

# Rationel Klinik

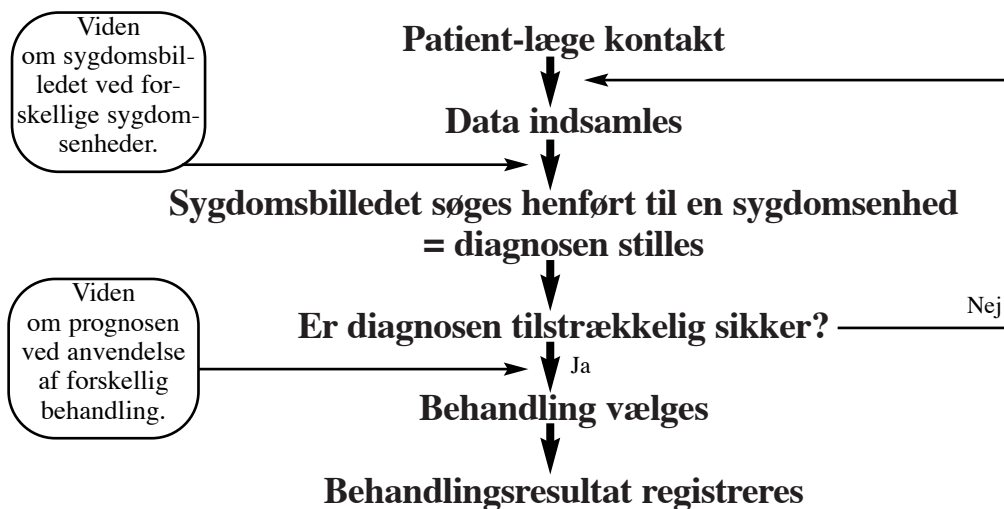
af

*Christian Lildal Carranza*



## RATIONEL KLINIK, KAPITEL 1

Den kliniske beslutningsproces:



Klinikerens ræsonnementer kan være *teoretiske* eller *empiriske*:

- **Teoretiske** ræsonnementer er deduktive slutninger fra teoretisk viden om sygdomsårsager og behandlingers virkningsmekanismer.
- **Empiriske** ræsonnementer er slutninger fra erfaringer fra tidligere ptt.

Beslutninger består af en *naturvidenskabelig* og en *humanistisk* komponent:

- **Naturvidenskabelig** komponent: De teoretiske og empiriske ræsonnementer.
- **Humanistisk** komponent: omfatter *empatisk-hermaneutisk* ræsonnementer, dvs. slutninger baseret på en forståelse af patienten som medmenneske og *etiske* ræsonnementer ud fra etiske normer.

Sygdomsdata:

- **Subjektive symptomer:** er patientens følelser og stemning.
- **Objektive symptomer:** er patientens eller de pårørendes observationer om patientens krop og dens produkter.
- **Kliniske fund:** er lægens observationer ved den objektive undersøgelse af patienten. De fund, der i virkeligheden er subjektive symptomer og først bemærkes af patienten ved den objektive undersøgelse, kan man kalde *provokeredes symptomer*.
- **Parakliniske fund:** er resultater af undersøgelser. De er enten *deskriptive* eller *kvantitative*.
- **Andre oplysninger:** f.eks. tidligere sygdomme, levevis, boligforhold mm.

Helhedsvurderinger af den kliniske tilstand:

Både til forskningsmæssig brug og i dagligt klinisk arbejde kan det være nyttigt at formalisere helhedsvurderinger ved angivelse af et *klinisk index*. Til dette formål kan man anvende simple rangskalaer eller mere eller mindre udviklede scorings-systemer. F.eks.: bedre/uændret/værre, NYHA (klassifikation af funktionskapaciteten ved hjerteinsufficiens), Apgar score (vurdering af nyfødte) og APACHE-score (anæstesiologisk vurdering af tilstanden og prognosen for kritisk syge patienter).

Iatrotropi-begrebet:

**Iatrotrope symptomer:** de subjektive eller objektive symptomer, som bevirker, at patienten går til læge. Disse er særdeles vigtige, da de i patientens øjne er det problem, som lægen skal løse.

**Ikke-iatrotrope symptomer:** de øvrige symptomer, som først kommer frem under udspørgningen.

**Iatrotropitærsklen:** grænsen for at patienten søger læge. En problemfyldt tilværelse kan sænke tærsklen.

Måleskalaer:

- **Nominalskala:** en "liste", som benyttes til at klassificere observationer i navngivne kategorier. Skalaen skal ideelt opfylde tre krav:

- 1) Hver kategori skal være *veldefineret*.
- 2) Skalaen skal være *eksklusiv*, så ingen observation kan registreres i mere en én kategori.
- 3) Klassifikationen skal være *ekshaustiv*, så at enhver observation kan klassificeres i en af kategorierne.

En skala, der består af 2 klasser, kaldes en *binær skala*.

• **Rangskala:** f.eks. Beauforts skala til måling af vindstyrke. Det giver ingen mening at beregne gennemsnit.

• **Intervalskala:** højeste måleniveau. Alle skalatrin er lige store. Den kan være:

- *Kontinuert*.
- *Diskontinuert* - hvis der kun bruges "hele tal".
- *Ratioskala* - undtagen for skalaer med et arbitrært nulpunkt.

En intervalskala kan altid reduceres til en rangskala eller en binær skala.

**Diagnostisk datatransformation:**

Konkretisering af et symptom bør normalt kun føre til, at klagen benævnes med et neutralt fagudtryk, dvs. et udtryk, som ikke indebærer, at man kender årsagen til klagen. Hvis man gør det modsatte foretager man en *diagnostisk datatransformation*, som kan medvirke til at give klinikerne skyklapper på. Er meget almindelig inden for radiologien.

## RATIONEL KLINIK, KAPITEL 2

**Reproducerbarhed:**

Upålidelige resultater, idet de svinger uforudsigeligt omkring den sande værdi. De frembyder en såkaldt *tilfældig variation*, men udregner man gennemsnittet, får man næsten den sande værdi. Kaldes også nøjagtigheden eller præcisionen.

**Rigtigheden:**

Upålidelige resultater, idet de hver gang er for høje/lave. De er med andre ord behæftet med en *systematisk fejl*, men de er reproducerbare. Kaldes også accuracy.

**Bias:**

Skævhed eller afvigelse fra den rette kurs.

**Standarddeviationen (s)** =  $\sqrt{(\sum(x-\bar{x})^2)/(N-1)}$

**AUC:**

- Gennemsnittet  $\pm$  1 standarddeviation udgør 2/3 af det samlede areal.
- Gennemsnittet  $\pm$  2 standarddeviationer udgør ca. 95% af det samlede areal.

Standarddeviationen bliver et mål for *reproducerbarheden* under den vigtige forudsætning, at fordelingen er gaussisk.

