

Eksamen IDR4000

Kandidatnummer

12/2/22

Table of contents

Preface	3
Deloppgave 1: Beskrivende statistikk, reliabilitet og validitet, verktøy for reproduserbar dataanalyse	4
Deloppgave 2: Laborasjonsrapport	5
Deloppgave 3: Vitenskapsfilosofi	6
Falsifikasjonisme	6
Hypotetisk deduktiv metode og abduksjon	7
Replikasjonskrisen	8
Deloppgave 4: Studiedesign	10
Introduksjon	10
Metode	11
Diskusjon	12
Konklusjon	13
Deloppgave 5: Analysere repeterte målinger	15
References	16

Preface

This is a Quarto book.

To learn more about Quarto books visit <https://quarto.org/docs/books>.

```
1 + 1
```

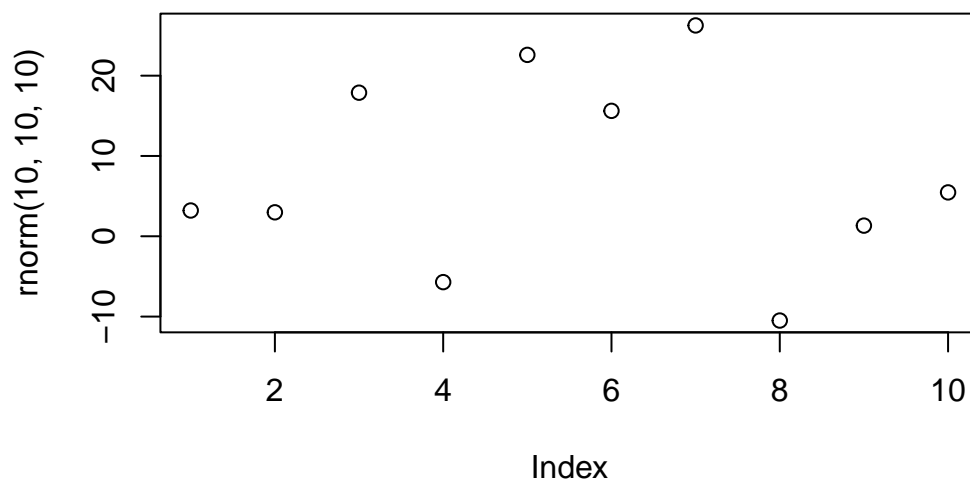
```
[1] 2
```

Forord:

Dette er min innlevering til eksamen, data og kode finnes her....

Deloppgave 1: Beskrivende statistikk, reliabilitet og validitet, verktøy for reproduserbar dataanalyse

```
plot(rnorm(10, 10, 10))
```



Figur 1: Eksempel

Deloppgave 2: Laborasjonsrapport

Deloppgave 3: Vitenskapsfilosofi

Falsifikasjonisme

Hva er Poppers falsifiserbarhetskriterium og hvilket spørsmål skal dette kriterium gi svar på? Hvorfor mener andre vitenskapsfilosofer (f.eks. Okasha) at vi ikke trenger å svare på dette spørsmålet? Hvem synes dere har rett?

Poppers falsifiserbarhetskriterium er basert på at en teori skal være falsifiserbar, men ikke falsifisert. Popper mente at dersom en teori skulle være vitenskapelig, måtte den også være mulig å avkrefte. Videre mente han at evidens som bekrefter en teori ikke eksisterer. Teorier som ble tilpasset data og som dermed ikke kunne motsies kalte Popper for «pseudo-science» («ikke-vitenskapelig vitenskap»). Popper kritiserte Karl Marx og Sigmund Freud sine teorier for å være «ikke-vitenskapelige». Karl Marx sine teorier ble kritisert av Popper for å tilpasses underveis. Dersom teoriene møtte motstand igjennom en test eller gjennom data så kunne teorien endres for å tilpasse seg motstanden i dataen. På denne måten kunne Marx fremdeles beholde teorien sin. Popper mente at Freuds teorier ikke var mulig å teste eller falsifisere og beskrev de dermed som myter og ikke som vitenskap. Poppers falsifiserbarhetskriterium skulle derfor gi svar på om en teori er vitenskapelig eller ikke. Utfordringen med å skille mellom vitenskapelige og «ikke-vitenskapelige» teorier kalte Popper for demarkasjonsproblemet (Okasha 2016; Popper 1969).

