

Arbetsrapport 2: Metaanalys baserad på primärstudier med Single Case Experimental Design

Analys, rapporteringsstandards och risk för bias

Innehåll

Arbetsrapport 2: Metaanalys baserad på primärstudier med Single Case Experimental Design	1
1. Inledning	2
2. Metaanalys baserad på data från SCED primärstudier	2
3. Effektmått	7
4. Programvaror.....	7
5. Antal studier	11
6. Heterogenitet	11
7. Publikationsbias	13
8. Stöd vid val av metaanalysstrategi.....	14
9. Rapporteringsstandards.....	17
10. Risk för bias-granskning	17
11. Medverkande.....	17
12. Referenser	18

1. Inledning

Single Case Experimental Design (SCED) är en forskningsansats som används för att studera effekter av en insats på individnivå se [1]. SCED kännetecknas av att enstaka fall (till exempel en person) undersöks under olika faser genom upprepade mätningar, där samma individ fungerar som sin egen kontroll.

Till skillnad från gruppbaseade studier bygger SCED på upprepade mätningar inom enskilda individer snarare än gruppmedelvärden. Detta kräver metaanalysmetoder som är anpassade för att hantera beroenden i data och variation både inom och mellan individer och studier.

Detta dokument ger en översikt över tillgängliga metoder för metaanalys av SCED primärstudier, inklusive centrala komponenter som effektmått, modelleringsstrategier, hantering av heterogenitet, risk för bias, och programvarustöd. Flödesscheman, tabeller och referenser till aktuell metodologisk litteratur tillhandahålls för praktisk vägledning.

2. Metaanalys baserad på data från SCED primärstudier

Vilken typ av metaanalys som kan genomföras beror bland annat på vilken sorts data som finns tillgänglig i primärstudierna; summerade effektstorlekar, fassammanfattade eller fullständiga rådata på individnivå, samt antal primärstudier som inkluderas i analysen. För en stabil och pålitlig SCED-metaanalys rekommenderas minst 30 primärstudier och minst 20 mätpunkter per individ [2].

I Faktaruta 1 beskrivs centrala komponenter att beakta vid genomförande av metaanalys. De vidareutvecklas därefter i detta kapitel (2.2 samt 2.3) samt i kapitlen 3–7. I kapitel 8 sammanfattas detta i tre tabeller som kan användas som stöd vid val av metaanalysstrategi.

Oavsett metaanalysmetod är en stor variation inbyggd i SCED primärstudier och heterogenitet är en förväntad egenskap. Det är inte troligt eller rimligt att anta att deltagardata som inkluderas gäller helt lika insats(er) eller utfall. Innan metaanalys genomförs bör därför en bedömning göras om sammanslagningen av insatser och utfall är rimlig.

Faktaruta 1 Centrala komponenter vid genomförande av metaanalys baserad på data från primärstudier med SCED.

- Datatillgång: Vilken typ av data som finns tillgängliga, summerade effektstorlekar eller rådata på individnivå.
- Antal primärstudier
- Metaanalysmetod: Den övergripande metod som används för att kombinera resultat från SCED primärstudier. Beroende av datatillgång.
- Estimeringsmetod: Tekniker för att skatta modellparametrar. Varierar med metaanalysmetod.
- Inferensmetod: Hur slutsatser dras från analyser (osäkerhetsuppskattningar, hypotesprövning). Varierar med metaanalysmetod.
- Datastruktur: Är analysen univariat (ett utfall) eller multivariat (flera utfall).
- Hierarkisk struktur: Är modellen enkel (unilevel) eller flernivå (multilevel).
- Effektmått: Avgöranden om lämpliga effektmått. Beroende av metaanalysmetod
- Heterogenitetshantering: Metoder för att identifiera och hantera variation mellan studier.
- Publikationsbiashantering: Sätt att identifiera och korrigera för publikationsbias.

2.2. Metaanalys-metoder

Det finns ett flertal olika metaanalys-metoder. Nedan presenteras ett urval.

Summary meta-analysis (SMA)

SMA [3] innebär att varje SCED-studie bidrar med en summerad (aggregerad) effektstorlek. Modellen kombinerar dessa med viktning baserat på varians.

När använda SMA:

- När endast sammanfattande effektstorlekar finns tillgängliga.
- För relativt homogena studier.
- När enkel syntes utan komplexa modellantaganden önskas.

Begränsningar:

- Ignorerar klustring inom studier.
- Kräver enhetliga effektstorlekar (till exempel endast SMD eller Tau-U).

Aggregate data meta-analysis (ADMA)

ADMA [4] utgår också från summerade effektstorlekar men tillåter meta-regression, där effektskattningar relateras till studienivåkaraktäristika.

När använda ADMA:

- När studieegenskaper (till exempel designtyp eller insatslängd) ska analyseras.
- När endast summerade data är tillgängliga men heterogenitet behöver förklaras.
- Vid olika designtyper eller studiedurationer att modellera.

Begränsningar:

- Ignorerar klustring inom studier.
- Kräver ofta studienivå-förklarande variabler.

Bayesian meta-analysis (BMA)

BMA är en ramverksteknik (kan vara både univariat, multivariat, multilevel) som kombinerar tidigare kunskap (priors) med observerade data för att uppskatta sannolikheten för olika effektstorlekar (posterior).

När använda BMA:

- Vid små urval med få studier.
- När tidigare kunskap kan formaliseras som priors.
- Vid komplex datastruktur med tidsberoende.
- När osäkerhet behöver kvantifieras med sannolikhetsfördelningar.
- Praktisk handledning i att anpassa Bayesianska modeller till SCED-data se [5].

Fördelar:

- Hanterar väl små urval, tidsberoende och heterogenitet [6] [7].
- Kan inkludera AR (1) -struktur för att hantera autokorrelation.
- Möjliggör detaljerad osäkerhetskvantifiering via posteriorfördelningar.
- Kan hantera modellosäkerhet genom modellvägning

Begränsningar:

- Kräver specifikation av prior-fördelningar.
- Beräkningsmässigt mer krävande.
- Analysen sker ofta i specialiserade program (Stan, JAGS).

