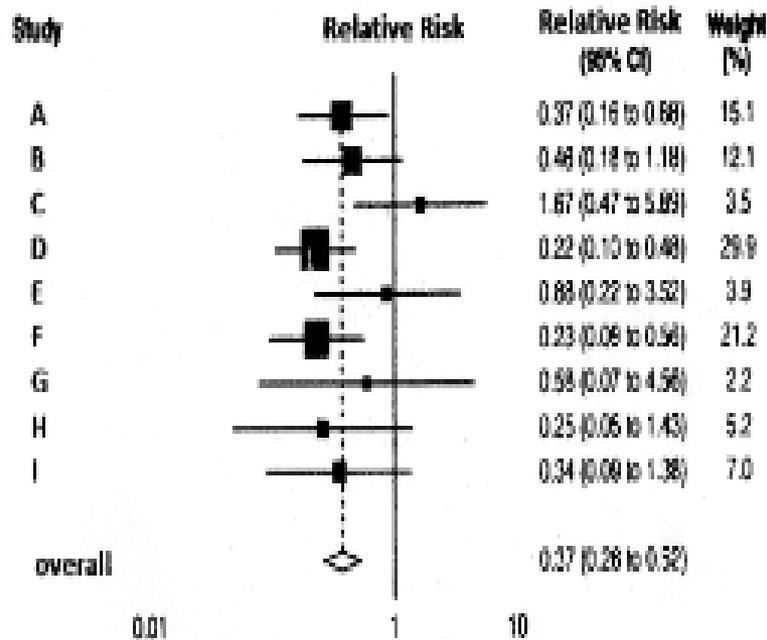
Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 7.55$  ( $p = 0.005$ ; fixed effect model)

Pada penelitian yang terlalu sedikit, dapat terjadi tidak bervariasi berdasarkan nilai p akan tetapi bervariasi berdasar nilai  $I^2$

Sebaliknya pada penelitian yang terlalu banyak, dapat terjadi bervariasi berdasar nilai p akan tetapi tidak bervariasi berdasar nilai  $I^2$

Jika terjadi hal demikian, yang harus “dimenangkan” adalah keputusan berdasar nilai  $I^2$

Walaupun demikian, karena keputusan berdasar nilai  $I^2$  bersifat subjektif, beberapa peneliti lebih menyukai pengambilan keputusan berdasar nilai p.



Test for heterogeneity  $\chi^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 2.55$  ( $p = 0.005$ ; fixed effect model)

Variasi total = variasi dalam penelitian + variasi antar penelitian

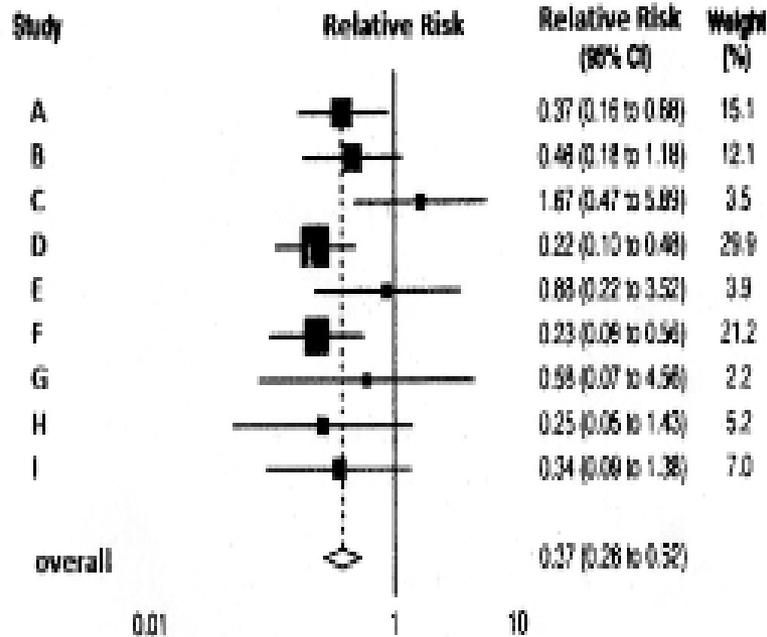
Variasi antar penelitian = variasi total – variasi dalam penelitian

Persen variasi antar penelitian = variasi antar penelitian / variasi total x 100%

Variasi total : variasi dalam penelitian apabila variasi antar penelitian ~ nol

Variasi antar penelitian dianggap nol apabila persen variasinya (12) kecil atau nilai p uji heterogenitas lebih besar dari 0,05

# 4. Apakah yang dimaksud dengan *fixed effect model*?



Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df=8,  $p=0.287$ ) ( $I^2 = 25.5\%$ )  
Test for overall effect  $Z=2.55$  ( $p=0.005$ ; fixed effect model)

Hasil uji heterogenitas akan menentukan model analisis untuk menghitung efek gabungan.

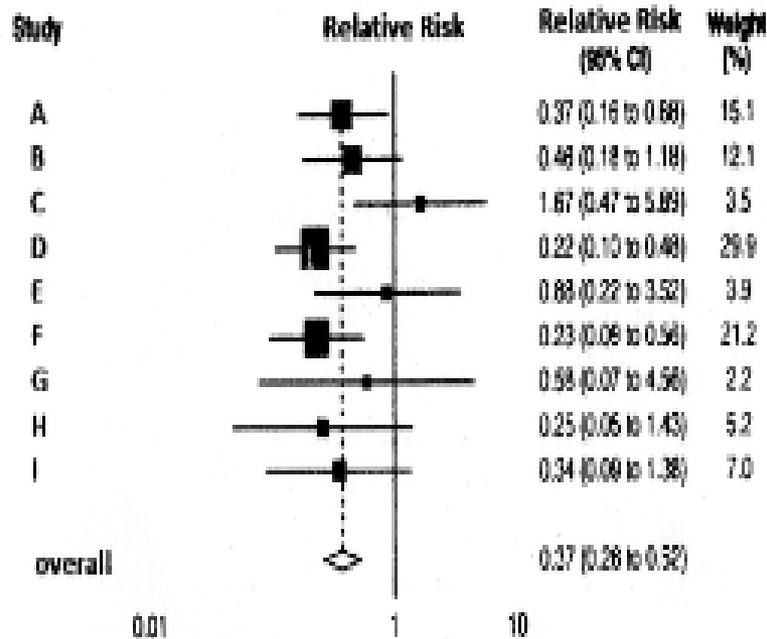
Apabila antar penelitian homogen (nilai  $p$  pada uji heterogenitas lebih besar dari 0,05 atau  $I^2$  kecil), model analisis yang digunakan adalah fixed effect model (FEM).

Sementara itu, apabila antar penelitian heterogen (nilai  $p$  pada uji heterogenitas kurang dari 0,05 atau  $I^2$  besar), model analisis yang digunakan adalah random effects model (REM)

Pada FEM, efek gabungan di hitung dengan mengabaikan variasi antar penelitian.

Sementara itu, pada REM, efek gabungan dihitung dengan melibatkan variasi antar penelitian.

# 5. Apakah yang dimaksud dengan *weight*(bobot)? Bagaimana menentukan dan interpretasi *weight*?



Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
Test for overall effect  $Z = 2.55$  ( $p = 0.005$ ; fixed effect model)

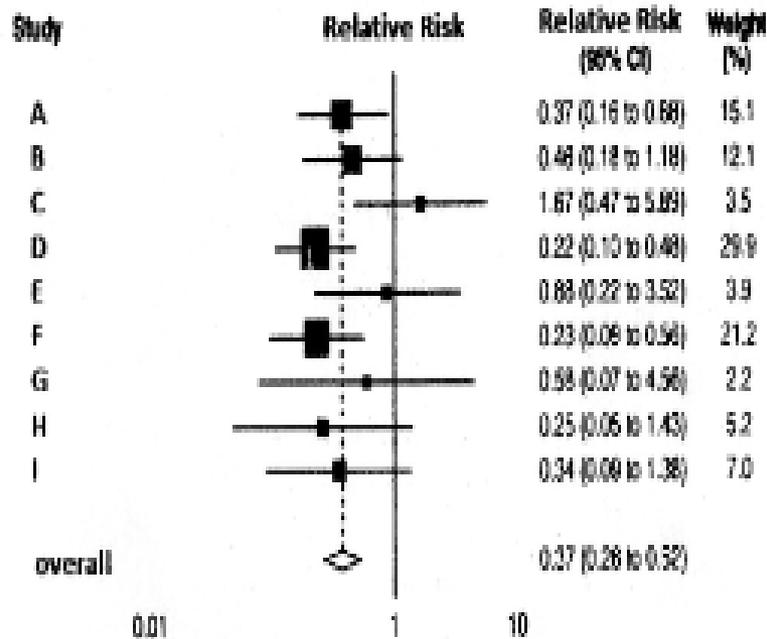
Weight (bobot) adalah peranan suatu hasil penelitian terhadap total efek.

