
Metaanalyser – statistisk alkymi for det 21. århundre?

REDAKSJONELT

IVAR SØNBØ KRISTIANSEN

Ivar Sønbo Kristiansen (f. 1947) er samfunnsmedisiner. Han har vært lektor ved den helseøkonomiske enheten ved Syddansk Universitet, Odense fra 1997.

Institut for Sundhedstjenesteforskning

– Sundhedsøkonomi

Syddansk Universitet, Odense

Winsløwparken 19,

DK-5000 Odense C

Medisinen har som fag vært handlingssøkende og ikke sannhetssøkende. Dette er i ferd med å endre seg, og i takt med dette hardner kampen om hva som er god medisin. Mens respekterte kollegers meninger før ble betraktet som sannhet, stiller vi nå med god grunn krav om vitenskapelig dokumentasjon. Den randomiserte, kliniske studie har lenge vært ”gullstandarden”, og de siste ti år er denne blitt supplert med metaanalyser.

Meta er gresk og betyr ”etter”. Metaanalysen bygger på den systematiske litteraturoversikten, men går videre ved at den lager en tallmessig syntese av alle studier som oppfyller spesifiserte kvalitetskrav. Fra den enkelte undersøkelse innhentes nøkkeltall om behandlingseffekt som gjerne uttrykkes som oddsratio eller relativ risiko. Usikkerheten i effektanslagene anskueliggjøres ved konfidensintervall. Jo flere pasienter i metaanalysen, jo smalere blir konfidensintervallet. Sett under ett kan mange, små enkeltstudier indikere en behandlingseffekt selv om studiene hver for seg ikke er ”signifikante”.

Antall publiserte metaanalyser øker raskt, og de er lett tilgjengelige på Internett bl.a. via Medline og Cochrane-databasen. Metaanalyser av randomiserte undersøkelser settes av mange øverst i hierarkiet over pålitelighet. Disse analysene har lagt grunnlaget for viktige behandlingstiltak som fibrinolytisk behandling ved hjerteinfarkt og steroidbehandling av premature. Imidlertid finnes det også kritiske røster. Feinstein sammenlikner metaanalyse med

alkymi – middelalderens mislykkede bestrebelser på å skape gull av uedle metaller (1). I det følgende utdypes noe av kritikken som har vært fremført mot metaanalysen som metode.

Det vitenskapelige prinsipp å sammenlikne likt med likt kan bli satt på harde prøver i metaanalyser. Dersom pasientgrupper eller behandling i de kliniske undersøkelser ikke er helt like, vil metaanalysen bygge på en syntese av ”epler og pærer”. Metaanalytikerene søker derfor å avsløre inhomogenitet i pasientmaterialet med statistiske tester, men disse er lite følsomme.

I tallsyntesen antar man typisk at effekten av en behandling er prosentvis den samme i ulike pasientgrupper selv om de har ulik utgangsrisiko for sykdom, men denne antakelsen holder ikke alltid (2). Et annet problem er at vi oftest vet lite om den *eksterne* validitet ved randomiserte undersøkelser. Man har ingen garanti for at en undersøkelse fra USA sier noe om behandlingseffekter i Norge eller Kina. Uten en høy grad av ekstern validitet for de studier som inngår i metaanalysen, gir denne liten mening fordi den nettopp søker å si noe generelt om behandlingseffekter.

En del randomiserte studier forblir upubliserte fordi de er små eller ikke gav de ønskede resultater. Andre publiseres på et språk metaanalytikerene ikke behersker. Alt dette kan gi metaanalysen skjevheter fordi man ikke får inkludert alle relevante studier. Metaanalytikerene stiller minstekrav om kvalitet for å inkludere studier, men vurderingene vil nødvendigvis bli preget av skjønn og valg av kvalitetskriterier (3). Dersom en metaanalyse domineres av en enkelt, stor studie av diskutabel kvalitet, vil metaanalysens konklusjon også bli diskutabel. Valg av endepunkter kan avgjøre om en behandling fremstår som effektiv eller ikke. To ulike metaanalyser av antibiotikabehandling ved luftveisinfeksjoner, basert på omtrent de samme randomiserte studier, kom til ulike resultater fordi de hadde valgt ulike kliniske effektmål (4). Tilfellet er ikke enestående. En metaanalyse fra 1996 konkluderte med at kalsiumtilskudd er gunstig ved preeklampsi, mens en senere analyse, basert på mange av de samme studier, bare tydet på effekt i høyrisikogrupper (5). Forklaringen var bl.a. ulik effekt i ulike pasientgrupper, altså et heterogenitetsproblem.