Multivariate meta-analysis (MMA)

Multivariate Meta-Analysis (MMA) [8] är en utvidgning av traditionell univariat metaanalys, där flera relaterade utfall analyseras samtidigt inom samma modell.

När använda MMA:

- När flera relaterade utfall ska analyseras samtidigt.
- Vid korrelation mellan utfallsmått.
- När vissa utfall saknas i vissa studier.
- När samtliga utfall rapporteras för samma individer eller enheter.

Fördelar:

- Tar hänsyn till korrelationen mellan utfall.
- Ökar precisionen i skattningar.
- Minskar bias vid saknade data för vissa utfall.
- Möjliggör gemensam inferens över flera utfall.

Begränsningar:

- Kräver information om korrelation mellan utfall.
- Mer komplex modellering än univariat metaanalys.
- Vanliga estimationsmetoder är Maximum Likelihood (ML) och Restricted Maximum Likelihood (REML).

Time-series meta-analysis (TSMA)

TSMA [2] används för att analysera nivå- och trendförändringar över tid inom faser (t.ex. ABAB-design).

När använda TSMA:

- Vid fokus på förändring över tid.
- När data har tydlig tidsstruktur.
- Vid många mätpunkter per individ.
- Kan vara univariat och multivariat beroende på om man modellerar flera utfall.

Fördelar:

- Hanterar seriekorrelation med autoregressiva modeller.
- Möjliggör analys av både nivå- och trendförändringar.

Begränsningar:

- Kräver strukturerad tidsdata.
- Mer komplex modellering.

Multi-Level Meta-Analysis (MLMA)

MLMA [9] beaktar den hierarkiska struktur som kännetecknar SCED genom att särskilja variation mellan mätningar, individer och studier.

När använda MLMA:

- När rådata är tillgängliga.
- Vid tillräckligt många studier (>30) och mätpunkter (≥ 20 per deltagare) [2].
- Robust Variance Estimation (RVE) kan användas om det är ett litet antal studier för att korrigera standardfel när varianskomponenter är svåra att estimeras [10] [11].
- När målet är att modellera moderatoreffekter på flera nivåer.

Fördelar:

- Särskiljer variation mellan mätningar, individer och studier [12] [13].
- Möjliggör modellering av moderatoreffekter på flera nivåer.

- Hierarkisk metaregression är en utvidgning av MLMA för att analysera moderatoreffekter. Tre-nivåmodellen har i simuleringar visat låg bias och hög precision vid minst 30 studier och ≥ 20 mätpunkter per individ [12].
- Ger mer exakta och tillförlitliga skattningar.

Begränsningar:

- Kräver tillgång till rådata.
- Kräver tillräckligt många studier och mätpunkter.

Individual participant data meta-analysis (IPD-MA)

IPD-MA [3] använder individdata från varje SCED-studie och möjliggör detaljerad modellering av tidsstruktur, förklarande variabler och individuella effekter.

När använda IPD-MA:

- När rådata på individnivå är tillgänglig.
- Vid behov av detaljerad modellering av tidsserier.
- När individuella förklarande variabler ska inkluderas i analysen.
- För att modellera icke-normaliserade utfall med GLMM (Generalized Linear Mixed Models).

Fördelar:

- Arbetar direkt med tidsseriedata för ökad flexibilitet.
- Möjliggör detaljerad modellering av individspecifika faktorer.
- Kan hantera autokorrelation med AR (1) -struktur.
- Kan använda GLMM för icke-normaliserade utfall.

Begränsningar:

- Kräver tillgång till rådata på individnivå.

2.3. Estimeringsmetoder

SCED-metaanalys kräver särskilda statistiska metoder för att hantera autokorrelation och beroendestrukturer. Här följer en kort översikt av olika estimeringsmetoder.

Robust variance estimation (eng. Robust variance estimation) (RVE)

- Beskrivning: Metod för att få korrekta konfidensintervall även när effektstorlekarna är beroende eller har felaktigt uppskattade varianser [10] [14].
- Användning i SCED: Lämplig för SCED-metaanalys där flera effektstorlekar kommer från samma studie eller individ. RVE tillåter inkludering av alla beroende effekter i en analys utan att exakt specificera korrelationsstrukturen [14]. Tiptons småsampelkorrigering [11] gör metoden mer tillförlitlig även med få studier. Nyligen har simulering [15] visat att en kombination av en tre-nivåers modell och RVE kan ge väl kalibrerade inferenser även om det finns autokorrelation i data. Implementeras i R via paketen robumeta och clubSandwich [10] [11] [16] [17].

Generaliserade minsta kvadratmetoden (Eng. Generalized least squares (GLS))

- Beskrivning: En form av viktad regressionsanalys där studier med högre precision får större vikt [18].
- Användning i SCED: Särskilt användbar för att hantera att mätningar inom SCED primärstudier som har olika precision. Mätpunkter med många observationer ges större vikt medan de med större osäkerhet får mindre inflytande.

Maximum likelihood (ML) och restricted maximum likelihood (REML)

- Beskrivning: Estimeringsmetoder för modeller med blandade effekter där parametrar uppskattas genom att maximera sannolikheten givet data [18].
- Användning i SCED: Lämpliga när man modellerar variation mellan studier med mixade effekter. REML är ofta att föredra i SCED-sammanhang eftersom den ger mindre snedvridna skattningar av varianskomponenter, vilket är viktigt när antalet studier är begränsat.

Bayesianska metoder, exempelvis Markov Chain Monte Carlo (MCMC)

- Beskrivning: Kombinerar tidigare kunskap (priors) med observerade data för att skapa uppdaterade sannolikhetsfördelningar.
- Användning i SCED: Värdefulla i SCED-kontext där stickproven ofta är små och datastrukturerna komplexa [7] [19]. Möjliggör modellering av hierarkiska strukturer och autokorrelation. Implementeras i R via paket som brms, med underliggande beräkningar i JAGS eller Stan [20] [21].

