

H. Risanto Siswosudarmo

UJI KLINIK SECARA RANDOM (UKR)

(RANDOMIZED CLINICAL TRIAL, RCT)



**Departemen Obstetrika dan Ginekologi
Fakultas Kedokteran UGM Yogyakarta**

BAB I. GAMBARAN UMUM

Tujuan:

Setelah menyelesaikan modul pertama ini peserta diharapkan mampu:

1. Mengenal jenis uji klinis secara random.
2. Mengenal sifat umum uji klinis secara random.
3. Mengidentifikasi unsur penting dalam uji klinis secara random.

Tahap pengujian sebuah obat.

I. Tahap praklinis

Sebelum suatu obat baru diujicobakan pada manusia, ia harus melewati uji coba laboratorium, baik secara *in-vitro* maupun *in vivo* pada binatang percobaan. Pada tahap ini uji coba suatu obat masih dalam tahap praklinis. Pada tahap ini akan diperoleh informasi tentang farmakologi dan toksikologi obat, yang meliputi:

- a. Efek farmakologi.
- b. Mekanisme dan titik tangkap kerja obat.
- c. Farmakokinetika, seperti absorpsi, distribusi, metabolisme, ekskresi dll.
- d. Dosis letal 50 (*LD50*)
- e. Efek teratogenik.
- f. Toksisitas.
- g. Kerusakan organ atau jaringan.
- h. Efek karsinogenik, dll.

Mengingat pada tahap ini tujuan pengujian sebuah obat begitu penuh bahaya dan risiko maka tahap ini hanya boleh dikerjakan di laboratorium atau menggunakan hewan coba.

II. Tahap Klinis

Sesudah melalui tahapan non klinis dan sebuah obat dinyatakan aman dan layak untuk diujicobakan pada manusia, maka obat tersebut baru boleh memasuki tahap uji coba pada manusia. Uji coba pada manusia inilah yang lazim disebut uji coba secara klinis (*clinical trial*). Karena sebuah obat yang aman pada binatang belum tentu aman bagi manusia, maka uji coba obat pada manusia juga harus melalui beberapa tahap. Uji coba sebuah obat baru pada manusia biasanya digolongkan pada empat fase, yaitu:

1. Uji Klinis Fase I

Ini adalah tahap awal penilaian suatu obat pada manusia. Tujuan utama adalah untuk mengetahui adanya efek samping yang tidak muncul pada hewan coba. Tujuan lain adalah untuk mengetahui apakah obat tersebut dapat diterima atau tidak oleh tubuh (**tolerated**). Di samping itu, juga untuk mengetahui bagaimana proses metabolisme dan **bioavailability** obat tersebut pada manusia. Hal yang harus diketahui misalnya adalah di mana absorpsinya dan berapa kecepatannya, bagaimana metabolismenya, ke mana distribusinya, lewat apa ekskresinya dan berapa kecepatannya. Dengan demikian pada uji coba klinis fase I yang penting adalah keamanan (**safety**) suatu obat baru dan bukan keampuhannya (**efficacy**) atau kemanjurannya (**effectiveness**). Subjek biasanya adalah relawan yang diberi obat dengan dosis bertingkat untuk mengetahui sejauh mana dosis obat tersebut dapat diterima. Biasanya, dosis yang diberikan pertama kali adalah 1/50 dosis minimal pada hewan coba yang masih memberikan efek, kemudian dinaikkan dengan kelipatan 1,5 atau 2 kalinya, sehingga muncul adanya efek farmakologik atau efek samping. Uji klinis fase I biasanya membutuhkan paling banyak 80 relawan.

Contoh 1.

Pada tahun 1984 dilakukan sebuah studi yang melibatkan 10 perempuan. Mereka diberi 250 mg kuinakrin hidroklorida dalam bentuk pelet yang diinsersikan kedalam tuba uterina melalui kavum uteri. Histerektomi dikerjakan 24 jam sesudahnya untuk mengetahui perubahan histologis pada tuba sampai seberapa jauh penutupan lumen tuba terjadi. Dihitung pula jumlah obat yang diabsorpsi selama periode 24 jam tersebut, efek samping yang timbul dan derajat keseriusannya. (Sumber: Family Health Internasional. A 24 Hours test of quinacrine. Internal Report, 1984).

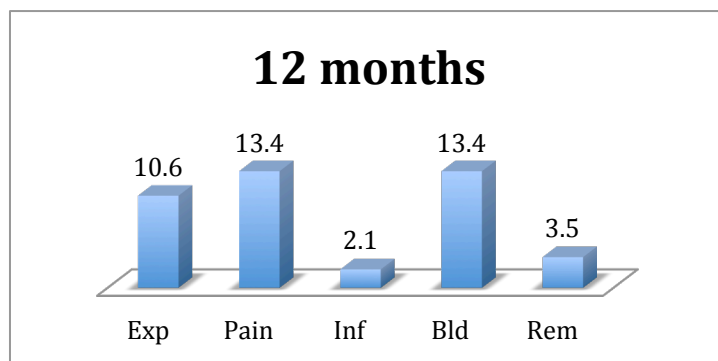
Uji klinis fase II

Pada fase ini yang penting adalah menilai efek terapeutik suatu obat, yaitu **keampuhan (efficacy)** atau **kemanjurannya (effectiveness)** di samping keamanannya (**safety**). Uji klinis fase II merupakan uji untuk melakukan penapisan (**screening**) yaitu untuk memilih jenis obat yang terbaik dari beberapa jenis obat yang ditawarkan. Pada fase ini ditentukan cara pemberian obat dan variasi dosis yang paling optimal. Yang terpenting pada fase ini adalah bahwa obat diberikan kepada sejumlah pasien tanpa

disertai kelompok kontrol, sehingga penelitian ini berupa penelitian deskriptif berupa seri kasus saja (**case series**). Oleh karena tanpa pembandingan (kelompok kontrol) kesimpulan yang dapat diambil dari fase II ini paling banter adalah bahwa obat ini mempunyai daya sembuh sekian persen. Kita belum bisa mengatakan bahwa obat ini lebih baik dari obat sebelumnya atau dari obat standar. Jumlah subyek pada fase II ini berkisar antara 100 sampai 200 orang.

Contoh 2.

Pada tahun 2011, sebuah alat untuk insersi IUD CuT 380A pascasalin dikembangkan, dan diberi nama R-inserter. Inserter ini berbeda dengan inserter yang tersedia di pasar dalam hal panjangnya (dari 19 cm menjadi 28 cm). Dengan memperpanjang inserter maka pemasangan IUD bisa dilakukan secara baku yakni **no touch and withdrawal**. Dengan cara pemasangan seperti itu maka kemungkinan infeksi dapat ditekan. Keprihatinan lain pada pemasangan IUD pascasalin adalah angka ekspulsi yang tinggi, yang bervariasi dari 4% sampai 44%. Inserter baru diuji cobakan pada 142 ibu-ibu pascasalin yang menginginkan IUD sebagai alat kontrasepsi yang memenuhi kriteria kelayakan (inklusi dan eksklusi). *Follow up* dikerjakan setiap bulan sampai 12 bulan. Hasilnya adalah sebagai berikut:



Exp = expulsion; Inf = infection; Bld = Bleeding; Rem = Removal (Siswosudarmo et al., 2014).

3. Uji klinis fase III

Setelah sebuah obat dinyatakan efektif, ia harus dibandingkan dengan obat standar yang lazim digunakan. Dengan demikian akan terlihat kemampuan relatifnya. Pasien harus memenuhi jumlah tertentu, yang biasanya cukup besar. Jumlah sampel sebaiknya dihitung untuk memenuhi tingkat kepercayaan dan kekuatan penelitian yang diinginkan. Pada uji klinis fase III ini yang terpenting adalah alokasi subyek ke dalam kelompok penelitian dan kontrol harus dilakukan secara random, artinya setiap subyek penelitian mendapat kesempatan yang sama untuk masuk ke dalam kelompok uji atau kelompok. Uji klinis fase III inilah yang lazim disebut dengan uji klinis secara random (UKR) atau ***randomized clinical trial (RCT)***. Karena pengelompokan subyek penelitian selalu dilakukan secara random, dan selalu menggunakan kelompok kontrol sebagai pembandingan, maka istilah RCT bisa juga berarti ***randomized controlled trial***. Dalam bahasa Indonesia, Penulis mengusulkan istilah **uji klinis secara random**, atau disingkat **UKR**.

Contoh 3

Setelah uji klinis fase II pada inserter baru (*R-inserter*) untuk IUD pascasalin dinyatakan layak digunakan (angka ekspulsi, infeksi dan keluhan lain tidak lebih tinggi dari yang disebut dalam kepustakaan), maka uji klinis fase III dikerjakan. Pada tahap ini *R-inserter* dibandingkan dengan cara pemasangan standar yakni menggunakan klem cincin (*ring forceps*). Sebanyak 208 subyek penelitian yang layak dialokasikan secara random ke dalam kelompok uji (*R-inserter* 104 orang) dan kelompok control (*ring forceps*, 104 orang). Hasilnya adalah sebagai berikut (Siswudarmo et al., 2016).

Twelve months cumulative events rate based on mode of insertion

Variables	Events Rate		%	RR (95% CI)	P value
	Yes	No			
Infection					
<i>R-inserter</i>	1	102	1 %	1.01 (0.06 – 15.93)	0.99
Ring forceps	1	103	1 %	1	
Expulsion					
<i>R-inserter</i>	5	98	4.9 %	1.26 (0.35 – 4.57)	0.72
Ring forceps	4	100	3.8 %	1	
Pain					
<i>R-Inserter</i>	8	95	7,8	0,62 (0,27 – 1,44)	0,26
Ring forceps	13	91	12,5	1	
Bleeding					
<i>R-inserter</i>	5	98	4.9 %	2.52 (0.50 – 12.72)	0.24
Ring forceps	2	102	1.9 %	1	
Removal					
<i>R-inserter</i>	4	99	3.9 %	0.67 (0.20 – 2.32)	0.53
Ring forceps	6	98	5.8 %	1	

4. Uji klinis fase IV

Setelah sebuah obat atau alat baru atau pengobatan baru diterima secara luas dan telah dipasarkan, mereka masih memerlukan pengawasan lebih lanjut untuk melihat efek samping jangka panjang, morbiditas dan atau mortalitasnya. Biasanya dilakukan dalam jangka lima tahun atau lebih. Uji klinis fase IV sering disebut juga pengawasan pasca pemasaran (*post marketing surveillance*).

Contoh 4.

Setelah pil kontrasepsi diterima secara luas dimasyarakat, timbul pertanyaan apakah perempuan yang memakai pil kontrasepsi mempunyai risiko kematian karena penyakit kardiovaskular yang berbeda dibandingkan dengan mereka yang tidak memakainya? Untuk menjawab pertanyaan ini diperlukan sebuah studi kohort yang melibatkan kelompok perempuan yang masih aktif melakukan hubungan seks. Satu kelompok terdiri atas perempuan yang menggunakan pil dan kelompok yang lain adalah perempuan yang tidak menggunakan pil. Pada

studi ini pengelompokan perempuan kedalam dua golongan tidak dilakukan secara random. (Sumber: Royal College of General Practitioners. Oral Contraceptives and Health, Pitman, 1974).

Gambaran umum sebuah uji klinis secara random (UKR)

Dalam UKR subjek dibagi dalam dua kelompok secara random. Kelompok yang mendapat mendapat pengobatan baru disebut kelompok perlakuan. Sedang kelompok lain yang mendapat pengobatan standar disebut kelompok kontrol. Kedua kelompok harus diikuti sejak awal pengobatan untuk mengetahui apakah efek yang terjadi berbeda. Saat mereka masuk penelitian dapat dianggap sebagai waktu yang sama yaitu waktu nol atau *base line*.

Kelompok perlakuan dan kelompok kontrol.

Kedua kelompok harus sebanding dalam semua aspek, yakni variabel luar (*confounding variables*) yang berpotensi mempengaruhi hasil yang diharapkan. Dengan demikian apabila mereka mendapat perlakuan dan menunjukkan adanya perbedaan hasil, maka perbedaan hasil tersebut benar-benar terjadi akibat pengaruh perlakuan bukan karena pengaruh variabel luar. Lazimnya dalam sebuah UKR kelompok perlakuan adalah mereka yang mendapat obat baru, atau cara pengobatan baru, atau variasi dosis baru. Kelompok kontrol adalah mereka yang mendapat obat standar, cara pengobatan standar atau dosis standar yang sudah lazim diberikan. Bila obat standar tidak ada, maka kelompok kontrol boleh terdiri dengan subyek yang mendapat plasebo. Kadang-kadang dua buah obat baru perlu dibandingkan sehingga penetapan mana kelompok kontrol dan mana perlakuan adalah sembarang.

Tidak semua UKR dapat menggunakan plasebo sebagai kontrol. Sebagai contoh pada penelitian obat kontrasepsi misalnya, pemakaian plasebo sebagai kontrol adalah tidak etis karena semua pasien yang masuk penelitian adalah mereka yang menghendaki adanya usaha pencegahan kehamilan. Demikian juga penelitian tentang keampuhan sebuah antibiotika baru harus dibandingkan dengan antibiotika standar karena sangat tidak etis kalau seseorang dengan diagnosis infeksi tertentu hanya diberikan plasebo, karena penyakit bisa bertambah parah. Hal serupa juga terjadi kalau kita melakukan uji coba obat antihipertensi.

UKR dapat pula membandingkan lebih dari dua kelompok, dan salah satunya dipakai sebagai kelompok kontrol. Contoh dalam penelitian kontrasepsi adalah uji untuk membandingkan beberapa jenis IUD seperti ML 250, CuT 250, CuT 380A dan Lippes Loop. Sebagai IUD standar selalu dipakai Lippes karena ia adalah IUD yang telah dipakai secara luas dan telah banyak dilakukan penelitiannya. Uji yang demikian menjadi sangat kompleks sehingga sebaiknya dihindari. Penulis menganjurkan untuk membuat penelitian dengan dua kelompok saja. Bila lebih, maka ambil salah satu sebagai kelompok pembandingnya.