Hvis der foreligger en systematisk fejl, dvs. analysen er behæftet med bias, vil det beregnede gennemsnit afvige fra den sande værdi. I så fald benyttes differensen mellem gennemsnit og sand værdi som mål for *rigtigheden*.

**Interobservatørvariation:** variation mellem vurderingen fra forskellige observatører af samme fænomen.

**Intraobservatørvariation:** variation mellem vurderingen af den samme observatør på forskellig tidspunkter af samme fænomen.

**Kappa-statistik:**

Tillader en vurdering af reproducerbarhedens størrelse uafhængigt af rigtigheden. Kappa kan i princippet variere mellem -1 og +1. Kappa er +1, hvis der er perfekt overensstemmelse, den er positiv, hvis overensstemmelsen ikke er perfekt, men større end den tilfældigt forventede, og den er 0, hvis overensstemmelsen netop er den tilfældigt forventede. En negativ kappaværdi er udtryk for en endnu lavere overensstemmelse, dvs. en systematisk uenighed mellem observatørerne.

**Kappa** =  $(P_0 - P_1)/(1 - P_1)$

Eksempel:

		2. vurdering:		
		Positive	Negative	Ialt
1. vurdering:	Positive	11	11	22
	Negative	3	35	38
	Ialt	14	46	60

		2. vurdering:		
		Positive	Negative	Ialt
1. vurdering:	Positive	5,13	16,87	22
	Negative	8,87	29,13	38
	Ialt	14	46	60

**Den observerede overensstemmelse  $P_0$ :**  $(11+35)/60 = 0,77$ .

Frekvensen af positive fund ville være  $22/60$  og  $14/60$ , og sandsynligheden for 2xpositiv ville være  $22/60 * 14/60$ . Da der var 60 observationer, ville man forvente  $22/60 * 14/60 * 60 = 5,13$ , der betegnedes som positive begge gange ved tilfældighedernes spil. Ligeledes:  $38/60 * 46/60 * 60 = 29,13$ .

**Den forventede tilfældige overensstemmelse  $P_1$ :**  $(5,13+29,13)/60 = 0,57$ .

**Kliniske fund:**

**Reproducerbarheden:** gennemgående ret lav.

**Rigtigheden:** kan kun vurderes, hvis man råder over en facitliste.

**Fordelingstyper:**

- Højre/venstre-skæv fordeling.
- Bimodal fordeling.
- Klokkeformet fordelingskurve, der ikke er gaussisk.
- Beskåret (trunkeret) fordeling.
- Histogram over resultater af et ringe antal forsøg.
- Gaussisk.

**Normal-områder:**

Omfatter ideelt set 95% af den raske befolkning, hvilket vil sige, at hver 20. raske person må forventes, at have en værdi, der enten er "abnormt" høj eller lav.

*Statistikeren:* definerer det normale som det almindelige og det abnorme som det sjældne, men dette synspunkt er indlysende utilstrækkeligt.

Meget svært at definere normalen for biologisk funktion (se Med. Filosofi kap. 4).

**Regression mod gennemsnittet:**

Hvis man i det daglige arbejde observerer et tilfældigt højt resultat, vil en ny måling med stor sandsynlighed give et lavere resultat (vise versa).

**Indikatorer:**

- *Diagnostiske indikatorer:* sygdomsdata, der taler for eller imod en eller anden diagnose.
- *Prognostiske indikatorer:* sygdomsdata, der tjener til at forudsige et sygdomsforløb. Hvis de tjener til at forudsige udviklingen af sygdom hos raske mennesker, benyttes ofte i stedet udtrykket *risikomarkører*.
- *Kliniske indikatorer:* sygdomsdata, der benyttes til at vurdere patientens kliniske tilstand på et givet tidspunkt eller til at monitorere forløbet af patientens sygdom.

**Somatisering:**

Personlige problemer kan sænke iatrotropitærsklen, og det kan let føre til, at der stilles en somatisk diagnose.

### RATIONEL KLINIK, KAPITEL 3

**Sygdoms-klassifikation:**

Ideelt: et sæt af kriterier, som altid er opfyldt af de patienter, der siges at have den pågældende sygdom, og som aldrig er opfyldt af dem, der ikke siges at have sygdommen.

Tjener til i koncentreret form at viderebringe klinisk viden og forskningsresultater.

**Sygdoms-  
enheder.:**

- **Symptomdiagnoser:** simpleste, benyttes som dels foreløbige diagnoser, dels som eksklusionsdiagnoser.

- **Syndromer:** definerede ved det samlede sygdomsbillede, idet det definitoriske kriterium er tilstedeværelsen af en bestemt konstellation af sygdomsdata. Ved redefinitioner afgrænser den gamle og den nye definition ikke ganske de samme patienter.

Komplekst syndrom: ikke alle patienterne frembyder samme kombination af symptomer og fund i modsætning til et simpelt syndrom.

- **Patoanatometisk definerede:** sygdomsnavnet er samtidig betegnelsen på en patoanatometisk læsion. De definitoriske kriterier kan inddeles:

Tilgængelige kriterier: forstås sådanne, hvis tilstedeværelse eller fravær kan påvises på det tidspunkt, hvor man søger at stille diagnosen.

Skjulte kriterier: er kendetegnet ved, at man ikke direkte kan be- eller afkræfte deres tilstedeværelse ved diagnosetidspunktet. Den teoretiske definition må suppleres med en arbejdsdefinition, der fortæller, hvornår man i praksis skal stille diagnosen. Den vil have form af en kompleks syndromdefinition.

Retrospektivt tilgængelige: skjulte kriterier, der bliver tilgængelige efter den diagnostiske beslutning.

- **Fysiologisk eller metabolisk definerede:** defineres ud fra kendskabet til patogenesen.

- **Ætiologisk definerede:** omfatter først og fremmest infektionssygdomme, genetisk definerede sygdomme, forgiftninger og sygdomme defineret ved en miljøfaktor.

**Ætiologi:**

Sygdomsprocessen sættes i gang af en kombination af genetiske faktorer og miljøfaktorer af enhver art. Disse kaldes tilsammen sygdommens *ætiologi*.

**Patogenesen:**

Herefter følger en kædereaktion af forandringer i organismen, som kan beskrives i fysiologiske, biokemiske, immunologiske eller morfologiske termer, og som tilsammen udgør sygdomsmekanismen eller *patogenesen*.

**Årsagsfaktorer:**

- **Nødvendig:** faktor X siges at være en *nødvendig* årsag til Y, hvis Y altid forudgås af X (dvs. Y ville ikke være indtrådt, hvis X ikke havde fundet sted).

- **Tilstrækkelig:** faktor X siges at være en *tilstrækkelig* årsag til Y, hvis X altid efterfølges af Y.

