

# Grunnlaget for evidensbasert medisin

Ian Mason, PHD, Occupational Health og medisinsk  
forfatter

I denne korte artikkelen skal vi se  
på følgende:

- Placebo vs. aktiv behandling
- Dobbeltblinde vs. enkeltblinde studier
- Randomisert vs. ikke-randomisert
- Multisenter- vs. enkeltcenterstudier
- Stor vs. liten studie
- Beregning av effekt og prøvestørrelse
- Prospektiv vs. retrospektiv studie
- Primært vs. sekundært endepunkt
- Metaanalyse
- Statistisk signifikans og p-verdi

## Ian Mason ser på historien til kliniske studier samt noen viktige aspekter som brukes i studier i dag.

Kan du huske første gang du ble lagt inn på sykehus? For meg var det på slutten av 1950-tallet da jeg måtte fjerne mandlene. Jeg var seks år gammel og kjemperedd. Det var imidlertid en liten trøst i at vi var en stor gjeng med barn på samme avdeling. På 50-tallet var det veldig populært å fjerne mandlene – et universalmiddel ved problemer i øvre luftveier. Ett av fire barn i Storbritannia fjernet mandlene, dvs. rundt 250 000 inngrep hvert år<sup>1</sup>. I tiårene som fulgte fikk man stadig flere bevis for at fjerning av mandlene hadde liten klinisk verdi<sup>2</sup>. Og derfor har antall barn som fjerner mandlene i Storbritannia i dag blitt redusert til rundt en tiendedel av det på som var tilfelle på 1950-tallet<sup>3</sup>. Pasientene velges nøye ut og behandles med den teknikken som har vist seg å være best egnet ut fra en rekke kliniske studier<sup>4</sup>.

Historien om storstilt fjerning av mandler er et godt eksempel på hvordan populære medisinske inngrep ble ansett å være arkaiske, ubrukelige eller direkte farlige. «Behandlinger» som årelating, å gi gravide klyster under fødsel eller foreskrive sengeleie for å fremskynde bedring etter et hjerteinfarkt, er klassiske eksempler på foreldet medisinsk praksis som var basert på kliniske anekdoter, misforståtte ekspertuttalelser eller utdatert undervisning på medisinske læresteder.

## En ny æra – evidensbasert behandling

I dag er vi opptatt av evidensbasert behandling. Kliniske studier med god design brukes for å gi objektiv informasjon om sikkerhet og effekt av helsebringende tiltak – enten det handler om legemidler, proteseimplantater eller andre medisinske innovasjoner. Nøye gjennomførte kliniske studier tilbyr ikke bare den raskeste og sikreste måten å fastslå om en behandling fungerer, men de er absolutt nødvendige for godkjenning av regelverk og for godkjenning av klinikere.



Evidensbasert medisin tok virkelig av tidlig på 1990-tallet, men startskuddet gikk mye tidligere enn det. En av de tidligste, kontrollerte kliniske studiene fant sted i 1747, da dr. James Lind bestemte seg for å finne ut om appelsiner og sitroner kunne hindre den invalidiserende (og potensielt dødelige) sykdommen skjørbuk blant sjøfolk på lange seilas. De grunnleggende prinsippene i hans eksperimentelle modell står fortsatt i dag. Under kontrollerte forhold sammenlignet han flere ernæringsmessige intervensjoner hos syke pasienter som var godt tilpasset alvorlighetsgraden av den lidelsen de led av. «De mest synlige og gode effektene skyldtes bruk av appelsiner og sitroner», skrev han.<sup>5</sup> Det tok et halvt århundre før funnene hans ble allemannseie, og den britiske marinen gjorde limejuice obligatorisk for sjømenn (og derfor kalles britene også fote «limeys»). Denne enkle innovasjonen styrket uten tvil den britiske marinen.

Senere ble den moderne kliniske studien forbedret skritt for skritt. I midten av det nittende århundre tok man i bruk en **placebo-** eller «lurebehandling», som **aktiv behandling** kunne sammenlignes med. Ytterligere to viktige utviklingstrekk fant sted på 1940-tallet. Den første var en tidlig dobbeltblindet, kontrollert studie (for å undersøke bruken av patulin mot vanlig forkjølelse)<sup>6</sup>.

Den viktige forskjellen mellom en **dobbeltblind og enkeltblind studie** er at deltakerne i en enkeltblind studie ikke vet hvilken behandling de får før studien er avsluttet. I en dobbeltblind studie vet verken legene eller deltakerne hvem som har motatt hvilken behandling før studien er avsluttet. Patulinstudien viste til slutt at pantulin ikke hadde noen positiv effekt på den aktive behandlingen, men det var likevel et nyttig funn. Det neste steget fant sted i 1946 da man for første gang brukte randomisering i en stor klinisk studie (for å evaluere bruken av streptomycin hos tuberkulosepasienter<sup>7</sup>).