Flere filosofer har kritisert Poppers falsifiserbarhetskriterium for å være altfor forenklet. En av de er Okasha (2016) som skriver i *Philosophy of Science: A Very Short Introduction* at det er flere eksempler på det som Popper kalte for «pseudo-science» har ført til flere viktige vitenskapelige funn. For eksempel ved oppdagelsen av planeten Neptun. To forskere på midten av 1800-tallet John Couch Adams fra England og Urbain Leverrier fra Frankrike oppdaget planeten Neptun ved å gjøre nettopp det Popper kritiserte Marxister for å gjøre. Adams og Leverrier forsket videre til tross for at Newtons gravitasjonsteori var feil i spådommen om Uranus sin bane i verdensrommet. I likhet med Marx og Freud tilpasset Adams og Leverrier sin teori og konkluderte med at det måtte finnes en ny planet som ikke var enda oppdaget. Neptun ble like etter oppdaget som bekreftet Newtons gravitasjonsteori nok en gang. (Okasha 2016). Okasha mente med blant annet dette eksempelet at Poppers forsøk på å avgrense vitenskap fra «pseudo-vitenskap» ikke kan stemme. Jeg mener selv at Poppers falsifiserbarhetskriterium har hatt en stor verdi for vitenskapsteori. Det utfordrer til kritisk tenkning og filosofering med en enkel og forståelig metode. Videre mener jeg at Okasha har rett i at Poppers falsifiserbarhetskriterium blir for enkelt og at det i tilfeller mener jeg at det

blir riktig og si at en teori alltid er feil eller gal (Okasha 2016). Det finnes en slags objektivitet ved at en teori skal kunne falsifiseres, men at den viktigste forskjellen ligger i hvor godt eller dårlig en teori er bekreftet. Det vil ikke være riktig å avskrive enhver teori som møter litt motstand, men jeg mener det bør være en viss balanse mellom det å avkrefte og bekrefte teorier. Dette ville også avskrevet flere viktige vitenskapelige funn. Spørsmålet om vi skal benytte oss av Poppers falsifiserbarhetskriterium blir derfor erstattet av spørsmålet om teorier er gode eller dårlige og hvorvidt det finnes empirisk data og statistikk før å bekrefte eller avkrefte teorier.

Hypotetisk deduktiv metode og abduksjon

Hva er strukturen på et bekreftende vitenskapelig argument ifølge den hypotetisk deduktive metode? Forklar ut fra Hempels artikkel, men bruk egne eksempler.

Hypotetisk deduktiv metode (HD-metoden) starter med at vi formulerer en teori eller hypotese. Deretter deduserer vi empiriske konsekvenser som følge av teorien eller hypotesen som testes gjennom eksperiment eller empiriske observasjoner. Dersom deduksjonen viser seg å stemme bekreftes teorien induktivt, til en viss grad. Det sies at teorien bekreftes til viss grad på grunn av at teoriene aldri kan bekreftes til å være 100% sanne. Hempel (1966) beskriver dette med at fremtidig relevant data kan gi nye funn enn det som tidligere har kommet frem. Dette kan beskrives med et eksempel fra styrketrening. For å oppnå økt maksimal styrke er det bedre å trene med få repetisjoner og tunge vekter enn å trene med mange repetisjoner på en lettere vekt. Nåværende forskning viser at dette stemmer, men vi kan ikke bevise at dette er sant ved 100% av tilfellene. Det kan komme forskning frem i tid som viser helt andre resultater enn det vi har i dagens litteratur. Med en slik induktiv påstand vil vi aldri være helt sikker på at det stemmer, men økt vitenskap og mer forskning vil enten styrke eller svekke påstanden. På denne måten kommer vi også frem til svakhetene ved Hempel sin HD-metode.

