
Hvor stor andel av publiserte forskningsresultat er 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. Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

METTE LANGAAS

Mette Langaas er dr.ing., professor i statistikk ved Institutt for matematiske fag, NTNU, og professor II ved avdeling for statistisk analyse, maskinlæring og bildeanalyse (SAMBA) ved Norsk Regnesentral. Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

Ikke alle publiserte forskningsresultat lar seg reprodusere – noen fordi funnene ikke er korrekte. Hvor omfattende er problemet?

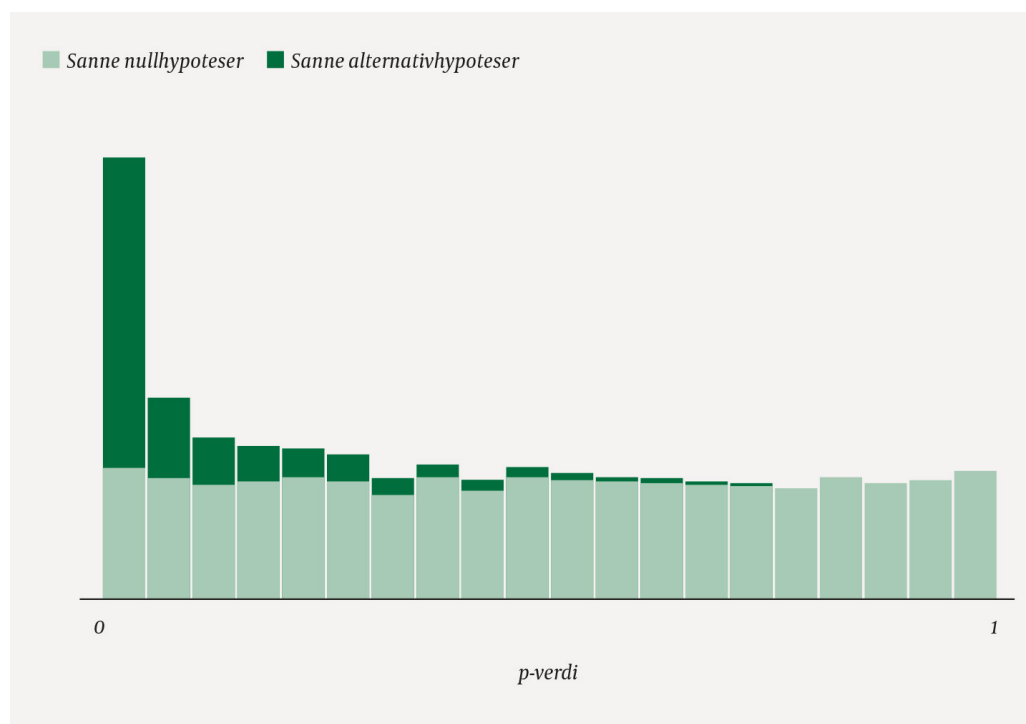
Flere forskere har forsøkt å beregne hvor ofte publiserte funn er feil. Tilnærmingene er til dels svært ulike.

Ulike metoder

Artikkelen «Why most published research findings are false» av John Ioannidis ble publisert i 2005, og fikk stor oppmerksomhet [\(1\)](#). Artikkelen var ikke basert på data, men postulerte en modell for andelen falskt positive funn blant publiserte positive funn, basert på følgende fire størrelser: andelen faktisk

sanne hypoteser blant de hypotesene som undersøkes, statistisk styrke, signifikansnivå (5 %) og skjevhet (*bias*). Skjevhet i denne sammenhengen betyr den andelen av studiene hvor hypotesen ville fremstå som sann selv om den ikke er det, f.eks. på grunn av publikasjonsskjevhet eller dårlig studiedesign. Ioannidis beregnet positiv prediktiv verdi, altså andelen sanne funn blant positive funn, for en rekke forskjellige kombinasjoner av disse fire størrelsene. For store randomiserte kontrollerte studier med adekvat styrke (80 %) anså han det som realistisk at andelen faktisk sanne nullhypoteser kan være 50 %, og at skjevheten er bare 10 %. Da blir positiv prediktiv verdi beregnet til 85 %. For utforskende observasjonelle studier med adekvat styrke på 80 %, andel sanne nullhypoteser på 9 % og skjevhet på 30 % fås en positiv prediktiv verdi på 20 %. Studier med lavere andel sanne nullhypoteser eller lavere styrke fører til enda lavere positiv prediktiv verdi ((1), tabell 4).

