

Mappeeksamen

Hermann Moen Kandidatnr.: 111

2021-12-02

Contents

1	Realabilitet	5
1.1	Introduksjon	5
1.2	Metode	5
1.3	Resultater	6
1.4	Diskusjon	8
2	Labrapport - cDNA synthesis using Superscript IV and general qPCR	9
2.1	Formål	9
2.2	Metode	9
2.3	Resultater	10
2.4	Diskusjon	12
3	Vitenskapsteori	13
3.1	Falsifikasjon	13
3.2	HD-metoden og abduksjon	14
3.3	Replikasjonskrisen	15
4	Study Designs	17
4.1	Hva undersøker disse studiene?	17
4.2	Hva kan forklare resultatene?	17
4.3	Hvilken sammenheng har hypotese med eksisterende litteratur i studiene?	18
4.4	Hvilke forskjeller er det i de metodiske valgene i studiene?	18
4.5	Hva er resultatene og konklusjonene	20

5 Høyere treningsvolum leder til bedre styke og muskelvekst i underekstremiteter hos utrente	21
5.1 Introduksjon	21
5.2 Deltakere og studieoversikt	22
5.3 Databehandling	22
5.4 Treningsprotokoll	23
5.5 Testprotokoll	23
5.6 Resultater	24
5.7 Diskusjon	24

Kapittel 1

Realabilitet

1.1 Introduksjon

Maksimalt oksygenopptak VO_{2max} ble først beskrevet av Hill og Lupton i 1923, og kan defineres som kroppens evne til å ta opp og forbruke oksygen per tidseenhet [Bassett and Howley, 2000, Hill and Lupton, 1923]. Innen toppidrett måles ofte det maksimale oksygenopptaket for å måle utøverens kapasitet opp mot arbeidskravet i den spesifikke idretten, og VO_{2max} kan i så måte også sees på som et mål på den aerobe effekten til utøveren [Bassett and Howley, 2000]. I Olympiatoppens testprotokoller benytter de flere definerte hjelpekriterier for å sikre at man faktisk har funnet deltakerens maksimale oksygenopptak [Tønnessen et al., 2017]. Følgende kriterier er beskrevet; platå i O_2 er oppnådd, økning i ventilasjon med utflating av O_2 verdi, RER-verdi over 1.10 (1.05 om laktatprofiltest er gjennomført i forkant) og blodlaktat over 8 [Tønnessen et al., 2017].

1.2 Metode

I forkant av testen ble testpersonen veid uten sko og 300g ble trukket av for vekten av klær. Denne vekten ble brukt i beregningen av maksimalt oksygenopptak ($ml\ kg^{-1}\ min^{-1}$). For å sikre intern validitet ble deltakerne bedt om å avstå fra anstrengende fysisk aktivitet dagen før test, standardisere måltidet i forkant av test samt avstå fra inntak av koffein under de siste 12 timene før testen [Halperin et al., 2015]. Både test 1 og 2 ble gjennomført på samme tid på døgnet under standardiserte forhold. Test 2 ble gjennomført 6 dager etter gjennomført test 1. Det ble ikke kontrollert for fysisk aktivitet mellom testdagene.

Det var 11 deltakere (tabell 1) i studien, der deltakerne gjennomførte en 10 minutter lang oppvarmingsprotokoll på tredemøllen (Woodway 4Front, Wisconsin,

USA), beskrevet for deltakerne i forkant av testen. Denne oppvarmingsprotokollen bestod av fem minutter på 11-13 i Borg 6-20 RPE skala [Borg, 1982], etterfulgt av 2x1min på starthastighet og stigning med 30 sekund pausene mellom dragene. Siste tre min var også 11-13 i borg på valgfri stigning. Etter oppvarming var det to min pause før testen begynte. Starthastighet (tabell 1) var satt til 8km/t, med stigning på 10.5% og 5.5% for henholdsvis menn og kvinner.

VO₂max ble målt ved hjelp av en metabolsk analysator med miksekammer (vyntus CPX, mixingchamber (Vyntus CPX, Jaeger-CareFusion, UK)). Forut for alle tester ble analysatoren gass og volumkalibrert. Analysatoren ble stilt inn til å gjøre målinger hvert 30sek, og VO₂max ble kalkulert gjennom å bruke snittet av de to høyeste påfølgende målingene av O₂. Underveis i testen mottok alle deltakerne en høylytt verbal oppmuntring fra testleder som var standardisert [Halperin et al., 2015]. Alle deltakerne gjennomførte også begge testene med samme testleder og med samme personer til stede i rommet for å redusere konfundering [Halperin et al., 2015].

For hvert medgåtte minutt av testen ble hastigheten på møllen økt med 1km/t, helt til utmattelse, hvor testen ble avsluttet. Deltakernes hjerterefrekvens ble også registrert under hele testen. Når testen ble avsluttet ble deltakerne bedt om å rapportere opplevd anstrengelse ved hjelp av Borg-skala [Borg, 1982]. Maksimal hjerterefrekvens under testen ble også registrert. Ett minutt etter avsluttet test ble hjerterefrekvens registrert, og det ble målt og analysert blodlaktat (Biosen C-line, EKF Diagnostics, Barleben, Germany).

Table 1.1: Deltakeroversikt

	Kvinner	Menn
N	4	7
Alder (år)	24.5 (1.29)	23.9 (1.77)
Vekt (kg)	58.9 (6.28)	74.8 (5.55)
Høyde (cm)	166 (2.99)	180 (3.1)
Verdier er gitt som gjennomsnitt og (Standardavvik)		

1.3 Resultater

Deskriptive data for disse deltakerne er vist i Tabell 1, i Figur 1 kan man se forskjellen mellom test 1 og 2 fordelt på kjønn. Det typiske målefeilet (typical error, [Hopkins, 2000]) fra test 1 til test 2 er utregnet til å være 4.04%.