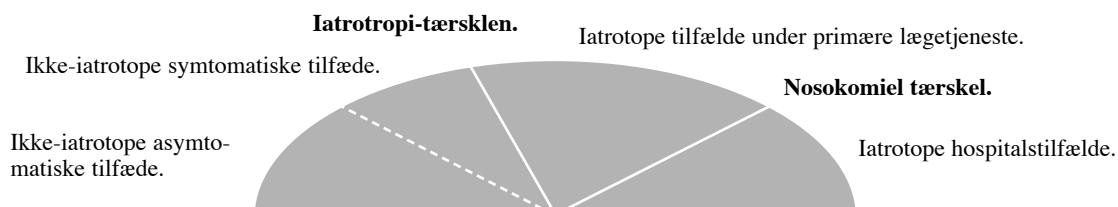
- **Medvirkende:** hvis en faktor X medfører en øget hyppighed af Y, selv om X ikke altid medfører Y, og selv om Y ikke altid er forudgået af X.

**Det kliniske  
spektrum:**

1) **Ikke-iatrotrope tilfælde:** pt. har ikke søgt læge, enten på grund af manglende symptomer, eller fordi symptomerne ikke har nået iatrotropitærsklen.

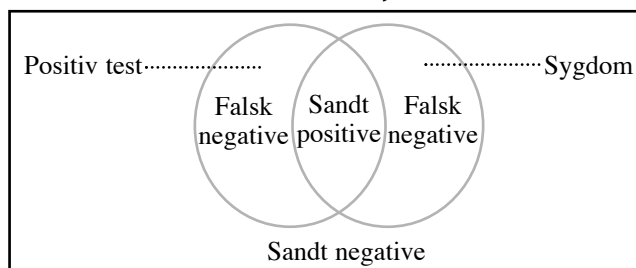
2) **Iatrotrope tilfælde,** der behandles af praktiserende læger.

3) **Iatrotrope tilfælde,** der har passeret den nosokomielle tærskel og ses af hospitalslæger.



## RATIONEL KLINIK, KAPITEL 4

Det  
diagnostiske  
univers:



		Sygdom		
		Til stede	Ikke til stede	
Diagnostisk test	Positive	Sandt positive	Falsk positive	Alle med positiv test
	Negative	Falsk negative	Sandt negative	Alle med negativ test
		Alle med sygdom	Alle uden sygdom	

De vandret beregnede hyppigheder betegnes **diagnostiske rater**, da klinikerer i den diagnostiske situation har brug for at vurdere sandsynligheden for, at en patient har eller ikke har sygdommen ud fra testresultatet.

De lodret beregnede hyppigheder betegnes **nosografiske rater**, da de først og fremmest tjener til at beskrive sygdommen.

Den **diagnostiske SP** sandsynlighed = sandt positive/alle positive =  $P(S+|T+)$ .

Den **diagnostiske FP** sandsynlighed = falsk positive/alle positive =  $P(S-|T+)$ .

Den **diagnostiske SN** sandsynlighed = sandt negative/alle negative =  $P(S-|T-)$ .

Den **diagnostiske FN** sandsynlighed = falsk negative/alle negative =  $P(S+|T-)$ .

Den **nosografiske SP** sandsynlighed = sandt positive/alle syge =  $P(T+|S+)$ .

Den **nosografiske FN** sandsynlighed = falsk negative/alle syge =  $P(T-|S+)$ .

Den **nosografiske SN** sandsynlighed = sandt negative/alle raske =  $P(T-|S-)$ .

Den **nosografiske FP** sandsynlighed = falsk positive/alle raske =  $P(T+|S-)$ .

Der er to metoder til vurdering af diagnostiske test.

**Direkte metode:**

Man udfører testen på patienter, der frembyder samme sygdomsbillede, hvorefter man be- eller afkræfter diagnosen ved hjælp af en anden, helt sikker metode. Så kan man udregne den diagnostiske SP sandsynlighed,  $P(S+|T+)$  og den diagnostiske SN sandsynlighed,  $P(S-|T-)$ .  $P(S+|T+)$ , fortæller klinikerer, hvor god testen er til at stille diagnosen, medens  $P(S-|T-)$  fortæller, hvor god testen er til at udlukke diagnosen.

**Inddirekte metode:**

= nosografiske metode. Man udgår fra to stikprøver. En gruppe med sygdommen og en gruppe uden sygdommen. Dernæst bestemmer man udfaldet af den diagnostiske test i disse to grupper, hvorefter man udregner den nosografiske SP rate,  $P(T+|S+)$  og den nosografiske SN rate,  $P(T-|S-)$ .

Denne metode har den ulempe, at man kommer til at ignorere de tilfælde, hvor den

diagnostiske test gav et usikkert resultat (T?), eller hvor det ikke lykkedes at stille en sikker diagnose (S?).

### Bayes' formel:

$$P(S+|T+) = P(T+|S+) * (P(S+)/P(T+))$$

$P(S+)$  er den ubetingede sandsynlighed for, at en patient fra dette univers har sygdommen (sygdomsprævalensen).  $P(T+)$  er den ubetingede sandsynlighed for, at en patient fra universet frembyder en positiv test.

$$P(S+|T+) = (P(T+|S+)*P(S+)) / (P(T+|S+)*P(S+)+P(T+|S-)*P(S-))$$

**Specificitet** = nosografisk SN rate, dvs.  $P(T-|S-)$ .

**Sensitivitet** = nosografisk SP rate, dvs.  $P(T+|S+)$ .

**Prædiktiv værdi af positiv test** = diagnostisk SP rate, dvs.  $P(S+|T+)$ .

**Prædiktiv værdi af negativ test** = diagnostisk SN rate, dvs.  $P(S-|T-)$ .

Specificiteten og sensitiviteten angives at være stabile mål for en tests evne til at diagnosticere en sygdom, medens de prædiktive værdier siges at afhænge af sygdomsprævalensen.

Det giver kun mening at diskutere, om testen er god eller dårlig til at diagnosticere sygdommen i et bestemt diagnostisk univers, dvs. hos patienter med et *veldefineret sygdomsbillede* fra en *bestemt del af det kliniske spektrum*.

### Diagnostik med skjult facitliste:

Der er problemer ved diagnosen af sygdomme med skjult definatorisk kriterium, da man aldrig kan være sikker på at diagnosen er korrekt. Man kan dog tage underbyggende tests, som kan være med til at sikre diagnosen.

Man kan vurdere resultaterne af den diagnostiske test over for resultaterne af en anden test, som man har større tillid til *guldstandard*. Guldstandard kan også bestå i en samlet vurdering af alle undersøgelsesresultater indhentet inden for et vist tidsrum.