Fixed effect model (FEM) och random effect-modell (REM)

- Beskrivning: Val mellan att anta en gemensam sann effekt för alla studier (FEM) eller att tillåta att effekter varierar mellan studier (REM) [22].
- Användning i SCED: REM rekommenderas nästan alltid för SCED-metaanalyser eftersom de tillåter variation i effekter mellan studier [18], vilket speglar verkligheten i SCED-forskning där kontext, design och deltagaregenskaper ofta varierar avsevärt [23].

För SCED-metaanalys är RVE och bayesianska metoder särskilt värdefulla tillskott till de mer klassiska ML- och GLS-baserade metoderna, eftersom de bättre hanterar den beroendestruktur och småskalighet som kännetecknar SCED-data [15] [14].

3. Effektmått

Det finns olika effektmått som kan användas vid statistisk analys av data från SCED primärstudier. Manolov och kollegor erbjuder en guide för att välja sådana [24].

4. Programvaror

Nedan presenteras en översikt över programvaror och verktyg som är tillämpliga för att genomföra SCED-metaanalys, inklusive deras styrkor, begränsningar och referenser.

R-paketet [18]

- lme4: Ett centralt paket för linjära blandade modeller [25]. Möjliggör modellering av mätningar inom fall och fall inom studier. Ger maximum likelihood-skattningar av fasta och slumpmässiga effekter. Rekommenderas för SCED-data.

- metafor: Används för två- och flernivåmetaanalyser, inklusive slumpmässiga effekter på flera nivåer [18]. Kan kombineras med clubSandwich för att hantera beroende effektstorlekar.
- brms: Ett Bayesianskt alternativ för hierarkiska modeller [20]. Stöder komplexa modellstrukturer och kan använda informativa prior-fördelningar. Bra alternativ vid små stickprov.
- clubSandwich: Paket för robust variansestimering (RVE) med småstickprovsjusteringar [26]. Används ofta tillsammans med metafor eller lme4.
- robumeta: Implementerar robust variansestimering enligt Hedges, Tipton och Johnson [16]. Mindre flexibel än clubSandwich men enklare att använda för en- och tvånivåmodeller.
- SCMA: Specialpaket för SCED-meta. Möjliggör beräkning och metaanalys av PND, PEM, Tau-U och standardiserade effektstorlekar [27].
- MultiSCED: Shiny-baserat verktyg som möjliggör flernivåmodellering av SCED-data via webbläsare. Använder lme4 i bakgrunden men kräver ingen programmering [28].

Statakommandon

- **mixed:** Motsvarighet till lme4 i R. Används för att modellera flernivåstrukturer, exempelvis mätningar inom individer och individer inom studier [29].
- **reghdfe:** Möjliggör absorption av högdimensionella fasta effekter. Användbart när man vill kontrollera för studie- och individnivå som fasta effekter [30].
- **robu (robmeta.ado):** Ett komplement i Stata för att genomföra robusta meta-analyser med beroende effektstorlekar. Implementerar RVE enligt Hedges m.fl. [31] [10].

Pythonbibliotek

- **statsmodels:** Stöder blandade modeller via MixedLM. Mindre specialiserat än R men användbart för att modellera aggregerade SCED-data [32].
- **pymer4:** Python-gränssnitt till lme4 i R. Möjliggör att köra hierarkiska modeller via R från Pythonmiljön. Används ofta inom neurovetenskap [33].

SPSS

- **Mixed Models-menyn och GENLINMIXED:** Möjliggör modellering av hierarkiska data, men saknar specialstöd för metaanalys av beroende effektstorlekar. SCED-analys kan genomföras via makron (t.ex. DHPS) [34].

SAS

- **PROC MIXED och PROC GLIMMIX:** Kraftfulla procedurer för multilevelmodellering. Stöd för Kenward-Roger-korrektioner och andra metoder lämpliga för småstickprovsanalyser. Saknar dock dedikerat metaanalysmodul (SAS Institute Inc. (2014). SAS/STAT® 13.2 User's Guide: The MIXED Procedure).

RevMan

- **Cochrane Collaboration's programvara för metaanalys.** Begränsat användbar för SCED, eftersom den inte stödjer hierarkiska modeller eller robusta SE. Kan endast användas om SCED-data är aggregerade på studienivå [35].

ESCI (Effect Size Calculator Interface)

- **ESCI är ett verktyg som erbjuds av Campbell Collaboration och används för att beräkna och rapportera effektstorlekar inom ramen för single case design.** Det är särskilt inriktat mot standardiserade effektmått som Tau-U och andra mått för icke-överlappande data.

Måluppfyllelse

Detta mått speglar förändringen i förhållande till ett uppsatt mål

- **Percent of Goal Obtained (PoGO) [36]:** Speglar måluppfyllelse. PoGO = 100 → mål uppnått, >100 → överskridit, <0 → negativ effekt.

Överlappning mellan baslinje och insats

För en bred översikt av overlap-baserade effektmått i single-case-forskning se [37].

- **Improvement Rate Difference (IRD) [38] [39] [40]:** Skillnad i förbättringsfrekvens mellan baslinje och insats. IRD=andel datapunkter i insatsfasen som visar förbättring jämfört med baseline - andel datapunkter i baslinjefasen som visar förbättring.

- Nonoverlap of All Pairs (NAP) [8] [41]: Mäter andel datapunktspar som inte överlappar mellan faser. Indikerar hur mycket de två faserna skiljer sig åt, men tar inte hänsyn till riktningen av förändringen.
- Percentage of Non-Overlapping Data (PND) [38] [42]: Mäter andel datapunkter i insatsfasen som överstiger den högsta datapunkten i baslinjen. $PND \geq 90\%$ anses som stark effekt, 70–89% som måttlig, osv. [43]
- Percentage of Data Points Exceeding the Median Difference (PND-Median/PEM) [38] [17] [43]: Mäter andelen datapunkter i insatsfas som överstiger baslinje-median
- Percentage of All Non-Overlapping Data (PAND) [38] [44]: En förbättrad version av PND där alla datapunkter används, inte bara den högsta från baseline.

