



Tidsskriftet

DEN NORSKE LEGEFORENING

Type I-feil og type II-feil

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

Anta at man vil prøve ut en ny type medisinsk behandling for å finne ut om denne har bedre effekt enn vanlig behandling (*treatment as usual*). Det er viktig at sannsynligheten for å komme til feil konklusjon er lav. Men det finnes to typer feilslutninger, og de regnes ikke som like viktige.

I utgangspunktet vet vi ikke hva som er sannheten. Det kan hende at de to behandlingene har samme effekt. Dette kalles nullhypotesen (H_0). Eller den nye behandlingen kan ha bedre effekt enn vanlig behandling. Dette kalles alternativhypotesen (H_A). Merk at dette gjelder for en overlegenhetsstudie (*superiority study*), slik det er beskrevet her. I en ikke-underlegenhetsstudie (*non-inferiority study*), derimot, vil nullhypotesen være at den nye behandlingen er dårligere enn vanlig behandling (1).

To typer feil

Dersom sannheten er at behandlingene har samme effekt, og man konkluderer med at man beholder H_0 , er dette en korrekt konklusjon. Hvis man derimot feilaktig konkluderer med å forkaste H_0 , gjør man en type I-feil. Dette er illustrert i Figur 1. Da vil man kunne ta i bruk den nye behandlingen, til tross for at den i virkeligheten ikke er bedre enn vanlig behandling.

| | | |
|---|---|--|
| | Sannhet H_0 er sann | Sannhet H_A er sann |
| Beslutning Beholde H_0 | Riktig konklusjon (sant negativ) | Gal konklusjon: Type II-feil (falskt negativ) |
| Beslutning Forkaste H_0 (og påstå H_A) | Gal konklusjon: Type I-feil (falskt positiv) | Riktig konklusjon (sant positiv) |

Figur 1 Vi ønsker i størst mulig grad å unngå en type I-feil, illustrert med rød farge, mens vi vanligvis ikke setter like sterke krav til sannsynligheten for type II-feil (gul farge). Dersom sannheten er at den nye behandlingen virkelig er bedre enn standard behandling, men ikke påviser dette, gjør man en type II-feil. Da vil man ikke ta i bruk den nye behandlingen, til tross for at den er bedre. Hvis man derimot konkluderer med å forkaste H_0 , gjør man det riktige.

Hypotesene er ikke likeverdige

Situasjonen er ikke likeverdig eller symmetrisk i H_0 og H_A . Hvis vi er i tvil, vil vi akseptere nullhypotesen. Det vil si, vi aksepterer at den *kan* være sann. Vi ønsker i størst mulig grad å unngå en type I-feil, som illustrert ved fargene i Figur 1. Når man planlegger en studie og den tilhørende statistiske analysen, vil man typisk kreve at sannsynligheten for en type I-feil skal være liten. Ofte kreves at sannsynligheten for type I-feil skal være maksimalt 5%. Denne grensen for sannsynligheten kalles signifikansnivået. Man stiller vanligvis ikke like sterke krav til sannsynligheten for type II-feil. Sannsynligheten for korrekt å forkaste H_0 når H_A er riktig, kalles statistisk styrke. Det er ganske vanlig å planlegge studier slik at statistisk styrke er minst 80 % eller 90 %. Da vil sannsynligheten for type II-feil bli inntil 20 % eller 10 %. Merk at denne asymmetrien også gjenspeiles i ordvalget: Vi aksepterer H_0 , eller vi forkaster H_0 (og påstår H_A).

Straffesaker i rettsvesenet har en asymmetri som er analog til dette: Der vil nullhypotesen være at tiltalte er uskyldig, mens alternativhypotesen er at tiltalte er skyldig. Hvis retten er i tvil, skal tiltalte ikke dømmes. Og hvis man er i tvil ved hypotesetesting, så aksepterer man at nullhypotesen kan være sann.

Alternativt kan man kalle type I-feil og type II-feil for henholdsvis falskt positive og falskt negative funn – se Språkspalten i dette nummeret av Tidsskriftet (2).

Relevant ved planlegging av en studie

Merk at sannsynlighet for type I-feil og sannsynlighet for type II-feil er relevante størrelser ved planlegging av en studie. Etter at studien og analysene er gjennomført, vil relevante mål på usikkerhet i forhold til konklusjonen være konfidensintervall og p-verdi (3).

LITTERATUR

1. Skovlund E. Hvordan vise likhet? Tidsskr Nor Legeforen 2017; 137. doi:10.4045/tidsskr.17.0668. [PubMed][CrossRef]

2. Lydersen S. Type I-feil og type II-feil – eller falskt positive og falskt negative funn? Tidsskr Nor Legeforen 2021; 141. doi: 10.4045/tidsskr.21.0118. [CrossRef]

3. Lydersen S. Statistisk styrke – før, men ikke etter! Tidsskr Nor Legeforen 2019; 139. doi: 10.4045/tidsskr.18.0847. [PubMed][CrossRef]

Publisert: 21. juni 2021. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.21.0013

© Tidsskrift for Den norske legeforening 2023. Lastet ned fra tidsskriftet.no 7. juni 2023.