Der er problemer ved diagnosen af kliniske syndromer, da:

- 1) Klinikeren vil ofte overveje, hvilken syndromdiagnose, der er den sandsynligste på det foreliggende grundlag, og herefter vil han ordinere de undersøgelser, der er nødvendige for at komplettere sygdomsbilledet med henblik på diagnosen af netop dette syndrom.
- 2) De fleste syndromer er ikke veldefinerede, og selv om de var det, ville man som kliniker ofte stå over for patienter, der kun delvis opfyldte en eller anden opstillet syndromdefinition.
- 3) Syndromdefinitioner ændrer sig. Unødige skred i diagnostiske kriterier har den meget uheldige konsekvens, at erfaringerne fra randomiserede behandlingsforsøg forældes. Hvis de diagnostiske kriterier ikke længere svarer til de indgangskriterier, der benyttes, da forsøget udførtes, kan man ikke længere anvende resultaterne.
- 4) Man kan ikke bedømme anvendeligheden af en test til syndromdiagnostik ved at beregne den diagnostiske SP og den diagnostiske SN sandsynlighed, da man selv bestemmer den diagnostiske værdi.

Når vi taler om diagnostiske og nosografiske sandsynligheder, anvender vi det traditionelle *frekventielle* eller *statistiske sandsynlighedsbegreb*, idet sandsynligheden er den "sande" hyppighed i populationen.

Det *dagligdags* eller *subjektive sandsynlighedsbegreb*, hvor sandsynligheden angiver vor tiltro til et eller andet fænomen. Kan godt baseres på den frekventielle.

### Diagnostisk strategi:

- *Deduktiv tænkning.*
- *probabilistiske overvejelser.*

## RATIONEL KLINIK, KAPITEL 5

Klinikerens tiltro til behandlingen kan begrundes ved:

- 1) *Deduktion fra teori.*
- 2) *Ukontrolleret erfaring.*
- 3) *Kontrolleret erfaring.*

**Fejlkilder ved  
ukontrolleret  
erfaring:**

- *Manglende kendskab til sygdommens spontanforløb:* patienter søger ofte hjælp, når deres symptomer er værst.

- *Måleusikkerhed og regression mod gennemsnittet:* en tilfældig høj/lav værdi vil rent statistisk ofte efterfølges af en lav/høj værdi.

- *Tilfældighedernes spil:* hvis erfaringen er baseret på et *begrænset patientmateriale* vil tilfældet kunne påvirke resultatet. Begrebet “**95% sikkerhedsinterval**” kan være med til at beskrive erfaringen, da *bredden af sikkerhedsintervallet afhænger af antallet af undersøgte patienter*. Et grøft skøn over øvre og nedre sikkerhedsgrænse fås ved at udgå fra følgende formel: **observerede helbredelsesprocent  $\pm 100/\sqrt{N}$** . Med ord: en helbredelsesprocent på X vil være en sjælden hændelse (5% = 1 ud af 20), hvis den sande helbredelsesprocent lå under laveste sikkerhedsgrænse eller over den højeste.

- **Bias:**

- ofte inkluderes ikke alle patienter og behandlingsindikationerne kan meget vel have ændret sig undervejs i undersøgelsen.
- man har en tilbøjelighed til at huske netop de tilfælde, som frembød et usædvanligt forløb.
- *psykologisk bias:* lægen ønsker, at det skal gå hans patient godt, og håber derfor, at den behandling, han giver, vil hjælpe.

- *Placeboeffekten:* en terapeutisk effekt af hele behandlingssituationen, inklusive læge-patient forholdet, dvs. en gavnlig effekt, som ikke kan forklares ved direkte indgriben i fysiologiske eller metaboliske processer.

Man kan tale om en “negativ” placeboeffekt eller en *noceboeffekt*, hvis patientens tilstand i stedet forværres.

Ontologisk:

- *Psyko-somatisk teori* er knyttet til sjæl-legemeteteorier, der forudsætter, at mentale fænomener som angst, sult, smerte og forventning kan spille en kausal rolle.
- En *betinget refleks:* placeboeffekten på smerter har et forløb, som svarer til forløbet efter indgift af analgetikum.

Epistemologisk:

- Muligheden for at erkende og måle en evt. placeboeffekt giver betydelige problemer.

Lægens holdning, lægens egen tiltro til en behandling, patientinformationene og mange andre forhold har en betydelig indflydelse på patientens oplevelse af sin tilstand og de subjektive symptomer.

Man må formode, at den læge, der søger at indleve sig i sin patients situation, viser interesse for patienten som medmenneske og i det hele taget giver patienten en grundig information, vil indhøste langt større placeboeffekter end den kollega, der udskriver en recept på et uvirksomt præparat.

**Konklusion:**

Den *ukontrollerede erfaring* er yderst upålidelig.

**Alternativ behandling:**• **Homøopati:**

**Grundidé:** sygdomme skal behandles med medicin, som hos raske fremkalder de symptomer, som kendetegner den pågældende sygdom. Homøopatisk medicin skal fortyndes så mange gange, at en patient, der får en dosis af det pågældende præparat, ikke kan være sikker på at indtage et eneste molekyle.

• **Kiropraktik:** Oprindeligt antog man at al sygdom var fremkaldt af små forskydninger (subluksationer) i rygsøjlen, og behandlingen bestod i at korrigere disse ved hjælp af manipulation. Nu bekender de sig til det etablerede videnskabssyn. Undersøgelser har kun vist beskeden effekt.

**Problemerne:**

- Behandlingens effekt er kun bevist ved at behandleren har erfaring for, at patienterne får det bedre.
- Det er ikke nok at "kineserne har haft en tusindårig erfaring" med behandlingen.
- Alternativ medicin er *holistisk*, dvs. tager hensyn til både legemlige og sjælelige faktorer, hvorimod lægevidenskaben er *reduktionistisk*, dvs. udelukkende anskuer mennesket som en biologisk organisme. Nogle læger interesserer sig for livkvalitet og fagets etiske aspekter. Nogle alternative teorier, går ud på, at alskens sygdom kan helbredes på én bestemt metode.
- Zoneterapeuter udsætter ikke deres grundlæggende teori for kritiske falsifikationsforsøg.
- Styrken ved den homøopatiske teori er, at det aldrig har været nødvendigt at ændre den.

**Forskellen mellem alternativ og videnskabelig er ikke:**

- Videnskabelig medicin nærmest per definition er det, som autoriserede læger udøver, medens alternativ medicin er det, som ikke-læger dyrker.
- Den alternative medicins teorier er forkerte, medens den etablerede lægevidenskabs teorier er sande - mange af de teorier, som læger tror på, bliver senere modbevist.
- De alternative behandlingsmetoder er uvirksomme, hvorimod de behandlingsmetoder, der anvendes af læger, er virksomme - læger anvender utvivlsomt mange virkningsløse behandlinger.