Bobot berbanding lurus dengan jumlah subjek penelitian.

Penelitian dengan jumlah subjek 100 akan mempunyai bobot yang lebih besar dibanding penelitian dengan jumlah subjek 50.

Selain jumlah subjek, bobot juga dipengaruhi oleh variasi data.

# .....5. Apakah yang dimaksud dengan *weight*(bobot)? Bagaimana menentukan dan interpretasi *weight*?



Test for heterogeneity  $\chi^2 = 9.7$  (df=8, p=0.287) ( $I^2 = 25.5\%$ )  
Test for overall effect  $Z = 2.55$  (p=0.005; fixed effect model)

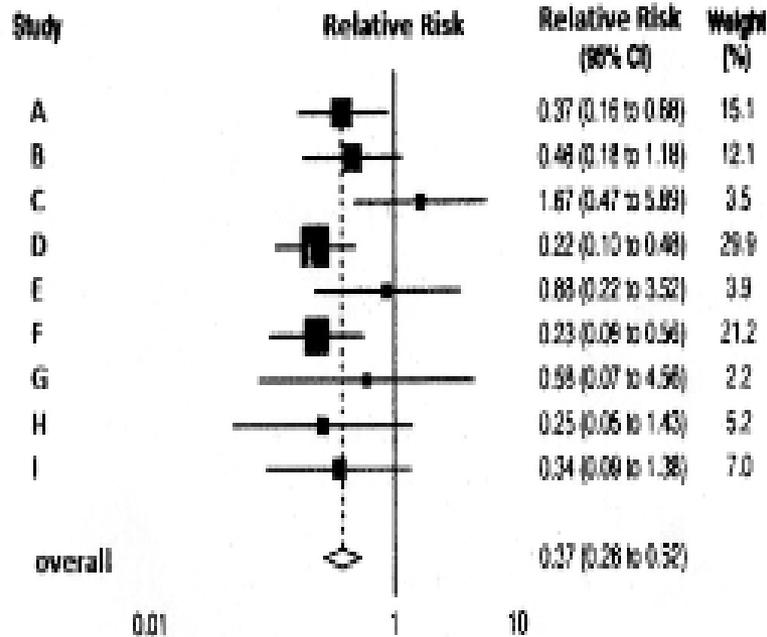
Bobot berbanding terbalik dengan variasi data.

Penelitian dengan data yang lebih bervariasi akan mempunyai bobot yang lebih kecil dibanding penelitian dengan variasi yang lebih kecil.

Misalnya, penelitian yang mempunyai simpang baku gabungan sebesar 20 akan mempunyai bobot yang lebih kecil dibandingkan penelitian dengan simpang baku gabungan 10.

Dengan demikian, secara matematis, bobot berbanding lurus dengan jumlah subjek dan berbanding terbalik dengan variasi data.

# .....5. Apakah yang dimaksud dengan *weight*(bobot)? Bagaimana menentukan dan interpretasi *weight*?



Test for heterogeneity  $\chi^2 = 9.7$  (df=8, p=0.287) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 2.55$  (p=0.005; fixed effect model)

Bobot =  $\frac{\text{jumlah subjek}}{\text{variasi}}$

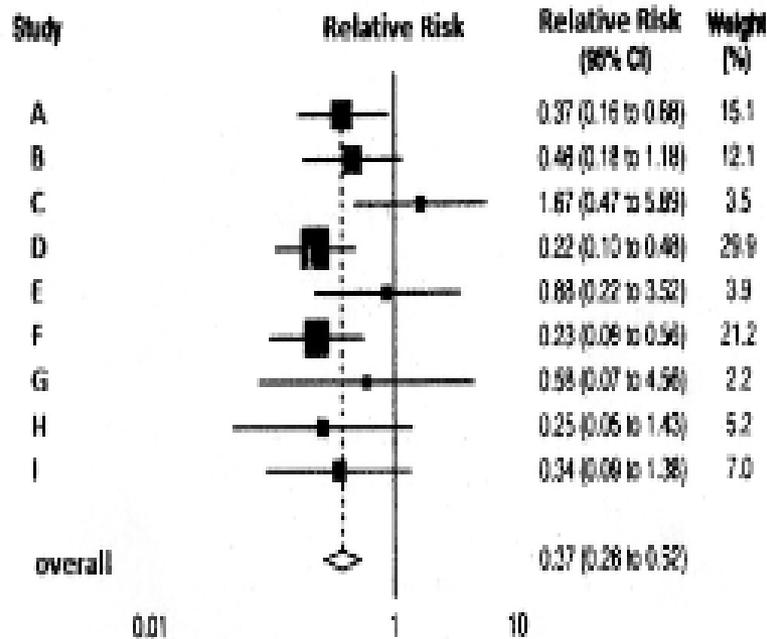
bobot inilah yang akan membedakan perhitungan rerata antara yang biasa dengan meta analisis.

Pada perhitungan efek gabungan biasa, perhitungan dilakukan tanpa memperhitungkan bobot.

Pada cara ini, nilai setiap penelitian (dengan lambang Y) akan dijumlahkan dibagi dengan jumlah penelitian (dengan lambang N)

$$\text{Rerata} = \frac{\sum Y}{n}$$

# .....5. Apakah yang dimaksud dengan *weight*(bobot)? Bagaimana menentukan dan interpretasi *weight*?



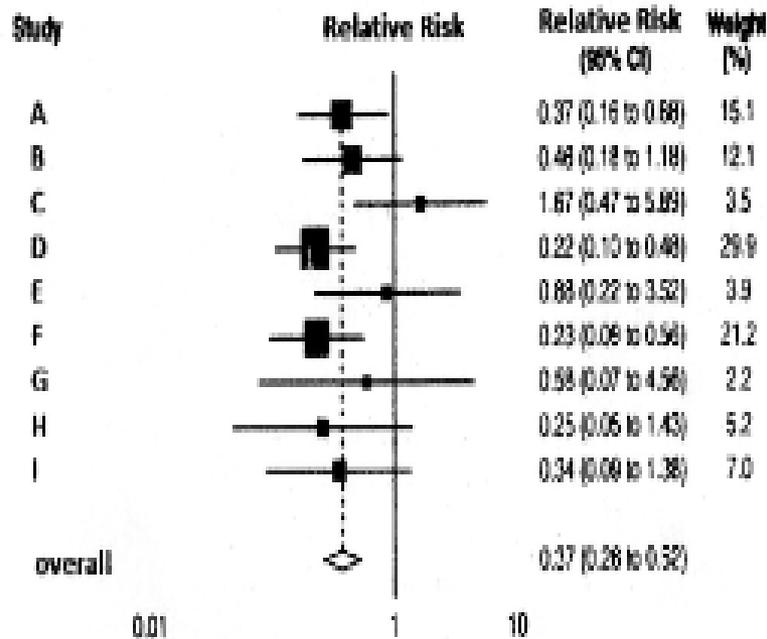
Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df=8, p=0.287) ( $I^2 = 25.5\%$ )  
 Test for overall effect  $Z = 2.55$  (p=0.005; fixed effect model)

Sementara itu, pada perhitungan efek gabungan dengan memperhitungkan bobot, nilai setiap penelitian (Y) akan dikalikan bobot (W).