Andre faktorer kan også påvirke resultatet til påstanden. For eksempel fysiske forutsetninger hos populasjonen det forskes på, forskjellig næringsinntak, søvnmengde og -rytme osv. Alle disse faktorene kan spille inn på effekten av styrketrening enten det er med tunge vekter og få repetisjoner eller om det er med lette vekter og mange repetisjoner. HD-metoden er alltid åpen for at det er andre forklaringer og slutter seg derfor bare til en teoris styrke eller svakhet ut fra det man har testet empirisk (Hempel 1966). HD-metoden er også svak ved spørsmål om sannsynlighet eller prevalens, ettersom at det ikke er mulig å dedusere noe om frekvens. Et eksempel kan være følgende teori: Sannsynligheten for å få et hoftebrudd øker i takt med alderen. Hvis man i dette tilfellet sier at sjansen for å få et hoftebrudd øker med 2% hvert år etter fylte 60 år så vil man i teorien ha en sannsynlighet for å få et hoftebrudd som øker lineært med alderen. Svakheten med HD-metoden i et slikt tilfelle er at det finnes flere teorier/faktorer som kan påvirke risikoen for å få et hoftebrudd enn bare økt alder, for eksempel aktivitetsstatus, kjønn, tidligere skadehistorikk og funksjonsnivå for å nevne noen. Dette kan påvirke frekvensen for å få hoftebrudd etter fylte 60 år slik at sannsynligheten ikke

er slik at den økes 2% per år. I dette tilfellet tar ikke HD-metoden hensyn til at ulike faktorer kan påvirke dataen.

Abduksjon er en annen teori som ligner den hypotetisk deduktive metoden. Abduksjon er også kjent som «slutning til den beste forklaringen». Den er ment å løse problemene til HD-metoden. Den logiske strukturen i abduksjon er lignende den i HD-metoden, men skiller seg fra at den slutter seg til den beste forklaringen. Den beste forklaringen anses å være god eller bedre når den enten forklarer flere forskjellige data, har høyere forklaringskraft, utgår fra færre årsaker/faktorer eller at den er tilsvarende en annen forklaring, men enklere. Abduksjon krever med andre ord at man sammenligner flere teorier for deretter å lande på den teorien som best forklarer teorien/hypotesen. Der abduksjon stiller flere teorier opp mot hverandre og sammenligner for å finne den teorien som passer best, tar HD-metoden kun for seg en teori om gangen for så å undersøke om dataen/funnet induktivt støtter teorien.

Replikasjonskrisen

Hva mener Alexander Bird er forklaringen til at mange resultat i noen vitenskaper ikke repliseres? Oppsummere Birds argument for dette. Sammenlign også Birds forklaring med noen av de andre forklaringene som Bird diskuterer i seksjon 4. Har Bird rett i at hans forklaring er bedre?

Bird (2020) beskriver at vitenskapen innen psykologi og medisin har vært utsatt for en krise. En krise kalt replikasjonskrisen. Dette er fenomen der flere tidligere vitenskapelige studier er forsøkt replisert, men der resultatene ikke har stemt overens med tidligere studier. Årsaken har blitt forsøkt forklart ved å peke på dårlig gjennomførte studier, tvilsomme forskningspraksiser, at det finnes en bias mot å publisere negative resultater og til og med forfalskning av resultater. Replikasjonskrisen dreier seg også om frykten for å grave opp i tidligere artikler og kritisere disse, ettersom dette ville svekke tilliten til vitenskapen i offentligheten.

Alexander Bird forklarer i sin artikkel i British Journal for the Philosophy of Science (Bird 2020) det han mener er en av utfordringene til at vitenskaper ikke repliseres. Bird mente at basisratefeilen (base rate fallacy) er en sannsynlig årsak til replikasjonskrisen. Han beskrev at basisratefeilen sannsynlig vil oppstå dersom man trekker en slutning om sannsynlighet for en gitt forekomst av et generelt fenomen. Feilen som gjøres er at man fokuserer spesifikt på resultatet av en test, men glemmer å ta over seg hvor hyppig forekomsten av gitt fenomen skjer uavhengig av testen som gjennomføres. Dette gir utslag på at en for eksempel konkluderer med at sannsynligheten for å få en sjelden sykdom er større enn det den i realiteten er, fordi man kan feiltolke falske positive eller falske negative resultater (type-I-feil og type-2-feil). I vitenskapen i dag er resultater typisk regnet som signifikant eller ikke signifikant basert på P-verdien. Grensen for P-verdien er som regel satt på 0.05 (5%) og blir en slags pekepinn på om noe er sant eller usant basert på om resultatet er over eller under grensen. P-verdien sier noe om sannsynligheten for at et resultat er falskt positivt (type-1-feil). Et problem med å bruke P-verdi er da kan man konkludere på en befolkning, men P-verdien i studiet vil bare si

noe om utvalget som det forskes på. Type-2-feil, kalles også falsk negativ, og betyr at man forkaster en alternativ forklaring, selv om den egentlig stemmer.

