

# Grundlaget for evidensbaseret medicin

Ian Mason, ph.d., forfatter til emner  
om arbejdsmiljø og arbejdsmedicin

## I denne korte artikel fremhæver vi følgende:

- Placebo vs. aktiv behandling
- Dobbeltblinde vs. enkeltblinde studie
- Randomiseret vs. ikke-randomiseret
- Multicenter vs. enkeltcenter
- Stort vs. lille studie
- Beregning af styrke og prøvestørrelse
- Prospektivt vs. retrospektivt studie
- Primært kontra sekundært endepunkt
- Metaanalyse
- Statistisk signifikans og p-værdien

## Ian Mason ser på historien om kliniske studier og nogle af de vigtigste elementer, der bruges i aktuelle studier.

Kan du huske din første hospitalsindlæggelse? Min var i slutningen af 1950'erne, hvor jeg skulle have fjernet mandler. Jeg var seks år gammel og skrækslagen. Men det var en trøst at være blandt en stor gruppe børn på samme afdeling. Tonsillektomi (fjernelse af mandler) var nemlig meget moderne i 50'erne – et kirurgisk universalmiddel til behandling af problemer i de øvre luftveje. Et ud af fire børn i Storbritannien fik fjernet mandler – en kvart million operationer om året<sup>1</sup>. I de følgende årtier viste akkumulerende evidens (dokumentation), at det at fjerne mandlerne var af tvivlsom klinisk værdi<sup>2</sup>. Derfor er antallet af børn, der får fjernet mandler i Storbritannien i dag omkring en tiendedel af antallet i 1950'erne<sup>3</sup>. Operationerne er nu nøje udvalgt efter retningslinjer fra veludførte kliniske studier<sup>4</sup>.

Historien om massefjernelsen af mandler er et godt eksempel på, hvordan engang så populære medicinske interventioner (indgreb) forekommer forældede, ubrugelige eller ligefrem farlige. Indgreb som f.eks. åreladning eller at ordinere sengeleje for at fremskynde restitution efter et hjerteanfald – klassiske eksempler på antikveret medicinsk praksis, der var baseret på kliniske anekdoter, vildledende ekspertudtalelser eller sørgeligt forældet undervisning på lægeuddannelsen.

## Evidensens tidsalder

I dag er det unægtelig blevet den evidensbaserede medicins tidsalder. Veldesignede kliniske studier bruges til at give objektiv information om sikkerheden og effektiviteten af sundhedsmæssige interventioner – uanset om det drejer sig om lægemidler, implantater eller andre medicinske nyskabelser. Omhyggeligt udførte kliniske studier tilbyder ikke kun den hurtigste og sikreste måde at afgøre på, om en intervention virker, men de er absolut nødvendige for myndighedernes godkendelse, og for at klinikere accepterer dem.

Evidensbaseret medicin tog fart i starten af 1990'erne,



men dens rødder ligger meget tidligere. Et af de tidligste kontrollerede kliniske studier fandt formentlig sted i 1747, da dr. James Lind besluttede at finde ud af, om appelsiner og citroner kunne forhindre den invaliderende (og potentielt dødelige) sygdom skørbug blandt sømænd på lange sørejser. Grundprincipperne i hans eksperimentelle model holder i dag. Under kontrollerede forhold sammenlignede han flere ernæringsmæssige interventioner hos syge patienter, der var godt matchet med hensyn til sværhedsgraden af deres skørbug. "De mest synlige og gode effekter blev bemærket ved brugen af appelsiner og citroner", skrev han<sup>5</sup>. Det tog et halvt århundrede for hans resultater at blive ført ud i livet og for den britiske flåde at gøre limejuice obligatorisk for sømænd i flåden (hvilket er grunden til, at briterne er kendt under betegnelsen "Limeys"). Denne enkle fornyelse styrkede den britiske flåde helt enormt.

Yderligere byggesten i det moderne kliniske studie blev langsomt tilføjet. I midten af det nittende århundrede tilføjede man brugen af **placebo** eller "snydemedicin", som en **aktiv behandling** kunne sammenlignes med. I 1940'erne skete der yderligere to vigtige udviklinger. Den første var et tidligt, dobbeltblindet, kontrolleret studie (for at undersøge brugen af patulin mod almindelig forkølelse)<sup>6</sup>.

Den vigtige forskel mellem et dobbeltblindet og et enkeltblindet studie er, at forsøgspersonerne i et enkeltblindet studie ikke ved, hvilken behandling de får, før studiet er afsluttet. I et dobbeltblindet studie ved hverken investigatorene eller forsøgspersonerne, hvem der har fået hvilken behandling før ved afslutningen af studiet. I øvrigt viste patulinstudiet ingen fordel for den aktive behandling – det var alligevel et nyttigt fund. Den anden udvikling fandt sted i 1946 og var den første anvendelse af randomisering i et større klinisk studie (til evaluering af streptomycin hos tuberkulosepatienter<sup>7</sup>).