Jumlah perkalian setiap penelitian dengan bobotnya ( $\sum WY$ ) kemudian dibagi dengan jumlah bobot  $\sum W$

$$\text{Rerata} = \frac{\sum W * Y}{\sum W}$$

## 6. Apakah yang dimaksud *test for overall effect*? Bagaimana cara memperoleh dan interpretasi?



Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df=8,  $p=0.287$ ) ( $I^2 = 25.5\%$ )  
Test for overall effect  $Z=2.55$  ( $p=0.005$ ; fixed effect model)

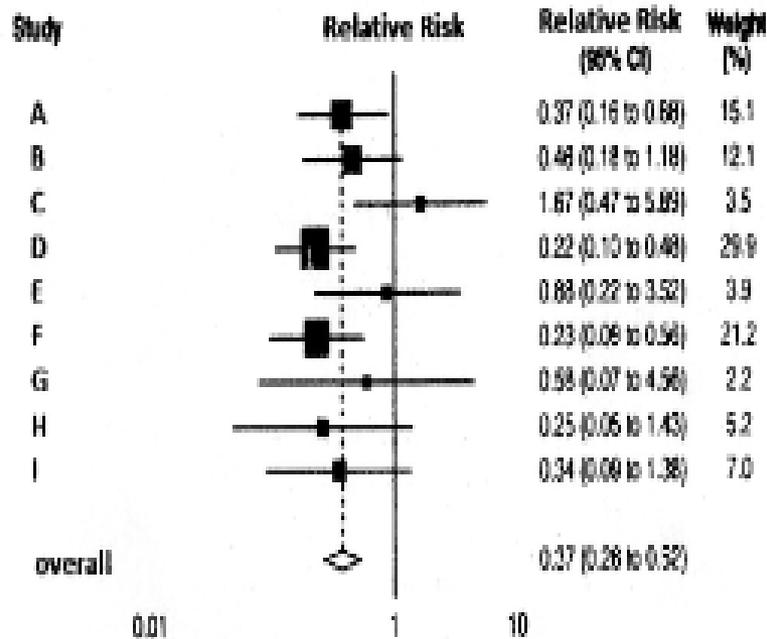
Overall effect adalah kesimpulan efek dari beberapa hasil penelitian (selanjutnya kita akan menyebutnya efek gabungan).

Test for overall effect adalah pengujian secara statistik apakah efek gabungan bermakna atau tidak.

Efek gabungan dikatakan bermakna secara statistik apabila nilai  $p$  kurang dari 0,05 dan interval kepercayaan efek gabungan tidak menyentuh garis vertikal.

.....6. Apakah yang dimaksud *test for overall effect*?

Bagaimana cara memperoleh dan interpretasi?



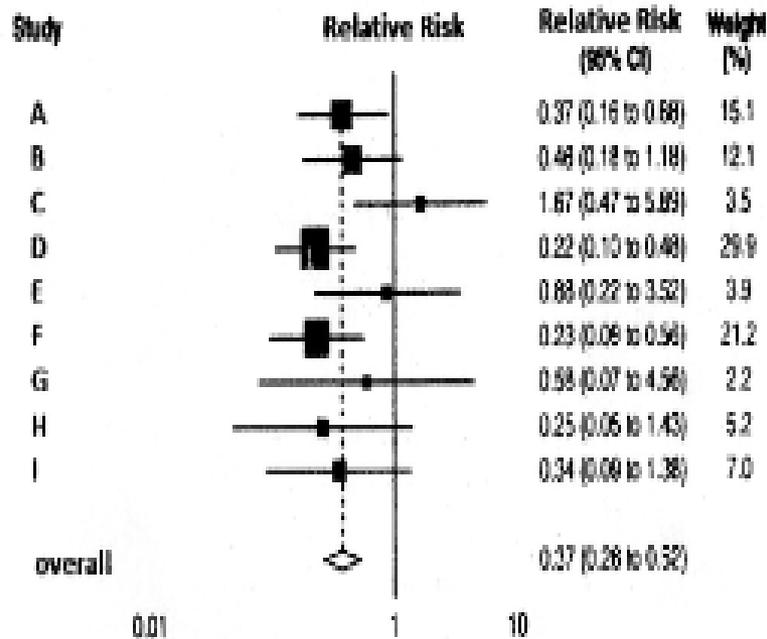
Test for heterogeneity  $\text{Chi}^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
Test for overall effect  $Z = 2.55$  ( $p = 0.005$ ; fixed effect model)

Pada gambar 1.1, nilai p efek gabungan adalah 0,05 dengan nilai efek gabungan sebesar 0,37 (0,26 sampai dengan 0,52).

Dengan demikian efek gabungan bermakna karena nilai p kurang dari 0,05 dan interval kepercayaan tidak melewati angka 1

# .....6. Apakah yang dimaksud *test for overall effect*?

## Bagaimana cara memperoleh dan interpretasi?



Test for heterogeneity  $Chi^2 = 9.7$  (df = 8,  $p = 0.287$ ) ( $I^2 = 25.5\%$ )  
Test for overall effect  $Z = 2.55$  ( $p = 0.005$ ; fixed effect model)

Selain secara statistik, efek gabungan harus diinterpretasikan secara klinis.

Interpretasi secara klinis bergantung pada efek gabungan yang diharapkan.

Efek gabungan yang diharapkan setiap peneliti mungkin berbeda-beda. Bagi peneliti yang mengharapkan efek gabungan sebesar 0,8 secara klinis efek gabungan penelitian ini (gambar 1..1) bermakna.

Akan tetapi, bagi peneliti yang mengharapkan efek gabungan sebesar 0,2 secara klinis efek gabungan penelitian tidak bermakna.

# INTERPRETASI EFEK GABUNGAN META- ANALISIS

---

# Tujuan

---

Setelah membaca bab ini, pembaca diharapkan mampu menjelaskan :

Interpretasi hasil penelitian secara statistic dengan menggunakan interval kepercayaan,

Interpretasi hasil penelitian secara statistic dengan menggunakan nilai p

Hubungan antara nilai p dan interval kepercayaan

Interpretasi hasil penelitian secara klinis berdasarkan effect size

Kombinasi kemungkinan interpretasi hasil penelitian secara statistic dan secara klinis

Table 2.1

Perbandingan hasil pengobatan A dan B berdasarkan keluaran kesembuhan, relaps, skor kualitas hidup, lama rawat, dan kadar kolesterol.



	obat A (n=500)	obat B (n=500)	P	efek (IK 95%)
<u>kesembuhan</u>				
<u>tidak sembuh</u>	100 (20)	50 (10)	<0,001	2,00(1,46 sd 2,47)
<u>sembuh</u>	400 (80)	450 (90)		
<u>relaps</u>				
<u>relaps</u>	175 (35)	150 (30)	0,521	1,17 (0,73 sd 1,87)
<u>tidak relaps</u>	325 (65)	350 (70)		
<u>kualitas hidup (0-100)</u>	65 ± 14	80 ± 15	<0,001	-15,0 (-16,8 sd -13,2)
<u>lama rawat (hari)</u>	10 ± 2	9 ± 2	0,001	0,5 (0,75 sd 1,25)
<u>kadar kolesterol (mg/dl)</u>	220 ± 50	215 ± 45	0,097	5,0 (-0,90 sd 10,9)

---

Variable kategorik disajikan dalam jumlah (n) dan persentase (%); variable numeric disajikan dalam rerata  $\pm$  simpang baku ; IK= Interval Kepercayaan; sd= sampai dengan; efek keluaran kategorik dinilai dengan resiko relative (RR) dan IK; efek skala numeric dalam selisih rerata dan IK.

1. Apa ukuran efek yang diteliti pada keluaran kesembuhan, relaps, kualitas hidup, lama rawat, dan kadar kolesterol?

---

Ukuran efek pada keluaran kesembuhan dan relaps adalah risiko relative (RR).

Ukuran efek pada keluaran kualitas hidup, lama rawat, dan kadar kolesterol adalah selisih rerata.

---

Keluaran kesembuhan dan relaps (tabel 2.3) adalah risiko relatif (RR).

Risiko relatif merupakan perbandingan probabilitas.

Dalam kasus ini, RR dihitung dengan cara membandingkan probabilitas tidak sembuh antara obat A dan obat B.

Probabilitas tidak sembuh pada obat A adalah 20% sedangkan pada obat B adalah 10%.

Dengan demikian, RR adalah  $20\%/10\%=2$ .

Risiko relatif adalah ukuran efek yang menggunakan konsep *perbandingan*.

Secara matematis, konsep perbandingan adalah x dibandingkan y atau x:y.

Effek lain yang menggunakan prinsip perbandingan adalah odds ratio (OR) dan hazard ratio (HR).

**Tabel 2.3.**

Ukuran efek untuk keluaran kesembuhan dan relaps

	<b>Obat A (n=500)</b>	<b>Obat B (n=500)</b>	<b>RR</b>
Kesembuhan	100 (20)	50 (10)	2,00
Tidak sembuh	400 (80)	450 (90)	
Sembuh			
Relaps			1,17
Relaps	175 (35)	150 (30)	
Tidak relaps	325 (65)	350 (70)	

Keluaran kualitas hidup, lama rawat, dan kadar kolesterol (tabel 2.4) adalah selisih rerata (*mean difference*). Selisih rerata merupakan hasil pengurangan rerata antar kelompok. Dalam kasus ini, selisih rerata dihitung dengan cara mengurangi rerata kelompok A dengan rerata pada kelompok B. Untuk keluaran kualitas hidup misalnya, selisih rerata adalah rerata kelompok A (80) dikurang kelompok B (60), yaitu 20. Secara matematis, konsep selisih adalah x dikurangi y atau  $x-y$ . Selisih rerata adalah salah satu efek yang menggunakan prinsip selisih. Efek lain yang menggunakan prinsip selisih adalah selisih proporsi.