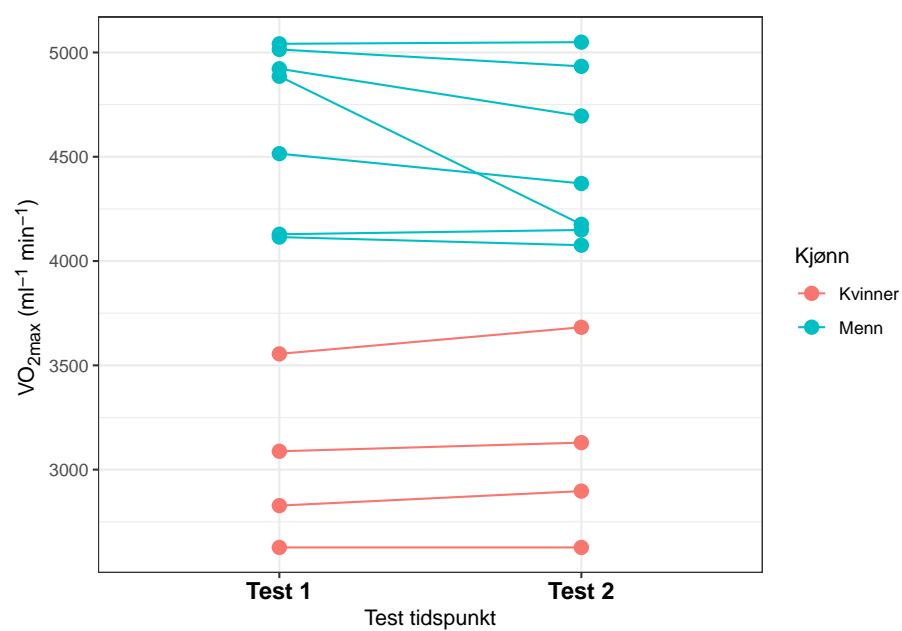


Figure 1.1: Figuren viser forskjellen i VO_{2max} (ml/min) i test 1 og test 2

1.4 Diskusjon

4,04%

Resultatet i vår realibilitetsstudie er funnet av typefeilen på 4,04. Ettersom testing av maksimalt oksygenopptak er en test som gjennomføres til utmattelse, vil man kunne forvente en viss variasjon i testresultatene ettersom opplevd anstrengelse kan påvirkes av flere ulike variabler [Halperin et al., 2015]. For å redusere støy vil flere faktorer være nyttig å ta hensyn til under slik testing. Som nevnt i metoden vil standardisering av matinntak, koffeininntak, utstyr og tidspunkt for gjennomføring av test være med på å kunne sikre intern validitet i resultatene. Eksempler er deltakernes kjennskap til testen, verbal oppmuntring og personer tilstede under testen er andre faktorer som potensielt kan bidra til støy [Halperin et al., 2015]. Felles for alle faktorer er at graden av påvirkning på resultatene muligens reduseres ved hjelp av en standardisert testprotokoll. Deltakerne - og testlederne, sin kjennskap til testen er en annen faktor som trolig påvirker resultatene i vårt prosjekt. I dette tilfellet fantes det enkelte deltakere som hadde gjennomført en liknende test flere ganger, og en kan da forvente en mindre grad av variasjon mellom resultatene på test 1 og test 2, sammenlignet med de deltakerne som gjennomførte testen for første gang på pretest. Dette fordi kjennskapen og kunnskapen de tilegnet seg på pre-test, trolig spiller inn på testresultatene. Typefeilen på 4.04% kan også tyde på at enkelte av disse resultatene kan være utsatt for støy av ulik sort [Hopkins, 2000].

Grunnen til at vi snakker om typefeil er at når vi ønsker å måle påvirkningen av trening på en gruppe individer er det viktig å kunne si noe om hva som er endring og hva som er støy (målefeil). Desto mindre støy en test innebærer jo bedre er målingen. Målet som brukes er typefeil. Hva som danner denne variasjonen som representeres ved typical error er multifaktorelt, men hoveddelen er som oftest biologisk [Hopkins, 2000].

For å måle typefeil har vi brukt within subject deviation metoden. Denne metoden påvirkes ikke av at gjennomsnittet endrer seg fra test til test [Hopkins, 2000]. Data for målinger i VO₂max fra fem sertifiserte Australske laboratorier fastslo ett gjennomsnitt på 2.2% for typefeil [Halperin et al., 2015]. Data fra det Australske institutt for sport har også fastslått at en typefeil på omtrent 2% er riktig for både maksimal og submaksimal O₂ [Clark et al., 2007, Robertson et al., 2010, Saunders et al., 2009]. Dette indikerer at med godt kalibrert utstyr og med utøvere som er godt vant med testingen vil en typefeil på 2% for det biologiske, og analytiske være riktig [Halperin et al., 2015]. Vår typefeil på 4.04% kan derfor tenkes å være et bilde hvordan det kan se ut med få deltakere, med ulikt utgangspunkt, men også uten skikkelig standardisering av treningshverdagen i forkant av testene. Det kan også tenkes at med et varierende nivå hos deltagerne kan enkelte oppleve en treningseffekt av test 1. Samtidig som andre kanskje ble slitne av å få en test inn i treningshverdagen.