Pada binatang percobaan, peneliti dapat dengan bebas melakukan pengendalian atas variabel pengacau secara ketat, tetapi tidak demikian halnya dengan penelitian yang menggunakan manusia sebagai subjek. Peneliti hanya dapat menganjurkan agar subjek penelitian menghindari obat-obat tertentu yang mengganggu hasil penelitian. Dengan demikian kita tidak mungkin memperoleh kelompok perlakuan murni dan kelompok kontrol murni, sehingga yang dibandingkan sebenarnya bukanlah pengobatan, tetapi hanya strategi pengobatan.

Semua subjek yang masuk dalam penelitian harus dianalisis pada akhir penelitian (*intention to treat*). Sebaiknya sebelum penelitian dimulai harus sudah diketahui dahulu kesulitan apa yang bakal dihadapi, cara mengatasinya dan hasil apa yang nanti akan diperoleh. Salah satu cara mengantisipasi jumlah sampel yang mungkin dropped out adalah dengan memperbesar jumlahnya (*inflating factor*). Pada umumnya *inflating factor* adalah 10%.

Protokol penelitian

Penelitian biasanya timbul dari sebuah pertanyaan penelitian (*research problem, research question*) yang merupakan kesenjangan antara apa yang ada dan apa yang seharusnya ada. Inilah yang dikenal dengan masalah yang akan diteliti untuk dicarikan jalan pemecahannya. Masalah ini kemudian dituangkan kedalam sebuah hipotesis spesifik yang selanjutnya diterjemahkan ke dalam langkah-langkah penelitian mulai dari tujuan sampai cara analisis datanya. Semua langkah diatas harus tertulis dalam sebuah dokumen yang lazim disebut protokol penelitian.

Dengan demikian protokol penelitian menggambarkan rencana penelitian dan berfungsi sebagai petunjuk pelaksanaan, rambu-rambu, dan sebagai alat komunikasi bagi semua personel yang terlibat dalam penelitian. Dapat pula dikatakan bahwa

protokol penelitian adalah perjanjian tertulis antara peneliti, subjek, dan masyarakat ilmiah.

Sebuah protokol penelitian setidaknya-tidaknyanya memuat hal-hal sebagai berikut:

1. Judul penelitian
2. Latar belakang dan rumusan masalah (pertanyaan penelitian).
3. Tujuan Penelitian
4. Manfaat penelitian
5. Hipotesis
6. Rancangan penelitian
7. Populasi dan Subyek Penelitian
 - a. Populasi penelitian
 - b. Subyek penelitian (Kriteria inklusi dan eksklusi, Besar sampel, Cara pengambilan sampel).
 - c. Cara randomisasi.
 - d. Cara penyamaran (blinding, bila diperlukan).
8. Tempat dan waktu (lama dan jadwal penelitian).
9. Bahan, obat dan alat.
10. Cara pengumpulan data.
11. Cara penilaian hasil.
12. Rencana analisis (jenis statistika).
13. Lampiran:
 - a. Borang penelitian.
 - b. Tim peneliti.
 - c. Kelayakan etika dan *informed consent*.
 - d. Ijin penelitian
 - e. *Sponsorship* dan rencana anggaran (bila diperlukan), dll.

Soal

I. Tulislah B (benar) atau S (salah)!

1. Pada uji klinis secara random, hanya strategi pengobatan yang dapat dibandingkan.
2. Subjek yang tidak mematuhi aturan dalam protokol harus dikeluarkan dari analisis supaya hasil penelitian benar.
3. Jumlah subjek yang sekiranya tidak mematuhi protocol sebaiknya diperkirakan sebelum penelitian dimulai.
4. Dalam sebuah UKR sebaiknya dibandingkan dua pengobatan yang berbeda.
5. Dalam sebuah UKR sebaiknya dibandingkan tiga kelompok atau lebih.
6. Dalam sebuah ujicoba antibiotika, kelompok kontrol sebaiknya mendapat plasebo.
7. Alokasi kedalam kelompok uji dan penelitian sebaiknya dilakukan secara bergantian sesuai nomer urut.
8. Kelompok kontrol umumnya adalah pengobatan standard.
9. Kelompok perlakuan harus selalu pengobatan yang lebih baru.
10. Alokasi random menjamin kedua kelompok seimbang.

II. Pilih satu yang paling benar.

1. Uji yang melakukan penyaringan terhadap pengobatan yang baik, menentukan cara pemberian obat dan menilai kemampuan potensial adalah:
 - A. Uji pra klinis.
 - B. Uji klinis fase I.
 - C. Uji klinis fase II.
 - D. Uji klinis fase III.
 - E. Uji klinis fase IV.
2. Yang mana dari uji klinis tersebut dibawah ini yang lazim disebut uji klinis secara random?
 - A. Uji pra klinis.
 - B. Uji klinis fase I.
 - C. Uji klinis fase II.
 - D. Uji klinis fase III.
 - E. Uji klinis fase IV.

3. Dalam suatu UKR, kelompok pengobatan baru biasanya dibandingkan dengan kelompok kontrol. Berikut adalah kelompok yang dapat dipakai sebagai kelompok kontrol, KECUALI:
- A. Plasebo.
 - B. Tanpa pengobatan sama sekali.
 - C. Obat standar.
 - D. Obat lain yang juga sama-sama baru.
 - E. Obat sejenis dengan dosis yang berbeda.
4. Saudara ingin melihat apakah asam mefenamat dapat mengurangi rasa sakit pada dismenorea. Saudara mengumpulkan 100 yang menderita dismenorea dan mendapat 500 mg asam mefenamat. Dari jumlah tersebut 90 orang mengatakan sakitnya reda. Penelitian ini masuk jenis uji yang mana?
- A. Uji praklinis.
 - B. Uji klinis fase 1.
 - C. Uji klinis fase 2.
 - D. Uji klinis fase 3.
 - E. Uji klinis fase 4.
5. Saudara ingin melihat apakah penggunaan pil KB lebih sedikit memberikan angka kehamilan dibanding IUD. Jenis penelitian uji klinis apa yang paling tepat?
- A. Uji praklinis.
 - B. Uji klinis Fase 1.
 - C. Uji klinis Fase 2.
 - D. Uji klinis Fase 3.
 - E. Uji klinis Fase 4.

III. Cocokkan pernyataan disebelah atas ini dengan uraian disebelah bawah

- A. Uji praklinis
 - B. Uji klinis fase 1
 - C. Uji klinis fase 2
 - D. Uji klinis fase 3
 - E. Uji klinis fase 4
1. Kelinci disuntik dengan progesteron, dikawinkan dan diikuti ada tidaknya kehamilan.
 2. Sebuah studi kecil membandingkan pemakaian salah satu dari dua dosis progestin untuk mengetahui kemampuan potensialnya.

3. Follow up terhadap sejumlah besar perempuan untuk mengetahui efek jangka panjang dari pemakaian pil kontrasepsi.
4. Lima puluh perempuan diberi salah satu 3 progestin, lalu dimonitor perubahan dalam kimia tubuh dan jaringan. Dimonitor juga apakah ovulasi dapat ditekan.
5. Dua ribu perempuan dibagi dalam dua kelompok, masing-masing diberi pil kontrasepsi dosis standar dan pil kontrasepsi dosis rendah. Mereka di-*follow up* selama satu tahun untuk membandingkan keampuhan relatif dari kedua jenis obat tersebut.

BAB II. SUBYEK PENELITIAN

Tujuan:

Setelah mengikuti modul ini peserta diharapkan mampu:

1. Mengidentifikasi populasi penelitian dan subyek penelitian
2. Menetapkan *eligibility criteria*, kriteria inklusi dan eksklusi
3. Menetapkan kriteria diagnosis

Suatu penelitian klinis dimaksudkan tidak lain adalah agar hasilnya dapat diterapkan secara luas dalam populasi umum. Oleh karena itu subyek dalam penelitian juga harus mencerminkan atau mewakili subyek dalam masyarakat pada umumnya. Sudah barang tentu bahwa ciri subyek dalam masyarakat adalah sangat heterogen, sehingga tidak mungkin subyek dalam penelitian diambil begitu saja dari subyek dalam masyarakat. Hanya subyek yang memenuhi kriteria kelayakan (*eligibility criteria*) sajalah yang dapat masuk sebagai subyek atau populasi penelitian. Selanjutnya tidak semua pasien yang layak mesti masuk dalam penelitian tetapi cukup sejumlah tertentu saja yang dibutuhkan yang secara statistika dianggap cukup mewakili. Bagian dari populasi penelitian yang dibatasi oleh *sample size* inilah yang disebut sampel penelitian.

Secara umum tujuan pemilihan subyek ini adalah untuk memperoleh jawaban yang benar baik bila terbukti ada perbedaan antara dua macam perlakuan yang diberikan atau bila terbukti tidak ada perbedaan (kepentingan metodologis). Tujuan lain adalah untuk menghindari adanya pasien yang berisiko tinggi terhadap perlakuan dalam arti penelitian tidak boleh mengakibatkan efek samping yang sangat serius bahkan fatal. Sebaliknya pasien yang akan mendapatkan plasebo harus dipertimbangkan betul bahwa plasebo tidak merugikan dirinya seperti misalnya penyakitnya tidak sembuh atau terjadi hal yang tidak diinginkan (kepentingan klinis). Contoh di mana plasebo tidak boleh digunakan adalah penelitian obat kontrasepsi, sitostatika, antibiotika, antihipertensi, anti asma dll.

Pemilihan pasien juga harus memenuhi kepentingan **etika** artinya peneliti harus menerangkan manfaat dan mudarat keikutsertaan pasien dalam penelitian. Sebagai contoh dalam penelitian obat kontrasepsi, plasebo sama sekali tidak boleh dipakai karena orang yang masuk dalam penelitian adalah mereka yang menginginkan kontrasepsi. Kalau ini terjadi maka peneliti dapat dituntut karena tidak jujur dalam memberikan *informed consent* kepada pasien. Demikian juga pasien harus orang yang betul betul memahami bahwa dirinya masuk dalam penelitian dengan segala risiko yang

mungkin akan terjadi. Penelitian pada orang yang kurang atau tidak kompeten harus melibatkan mereka yang bertanggung jawab (misal bayi dan anak, harus melibatkan orang tua atau walinya)

Sebuah penelitian harus secara jelas menyebutkan:

1. Siapa yang menjadi subyek penelitian (*eligibility criteria*).
2. Bagaimana mereka dipilih dan dikelompokkan (*random allocation*).
3. Apa kriteria hasil dan bagaimana peneliti melakukan penilaian terhadap hasil (*blinding*). Penting juga menyebutkan siapa yang melakukan penilaian terhadap hasil penelitian. Apakah penilai (observer) terdiri atas seorang saja atau lebih dari seorang.
4. Bagaimana peneliti membagi obat atau perlakuan (*concealment*).
5. Berapa jumlah sampelnya (*sample size*).

Pertanyaan pertama dimaksudkan untuk melihat pada kelompok mana sebenarnya suatu pengobatan itu ditujukan. Subyek yang sangat heterogen menyusahkan peneliti mengambil kesimpulan dari hasil penelitiannya oleh karena: (1) Subyek yang sangat heterogen akan mempengaruhi komparabilitas kelompok yang akan dibandingkan, (2) Respons pengobatan terhadap heterogenitas ini sangat berbeda.

Untuk menjawab pertanyaan ini perlu dibuat kriteria kelayakan (*eligibility criteria*), yakni populasi yang telah dibatasi dengan kriteria inklusi dan eksklusi. Variabilitas subyek biasanya menyangkut beberapa faktor seperti: (1) Berat ringannya penyakit, (2) Ciri biologis dan demografis seperti umur, jenis kelamin, pendidikan, tempat tinggal, pekerjaan, berat badan, tinggi badan, dll. Untuk menilai berat ringannya penyakit harus disebutkan dengan jelas cara dan kriteria diagnosis yang akan memudahkan para peneliti untuk menentukan pasien mana yang dapat masuk ke dalam penelitian.

Pertanyaan kedua dimaksudkan untuk menghindari pembagian kelompok yang tidak seimbang atau agar pembagian subyek ke dalam kedua kelompok benar-benar sebanding dalam semua hal terutama variabel luar (*confounding variables*) yang berpotensi mengganggu hasil. Karena dalam sebuah UKR subyek masuk ke dalam penelitian secara berurutan, maka randomisasi harus dikerjakan pada saat melakukan alokasi subyek ke dalam kelompok, baik kelompok perlakuan maupun kelompok kontrol. Dengan alokasi secara random maka kemungkinan subyek masuk ke dalam kelompok uji dan kelompok kontrol adalah sama, sehingga semua variabel luar yang berpotensi mengganggu hasil telah terdistribusi secara ekuil. Jadi randomisasi menjamin

terhindarnya bias pengganggu (*confounding bias*). Karena ada berbagai cara alokasi random maka cara randomisasinya juga harus disebutkan secara jelas (lihat bab randomisasi) apakah dengan tabel bilangan random atau dengan bilangan random yang dihasilkan oleh komputer (*computer generated random numbers*).

Pertanyaan ketiga dimaksudkan untuk keseragaman dalam menilai ada tidaknya suatu hasil. Bila penilai (*observer*) lebih dari satu orang maka mereka harus mempunyai keseragaman atau kriteria yang sama dalam menentukan adanya hasil tersebut. Agar peneliti atau penilai tidak cenderung memihak, maka peneliti atau penilai harus tidak mengetahui dari kelompok mana pasien yang dinilai berasal. Cara demikian dikenal dengan penyamaran (*blinding*). Jadi dengan adanya *blinding* maka **bias penilaian** atau **bias pengukuran** (*measurement bias*) bisa dihindari.

Pertanyaan keempat adalah usaha untuk menghindari bias pemilihan (***selection bias***) yaitu pemilihan pasien yang cenderung memihak (misal hanya memasukkan yang ringan saja atau sebaliknya yang berat saja). Dengan *concealment* (amplop tertutup yang tidak tembus pandang) maka peneliti tidak mengetahui apakah pasien A mendapat obat X atau obat Y, dan demikian pula seterusnya.

Pertanyaan kelima mempengaruhi kekuatan penelitian (***power of the study***) yang penting dalam perhitungan statistika. Para ahli epidemiologi biasanya sangat berkepentingan dengan besar sampel ini. Makin besar kekuatan penelitian yang kita inginkan, makin besar pula jumlah sampel yang diperlukan. Akhirnya semua yang disebut di atas ditambah dengan bentuk rancangan penelitian adalah penting karena selain menentukan mutu penelitian yang kita kerjakan juga memungkinkan peneliti lain melakukan evaluasi atas hasil penelitian kita.