Trendjusterande mått

- Tau (AB): Mellanfasmått. Rangkorrelationsmått mellan baslinje (fas A) och insats (fas B) som visar andelen förbättrade parvisa jämförelser mellan faserna. Värdet varierar från -1 till 1; ett positivt Tau (AB) indikerar förbättring under insats, ett negativt indikerar försämring, och 0 innebär ingen nettoeffekt [45].
- Tau (U): Inomfasmått. Rangkorrelation för monoton trend inom en enskild fas, oftast använd för att mäta och justera oönskad trend i baslinjen (fas A) innan insats. Värdet varierar från -1 till 1; ett positivt Tau (U) betyder att det finns en uppåtgående trend (förbättring) inom fasen, ett negativt betyder en nedåtgående trend (försämring), medan 0 indikerar ingen trend [46].
- Tau-U: Kombination av mellanfas- och inomfasmått. Ett effektstorleksmått som kombinerar förbättringen mellan fas A och B (Tau AB) med eventuell trend under insatsens fas, med möjlighet att korrigera för oönskad baslinjetrend. Tau-U varierar från -1 till 1; positiva värden tyder på en total förbättring av beteendet efter justering för baslinjetrend, negativa värden tyder på total försämring, och 0 innebär ingen nettovariation [40] [46] [45].

Förändring i medelvärden

Absolut och standardiserad medelvärdesskillnad

- Log Response Ratio (LRR) [47]: Log-transformerad proportionell förändring i medelvärde. Positiv = ökning, negativ = minskning, 0 = ingen förändring. Om baslinjedata är nära noll kan logaritmttransformation ge missvisande värden
- Mean Difference (MD): Medelvärdesskillnader mellan faser.
- Standardised Mean Difference (SMD) [48]: Medelvärdesskillnader mellan faser justerad för standardavvikelse. Hedges' g, en korrigerad version av Cohen's d, anpassad för små urval [49].

Relativ förändring i medelvärde uttryckt i procent

- Percent Change: Procentuell förändring i medelvärde mellan faserna.

$$(M_{\text{Insats}} - M_{\text{Baselinje}}) / M_{\text{Baseline}} \times 100$$
Mått som används i SCED primärstudier där medelvärden rapporteras [50] [51] [52].

5. Antal studier

Det finns metodlitteratur som berör hur omfattningen av dataunderlag i form av antal SCED primärstudier, antal individer samt antal mätpunkter påverkar vilken typ av metaanalys som kan genomföras och med vilken tillförlitlighet. I tabell 1 presenteras exempel på omfattning av dataunderlag i förhållande till datatyp. För beskrivningar och rekommendationer hänvisas direkt till primärkällorna [53] [17] [54] [55] samt [56].

Tabell 1 Beskrivningar av antal SCED primärstudier, antal individer samt antal mätpunkter som utgör en form av tröskelvärden för olika metaanalyser.

Datatyp	Antal studier	Referens
Summerade effektmått	<10 studier	[54]
	≥10 studier	[54]
	≥20–30 studier	[54] [55]
Fassammanfattad data	<10 studier / <30 individer	[53] [17]
	<10 studier / ≥ 30 individer	[53]
	≥10 studier / <30 individer	[53]
	≥10 studier / ≥ 30 individer	[53] [17]
Rådata på individnivå	<30 studier / <20 mätpunkter	[53] [56]
	≥30 studier / <20 mätpunkter	[53] [17]
	<30 studier / ≥ 20 mätpunkter	[53]
	≥30 studier / ≥ 20 mätpunkter/individer	[53] [56]

6. Heterogenitet

Heterogenitet är ett centralt begrepp i SCED-metaanalys. Insatser och utfall är ofta individuellt anpassade och varierar därmed mellan studier. Heterogenitet avser variationen i effektstorlekar mellan studier (och mellan individer inom studier) utöver ren slumpvariation. Man skiljer ofta mellan klinisk heterogenitet (skillnader i population, insats, miljö etcetera) och statistisk heterogenitet (observerad variation i effektstorlekar mellan enheter). Statistisk heterogenitet manifesteras när effektstorlekar varierar mer än vad som kan förklaras av slump.

6.1. Heterogenitet i SCED-metaanalys

SCED-metaanalyser har en inbyggd komplexitet då data är hierarkiskt strukturerade: upprepade mätningar är inom varje enskild deltagare, och deltagare är inom studier [12]. Variation kan uppstå både på individnivå och på studienivå (skillnader i protokoll, utfallsmått, miljö etcetera). Om variationen ligger främst på individnivå kan det tyda på att deltagarnas svar på insatsen skiljer sig trots liknande betingelser, vilket i sin tur kan påverka den uppskattade genomsnittseffekten och dess tolkbarhet. När man kombinerar många SCED primärstudier får man ofta en tre-nivå struktur: nivå 1 = mätningar, nivå 2 = deltagare (fall), nivå 3 = studie [12]. Detta innebär att heterogenitet kan analyseras på flera nivåer – både inom en studie (mellan deltagare) och mellan olika studier.

6.1.1. Identifiering av heterogenitet

Nedan presenteras olika sätt att identifiera heterogenitet.

Statistiska mått

Statistiska mått kan användas för att bedöma statistisk heterogenitet:

- I^2 : Andelen av den totala variationen i observerade effektstorlekar som beror på faktisk variation mellan studier snarare än slump [57].
- τ^2 : Skattad varians mellan studiers sanna effektstorlekar.
- Cochran's Q: Testar hypotesen att alla studier delar en gemensam effektstorlek. Ett signifikant Q indikerar närvaro av statistisk heterogenitet.

Dessa mått har dock begränsad tillförlitlighet vid små sampelstorlekar, vilket ofta är fallet i SCED-metaanalyser [10] [12]. Vid binära utfall kan BESD (Binomial Effect Size Display) eller liknande metoder användas för att undersöka heterogenitet [37] [58]. Skillnader i binomial-fördelningarna mellan studierna kan indikera underliggande variation.

Visualiseringsverktyg

- Skogsdiagram (forest plot): Visar individuella studiers effektstorlekar och deras konfidensintervall, vilket ger en visuell indikation på heterogenitet.
- Trattdiagram (funnel plot): Används främst för att undersöka publikationsbias, men asymmetri kan också indikera heterogenitet eller systematiska avvikelser [59]

6.1.2. Hantering av heterogenitet

Heterogenitet kan hanteras på olika sätt, nedan följer en beskrivning av några olika metoder som kan användas.