Kapittel 2

Labrapport - cDNA synthesis using Superscript IV and general qPCR

2.1 Formål

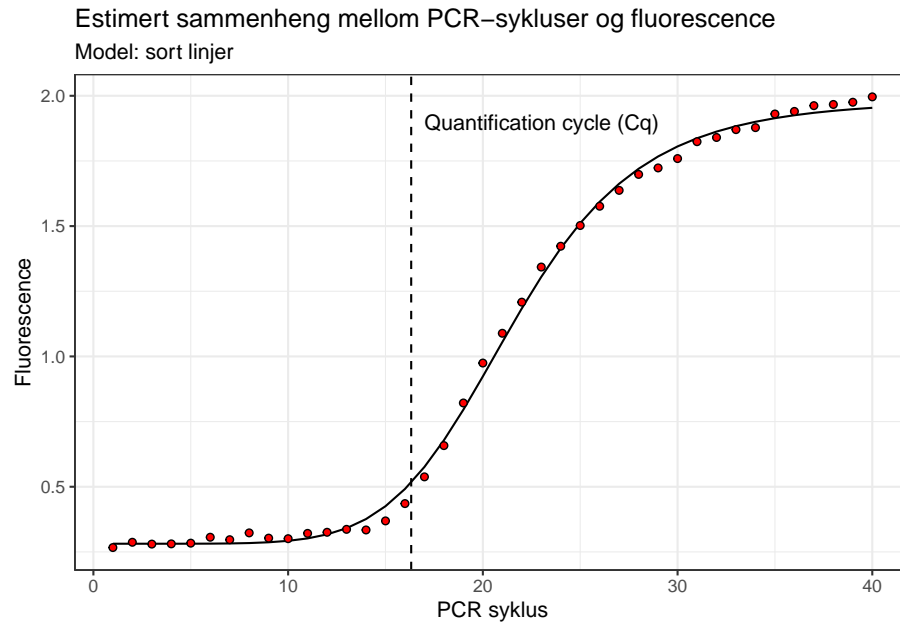
RNA-overflodsanalyse er gjort ved hjelp av syntese av komplementært DNA fra enkelttrådet RNA. Vi ønsker å amplifisere opp bestemte proteiner ved hjelp av bestemte primere og qPCR. Vi ønsker å få frem en cQ-verdi for å kunne evaluere gen-opphopningen, og sammenlikne mål-genene med referanse gener.

2.2 Metode

Vi hentet cDNA fra 3 forsøkspersoner. Dette er cDNA hentet fra testene som ble gjennomført i uke 0 og uke 2. Alle prøver er fra venstre ben. Det ble laget en 5 fortynningsserier fra disse prøvene. Dette ble fortynnet ved hjelp av DEPC-behandlet vann, i følgende serie 1:10, 1:50, 1:250, 1:1250, 1:6250, 1:31250, 1:156250. Vortex ble brukt mellom hver fortynningsfase.

Det ble deretter laget sju forskjellige master mixer ved hjelp av 3 referansegener (REEP5, CHMP2A, B2M) og 4 målgener (MyHC I, 2A, 2X, rRNA 475). Mastermix bestod av 5 µl sybr green, 1 µl valgt gen, 2 µl DEPC-behandlet vann, 2 µl fortynnet cDNA. Deretter ble det fylt 71 brønner i en qPCR-reaksjonsplate ed henholdsvis 2 µl prøve, og 8 µl med mastermix. Reaksjonsplaten med brønner ble dekt med plastfilm, og sentrifugerte 1 minutt på 1200 omrdreininger (rpm), før PCR protokoll ble gjennomført.

En PCR protokoll ble på forhånd forberedt i QuantStudio5. PCR protokollen bestod av 50 grader i 2 minutter, og 95 grader i 2 minutter, før den kjørte 40 sykluser bestående av 1 sekund på 95 grader celsius, og 30 sekunder på 60 grader celsius.



2.3 Resultater

Modellen viser sammenhengen mellom antall sykluser og fluorescence. Flere PCR-sykluser gir flere kopier, og dermed også en økt konsentrasjon i prøven. På denne måten kan vi bruke fluorescence til å si noe om hvor mange sykluser som må til for å oppnå en bestemt terskelverdi (Cq-verdi). Med primerne vi benyttet i forsøket var det ønskelig med et sted mellom 10 og 40 sykluser for å sikre at vi oppnådde terskelverdien. Det ble derfor kjørt 40 sykluser. Ved flere sykluser øker trolig sannsynligheten for falske positive.

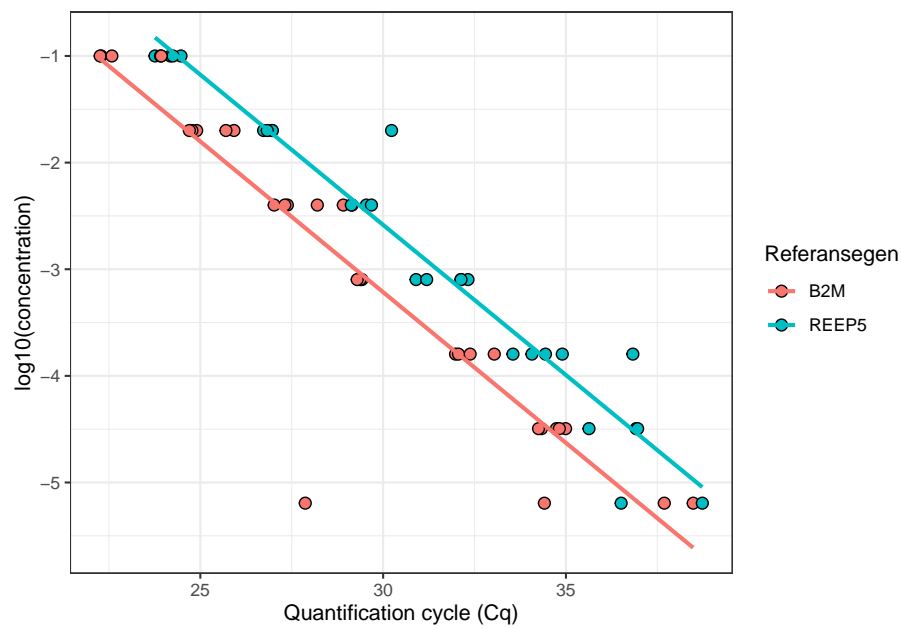
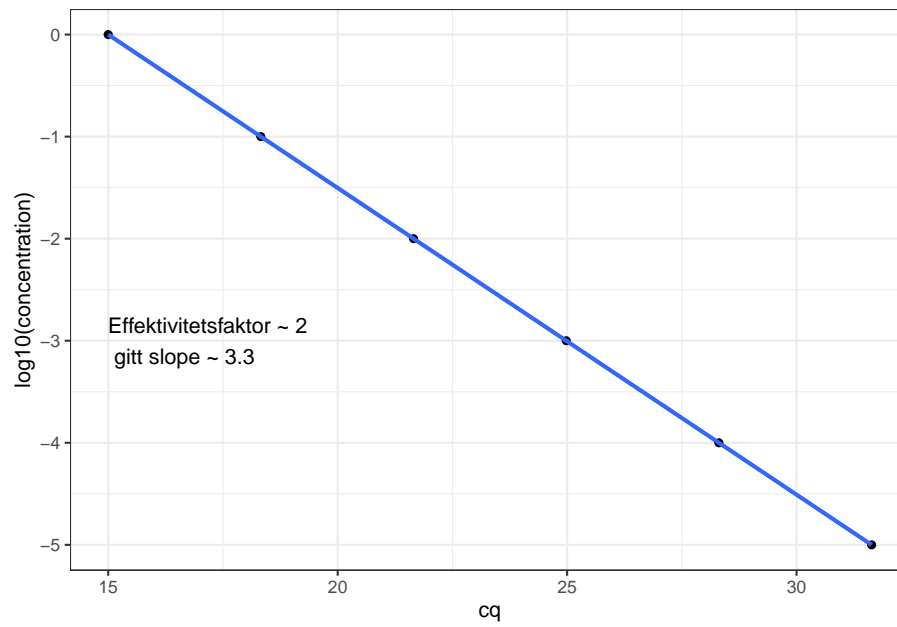