**- men:**

Den afgørende forskel består først og fremmest i udøvernes holdning til deres teorier og behandlinger. Popper: "*Det der mere end noget andet karakteriserer videnskabsprocessen, må være videnskabsmandens kritiske holdning.*"

*Den alternative medicin er et anakronistisk islæt i den moderne sygdomsbehandling.*

**RATIONEL KLINIK, KAPITEL 6****Det randomiserede behandlingsforsøg:**

- **Patientmaterialet:** 3 betingelser skal opfyldes, før det gavner ens egne patienter:
  - 1) Indgangskriterierne må være veldefinerede.
  - 2) Patienterne i forsøget må udgøre en tilfældig stikprøve af alle patienter, der opfyldte indgangskriterierne.
  - 3) Ens egne patienter må ligne de patienter, der indgik i forsøget.

*Indgangskriterier:* f.eks. symptomer, kliniske fund og parakliniske fund.

*Eksklusionskriterier:* f.eks. graviditet eller komplicerede sygdomme.

Patientmaterialet bør i videst muligt omfang omfatte alle patienter, der i forsøgsperioden opfyldte kriterierne for indgang i forsøget, da det er betingelsen for, at materialet er repræsentativt (for at undgå *stikprøvebias*).

Stikprøvestørrelsen, skal være fastlagt før forsøgets start.

- **Randomiseringen:** tilfældig fordeling af patienterne på de to behandlingsgrupper

ved en eller anden form for lodtrækning, for at tilstræbe, at patienterne i de to grupper i videst muligt omfang er sammenlignelige. Jo større patienttallet er, jo mere vil forskellene grupperne imellem blive udjævnet.

*Blokrandomisering:* tjener til at sikre, at undergrupper af patientmaterialet er ligeligt repræsenteret i de to behandlingsgrupper.

*Stratificering:* randomiseringen er sådan indrettet, at behandlingsgrupperne bliver sammenlignelige, hvad angår en eller flere faktorer, som skønnes at have særlig betydning for prognosen.

Forsøg, hvor randomiseringen ikke er udført ved hjælp af en tilfredsstillende metode, tenderer til at overvurdere effekten af en ny behandling.

- **Valg af behandling i kontrolgruppen:** en ny behandling af en bestemt sygdom skal sammenlignes med den behandling, der hidtil har vist at have bedst effekt.

Normalt er der ingen grund til at inføre den nye, hvis den ikke er bedre end den hidtidige, ved mindre at den nye har færre bivirkninger eller er billigere end den gængse.

- **Blinding:**

- *simple dobbeltblinde forsøg:* den ene gruppe patienter får den nye tabletbehandling, A, og den anden får den gængse, B. Hverken patienterne eller de behandlende læger ved, hvem der får hvad.

- *dobbeltblindt med åben terapeutisk styring:* en læge, der ikke har kontakt med patienterne, styrer terapien ved om nødvendigt at øge eller sænke dosis og samtidig på skrømt ændre doseringen af blindtabletter til patienterne i kontrolgruppen.

- *enkeltblinde forsøg:* hvor patienten, men ikke lægen er blindet, eller hvor lægen, men ikke patienten er blindet.

- *åbne forsøg:* bør kun anvendes, når andre muligheder er udelukkede.

*Double dummy teknik:* behandling A + placebo B kontra behandling B + placebo A, for at modvirke at pillernes udseende kan årsage bias.

- **Parret design:**

*Gruppesammenligningsforsøg:* patienterne er fordelt på to grupper, der får forskellig behandling.

- **Uparret design:**

*Overkrydningsforsøget:* patienten er sin egen kontrol, idet hver patient prøver både den nye og den gængse behandling i to på hinanden følgende perioder.

Kræver at det drejer sig om en sygdom med et langvarigt forløb og en behandling, hvis virkning indtræder hurtigt og ophører ved seponeringen (symptomatisk behandling).

*Efterslæbseffekt:* det fænomen, at behandlingseffekten ikke ophører ved seponeringen, men fortsætter et stykke tid herefter.

*Periodeeffekt:* en generel tendens til, at patienterne har det bedst i den anden periode. De søgte læge, da de havde det værst, og får det spontant bedre uanset behandlingen.

- **Effekt mål:** skal være så **pålidelige** som muligt, hvad angår både reproducerbarhed og rigtighed, men herudover må man som læser overveje deres **kliniske relevans**.

*Surrogat mål:* man vælger ikke det klinisk relevante effekt mål, men som erstatning herfor vælger man et andet effekt mål, som man formoder er en tilstrækkelig god indikator for det kliniske relevante mål.

Artikler om randomiserede forsøg bør ikke kun indeholde oplysninger om de registrerede bivirkninger, men også om deres sværhedsgrad.

- **Stopperegul:** en regel for, hvornår man skal ophøre med at inkludere flere patienter i forsøget.

Hvis man sammenligner to behandlinger, A og B, der er lige effektive, og gør resultaterne op, hver gang en patient har gennemført behandlingen, vil der med stor sandsynlighed være tidspunkter, hvor den nye behandling har givet de bedste resultater, blot som følge af tilfældighedernes spil.

Hvis der er tale om et dobbelblindt forsøg kan man etablere en uafhængig, ikke-blindet sikkerhedskomite. Denne kan så regelmæssigt inspicere forsøgsresultaterne og stoppe forsøget, hvis forskellen mellem behandlingsresultaterne er særlig stor, eller hvis den ene behandling har givet særlig mange bivirkninger.

- **Drop-outs:** frafald, f.eks. patienter, der frembød bivirkninger og derfor måtte tages ud af forsøget, patienter, der ikke mødte til de aftalte kontroller, og patienter, der ikke tog den ordinerede medicin.

*Beslutninger af denne art skal altid taget "blindet".*

Resultaterne kan derefter opgøres på to måder:

- *per protokol-analyse:* som udelukkende omfatter de patienter, der fuldførte forsøget som planlagt, men der opstår den mulighed, at resultatet bliver biased på grund af drop-outs.

- *intention-to-treat analyse:* som består i, at man ved analysen medtager alle de randomiserede patienter, dvs. alle de patienter, som man agtede at behandle, uanset om de gennemførte forsøget eller ej.

I praksis vil det undertiden være bedst at foretage en per protokol-analyse og flere intention-to-treat analyser, og er der så mange drop-outs, at disse analyser giver forskelligt resultat, kan man i værste fald være nødt til at acceptere, at forsøget var mislykket.

### Hypotese-testning:

**Nulhypotesen:** de to behandlinger er lige gode, hvilket betyder, at den observerede forskel mellem behandlingsresultaterne tilskrives tilfældighedernes spil.

**Alternativ hypotesen:** de to behandlinger har forskellig effekt.

Den ukendte sandhed og klinikerens valg skaber i alt fire muligheder:

- 1) *Forkasning af en falsk nulhypotese - ingen fejl.* Klinikeren handler som om A er bedre end B og sandheden er, at A er bedre end B.
- 2) *Forkastning af en sand nulhypotese - type I fejl.* Klinikeren handler som om A er bedre end B og sandheden er, at de to er lige gode. "**Optimistfejl**".
- 3) *Acceptering af en falsk nulhypotese - type II fejl.* Klinikeren handler som om de to behandlinger er lige gode og sandheden er at A er bedre end B. "**Pessimistfejl**".
- 4) *Acceptering af en sand nulhypotese - ingen fejl.* Klinikeren handler som om de to behandlinger er lige gode og sandheden er at de to er lige gode.