**Tabel 2.4.**

Ukuran efek untuk keluaran kualitas hidup, lama rawat, dan kadar kolesterol

Kesembuhan	<b>Obat A</b> (n=500)	<b>Obat B</b> (n=500)	<b>Perbedaan rerata</b>
Kualitas hidup (0-100)	65±14	80±15	-15,0
Lama rawat (hari)	10±2	9±2	0,5
Kadar Kolesterol (mg/dl)	220±50	215±45	5,0

Ukuran efek yang mempunyai prinsip pembagian secara matematis adalah  $x:y$ .

Ukuran efek yang mempunyai prinsip selisih secara matematis adalah  $x-y$ .

2. Metode inferensi secara statistik apakah yang digunakan pada penelitian- penelitian tersebut ?

---

**Jawab:**

Metode inferensi statistik yang digunakan dalam penelitian tersebut adalah nilai  $p$  dan interval kepercayaan.

**Penjelasan:**

Inferensi adalah teknik untuk melakukan penarikan kesimpulan dari sampel ke populasi.

Secara statistik, inferensi dilakukan dengan dua metode, yaitu dengan menghitung nilai  $p$  dan nilai interval kepercayaan.

Pada tabel 2.1 misalnya, inferensi dilakukan dengan menghitung nilai  $p$  (lihat kolom  $p$ ) dan interval kepercayaan (lihat kolom interval kepercayaan).

Beberapa jurnal menginginkan inferensi dilakukan dengan nilai  $p$ , beberapa yang lain dengan interval kepercayaan.

Tidak sedikit pula jurnal yang menghendaki kedua teknik ini digunakan secara bersamaan.

3. Apakah informasi yang diberikan oleh nilai  $p$ ?

**Jawab:**

---

Nilai  $p$  memberikan informasi peranan faktor probabilitas untuk memperoleh efek atau lebih ekstrim lagi apabila hipotesis nol benar.

**Penjelasan:**

Perhatikan keluaran kesembuhan (tabel 2.5).

Nilai  $p$  pada keluaran kesembuhan adalah  $<0,001$ .

Perbedaan proporsi tidak sembuh antara obat A dan obat B adalah 20% dengan risiko relatif sebesar 2.

Nilai  $p$  sebesar  $<0,001$  artinya peranan faktor probabilitas untuk memperoleh perbedaan proporsi 20% atau lebih dari 20% adalah kurang dari 0,1% bila tidak ada perbedaan kesembuhan antara obat A dan obat B.

Nilai  $p$  sebesar  $<0,001$  dapat juga diartikan peranan faktor probabilitas untuk memperoleh risiko relatif 2 atau lebih besar dari 2 adalah kurang dari 0,1% bila tidak ada perbedaan kesembuhan antara obat A dan obat B.

**Tabel 2.5.**

Inferensi dengan nilai  $p$  pada keluaran kesembuhan

Kesembuhan	Obat A (n=500)	Obat B (n=500)	$p$
Tidak sembuh	100 (20)	50 (10)	<0,001
Sembuh	400 (80)	450 (90)	

Sekarang perhatikan keluaran kadar kolesterol (tabel 2.6). Nilai  $p$  pada keluaran kadar kolesterol adalah <0,097 dengan perbedaan rerata 5 mg/dl. Nilai  $p$  sebesar 0,097 artinya peranan faktor probabilitas untuk memperoleh perbedaan rerata 5 mg/dl atau lebih besar dari 5 mg/dl adalah 9,7% bila tidak ada perbedaan kadar kolesterol antara obat A dan obat B.

**Tabel 2.6.**

Inferensi dengan nilai  $p$  pada keluaran kualitas hidup, lama rawat, dan kadar kolesterol

	<b>Obat A</b> <b>(n=500)</b>	<b>Obat B</b> <b>(n=500)</b>	<b>p</b>
Kadar kolesterol mg/dl)	220±50	215±45	0,097

Nilai  $p$  memberikan informasi peranan faktor probabilitas untuk memperoleh efek atau lebih ekstrim lagi untuk apabila hipotesis nol benar.

4. Apakah kesimpulan yang diberikan oleh nilai  $p$  ?

---

**Jawab:**

Hipotesis nol ditolak apabila nilai  $p$  kurang dari 0,05. Hipotesis nol diterima apabila nilai  $p$  lebih dari 0,05.

---

Ujistatistik adalah uji yang dilakukan terhadap hipotesis nol.

Hipotesis nol adalah hipotesis yang menyatakan tidak ada perbedaan atau tidak ada hubungan.

Hipotesis nol **diterima** apabila nilai  $p$  pada uji statistik lebih besar dari batas kemaknaan.

Sebaliknya, hipotesis nol **ditolak** apabila  $p$  pada uji statistik lebih kecil dari batas kemaknaan.

Batas kemaknaan pada contoh kasus ini adalah 5%.

# Tabel 2.7

## Kesimpulan nilai p pada keluaran kesembuhan

A	Obat A (n=500)	Obat B (n=500)	P	Kesimpulan
Kesembuhan				Karena nilai $p < 0,005$ , hipotesis nol ditolak.
Tidak Sembuh	100 (20)	50 (10)	<0,001	
Sembuh	400 (80)	450 (90)		

Perhatikan keluaran kesembuhan (table 2.7).

Hipotesis nol pada uji statistik untuk membandingkan kesembuhan adalah *tidak ada perbedaan kesembuhan antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.*

Karena nilai  $p < 0,05$ , hipotesis nol ditolak.

Dengan demikian, secara statistik terdapat perbedaan kesembuhan antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.

Secara populer, hasil ini sering disebut 'secara statistik bermakna'.

# Tabel 2.8

## Kesimpulan nilai p pada keluaran kesembuhan

---

	Obat A (n=500)	Obat B (n=500)	p	Kesimpulan
Relaps				Karenilai $p < 0,005$ , hipotesisnolditerima.
Relaps	175 (35)	150 (30)	0,521	
Tidakrelaps	325 (65)	350 (70)		

# 5. Apakah informasi yang diberikan oleh interval kepercayaan?

Interval kepercayaan memberikan informasi prediksi nilai pada populasi.

Hakekat suatu penelitian adalah untuk mengetahui nilai pada populasi.

Pada sebagian besar masalah, nilai pada populasi tidak dapat secara langsung diketahui melainkan diprediksikan dari nilai yang kita peroleh pada subjek penelitian.

Cara memprediksi nilai pada populasi adalah dengan mengurangi dan menambahkan nilai pada subjek dengan perkalian antara deviat baku alpha dan standar error (lihat persamaan 2.1)

$$IK \text{ min} = S - Z\alpha * SE$$

$$IK \text{ mak} = S + Z\alpha * SE$$

$$IK = S \pm Z\alpha * SE$$

Bila pada interval kepercayaan konsep perbandingan ada angka 1, berarti tidak ada perbedaan dari nilai yang sedang dibandingkan.

Secara matematis, jika  $x:y=1$ , maka  $x=y$ .

---

S = nilai pada subjek

IKmin = interval kepercayaan minimal

IKmak = interval kepercayaan maksimal

$Z\alpha$  = deviat baku alpha. Bila alpha 5%,  $Z\alpha=1,96$

SE = standar error

**Catatan:** bagian paling sulit dari rumus ini adalah pada standar error karena masing-masing ukuran statistic mempunyai rumus standar error yang berbeda.

Kita cukup tahu sampai rumus interval kepercayaan saja.

Yang penting, kita mengetahui prinsip bagaimana interval kepercayaan dihitung.