Figure 2.1: Efficiency calculations made from serial dilution of cDNA

2.4 Diskusjon

Cq-verdien sier noe om hvor mange PCR-sykluser som trengs for å detektere ulike målgen[Kuang et al., 2018]. En høyere Cq-verdi indikerer altså at mengden RNA må dobles flere ganger for å detektere en terskelverdi av target. En lavere Cq-verdi indikerer at terskelverdien oppnås ved færre PCR-sykluser, altså at konsentrasjonen av target er høyere. En lavere Cq-verdi ved uke 2, sammenlignet med uke 0(baseline) som i forsøket, indikerer høyere konsentrasjon ved uke 2 enn ved uke 0. Dermed en effekt av intervensjonen, avhengig av funksjonen til målgenet vi underøkser.

Kapittel 3

Vitenskapsteori

3.1 Falsifikasjon

Falsifiseringsprinsippet Karl Raimund Popper (1902-1994), var en Østerisk-Britisk filosof. Popper hadde siden høsten 1919 bearbeidet spørsmålet om hva som skulle til for å kunne danne et skille mellom vitenskap og pseudo-vitenskap. På den tiden(starten av og et stykke inn på 1900tallet) var det en aksept for at skilnaden mellom vitenskap og pseudo-vitenskap, var den empiriske fremgangsmetoden [Popper, 2002]. Den empiriske fremgangsmetoden går fra observasjoner, eller eksperimenter til å danne teorier ved hjelp av induktive argumenter. Et induktivt argument bygger på at premissene som er satt er sanne, og ved nok observasjoner av disse premissene, vil konklusjonen sannsynligvis være sann, eller man kan si at man har styrket bekreftelse av en teori. Denne tankerekkefølgen ble kritisert av David Hume (1711- 1776) allerede på 1700tallet. Hume stilte spørsmål for hvordan man kan slå fast at induksjon er pålitelig, og svaret ble ved hjelp av et argument. Hvilket argument skulle brukes? Induktiv metode kan ikke garanteres for ved et deduktivt argument, da det ikke finnes et deduktivt argument for at induksjon vil fungere i fremtiden. Et induktivt argument bygger på erfaringer, men kan heller ikke slå fast induksjonens sikkerhet. Dette var kronglete å holde styr på i hodet, men var et fundamentet i kritikken av induksjonismen langt inn på 1900tallet.

Popper som i utgangspunktet var ute etter å skille vitenskap fra pseudo-vitenskap viste om denne svakheten ved induktive argumenter som var det som var det opprinnelige skille. Ved hjelp av datidens største og mest fremgangsrike vitenskapsteorier lette Popper etter forskjeller. Einsteins relativitetsteori, marx historiske teorier, Freuds psyko-analyser, og Alfred Adlers individuelle psykologi var de mest omtalte teoriene på den tiden. Når han etter hvert oppdaget at flere av punktene i de forskjellige teoriene ble bekreftet, ble han også mens tiden gikk mindre og mindre imponert av Marx, Freud og Adler.

Det var her han startet å se etter forskjeller. De andre teoriene hadde mange bekreftelser, og hvert eneste tilfelle kunne på en eller annen måte forklares ved hjelp av teoriene. Einsteins relativitetsteori kunne derimot motsies ved hjelp av en eksakt måling. Ut av dette vokste Poppers falsifiserbarhetskriterium [Popper, 2002].

Falsifiserbarhetskriteriumet vokste altså frem av at en teori ikke burde bekreftes, men avkreftes, eller falsifiseres som det kalles vitenskapelig. Teorien må derfor være så konkret at den kan måles, og dermed falsifiseres. Et eksempel kan være: «det kommer til å regne i morgen». Dette utsagnet vil i løpet av morgendagen kunne falsifiseres, hvis det blir oppholdsvær. Om det skulle regne, vil det komme en ny morgendag, der teorien kan falsifiseres. Dette vil skille pseudo-vitenskap og vitenskap i og kalles for Poppers demarkasjonskriterium. Vitenskap vil derfor være alle teorier som kan falsifiseres. Når noe er falsifisert vil det vokse frem en ny falsifiserbar teori, og vitenskapen som en enhet vil etter hvert nærme seg en sannhet. Med dette kriterium mente også popper at induksjonsproblemet også var løst. Dette mente han ved at alle vitenskapelige teorier er deduktive argumenter som falsifiserer, og at vitenskapen dermed ikke behøver bruke induktive argumentasjon. Denne tankerekkefølgen gjør at man ikke kan stole på at noen ting vil være sant, eller ikke engang at det kan argumenteres for at det er sannsynlig å være sant.