Kriteria kelayakan (*eligibility criteria*)

Kriteria kelayakan adalah kriteria tertentu yang membatasi subyek agar mereka homogen. Di samping itu kriteria ini juga membatasi nilai ekstrim tertentu yang sangat mengganggu, seperti misalnya umur lanjut, stadium terminal, dll. Keikut-sertaan pasien yang berulang-ulang dalam suatu uji klinis sebaiknya juga dihindari. Adanya batasan kriteria kelayakan yang jelas sangat memudahkan para peneliti untuk merekrut pasien dalam sebuah penelitian multisenter.

Penentuan kriteria kelayakan hendaknya realistis artinya betul mewakili populasi yang ada dalam masyarakat sehingga hasil penelitian dapat digeneralisasikan dan diterapkan secara luas dalam populasi umum. Pembatasan kriteria kelayakan yang

terlalu ketat memang menghasilkan populasi penelitian yang sangat homogen sehingga secara metodologis hasilnya mungkin sangat memuaskan. Hasil penelitian semacam ini dikatakan mempunyai validitas dalam (***internal validity***) yang sangat kuat tetapi sangat mungkin sukar diterapkan dalam populasi karena populasi penelitian tidak mewakili populasi umum. Dengan demikian penelitian yang sangat ketat dalam pembatasan kriteria kelayakan hanya berarti bagi penelitian itu sendiri bukan untuk masyarakat luas. Penelitian semacam ini dikatakan mempunyai validitas dalam yang kuat tetapi validitas luar (***external validity***) yang rendah. Sebaliknya penelitian yang tidak mempunyai batasan kriteria kelayakan yang jelas hasilnya diragukan.

Pada dasarnya kriteria kelayakan terdiri atas dua unsur yaitu kriteria inklusi (subyek yang bisa masuk) dan kriteria eksklusi (subyek yang tidak bisa masuk). Di samping itu, kalau menyangkut sebuah penyakit maka untuk memperoleh keseragaman, perlu pula adanya kriteria diagnosis yang jelas.

Kriteria inklusi adalah satu batasan yang menyebutkan syarat tertentu yang membolehkan pasien masuk dalam penelitian yang meliputi:

1. Ciri biologis dan demografis seperti umur, jenis kelamin, berat badan, tinggi badan, pendidikan, pekerjaan, tempat tinggal, dll.
2. Ciri penyakit, tanda dan lama gejala diderita, hasil laboratorium, stadium penyakit, dll.

Kriteria eksklusi adalah batasan tertentu yang melarang pasien masuk ke dalam penelitian yang meliputi:

1. Tidak memenuhi kriteria inklusi.
2. Terdapat kontraindikasi dengan obat yang akan diuji.
3. Menderita penyakit lain yang dapat mengganggu hasil atau sedang dalam pengobatan untuk penyakit yang lain.
4. Kehamilan (umumnya selalu), kecuali memang uji obat untuk perempuan hamil.
5. Subyek yang sangat mungkin tidak dapat mengikuti jadwal *follow-up* dengan baik, misalnya rumah jauh, bukan penduduk tetap, dll.

Catatan: kriteria eksklusi bukan lawan kriteria inklusi, sehingga kalau sudah disebut dalam kriteria inklusi jangan disebut kembali dalam kriteria inklusi hanya dengan menyebut sesuatu yang berlawanan. Misal bila disebut dalam kriteria inklusi batasan umur adalah antara 20 sampai 40 tahun, maka jangan menyebut dalam kriteria eksklusi adalah mereka yang berumur kurang dari 20 tahun dan lebih dari 40 tahun.

Untuk jelasnya diberikan beberapa contoh sebagai berikut:

Contoh 6.

Seorang peneliti ingin melakukan uji klinis secara random untuk menilai apakah risiko terkena infeksi panggul pada pemakaian IUD tanpa benang (sebagai kelompok perlakuan) berbeda dibandingkan dengan IUD dengan benang (sebagai kelompok kontrol). Penelitian dilakukan selama kurun waktu satu tahun.

Kriteria inklusi

Populasi penelitian adalah perempuan sehat yang ingin menggunakan IUD sebagai alat kontrasepsi dan bersedia melakukan *follow-up* selama 12 bulan.

Masing-masing harus memenuhi:

1. Berumur antara 20-40 tahun.
2. Belum pernah menggunakan IUD.
3. Bersuami dan aktif melakukan hubungan seksual.
4. Dalam masa interval atau sekurang-kurangnya 42 hari pascasalin.
5. Tidak ada riwayat radang panggul sebelumnya.

Kriteria eksklusi

Perempuan dengan kondisi sebagai berikut tidak dimasukkan ke dalam penelitian:

1. Abnormalitas uterus yang mengakibatkan kelainan rongga uterus dan kanalis servikalis (misal hiperretrofleksi, mioma uteri, dll).
2. Terdapat massa tumor di adnexa.
3. Perdarahan uterus yang abnormal, menstruasi yang tidak teratur, riwayat hiperplasia endometrii, polips dan karsinoma.
4. Riwayat endometritis pascasalin atau abortus infeksius.
5. Sedang dalam pengobatan antibiotika atau riwayat menggunakan antibiotika dalam dua minggu terakhir.
6. Sedang menderita radang panggul.
 - a. Panas $> 38^{\circ}\text{C}$.
 - b. Nyeri tekan di daerah pelvis.

- c. *Discharge* vagina purulent, dan/atau pembesaran kelenjar Bartholini.

Catatan:

1. Sebelum IUD dipasang, dilakukan anamnesis, pemeriksaan fisik, pemeriksaan dalam, kultur gonokokus dari serviks dan kripte anus dan kultur klamidia dari serviks.
2. Vaginitis dan servitis yang ringan tidak termasuk dalam penyakit radang panggul.

Contoh 7

Seorang peneliti ingin mengetahui apakah penambahan asam klavulanat pada amoksisilin menambah keefektifan obat tersebut dalam pengobatan penyakit radang panggul (PRP).

Kriteria inklusi

1. Perempuan berumur 15 – 40 tahun yang menderita penyakit radang panggul akut atau kronis yang datang ke rumah sakit.
2. Bersedia mengikuti protokol penelitian dengan baik termasuk menandatangani informed consent, dll

Kriteria eksklusi

1. Sedang dalam pengobatan antibiotika atau pernah mendapat antibiotika dalam dua minggu terakhir ini untuk sebab apapun.
2. Memerlukan antibiotika lain sebagai obat tambahan maupun atas indikasi penyakit lain yang menyertainya.
3. Hamil.
4. Terdapat kontraindikasi terhadap obat.
5. Kultur *discharge* vagina menunjukkan adanya jamur.
6. Herpes genitalis.
7. Akseptor IUD.

Kriteria diagnosis

1. Penyakit radang panggul akut dikatakan ada bila terdapat sekurang-kurangnya tiga dari lima tanda klinis dan satu tanda laboratoris sebagai berikut:
 - a. Tanda dan gejala klinis: temperatur > 38°C, sakit perut bagian bawah, terdapat discharge vagina yang berbau busuk, sakit pada pemeriksaan dalam

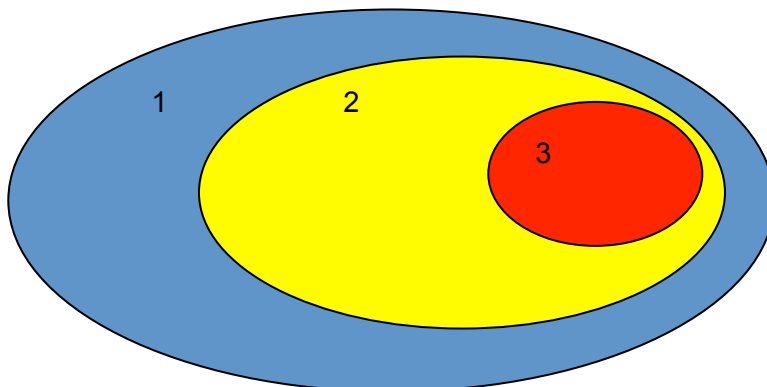
yaitu di daerah uterus, parametrium dan adneksa disertai atau tidak dengan gejala **muscular defense**, dan hiperemia vagina, serviks dengan atau tanpa pembesaran kelenjar Bartholini.

- b. Tanda laboratorium: KED > 20 mm/jam (Wintrobe) atau > 30 mm/jam (Westergreen), angka leukosit > 10.000/dL
 - c. Kultur sekret vagina menunjukkan adanya bakteri.
2. Penyakit radang panggul kronis ditegakkan berdasarkan adanya 5 temuan dari tujuh tanda dan gejala sebagai berikut:
- a. Terdapat riwayat PRP yang berulang.
 - b. Rasa tidak enak di perut bagian bawah.
 - c. *Discharge* vagina yang berbau busuk.
 - d. Rasa sakit yang kambuh-kambuhan di perut bagian bawah.
 - e. Rasa nyeri pada saat serviks digerakkan.
 - f. Parametrium kaku dan pendek, disertai nyeri atau tidak pada penekanan.
 - g. Terdapat massa tumor atau infiltrat di adneksa yang sakit tekan.

Cara memilih subyek

Subyek yang diambil dari populasi penelitian biasanya dilakukan tanpa random. Umumnya mereka diambil dari rumah sakit, klinik atau praktek pribadi yang telah memenuhi kriteria kelayakan. Alokasi random menjadi kelompok perlakuan dan kontrol dilakukan setelah subyek memenuhi kriteria kelayakan dan bersedia mengikutinya. Dengan demikian peneliti masih menghadapi problema yang menyangkut generalisasi dari subyek yang ada dalam penelitian ke dalam populasi penelitian dan selanjutnya ke populasi sesungguhnya yang lebih luas.

Secara umum pemilihan subyek penelitian dapat dilukiskan pada skema sebagai berikut:



1. Populasi umum (area biru) yakni di mana penelitian dikerjakan dan subyek akan direkrut. Belum diberikan kriteria pembatasan.
2. Populasi penelitian (area kuning), yakni bagian dari populasi umum yang telah dibatasi oleh kriteria kelayakan yakni inklusi dan eksklusi
3. Sampel penelitian (area merah), yakni subyek yang layak (*eligible*) yang telah dibatasi oleh *sample size*. Pada populasi inilah penelitian dikerjakan.

BAB III. BENTUK RANCANGAN PENELITIAN (*STUDY DESIGN*)

Tujuan:

Setelah menyelesaikan modul ini anda diharapkan mampu:

1. Mengetahui beberapa bentuk rancangan pada sebuah UKR.
2. Mengetahui kelebihan dan kekurangan setiap bentuk rancangan.
3. Memilih bentuk rancangan tertentu sesuai dengan tujuan penelitian anda.

Rancangan Dasar

Pemilihan bentuk rancangan yang akan dipakai dalam sebuah uji klinis tergantung pada beberapa hal seperti jenis perlakuan, obat yang akan diberikan, jenis penyakit, *outcome* yang dinilai, tersedianya subyek yang cukup, dana yang tersedia, waktu yang diperlukan dan sebagainya. Pada dasarnya dalam uji klinis fase III ini peneliti ingin membandingkan antara dua perlakuan misalnya obat A dengan obat B atau obat A dengan plasebo. Oleh karena itu yang terpenting dalam membuat rancangan adalah tersedianya dua kelompok yang akan dibandingkan.

Ada tiga jenis rancangan dasar yaitu (1) rancangan paralel, (2) rancangan silang (***cross over***) dan (3) rancangan pasangan. Dalam praktik yang paling populer dan banyak dipakai hanyalah bentuk pertama yakni **rancangan paralel**.

1. Rancangan paralel

Merupakan rancangan dasar UKR yang paling banyak dipakai. Nama lain adalah rancangan antar subyek (*between subject trial*). Pada bentuk ini setiap kelompok memperoleh perlakuannya sendiri-sendiri lalu dibandingkan hasilnya. Sebagai contoh jika kita ingin membandingkan kemanfaatan obat A dan obat B maka dilakukan langkah-langkah sebagai berikut:

Mula-mula seluruh subyek yang telah memenuhi kriteria kelayakan dialokasikan secara random ke dalam dua kelompok yakni kelompok perlakuan dan kelompok kontrol. Kelompok pertama akan menerima obat A saja sebagai kelompok perlakuan sedangkan kelompok kedua menerima obat B saja sebagai kelompok kontrol.

Penempatan subyek secara random (***random allocation***, atau ***random assignment***) ke dalam salah satu kelompok dimaksudkan agar setiap individu mempunyai kesempatan yang sama untuk mendapatkan baik obat A ataupun obat B. Dengan demikian distribusi cirri-ciri variabel yang relevan pada masing-masing kelompok

bisa sama sehingga homogenitas subyek di antara keduanya dapat dicapai. Cara ini menjamin komparabilitas antara dua kelompok tersebut dan mengurangi bias yang mungkin terjadi akibat perbedaan ciri-ciri subyek (variabel luar) yang berpotensi mengganggu. Seperti telah disebut di depan alokasi random menghindari bias pengganggu (*confounding bias*). Selanjutnya hasil atau pengaruh masing-masing perlakuan dibandingkan antara kedua kelompok.

Beberapa kelebihan rancangan paralel ini adalah:

1. Prosedur yang ditempuh lebih sederhana baik mengenai cara pengorganisasian maupun analisisnya.
2. Subyek yang ada dalam tiap kelompok dapat dianggap homogen (komparabel), mengingat bahwa mereka ditempatkan secara random ke dalam kelompoknya masing-masing sehingga memiliki kesempatan yang sama dalam memperoleh perlakuan. Walaupun ada variabel luar yang mungkin mengganggu mereka telah terbagi secara imbang ke dalam dua kelompok tersebut.
3. Dapat digunakan untuk semua keadaan penyakit baik akut maupun kronis dan juga dapat dipakai untuk menilai semua jenis *outcome* yang ingin diteliti.
4. Tidak dipengaruhi oleh perlakuan lain yang sedang dibandingkan.