For å minimere risikoen for skjevhet i en komparativ studie, kan pasienter som samtykker i å delta bli tilfeldig allokert til ulike behandlingsgrupper. I en studie som sammenligner to behandlinger, kan man sammenligne tildelingsprosessen med å trekke lodd for å avgjøre hvilken behandling en pasient skal få. Randomisering balanserer også eventuelle andre faktorer som kan påvirke behandlingseffekten, som alder, kjønn eller vekt. En **randomisert studie** anses alltid å være mer robust enn en **ikke-randomisert studie**.

## Andre egenskaper ved moderne kliniske studier

Moderne randomiserte, kontrollerte kliniske studier kan ha flere andre viktige funksjoner. Mange studier er i dag multisenterstudier (de utføres på mer enn ett sted), mens en **enkeltcenterstudie** gjennomføres på ett sted. Selv om en **multisenterstudie** kan høres mer «troverdig» ut, er ikke dette nødvendigvis tilfelle i en enkeltcenterstudie er det enklere å konsekvent kontrollere variabler som kan påvirke utfallet. I tillegg blir det noen ganger utført multisenterstudier for å rekruttere en mer mangfoldig pasientpopulasjon, eller fordi det er den eneste mulige metoden for å rekruttere et tilstrekkelig stort antall personer for å evaluere en liten behandlingseffekt på en pålitelig måte. En **stor studie** er ikke nødvendigvis mer troverdig enn en **mindre studie**, og en god statistiker vil på forhånd ha spesifisert antall deltakere som trengs for å gi et pålitelig resultat i begge tilfeller ved såkalt beregning av effekt- og prøvestørrelse.

En **prospektiv studie** er utformet på forhånd for å svare på ett eller flere forskningsspørsmål, mens en **retrospektiv studie** gjennomgår eksisterende informasjon, for eksempel pasientregisterdata eller elektroniske helsejournaler. Derfor følger en prospektiv studie pasientene inn i fremtiden, mens en retrospektiv studie ser på data fra fortiden. Selvfølgelig kan en studie også være begge deler hvis en forsker gjennomgår data fra en retrospektiv studiegruppe og deretter følger denne pasientgruppen inn i fremtiden.

Jeg har allerede nevnt bruk av placebo som et kontrollparameter for den behandlingen som studeres. Andre parametre omfatter den beste eksisterende behandlingen (også kalt «gullstandard»). I studier av medisinsk utstyr kalles placebo ofte lurenheter. God kontroll gir en studie troverdighet.

Det er ikke alltid mulig å designe en blindstudie, for eksempel hvis medisinsk utstyr eller sårbandasjer skal sammenlignes, og både pasienten og studielegen tydelig kan se dem. I slike tilfeller kan man velge en såkalt overkrysningsstudie. I denne typen studie får forsøkspersonen begge behandlinger i tilfeldig rekkefølge.

## Endepunkter

I noen studier omtales såkalte **primære og sekundære endepunkter**<sup>8</sup>. Det primære endepunktet i en klinisk studie er det primære målet for studien der man ser om behandlingen virker. Sekundære endepunkter er ytterligere endepunkter, fortrinnsvis også forhåndsspesifiserte, som ikke var utgangspunkt for studien, men som kan gi nyttig informasjon eller hint om veier til videre forskning.

En annen type studie som har blitt mer utbredt de siste årene er **metaanalyse**<sup>9</sup>. Dette innebærer bruk av statistiske metoder for å kombinere resultater fra publiserte studier som adresserer det samme spørsmålet. Resultatene samles i en systematisk gjennomgang som kombinerer resultatene for å forbedre presisjonen i estimatene for behandlingseffekt og vurdere om behandlingseffektene er like i lignende situasjoner. Cochrane-studier er eksempler på systematiske gjennomganger som generelt sett anses som den høyeste standarden innen evidensbasert behandling. Selvfølgelig er solid statistikk og robuste prøveresultater bare en del av historien. Man har i økende grad lagt vekt på behovet for

å kombinere kritisk vurdering av evidens med pasientens verdier og preferanser gjennom delt beslutningstaking<sup>10</sup>.

## Statistisk signifikans og klinisk relevans

Tolkningen av «statistisk signifikans» blir ofte feiltolket som «klinisk viktig». **Statistisk signifikans** kvantifiserer sannsynligheten for at en studies resultater skyldes tilfeldighet. Klinisk signifikans eller relevans viser derimot til den faktiske behandlingseffekten. I en vellykket studie ønsker du å se en forskjell mellom intervensjons- og kontrollgruppen som er klinisk relevant, noe som sannsynligvis vil påvirke medisinsk praksis og statistisk signifikans, slik at resultatet ikke bare skyldes tilfeldighet. For å måle statistisk signifikans brukes en p-verdi. **p-verdien** gir en indikasjon på om resultatet skyldes en tilfeldighet eller en reell behandlingseffekt. Vanligvis sies det at hvis p-verdien er mindre enn 0,05, er statistisk signifikans oppnådd (se Tabell 1) og at muligheten for at resultatet skyldes en tilfeldighet er mindre enn 5 %<sup>11</sup>.