### Signifikanstest:

Tjener til at beregne sandsynligheden for at opnå de observerede resultater eller endnu mere ekstreme resultater.

$P = 0,0006$ : hvis nulhypotesen er sand, vil der være en sandsynlighed på 0,6% for at opnå en forskel mellem helbredelsesprocenterne på 10% (forskellen mellem de to helbredelsesprocenter) eller mere.

**Statistisk signifikant:** hvis P-værdien er mindre end 5%. Kan være relevant at bruge 1% i stedet.

Sandsynligheden for at opnå de observerede resultater, hvis nulhypotesen er sand = p-værdien, der er en frekventiel sandsynlighed.

Diagnostisk vigtigere er: sandsynligheden for at nulhypotesen er sand, givet de opnåede resultater = min tiltro til nulhypotesen = subjektiv sandsynlighed.

*Tommelfingerregel:* man skal fortsat nære størst tiltro til nulhypotesen, hvis P-værdien er større end forhåndstiltroen til behandlingens effekt.

### Terapeutisk gevinst:

Forskellen i "helbredelsesprocent" ved A i forhold til B.

Beregning af 95% sikkerhedsgrænser for den observerede forskel:

$$TG \pm 2 * \sqrt{((p1(1-p1))/N + (p2(1-p2))/N)}$$

TG er den observerede terapeutiske gevinst (p1-p2).

Beregning af sikkerhedsgrænser på den terapeutiske gevinst er særligt afslørende, hvis patientantallet er lille.

**Number needed to treat:** NNT. En angivelse af det nødvendige antal, der skal behandles for at helbrede én patient mere. Beregnes som den *reciprokke værdi af den terapeutiske gevinst*.

### Feljkilder/ problemer ved randomiserede forsøg:

- Fejl i forsøgsdesignet. Specielt ved randomiseringen eller blindingen.
- Hvis sygdommen eller tilstandene er så sjældne, at det er umuligt at samle et tilstrækkeligt antal patienter.
- Behandlingen ikke virker lige godt på alle patienter med en bestemt sygdom.
- Maligne sygdomme, som i dag altid er dødelige, og hvor patienter vil være utilbøjelige til at acceptere en lodtrækning.
- Skævvridning af den kliniske forskningsindsats som følge af en kommercialisering.
- Kræver altruisme - pt. er individualismen i højsædet + ulyst til altruisme, hvis der er et kommercielt islæt i forsøget.
- Forskende klinikere behandler de deltagende patienter som forsøgsobjekter og ikke som samarbejdspartnere.

### Problemer ved subgruppe-analyser:

- Man vil på grund af tilfældighedernes spil finde forskelle undergrupperne imellem, blot man kigger grundigt nok.

### Meta-analyser:

resultaterne af en systematisk forskning i forskningsresultater.

Ved en speciel statistisk procedure vægtes de største forsøg mest.

### Problemer:

- Er de foretagne forsøg tilstrækkeligt sammenlignelige?
- Hvilke kvalitetskrav vil man stille?
- *Publikationsbias:* der er størst sandsynlighed for, at forsøg, der har givet et signifikant resultat bliver publiceret.
- Der indgår subjektive skøn ved udførelsen af analysen.
- Forskellige forfattere benytter forskellige statistiske metoder, og nogle af dem begår statiske fejl.

## RATIONEL KLINIK, KAPITEL 7

### Grundlaget for kliniske beslutninger:

#### Naturvidenskabelig tænkning:

##### 1) *Den teoretiske komponent*

Slutninger fra teoretisk viden om sygdomme og sygdomsmekanismer.

##### 2) *Den empiriske komponent*

Slutninger ud fra erfaring med tidligere patienter.

- **Ukontrolleret erfaring.**
- **Kontrolleret erfaring.**

#### Humanistisk tænkning:

##### 3) *Den empatisk-hermeneutiske komponent*

Slutninger ud fra indlevelsen i patientens situation og en forståelse af patienten som medmenneske.

##### 4) *Den etiske komponent*

Slutninger ud fra en etisk vurdering af de mulige handlinger og deres konsekvenser.

### Empati: Hermeneutisk:

Indføling

Empati forudsætter viden, som tillader én at forstå og fortolke det andet menneskes ytringer, hermeneutisk viden. *Den hermeneutiske viden* er således den baggrundviden, der skal til for at leve sig ind i et andet menneskes situation.

### Normer:

- 1) **Tekniske normer:** handler om effektivitet eller egnethed til et bestemt formål uden at sige noget om dette formåls "godhed" i etisk forstand.
- 2) **Juridiske normer:** handler om lands lov og ret, og ser man stort på disse normer, risikerer man at blive straffet.
- 3) **Etiske eller moralske normer:** handler om, hvad der *i sig selv* er godt eller acceptabelt, og bryder man disse normer, er straffen andres moralske indignation eller ens egen dårlige samvittighed.

### Emotivisme:

Det nytter ikke at diskutere etiske spørgsmål, da det jo til syvende og sidst er et spørgsmål om smag og behag.

### Konsistenskravet:

*Ens tilfælde skal behandles ens!*

Hvis man ønsker at handle forskelligt i to situationer, må man være i stand til at angive den etisk relevante forskel.

### Den etiske beslutningsproces:

- **Umiddelbar intuition:** menneskesyn, evnen til at føle medlidenhed, respekt, omsorg.

- **Det faktuelle grundlag:** både naturvidenskabelige fakta og lægens viden om patienten og dennes ønsker. Også fremtidige fakta bør inddrages. Vær opmærksom på *skråplansargumenter*, som går ud på, at selvom en handling kan betragtes som moralsk acceptabel, bør den alligevel ikke tillades, da det uvægeligt ville føre til andre handlinger, som ikke anses for moralsk acceptable.

- **Konsekvensbetragtninger og deontologiske betragtninger:**

Etiske konsekvensbetragtninger: sammenligne "godheden" af konsekvenserne af forskellige handlinger.

**Utiliteten** = "nyttens" af forskellige konsekvenser.

- **Individrettede:**
  - **patientrettede:** bedste konsekvens for pt.
  - **andre individrettede:** bedste konsekvens for lægen, familien, kolleger mm.
- **Generelle:** beslutningens konsekvens for sundhedsvæsenet og fremtidige patienter.

Deontologiske betragtninger: moralske værdi af forskellige typer af handlinger f.eks.: pligt til at værne liv, pligt til ikke at lyve, pligt til at respektere pt.'s autonomi.