Dette er en av punktene som kritiseres av andre filosofer. Altså er det ikke mulig å si at noe er mer sant en annet. Popper kritiseres av andre av at det ikke er forskjell på godt bekreftede og dårlig bekreftede teorier. Det gjør også det mulig å sette opp nye teorier, eller endre teorier sånn at en teori ikke er falsifisert, da man ikke har noen form for hierarki over hvilke teorier som er best bekreftet. Et annet argument som ofte trekkes frem i denne sammenhengen er at det ikke trenger å være et klart skille mellom hva som er ekte vitenskap, og hva som er pseudo-vitenskap. Det vil i mine øyne ikke være så viktig da hvor godt noe er bevist eller bekreftet bør veie tyngre enn at det er formulert en hypotese som kan motsies. I mine øyne er en god vitenskapelig teori noe som er godt bekreftet. Et argument mot denne definisjonen på god vitenskapelig teori, er at den er en lite konkret og veldig vag definisjon, som er vanskelig å standardisere, og dermed dårlig i praksis. Denne definisjonen på en god teori, betyr at i praksis at teorien er bygget i bunn på induktive argumenter. Det er flere forskjellige metoder å benytte seg av induktiv argumentasjon, fire kjente måter er: Hypotetisk deduktiv metode, naiv induktivisme, abduksjon, og bayesianisme

3.2 HD-metoden og abduksjon

En av de som støttet induktiv argumentasjon for vitenskapelig teorier var Carl Hempel (1905- 1997), han var naturlig nok (delte nesten et helt århundre sammen, med motsatte meninger) en av Poppers kritikere. Han kritiserte også

naiv induktivisme, som baserte seg rundt to påstander. Den første er at en teori kan støttes induktivt når den kan trekkes ut fra et utvalg, og den andre er at teorien da bør formuleres etter at alle dataene er innsamlet. Hempel hang seg spesielt opp i det siste og mente at det var umulig å trekke noen konklusjoner før tidens slutt, da det alltid vil dukke opp nye data. Ei heller ville det være mulig å innhente all data som eksisterer i dag. [Hempel, 1966, s. 11] Hempel var av en annen oppfatning av hvordan en teori kunne bekreftes. Hempel mente dette var den hypotetisk deduktive metoden.

Den hypotetisk deduktive metoden baserer seg på fire steg. Det første steget er steget der det dannes eller formuleres en hypotese, teori etc. Dette steget gjøres ved hjelp av et «educated guess», eller «utdannet gjetning» som oversettelsen blir. Det neste steget er deduksjon, som i praksis betyr å utlede empiriske konsekvenser. Her legges grunnlaget for de neste to stegene da disse empiriske konsekvensene deduseres. I steg tre blir disse deduserte empiriske konsekvensene testet ut, som dirkte leder til steg fire. Hvis disse empiriske konsekvensene viser seg å være sanne er denne teorien induktivt bekreftet, eller induktivt styrket bekreftelse av teorien [Hempel, 1966, s.12-13]. Den hypotetisk deduktive metoden kan forklares som en vei med to kjøreretninger, fra data til teori går det induktive argumentet, som er bygd på deduktivt argument som går motsatt vei, altså fra teori til data. Problemet knyttet til Hypotetisk deduktiv metode er at ikke alt kan deduseres, samt at enklere teorier kan passe like bra eller bedre. Om man følger Ockhams barberkniv ønsker man å skjære bort alt unødvendig, da den enkleste teorien som oftest vil være rett.

Abduksjon er en liknende modell til den hypotetisk deduktive metoden, men har et par vir og vendinger som gjør dem klart forskjellige allikevel. Der den hypotetisk deduktive metoden vil undersøke en hypotese kan en ved hjelp av abduksjon undersøke flere teorier. Så lenge teoriene gir en forklaring på fenomenet kan de sammenliknes, og evalueres. Deretter slutte seg til den beste teorien som forklarer et fenomen. Er det ingen annen teori som gir like god eller bedre forklaring, har teorien induktivt styrket bekreftelse eller er bekreftet. Dette kan oppfattes som subjektivt, for hva er egentlig en bedre forklaring når begge teorier forklarer fenomenet. I Abduksjon vil enkelhet tas hensyn til og hvis tre teorier forklarer det samme fenomenet vil det enkleste induktivt bekreftes.

3.3 Replikasjonskrisen

En viktig faktor i vitenskapens unione søken mot et bedret kunnskapsbilde er åpenhet, det skal kunne testes av andre forskere at man får samme data fra samme forsøk. Dette er for å kunne styrke eller svekke et fenomens teori, og kanskje kunne oppdage en ny teori. For at dette skal kunne skje må forskningen være repliserbar. Dette er ikke vært gjennomført godt nok i de siste årene. Et økt fokus har gjort at det har kommet tall på problemene. 1576 forskere fikk

spørsmål om det var en signifikant krise når det kommer til reproduserbarhet i forskning, i en undersøkelse for journalen Nature, der 52% av deltakerne var enige i det [Baker, 2016].

Forklaringen på hva som står bak denne krisen er bestående av en hel haug med små og store bidragsytende faktorer. Den minst beskuende og tillitsfulle forklaringen ligger i statistikken. Sann at replikasjonskrisen kan være et produkt av to statistiske faktorer. De fleste hypoteser vi tester er usanne, og nivået for statistisk signifikans er for lavt ($p < 0,05$). Dette forklarer bird ved hjelp av basefrekvensfeilen. Birds forklaring på replikasjonskrisen er at basefrekvensfeilen vil oppstå når man trekker en slutning på et fenomen basert på en undersøkelse, men ikke hensyntar hvor normalt fenomenet er i utgangspunktet. Han mener dette er en av grunnene for at det er så store problemer med å replisere studier.

Det forekommer tvilsomme forskningspraksiser, som f.eks tuklinger med data, p-hacking, fjerning av materiale. Det er mange små studier, som fører til lav statistisk styrke. Det publiseres ikke like ofte studier med kjedelige eller negative resultater [Bird, 2020]. Det er også kraftig motivasjon for å få frem data med overraskende, eller nyttig resultat, som kan føre til at det blir gjort små ting som gjør at data ikke kan repliseres [Romero, 2017].

For meg høres dette ut som et kombinert problem. Kanskje er verden sann at vi ikke presterer den beste forskningen i ett frittstående konkurransebasert miljø. Det kan være at det er enklere å få navnet opp og frem med noe kontroversielt enn med noe som støtter en allerede eksisterende teori. I det idrettsvitenskaplige fagfeltet er det ofte små utvalg i intervensjoner som gjennomføres, og med relativt kort varighet i forhold til adaptasjonene som skal måles på, som gjør at det krever at mye skal stemme for å få et signifikant resultat, som kan forsterke publikasjonsskjevheten, og dårlig forskningspraksis for å kunne få signifikante resultater.