For at minimere risikoen for bias (skævhed) i et komparativt studie kan patienter, der indvilliger i at deltage, randomiseres til forskellige behandlingsgrupper. I forbindelse med et studie, der sammenligner to interventioner, svarer tildelingsprocessen til at slå plat eller krone for at afgøre, hvilken intervention en patient får. Randomiseringen afbalancerer også eventuelle andre faktorer, der kan påvirke behandlingseffekten, f.eks. alder, køn eller vægt. Et **randomiseret studie** anses altid for at være mere robust end et **ikke-randomiseret studie**.

## Andre karakteristika ved moderne kliniske studier

Moderne randomiserede, kontrollerede kliniske studier kan have flere andre vigtige karakteristika. Mange studier i dag er multicenterstudier (der udføres på mere end ét forsøgscenter), hvorimod et **enkeltcenterstudie** udføres på ét forsøgscenter. Selvom et **multicenterstudie** måske lyder mere " troværdigt", er dette ikke nødvendigvis tilfældet. I et enkeltcenterstudie er det nemmere konsekvent at kontrollere variabler, der kan påvirke udfaldet. Derudover udføres der sommetider multicenterstudier for at rekruttere en mere forskelligartet patientpopulation, eller fordi det er den eneste realistiske metode til at rekruttere et tilstrækkeligt antal forsøgspersoner til pålidelig vurdering af en lille behandlingseffekt. Et **stort studie** er ikke nødvendigvis mere troværdigt end et **mindre**. En god statistiker vil på forhånd have specificeret det antal patienter, der er nødvendigt for at give et pålideligt resultat, i begge tilfælde ved hjælp af såkaldt styrke- og studiestørrelsesberegning.

Et **prospektivt studie** er på forhånd designet til at besvare et eller flere forskningsspørgsmål, mens et **retrospektivt studie** gennemgår eksisterende oplysninger, såsom patientregisterdata eller elektroniske patientjournaler. Et prospektivt studie følger derfor patienterne ind i fremtiden, og et retrospektivt kigger på data fra deres fortid. Et studie kan selvfølgelig være begge dele – hvis en forsker gennemgår data fra en retrospektiv studiegruppe og derefter følger denne patientkohorte ind i fremtiden.

Jeg har allerede nævnt brugen af placebo som komparator til kontrol af den intervention, der undersøges. Andre komparatorer omfatter den bedste eksisterende behandling (eller "guldstandard"). I tilfælde af studier med medicinsk udstyr kaldes placeboudstyr ofte for placebo. En passende kontrol tilføjer troværdighed til et studie.

Det er ikke altid muligt at designe et blindet studie, f.eks. hvis medicinsk udstyr eller sårbandager sammenlignes, og forsøgspersonen og den, der foretager eksperimentet, tydeligt kan se disse. I sådanne tilfælde kan der anvendes et overkrydsningsstudie. I denne type studie får forsøgspersonen begge behandlinger i tilfældig rækkefølge.

## Endepunkter

Nogle studier er beskrevet med **primære og sekundære endepunkter**<sup>8</sup>. Det primære endepunkt for et klinisk studie er det endepunkt, for hvilket studiet har statistisk styrke. Sekundære endepunkter er ekstra endepunkter, helst også forudspecificerede, for hvilke studiet måske ikke har nogen styrke, men som kan give nyttige oplysninger eller spor om veje til yderligere forskning.

En anden type studie, der er blevet mere udbredt i de seneste år, er **metaanalysen**<sup>9</sup>. Der er tale om anvendelse af statistiske metoder til at kombinere resultater fra offentliggjorte studier, der behandler samme spørgsmål. Resultaterne samles til en systematisk gennemgang, der kombinerer resultaterne for at forbedre nøjagtigheden af estimater for behandlingseffekt og vurdere, om behandlingseffekterne er de samme i ens situationer. Cochrane Reviews er eksempler på systematiske gennemgange, der generelt er anerkendt som den højeste standard inden

for evidensbaseret sundhedspleje. Naturligvis er solid statistik og robuste studieresultater kun en del af historien. Betydningen af behovet for at kombinere kritisk vurdering af evidensen med patientens værdier og præferencer er i stigende grad blevet anerkendt gennem fælles beslutningstagning<sup>10</sup>.