Av andre forklaringer som kan ha bidratt til replikasjonskrisen skriver Bird (2020) om at flere studier har lav statistisk, publiseringsbias og tvilsom forskningspraksis og svindel. Flere har kommet med forklaringen om at lav statistisk styrke er en av årsakene til replikasjonskrisen. Studier med få forsøkspersoner blir spesielt trukket frem, da det er vanskeligere å replisere resultater gjort på små populasjoner. Publiseringsbias eller publiserings-skjevhet er det at det er for mange studier med positive resultater er publisert, når man sammenligner med studier med negative resultater. Det er et stort press fra flere kanter som alle ønsker å publisere nye funn og positive resultater, og ikke negative resultater. Dette kan føre til økt ønske om å replisere studier med positive funn og at igjen kan øke forekomsten av at falske positive resultater blir publisert. Tvilsom forskningspraksis er også et av områdene det gjøres mye feil i. Dette kan være alt fra ubevisst skjevhet og p-hacking, der man leter bevisst etter positive resultater i egen studie, til helt bevisst forfalskning av data. Presset på å publisere er så høyt at dette kan føre til at forskningen gjennomføres etter standarden den burde. Det er også store økonomiske faktorer som påvirke ønsket om å publisere. Det kan være årsaker der større studier koster langt mer penger å gjennomføre og derfor tar man til takke med mindre studier og publiserer heller resultater og funn hyppigere.

Det er lett å si at Bird har rett til at basisratefeilen er en av årsakene til replikasjonskrisen. Dette er heller ikke hele årsaken da vi vet et at det finnes skjevhet der positive eller negative resultater ikke publiseres like ofte. Ønsket om å finne helt nye funn er så stort at forskere kan ignorere bevisst og ubevisst sannsynligheten for at de har falske positive resultater. Det er veldig mange studier som er gjennomført med lav statistisk styrke som følge av få forsøkspersoner, dette fører til at det er vanskelig å replisere disse studiene. Det økonomiske presset eller insentivene som finnes er også store, gjerne for å finne positive resultater, som kan gi disse resultatene for mye vekt, men også føre til økt hyppighet av forfalskning og svindel.

Deloppgave 4: Studiedesign

Introduksjon

Styrketrening er den viktigste intervensjonen for å øke muskelmasse blant mennesker. Teorien sier at volumet med styrketrening spiller en signifikant rolle i kroniske muskulære endringer og tilpasninger som for eksempel økt styrke og økt muskelmasse (Kraemer and Ratamess 2004). Ivaretagelse av muskelstyrke og muskelmasse er viktig for alle. For populasjonen i sin helhet er det viktig med tilstrekkelig styrke for å kunne utføre daglige oppgaver uten utfordringer. Å være fysisk sterk har også flere positive helseeffekter, det er til og med god evidens på at økt muskulær helse gir lavere dødelighet (Ruiz et al. 2008).

Det er tidligere foreslått at styrketrening med høyt treningsvolum, med moderat til høy intensitet og korte hvilepauser gir først og fremst muskelbyggende effekt i form av hypertrofi, deretter økt utvikling av styrke (Baechle, Earle, and Conditioning Association (U.S.) 2008). På den andre siden vil et lavere treningsvolum, men tyngre vekter/økt intensitet og lengre hvilepauser primært være et program for å øke muskelstyrke først og fremst. Dette er i tråd med hvordan kroppsbyggere trener i faser der de ønsker å maksimere muligheten for hypertrofi og i studier som rapporterer om økt hypertrofi i tråd med at treningsvolumet øker (Schoenfeld, Ogborn, and Krieger 2017).