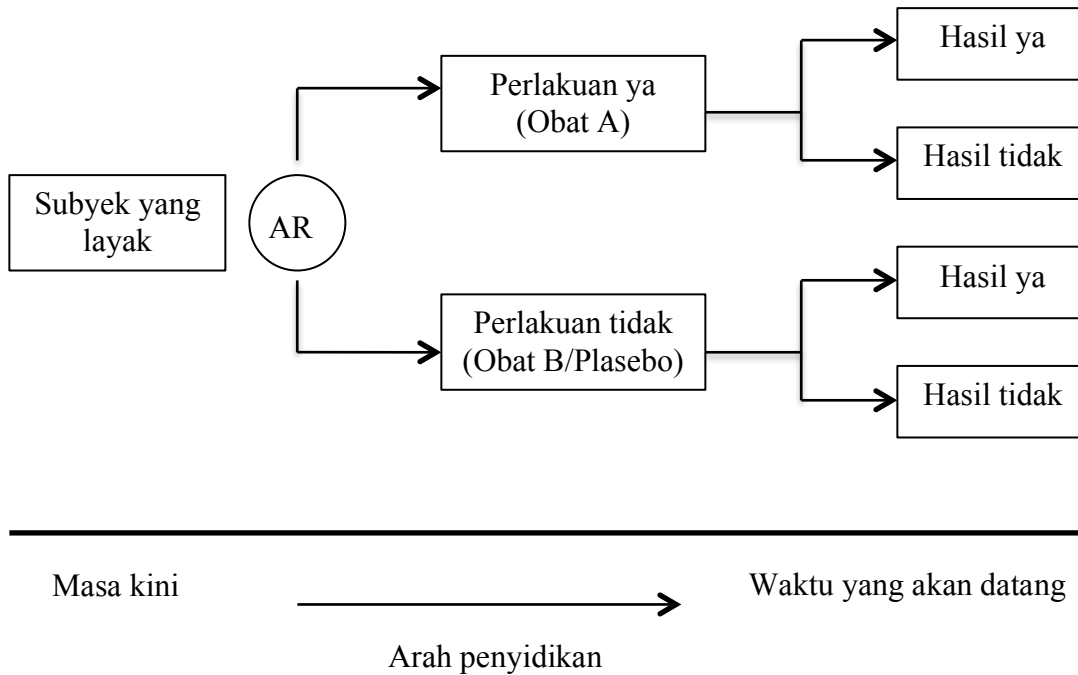
Meskipun demikian rancangan ini tidak luput dari kekurangan yaitu antara lain:

1. Keanekaragaman antar individu sulit dikendalikan meskipun alokasi secara random telah dikerjakan.
2. Diperlukan jumlah sampel yang cukup besar untuk memperoleh taraf signifikansi dan kekuatan penelitian yang diinginkan.

Untuk menghindari butir (1) di atas maka perlu dilakukan stratifikasi sebelum dilakukan alokasi jenis terapi. Stratifikasi dimaksudkan untuk mengelompokkan penderita berdasarkan derajat penyakit yang sama sehingga dapat lebih menjamin validitasnya pada saat membandingkan hasil atau pengaruh dari masing-masing derajat penyakit.

Secara skematis bentuk rancangan paralel dapat dilihat pada gambar berikut:

Skema uji klinis secara random, bentuk parael



Contoh 8.

Sebuah penelitian dirancang untuk membandingkan angka ekspulsi IUD CuT 380A yang dipasang segera pascasalin dengan R-inserter dan yang dipasang dengan *ring forcps*. Sejumlah 208 ibu pascasalin yang memenuhi kriteria kelayakan dirandom dalam 2 kelompok. Kelompok pertama dipasang IUD CuT 380 A menggunakan *R-inserter* dan kelompok kedua menggunakan *ring forcps*. Follow up dikerjakan setiap bulan sampai 12 bulan untuk melihat angka ekspulsi, angka infeksi, angka perdarahan, angka pencabutan dan angka kelangsungan pemakaian (Siswosudarmo et al., 2016).

Contoh 9.

Karena pemakain antibiotika yang irrasional makin meningkat, maka resistensi amoksisilin terhadap kuman gonorea juga meningkat dengan membentuk enzim beta laktamase. Sebuah penelitian dirancang untuk mengetahui apakah penambahan asam klavulanat, yang dapat menghancurkan enzim betalaktamase, pada amoksisilin dapat menaikkan angka kesembuhan penyakit urethritis gonorea (PID). Untuk ini sejumlah

penderita yang memenuhi kriteria penelitian dirandom dalam dua kelompok yang masing-masing mendapat amoksisilin plus asam klavulanat (sebagai kelompok perlakuan), dan ampisilin saja (sebagai kelompok kontrol). Pengobatan dilakukan dalam 6 hari dengan dosis yang telah ditentukan. Penderita diminta untuk mencatat keluhan yang ada setiap harinya. Pada hari ke 6 dilakukan pemeriksaan *swab* uretra dan kultur urin untuk menilai kesembuhannya.

Hasil yang diharapkan antara lain adalah:

- a. Berapa persen penderita yang secara subyektif dan obyektif mengalami perbaikan pada tiap kelompok.
- b. Apakah perbedaan tersebut secara klinis dan statistika bermakna.
- c. Berapa persen penderita yang mengalami keluhan efek samping (misal rasa mual, nyeri lambung dll) pada masing-masing kelompok.

2. Rancangan silang (*Cross over design*)

Pada bentuk ini setiap kelompok menerima kedua perlakuan secara bergantian setelah satu periode tertentu yang disebut dengan periode istirahat (*wash out period*). Sebagaimana pengambilan pasien (subyek) pada setiap UKR maka subyek yang memenuhi kriteria kelayakan dibagi dalam dua kelompok secara random. Pada periode pertama kelompok pertama mendapat perlakuan obat A misalnya, dan kelompok kedua mendapat perlakuan obat B. Kedua kelompok ini diamati hasilnya selama periode tertentu kemudian perlakuan dihentikan untuk waktu tertentu. Selanjutnya pada periode kedua, kelompok pertama mendapat obat B dan kelompok kedua mendapat obat A dan mereka diamati untuk waktu yang sama dengan periode pertama. Hasil masing-masing pengamatan (obat A dan obat B) dibandingkan dan dianalisis.

Yang dimaksud dengan periode istirahat adalah interval waktu tertentu untuk menghilangkan pengaruh perlakuan pertama sehingga sewaktu subyek masuk dalam perlakuan kedua seolah-olah ia adalah subyek baru. Demikian pula sebaliknya subyek dari perlakuan kedua yang akan masuk dalam perlakuan pertama. Berapa lama waktu istirahat ini sangat tergantung pada jenis obat yang diteliti. Untuk pil kontrasepsi misalnya, interval satu kali mensis dianggap sudah cukup sebagai waktu istirahat. Untuk suntik KB, maka waktu istirahat paling tidak empat sampai enam bulan setelah suntikan terakhir.

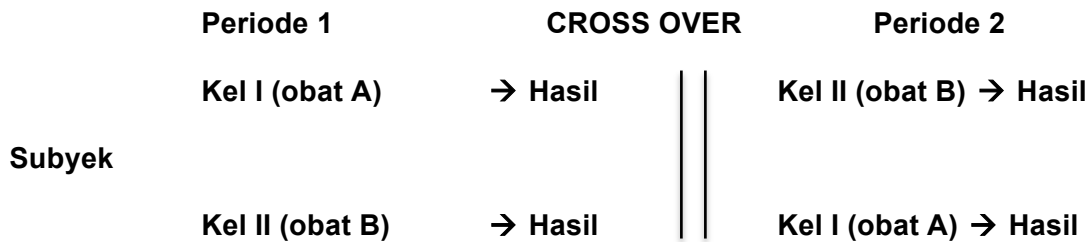
Keunggulan rancangan silang ini antara lain adalah:

1. Jumlah subyek yang diperlukan relatif lebih sedikit dibanding dengan rancangan parallel sebab setiap subyek paling tidak digunakan dua kali. Hal ini sangat bermanfaat terutama untuk penelitian kasus-kasus yang jarang.
2. Rancangan silang menjamin perbandingan perlakuan yang lebih tepat karena keanaka-ragaman antar individu dapat dikendalikan. Masing-masing subyek dibandingkan dengan dirinya sendiri di antara fase-fase uji klinis yang berbeda. Sebagai contoh penderita asthma menunjukkan variasi yang sangat besar dalam hal fungsi paru-parunya (*Force Expiratory Volume/FEV*)

Beberapa kelemahannya antara lain adalah:

1. Ada kemungkinan obat pertama mempunyai pengaruh besar terhadap kesembuhan pasien sehingga pemberian obat yang ke dua sudah tidak berpengaruh lagi. Salah satu cara mengatasinya adalah dengan memberikan waktu istirahat tadi.
2. Rancangan silang hanya cocok digunakan untuk penyakit-penyakit yang bersifat kronik stabil dengan kemungkinan kekambuhan yang tinggi dan ketergantungan terhadap obat juga besar. Contohnya adalah artritis reumatoid, hipertensi, diabetes mellitus dan lain-lain. Juga cocok untuk penelitian KB seperti pil, suntik dll.
3. Rancangan sama subyek tidak dapat dinilai bila waktu uji yang dijadwalkan ternyata lebih pendek dari riwayat alamiah penyakit, sehingga hasil terapi tidak mencerminkan keadaan yang sebenarnya.
4. Dibutuhkan waktu pemantauan yang lebih lama dan diperlukan pemeriksaan-pemeriksaan ulang untuk membuktikan bahwa setelah periode istirahat kondisi pasien sudah kembali ke keadaan semula seperti sebelum intervensi.
5. Yang lebih sulit lagi ialah bila pasien kemudian dinyatakan *drop out* dari penelitian sebelum seluruh intervensi diberikan.
6. Tidak bisa digunakan untuk pengujian yang:
 - a. Perlakuannya yang berupa tindakan operatif.
 - b. Hasil yang diperoleh berupa kematian (misal trial tentang obat anti-kanker)
 - c. Kesembuhannya sempurna atau permanent (misal antibiotika)

Secara skematis bentuk rancangan silang adalah sebagai berikut:



Contoh 10.

Dalam mengikuti program keluarga berencana seorang perempuan hampir tidak mungkin minum satu macam pil secara terus menerus. Pergantian pil menimbulkan kekhawatiran akan timbulnya efek samping yang dapat menyebabkan ibu enggan meneruskan pemakaian pil. Kekhawatiran ini timbul setelah lahirnya generasi pil kontrasepsi baru yang mengandung estrogen dosis rendah. Untuk keperluan ini dirancang suatu uji klinis *cross over* untuk menilai penerimaan penggantian pil dari estrogen dosis standar ke estrogen dosis rendah dibandingkan dengan penggantian pil dari estrogen dosis rendah ke pil estrogen dosis standar. Setelah tiga bulan pemakaian mereka dipindah ke pil berikutnya juga untuk waktu tiga bulan. Satu kali menses dianggap sebagai *wash out period* untuk dapat masuk dalam pengobatan kedua.

BAB IV. RANDOMISASI, PENYAMARAN DAN CONCEALMENT

Randomisasi

Penelitian yang benar selalu menuntut adanya kelompok control yang berfungsi sebagai pembandingan terhadap perlakuan baru yang akan diujicobakan. Randomisasi adalah cara terpilih untuk menempatkan subyek ke dalam kelompok perlakuan dan kelompok kontrol. Subyek yang dialokasikan secara random mempunyai kesempatan yang sama untuk masuk ke dalam kelompok perlakuan ataupun kelompok kontrol bila jumlah sampelnya seimbang.

Randomisasi menghasilkan kelompok penelitian yang sebanding baik dalam faktor risiko yang diketahui maupun faktor risiko yang tidak diketahui, dan menjamin bahwa uji statistiknya akan menghasilkan tingkat kepercayaan yang sah dan terpercaya (*valid and reliable*). Tanpa randomisasi maka peneliti cenderung memilih atau menolak obat atau pengobatan tertentu dan akan berusaha baik secara sadar maupun tidak, untuk memasukkan atau memindahkan subyek ke dalam kelompok tertentu. Dengan demikian randomisasi menghindari adanya bias pengganggu (*confounding bias*), yakni variabel luar yang berpotensi mempengaruhi hasil.

Cara randomisasi. Sekali lagi diingatkan bahwa randomisasi adalah cara terpilih untuk menempatkan subyek ke dalam kelompok perlakuan dan kelompok kontrol, sehingga setiap subyek mempunyai kesempatan yang sama untuk masuk ke dalam kelompok perlakuan ataupun kelompok control. Ini tidak diketahui lebih dahulu baik oleh peneliti maupun oleh subyek yang bersangkutan. Dengan demikian ada cara-cara yang dulu dianggap sebagai cara randomisasi sekarang tidak lagi dapat diterima. Cara tersebut antara lain adalah:

1. Cara bergantian, misal A, B, A, B dst.
2. Menggunakan nomer ganjil dan nomer genap.
3. Menggunakan tanggal lahir (tanggal genap A, ganjil B dst).
4. Cara mengocok kartu.
5. Melempar koin, atau mata uang.

Cara tersebut telah ditinggalkan karena selain urutan perlakuan dapat ditebak, cara itu tidak dapat di *re-checked* dan tidak *reproducible*.

Randomisasi tersembunyi. Dengan randomisasi tersembunyi, peneliti tidak mengetahui perlakuan apa yang akan diberikan pada subyek berikutnya. Untuk maksud

ini biasanya digunakan tabel bilangan random (*random number table*) atau bilangan random yang dibuat dengan komputer (*computer generated random number*). Nomor yang telah dipilih dimasukkan dalam amplop tertutup yang tidak tembus pandang dan si pembuat nomer random tidak boleh terlibat dalam pengumpulan subyek, pemberian perlakuan dan evaluasi. Pada waktu membuat laporan penelitian, cara randomisasi harus disebutkan secara jelas, agar pembaca tidak ragu lagi dan agar penelitian anda *reproducible*.

Cara randomisasi.

1. Randomisasi dengan menggunakan table bilangan random.

Tabel bilangan random terdapat dalam banyak buku statistika. Berikut ini kami sampaikan sebuah contoh tabel bilangan random dan cara memakainya. Pertama, kita telah mempunyai daftar blok yang terdiri atas urutan nomer dan kombinasi huruf seperti berikut:

Tabel 1. Daftar blok, terdiri atas 6 huruf per blok.

