
Individuell variasjon

MEDISIN OG TALL

ARE HUGO PRIPP

apripp@ous-hf.no

Are Hugo Pripp er forsker og biostatistiker ved Oslo senter for biostatistikk og epidemiologi, Forskningsstøtteavdelingen, Oslo universitetssykehus. Han er professor II ved Fakultet for helsevitenskap, OsloMet – storbyuniversitetet. Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

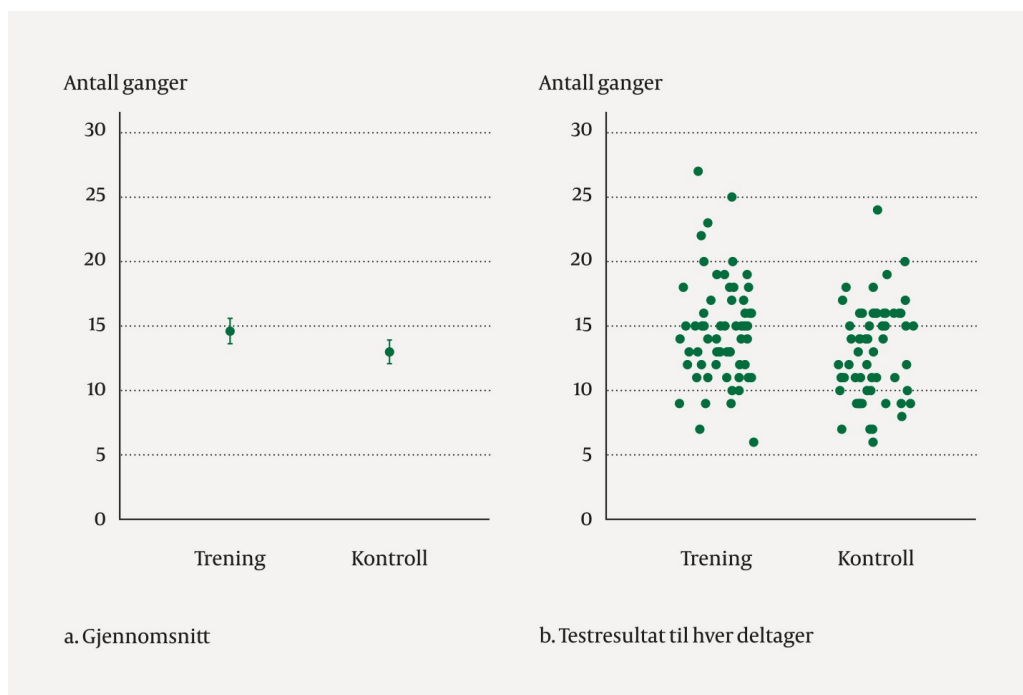
En medisinsk behandling med statistisk signifikant effekt garanterer ikke en betydningsfull effekt for den enkelte pasient. Den individuelle variasjonen hos pasientene er ofte større og betyr mer enn den gjennomsnittlige effekten på gruppenivå.

Statistikk er mer opptatt av gruppen enn av individet. Utfallsmål som gjennomsnittlig forskjell, relativ risiko og oddsratio, med tilhørende konfidensintervaller og *p*-verdier, sier noe om effekten på gruppenivå. Signifikante resultater er vel så ofte et resultat av mange studiedeltagere som av betydningsfulle behandlingseffekter for hver enkelt pasient. Vi bør derfor undersøke i hvilken grad de statistiske analysene forklarer den individuelle variasjonen.

Statistisk forklaringsevne

I en randomisert studie med benskjøre eldre kvinner gjennomførte intervensjonsgruppen et treningsopplegg og ble sammenlignet med en kontrollgruppe (1). Et av utfallsmålene ved oppfølging etter tre måneder var antall ganger deltagerne kunne reise seg fra en stol i løpet av 30 sekunder. Treningsgruppen reiste seg i gjennomsnitt 1,6 ganger mer enn kontrollgruppen, med et gjennomsnitt på hhv. 14,6 og 13,0 ganger for de to gruppene (figur 1a). Konfidensintervallene gjenspeiler den statistisk signifikante forskjellen mellom gruppene, og viser at trening hjalp deltagerne «opp av stolen». På individnivå var det likevel stor variasjon, og mange i kontrollgruppen hadde bedre resultat enn deltagere i treningsgruppen (figur 1b). På individnivå virker forskjellen, og dermed treningseffekten, mindre enn på gruppenivå.