Kapittel 4

Study Designs

I denne oppgaven om studie design vil jeg bruke QALMRI metoden [Brosowsky et al., 2020] for å bryte ned artiklene fra fem forskjellige originale studier som undersøker det samme i et vitenskapelig område. De utvalgte studiene for denne analysen er [Johnsen and van den Tillaar, 2021], [Brigatto et al., 2019], [Gentil et al., 2018], [Saric et al., 2019], og [Lasevicius et al., 2019].

4.1 Hva undersøker disse studiene?

Den overordnede tematikken i mine utvalgte studier er treningsfrekvens, og hvilke påvirkninger den har på adaptasjonene. Treningsfrekvens beskrives i litteraturen som hvor mange treningsøkter som gjennomføres i en gitt tidsperiode, men også som hvor mange ganger en muskelgruppe blir trent i løpet av en periode eller hvor mange økter som blir gjennomført [KRAEMER and RATAMESS, 2004, Schoenfeld et al., 2016]. Alle disse studiene kontrollerer volum på den nevnte definisjonen.

Studiene utforsket tilpasningene til styrketrening med forskjellig antall treningsøkter med likt ukentlig treningsvolum. De ønsket altså besvare hvilken treningsfrekvens som ga de beste tilpassningene [Johnsen and van den Tillaar, 2021, Brigatto et al., 2019, Gentil et al., 2018, Saric et al., 2019, Lasevicius et al., 2019].

4.2 Hva kan forklare resultatene?

Alle disse studiene hadde en forklaring for hvorfor studien ble gjort formulert som et form for spørsmål, men det var også hypoteser i introduksjonen. I introduksjonen forklares bakgrunnen for studien, og hvilke funn som kan forventes

basert på tidligere fagfellevurdert forskning. For eksempel skrev [Brigatto et al., 2019] at basert på tidligere forskning ville de forvente bedre adaptasjoner i gruppen som trente 2x 8 sett per uke enn i gruppen som trente 1x 16 set per uke. Dette skilte seg fra studiene som sammenlinket høyere frekvenser. [Saric et al., 2019] som sammenlinket 3 vs 6 økter per uke med likt ukentlig treningsvolum, formulerte en hypotese om at det ville være like adaptasjoner, basert på studiens design og tidligere forskning med kontrollert ukentlig volum og liknende protokoller. Dette er en styrke i alle studiene, og de fokuserer på liknende intervensjoner, og forklarer begrensningene når de sammenlikner med mindre liknende protokoller. Det kan også være andre ting en den tenkte årsaksammenheng som kan forklare hypotesen, det kan være bias, sjanse, virkning-årsak, og forvirrende faktorer [Hulley, 2013]

4.3 Hvilken sammenheng har hypotese med eksisterende litteratur i studiene?

Logikken i veien til en hypotese eller spørsmål i disse studiene som undersøker effekten av forskjellige treningsfrekvenser baseres i alle disse studiene på tidligere evidens fra andre fagfellevurderte artikler [Brigatto et al., 2019, Gentil et al., 2018, Johnsen and van den Tillaar, 2021, Saric et al., 2019, Lasevicius et al., 2019]. Selv om disse studiene undersøker det samme overordnede spørsmålet, gjør de metodiske forskjellene i studiene at formuleringen av hypotesene blir forskjellig som vist i avsnittet over.

Studiene som har undersøkt høyere frekvenser opp mot hverandre [Johnsen and van den Tillaar, 2021, Saric et al., 2019] bruker også den samme vitenskaplige evidensen for å komme frem til sin hypotese som studiene som undersøker lavere frekvenser [Brigatto et al., 2019, Gentil et al., 2018, Lasevicius et al., 2019]. En styrke ved Studiene med høyere frekvenser, er at de i tillegg legger vekt på at det er mindre forsket på, og støtter seg derfor på generell evidens om at likt treningsvolum gir like adaptasjoner [Johnsen and van den Tillaar, 2021, Saric et al., 2019].

4.4 Hvilke forskjeller er det i de metodiske valgene i studiene?

Alle disse studiene har en lik type design. Studiene som undersøkes kalles longituduelle studier. Den eksperimentielle tilnærmingen baserer seg på to randomiserte grupper, som følger to forskjellige forutbestemte protokoller for å se om det gir utslag på pre til post test resultater. Randomiserte kontrolerte forsøk som dette kalles, gjøres for å undersøke en bestemt variabel, og er en streng rigid metode for å finne årsak-virkning mellom variabelen og utfallet [Sibbald

4.4. HVILKE FORSKJELLER ER DET I DE METODISKE VALGENE I STUDIENE?19

and Roland, 1998]. Dette gjøres ved at en gruppe mottar en type behandling eller trening, og sammenliknes opp mot en annen gruppe som ofte kalles kontrollgruppe, som enten mottar annen behandling/trening, eller ingenting. Randomiserte kontrollerte studier er dyre, krever mere tid, men gir til gjengjeld, sammen med tidligere evidens, bedre forklaring på årsak-virkning enn kohort studier og case-controll design[Hulley, 2013].

Treningsfrekvensen er variabelen som undersøkes i disse studiene, og basert på tidligere forskning må de ha et likt totalt treningsvolum i begge gruppene, for å kunne undersøke effekten av frekvensen best mulig. Derfor er måten de gjennomfører randomisert kontrollerte studier her, ved å dele en gruppe mennesker fra samme segment, gjennom en like lang treningsintervensjon, med likt ukentlig volum, men forskjellig frekvens. Dette er hensyntatt og er en styrke i alle studiene. Det som varierer i disse studienes metoder er varighet på intervensjon, testmetoder, antall deltakere, frekvens som nevnt tidligere og til en viss grad hvilket samfunns segment forsøkspersonene tilhører.