Jager og Leek estimerte i 2014 andelen falskt positive funn basert på data (2). De leste elektronisk alle de 77 430 publikasjonene i The Lancet, The Journal of the American Medical Association, The New England Journal of Medicine, BMJ og The American Journal of Epidemiology i 2000, 2005 og 2010. Analysene bygger på det faktum at når nullhypotesene er sanne, vil *p*-verdiene være jevnt fordelt fra 0 til 1, mens når alternativhypotesene er sanne, vil *p*-verdiene være skjevt fordelt mot 0. Dette er illustrert i figur 1.



Figur 1 Typisk fordeling av *p*-verdier for sanne nullhypoteser (lysegrønt) og sanne alternativhypoteser (mørkegrønt).

Jager og Leek estimerte andelen falskt positive funn (*science-wise false discovery rate*) til å være 14 %. Sammen med artikkelen deres ble det publisert diskusjonsinnlegg fra flere forskere. Avslutningsvis kom et tilsvarende fra Jager og Leek (3) der de skrev at estimatet på 14 % antakelig er optimistisk, men at andelen trolig ikke er over 50 %, i alle fall ikke for studier som er godt planlagt og godt gjennomført.

Forskere i gruppen Open Science Collaboration brukte en annen fremgangsmåte for å studere reproduktibilitet (4). De identifiserte 100 studier publisert i 2008 i tre forskjellige psykologitidsskrift. Disse studiene ble gjentatt i nye studier med nye deltakere, med en design mest mulig lik den opprinnelige og med planlagt statistisk styrke på minst 80 %. Dette var et meget omfattende arbeid, og hele 274 personer er listet som forfattere. Hva fant man så? Den estimerte effekten målt ved korrelasjonskoeffisienten var i originalstudiene i gjennomsnitt 0,403 (standardavvik 0,188), og i de gjentatte studiene bare 0,197 (0,257). Av de opprinnelige studiene rapporterte i alt 97 % statistisk signifikant effekt (p -verdi $< 0,05$), mot bare 36 % av de gjentatte studiene. Etter å ha kombinert de opprinnelige og de gjentatte studiene ble 68 % statistisk signifikante.

En fallitterklæring?

Disse tre studiene brukte svært ulike metoder. Ioannidis laget en modell som bygger på en rekke antakelser i forskjellige studiedesign. Antakelsene kan synes realistiske, men det er en svakhet at de ikke bygger direkte på data. Jager og Leek baserte seg på rapporterte p -verdier og gjorde beregninger basert på forventede fordelinger av p -verdier når nullhypotesene er sanne eller usanne. Open Science Collaboration gikk enda grundigere til verks: De gjentok 100 studier. Studiene til Jager og Leek og Open Science Collaboration bygger på omfattende empiri. Litt avhengig av hvordan man vektlegger resultatene, kan man si at andelen falskt positive funn ble estimert til henholdsvis 14 % og 29 % ($97 \% - 68 \% = 29 \%$) i disse studiene. Dette må uansett sies å være høye andeler.

Er dette en fallitterklæring for forskningen? Absolutt ikke. Men det belyser viktigheten av at studier er godt planlagt, gjennomført og rapportert. Og at det er viktig å forsøke å reproducere publiserte studier. Den samlede evidensen fra flere studier, gjerne i en systematisk oversikt eller metaanalyse, vil være vesentlig mer pålitelig enn den fra en enkelt studie.

LITTERATUR

1. Ioannidis JPA. Why most published research findings are false. PLoS Med 2005; 2: e124. [PubMed][CrossRef]
2. Jager LR, Leek JT. An estimate of the science-wise false discovery rate and application to the top medical literature. Biostatistics 2014; 15: 1–12. [PubMed][CrossRef]
3. Jager LR, Leek JT. Rejoinder: An estimate of the science-wise false discovery rate and application to the top medical literature. Biostatistics 2014; 15: 39–45. [PubMed][CrossRef]
4. Open Science Collaboration. Estimating the reproducibility of psychological science. Science 2015; 349: aac4716. [PubMed][CrossRef]

Publisert: 7. desember 2021. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.21.0703
Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 23. mars 2026.