Informasjon om den individuelle variasjonen i dataene og hvor mye den statistiske analysen forklarer denne, er viktig. Et vanlig mål er determinasjonskoeffisienten, ofte kalt R^2 . Vi bruker den typisk ved regresjonsanalyser, og den forteller oss hvor mye av variasjonen i utfallsmålet som er forklart av regresjonsmodellen. Den går fra 0 ved ingen forklaring til 1 ved perfekt eller 100 % forklaring av variasjonen. I eksempelet i figur 1 var $R^2 = 0,04$. Altså er 4 % av variasjonen mellom studiedeltagerne forklart av den gjennomsnittlige forskjellen mellom de to gruppene. Er utfallsmålet binært (f.eks. syk eller frisk) og dataene analysert med logistisk regresjon, finnes varianter av R^2 som f.eks. Tjurs R^2 , McFaddens R^2 og Nagelkerkes R^2 (2). For en logistisk regresjonsmodell er det også relevant å si hvilken evne modellen har til å skille mellom de to gruppene, f.eks. mellom syke og friske. Et vanlig mål for dette er det såkalte arealet under kurven, der 0,5 betyr ingen evne utover tilfeldigheter, og 1 betyr en perfekt evne.



Figur 1 I en intervensjonsstudie (1) reiste treningsgruppen seg i gjennomsnitt 1,6 (95 %-KI 0,2 til 3,0) ganger mer enn kontrollgruppen fra en stol i løpet av 30 sekunder, med a) gjennomsnitt på hhv. 14,6 og 13,0 ganger, men b) på individnivå var det stor variasjon, og mange i kontrollgruppen hadde bedre resultat enn deltagere i treningsgruppen, som illustrert i dette såkalte *jitter*-plottet, der punktene representerer individuelle observasjoner.

Det finnes også andre avanserte statistiske mål for hvor godt den statistiske modellen er tilpasset dataene. Akaikes informasjonskriterium (AIC, *Akaike information criterion*) og bayesiansk informasjonskriterium (BIC, *Bayesian information criterion*) er to mål som tar hensyn til både tilpasningen av dataene og modellens kompleksitet. Tolkningen og bruken er vanskeligere enn for R^2 , og disse målene benyttes i hovedsak til å sammenligne forskjellige statistiske modeller basert på det samme datasettet.

Hva kan din statistiske modell gjøre for andres data?

Jo mer kompleks den statistiske modellen er, desto bedre kan vi tilpasse den til dataene som er brukt i beregningen av modellen. Dilemmaet er at vi bruker de samme dataene for å beregne den statistiske modellen som for å vurdere hvor god den er, noe som ofte

fører til en overtilpasning av modellen. Bruker vi den overtilpassete statistiske modellen på nye data, fungerer den dårlig selv om den var god for de opprinnelige dataene. En statistisk modell bør fungere godt med nye data, f.eks. slik at den kan forutsi prognosen til nye pasienter. Ideelt sett bør vi teste ut den statistiske modellen med et nytt datasett, såkalt ekstern validering. Hvis vi ikke har tilgang til et nytt datasett, kan vi prøve ut modellen med nye kombinasjoner av de opprinnelige dataene med metoder som bootstrapping eller kryssvalidering, såkalt intern validering (3).

Legevitenskap eller legekunst

Selv med en så god statistisk analyse som vitenskapelig mulig er det ofte en betydelig individuell variasjon som vi ikke kan forutsi eller forklare. Den uforklarte forskjellen mellom enkeltpasienter blir i den statistiske modellen en del av feilledet (eng. *residual term*) og vitenskapelig lett sett på som tilfeldig variasjon i dataene. I behandlingen av den enkelte pasient blir da kanskje legekunst vel så viktig som legevitenskap.

REFERENCES

1. Stanghelle B, Bentzen H, Giangregorio L et al. Physical fitness in older women with osteoporosis and vertebral fracture after a resistance and balance exercise programme: 3-month post-intervention follow-up of a randomised controlled trial. BMC Musculoskelet Disord 2020; 21: 471. [PubMed][CrossRef]
2. Hughes G, Choudhury RA, McRoberts N. Summary Measures of Predictive Power Associated with Logistic Regression Models of Disease Risk. Phytopathology 2019; 109: 712–5. [PubMed][CrossRef]
3. Steyerberg EW. Validation of Prediction Models. I: Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Cham: Springer International Publishing, 2019: 329–44.

Publisert: 21. november 2022. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.22.0532

Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 14. juli 2026.