Den deontologiske betragtning gælder især ved fravald af behandling, men er mindre

<b>Livskvalitet:</b>	<p>stærk, når det gælder tilvalg af behandling eller ønske om særbehandling. er til syvende og sidst noget subjektivt, men det har alligevel en objektiv, observerbar side, da mennesker har meget til fælles. Pseudomatematisk:</p> $\text{subjektiv livskvalitet} = \text{objektiv livskvalitet} / \text{forventning}$
<b>Autonom:</b>	<ul style="list-style-type: none"> <li>- <i>beslutning</i>: en beslutning, som er taget efter moden overvejelse uden ydre pression.</li> <li>- <i>person</i>: en person, der er i stand til at tage autonome beslutninger.</li> <li>- <i>ikke-autonom person</i>: en person, der på grund af umodenhed, sygdom eller handicap er ude af stand til at gøre det.</li> </ul>
<b>Paternalisme:</b>	<p>en person A handler paternalistisk over for B, hvis A vælger den beslutning, der efter hans mening gavner B bedst, uanset om B har givet sit samtykke eller ej.</p> <ul style="list-style-type: none"> <li>• <b>Genuin</b>: paternalistisk handlen over for ikke-autonome personer.</li> <li>• <b>Ikke-ønsket</b>: paternalistisk handlen overfor autonome personer, uden dennes fulde samtykke. Moralsk uacceptabel handlemåde.</li> <li>• <b>Ønsket</b>: paternalistisk handlen overfor autonome personer, som ikke ønsker at blive draget ind i beslutningsprocessen.</li> </ul>
<b>Selektivt tab af autonomi:</b>	<p>Personer, som er fuldt autonome på visse punkter, men på enkelte punkter mister sin autonome beslutningsdygtighed. F.eks.: stofmisbrugere.</p>
<b>Defensiv medicin:</b>	<p>Beslutninger, som har til formål at beskytte lægen mod sagsanlæg.</p>
<b>Alternativ nytte:</b>	<p>Da sundhedsvæsenet har et fastlagt loft over udgifterne, bliver den egentlige omkostning <i>den alternative nytte</i> (= <i>the opportunity cost</i>). Omkostningen drejer sig således ikke blot om penge, men om det man ellers kunne have fået for pengene.</p>
<b>RATIONEL KLINIK, KAPITEL 8</b>	
<b>Prospektive undersøgelser:</b>	<p>P(virkninglårsag).          Starter med årsagen og ræsonnerer fra årsagen fremad i tiden til en formodet virkning.</p> <p><u>Den prospektive kohorteundersøgelse:</u>          Udvalger en eller flere grupper personer, som opfylder visse kriterier og følger dem over tid. Dataindsamlingen starter ved undersøgelsens start.          De prospektive kohorteundersøgelser, især de randomiserede, er langt de pålideligste.</p>
<b>Retrospektive undersøgelser:</b>	<p>P(årsaglvirkning).          Starter med virkningen og ræsonnerer fra virkningen til formodet årsag bagud i tiden.</p> <p><u>Den retrospektive case-kontrolundersøgelse:</u>          Udvalger en gruppe personer, der frembyder den virkning, man er interesseret i, og registrerer hyppigheden af en formodet årsagsfaktor. Samtidig udvælges en gruppe kontroller og hyppigheden af den formodede årsag registreres hos disse. Typisk anvendes allerede indsamlede data.          Langt lettere at udføre. Kan bla. anvendes ved undersøgelser af årsagerne til sjældne bivirkninger. Man kan undersøge for flere årsagsfaktorer samtidig.</p>
<b>Problemer:</b>	<ul style="list-style-type: none"> <li>• Det er svært at sikre sig, at gruppen af cases og gruppen af kontroller er sammenlignelige.</li> <li>• De retrospektivt indsamlede oplysningers pålidelighed f.eks. <i>erindringsbias</i>.</li> <li>• Manglende randomisering eller blinding.</li> <li>• Case-kontrolundersøgelsen bør ikke anvendes til at vurdere den eventuelle positive effekt af vore behandlinger, idet den reelle effekt af de fleste af de behandlinger, vi bruger, er langt mindre end størrelsen af den bias, case-kontrolundersøgelser ofte er behæftet med.</li> </ul>

*Confounding*: en fælles årsag eller ledsagefaktor. Gælder både for prospektive og retrospektive undersøgelser.

**Tværsnitsundersøgelser:**

Undersøger den statistiske sammenhæng mellem de indsamlede data uden at kunne afgøre, hvad der forårsager hvad. Man har på et givet tidspunkt indsamlet forskellige oplysninger.

Problemer:

- Langt det svageste design, når det gælder udredningen af kausale sammenhænge.
- Man må ikke benytte *den longitudinelle tolkning af tværsnitsdata*, da det drejer sig om tværsnit af vidt forskellige populationer.
- Fejlslutninger ud fra falske associationer i hospitalsmaterialer (Berkson's fallacy), da man ikke kan afgøre den kausale sammenhæng.

**Historiske kohorteundersøgelser:**

En retrospektiv dataindsamling kombineret med kohortelogik, dvs. slutning fra årsag (operation) til virkning (operationsresultat). Tillader beregning af P(virkning|årsag), men har også nogle af den retrospektive undersøgelses svagheder.

		Virkning		
		+	-	
Årsag	+	a	b	a+b
	-	b	d	c+d
		a+d	b+c	

Statistisk behandling af den prospektive kohorteundersøgelse:

**Relativ risiko:** RR. Udregnes ved data på en binær nominalskala (helbredt/ikke-helbredt).

$$RR = (a/(a+b))/(c/(c+d))$$

Man læser tabellen vandrat, når det drejer sig om en sammenlignende kohorteundersøgelse.

Yderst utilfredsstillende, da man ikke får noget at vide om den absolutte risiko.

Statistisk behandling af den retrospektive case-kontrolundersøgelse:

**Odds ratio:** OR.  $OR = (a/c)/(b/d) = (ad)/(bc)$

Odds ratio er en god tilnærmelse til den relative risiko, hvis hyppigheden af virkningen er lille, dvs.  $a \ll a+b$  og  $c \ll c+d$ .

**Deskriptiv statistik:**

Den simpleste form for statistik. Den indeholder tabeller, som beskriver patientmaterialet (kønsfordeling, alder, sygdommens sværhedsgrad osv.) og de opnåede resultater.

**Kvartiler:**

Kan benyttes ved rangstillede observationer. Deler stikprøven i fire lige store dele. De tre kvartiler benævnes den **laveste kvartil**, **medianen** (eller midtvejsværdien) og den **højeste kvartil**. Området mellem den laveste og den højeste kvartil benævnes det **interkvartile område**.

**Fraktiler:**

Median, kvartiler, deciler og percentiler. Man definerer normalområdet som de værdier, der ligger mellem 2,5 og 97,5 percentilen.