Nomer	Urutan	Nomer	Urutan
00-04	AAABBB	50-54	BAAABB
05-09	AABABB	55-59	BAABAB
10-14	AABBAB	60-64	BAABBA
15-19	AABBBA	65-69	BABAAB
20-24	ABAABB	70-74	BABABA
25-29	ABABAB	75-79	BABBAA
30-34	ABABBA	80-84	BBAAAB
35-39	ABBAAB	85-89	BBAABA
40-44	ABBABA	90-94	BBABAA
45-49	ABBBAA	95-99	BBBAAA

Tabel 2. Daftar bilangan random

41	82	79	37	00	45	98	54	52	89	26	34
66	18	79	82	22	18	61	90	90	63	78	57
42	34	00	49	97	53	33	16	26	91	37	58
90	84	22	16	26	96	54	11	01	96	58	91
33	43	01	32	58	39	19	54	56	37	23	38
17	89	37	04	73	32	13	45	59	03	91	08
36	87	98	73	77	64	75	19	05	61	11	64
15	58	19	68	95	47	25	69	11	90	26	19
45	52	27	35	86	81	16	20	37	60	39	35
72	72	81	84	36	58	05	10	70	50	31	04
06	68	52	50	39	85	92	28	18	89	64	37
95	73	30	92	26	49	54	30	41	21	06	62
96	23	16	46	15	51	60	31	55	27	84	14
34	96	33	68	48	22	40	17	43	25	33	31
07	19	94	46	17	51	03	73	99	89	28	44
37	08	08	46	56	76	29	48	33	87	70	79
18	01	67	19	29	49	58	67	08	56	27	24
08	79	18	78	00	32	86	74	78	55	55	72

Cara menggunakan

1. Tentukan jumlah sampel anda, misal 60 (pelakukan dan kontrol), maka anda membutuhkan 10 blok ($10 \times 6 = 60$).
2. Tentukan sebuah nomer sembarang yang ada dalam table bilangan random, misal nomer 33 (dalam contoh diberi warna merah dan dicetak tebal).
3. Bergeraklah anda ke arah sembarang dari nomer 33, boleh ke bawah, boleh mendatar, boleh diagonal. Ambillah contoh ke bawah, maka anda memperoleh urutan sbb:

33 54 19 13 75 25 16 05 92 54

4. Gantilah angka tersebut dengan huruf yang ada dalam daftar blok. Angka 33 terletak dalam blok 30 – 34, dengan urutan ABABBA, dan demikian seterusnya,

Sehingga anda memperoleh urutan huruf sebagai berikut:

No. random	Urutan huruf
33	ABABBA
54	BAAABB
19	AABBBA
13	AABBAB
75	BABBAA
25	ABABAB
16	AABBAA
05	AABABB
92	BBABAA
54	BAAABB

5. Ambilah 60 amplop, dan berilah nomer urut 1 sampai 60. Masukkan huruf-huruf tersebut di atas secara berurutan sejak A pertama dari baris paling atas (sesuai nomer 33) sampai B akhir dari baris paling bawah (sesuai nomer 54). Hasil urutannya adalah sebagai berikut:

No urut amplop	Isi random
1	A
2	B
3	A
4	B
5	B
..	..
55	B
56	A
57	A
58	A
59	B
60	B

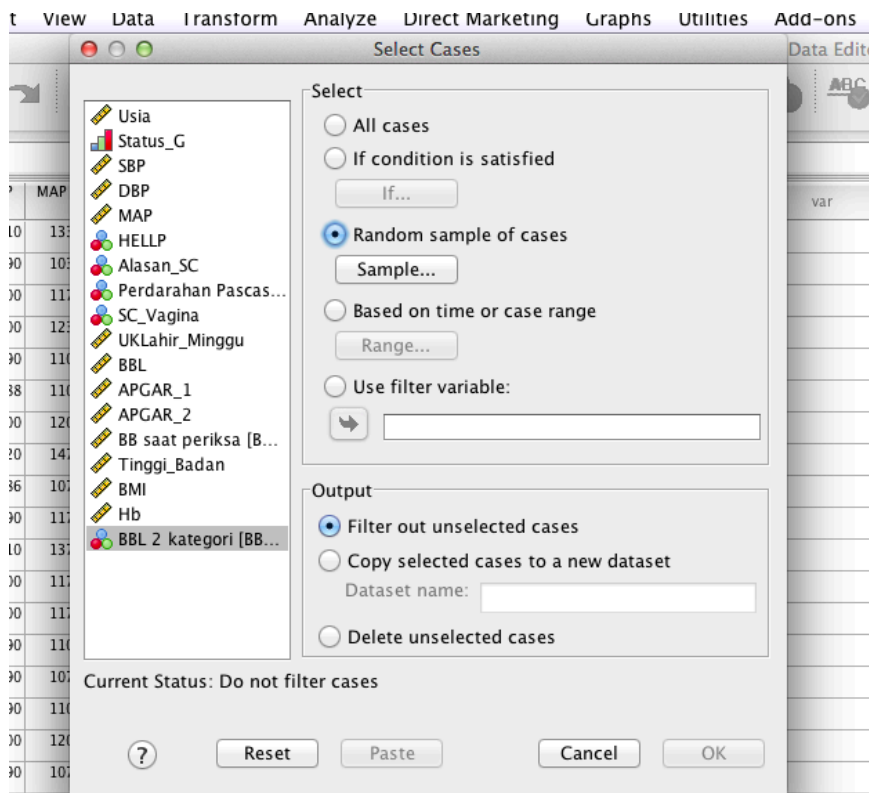
Perhatikan bahwa amplop harus tidak tembus pandang dan dilem, sehingga isi amplop baru boleh diketahui bila telah ada subyek yang memenuhi kriteria kelayakan.

6. Langkah ke enam adalah apabila anda mendapatkan subyek yang memenuhi kriteria inklusi bukanlah amplop nomer 1, dan lihatlah perlakuan apa yang akan ia dapatkan. Subyek yang datang kedua dan seterusnya mendapat nomer urut 2,3 dst sampai subyek yang ke 60.

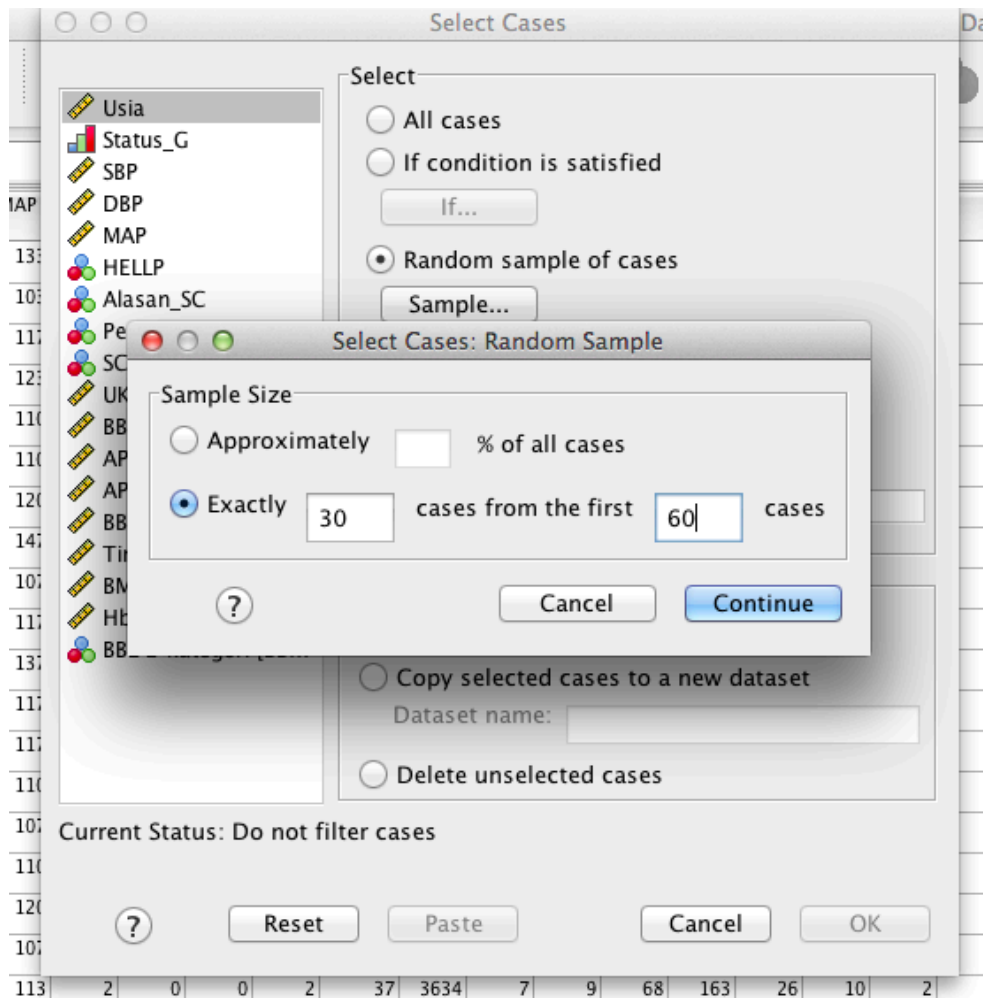
2. Randomisasi dengan bilangan random dari komputer.

Cara ini lebih mudah dan anda dapat menggunakan program SPSS untuk membantu anda. Berikut adalah langkah-langkah untuk memperoleh bilangan random menggunakan program SPSS. Misal anda membutuhkan 30 subyek untuk kelompok perlakuan dan 30 kontrol.

- a. Buka sembarang data SPSS, pilihlah 60 kasus pertama dari data yang tersedia.
- b. Klik **data** → **select cases** → **random sample of cases** → maka anda akan masuk pada jendela sbb:



- c. Klik **sample** → **exactly** → isikan pada kotak **30** from the first **60** lalu, maka anda akan masuk dalam jendela sbb:



d. Selanjutnya klik **continue** maka anda akan memperoleh bilangan random sbb:

	Usia	filter_\$
1	32	1
2	33	0
3	21	0
4	35	0
5	35	1
6	24	1
7	30	1
8	37	1
9	37	0
10	29	1
11	32	0
12	38	1
13	23	1
14	19	1
15	26	0
16	41	0
17	25	0
18	24	0
19	24	1
20	33	1

Dari daftar bilangan di atas, anda tinggal menentukan angka yang dicoret sebagai perlakuan dan yang tidak dicoret sebagai kontrol, atau sebaliknya sama saja. (Filter 0 sebagai perlakuan atau filter 1 sebagai kontrol terserah saja, karena jumlahnya seimbang yakni 15 dan 15).

Catatan: Jika random ini anda ulang, maka anda akan mendapatkan nomer baru lagi dan tidak pernah sama. Jadi alokasi mereka benar-benar secara random.

Penyamaran (*Blinding*)

Dalam sebuah UKR sering terjadi kesalahan yang bersifat sistematis (*systematic error*) yang dikenal sebagai bias. Bias adalah kesalahan sistematis yaitu perbedaan antara hasil yang seharusnya dengan dengan hasil yang diperoleh, yang disebabkan

oleh adanya faktor-faktor yang sifatnya bukan kebetulan. Bias dapat disebabkan oleh faktor-faktor yang disadari maupun yang tidak disadari. Ia dapat terjadi kapan saja baik sejak membuat rancangan, pelaksanaan, analisis maupun interpretasi hasil.

Untuk menghindarinya, secara umum dianjurkan agar peneliti dan subyek tetap dalam keadaan tidak mengetahui ke dalam kelompok mana sebenarnya subyek berada. Keadaan tidak mengetahui inilah yang dikenal dengan istilah penyamaran (*blinding*). Penyamaran harus dilakukan sejak alokasi random, pemberian perlakuan, **penilaian, klasifikasi, analisis dan interpretasi hasil.**

Pada umumnya dikenal tiga penyamaran yakni penyamaran tunggal (*single blinding*), penyamaran ganda (*double blinding*) dan penyamaran rangkap tiga (*triple blinding*). **Single blinding** adalah keadaan di mana hanya subyek atau peneliti yang tidak mengetahui ke dalam kelompok mana subyek berada, sedang **double blinding** terjadi bila keduanya (subyek dan peneliti) tidak mengetahuinya. **Triple blinding** diperoleh bila evaluator adalah bukan peneliti dan baik subyek, peneliti dan evaluator tidak mengetahui ke dalam kelompok mana subyek berada, atau perlakuan apa yang diterima oleh masing-masing subyek.

Penyamaran peneliti terhadap perlakuan adalah penting karena bila si peneliti mengetahui bahwa obat yang diterima pasien diduga kurang efektif, ia akan berusaha memberi obat tambahan atau akan melakukan pengontrolan lebih sering dari yang seharusnya seperti yang ada dalam protokol. Penyamaran subyek terhadap perlakuan juga amat penting sebab bila subyek mengetahui jenis obat yang ia terima dan ia menduga obat tersebut tidak efektif atau obat tersebut tidak populer maka ia akan merasa sulit mendapatkan kesembuhan. Hal ini akan menjadi sangat berpengaruh bila *outcome* yang diharapkan merupakan keluhan yang bersifat subyektif, seperti penelitian obat analgetika, tranquiliser dll. Bahaya yang kedua adalah tingginya angka *drop out* disebabkan pasien merasa tidak ada manfaatnya masuk dalam penelitian bila mereka hanya mendapatkan pengobatan yang biasa saja.

Sudah barang tentu tidak semua UKR dapat dibuat tersamar, karena beberapa perlakuan dapat dengan mudah dikenali baik oleh subyek maupun oleh peneliti. Sebagai contoh bila anda ingin membandingkan metode operasi dengan metode non operasi, maka *blinding* tidak mungkin dilakukan. Contoh lain adalah bila anda ingin membandingkan kemampuan obat oral dengan obat suntikan, seperti misalnya membandingkan *effectiveness* pil KB dengan DMPA. Dalam hal ini *blinding* masih dapat dilakukan terhadap evaluator yakni orang yang mengamati subyek pada saat mereka

datang berkunjung, dengan catatan si evaluator tidak berusaha bertanya kepada para akseptor.

