



Balanserte eller ubalanserte utvalg?

MEDISIN OG TALL

STIAN LYDERSEN

E-post: stian.lydersen@ntnu.no

Stian Lydersen er dr.ing. og professor i medisinsk statistikk ved Regionalt kunnskapssenter for barn og unge – psykisk helse og barnevern, Institutt for psykisk helse, Fakultet for medisin og helsevitenskap, Norges teknisk-naturvitenskapelige universitet.

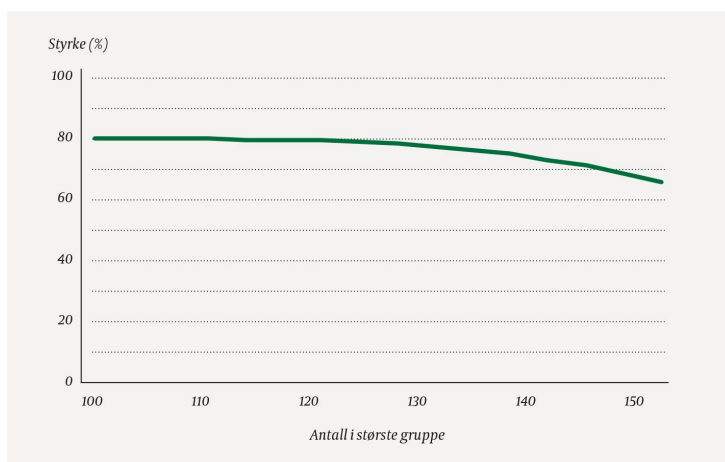
Forfatter har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

Når man skal sammenligne to grupper, planlegger man ofte å inkludere to som er like store. Men en moderat ubalanse behøver ikke bety nevneverdig redusert statistisk styrke. Og i visse sammenhenger kan ubalanse faktisk gi økt statistisk styrke.

I randomiserte kontrollerte studier planlegger man vanligvis at gruppene skal være like store. Men med enkelte randomiseringsprosedyrer, særlig i multisenterstudier, vil den ene gruppen kunne bli noe større enn den andre, selv om man randomiserer 1:1. Dette viser seg imidlertid å være uproblematisk, fordi statistisk styrke ikke blir særlig redusert ved moderat ubalanse (1, s. 46). I andre studier, for eksempel i kasus-kontroll-studier, kan antall personer tilgjengelig for den ene gruppen være begrenset. Da kan styrken økes en god del ved å inkludere flere i den andre gruppen.

Når er moderat ubalanse uviktig?

Ta for eksempel en randomisert kontrollert studie som er planlagt med to grupper à 100 pasienter, dvs. totalt 200 pasienter. Ved normalfordelt utfallsvariabel og en effektstørrelse på 0,4 standardavvik vil studien få en statistisk styrke på 80,4 % ved signifikansnivå 5 %. Hvis man i stedet inkluderer henholdsvis 110 og 90 pasienter i de to gruppene, som er en ganske stor tilfeldig ubalanse ved 1:1-randomisering, ville styrken bli 80,0 % under de samme forutsetningene. En slik ubalanse har altså en neglisjerbar effekt på styrken. Selv en 2:1-randomisering med 133 og 67 i de to gruppene ville gi en beskjeden reduksjon i styrke til 75,5 %. Først når den ene gruppen er mye større enn den andre, ville man få en vesentlig reduksjon i styrke. For eksempel ville gruppestørrelser på 150 og 50 gi en styrke på bare 68,4 % (fig 1).