Multilevel-modellering

Multilevel meta-analys (hierarkisk linjär modellering, HLM) är den metod som rekommenderas för att hantera heterogenitet i SCED-synteser [12] [60]. En tre-nivåmodell möjliggör att man explicit modellerar variation på både deltagarnivå (nivå 2) och studienivå (nivå 3), genom att inkludera slumpmässiga effekter på båda nivåerna. Om till exempel den största variationen ligger på studienivå, tyder det på att kontextuella olikheter är viktiga. Om variationen ligger främst på individnivå kan det tyda på att deltagarnas svar på insatsen skiljer sig trots liknande betingelser.

Moderatoranalys

För att förklara heterogenitet är moderatoranalyser centrala. Här inkluderas förklarande variabler som kan påverka effektstorleken. I SCED-metaanalyser kan moderatorer placeras på flera nivåer i en multilevel-modell – t.ex. insatsskaraktäristika på studienivå eller deltagaregenskaper på individnivå [60] [61]. Moeyaert et al. har visat att multilevel-modeller med moderatorer effektivt kan förklara variabilitet på både fall- och studienivå [60] [61]. Meta-regression eller regressionskomponenter i en HLM används för detta syfte. Exempel på moderatorer är typ av design (AB vs. ABAB), antal mätpunkter, längd på insats, deltagarens ålder, med mera. Om heterogeniteten kvarstår efter att rimliga moderatorer beaktats, bör man tolka den sammanvägda effekten med försiktighet då den kan dölja viktiga variationer i kontext.

Subgruppsanalys

Subgruppsanalyser kan användas för att undersöka potentiella källor till heterogenitet, till exempel om olika designfaktorer (antal mätpunkter, baslinjens längd etcetera) påverkar resultaten. Till skillnad från kontinuerliga moderatorer delas här studier upp i kategorier (till exempel kort vs. lång insats) och analyseras separat.

7. Publikationsbias

Publikationsbias uppstår när studier med signifikanta eller positiva resultat har högre sannolikhet att publiceras än studier med noll- eller negativa fynd. Detta kan leda till en överskattning av insatseffekters storlek i en meta-analys.

7.1 Publikationsbias i SCED primärstudier

I SCED primärstudier, där antalet deltagare (N) ofta är litet och visuella analysmetoder är vanliga, kan icke-signifikanta utfall ibland förbli opublicerade. Detta fenomen kan också förstärkas av de mer varierande och komplexa design- och analysmetoder som används i SCED jämfört med gruppstudier. Eftersom SCED-primärstudier ofta rapporterar enskilda fall och individuella datapunkter, kan variationen i rapportering och analys påverka förekomsten av publikationsbias. Dessutom kan publikationsbias få större konsekvenser i SCED-metaanalyser eftersom varje enskilt fall har en betydande inverkan på det sammanlagda resultatet.

Det är viktigt att både identifiera och hantera publikationsbias för att säkerställa tillförlitliga slutsatser.

7.1.1. Identifiering av publikationsbias

Trattdiagram (Funnel Plots)

Det vanligaste sättet att undersöka publikationsbias är genom ett trattdiagram där varje studie, fall eller effektstorlek plottas mot dess standardfel (eller N). Om trattdiagrammet är symmetriskt (små studier är jämnt spridda kring noll och effekt) indikerar det frånvaro av stark bias. Om diagrammet däremot är asymmetriskt (till exempel saknas små studier med negativa eller obetydliga effekter) kan detta tyda på publikationsbias [59].

För att statistiskt testa asymmetri i trattdiagram används Egger's regressions test som undersöker om det finns en linjär relation mellan effektstorlek och standardfel [59].

Identifiering av opublicerade studier

Granskning av primärstudiernas protokoll kan ge insikter om eventuell selektiv rapportering. Detta kan inkludera jämförelse mellan förregistrerade utfallsmått och de som faktiskt rapporterats i den slutliga studien.

7.1.2. Hantering av publikationsbias

Förebyggande åtgärder mot publikationsbias

Den bästa strategin är att proaktivt minska publikationsbias vid litteratursökningen. Detta inkluderar:

- Utförliga sökningar som inkluderar grå litteratur (till exempel opublicerade avhandlingar, konferensposter, rapporter).
- Kontakt med experter inom området för att hitta opublicerade data.
- Inkludering av både publicerade och opublicerade SCED primärstudier för att skapa en mer balanserad bild [56].

Statistiska korrigeringar

Genom att tillämpa statistiska korrigeringar kan man delvis justera för publikationsbias och därigenom öka tillförlitligheten i resultaten. Ett exempel på en sådan korrigering är Trim and Fill-metoden, som justerar snedfördelade data genom att "trimma" bort de mest extrema studierna och fylla i med uppskattade studier för att skapa symmetri [59]. Denna metod kräver dock ett relativt stort antal studier för att vara robust.

Subgruppsanalys

Subgruppsanalyser kan användas för att jämföra publicerade och opublicerade studier. Om det finns en systematisk skillnad i effektstorlek mellan dessa grupper kan detta vara en indikator på publikationsbias [59].

8. Stöd vid val av metaanalysstrategi

I tabellerna 2–4 sammanfattas centrala komponenter i metaanalys baserad på data från SCED primärstudier. De presenteras där per datatyp och förklaringar till förkortningar hittar du under respektive tabell. Tabeller är tänka att användas som stöd vid val av metaanalysstrategi. För kritiska beslut bör man konsultera primärkällorna direkt.

Tabell 2 Beskrivning av centrala komponenter i metaanalys baserad på summerade effektstorlekar från SCED primärstudier.