Table 2.1

Perbandingan hasil pengobatan A dan B berdasarkan keluaran kesembuhan, relaps, skor kualitas hidup, lama rawat, dan kadar kolesterol.

	obat A (n=500)	obat B (n=500)	P	efek (IK 95%)
<u>kesembuhan</u>				
<u>tidak sembuh</u>	100 (20)	50 (10)	<0,001	2,00(1,46 sd 2,47)
<u>sembuh</u>	400 (80)	450 (90)		
<u>relaps</u>				
<u>relaps</u>	175 (35)	150 (30)	0,521	1,17 (0,73 sd 1,87)
<u>tidak relaps</u>	325 (65)	350 (70)		
<u>kualitas hidup (0-100)</u>	65 ± 14	80 ± 15	<0,001	-15,0 (-16,8 sd -13,2)
<u>lama rawat (hari)</u>	10 ± 2	9 ± 2	0,001	0,5 (0,75 sd 1,25)
<u>kadar kolesterol (mg/dl)</u>	220 ± 50	215 ± 45	0,097	5,0 (-0,90 sd 10,9)

Perhatikan keluaran kesembuhan.

Nilai RR dan IK pada keluaran kesembuhan 2,00 (1,46 sampai dengan 2,74).

Artinya, nilai RR pada sampel adalah 2 sementara prediksi RR pada populasi adalah antara 1,46 sampai dengan 2,74.

Angka tersebut diperoleh dari:

$$IK = S \pm Z\alpha * SE = 2 \pm 1,96 * SE$$

Sekarang perhatikan keluaran skor kualitas hidup.

Perbedaan rerata dan IK adalah -15,0 (-16.8 sampaidengan -13.2).

Artinya, selisih kualitas hidup antara obat A dan B adalah -15 sementara prediksi selisih kualitas hidup antara obat A dan obat B pada populasi adalah antara -16,8 sampai dengan -13,2 (tanda minus menunjukkan A lebih besar dari B). Angka tersebut diperoleh dari:

$$IK = S \pm Z\alpha * SE = -15 \pm 1,96 * SE$$

## 6. Mengapa angka 1 penting dalam interpretasi interval kepercayaan pada ukuran yang menggunakan prinsip perbandingan?

---

Bila pada interval kepercayaan konsep perbandingan ada angka 1, berarti tidak ada perbedaan dari nilai yang sedang dibandingkan.

Secara matematis, jika  $x:y=1$ , maka  $x=y$ .

# Tabel 2.9

## Interval kepercayaan untuk keluaran kesembuhan

	Obat A (n=50)	Obat B (n=50)	Efek (IK95%)
<b>Kesembuhan</b>			
<b>Tidak Sembuh</b>	100 (20)	50 (10)	2,00 (1,46 sd 2,74)
<b>Sembuh</b>	400 (80)	450 (90)	

Perhatikan keluaran kesembuhan.

Kita sepakati proporsi kesembuhan kelompok A disebut  $x$  sementara proporsi kesembuhan kelompok B adalah  $y$ .

Bila hasil pembagian antara  $x$  dan  $y$  pada populasi sama dengan 1, kita bisa mengatakan bahwa  $x$  sama dengan  $y$  atau tidak ada perbedaan antara  $x$  dan  $y$ .

Sebaliknya, bila pada populasi pembagian antara  $x$  dan  $y$  tidak sama dengan satu, kita bisa mengatakan bahwa  $x$  berbeda dengan  $y$ .

Pada konsep pembagian berlaku:

Jikax :  $y = 1$ , maka  $x = y$

Jika  $x = y$ , maka  $x : y = 1$

Perhatikan kembali keluaran kesembuhan (table 2.9).

Nilai RR pada populasi adalah 1,46 sampai dengan 2,74.

Interval kepercayaan tidak melewati angka 1.

Dengan demikian, kita dapat mengatakan bahwa kesembuhan kelompok A dan B berbeda.

## Tabel 2. 10.

### Interval kepercayaan untuk kelurahan relaps

	Obat A (n=500)	Obat B (n=500)	Efek (IK95%)
Relaps			
Relaps	175(35)	150(30)	1,17 (0,73 sd 1,87)
Tidak relaps	325(65)	350(70)	

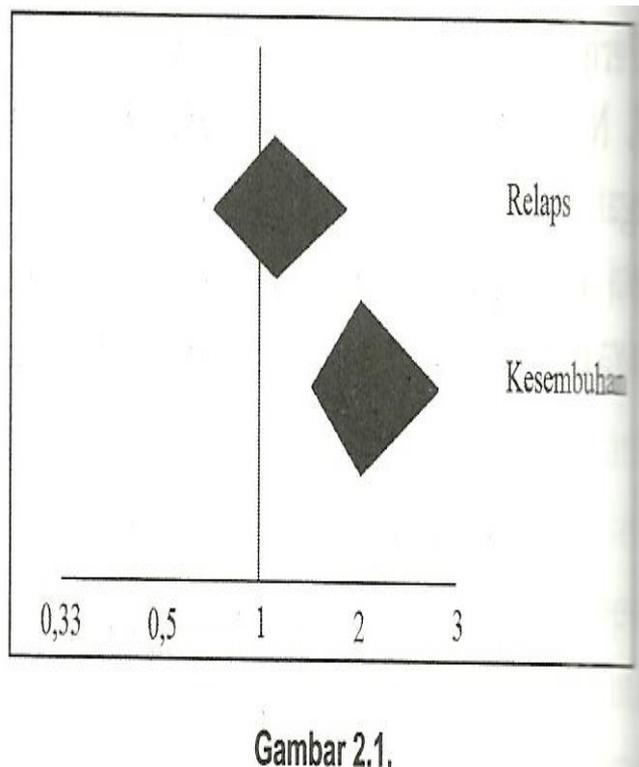
Pada keluaran relaps (tabel 2.10), nilai RR pada populasi adalah (0,73 sd 1,87).

Interval kepercayaan melewati angka 1.

Dengan demikian, kita dapat mengatakan bahwa secara statistik tidak terdapat perbedaan relaps antara kelompok A dan kelompok B.

## Gambar 2.1.

Perbandingan kesembuhan dan relaps antara obat A dengan obat B dalam bentuk grafik.



Gambar 2.1.

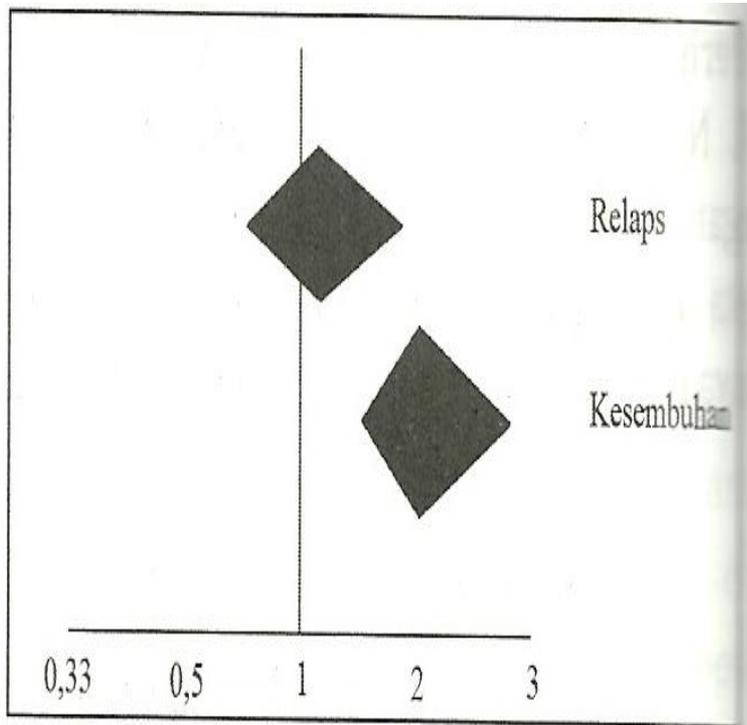
Pada meta-analisis, efek gabungan dan interval kepercayaan divisualisasikan dalam gambar berbentuk wajik (gambar 2.1).

Perhatikan bahwa titik tengah wajik merupakan efek gabungan sementara ujung kiri dan kanan adalah nilai minimal dan maksimal dari interval kepercayaan.

Gambar dilengkapi dengan garis vertikal pada nilai 1 untuk memudahkan penarikan kesimpulan.

## Gambar 2.1.

Perbandingan kesembuhan dan relaps antara obat A dengan obat B dalam bentuk grafik.



Gambar 2.1.

Perhatikan bahwa untuk keluaran kesembuhan, interval kepercayaan tidak melewati angka 1.

Sementara untuk relaps interval kepercayaan melewati angka 1.

7. Mengapa angka nol penting dalam interpretasi interval kepercayaan pada ukuran yang menggunakan prinsip selisih?

---

Bila pada interval kepercayaan konsep selisih ada angka nol, berarti tidak ada perbedaan dari nilai yang sedang dibandingkan.

Secara matematis, jika  $x-y=1$ , maka  $x=y$ .