Når metaanalyser som er bygd på mindre kliniske studier, følges opp med store, randomiserte undersøkelser, er konklusjonene ikke alltid de samme som i metaanalysen. Et klassisk eksempel er bruk av magnesium ved akutt hjerteinfarkt. En randomisert studie på 58 000 pasienter kunne ikke bekrefte den gunstige effekten man hadde funnet i en metaanalyse, formodentlig fordi den sistnevnte var utsatt for publikasjonsskjevhet (6). LeLorier og medarbeidere sammenliknet 12 store randomiserte studier med 19 forutgående metaanalyser og fant at de sistnevnte bare gav korrekt ”fasit” i to av tre tilfeller (7). Det er derfor et tankekors at det oftest bare er sterke, kommersielle aktører som kan finansiere de store randomiserte undersøkelser.

Metaanalyser kan gi feilaktige resultater når de bygger på metodemessige forutsetninger som ikke er oppfylt. Man gjør derfor klokt i å vurdere metaanalyser med samme kritiske blikk som annen dokumentasjon. Bailar går så langt som å hevde at den som *gjennomfører* metaanalyser bør ha like godt kjennskap til forskningsfeltet som den som gjør originalstudier (8). Han mener

også de som *bruker* metaanalyser bør kjenne både grunnlagsstudiene og den kliniske problemstillingen. Han advarer altså mot å bruke metaanalysen som en snarvei til innsikt i kompliserte medisinske spørsmål.

Til tross for kvalitetskontroll kan metaanalyser inneholde klare feil (9). Forhåpentlig vil internasjonale retningslinjer bedre kvaliteten (10), men de kan ikke løse problemene knyttet til den store arbeidsinnsats som ligger bak gode metaanalyser. For det første tenderer de til å være foreldet når blir publisert. For det andre kan den store primære arbeidsinnsats "skremme" metaanalytikerne fra å gjøre nødvendige oppdateringer.

En systematisk oversikt over kliniske studier er åpenbart bedre enn en usystematisk. Når den systematiske oversikt suppleres med en syntese av usikre tall, bør man imidlertid kjenne de studier som går inn den "svarte boks" (syntesen) og hva som skjer i "boksen". Metaanalysen er ikke statistisk alkymi, men den er heller ikke de vises sten. Brukt med kritikk og fornuft kan den være til stor nytte både for klinikere og helsemyndigheter.

LITTERATUR

1. Feinstein AR. Meta-analysis: statistical alchemy for the 21st century. *J Clin Epidemiol* 1995; 48: 71 – 9.
2. Sharp SJ, Thompson SG, Altman DG. The relation between treatment benefit and underlying risk in meta-analysis. *BMJ* 1996; 313: 735 – 8.
3. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999; 282: 1054 – 60.
4. Lindbæk M, Hjortdahl P. How do two meta-analyses of similar data reach opposite conclusions? *BMJ* 1999; 318: 873 – 4.
5. DerSimonian R, Levine RJ. Resolving discrepancies between a meta-analysis and a subsequent large controlled trial. *JAMA* 1999; 282: 664 – 70.
6. ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. *Lancet* 1995; 345: 669 – 85.
7. LeLorier J, Gregoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med* 1997; 337: 536 – 42.
8. Bailer JC. The practice of meta-analysis. *J Clin Epidemiol* 1995; 48: 149 – 57.
9. Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M et al. Systematic reviews and meta-analyses on treatment of asthma: critical evaluation. *BMJ* 2000; 320: 537 – 40.
10. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet* 1999; 354: 1896 – 900.

Publisert: 20. september 2000. Tidsskr Nor Legeforen.

© Tidsskrift for Den norske legeforening 2026. Lastet ned fra tidsskriftet.no 15. februar 2026.