Alle studiene har beskrivende data på sin deltaker masse. [Saric et al., 2019] sin studie hadde en deltaker masse bestående av 30 styrketrente menn, (gjennomsnitt+Standardavvik) alder = 22.6 +- 2.1 år, høyde = 183,1+-6,0cm, og kroppsvekt = 87,2+-11,6kg. Dette er en styrke med denne studien, da det er et froholdsvist høyt nivå på deltakerne, og et relativt stort utvalg sammenliknet med andre studier. Svakheten med denne studien er at den bare hadde en treningsintervensjon med varighet på 6 uker.

Stuiden til [Saric et al., 2019] undersøkte muskelstyrke ved hjelp av 1RM, muskulær utholdenhet med RM på 60% av 1RM, og muskeltykkelse ved hjelp av ultralyd. De samme testene pre og post ble gjort i studien til [Brigatto et al., 2019], men som testet lavere frekvenser mot hverandre, og med en lengere intervensjon, som er en av styrkene i den studien. Disse studiene undersøkte begge frekvens på styrketrente menn, men med såpass store metodiske variasjoner at hypotesene er forskjellige.

I alle studiene brukes det flere forskjellige statistiske tester. Frekventiske tester er brukt i samtlige, som feks Shapiro-wilk test [Brigatto et al., 2019, Johnsen and van den Tillaar, 2021, Lasevicius et al., 2019], t-test Brigatto et al. [2019], og 2x2 repetert ANOVA [Brigatto et al., 2019, Johnsen and van den Tillaar, 2021] . Den andre måten er å estimere en hypoteses sansynlighet ved hjelp av bayesisk sansynlighet, som i tillegg til frekventistisk statistikk blir gjort i [Saric et al., 2019].

Alle de evaluerte studiene har også beregnet effekt størrelse fra pre-post test resultatene [Brigatto et al., 2019, Gentil et al., 2018, Johnsen and van den Tillaar, 2021, Saric et al., 2019, Lasevicius et al., 2019].

4.5 Hva er resultatene og konklusjonene

I alle studiene besvares hypotesen, men det presiseres også at det trengs mer forskning på området for å kunne konkludere. Det gis også forklaringer på betydningen av metodiske svakheter i studiene som for eksempel små utvalg, og kort varighet som . Disse studiene viser også at den metodiske utformingen er viktig da frekvensens påvirkning kan gi signifikant p-verdi i en type studie, med lavere frekvenser som sammenliknes, men ikke i andre. Derfor er også konklusjonene litt forskjellige, og gjelder kun for frekvensene som sammenliknes, og ikke treningsfrekvens generelt [Johnsen and van den Tillaar, 2021, Brigatto et al., 2019, Gentil et al., 2018, Saric et al., 2019, Lasevicius et al., 2019].

Kapittel 5

Høyere treningsvolum leder til bedre styke og muskelvekst i underekstremiteter hos utrente

5.1 Introduksjon

Styrketrening er anerkjent som en god metode til å forbedre muskelstyrke, evnen til kraftproduksjon, skape muskelhypertrofi, samt forbedre en rekke helsemarkører [Kraemer et al., 2002]. Kunnskap om hvordan man best tilrettelegger styrketrening for å opnå de ønskede adaptasjonene viktig, da styrketrening består av en rekke variabler som kan manipuleres ut ifra behov og gi ønsket adaptasjon [KRAEMER and RATAMESS, 2004, Bird et al., 2005]. Disse variablene som kan påvirke adaptasjoner er volum, frekvens, pauselengde, øvelsesutvalg, og øvelsesrekkefølge [Ame, 2009].

Treningsvolum og den tilknyttede dose-reponsen er således interessant for å øke effektiviteten av trening. De akutte responsene på forskjellige treningsvolum er undersøkt, og myofibrilær muskelproteinsyntesen ble signifikant økt når deltakere gjennomførte 3set på 70% av 1rm, vs 1 sett målt 5 og 29 timer etter trening [Burd et al., 2010]. Når det kommer til lengere varende intervensjonsstudier er det konflikt relatert til treningsvolumets påvirkning. Noen studier har vist signifikant bedre tilpassninger til høyere volum enn lavere volum [RØNNESTAD et al., 2007, STARKEY et al., 1996, Radaelli et al., 2015,]. An-

dre studier har vist at det ikke skiller seg i tilpassninger til styrketrening med forskjellig volum[Bottaro et al., 2011, ?, McBride et al., 2003]. Treningssadaptasjoner er avhengig av en rekke biologiske variabler. Målet med denne studien er å fjerne så mye biologisk påvirkning som mulig og undersøke effekten av volumet alene så godt som mulig.

5.2 Deltakere og studieoversikt

Deltakerne i denne studien bestod av ikke røykende kvinner og menn mellom 18 og 40år. Deltakeren kunne ikke ha trent styrke oftere enn en gang per uke de siste 12 månedene, ikke være inntolerante til lokalbeøvelse, ha en skade som påvirket muskelstyrke negativt, eller ta reseptbelagte legemidler som kan påvirke adaptasjonene til treningen. 7 deltakere som startet på studien kunne ikke gjennomføre tilstrekkelig, og ble ekskludert se tabell 2.

Table 5.1: Deltakeroversikt

	Kvinner		Menn	
	Inkludert	Ekskludert	Inkludert	Ekskludert
N	18	4	16	3
Alder (år)	22 (1.25)	22.9 (1.57)	23.6 (4.11)	24.3 (1.46)
Kroppsvekt (kg)	64.4 (10.4)	64.6 (9.71)	75.8 (10.7)	88.2 (22.4)
Høyde (cm)	168 (6.87)	166 (7.59)	183 (5.88)	189 (4.58)
Kroppsfett (%)	34.1 (5.64)	28.8 (8.69)	20.4 (5.99)	24.3 (15.3)

Verdier er Gjennomsnitt og standard avvik

I denne studiene ble deltakerne testet i pre og post for styrke, og forskjellige mål på muskelvekst. Testene for styrke var isokinetsik og isometrisk unilateral kneekstensjon, og unilateral beinpress. Målene for muskelvekst var måling av kroppsmassesammensetning gjort med DEXA-scan.

