

Er effekten liten eller stor?

MEDISIN OG TALL

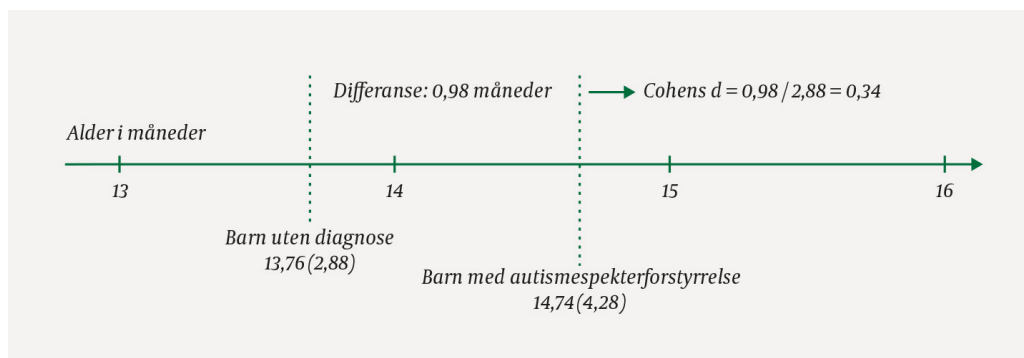
STIAN LYDERSEN

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.

Hvordan tallfester vi resultatet av en studie? Er det effekten målt på originalskalaen eller en standardisert effektstørrelse som er mest relevant?

Reindal og medarbeidere (1) studerte alderen til barn da de gikk uten støtte første gang. For barna som fikk diagnosen autismspekterforstyrrelse, var den gjennomsnittlige alderen (standardavvik) 14,74 (4,28) måneder. For barna uten autismspekterforstyrrelse var den 13,76 (2,88) måneder. Forskjellen var altså $14,74 - 13,76 = 0,98$ måneder. Dette kalles effekt målt på originalskalaen, eller ustandardisert effektstørrelse. I tillegg oppgir forfatterne en standardisert effektstørrelse i form av differansen delt på standardavviket i sammenlikningsgruppen, altså $0,98 / 2,88 = 0,34$. Denne kan tolkes som antall standardavvik i forskjell mellom gruppene (se figur 1). Hvilken av disse størrelsene er mest relevant?



Figur 1 Gjennomsnitt for alder (standardavvik) da de gikk uten støtte første gang, hos 376 barn med diagnosen autismspekterforstyrrelse og 114 uten diagnosen (1). Differansen var 0,98 måneder, som tilsvarte Cohens $d = 0,34$.

Hva er effektstørrelse?

Termen *effektstørrelse* er ikke presis. Noen forfattere bruker den i betydningen Cohens d eller et relatert mål som Glass' delta eller Hedges' g (2). Disse er differansen mellom to gjennomsnitt, delt på et standardavvik, og er eksempler på standardiserte effektstørrelser. Andre eksempler på standardiserte effektstørrelser er Pearsons korrelasjonskoeffisient, standardisert regresjonskoeffisient i lineær regresjon og partiell eta kvadrert i variansanalyser (ANOVA).

Det er ikke uvanlig å oppgi standardiserte effektstørrelser innen adferdsvitenskapene. Men hvilken rolle har de egentlig? Forskere som oppgir standardiserte effektmål, siterer vanligvis boka *Statistical Power Analysis for the Behavioral Sciences* av Jacob Cohen (1923–98) (3, 4). Cohen innfører i denne boka standardiserte effektstørrelser som grunnlag for å beregne statistisk styrke eller utvalgsstørrelse i en fremtidig studie. Men han skriver ikke noe om bruk av standardiserte effektstørrelser ut over denne anvendelsen.

Hva som er en relevant effektstørrelse i en studie etter at den er gjennomført, er avhengig av konteksten. Eksempler på ustandardiserte effektstørrelser er differanse mellom to gjennomsnitt, ustandardisert regresjonskoeffisient, oddsforhold (oddsratio) og risikodifferanse. Flere forfattere anbefaler generelt å rapportere ustandardiserte effektstørrelser (5, 6). En bredere diskusjon av ustandardiserte og standardiserte effektstørrelser finnes i artiklene (7) og (8).

Cohen selv klassifiserer Cohens d som liten, moderat, eller stor hvis den er lik henholdsvis 0,2, 0,5 eller 0,8 (4, s. 26). Andre forfattere klassifiserer standardiserte effektstørrelser i intervall, og til dels noe annerledes enn Cohen, se f.eks. (4, s. 79–80) og (9), s. 123). Klassifisering av standardiserte effektstørrelser kan være nyttig ved beregning av styrke og utvalgsstørrelse, men anses av flere forfattere å være lite relevant for å klassifisere observert effekt i en gjennomført studie ((5, 8).

Ustandardisert er lett å forstå

Når man oppgir at differansen i alder for å gå uten støtte mellom to grupper er 0,98 måneder, er dette lett å forstå. Gir den standardiserte effektstørrelsen Cohens $d = 0,34$ noen klinisk relevant tilleggsinformasjon? Standardiserte effektstørrelser kan som nevnt være nyttige som grunnlag for beregning av statistisk styrke eller utvalgsstørrelse i en fremtidig studie, og i tillegg kan de være nyttige som datagrunnlag i metaanalyser (2), men ut over dette synes de å være lite relevante.

LITTERATUR

1. Reindal L, Nærland T, Weidle B et al. Age of First Walking and Associations with Symptom Severity in Children with Suspected or Diagnosed Autism Spectrum Disorder. *J Autism Dev Disord* 2019 doi: 10.1007/s10803-019-04112-y. [PubMed][CrossRef]
2. Grissom RJ, Kim JJ. Effect sizes for research. Univariate and multivariate applications. 2. utg. New York: Routledge, 2012.
3. Cohen J. Statistical power analysis for the behavioral sciences. 1. utg. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.
4. Cohen J. Statistical power analysis for the behavioral sciences. 2. utg. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1988.
5. Pek J, Flora DB. Reporting effect sizes in original psychological research: A discussion and tutorial. *Psychol Methods* 2018; 23: 208–25. [PubMed][CrossRef]
6. Baguley T. Standardized or simple effect size: what should be reported? *Br J Psychol* 2009; 100: 603–17. [PubMed][CrossRef]
7. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen* 2012; 141: 2–18. [PubMed][CrossRef]
8. Kelley K, Preacher KJ. On effect size. *Psychol Methods* 2012; 17: 137–52. [PubMed][CrossRef]
9. Campbell MJ, Swinscow TDV. Statistics at Square One. 11. Utg. Wiley-Blackwell, 2009.

Publisert: 17. februar 2020. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.19.0665
Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 30. mars 2026.