Datatillgång		Summerade effektstorlekar		
Hierarkisk struktur		Unilevel/1-nivå		
Datastruktur	Univariat	Uni-/Multivariat	Univariat	
Antal primärstudier	>20–30 studier		>10 studier	<10 studier
Metaanalys-metod	SMA [3]	ADMA [13]	SMA [3]	BMA [5] [21]

Estimerings- metod	GLS, FEM/REM	GLS, FEM/REM, Meta-regression, RVE [10] [11]	GLS, FEM/REM	Bayesian MCMC
Inferens- metod	Viktat medelvärde, KI, p-värde	Meta-regression (med/utan RVE), KI, p-värde	Viktat medelvärde, KI, p-värde	Posteriorfördelning, credibility intervals
Effektmått	Percent change, LRR, MD, SMD			

Förklaring av förkortningar i tabell 2. SMA=Summary meta-analysis, ADMA= Aggregate data meta-analysis, BMA= Bayesian meta-analysis. GLS= Generalized least squares, FEM=Fixed effect Model, REM=Random effect Model, RVE=Robust variance estimation, MCMC=Markov Chain Monte Carlo, KI=Konfidensintervall, LRR=Log response ratio, MD=Mean difference, SMD=Standardized mean difference

Tabell 3 Beskrivning av centrala komponenter i metaanalys baserad på fassammanfattade rådata från SCED primärstudier.

Datatillgång		Fassammanfattade rådata	
Hierarkisk struktur		Multilevel/2-nivå	
Datastruktur		Multivariat	Univariat
Antal primärstudier		>10 studier/> 30 individer	<10 studier/ <30 individer
Metaanalys-metod	BMA [5] [21]	MLMA [2] [12]	BMA [5] [21]
Estimerings-metod	Bayesian MCMC	ML, REML, RVE, GLS	Bayesian MCMC
Inferens-metod	Posteriorfördelning, credibility intervals	KI, I^2 , τ^2 , modellering av flera nivåer (med/utan RVE)	Posteriorfördelning, credibility intervals
Effektmått		Percent change, LRR, MD, SMD	
Programvara	Stan, JAGS, R (brms)	R (lme4, metafor, clubSandwich), Stata (mixed), SPSS	Stan, JAGS, R (brms)

Förklaring av förkortningar i tabell 3. BMA= Bayesian meta-analysis, MLMA=Multi-level meta-analysis, MCMC=Markov chain Monte Carlo, ML=Maximum likelihood, REML= Restricted maximum likelihood, RVE=Robust variance estimation, GLS= Generalized least squares, KI=Konfidensintervall, LRR=Log response ratio, MD=Mean difference, SMD=Standardized mean difference

Tabell 4 Beskrivning av centrala komponenter i metaanalys baserad på fullständiga rådata från SCED primärstudier.

Datatillgång		Fullständiga rådata			
Hierarkisk struktur		Multilevel/2–3 nivå			
Datastruktur		Univariat		Multivariat	
Antal primärstudier		>30 studier/>20 mätpunkter		<30 studier/ <20 mätpunkter	
Metaanalys-metod	TSMA [2]	MLMA [2] [12]	MMA [8]	IPD-MA [3]	BMA [5] [21]
Estimerings-metod	GLS med AR (1)	ML, REML, RVE, GLS	ML, REML, RVE	GLS med AR (1), Bayesiansk MCMC, GLMM	Bayesian MCMC
Inferens-metod	Trendanalys, nivå- och slopeskillnader, KI	KI, I^2 , τ^2 , modellering av flera nivåer	Gemensam inferens, KI, modellering av flera utfall	Regression, posteriorer, KI	Posteriorfördelning, credibility intervals
Effektmått		Percent change, LRR, MD, SMD, PoGo, IRD, NAP, PND, PEM, PAND, Tau-(AB), TAU (U), Tau-U		Percent change, LRR, MD, SMD,	

Förklaring av förkortningar i tabell 4. TSMA= Time-series meta-analysis, MLMA=Multi-level meta-analysis, MMA= Multivariate meta-analysis, IPD-MA= Individual participant data meta-analysis, BMA= Bayesian meta-analysis, KI=Konfidensintervall, LRR=Log response ratio, MD=Mean difference, SMD=Standardized mean difference, PoGo=Percent of Goal Obtained, IRD=Improvement Rate Difference, NAP=Nonoverlap of All Pairs, PAND=Percentage of All Non-Overlapping Data, PEM=Percentage of Data Points Exceeding the Median Difference, PAND=Percentage of Non-Overlapping Data

9. Rapporteringsstandarder

PRISMA 2020 är tillämplig för att rapportera resultat från metaanalyser baserad på data från SCED primärstudier. Se [PRISMA Statement](#).

10. Risk för bias-granskning

För att granska risk för bias i studier som gjort metaanalys baserat på data från SCED primärstudier kan ROBIS [62], med avsteg från kriteriet: 4.4 skillnaden mellan studierna (heterogenitet) användas. Även AMSTAR 2 [63] (Assessment of Multiple Systematic Reviews) är tillämplig.