## Statistisk signifikans og klinisk relevans

Fortolkningen af 'statistisk signifikans' fejlfortolkes ofte som 'klinisk vigtig'. **Statistisk signifikans** kvantificerer (sætter tal på) sandsynligheden for, at et studies resultater skyldes tilfældigheder. Klinisk signifikans eller relevans henviser derimod til størrelsen af den faktiske behandlingseffekt. I et vellykket studie vil man gerne se en forskel mellem interventions- og kontrolgruppen, som er klinisk relevant, noget, der sandsynligvis vil påvirke medicinsk praksis, og statistisk signifikant, så resultatet ikke kun skyldes tilfældigheder. Der anvendes en p-værdi til at måle den statistiske signifikans. **P-værdien** giver en indikation af, om resultatet skyldes tilfældighed eller en reel behandlingseffekt. Almindeligvis bruges det til at sige, at hvis p-værdien er mindre end 0,05, er der opnået statistisk signifikans (se tabel 1), og at resultatet skyldes en tilfældighed er mindre end 5 %<sup>11</sup>.

Endelig udføres visse studier med henblik på kommercielle formål. Hvis en virksomhed har til hensigt at fremsætte påstande, der sammenligner et af dens produkter med et andet produkt, er det vigtigt at basere påstandene på resultaterne af en undersøgelse, der direkte sammenligner disse produkter. Det anses generelt for uacceptabelt at fremsætte sådanne påstande, medmindre der foreligger en sådan dokumentation (evidens) fra et sammenlignende studie.

Denne korte artikel kan kun strejfe overfladen på det kliniske studiedesign – der er skrevet hele lærebøger om emnet. Hvis din interesse er blevet vakt, har Royal College of Physicians of Edinburgh oprettet et fascinerende websted, der hylder dr. James Lind (forfatteren af *The Scurvy Study*). Den kan findes på [www.jameslindlibrary.org](http://www.jameslindlibrary.org). I en række fremragende og frit tilgængelige artikler illustrerer webstedet udviklingen af metoder til at vurdere effekterne af behandlinger – fra at opstille principperne for test af behandlinger for tusind år siden til systematiske analyser af erfaringerne fra titusindvis af patienter i dag.

### Referencer:

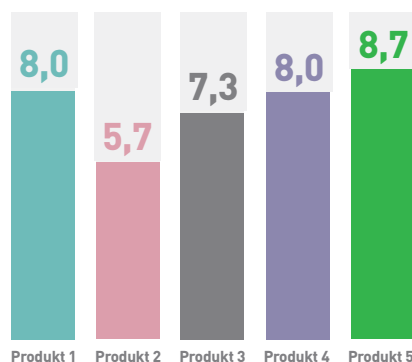
1. Illingworth RS Proceedings of the Royal Society of Medicine Vol 54 May 1961:393-399
2. Dwyer-Hemmings L. 'A Wicked Operation'? Tonsillectomy in Twentieth-Century Britain. *Med Hist.* 2018 Apr; 62(2): 217-241.
3. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/coa.13707>
4. Akural El et al Post-tonsillectomy 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.

Tabel 1. P-værdi og de relaterede signifikansniveauer

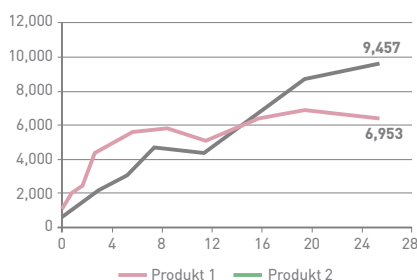
Signifikansniveau	Specifikation
$p > 0,05$	ikke signifikant
$p \leq 0,05$ (5 %)	signifikant
$p \leq 0,01$ (1 %)	meget stor
$p \leq 0,001$ (0,1 %)	meget signifikant

Der er mange måder at vise resultaterne fra et komparativt studie på. Et diagram bør indeholde tilstrækkeligt med oplysninger til, at det ikke behøver yderligere tekst for at blive forstået, og du bør vælge det diagram, der er bedst egnet til den sammenligning, du har foretaget.

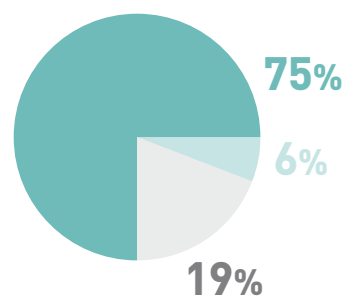
Hvis det er relevant, skal du også tilføje antallet af forsøgspersoner, der er inkluderet som "n=", samt p-værdien for hver sammenligning, du foretager.



Søjlediagram



Linjediagram



Cirkeldiagram

Læs mere på [www.molnlycke.dk](http://www.molnlycke.dk)