Soal

Pilihlah satu jawaban yang paling betul

1. Sebut cara randomisasi yang dianjurkan
 - a. Bergantian menggunakan nomer ganjil dan genap
 - b. Menggunakan tanggal lahir
 - c. Mengocok kartu
 - d. Menggunakan table bilangan random
2. Sebut jenis penyamaran bila anda ingin membandingkan obat anti hipertensi baru dengan standard
 - a. Single blind, pasien
 - b. Double blind, pasien dan dokter
 - c. Double blind, pasien, dokter dan evaluator
 - d. Triple blind, pasien, dokter dan evaluator
3. Penyamaran apa yang dianjurkan bila obat suntik akan dibandingkan dengan obat oral
 - a. Single blind, pasien
 - b. Single blind, dokter
 - c. Double blind pasien dan dokter
 - d. Double blind, dokter dan evaluator
4. Sebut satu faktor terpenting untuk mengurangi bias
 - a. Alokasi subyek secara random
 - b. Subyek dan peneliti harus tetap blinded terhadap perlakuan
 - c. Menggunakan placebo
 - d. Alokasi bubyek lewat telepon
5. Tujuan utama randomisasi adalah sbb:
 - a. Memberi kesempatan yang sama kepada setiap subyek untuk masuk ke dalam kelompok perlakuan atau kelompok control.
 - b. Memudahkan analisis statistik.
 - c. Menghindari *drop out*.
 - d. Menghindari bias pengukuran

BAB V. BESAR SAMPEL (*SAMPLE SIZE*)

Pada saat menyusun sebuah UKR besar sampel (banyaknya subjek penelitian atau *sample size*) harus ditentukan. Besar sampel sangat mempengaruhi jalannya penelitian yaitu dalam hal biaya, tempat, waktu penelitian dan jumlah tenaga yang diperlukan agar hasil yang diperoleh menjadi *valid* dan *reliable*. Dari sisi metodologi, besar sampel mempengaruhi kekuatan penelitian dan ini sangat mempengaruhi kualitas hasil penelitian.

Bila jumlah sampel tidak diperhitungkan terlebih dahulu ia bisa terlalu kecil sehingga kesimpulan penelitian sukar ditarik, atau ia tidak bisa mendeteksi adanya perbedaan yang seharusnya ada. Ia bisa juga terlalu besar yang berarti kita memboroskan uang, waktu dan tenaga untuk memperoleh kesimpulan yang sama dengan penelitian yang jumlah sampelnya bisa lebih kecil.

Menentukan berapa jumlah sampel yang diperlukan sering merupakan kompromi antara keperluan ilmiah, sumber daya dan jumlah subjek yang tersedia. Masalah etika sering juga membatasi penentuan jumlah sampel. Dalam menentukan jumlah sampel pertanyaan kunci seperti berikut harus dipahami betul.

Contoh 12.

Dalam pil KB komponen estrogen diduga mempunyai hubungan erat dengan timbulnya bermacam-macam efek samping. Makin rendah dosis estrogen mestinya makin memperkecil frekuensi dan durasi efek samping (seperti nausea, muntah dan sakit kepala). Untuk membuktikan hal itu dirancang sebuah UKR untuk membandingkan Pil KB standar (1,0 mg norethindrone dengan μg mestranol dengan pil KB dosis rendah (1,0 mg norethindrone dengan 30 μg ethinyl estradiol)). Pada saat merancang penelitian ini beberapa pertanyaan perlu dijawab sebelum jumlah sampel ditentukan.

Sebelum menentukan besar sampel ada beberapa pertanyaan yang harus diketahui:

1. Apakah tujuan primer penelitian ini?

Tujuan penelitian ini adalah untuk melihat apakah pil KB dosis rendah mempunyai angka kelangsungan (*continuation rate*) yang lebih baik dibandingkan dengan pil KB

dosis standard. Hal ini didasarkan kenyataan bahwa pil KB dosis rendah mempunyai angka kelainan perdarahan yang lebih sedikit sehingga angka kelangsungannya lebih tinggi dibanding pil KB dosis standar. Penelitian ini juga ingin mengetahui alasan penghentiannya, tetapi alasan tersebut bukan merupakan tujuan utama penelitian ini.

2. Apakah hasil primer yang akan diukur?

Hasil primer yang akan diukur adalah angka kelangsungan pemakaian atau dapat juga angka diskontinuitas selama satu tahun dari pemakaian kedua pil KB tersebut.

3. Berapa proporsi hasil positif yang anda harapkan dengan pil KB dosis standar?

Pil KB dosis standar di sini adalah kelompok kontrol yang menurut penelitian lampau angka kontinuitas selama 12 bulan adalah 60%. Dengan demikian angka diskontinuitasnya pada 12 bulan adalah 40%. Dalam contoh ini kita berbicara tentang angka kontinuitas meskipun sama saja bila kita menggunakan angka diskontinuitas.

4. Berapakah perbedaan (selisih proporsi) yang ingin dideteksi antara hasil pada perlakuan dan kontrol dan berapakah tingkat kepercayaannya (*level of confidence*?).

Pada pemakaian pil KB dosis rendah diharapkan terdapat kenaikan angka kontinuitas dibandingkan dengan pil KB dosis standar. Meskipun demikian mungkin juga terjadi penurunan angka kontinuitas oleh karena naiknya kejadian *breakthrough bleeding*. Peneliti ingin mendeteksi adanya perbedaan angka kontinuitas paling tidak 10% untuk bisa dikatakan bahwa secara klinis perbedaan tersebut memang ada. Berdasarkan adanya angka kontinuitas pada pemakaian pil KB dosis rendah adalah 60%, maka angka kontinuitas pada pil KB baru dosis rendah adalah serendah-rendahnya 50% atau setinggi-tingginya 70%. Selanjutnya kemungkinan terjadinya perbedaan sebesar 10% tersebut setidaknya-tidaknya adalah 90% (*level of significance*) sehingga bisa dikatakan bahwa perbedaan tersebut secara statistika adalah bermakna.

5. Bila hasil kedua perlakuan tersebut sebenarnya memang tidak berbeda berapakah besarnya risiko yang mungkin kita peroleh bila penelitian kita mengatakan bahwa secara statistik ada perbedaan?

Peneliti mengharapkan bahwa kemungkinan paling tinggi hanyalah 5% yang mengatakan ada perbedaan padahal sebenarnya tidak ada, artinya kemungkinan salah dengan mengatakan ada perbedaan hanyalah 5%.

6. Berapakah proporsi subyek (perempuan) yang telah masuk ke dalam penelitian yang mungkin akan *drop out* oleh sebab selain hal yang sedang diteliti?

Karena kontinuitas adalah hasil yang sedang diteliti maka semua alasan untuk meninggalkan penelitian dianggap sebagai diskontinuitas.

Ada dua jenis data yang biasa dipakai sebagai dasar penghitungan besar sampel, yakni data **dikotom** dan data **kontinyu**.

Data dikotom

Data dikotom adalah data yang hanya terdiri atas dua kategori misal **mati** dan **hidup**, **ya** dan **tidak** dan sebagainya. Dalam penelitian kontrasepsi sebagian besar hasil yang dinilai bersifat dikotomik misal **kehamilan ya** atau **tidak**, **ekspulsi ya** atau **tidak**, **perdarahan ya** atau **tidak** dan sebagainya. Data ini biasanya disajikan dalam persen.

Ada lima hal yang harus diperhatikan bila kita menghitung besar sampel untuk data dikotom yaitu P_1 , P_2 , alfa (α), beta (β) dan pasien yang diperkirakan hilang dari pengamatan (f) yang semuanya telah dibicarakan di atas.

P_1 adalah persentase jawaban **ya** atau hasil **positif** yang diharapkan dari kelompok pertama, biasanya kelompok yang dianggap sebagai kontrol. Nilai P_1 ini biasanya didapat berdasarkan penelitian orang lain dan biasanya diambil dari studi kepustakaan. Pada contoh di atas P_1 adalah 60%.

P_2 adalah persentase jawaban **ya** dari hasil yang diharapkan pada kelompok perlakuan yang sedang diteliti. Besarnya P_2 tergantung pada pertimbangan peneliti, misalnya apakah P_2 akan dibuat separoh P_1 atau selisih 10% seperti contoh di atas. Sering pula besarnya P_2 ini disesuaikan dengan kepentingan klinis hasil penelitian, artinya kalau ada perbedaan antara perlakuan dengan kontrol apakah perbedaan tersebut secara klinis mempunyai arti. Pada contoh di atas maka P_2 adalah 50% atau 70% (selisih 10% lebih besar atau lebih kecil dari P_1).

Alfa (α), adalah tingkat kemaknaan (*level of significance*) atau disebut juga kesalahan tipe I, yang digunakan untuk mengatakan sampai seberapa jauh ada atau tidak ada perbedaan antara sekelompok kontrol dengan kelompok perlakuan. Jadi α digunakan untuk mengatakan bahwa ada perbedaan, padahal sebenarnya tidak ada. Pada contoh diatas α adalah 0,05 artinya bila peneliti mengatakan ada perbedaan, maka yang 5% adalah salah (atau sebenarnya tidak beda). Dengan kata lain kemungkinan salah dengan mengatakan ada perbedaan adalah 5% karena yang 5% ini sebenarnya tidak ada perbedaan. Besarnya nilai α biasanya bervariasi dari 0,01; 0,05 dan 0,10. Dalam penelitian biomedis, umumnya dipakai 0,05 atau 5%.

Beta (β) atau kesalahan tipe II digunakan untuk mengatakan tidak ada perbedaan antara kelompok kontrol dan perlakuan padahal sebenarnya ada. Kekuatan penelitian adalah $1 - \beta$ yaitu tingkat kepastian untuk mendeteksi suatu perbedaan kalau perbedaan tersebut memang ada. Pada contoh diatas β adalah 0,1 atau 10%. Biasanya nilai β adalah 0,05, 0,1 dan 0,20 atau kekuatan penelitian menjadi 95%, 90% atau 80%. Artinya peneliti ingin 95%, 90% atau 80% yakin bahwa kalau ada perbedaan, maka perbedaan tersebut pasti terdeteksi.

Pengertian kesalahan tipe I dan tipe II dapat dijelaskan lagi sebagai berikut: Perhatikan Tabel 2 X 2 berikut ini:

Exposure	Effect	
	Yes	No
Yes	Yes and Yes. (Tru yes)	Yes and No. (Type I error = α)
No	No and Yes (Type II error = β)	No and No (True no)

Note:

1. *Type I error* (α): Nilai α 5%, artinya, kalau kita memperoleh sebuah efek (ada perbedaan), maka kita 95% yakin bahwa efek tersebut benar karena pengaruh perlakuan dan bukan secara kebetulan. Hanya 5% efek ini terjadi secara kebetulan, yang sebenarnya tidak ada efek.
2. *Type II error* (β): Kita tidak mendapatkan adanya efek (tidak mendapatkan perbedaan) padahal sebenarnya ada. Kalau β 0,2 artinya bila kita **tidak**

mendapatkan efek (perbedaan) maka kita yakin bahwa 80% memang tidak ada efek dan hanya 20% yang memang ada efek (ada perbedaan). Jadi kalau penelitian kita tidak mendapatkan adanya perbedaan, sangat mungkin hal ini terjadi karena penelitian kita tidak cukup besar (*sample size is not big enough*) untuk mendeteksi adanya efek (perbedaan) tersebut.

F adalah proporsi subyek yang diperkirakan akan meninggalkan penelitian sebelum penelitian berakhir oleh sebab yang bukan merupakan hasil yang sedang diteliti. Besarnya nilai *f* ditentukan berdasarkan pengalaman penelitian orang lain terdahulu atau menurut pengalaman Peneliti itu sendiri. Pada umumnya besarnya angka *loss to follow up* atau *drop out* tidak boleh lebih dari 20%. Pada contoh diatas *f* adalah 0%, karena semua pasien yang *loss to follow up* dan atau *drop out* dianggap diskontinyu.

Uji satu sisi (ekor) dan uji dua sisi (ekor)

Jika kita membandingkan perlakuan dan kontrol dan kita ingin mengetahui bahwa perlakuan yang baru lebih baik dari kontrol, maka perbandingan semacam ini disebut perbandingan satu sisi. Mungkin juga seorang peneliti hanya ingin mengetahui apakah perlakuan baru tersebut berbeda dengan kontrol pada tingkat perbedaan tertentu tanpa mempedulikan apakah perbedaan tersebut lebih baik atau lebih jelek (lebih besar atau lebih kecil) dibandingkan kontrol. Uji semacam ini disebut uji dua sisi.

Contoh di atas adalah uji dua sisi sebab peneliti hanya ingin mengetahui apakah pil KB dosis rendah berbeda paling tidak 10% dibandingkan dengan pil KB dosis standar dalam hal angka kelangsungan, tanpa menentukan bahwa perlakuan pil KB dosis rendah lebih tinggi atau lebih rendah dalam hal kelangsungan pemakaian. Dengan mengambil uji 2 sisi, jumlah sampel yang dibutuhkan lebih besar karena nilai Z_{α} dua sisi lebih besar dibanding nilai Z_{α} satu sisi (lihat tabel α di bawah).

Rumus besar sampel

Besar sampel untuk masing-masing kelompok (data diskrit) adalah:

$$N \text{ (per grup)} = \frac{1}{(1-f)} \times 2 \frac{(Z_{\alpha} + Z_{\beta})^2 \cdot p(100-p)}{(p_1 - p_2)^2}$$

Di mana:

- p_1 = proporsi (prosen) hasil positif pada kelompok kontrol
- p_2 = proporsi (prosen) hasil positif pada kelompok perlakuan
- p = $(p_1 + p_2)/2$

- α = kesalahan tipe I
- β = kesalahan tipe II
- f = proporsi subyek yang diperkirakan *loss to follow up*

Nilai Z_α dan Z_β dapat dilihat pada tabel sebagai berikut:

Tabel α dan nilai Z_α

α	Z_α	
	1 sisi	2 sisi
0,1	1,28	1,65
0,05	1,65	1,96
0,025	1,96	2,24
0,01	2,33	2,58

Tabel β dan nilai Z_β

β	Z_β	Kekuatan
0,50	0,00	0,50
0,40	0,25	0,60
0,30	0,53	0,70
0,20	0,84	0,80
0,15	1,03	0,85
0,10	1,28	0,90
0,05	1,65	0,95
0,025	1,98	0,975
0,01	2,33	0,99

Pada contoh diatas data yang diperlukan untuk menghitung besar sampel adalah:

- P_1 = 60%
- P_2 = 50% atau 70%
- α = 0,05
- β = 0,10

$f = 0$, karena semua kasus drop out dianggap diskontinyu

Dengan memakai uji dua sisi maka nilai p_2 harus diambil yang mendekati 50% supaya menghasilkan sampel yang terbesar.