I denne oppgaven skal det undersøkes hvordan ulike mengder og ulike intensiteter med styrketrening påvirker utviklingen av styrke og muskelvekst hos trente individer. Det er valgt ut fem artikler for gjennomgang av disse undersøkelsene (Brigatto et al. 2022; Mangine et al. 2015, 2016; Schoenfeld et al. 2015, 2019). For å finne frem til nettopp disse 5 studiene ble det brukt et generelt litteratursøk etter originale studier rettet mot treningsvolum og intensitet, begrenset til de siste 10 årene. 2 av studiene undersøkte lavt treningsvolum versus høyt treningsvolum, på ulike intensitetsområder (Schoenfeld et al. 2015, 2019). I de to neste artiklene undersøkte forfatterne (Mangine et al. 2015, 2016) effekten av et visst treningsvolum versus treningsintensitet i to ulike studier, med litt forskjellige målemetoder. I det siste av fem studier så Brigatto et al. (2022) på effekten av ulike treningsvolum for å undersøke nærmere dose-respons forholdet til styrketrening. Oppgaven inkluderer en vurdering av studiedesign, likheter og ulikheter i studiene og styrker og svakheter med de forskjellige studiene.

Metode

Schoenfeld et al. (2015) brukte et randomisert studiedesign der 24 menn ble parett etter styrke ved pre-test for deretter å bli tilfeldig plassert i en av to grupper. De hadde en gjennomsnittsalder på 23 år, hadde ingen nåværende muskel- og skjelettskader og var «rene utøvere». Alle deltagerne var erfarne i styrketrening, det vil si at de har trent minimum 3 økter i uken det siste året og regelmessig trent knebøy og benkpress. Studien bestod av en gruppe som trente på lav intensitet (30-50% av 1RM) til utmattelse og den andre gruppen trente 8-12 repetisjoner på høy intensitet (70-80% av 1RM). Begge gruppene trente samme øvelser, men med ulike vekter og repetisjoner. Tester ble gjennomført før og etter (pre og post) intervensjonsperioden (8 uker) med måling av maksimal styrke i form av 1RM i Knebøy og benkpress og måling av muskulær utholdenhet i benkpress ved maks antall repetisjoner utført på 50% av 1RM. Det ble også målt tykkelse av muskulatur ved å bruke ultralyd. Tykkelsen ble målt på triceps, biceps og quadriceps. En parett t-test ble brukt for å undersøke forskjeller fra før og etter intervensjon. Multipel regresjonsanalyse med mål post-intervensjon som avhengig variabel og pre-test verdiene som kovariater ble brukt for å vurdere forskjeller mellom gruppene. Forutsetningene i regresjonsanalysene ble kontrollert. Signifikansnivået ble satt til $P < 0.05$.

Mangine et al. (2016) brukte et randomisert studiedesign der 33 fysisk aktive menn ble rekruttert for å sammenligne to ulike intervensjoner. Alle deltagerne med erfaring i styrketrening tilsvarende minimum 2 år (gjennomsnitt på 5.7 års erfaring), uten noen nåværende muskel- og skjelettskader. Intervensjonen var et 8 ukers styrketreningsprogram etter et 2 ukers styrketreningsprogram likt for alle deltagerne. Gruppene ble delt i to grupper der en gruppe gjennomførte trening på høy intensitet og lavt treningsvolum, og den andre gruppen gjennomførte trening på høyt volum og moderat intensitet. Utover dette har ikke forfatterne utdypet metoden de brukte for å rekruttere deltagere. Deltagerne ble testet i 1RM i knebøy og benkpress, med måling av hastigheten på stangen i den konsentriske fasen. Nøyaktige samme maksimale vekt ble brukt på pre- og post-tester, slik at forfatterne kunne utforske om hastigheten endret seg på samme vekt etter intervensjonsperioden. Kraftproduksjon ved isometrisk kontraksjon i andredraget (midten av lårene, i markløft/frivending) ble også målt ved hjelp av en kraftplate og en fastspent stang. Dataen ble analysert med kovariansanalyse (ANCOVA) etter intervensjonsperioden. Dataen fra pre-test ble brukt som kovarians. En parett t-test ble brukt for å undersøke om det var en signifikant endring fra pre- til post-test. Forskjellen mellom gruppene ble analysert ved å regne ut effektstørrelsen. Signifikansnivået ble satt til $P < 0.05$.