Til slutt gjennomføres det tester for kommersielle formål. Hvis et selskap ønsker å fremsette påstander som sammenligner ett av produktene deres med et annet produkt, er det viktig å basere påstandene på resultatene fra en studie som direkte sammenligner disse produktene. Det anses generelt uakseptabelt å fremsette slike påstander med mindre det foreligger slike bevis fra en sammenlignende studie.

Denne korte artikkelen kan ikke gå i dybden når det gjelder kliniske tester, og det er skrevet hele lærebøker om emnet. Hvis du er interessert i å lære mer, finnes det et fascinerende nettsted som omhandler arbeidet til dr. James Lind (forfatteren av The Scurvy Study), opprettet av Royal College of Physicians of Edinburgh. Du finner det på [www.jameslindlibrary.org](http://www.jameslindlibrary.org). I en serie med gode og fritt tilgjengelige artikler illustrerer nettstedet utviklingen av metoder for å vurdere effekter av behandling – fra å fastsette prinsippene for testing av behandlinger for tusen år siden til systematiske analyser av erfaringene til titusenvis av pasienter i dag.

### Referanser:

1. Illingworth RS Proceedings of the Royal Society of Medicine Vol 54 May 1961:393-399
2. Dwyer-Hemmings L. 'A Wicked Operation'? Tonsillectomy in Twentieth-Century Britain. Med Hist. 2018 Apr; 62(2): 217-241
3. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/coa.13707>
4. Akural El et al Post-tonsillectomy pain: a prospective, randomised and double-blinded study to compare an ultrasonically activated scalpel technique with the blunt dissection technique. Anaesthesia 2001 Nov;56(11):1045-50.
5. Lind J. 'A Treatise of the Scurvy.' Edinburgh 1753
6. Bhatt A. Evolution of Clinical Research. Perspect Clin Res. 2010 Jan-Mar; 1(1): 6-10
7. MRC Streptomycin in TB Trials Committee. BMJ 1948;2:769-83
8. Vetter TR. Defining the Primary Outcomes and Justifying Secondary Outcomes of a Study: Usually, the Fewer, the Better. Anesth Analg. 2017 Aug;125(2):678-681.
9. Haidich AB. Meta-analysis in medical research. Hippokratia. 2010 Dec; 14(Suppl 1): 29-37
10. Djulbegovic B. Progress in evidence-based medicine: a quarter century on. Lancet. 2017 Jul 22;390(10092):415-423.
11. Common pitfalls in statistical analysis: Clinical versus statistical significance. Priya Ranganathan, C. S. Pramesh, and Marc Buyse. Perspect Clin Res. 2015 Jul-Sep; 6(3): 169-170

Les mer på [www.molnlycke.no](http://www.molnlycke.no)

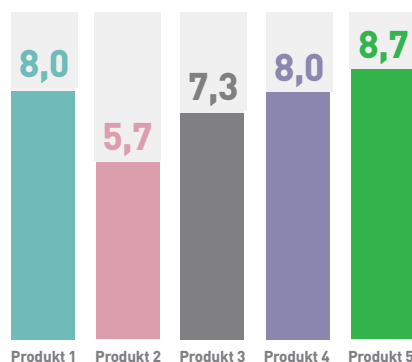
Molnlycke Health Care AS, Postboks 6229 Etterstad, Brynslåleén 4, 0667 Oslo, Norge. Tlf. 22 70 63 70. Varemerker, navn og logo knyttet Molnlycke er registrert globalt av ett eller flere av selskapene i konsernet Molnlycke Health Care. ©2022 Molnlycke Health Care AB. Med enerett. NOSU1272203

Tabell 1. P-verdi og de relaterte signifikansnivåene

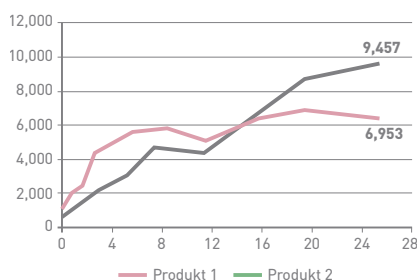
Signifikansnivå	Spesifikasjon
$p > 0,05$	ikke signifikant
$p \leq 0,05$ (5 %)	signifikant
$p \leq 0,01$ (1 %)	svært signifikant
$p \leq 0,001$ (0,1 %)	ekstremt signifikant

Det finnes mange måter å vise resultatene fra en sammenlignende studie på. En graf bør inneholde nok informasjon til at den ikke trenger ytterligere tekst for å bli forstått, og du bør velge den grafen som er best egnet til den sammenligningen du har gjort.

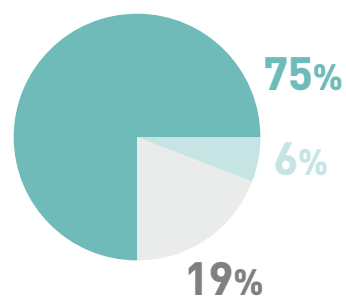
Hvis relevant, kan du også legge til antall pasienter som er inkludert som «n=», samt p-verdien for hver sammenligning du gjør.



Stolpediagram



Linjediagram



Kagediagram