**Stikprøvekaraktæristikker:**

Hyppighed, gennemsnit, standarddeviationen osv., da de repræsenterer numeriske størrelser, som karakteriserer stikprøven.

*Karakteristik:*

Hyppighed  
Median og fraktiler  
Gennemsnit  
Standarddeviation

*Måleniveau:*

Nominalskala  
Rangskala  
Intervalskala  
Intervalskal og gaussisk fordeling

<b>Parameter:</b>	<p>Det er altid muligt at reducere målniveauet.          En populationskarakteristik, som man aldrig kan bestemme med sikkerhed, men kun skønne over.</p>
<b>Standard error of the mean SE / SEM:</b>	<p>Standardfejlen på gennemsnittet. Beregnes:  <math display="block">SE = SD/\sqrt{N}</math>         Man antager at gennemsnit <math>\pm 2</math> SE angiver 95% sikkerhedsgrænserne for gennemsnittet. <math>2SD = 2SE * \sqrt{N}</math> kan nogle gange være brugbar.</p>
<b>Hypotetisk deduktiv forskning:</b>	<p>Standardfejlen kan med god tilnærmelse benyttes til at beregne sikkerhedsgrænserne for gennemsnittet, hvis den oprindelige fordeling ikke afviger alt for meget fra en gaussisk fordeling, og hvis antallet af patienter ikke er alt for lille. Den kan beregnes udfra SD selvom fordelingen er højreskæv og SD derfor ikke kan benyttes til at beskrive fordelingen med.</p> <ol style="list-style-type: none"> <li>1) Opstilling af en hypotese om et årsags-virkningsforhold.</li> <li>2) Udførelse af passende undersøgelse.</li> <li>3) Foreneligheden af resultaterne med hypotesen testes vha. en passende statistisk metode.</li> </ol>
<b>Pseudo-hypotetisk deduktiv logik:</b>	<p>Man foretager talrige observationer, og når man observerer noget usædvanligt, foretager man en signifikanstest og konkluderer, at de har opdaget en ny regelmæssighed i naturen. De sammenhænge, der påvises, skal ikke ignoreres, men de tjener kun til at generere hypoteser, som kan testes i nye undersøgelser med andre patienter.</p>
<b>Forkert stikprøveenhed:</b>	<p>En patient figurerer flere gange i undersøgelse, hvilket medfører at observationerne ikke er indbyrdes uafhængige. F.eks:</p> <ul style="list-style-type: none"> <li>- En patient som en gang har være helbredt bliver det med en øget sandsynlighed igen næste gang.</li> <li>- Den samme patient kan godt have flere bivirkninger =&gt; figurerer flere gange.</li> <li>- Befolkningsundersøgelser, hvor man regner med personår. 1 ambassadør i et år = 26 turister i 2 uger.</li> </ul>
<b>Beregning af patientantal:</b>	<ol style="list-style-type: none"> <li>1) Hvor stor må en type I fejlsrisiko være? Konventionelt 5%.</li> <li>2) Hvor stor en forskel mellem helbredelseshyppighederne er forskeren villig til at overse? Hvis det drejer sig om overlevelse eller død, er små terapeutiske gevinster af stor betydning, og den slags forsøg kræver derfor meget store patientantal.</li> <li>3) Hvor stor må type II fejlsrisikoen være? Bør normalt være som hos type I. Man risikerer at bortkaste signifikante undersøgelser og at godkende insignifikante.</li> </ol>
<b>P-værdier:</b>	<p><u>Definition:</u> sandsynligheden for ved tilfældighedensnes spil at opnå en forskel som den observerede eller en endnu større forskel, og hermed forstås en forskel enten i den ene behandlings favør eller i den andens favør.          Man skal medindrage sin <i>a priori</i> tiltro til, at nulhypotesen er sand.</p>
<b>Statistiske test:</b>	<ul style="list-style-type: none"> <li>• <math>\chi^2</math> anvendes til sammenligning af de to hyppigheder, dvs. observationer på en <i>binær skala</i>.</li> <li>• <b>Rangsumtest</b> for uparrede data anvendes til sammenligning af to fordelinger målt på en <i>rangskala</i>.</li> <li>• <b>t-test</b> for uparrede data anvendes til sammenligning af to fordelinger målt på en <i>intervalskala</i>, forudsat at begge fordelinger er tilnærmet <i>gaussiske med sammen standarddeviation</i>.          Hvis patientantallet er begrænset og fordelingen meget skæv, skal man rangstille data og anvende en rangsumtest.          Den tester forskellen mellem gennemsnitsværdierne i de to grupper, og gennemsnittet som statistisk mål er meningsløst, hvis skalaintervallerne ikke er lige store.</li> </ul>

**Korrelations-  
analyse:**

- **Parrede problemstillinger** f.eks. overkrydsningsforsøg, hvor man har to observationer på samme patient benyttets fortegnstest, parret rangsumtest og parret t-test, afhængigt af måleskalaen.

**Totalvariationen:** variationen omkring gennemsnittet.

Bestemmes i princippet ved at måle de lodrette afstande mellem de enkelte punkter og gennemsnittet, kvadrere disse afstande og beregne summen.

**Residualvariationen:** variationen omkring regressionslinien.

Bestemmes i princippet ved at måle de lodrette afstande mellem punkterne og regressionslinien, hvorefter man kvadrerer og summerer dem.

**Forklaret variation = totalvariation - residualvariation**

Perfekt korrelation: residualvariationen være 0, hvilket vil sige, at hele totalvariationen er forklaret.

Ingen korrelation: en horisontal regressionslinien på højde med gennemsnittet ville medføre at variationen omkring denne bliver som variationen omkring gennemsnittet. Residualvariationen = totalvariationen, og intet ville være forklaret.

**Korrelationskoefficienten  $r$ :** benyttes som mål for korrelationens styrke.

**Determinationskoefficienten  $r^2$ :** benyttes som mål for korrelationens styrke.

$$r^2 = \text{forklaret variation} / \text{totalvariation}$$

**Korrelationskoefficienten** varierer fra -1 (perfekt negativ korrelation) over 0 (ingen korrelation) til +1 (perfekt positiv korrelation). Den er en numerisk karakteristik af den pågældende stikprøve, og tilfældighedernes spil kan medføre, at den er større eller mindre end populationens "sande".

Den konventionelle korrelationsanalyse bygger på den forudsætning, at der er tale om målinger på en intervallskal, og at punkterne spreder sig gaussisk omkring regressionslinien.

**Problemer:**

- **Mus-elefantfænomenet:** selvom at der ikke er nogen positiv korrelation mellem alder og længde af voksne mus eller af voksne elefanter, ville man hvis man undersøgte 50 mus og 50 elefanter med en korrelationsanalyse få en uhyre lille p-værdi.