## Tabel 2.11.

Interval kepercayaan untuk keluaran kualitas hidup, lama rawat, dan kadar kolesterol

	Obat A (n=500)	Obat B (n=500)	Efek (IK95%)
Kualitas hidup (0-100)	65+-14	80+-15	-15,0 (-16,8 sd -13,2)
Lama rawat (hari)	10+-2	9+-2	0,5 (0,75 sd 1,25)
Kadar kolesterol (mg/dl)	220+-50	21,5+- 45	5,0 (-0.90 sd 10,9)

Perhatikan keluaran skor kualitas hidup.

Kita sepakati kualitas hidup kelompok A pada populasi disebut x sementara kualitas hidup kelompok B pada populasi disebut y sementara kualitas hidup kelompok B pada populasi adalah y.

Bila hasil pengurangan x dengan y pada populasi sama dengan nol, kita bisa mengatakan bahwa x berbeda dengan y.

Pada konsep pengurangan berlaku:

Bila  $x-y=0$ , maka  $x=y$

Bila  $x=y$ , maka  $x-y=0$

Perhatikan kembali keluaran kualitas hidup (tabel 2.11). perbedaan rerata pada populasi adalah -15 sampai dengan -13,2.

Interval kepercayaan tidak melewati angka nol.

Dengan demikian, kita dapat mengatakan bahwa kualitas hidup antara kelompok A dan B berbeda.

## Tabel 2.11.

Interval kepercayaan untuk keluaran kualitas hidup, lama rawat, dan kadar kolesterol

	Obat A (n=500)	Obat B (n=500)	Efek (IK95%)
Kualitas hidup (0-100)	65+-14	80+-15	-15,0 (-16,8 sd -13,2)
Lama rawat (hari)	10+-2	9+-2	0,5 (0,75 sd 1,25)
Kadar kolesterol (mg/dl)	220+-50	21,5+- 45	5,0 (-0.90 sd 10,9)

Pada keluaran lama rawat, perbedaan rerata pada populasi adalah 0,75 sampai dengan 1,25.

Interval kepercayaan tidak melewati angka nol.

Dengan demikian, kita dapat mengatakan bahwa lama rawat kelompok A dan B berbeda.

Pada keluaran kadar kolesterol, perbedaan rerata pada populasi adalah -0,90 sampai dengan 10,9.

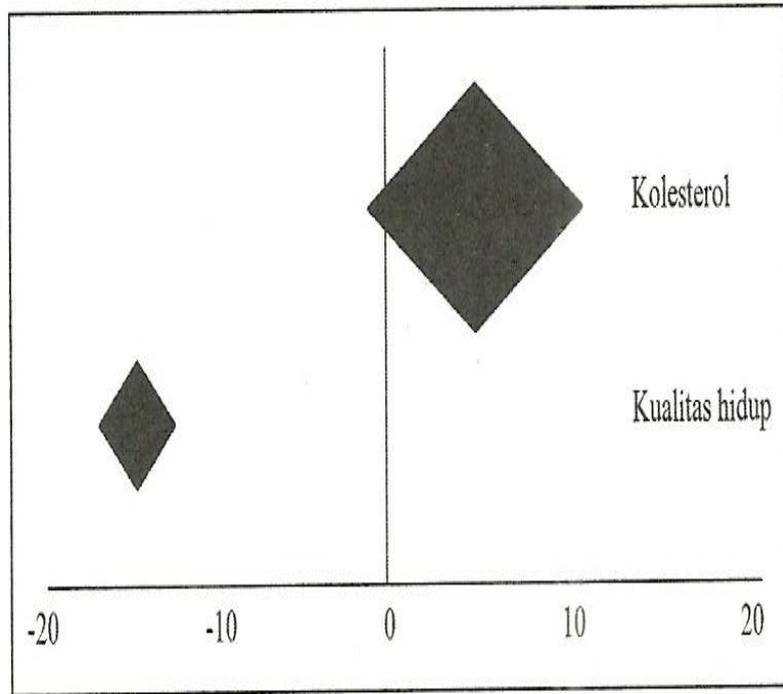
Interval kepercayaan melewati angka nol.

Dengan demikian, kita dapat mengatakan bahwa kadar kolesterol kelompok A dan B sama.

## Gambar 2.2.

Perbandingan kualitas hidup dan kadar kolesterol antara obat A dengan B dalam bentuk grafik.

Perhatikan bahwa untuk keluaran kualitas hidup. Interval kepercayaan melewati angka nol.



Gambar 2.2.

Pada meta-analisis, efek gabungan dan interval kepercayaannya divisualisasikan dalam gambar berbentuk wajik (gambar 2.2).

Perhatikan bahwa titik tengah wajik merupakan efek tengah gabungan sementara ujung kiri dan kanan adalah nilai minimal dan maksimal dari interval kepercayaan.

Gambar dilengkapi dengan garis vertikal pada nilai 0 untuk memudahkan penarikan kesimpulan.

## 8. Apakah kesimpulan yang diberikan oleh interval kepercayaan?

---

Pada konsep selisih, hipotesis nol ditolak apabila interval kepercayaan tidak melewati angka nol.

Pada konsep selisih, hipotesis nol diterima apabila bila interval kepercayaan melewati angka nol.

Pada konsep perbandingan, hipotesis nol ditolak bila interval kepercayaan tidak melewati angka satu.

Pada konsep perbandingan, hipotesis nol diterima bila interval kepercayaan melewati angka satu.

Uji statistik adalah uji yang dilakukan terhadap hipotesis nol.

Hipotesis nol adalah hipotesis yang menyatakan tidak ada perbedaan atau tidak ada hubungan.

Hipotesis nol diterima apabila interval kepercayaan melewati angka nol (pada konsep selisih) atau melewati angka satu (pada konsep pembagian).

Sebaliknya, hipotesis nol ditolak apabila interval kepercayaan tidak melewati angka nol (pada konsep selisih) atau tidak melewati angka satu (pada konsep pembagian).

## Tabel 2.12.

### Interval kepercayaan untuk keluaran kesembuhan dan relaps

	Obat A (n=500)	Obat B (n=500)	Efek (IK95%)
<b>Kesembuhan</b>			
<b>Tidak sembuh</b>	100(20)	50(10)	2,00 (1,46 sd 2,74)
<b>Sembuh</b>	400(80)	450(90)	
<b>Relaps</b>			
<b>Relaps</b>	175(35)	150(30)	1,17 (0,73 sd 1,87)
<b>Tidak relaps</b>	325(65)	350(70)	

Perhatikan keluaran kesembuhan (tabel 2.12).

Hipotesis nol pada uji statistik untuk membandingkan kesembuhan adalah *tidak ada perbedaan kesembuhan antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.*

Karena interval kepercayaan tidak melewati angka satu, hipotesis nol ditolak.

Dengan demikian, secara statistik terdapat perbedaan kesembuhan antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.

Secara populer, hasil ini sering disebut 'secara statistik bermakna'.

## Tabel 2.12.

### Interval kepercayaan untuk keluaran kesembuhan dan relaps

	Obat A (n=500)	Obat B (n=500)	Efek (IK95%)
Kesembuhan			
Tidak sembuh	100(20)	50(10)	2,00 (1,46 sd 2,74)
Sembuh	400(80)	450(90)	
Relaps			
Relaps	175(35)	150(30)	1,17 (0,73 sd 1,87)
Tidak relaps	325(65)	350(70)	

Perhatikan keluaran relaps (tabel 2.12).

Hipotesis nol pada uji statistik untuk membandingkan relaps adalah *tidak ada perbedaan relaps antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.*

Karena interval kepercayaan melewati angka satu, hipotesis nol diterima.

Dengan demikian, secara statistik tidak terdapat perbedaan relaps antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.

Secara populer, hasil ini sering disebut 'secara statistik tidak bermakna'.

## Tabel 2.13.

Interval kepercayaan untuk keluaran kualitas hidup, lama rawat, dan kadar kolesterol

	Obat A (n=500)	Obat B (n=500)	Efek (IK 95%)
Kualitas hidup (0-100)	65±14	80±15	-15,0 (-16,8 sd -13,2)
Lama rawat (hari)	10±2	9±2	0,5 (0,75 sd 1,25)
Kadar kolesterol (g/dl)	220±50	215±45	5,0 (-0,90 sd 10,9)

Perhatikan keluaran kualitas hidup (tabel 2.13).

Hipotesis nol pada uji statistik untuk membandingkan kualitas hidup adalah *tidak ada perbedaan kualitas hidup antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.*

Karena interval kepercayaan tidak melewati angka nol, hipotesis nol ditolak.