Dengan menggunakan rumus diatas,

$$\begin{aligned} \text{N per grup} &= \frac{1}{(1-0)} \times 2 \frac{(1,96 + 1,28)^2 \times 2 \times 55 \times (100 - 55)}{(60 - 50)^2} \\ &= 520 \end{aligned}$$

Jadi besar sampel seluruhnya = 1040

Tugas: Cobalah hitung berapa besar sampel pada contoh di atas bila menggunakan P_2 70%.

Jawab: 47, per grup atau 94 seluruhnya.

Data kontinyu

Data kontinyu adalah data yang dapat diukur (mempunyai satuan) misalnya tinggi badan (dalam cm), berat badan (dalam kg), kadar hemoglobin (dalam gram/dL). Ciri lain data kontinyu adalah data yang mempunyai rata-rata (*mean*) dan simpang baku (*standard deviation, atau SD*). Pada pemakaian pil KB, peneliti mungkin juga tertarik pada perubahan tekanan darah dan fungsi hati atau jumlah produksi air susu dan konsentrasi estrogen atau progesteron dalam air susu. Data di atas adalah data kontinyu (data numerik).

Perhitungan besar sampel pada data kontinyu pada dasarnya sama saja dengan data dikotom hanya saja yang dibandingkan di sini bukan hasil yang berupa persentase (*rate*), tetapi hasil yang berupa rata-rata (*mean*). Di samping itu dibutuhkan juga data tentang *variance* atau deviasi standar dari kelompok kontrol. Nilai α , β dan f sama seperti pada data dikotom.

Rumus besar sampel untuk data kontinyu adalah sebagai berikut:

$$N \text{ (per grup)} = \frac{1}{1-f} \times 2 \frac{(Z_{\alpha} + Z_{\beta})^2 x (SD)^2}{(x_1 - x_2)^2}$$

Di mana:

SD adalah *standard deviation* kelompok kontrol, diperoleh dari studi kepustakaan.

X_1 adalah *mean* kelompok kontrol, diperoleh dari studi kepustakaan.

X_2 adalah *mean* kelompok perlakuan, ditentukan oleh peneliti.

Rumus ini dapat digunakan untuk perbandingan satu sisi maupun dua sisi dengan memilih nilai α untuk perbandingan satu atau dua sisi. Nilai X_2 (*mean* kelompok perlakuan) adalah sembarang artinya boleh mengambil nilai yang tinggi maupun nilai yang rendah.

Contoh 13

Seorang peneliti ingin melihat apakah penambahan zat besi dalam pil kontrasepsi dapat mengurangi kemungkinan anemia akibat darah yang keluar pada saat menstruasi. Untuk itu didesain sebuah UKR untuk membandingkan pil KB dosis standar yang berisi 21 pil KB aktif dan plasebo dengan pil KB serupa yang tidak berisi plasebo tetapi mengandung butir pil yang berisi 7 μg ferrous fumarat. Penelitian ini dirancang secara **double blind**.

Tujuan penelitian ini adalah untuk melihat efek pemberian zat besi terhadap kadar hemoglobin, hematokrit, dan besi serum. Hanya perempuan dengan kadar hemoglobin kurang dari 10 g/dL yang dapat masuk dalam penelitian ini. Mereka diikuti selama 6 bulan dengan pengambilan darah pada bulan ke 1, 3 dan ke 6. Kadar hemoglobin dan hematokrit pada populasi umum tidak diketahui tetapi kadar besi serum diketahui. Berdasarkan penelitian terdahulu diketahui bahwa pada pemakai pil KB dosis standar dalam bulan pertama terdapat kenaikan kadar besi serum sebesar 4 $\mu\text{g/dL}$ dengan deviasi standar nya juga 4 $\mu\text{g/dL}$.

Dalam penelitian ini peneliti ingin melihat apakah pemberian zat besi tambahan dapat menaikkan perubahan besi serum menjadi dua kalinya yakni menjadi 8 µg/dL. Ditentukan α 0,05; β 0,10 dan diperkirakan 15% dari para akseptor akan *drop out* sebelum penelitian berakhir.

Jumlah sampel yang dibutuhkan dihitung dengan rumus sbb:

$$\begin{aligned} N \text{ per grup} &= \frac{1}{(1-0,15)} \times 2 \frac{(1,68 + 1,28)^2 \times 4^2}{(4-8)^2} \\ &= 21 \end{aligned}$$

Jadi jumlah seluruh subjek yang dibutuhkan adalah 42

Soal: Hitunglah berapa jumlah subjek yang diperlukan bila peneliti pada contoh diatas ingin mendeteksi adanya perbedaan besi serum pada kelompok perlakuan 2 µg lebih besar dibandingkan dengan kelompok kontrol? (nilai α , β dan f adalah sama).

Jawab: N per grup 81 atau jumlah seluruh subjek 162.

3. Melakukan uji untuk menunjukkan tidak ada perbedaan.

Beberapa penelitian dirancang untuk membuktikan bahwa suatu obat baru adalah sama baik dengan obat standar yang sudah ada. Penelitian tersebut dilakukan bila obat yang baru memiliki beberapa kelebihan lain seperti misalnya ia lebih murah atau efek sampingnya lebih sedikit. Dalam hal ini cara menentukan besar sampel berbeda dengan penghitungan besar sampel untuk membuktikan bahwa dua buah hasil adalah berbeda.

Rumus besar sampel untuk data dikotom adalah sebagai berikut:

$$N \text{ (per grup)} = \frac{1}{1-f} \times 2 \frac{(Z_{\alpha} + Z_{\beta})^2 \cdot p(100-p)}{d^2}$$

Di mana:

P adalah prosentase “ya” pada kelompok kontrol.

D adalah perbedaan prosentase terbesar dari hasil “ya” antara kelompok kontrol dan kelompok perlakuan yang **secara klinis atau dalam kenyataan masih dianggap seimbang (tidak berbeda)**

Nilai α , β , dan f sama seperti contoh di atas.

Contoh 14

Dalam sebuah program KB yang disponsori pemerintah, suatu pil KB dosis standar telah dipakai dengan hasil memuaskan selama lima tahun. Suatu pil KB lain akan diperkenalkan kepada masyarakat karena harganya 20% lebih murah.

Angka kehamilan pada pil KB dosis standar yang sudah ada adalah 2 per 100 perempuan. Apabila pil KB yang baru menunjukkan angka kehamilan tidak lebih dari 4 per 100 perempuan maka ia secara klinis atau dalam kenyataannya dianggap tidak jauh berbeda dengan pil KB dosis standar yang sudah ada. Bila ditentukan α 0,05 dan β 0,10 dan kira-kira 20% dari perempuan yang masuk dalam penelitian akan *drop out*, maka berapa jumlah subjek yang dibutuhkan?

Jawab:

$$N \text{ (per grup)} = \frac{1}{(1-0,2)} \times 2 \frac{(1,96 + 1,28)^2 \times 2(100-2)^2}{(2-4)^2} = 1286$$

Jadi jumlah seluruh kasus adalah 2572

Soal.

Hitunglah jumlah subjek yang dibutuhkan bila angka kehamilan pil KB yang baru dituntut tidak lebih dari 3 per 100 perempuan untuk bisa dikatakan bahwa pil KB yang baru sama efektif dengan pil KB yang lama.

Jawab: 5144 per grup atau 10288 untuk seluruh kasus.

Catatan:

Perhatikan bahwa makin kecil perbedaan yang akan dideteksi makin besar jumlah sampel yang dibutuhkan. Di sinilah seorang peneliti harus pandai menyesuaikan angka-

angka sehingga kepentingan klinis, kelayakan (dana dan sumber daya) dan ilmiah sebuah penelitian terwakili semuanya dan penelitian menjadi layak untuk dikerjakan.

4. Besar sampel untuk variabel hasil yang lebih dari satu.

Dalam sebuah UKR biasanya variabel hasil yang akan diteliti memang lebih dari satu tetapi tujuan utama penelitian adalah hanya satu, yakni hasil yang paling penting atau paling utama. Bila semua variabel hasil dianggap penting maka besar sampel harus dihitung untuk setiap variabel hasil dan akhirnya dipilih jumlah sampel yang terbanyak.

Contoh 14

Peneliti ingin membandingkan sebuah IUD baru dengan IUD standar dalam hal angka kehamilan dan angka ekspulsi. Diketahui angka ekspulsi IUD standar adalah 10 per 100 perempuan dalam satu tahun, dan angka kehamilannya 4 per 100 perempuan per tahun. Selanjutnya IUD baru diharapkan dapat menurunkan angka ekspulsi dan angka kehamilan menjadi separohnya. Dengan mengambil α sebesar 0,05, dan kekuatan penelitian 80%, lalu 20% akseptor diperkirakan akan *loss to follow up* dalam satu tahun, berapa jumlah sampel yang diperlukan?

Jawab:

Rumus yang digunakan adalah rumus untuk data dikotom sebagai berikut:

Untuk angka kehamilan:

$$P_1 \text{ (kontrol)} = 4\%; P_2 \text{ (perlakuan)} = 2\%, \text{ maka } p = (4+2)/2 = 3\%$$

Untuk angka ekspulsi:

$$P_1 \text{ (kontrol)} = 10\%; P_2 \text{ (perlakuan)} = 5\%, \text{ maka } p = (10+5)/2 = 7,5\%$$

Telah ditentukan bahwa α 0,05 maka $Z_\alpha = 1,65$ (test satu sisi), dan β 0,20 maka $Z_\beta = 0,84$. Bila pasien yang diperkirakan *loss to follow up* (nilai f) adalah 20%, maka dengan menggunakan rumus diatas diperoleh jumlah sampel sbb:

$$N \text{ untuk kehamilan} = 1128 \text{ per grup}$$

$$N \text{ untuk ekspulsi} = 431 \text{ per grup}$$

Dalam hal seperti ini kita harus mengambil angka terbesar supaya keyakinan kita mencapai 80% untuk dapat mendeteksi adanya penurunan angka kehamilan sampai

separuhnya dan lebih dari 80% keyakinan kita untuk dapat mendeteksi penurunan angka ekspulsi menjadi separuhnya.

Keuntungan dan kerugian sebuah UKR

Keuntungan:

1. Adanya randomisasi menjamin alokasi subjek ke dalam kelompok perlakuan dan kontrol menjadi seimbang dan tidak dipengaruhi baik oleh personnel penelitian maupun subjeknya sendiri.
2. Dengan randomisasi variabel luar yang mengganggu akan terdistribusi secara seimbang ke dalam kelompok perlakuan dan kontrol.
3. Dengan randomisasi analisis statistiknya lebih valid, karena secara teoritis analisis statistik didasarkan atas pengelompokan subjek secara random.
4. Dalam UKR dimungkinkan adanya standarisasi tentang kriteria kelayakan, pemberian perlakuan dan kriteria hasil penelitian.
5. Dengan penyamaran, penilaian hasil penelitian lebih obyektif dan tidak memihak.
6. Dengan *concealment*, maka bias seleksi dapat dihindari.

Kerugian:

1. Rancangan dan pelaksanaannya kompleks sehingga mahal.
2. *Representativenessnya* kurang karena subjek yang secara sukarela masuk ke dalam penelitian mungkin berbeda dengan populasi ke mana penelitian ditujukan.
3. UKR dihadapkan pada masalah etika. Bolehkah sebuah obat atau perlakuan tertentu yang mungkin membahayakan atau setidaknya kurang bermanfaat diberikan pada sekelompok subjek? Bolehkan sekelompok subjek menolak perlakuan yang diberikan setelah mereka merasa tidak bermanfaat?
4. Tidak jarang penelitian memerlukan waktu yang cukup lama untuk memperoleh hasil yang diharapkan.

Bab VI. ANALISIS PADA SEBUAH UKR

Sebuah keunggulan pada UKR adalah adanya randomisasi, yang menjamin komparabilitas antara kedua kelompok sehingga semua variabel luar yang berpotensi sebagai variabel pengganggu telah terdistribusi secara seimbang ke dalam kedua kelompok, sehingga mereka tidak lagi berfungsi sebagai variabel pengganggu. Dengan demikian kalau ada perbedaan dalam hasil, maka hasil yang diperoleh adalah semata karena perlakuan dan bukan karena perbedaan dalam kelompok. Meskipun demikian, tahap pertama dalam melakukan analisis adalah menilai apakah kedua kelompok sebanding, atau tidak. Tabel univariat tidak perlu lagi ditampilkan.

Berikut adalah contoh tabel komparabilitas antar kelompok baik untuk variable dengan data diskrit maupun variable dengan data kontinyu.

Table Comparability between groups*)

	<i>R-inserter</i>		Ring forceps		P value
	N	%	N	%	
Parity					
Primipara	53	51	49	47	0.67
Multipara	51	49	55	53	
Time of insertion					
≤10 minutes	100	96.2	97	93.8	0.54
> 10 minutes	4	3.8	7	6.7	
Age (years)					
N	104		104		0.07
Mean ± SD	27.38 ± 5.34		28.59 ± 5.99		

*) Siswosudarmo et al. 2014.

Perhatikan bahwa komparabilitas ditentukan oleh nilai P saja, dan tidak ada RR maupun *confidence interval*. Nilai P diperoleh dengan melakukan uji chi square kalau datanya diskrit dan uji t (t-test) kalau datanya bersifat kontinyu. (Selanjutnya lihat bawah).

Sebagaimana penelitian yang lain, hasil (*outcome*) pada sebuah UKR selalu dalam dua bentuk: data diskrit dan/atau data kontinyu. Dua uji yang dapat digunakan adalah menggunakan uji **chi square** dan **uji t (t-test)**.

Uji Chi Square.