Brigatto et al. (2022) brukte et randomisert studiedesign der 27 mannlige deltagerne ble parett etter styrkenivå og deretter tilfeldig plassert i en av tre like store grupper. Der gruppene gjennomførte henholdsvis, 16, 24 og 32 ukentlig sett med trening per muskelgruppe. Deltagerne hadde minimum 1 års erfaring med styrketrening tre ganger per uke og ble rekruttert gjennom universitetets treningsstudio. Det ble på forhånd gjort beregninger på utvalgsstørrelsen med basert på en pilotstudie, med en effektstørrelse på 0.75, alfanivå på $P < 0.05$ og statistisk styrke på 0.80. Deltagerne ble testet med 1RM i knebøy og benkpress og ved å måle

muskeltykkelse ved bruk av ultralyd på biceps, triceps og vastus lateralis. Det var en utdannet tekniker som utførte alle ultralydene. Forfatterne bekreftet normaliteten og homogeniteten til variansene ved å bruke Shapiro-Wilk og Levene tester. En variansanalyse med gjentatte mål (ANOVA) ble brukt for å sammenligne 1RM og hypertrofi før og etter intervensjonen på tvers av studiegruppene.

Schoenfeld et al. (2019) brukte et randomisert studiedesign med 45 mannlige deltagere som meldte seg frivillig. Deltagerne var studenter på et universitet som var erfarne innen styrketrening med gjennomsnittlig erfaring på 4.4 år. Det ble på forhånd gjort beregninger på utvalgsstørrelsen med en effektstørrelse på 0.25, med en alpha-verdi på 0.05 og en statistisk styrke på 0.80. Disse beregningene identifiserte at studien hadde behov for minimum 36 deltagere. Øvrige deltagere var rekruttert i tilfelle noen trakk seg underveis. Deltagerne ble deretter tilfeldig delt in i 3 like store grupper der en gruppe gjennomførte 1 sett per øvelse per økt, en gruppe trente 3 sett per øvelse per økt og den siste gruppen trente 5 sett per øvelse per økt. Det ble målt muskeltykkelse med ultralyd på biceps, triceps og quadriceps av en enkelt utdannet tekniker på alle deltagere. Maksimal styrke ble testet med 1RM i knebøy og benkpress og muskulær utholdenhet i benkpress ble testet med 50% av 1RM til utmattelse. Statistiske analyser ble gjennomført ved å bruke både frekventistisk og Bayesianisk tilnærming. Den første brukte en ANCOVA på endringen fra pre-test til post-test, med separate grupper som faktor og med verdiene fra pre-test som kovarians. I den Bayesianiske tilnærmingen ble det brukt en JZS Factor ANCOVA med standardiserte skalaer. Signifikansnivået ble satt til $P < 0.05$.

Mangine et al. (2015) brukte et randomisert studiedesign med 33 fysisk aktive menn. Minimum 2 års erfaring med regelmessig styrketrening var et av inklusjonskriteriene (samme studiepopulasjon som Mangine et al. 2016). Alle deltagere gjennomførte samme treningsprogram på 2 uker før de ble tilfeldig delt inn i en av 2 like store grupper. Den ene gruppen trente med et høyt volum på 10-12 repetisjoner på 70% av 1RM, mens den andre gruppen trente på en høy intensitet på 3-5 repetisjoner på 90% av 1RM. Begge gruppene trente i 8 uker. Deltagerne ble testet og målt med maksimal styrke med 1RM i knebøy og benkpress, ultralydmålinger på triceps, quadriceps og pectoralis major for å måle muskeltykkelse og DXA skann for å måle kroppsmasse og endringer. Alle DXA undersøkelser ble utført av samme utdannede radiograf. Blodprøver ble tatt ved 4 tidspunkt på pre- og post-testing, med gjentakende tidspunkt på dagen. Dataen ble analysert med kovariansanalyse (ANCOVA) etter intervensjonsperioden. Dataen fra pre-test ble brukt som kovarians. En paret t-test ble brukt for å undersøke om det var en signifikant endring fra pre- til post-test. Forskjellen mellom gruppene ble analysert ved å regne ut effektstørrelsen. Signifikansnivået ble satt til $P < 0.05$.