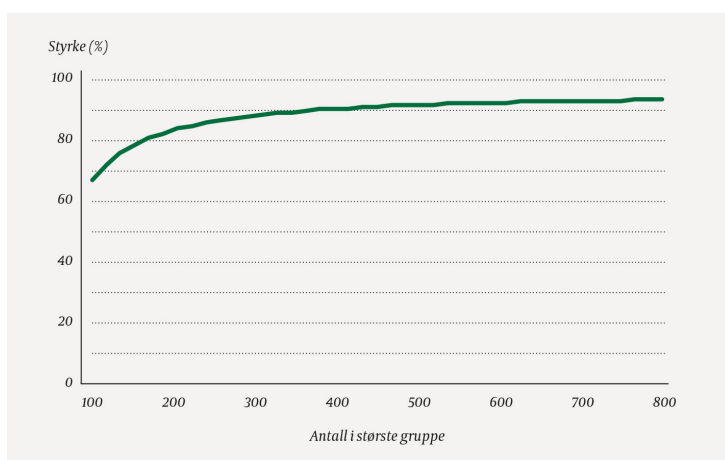


Figur 1 Statistisk styrke for sammenligning av to grupper med t-test. Antallet i de to gruppene er til sammen 200. Beregnet ved effektstørrelse på 0,4 standardavvik og signifikansnivå 5 %

Når er ubalanse gunstig?

I enkelte situasjoner kan det være en begrensning på hvor mange personer man kan ha i den ene behandlingsgruppen, f.eks. fordi denne behandlingen er meget kostbar, mens det er billigere eller enklere å inkludere flere i kontrollgruppen. Da kan man øke den statistiske styrken betraktelig ved å øke antall personer i kontrollgruppen. La oss si at kun 70 personer kan inkluderes i behandlingsgruppen. Ved å inkludere 70 personer i kontrollgruppen, ville man under samme forutsetninger som ovenfor få en styrke på bare 65,2 %. Vanligvis ønsker man en statistisk styrke på minst 80 % for å gjennomføre en studie. Økes kontrollgruppen til 140, får man en styrke på 77,6 %, mens en kontrollgruppe på 210 gir en styrke på 82,3 %.

Tilsvarende problemstilling kan være aktuell i en kasus-kontroll-studie av en sjelden sykdom der det vil være få tilgjengelige pasienter. Kontrollgruppen vil være personer uten denne sykdommen. Man registrerer hvor mange i hver gruppe som hadde vært utsatt for en bestemt eksponering, for å undersøke om det er en sammenheng mellom eksponeringen og sykdommen. La oss si at kun 100 pasienter er tilgjengelig og at vi forventer at 10 % av disse og 2 % av kontrollpersonene har vært eksponert. Dersom vi inkluderer 100 kontrollpersoner i studien, vil statistisk styrke bli bare 66,6 % ved signifikansnivå 5 %. Men ved å øke antall kontrollpersoner til 200 eller 400 blir styrken 82,4 % eller 89,8 %, som normalt anses som tilstrekkelig for å gjennomføre en studie. Dette er illustrert i figur 2. En tommelfingerregel sier at det oftest er lite å vinne på å øke den største gruppen til mer enn 4–5 ganger den minste gruppen (1, s. 124). Dette ser vi et eksempel på i figur 2, da grafen nærmest «flater ut» ved 400–500 i den største gruppen.



Figur 2 Statistisk styrke for sammenligning av to grupper med Pearsons khikvadrattest ved signifikansnivå 5 %. I dette eksemplet er det 100 i den minste gruppen, og sannsynligheten i den minste og største gruppen er hhv. 10 % og 2 %

Konklusjon

Vi har sett på to forskjellige situasjoner: Når det totale antallet i en studie er gitt, vil ikke styrken bli vesentlig redusert ved noe ubalanse i antallet mellom de to gruppene. Når antallet i den ene gruppen er begrenset, kan styrken økes vesentlig ved å øke antallet i den andre gruppen.

LITTERATUR:

1. van Belle G. Statistical rules of thumb. 2. utg. Hoboken, NJ: Wiley, 2008.

Publisert: 17. september 2018. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.18.0539
© Tidsskrift for Den norske legeforening 2020. Lastet ned fra tidsskriftet.no