11. Medverkande

Susanna Larsson Tholén, projektledare SBU, rapportens författare

Fredrik Tholander, projektledare SBU

Maria Ahlberg, projektadministratör

Anna Attergren Granath, projektadministratör

Sakgranskning: SBU:s metodstöd samt SBU:s metodgrupp

12. Referenser

1. Larsson Tholén S, Dagerhamn, J., Silverstein, R.;. Arbetsrapport: Single Case Experimental Design. Metodik, analys, risk för bias-bedömning, överförbarhet och rapporteringsstandarder. <https://www.sbu.se/374>; 2025.
2. Van den Noortgate W, Lopez-Lopez JA, Marin-Martinez F, Sanchez-Meca J. Three-level meta-analysis of dependent effect sizes. Behav Res Methods. 2013;45(2):576-94. Available from: <https://doi.org/10.3758/s13428-012-0261-6>
3. Shadish WR, Rindskopf DM. Methods for evidence-based practice: Quantitative synthesis of single-subject designs. New Directions for Evaluation. 2007;2007(113):95-109. Available from: <https://doi.org/10.1002/ev.217>
4. Polanin JR, Hennessy EA, Tanner-Smith EE. A Review of Meta-Analysis Packages in R. Journal of Educational and Behavioral Statistics. 2017;42(2):206-42.
5. Natesan P. Fitting Bayesian Models for Single-Case Experimental Designs: A Tutorial. Methodology. 2019;15:147-56. Available from: <https://doi.org/10.1027/1614-2241/a000180>
6. Swaminathan H, Rogers HJ, Horner RH. An effect size measure and Bayesian analysis of single-case designs. J Sch Psychol. 2014;52(2):213-30. Available from: <https://doi.org/10.1016/j.jsp.2013.12.002>
7. Rindskopf D. Nonlinear Bayesian analysis for single case designs. J Sch Psychol. 2014;52(2):179-89. Available from: <https://doi.org/10.1016/j.jsp.2013.12.003>
8. Jackson D, Riley R, White IR. Multivariate meta-analysis: potential and promise. Stat Med. 2011;30(20):2481-98. Available from: <https://doi.org/10.1002/sim.4172>
9. Van den Noortgate W, Onghena P. Hierarchical linear models for the quantitative integration of effect sizes in single-case research. Behav Res Methods Instrum Comput. 2003;35(1):1-10. Available from: <https://doi.org/10.3758/bf03195492>
10. Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. Res Synth Methods. 2010;1(1):39-65. Available from: <https://doi.org/10.1002/jrsm.5>
11. Tipton E. Small sample adjustments for robust variance estimation with meta-regression. Psychol Methods. 2015;20(3):375-93. Available from: <https://doi.org/10.1037/met0000011>
12. Moeyaert M, Dehghan-Chaleshtori M, Xu X, Yang P. Single-case design meta-analyses in education and psychology: a systematic review of methodology. Front Res Metr Anal. 2023;8:1190362. Available from: <https://doi.org/10.3389/frma.2023.1190362>
13. Moeyaert M, Ugille M, Natasha Beretvas S, Ferron J, Bunuan R, Van den Noortgate W. Methods for dealing with multiple outcomes in meta-analysis: a comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. International Journal of Social Research Methodology. 2016;20(6):559-72. Available from: <https://doi.org/10.1080/13645579.2016.1252189>
14. Cheung MW. A Guide to Conducting a Meta-Analysis with Non-Independent Effect Sizes. Neuropsychol Rev. 2019;29(4):387-96. Available from: <https://doi.org/10.1007/s11065-019-09415-6>
15. Chen M, Pustejovsky JE. Multilevel meta-analysis of single-case experimental designs using robust variance estimation. Psychol Methods. 2024;29(3):537-60. Available from: <https://doi.org/10.1037/met0000510>
16. Fisher Z, Tipton E. robumeta: An R-package for robust variance estimation in meta-analysis. 2015.
17. Pustejovsky JE, Tipton E. Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. Prev Sci. 2022;23(3):425-38. Available from: <https://doi.org/10.1007/s11121-021-01246-3>

18. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*. 2010;36(3):1 - 48. Available from: <https://doi.org/10.18637/jss.v036.i03>
19. Rindskopf D. Bayesian analysis of data from single case designs. *Neuropsychol Rehabil*. 2014;24(3-4):572-89. Available from: <https://doi.org/10.1080/09602011.2013.866903>
20. Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*. 2017;80(1):1 - 28. Available from: <https://doi.org/10.18637/jss.v080.i01>
21. van de Schoot R, Depaoli S, King R, Kramer B, Märtens K, Tadesse MG, et al. Bayesian statistics and modelling. *Nature Reviews Methods Primers*. 2021;1(1):1. Available from: <https://doi.org/10.1038/s43586-020-00001-2>
22. Dettori JR, Norvell DC, Chapman JR. Fixed-Effect vs Random-Effects Models for Meta-Analysis: 3 Points to Consider. *Global Spine J*. 2022;12(7):1624-6. Available from: <https://doi.org/10.1177/21925682221110527>
23. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods*. 2010;1(2):97-111. Available from: <https://doi.org/10.1002/jrsm.12>
24. Manolov R, Moeyaert M, Fingerhut JE. A Priori Justification for Effect Measures in Single-Case Experimental Designs. *Perspect Behav Sci*. 2022;45(1):153-86. Available from: <https://doi.org/10.1007/s40614-021-00282-2>
25. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*. 2015;67(1):1 - 48. Available from: <https://doi.org/10.18637/jss.v067.i01>
26. Pustejovsky JE, and Tipton E. Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models. *Journal of Business & Economic Statistics*. 2018;36(4):672-83. Available from: <https://doi.org/10.1080/07350015.2016.1247004>
27. Bulté I, Onghena P. The Single-Case Data Analysis Package: Analysing Single-Case Experiments with R Software. *Journal of Modern Applied Statistical Methods*. 2013;12(2):450-78. Available from: <https://doi.org/10.22237/jmasm/1383280020>
28. Declercq L, Cools W, Beretvas SN, Moeyaert M, Ferron JM, Van den Noortgate W. MultiSCED: A tool for (meta-)analyzing single-case experimental data with multilevel modeling. *Behav Res Methods*. 2020;52(1):177-92. Available from: <https://doi.org/10.3758/s13428-019-01216-2>
29. StataCorp. Multilevel Mixed-Effects Reference Manual Stata 18. College Station, TX: Stata Press.; 2023.
30. Correia S. Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator. Working Paper. 2016.
31. Hedberg E. ROBUMETA: Stata module to perform robust variance estimation in meta-regression with dependent effect size estimates. 2011.
32. Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*. 2010;2010.
33. Jolly. Pymer4: Connecting R and Python for Linear Mixed Modeling. *Journal of Open Source Software*. 2018;3(31):862.
34. Corp. I. IBM SPSS Statistics for Windows, Version 28.0. . Armonk, NY: IBM Corp; 2021.
35. The Cochrane Collaboration. Review Manager (RevMan). Version 5.4.1.,. 2020.
36. Ferron J, Goldstein H, Olszewski A, Rohrer L. Indexing effects in single-case experimental designs by estimating the percent of goal obtained. *Evidence-Based Communication Assessment and Intervention*. 2020;14(1-2):6-27. Available from: <https://doi.org/10.1080/17489539.2020.1732024>