Diskusjon

Alle studiene gjennomførte et randomisert studiedesign, men Brigatto et al. (2022) skriver sin studie som en randomisert longitudinell studie. Ettersom ingen av studiene har en kontroll-

gruppe, men sammenligner to eller flere ulike intervensjoner er det heller ingen av forfatterne som hevder at de gjennomfører en randomisert kontrollstudie. Alle studiene er gjennomført med en 8 ukers treningsperiode. Antall forsøkspersoner er fordelt fra 18 til 45 forsøkspersoner. Alle forsøkspersoner er unge menn i aldersgruppen 18-35 år og alle er kjent med styrketrening, med litt ulike definisjoner på antall år for å regnes som kjent med styrketrening. Alle studiene beskriver hvordan forsøkspersonene er rekruttert, Mangine et al. (2016) har en begrenset beskrivelse av rekruttering av forsøkspersoner. Tre av fem studier er helt randomiserte og de siste studiene bruker stratifisert blokkert randomisering (Hulley 2013).

Alle fem studier bruker 1RM målinger i knebøy og markløft, noe som bidrar til en enkel sammenligning av forskjellige studier til tross for ulike spørsmål mål med studiene. Utover det varierer målemetoden litt fra studie til studie. Det er ikke i noen av studiene nevnt noe om teknikker som gjennomfører ultralyd- eller andre målinger er blindet. Det er også relativt store variasjoner i hvordan målinger med ultralyd er gjennomført. Liggende og stående målinger er ikke nødvendigvis sammenlignbare, og det ville vært en fordel i denne analysen at alle målingene ble gjennomført i samme utgangsstilling. Øvelsene som brukes i de forskjellige studiene samsvarer i all hovedsak godt i de forskjellige studiene, men kun en eller to øvelser som varierer fra studie til studie. Felles for studiene er inklusjon av typiske øvelser som brukes i styrketreningsprogram og av kroppsbyggere (Baechle, Earle, and Conditioning Association (U.S.) 2008).

To av de fem studiene beskriver at de på forhånd har beregnet utvalgsstørrelsen som er nødvendig for å trekke konklusjoner, men de øvrige tre studiene beskriver ikke dette i sin statistikk. De forskjellige studiene bruker ANCOVA der data samlet fra pre-test ble brukt som kovarians. Brigatto et al. (2022) bruker ANOVA for å sammenligne 1RM og hypertrofi før og etter intervensjonen på tvers av studiegruppene. Tre av fem studier beskriver at de også brukte en paret t-test for å undersøke om det var en signifikant endring fra pre- til post-test innad de forskjellige gruppene.

Alle studiene svarer på egen problemstilling gjennom analyse av endringer i pre og post-verdier, både innad i egne grupper og på tvers av ulike grupper. Alle studier, bortsett fra Mangine et al. (2015) går grundig gjennom studiene sine begrensninger, sistnevnte beskriver ikke begrensninger studiet kan ha hatt. Alle artikler er spesifikke på sine konklusjoner til sin populasjon (unge, trente menn), men ikke alle går dypere gjennom mulig inferens mot øvrig populasjon.

Konklusjon

Alle studiene adresserte hypotesen deres i konklusjonen og diskuterer resultatene som førte til dette. Schoenfeld et al. (2015), Schoenfeld et al. (2019) og Brigatto et al. (2022) foreslår alle videre forskning innenfor teamet. Alle spesifiserer også behovet for studier på treningsvolum og effekt av styrketrening mot en bredere befolkning, for eksempel kvinner og eldre. Mangine et al. (2015) og Mangine et al. (2016) er to studier som kommer fra nøyaktig samme populasjon og lik intervensjon, men med ulike mål og hypoteser. Ingen av de studiene kommer med forslag

til videre forskning og har heller ikke et stort søkelys på inferens i sin artikkel. Videre studier undersøke og sammenligne effekten unge, trente menn har på styrketrening med eldre og/eller kvinner. Dette kan i fremtiden gi oss mer kunnskap om treningsprogrammering og spesifikk inkludering av trening på en gitt intensitet eller et gitt treningsvolum.