Dengan demikian, secara statistik terdapat perbedaan kualitas hidup antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.

Secara populer, hasil ini sering disebut 'secara statistik bermakna'.

## Tabel 2.13.

Interval kepercayaan untuk keluaran kualitas hidup, lama rawat, dan kadar kolesterol

	Obat A (n=500)	Obat B (n=500)	Efek (IK 95%)
Kualitas hidup (0-100)	65±14	80±15	-15,0 (-16,8 sd -13,2)
Lama rawat (hari)	10±2	9±2	0,5 (0,75 sd 1,25)
Kadar kolesterol (g/dl)	220±50	215±45	5,0 (-0,90 sd 10,9)

Perhatikan keluaran kolesterol (tabel 2.13).

Hipotesis nol pada uji statistik untuk membandingkan kolesterol adalah tidak ada perbedaan kolesterol antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.

Karena interval kepercayaan melewati angka nol, hipotesis nol diterima.

Dengan demikian, secara statistic tidak terdapat perbedaan kolesterol antara subjek yang memperoleh obat A dibandingkan subjek yang menggunakan obat B.

Secara populer, hasil ini sering disebut 'secara statistik tidak bermakna'.

# 9. Buktikan bahwa nilai p dan interval kepercayaan memberikan kesimpulan yang sama walaupun informasi yang diberikan berbeda!

Tabel 2.14.

Kesamaan kesimpulan nilai p dan interval kepercayaan untuk keluaran kesembuhan dan relaps

	P	Efek (IK95%)
<b>Kesembuhan</b> Tidak sembuh Sembuh	<0,001	2,00 (1,46 sd 2,74)
<b>Relaps</b> Relaps Tidak relaps	0,521	1,17 (0,73 sd 1,87)

Perhatikan keluaran kesembuhan (tabel 2.14).

Nilai  $p < 0,001$  dan IK antara 1,46 sampai dengan 2,74 memang memberikan informasi berbeda.

Nilai p memberikan informasi peranan faktor probabilitas untuk memperoleh perbedaan proporsi 20% atau lebih dari 20% adalah kurang dari 0,1% bila tidak ada perbedaan kesembuhan antara obat A dan obat B.

Sementara itu, IK memberikan informasi prediksi RR pada populasi adalah 1,46 sampai dengan 2,74.

Walaupun informasi berbeda, kesimpulan yang diberikan sama, yaitu *bermakna*.

Menurut nilai p bermakna karena nilai p kurang dari 0,05.

Menurut IK bermakna karena pada interval kepercayaan tidak ada angka satu.

## Tabel 2.14.

Kesamaan kesimpulan nilai p dan interval kepercayaan untuk keluaran kesembuhan dan relaps

	P	Efek (IK95%)
<b>Kesembuhan</b> Tidak sembuh Sembuh	<0,001	2,00 (1,46 sd 2,74)
<b>Relaps</b> Relaps Tidak relaps	0,521	1,17 (0,73 sd 1,87)

Perhatikan keluaran relaps (tabel 2.14). Nilai  $p=0,571$  dan IK antara 0,73 sampai dengan 1,87 memang memberikan informasi berbeda.

Nilai p memberikan informasi peranan faktor probabilitas untuk memperoleh RR 1,17 atau lebih dari 1,17 adalah 57,1% bila tidak ada perbedaan relaps antara obat A dan obat B.

Sementara itu, IK memberikan informasi prediksi RR pada populasi adalah 0,73 sampai dengan 1,87.

Walaupun informasi berbeda, kesimpulan yang diberikan sama, yaitu *tidak bermakna*.

Menurut nilai p tidak bermakna karena nilai p lebih dari 0,05.

Menurut IK tidak bermakna karena pada interval kepercayaan ada angka satu.

---

Bila nilai  $p$  menyimpulkan bermakna, IK akan menyimpulkan bermakna.

Bila nilai  $p$  menyimpulkan tidak bermakna, IK akan menyimpulkan tidak bermakna.

Bila IK menyimpulkan bermakna, nilai  $p$  akan menyimpulkan bermakna.

Bila IK menyimpulkan tidak bermakna, maka nilai  $p$  akan menyimpulkan tidak bermakna.

## Tabel 2.15

Kesamaan kesimpulan nilai p dan interval kepercayaan untuk keluaran kualitas hidup, lama rawat, dan kadar kolesterol

	p	Efek (IK95%)
Kualitas hidup (0-100)	<0,001	-15,0 (-16,8 sd -13,2)
Lama rawat (hari)	0,001	0,5 (0,75 sd 1,25)
Kadar kolesterol (g/dl)	0,097	5,0 (-0,90 sd 10,9)

Perhatikan keluaran kadar kolesterol (tabel 2.15).

Nilai  $p=0,0097$  dan IK antara -0,90 sampai dengan 10,90 memang memberikan informasi berbeda.

Nilai p memberikan informasi peranan faktor probabilitas untuk memperoleh perbedaan rerata -0,90 atau lebih dari -0,90 adalah 9,7% bila tidak ada perbedaan kolesterol antara obat A dan obat B.

Sementara itu, IK memberikan informasi prediksi selisih rerata pada populasi adalah -0,90 sampai dengan 10,90.

Walaupun informasi berbeda, kesimpulan yang diberikan sama, yaitu *tidak bermakna*. Menurut nilai p tidak bermakna karena nilai p lebih dari 0,05.

Menurut IK tidak bermakna karena pada interval kepercayaan ada angka nol.

# 10. Apa kelemahan interpretasi penelitian yang hanya berdasarkan statistik?

---

Interpretasi secara statistik (nilai p dan interval kepercayaan) dipengaruhi oleh jumlah subjek dan variasi data.

Pada efek yang sama, nilai p akan semakin kecil dengan semakin bertambahnya jumlah subjek.

Sementara itu, Interval kepercayaan akan semakin sempit seiring dengan bertambahnya subjek.

# Tabel 2.16.

Ilustrasi untuk membuktikan bahwa nilai  $p$  dan interval kepercayaan dipengaruhi oleh jumlah subyek

	Obat A	Obat B	P	Efek (IK95%)
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	100 (20)	50 (10)	<0.001	2.00 (1.46 sd 2.74)
<b>sembuh</b>	400(80)	450 (90)		
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	50 (20)	25 (10)	0.002	2.00 (1.28 sd 3.13)
<b>sembuh</b>	200 (80)	225 (90)		
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	20 (20)	10 (10)	0.048	2.00 (0.99 sd 4.05)
<b>sembuh</b>	80 (80)	90 (90)		
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	10(20)	5 (10)	0.161	2.00 (0.74 sd 5.43)
<b>sembuh</b>	50 (80)	50 (90)		
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	2 (20)	1 (10)	0,531	2.00 (0.21 sd 18.69)
<b>sembuh</b>	8 (80)	10 (90)		

Variable kategorik disajikan dalam jumlah dan presentase; IK=interval kepercayaan; sd=sampai dengan;efek keluaran kategorik dinilai dengan risiko relative(RR) dan IK.

Pada table 2.16,terdapat lima hasil penelitian yang membandingkan kesembuhan obat A dan obat B.

Setiap penelitian mempunyai jumlah subjek yang berbeda.

Penelitian pertama mempunyai jumlah subjek terbanyak sementara penelitian kelima terkecil.

Proporsi tidak sembuh pada pada kelompok A pada semua penelitian adalah 20 % sedangkan kelompok B 10 %.

Dengan demikian, nilai risiko relative semua penelitian sama , yaitu 2

Tabel 2.16.

Ilustrasi untuk membuktikan bahwa nilai  $p$  dan interval kepercayaan dipengaruhi oleh jumlah subyek

	Obat A	Obat B	P	Efek (IK95%)
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	100 (20)	50 (10)	<0.001	2.00 (1.46 sd 2.74)
<b>sembuh</b>	400(80)	450 (90)		
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	50 (20)	25 (10)	0.002	2.00 (1.28 sd 3.13)
<b>sembuh</b>	200 (80)	225 (90)		
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	20 (20)	10 (10)	0.048	2.00 (0.99 sd 4.05)
<b>sembuh</b>	80 (80)	90 (90)		
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	10(20)	5 (10)	0.161	2.00 (0.74 sd 5.43)
<b>sembuh</b>	50 (80)	50 (90)		
<b>Kesembuhan</b>				
<b>Tidak sembuh</b>	2 (20)	1 (10)	0,531	2.00 (0.21 sd 18.69)
<b>sembuh</b>	8 (80)	10 (90)		

Perhatikan nilai  $p$ .

