
Overkrysningsforsøk

MEDISIN OG TALL

STIAN LYDERSEN

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 (RKBU Midt-Norge) ved Institutt for psykisk helse, NTNU.

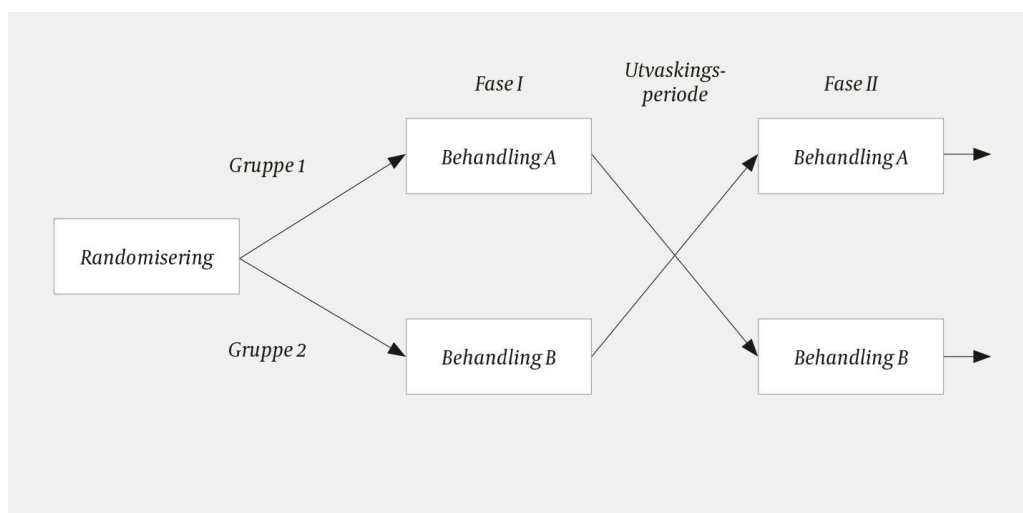
Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

De fleste randomiserte kontrollerte forsøk gjennomføres med parallelle grupper, men enkelte behandlinger kan studeres mer effektivt i et overkrysningsforsøk.

I en randomisert studie med parallelle grupper randomiseres deltakerne enten til behandling A eller til behandling B. I et overkrysningsforsøk gjennomgår hver deltaker både behandling A og behandling B.

Forskjellig behandling i to faser

La oss starte med et eksempel hvor vi undersøkte effekten av et probiotikum ved irritable tarm-syndrom i et dobbeltblindt placebokontrollert forsøk (1). Dette ble gjennomført som et overkrysningsforsøk, som illustrert i figur 1. Et slikt forsøk gjennomføres i to faser: Etter inklusjon blir deltakerne randomisert til gruppe 1 eller gruppe 2. I fase I får gruppe 1 behandling A og gruppe 2 behandling B. I fase II blir det omvendt: Da får gruppe 1 behandling B og gruppe 2 behandling A. Mellom de to fasene lar man det gå en viss tid uten behandling, for å vaske ut effekten av behandlingen fra fase I.



Figur 1 Et overkrysningsforsøk med to behandlinger og to faser.

Et slikt design kan ikke brukes for enhver type behandling: Det kan for eksempel ikke brukes for å sammenlikne kurativ effekt av to behandlinger. Men designet kan være godt egnet for å sammenlikne symptomreducerende effekt av behandlinger for kroniske plager. Den store fordelen ved designet er at hver deltaker er sin egen kontroll. Variasjonen mellom deltakerne er ofte høy sammenliknet med variasjonen innen den enkelte deltaker. Når hver deltaker er sin egen kontroll, reduseres den tilfeldige variasjonen. Man oppnår dermed høyere statistisk styrke, og forsøket kan gjennomføres med færre deltakere enn ved parallelle grupper. Men alle deltakerne må gjennomføre to faser, slik at dette kan ta lengre tid, noe som også kan føre til større risiko for frafall.

Parede data

Når man analyserer resultatene, må man ta hensyn til at dataene er parett, fordi resultatene av behandling A og behandling B vil være positivt assosiert innen den enkelte deltaker. Hvis utfallsvariabelen er kontinuerlig, vil en enkel analysemetode kunne baseres på differansen mellom utfallsvariablene ved behandling A og behandling B. En parett *t*-test med tilhørende konfidensintervall kan være aktuell. En av variablene i den nevnte studien var summen av mageplager målt ved en skår. De 16 deltakerne som fullførte studien, hadde en gjennomsnittsskår (standardavvik) på 6,44 (1,81) i den aktive perioden og 5,35 (1,77) i placeboperioden. Forskjellen i skår mellom aktiv periode og placebo var $-1,09$ (1,47), i favør av placebo. Korrelasjonen mellom skårene i de to periodene var 0,66. En analyse basert på en parett *t*-test gav et 95%-konfidensintervall på $-1,87$ til $-0,31$ og en *p*-verdi på 0,010.

I noen overkrysningsforsøk vil det være hensiktsmessig å bruke en mer kompleks analysemodell enn en parett *t*-test. Det kan for eksempel være aktuelt å ta hensyn til en periodeeffekt, som at utfallsvariabelen i fase II kan være systematisk høyere enn i fase I. Vi har så langt beskrevet overkrysningsforsøk med to behandlinger og to faser. Designet kan også generaliseres til flere behandlinger eller flere faser. Slikt kan håndteres i en lineær blandet effekt-modell (2, s. 61–72). Hvis utfallsvariabelen er kategorisk, må man bruke andre metoder, noe vi kommer tilbake til i en senere artikkel i denne spalten.

Høyere styrke enn ved parallelle grupper

Hva hvis det nevnte forsøket hadde vært gjennomført med to parallelle grupper med 16 deltakere i hver gruppe isteden? Hvis resultatene ovenfor hadde stammet fra et parallellgruppeforsøk, ville en to-utvalgs t -test gitt et 95%-konfidensintervall på $-2,38$ til $0,20$ og en p -verdi på $0,095$, altså et bredere konfidensintervall enn overkrysningsforsøket, selv om det var tenkt med $16 \times 2 = 32$ deltakere, altså dobbelt så mange som i overkrysningsforsøket.

Styrken i et overkrysningsforsøk kan også illustreres ved å se på hvor mange deltakere man trenger. La oss si at man planlegger en studie der utfallsvariabelen har et standardavvik på $1,5$, og man ønsker å kunne påvise en gjennomsnittlig forskjell på $1,0$. Hvis man planlegger et overkrysningsforsøk, og korrelasjonen antas å være $0,50$, trenger man totalt 20 deltakere for å oppnå en statistisk styrke på 80% ved signifikansnivå 5% . Hvis man derimot planlegger et forsøk med parallelle grupper, trengs 37 deltakere i hver gruppe, totalt 74 deltakere.

Det er begrenset hvilke behandlinger og forskningsspørsmål som egner seg for en overkrysningsstudie. Men i de tilfellene der det er mulig, er dette et meget effektivt design som kan vurderes.

REFERENCES

1. Ligaarden SC, Axelsson L, Naterstad K et al. A candidate probiotic with unfavourable effects in subjects with irritable bowel syndrome: a randomised controlled trial. *BMC Gastroenterol* 2010; 10: 16. [PubMed][CrossRef]
2. Twisk JWR. Analysis of data from randomized controlled trials. A practical guide. Cham: Springer, 2021.

Publisert: 2. mai 2022. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.22.0013
Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 7. juli 2026.