Deloppgave 5: Analysere repeterte målinger

References

- Baechle, Thomas R., Roger W. Earle, and National Strength & Conditioning Association (U.S.), eds. 2008. *Essentials of Strength Training and Conditioning*. 3rd ed. Champaign, IL: Human Kinetics.
- Bird, Alexander. 2020. "Understanding the Replication Crisis as a Base Rate Fallacy." *The British Journal for the Philosophy of Science*, December, 000–000. <https://doi.org/10.1093/bjps/axy051>.
- Brigatto, Felipe A., Leonardo Emmanuel de Medeiros Lima, Moisés D. Germano, Marcelo S. Aoki, Tiago V. Braz, and Charles R. Lopes. 2022. "High Resistance-Training Volume Enhances Muscle Thickness in Resistance-Trained Men." *The Journal of Strength & Conditioning Research* 36 (1): 22–30. <https://doi.org/10.1519/JSC.0000000000003413>.
- Hempel, Carl G. 1966. *Philosophy of Natural Science*. Prentice Hall.
- Hulley, Stephen B., ed. 2013. *Designing Clinical Research*. 4th ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins.
- Kraemer, William J., and Nicholas A. Ratamess. 2004. "Fundamentals of Resistance Training: Progression and Exercise Prescription." *Medicine & Science in Sports & Exercise* 36 (4): 674–88. <https://doi.org/10.1249/01.MSS.0000121945.36635.61>.
- Mangine, Gerald T., Jay R. Hoffman, Adam M. Gonzalez, Jeremy R. Townsend, Adam J. Wells, Adam R. Jajtner, Kyle S. Beyer, et al. 2015. "The Effect of Training Volume and Intensity on Improvements in Muscular Strength and Size in Resistance-Trained Men." *Physiol Rep* 3 (8): e12472. <https://doi.org/10.14814/phy2.12472>.
- Mangine, Gerald T., Jay R. Hoffman, Ran Wang, Adam M. Gonzalez, Jeremy R. Townsend, Adam J. Wells, Adam R. Jajtner, et al. 2016. "Resistance Training Intensity and Volume Affect Changes in Rate of Force Development in Resistance-Trained Men." *Eur J Appl Physiol* 116 (11-12): 2367–74. <https://doi.org/10.1007/s00421-016-3488-6>.
- Okasha, Samir. 2016. *Philosophy of science: a very short introduction*. Second edition. Vol. 67. Very short introductions. Oxford: University Press.
- Popper, Karl R. 1969. *Conjectures and Refutations: The Growth of Scientific Knowledge*. 3rd ed. (revised). London: Routledge & K. Paul.
- Ruiz, J. R., X. Sui, F. Lobelo, J. R. Morrow, A. W. Jackson, M. Sjostrom, and S. N. Blair. 2008. "Association Between Muscular Strength and Mortality in Men: Prospective Cohort Study." *BMJ* 337 (jul01 2): a439–39. <https://doi.org/10.1136/bmj.a439>.
- Schoenfeld, Brad J., Bret Contreras, James Krieger, Jozo Grgic, Kenneth Delcastillo, Ramon Belliard, and Andrew Alto. 2019. "Resistance Training Volume Enhances Muscle Hypertrophy but Not Strength in Trained Men." *Medicine & Science in Sports & Exercise* 51 (1): 94–103. <https://doi.org/10.1249/MSS.0000000000001764>.

- Schoenfeld, Brad J., Dan Ogborn, and James W. Krieger. 2017. "Dose-Response Relationship Between Weekly Resistance Training Volume and Increases in Muscle Mass: A Systematic Review and Meta-Analysis." *Journal of Sports Sciences* 35 (11): 1073–82. <https://doi.org/10.1080/02640414.2016.1210197>.
- Schoenfeld, Brad J., Mark D. Peterson, Dan Ogborn, Bret Contreras, and Gul T. Sonmez. 2015. "Effects of Low- Vs. High-Load Resistance Training on Muscle Strength and Hypertrophy in Well-Trained Men." *Journal of Strength and Conditioning Research* 29 (10): 2954–63. <https://doi.org/10.1519/JSC.0000000000000958>.