Dengan perbedaan efek yang sama, yaitu  $RR=2$ , nilai  $p$  semakin besar seiring dengan berkurangnya jumlah subjek.

Perhatikan juga kolom interval kepercayaan.

Interval kepercayaan semakin lebar seiring dengan berkurangnya jumlah subjek.

Dengan demikian, terbukti bahwa statistic ( nilai  $p$  dan IK) bergantung pada jumlah subjek, dimana nilai  $p$  dan lebar IK berbanding terbalik jumlah subjek.

Nilai  $.p$  & IK  $\approx \frac{1}{2}$

Selain jumlah subjek, statistic juga dipengaruhi oleh variasi data.

Semakin data bervariasi, semakin besar nilai  $p$  dan semakin lebar interval kepercayaannya.

Tabel 2.17.

Ilustrasi untuk membuktikan bahwa nilai  $p$  dan interval kepercayaan dipengaruhi oleh variasi

	Obat A	Obat B	P	Efek (IK95%)
	(n=50)	(n=50)		
Kualitas hidup (0-100)	75+-10	80+-10	0.0141	-5 (-9.99 sd 1.03 )
Kualitas hidup (0-100)	75+-20	80+-20	0.2143	-5 (-12.94 sd 2.94 )
Kualitas hidup (0-100)	75+-25	80+-25	0.3198	-5 ( -14,92 sd 4.92 )
Kualitas hidup (0-100)	75+-30	80+-30	0.4067	-5 (-16,91 sd 6.91 )

Variabel numeric disajikan dalam rerata kurang lebih simpang baku; Ik=interval kepercayaan; sd + sampai dengan; efek skala numeric dalam selisih rerata dan IK.

Pada table 2.17, terdapat lima hasil penelitian yang membandingkan kesembuhan obat A dan obat B.

Setiap penelitian mempunyai jumlah subjek yang sama.

Pembeda antara penelitian adalah pada variasinya yang dapat dilihat dari simpang baku.

Penelitian pertama mempunyai variasi terkecil Sementara penelitian kelima mempunyai variasi terbesar.

Selisih kualitas hidup antar kelompok A dan kelompok B pada semua penelitian adalah -5.

Tabel 2.17.

Ilustrasi untuk membuktikan bahwa nilai  $p$  dan interval kepercayaan dipengaruhi oleh variasi

	Obat A	Obat B	P	Efek (IK95%)
	(n=50)	(n=50)		
Kualitas hidup (0-100)	75+-10	80+-10	0.0141	-5 (-9.99 sd 1.03 )
Kualitas hidup (0-100)	75+-20	80+-20	0.2143	-5 (-12.94 sd 2.94 )
Kualitas hidup (0-100)	75+-25	80+-25	0.3198	-5 ( -14,92 sd 4.92 )
Kualitas hidup (0-100)	75+-30	80+-30	0.4067	-5 (-16,91 sd 6.91 )

Perhatikan nilai  $p$ .

Dengan perbedaan efek yang sama, yaitu -5, nilai  $p$  semakin besar seiring dengan bertambahnya variasi, Perhatikan juga kolom interval kepercayaan.

Interval kepercayaan semakin lebar seiring dengan bertambah besarnya variasi.

Dengan demikian, terbukti bahwa statistic bergantung pada variasi data, dimana nilai  $p$  dan lebar IK berbanding lurus dengan variasi.

Nilai  $p$  & IK = variasi

Karena kelemahan ini, suatu penelitian bisa jadi secara statistic bermakna karena jumlah subjek yang terlalu besar atau variasi terlalu kecil.

Sebaliknya, suatu penelitian bisa jadi secara statistic tidak bermakna karena jumlah subjek sedikit atau variasi terlalu besar

# 11. Jelaskan apa yang dimaksud interpretasi hasil penelitian secara klinis!

---

Jawab :

Interpretasi hasil penelitian secara klinis adalah interpretasi yang didasarkan pada perbandingan antara efek yang diperoleh dalam penelitian dengan efek minimal yang diharapkan.

Penjelasan:

Interpretasi hasil penelitian secara klinis adalah interpretasi yang didasarkan pada perbandingan antara efek yang diperoleh dalam penelitian dengan efek minimal yang diharapkan.

Efek minimal yang diharapkan disebut efek size.

Secara klinis, hasil penelitian bermakna apabila efek hasil penelitian lebih besar atau sama dengan efek size.

Sebaliknya, hasil penelitian secara klinis tidak bermakna bila efek hasil penelitian lebih kecil dari efek size.

## Tabel 2.18.

### Perbandingan efek yang secara klinis diharapkan dengan efek yang diperoleh

	Efek yang diharapkan	Efek yang diperoleh
<b>Kesembuhan</b>	RR=2	RR=2
<b>Relaps</b>	RR=1,5	RR=1,17
<b>Kualitas hidup (0-100)</b>	Skor 10	Skor 15,0
<b>Lama rawat (hari)</b>	2 hari	0,5 hari

Perhatikan kembali table 2.18. Untuk keluaran kesembuhan, penelitian memperoleh RR sebesar 2 yang sama dengan efek size yang diharapkan peneliti.

Dengan demikian, efek keluaran relaps, penelitian memperoleh RR sebesar 1.17 yang lebih kecil daripada efek size yang diharapkan (1,5).

Dengan demikian, keluaran relaps secara klinis tidak bermakna.

Untuk keluaran skor kualitas hidup, penelitian memperoleh perbedaan rerata sebesar 15 yang lebih besar daripada efek size yang diharapkan (10).

Dengan demikian, keluaran skor kualitas hidup secara klinis bermakna.

Untuk keluaran lama rawat sebesar 0,5 hari yang lebih kecil daripada efek size yang diharapkan (2 hari).

Dengan demikian, keluaran lama rawat secara klinis tidak bermakna.

12. Jelaskan empat kemungkinan kombinasi interpretasi hasil penelitian !

---

**Jawab :**

Empat kemungkinan kombinasi interpretasi hasil penelitian :

Secara statistik bermakna, secara klinis bermakna

Secara statistik bermakna, secara klinis tidak bermakna

Secara statistik tidak bermakna, secara klinis bermakna

Secara statistik tidak bermakna, secara klinis tidak bermakna

	Efek yang diharapkan	Efek yang diperoleh
<b>Kesembuhan</b>	RR=2	RR=2
<b>Relaps</b>	RR=1,5	RR=1,17
<b>Kualitas hidup (0-100)</b>	Skor 10	Skor 15,0
<b>Lama rawat (hari)</b>	2 hari	0,5 hari

	P	Efek (IK95%)	Kesimpulan statistik	Kesimpulan Klinis
<b>Kesembuhan</b>	<0,001	2,00 (1,46 sd 2,74)	Bermakna	Bermakna
<b>Relaps</b>	0,521	1,17 (0,73 sd 1,87)	Tidak Bermakna	Tidak Bermakna
<b>Kualitas hidup (0-100)</b>	<0,001	-15,0 (-16,8 sd -13,2)	Bermakna	Bermakna
<b>Lama rawat (hari)</b>	0,001	0,5 (0,75 sd 1,25)	Bermakna	Tidak Bermakna
<b>Kadar kolesterol (g/dl)</b>	0,097	5,0 (-0,90 sd 10,9)	Tidak Bermakna	Tidak Bermakna

### Penjelasan :

Perhatikan kembali tabel 2.19.

Untuk keluaran kesembuhan dan skor kualitas hidup, interpretasi yang lengkap adalah secara statistik bermakna, secara klinis bermakna.

Selanjutnya, untuk relaps dan kolesterol, interpretasi yang lengkap adalah secara statistik tidak bermakna, secara klinis tidak bermakna.

Untuk keluaran lama rawat, interpretasi yang lengkap adalah secara statistik bermakna, secara klinis tidak bermakna.

Kombinasi interpretasi terhadap hasil penelitian, dengan demikian, ada empat variasi.

**TABEL 2.20**

KOMBINASI INTERPRETASI SECARA STATISTIK DAN SECARA KLINIS

		Klinis	
		Bermakna	Tidak Bermakna
Statistik	Bermakna	A	B
	Tidak Bermakna	C	D

- A. Secara statistik bermakna, secara klinis bermakna
- B. Secara statistik bermakna, secara klinis tidak bermakna
- C. Secara statistik tidak bermakna, secara klinis bermakna
- D. Secara statistik tidak bermakna, secara klinis tidak bermakna