37. Parker RI, Vannest KJ, Davis JL. Effect size in single-case research: a review of nine nonoverlap techniques. *Behav Modif.* 2011;35(4):303-22. Available from: <https://doi.org/10.1177/0145445511399147>
38. Byiers BJ, Reichle J, Symons FJ. Single-subject experimental design for evidence-based practice. *Am J Speech Lang Pathol.* 2012;21(4):397-414. Available from: [https://doi.org/10.1044/1058-0360\(2012/11-0036\)](https://doi.org/10.1044/1058-0360(2012/11-0036))
39. Parker RI, Vannest KJ, Brown L. The improvement rate difference for single-case research. *Exceptional Children.* 2009;75(2):135-50.
40. Dowdy A, Peltier C, Tincani M, Schneider WJ, Hantula DA, Travers JC. Meta-analyses and effect sizes in applied behavior analysis: A review and discussion. *J Appl Behav Anal.* 2021;54(4):1317-40. Available from: <https://doi.org/10.1002/jaba.862>
41. Parker RI, Vannest K. An improved effect size for single-case research: nonoverlap of all pairs. *Behav Ther.* 2009;40(4):357-67. Available from: <https://doi.org/10.1016/j.beth.2008.10.006>
42. Scruggs TE, Mastropieri MA, Casto G. The quantitative synthesis of single-subject research: Methodology and validation. *RASE: Remedial & Special Education.* 1987;8(2):24-33. Available from: <https://doi.org/10.1177/074193258700800206>
43. Carlin MT, Costello MS. Statistical Decision-Making Accuracies for Some Overlap- and Distance-based Measures for Single-Case Experimental Designs. *Perspect Behav Sci.* 2022;45(1):187-207. Available from: <https://doi.org/10.1007/s40614-021-00317-8>
44. Parker RI, Hagan-Burke S, Vannest K. Percentage of all non-overlapping data (PAND) an alternative to PND. *The Journal of Special Education.* 2007;40(4):194-204.
45. Parker RI, Vannest KJ, Davis JL, Sauber SB. Combining nonoverlap and trend for single-case research: Tau-U. *Behav Ther.* 2011;42(2):284-99. Available from: <https://doi.org/10.1016/j.beth.2010.08.006>
46. Lee JB, Cherney LR. Tau-U: A Quantitative Approach for Analysis of Single-Case Experimental Data in Aphasia. *Am J Speech Lang Pathol.* 2018;27(1s):495-503. Available from: https://doi.org/10.1044/2017_ajslp-16-0197
47. Pustejovsky JE. Using response ratios for meta-analyzing single-case designs with behavioral outcomes. *Journal of School Psychology.* 2018;68:99-112. Available from: <https://doi.org/https://doi.org/10.1016/j.jsp.2018.02.003>
48. Shadish WR, Rindskopf DM, Hedges LV. The state of the science in the meta-analysis of single-case experimental designs. *Evidence-Based Communication Assessment and Intervention.* 2008;2(3):188-96. Available from: <https://doi.org/10.1080/17489530802581603>
49. Hedges LV, Pustejovsky JE, Shadish WR. A standardized mean difference effect size for multiple baseline designs across individuals. *Res Synth Methods.* 2013;4(4):324-41. Available from: <https://doi.org/10.1002/jrsm.1086>
50. Kazdin AE. *Single-case research designs: Methods for clinical and applied settings*, 2nd ed. New York, NY, US: Oxford University Press; 2011.
51. Smith JD. Single-case experimental designs: a systematic review of published research and current standards. *Psychol Methods.* 2012;17(4):510-50. Available from: <https://doi.org/10.1037/a0029312>
52. Manolov R, Moeyaert M. Recommendations for Choosing Single-Case Data Analytical Techniques. *Behav Ther.* 2017;48(1):97-114. Available from: <https://doi.org/10.1016/j.beth.2016.04.008>
53. Moeyaert M, Ugille M, Ferron JM, Onghena P, Heyvaert M, Beretvas SN, et al. Estimating intervention effects across different types of single-subject experimental designs: Empirical illustration. US: Educational Publishing Foundation; 2015. p. 50-63.
54. Borenstein M, Hedges, L., Higgins, J., & Rothstein, H. *Introduction to Meta-Analysis.* <https://www.perlego.com/book/2762574/introduction-to-metaanalysis-pdf>; Wiley; 2021.

55. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-58. Available from: <https://doi.org/10.1002/sim.1186>
56. Fernandez-Castilla B, Jamshidi L, Declercq L, Beretvas SN, Onghena P, Van den Noortgate W. The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behav Res Methods*. 2020;52(5):2031-52. Available from: <https://doi.org/10.3758/s13428-020-01373-9>
57. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557-60. Available from: <https://doi.org/10.1136/bmj.327.7414.557>
58. Chen M, Pustejovsky JE, Klingbeil DA, Van Norman ER. Between-case standardized mean differences: Flexible methods for single-case designs. *J Sch Psychol*. 2023;98:16-38. Available from: <https://doi.org/10.1016/j.jsp.2023.02.002>
59. Dowdy A, Hantula DA, Travers JC, Tincani M. Meta-Analytic Methods to Detect Publication Bias in Behavior Science Research. *Perspect Behav Sci*. 2022;45(1):37-52. Available from: <https://doi.org/10.1007/s40614-021-00303-0>
60. Moeyaert M, Akhmedjanova D, Ferron J, Beretvas SN, Noortgate Wvd. Effect size estimation for combined single-case experimental designs. *Evidence-Based Communication Assessment and Intervention*. 2020;14(1-2):28-51. Available from: <https://doi.org/10.1080/17489539.2020.1747146>
61. Moeyaert M, Yang P. Assessing generalizability and variability of single-case design effect sizes using two-stage multilevel modeling including moderators. *Behaviormetrika*. 2021;48(2):207-29. Available from: <https://doi.org/10.1007/s41237-021-00141-z>
62. Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016;69:225-34. Available from: <https://doi.org/10.1016/j.jclinepi.2015.06.005>
63. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008. Available from: <https://doi.org/10.1136/bmj.j4008>