Uji ini dipakai untuk membandingkan dua buah proporsi (data diskrit). *Dummy table* yang dipakai adalah tabel kontingensi 2 X 2 sebagai berikut:

Variabel	Effect		Percent of effect yes	RR (95% CI)	P Value
	Yes	No			
Exposed	a	b	a/(a+b)	[a/(a+b)]/ [c/(c+d)]	
Control	c	d	c/(c+d)	1	-

Risiko relatif

Dari uji *chi square*, kita mendapatkan 4 informasi penting yakni:

- Persentase **kejadian ya** pada kelompok **terpapar** (*exposed*) dan persentase **kejadian ya** pada kelompok **kontrol**.
- Risiko relatif (RR)**, yakni rasio (perbandingan) antara persentase **kejadian ya** pada kelompok **terpapar** (*exposed*) dengan persentase **kejadian ya** pada kelompok **kontrol**.
- Confidence interval (CI) dari nilai RR**. Perhatikan rentang CI apakah menyeberangi angka SATU atau tidak. Jika CI menyeberangi angka satu, maka secara statistika, kedua proporsi tersebut tidak berbeda secara bermakna. Jika CI tidak menyeberangi angka SATU, maka kedua proporsi dikatakan berbeda secara bermakna.
- Nilai P**. Para ahli statistika sepakat bila $P < 0,05$ maka beda proporsi kedua kelompok adalah bermakna.

Perhatikan bahwa risiko relatif adalah rasio antara persentase kejadian ya pada kelompok perlakuan (*exposed*) dengan persentase kejadian ya pada kelompok kontrol, sehingga rumus RR adalah:

$$RR = \frac{a/(a+b)}{c/(c+d)}$$

dengan angka 1 sebagai *referent* atau acuan (lihat contoh).

Rumus 95% *confidence interval* adalah:

$$95\% \text{ CI} = \text{Exp} \{ \text{Ln}(\text{RR}) - 1,96 * \text{SE Ln}(\text{RR}) \} \text{ sampai } \text{Exp} \{ \text{Ln}(\text{RR}) + 1,96 * \text{SE Ln}(\text{RR}) \}$$

di mana: $SE \ln(RR) = \sqrt{\frac{1}{a} + \frac{1}{b} - \frac{1}{a+b} - \frac{1}{c+d}}$

Jika satu atau sel mengandung angka nol, maka perhitungan menjadi sulit, atau jalan ke luarnya setiap sel ditambah dengan 0,5 (a, b, c, d).

Number needed to treat (NNT).

NNT menggambarkan berapa jumlah pasien yang dibutuhkan untuk mendapat satu *outcome* tambahan. Makin sedikit NNT makin baik hasil terapi. Untuk mendapatkan NNT perlu diketahui berapa perubahan *incidence rate* dari kelompok kontrol ke kelompok perlakuan. Perubahan ini disebut *attributable risk* (AR). Jadi AR adalah insidensi hasil positif pada kelompok perlakuan dikurangi insidensi hasil positif pada kelompok kontrol.

Rumusnya adalah:

$$AR = \frac{a}{a+b} - \frac{c}{c+d}$$

$$NNT = \frac{1}{AR}$$

Number needed to harm (NNH).

NNH menggambarkan berapa jumlah pasien yang diberi terapi yang akan memunculkan tambahan satu efek samping. Kebalikan dari NNT, maka makin besar NNH, makin baik hasil terapi.

$$NNH = \frac{1}{AR}$$

Berikut adalah tabel untuk menggambarkan perbandingan dua proporsi, yakni bagaimana pengaruh *R_inserter* dibanding dengan *ring forceps* dalam kejadian ekspulsi dan *removal* pada pemasangan IUD pascasalin.

TABLE Twelve months cumulative events rate based on mode of insertion

Variables	Events Rate		Percent yes	RR (95% CI)	P value
	Yes	No			
Expulsion					
<i>R-inserter</i>	5	98	4.9 %	1.26 (0.35 – 4.57)	0.72
Ring forceps	4	100	3.8 %	1	
Removal					
<i>R-inserter</i>	4	99	3.9 %	0.67 (0.20 – 2.32)	0.53
Ring forceps	6	98	5.8 %	1	

Berapa NNH pada kedua keadaan di atas?

Berdasar rumus NNH maka:

1. Untuk ekspulsi, diperoleh $1/(4,9\%-3,8\%) = 99,99$ atau 100 orang; artinya kejadian tambahan satu kasus ekspulsi, baru akan terjadi setelah 100 kasus dipasang IUD dengan *R-inserter*.
2. Untuk kasus removal, diperoleh $1/(5,8\%-3,9\%) = 52,6$ atau 53 orang, artinya kejadian tambahan 1 kasus *removal* baru akan terjadi setelah 53 orang dipasang IUD dengan *ring forceps*.

Uji t (t-test).

Uji t digunakan untuk membandingkan dua data kontinyu, yakni membandingkan dua *means* (rata-rata). Yang dilihat adalah beda rata-rata (*mean difference*). Karena yang dibandingkan adalah dua buah rata-rata, maka perbedaan kedua *means* dikatakan bermakna bila CI menyeberangi angka NOL, dan tidak bermakna bila CI melewati angka NOL.

Uji t memerlukan syarat yakni distribusi data sebaiknya normal. Ciri data yang berdistribusi normal antara lain adalah:

1. Bentuknya seperti lonceng
2. Mean, median dan mode berdekatan
3. $SD < 30\%$ dari *mean* atau koefisien varians (*rasio SD/means*) $< 30\%$.
4. Rasio kemiringan atau *skewness* (*skewness/SE skewness*) antara -2 sampai +2.
5. Rasio kurtosis (*kurtosis/SE kurtosis*) antara -2 sampai +2
6. Pada *box plot*, garis terletak ditengah
7. Normal QQ plots, data menyebar sekitar garis
8. Pada uji Kolmogorov-Smirnov, $p > 0.05$

Dalam praktek, perhatikan 3 butir pertama saja untuk melihat normalitas distribusi data.

Ada dua bentuk uji t, yakni ***independent sample t-test*** dan ***paired sample t-test***. Yang pertama digunakan jika kedua set data berbeda dalam subyeknya, sedang yang kedua digunakan bila kedua set data mempunyai subyek yang sama tetapi berbeda dalam waktu pengamatan. Rumus t-test adalah sebagai berikut:

Beda *means* dibagi *standard error* beda *means*:

$$t = \frac{x_1 - x_2}{\sqrt{\frac{SD1^2}{N1} + \frac{SD2^2}{N2}}}$$

di mana:

x1 adalah *mean* pada kelompok 1

X2 adalah *mean* pada kelompok 2

SD1 adalah *standard deviation* kelompok 1

SD2 adalah *standard deviation* kelompok 2

N1 adalah jumlah pasien kelompok 1

N2 adalah jumlah pasien kelompok 2

Bandingkan nilai t yang diperoleh ini, dengan nilai t yang ada dalam tabel untuk melihat signifikansinya. Makin besar nilai t makin kecil nilai P sebagaimana terlihat dalam table berikut.

Tabel Konversi nilai t ke dalam nilai P untuk

Nilai t	Nilai P
0,67	0,5
1,28	0,2
1,64	0,1
1,96	0,05
2,33	0,02
2,58	0,01
3,29	0,001

Menghitung *confidence interval* untuk beda *means*.

Confidence interval beda *means* dihitung dengan rumus:

$$95\% \text{ CI} = (X1 - X2) \pm 1,96 \times \sqrt{\frac{SD1^2}{N1} + \frac{SD2^2}{N2}}$$

Dalam praktek, gunakan paket pengolah data (misal SPSS) maka anda langsung mendapatkan nilai P dan *confidence intervalnya*.

Berikut adalah contoh analisis statistika menggunakan ***independent sample t-test***.

Tabel... Pengaruh pemberian asam traneksamat terhadap jumlah perdarahan kala IV, penurunan kadar hemoglobin dan kadar hematokrit.

Variabel	N	Mean	SD	Δ Mean	95%CI	Nilai P
Perdarahan kala IV (ml)						
Asam traneksamat	99	107,24	70,70	-8,86	-29,62-11,89	0,40
Plasebo	99	116,11	77,25			
Penurunan kadar Hb (gdL)						
Asam traneksamat	99	0,99	1,13	-0,06	-0,35-0,22	0,66
Plasebo	99	1,05	0,93			
Penurunan kadar hematocrit (%)						
Asam traneksamat	99	4,06	3,73	-0,52	-1,63-0,59	0,36
Plasebo	99	4,58	4,18			

Uji Mann Whitney (*Mann Whitney U test*).

Uji ini digunakan untuk membandingkan dua set data independent yang **tidak memenuhi syarat untuk uji t**, karena:

1. Data berdistribusi tidak normal.
2. Variansi kedua set data sangat berbeda.
3. Skala pengukurannya adalah ordinal.
4. Jumlah sampel terlalu kecil.

Uji ini serupa dengan ***independent sample t-test*** bila distribusi datanya berdistribusi normal atau *sample size* nya cukup besar.

Contoh: Sebuah penelitian ingin melihat apakah pemberian steroid pada ibu hamil yang menderita *autoimmune thrombocytopenia* dapat meningkatkan jumlah trombosit pada bayi yang dilahirkan?

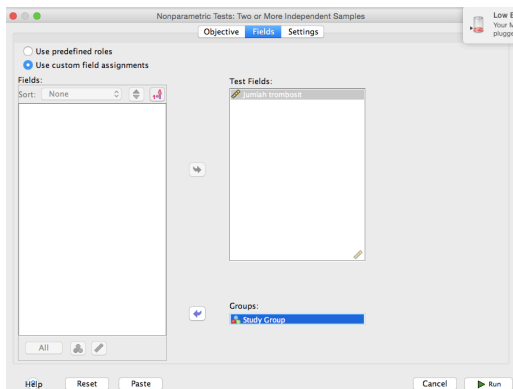
Sebanyak 12 ibu mendapat steroid dan kontrolnya 7 ibu tidak. Trombosit diukur segera setelah bayi lahir. Hasilnya adalah sebagai berikut:

Tabel Jumlah trombosit pada bayi baru lahir (dalam ribuan/ μ L)

Kelompok I: (mendapat steroid)	Kelompok II, kontrol (tidak mendapat steroid)
120	12
124	20
215	112
90	32
67	60
126	40
95	18
190	
180	
135	
399	
65	

Data di atas menunjukkan distribusi yang tidak normal dengan sampel yang terlalu kecil. Karena uji t tidak memungkinkan maka dilakukan uji **Man Whitney U test**. Langkah-langkah dalam SPSS adalah sbb:

1. Buka data di SPSS
2. Klik **Analyze** → klik **Nonparametric test** → klik **Independent samples** → klik **Filed**
3. Pindahkan: **Study group** dalam kotak **Groups** dan **Jumlah trombosit** dalam kotak **Test field**, maka akan muncul jendela sbb:



4. Selanjutnya klik **Run** maka anda akan mendapatkan gambar berikut:

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Jumlah trombosit is the same across categories of Study Group.	Independent-Samples Mann-Whitney U Test	.036 ₁	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

₁ Exact significance is displayed for this test.

Gambar Output *Independent-samples Mann Whitney U test*.

Perhatikan *null hypothesis* ditolak dengan nilai **P 0,036** yang berarti pemberian steroid dapat **meningkatkan jumlah trombosit secara bermakna**.

Wilcoxon rank test.

Uji ini menggantikan *paired sample t-test* karena datanya berdistribusi tidak normal, atau jumlah sampel yang kecil atau bila datanya berskala ordinal. Perhatikan contoh berikut: Seseorang dengan prolaps uteri, titik BA mengalami penurunan. Dengan melakukan operasi histerktomi vaginal, maka titik BA akan bergeser ke atas. Seberapa besar perubahan titik BA terjadi setelah dilakukan tindakan operatif?

Seorang residen obgin ingin menguji hipotesis bahwa dengan operasi, titik BA akan berubah secara signifikan dari yang tadinya di bawah introitus vagina menjadi masuk ke dalam, berada di atas introitus. Dalam satu tahun diperoleh sampel 20 orang, dengan data sebagai berikut:

Tabel ... Perubahan posisi titik BA, 2 minggu setelah operasi

No Kasus	Letak titik BA preop (cm)	Letak titik BA postop (cm)
1	7	-1
2	1	-3
3	1	-1
4	5	0
5	5	-2
6	4	-3
7	2	-3
8	2	-1
9	2	-2
10	5	-1
11	3	-2
12	0	-3
13	2	0
14	1	0
15	2	-1
16	5	-3
17	6	-3
18	2	-3
19	5	-4
20	4	0
<i>Mean</i>	3,20	-1,80
<i>SD</i>	1,96	1,28

Data tersebut berdistribusi tidak normal dengan sampel yang relatif kecil. Meskipun demikian, uji t bisa saja dikerjakan kalau kita ingin melihat seberapa besar perubahan letak titik BA, yakni dengan membandingkan nilai *mean* keduanya. Karena data tersebut sama subyek, maka digunakan *paired sample t-test*.

Dengan *paired sample t-test* diperoleh perbedaan *mean* sbb:

Variabel	N	Mean	SD	Δ Mean	95%CI	P
Posisi BA (cm)						
Pre-operasi	20	3,20	1,96	5,00	3,88 - 6,12	0,00
Post-operasi	20	-1,80	1,28			

Kalau kita menggunakan analisis non parametrik, maka *Wilcoxon sign rank test* adalah jenis uji statistika yang tepat karena uji ini digunakan untuk dua data yang berpasangan. Bagaimana langkah-langkah dalam SPSS?

1. Buka data di SPSS
2. Klik **Analyze** → klik **Nonparametric test** → klik **Related samples** → klik **Filed**
3. Pindahkan: **BA pre** dan **BA post 2 minggu** dalam kotak kotak **Test fields**.
4. Klik **Run** maka akan muncul hasil sbb:

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between BA Pre and BA post 2 minggu equals 0.	Related-Samples Wilcoxon Signed Rank Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Ternyata *null hypothesis* ditolak, artinya terjadi perubahan letak titik BA secara bermakna.

Kepustakaan

1. Friedman LM, Furberg CG And Demets DL. Fundamental Of Clinical Trials. John Wright PSC, Inc., Boston 1983
2. Pocock SJ. Clinical Trial, John Wiley & Sons, New York, 1983
3. Higgins JE. Introduction to Clinical Trials Part I To IV. Family Health International. Research Triangle Park, NY, 1987
4. Siswosudarmo R, Kurniawan K, Suwartono H, Alkaff TR, Anggraeni M. THE USE OF NEW INSERTER (R_INSERTER) FOR DELIVERING CuT-380A IUD DURING POSTPARTUM PERIOD. Phase II Clinical Trial. J Kesehatan Reproduksi. 2014;1(3):189–95).