5.3 Databehandling

All deskriptiv data er presntert som gjennomsnitt og Standard avvik gjennomsnitt(standard avvik), hvis ikke annet opplyst. P-verdien er beregnet gjennom en ANCOVA modell på endringen fra pre til post i de to gruppene. All databehandling er gjort i Rstudio.

5.4 Treningsprotokoll

Treningsintervensjonen varte i 12 uker, og bestod av 2-3 ukentlige treningsøkter. Øktene var fullkroppss styrkeøkter, der man gjennomførte ett sett på det ene beinet og flere sett på det andre beinet, samt alternerende overkroppssøvelser. Før øktene ble det gjennomført en standardisert oppvarming bestående av armhevinger, sit-ups, rygghev, og knebøy med kroppsvekt, samt 1 sett med 50% av 1RM for hver øvelse.

Underkroppssøvelsene ble gjennomført i følgende rekkefølge: unilateral beinpress, kneekstensjon, og knefleksjon. Etter underkroppssøvelsene ble det gjort bilateral benkpress, nedtrekk, og enten skulderpress eller sittende roing(alternerende annenhver økt). Pauselengde var satt til å være mellom 90 og 180 sekunder. Det var 3 ukentlige økter. Intensiteten ble gradvis økt underveis i treningsintervensjon, 10RM uke 1-2, 8RM uke 3-5, og 7RM for de siste 7 ukene. 91%(10%) av treningsøkterne ble gjennomført under oppsyn av forskningsmedarbeider, men de ble også bedt om å føre detaljert treningslogg for å ha full kontroll. Fra den 9ende treningsøkta ble det innført en økt med redusert belastning hver uke, som bestod av 90% av belastningen på den foregående treningsøkta, men med samme repetisjonsantall. Øktene ed maksimal innsats ble holdt med minst 48timers mellomrom, og de øktene med redusert intensitet ble holdt med inst 24 timers avstand fra de andre øktene

5.5 Testprotokoll

5.5.1 Muskelstyrke

Styrken i isokinetisk og isometrisk unilateral kneekstensjon ble målt i et dynamometer(Cybex 6000, Cybex International, Medway, MA, USA). Deltakerne ble målt i maksimal isokinetisk dreiemoment ved tre forskjellige hastigheter(60°, 120°, og 240° s-1). Etter isokinetisk test, ble maksimal kontraksjonsmoment med 30°knepleson(full ekstensjon beskrives her som 90°).

Maksimal styrke ble målt som 1 repetisjons maks (1RM), i unilateral beinpress, og kneekstensjon. Testingen for hver øvelse startet med en standardisert oppvarming bestående av 10 reps på 50%av estimert 1RM(E1RM), 6repetisjoner på 75% av E1RM, og 3 repetisjoner på 85% av E1RM. Deretter ble 1RM funnet ved at vekten ble økt frem til de ikke lengere kunne gjennomføre repetisjonen med full bevegelsesutslag. Forsøkspersonene ble gitt 4-6 forsøk på å finne 1RM.

På alle styrketester ble pre test gjennomført to ganger for alle forsøkspersoner, med minimum 4 dager mellom hver gang. Det maksimale resultatet fra de to testene av hver øvelse ble brukt i analysen. Resultatene fra muskelstyrke i pre test(1RM, isokinetisk, og isometrisk) ble slått sammen til et kombinert resultat. Det kombinerte resultatet bestod av en gjennomsnittelig verdi av alle styrketestene, hvor hver test ble gitt lik betydning.

5.5.2 Mager masse

For å måle mager masse ble kroppsmassesammensetning målt i dual-energy X-ray absorptiometry (DXA) (Lunar Prodigy, GE Healthcare, Oslo, Norway), både pre og post treningsintervensjon. Deltakerne ble bedt om å faste 2 timer, og avstå fra intensiv fysisk aktivitet i 48 timer i forkant av DXA.

5.6 Resultater

12 uker med styrketrening resulterte i en signifikant økning i muskelstyrke (kombinert resultat) hos begge grupper (fig 5.2) hhv. 31 ± 14 (14.20)% for multiple- sett ($P=0.001$) og 25 ± 13 (12.90)% for multiple- sett ($P=0.001$) (fig. 5.3). Forskjellen i styrkeøkningen mellom gruppene var signifikant på ($P=0.02988$). Endringen i mager masse var også signifikant, hhv. $3,3 \pm 4\%$ for multiple- sett ($P=0.001$) og $2 \pm 4\%$ for single- sett ($P=0.001$) men her var det ikke signifikant forskjell mellom gruppene ($P=0.193$, fig 5.1).

5.7 Diskusjon

Hovedfunnene i denne studien er at høyere volum gir bedre effekt enn lavere volum for muskelstyrke, og samsvarer bra med tidligere studier med liknende treningsprotokoller [RØNNESTAD et al., 2007, STARKEY et al., 1996, Radaelli et al., 2015]. Endringer i muskelstyrke er vist i en metaanalyse bestående av 61 randomiserte kontrollerte studier, med varighet på minimum 4 uker å ha et gradert dose-respons forhold [Ralston et al., 2017].

I denne studien får deltakerne trent i snitt 15,6 sett per muskelgruppe i låret (quadriceps, og hamstring) per uke på beinet som trener flere sett per øvelse, og 5,2 sett på beinet som trente ett sett per øvelse. Dermed kunne man kanskje tenke seg at det også ville skille signifikant mellom protokollene i endringene av mager masse, som vist i en metaanalyse der de fant at hvert ekstra sett ledet til 0,36% bedre adaptasjoner [Schoenfeld et al., 2016]. Det er derfor viktig å huske at deltakerne her er utrente, og resultatene av ett sett kan være velig gode, og gi ikke signifikante forskjeller [Bottaro et al., 2011, ?, McBride et al., 2003]. Det kunne derfor vært interessant å se på godt trente personer gjennomføre nøyaktig samme protokoll, og gi mere konkret evidens til å si noe om et gradert dose respons forhold.

5.7.1 Konklusjon

I denne studien ser vi at det er mest effektivt å trene med høyere treningsvolum for styrke i underkroppen, men at det ikke er nødvendig for muskelveksten. Det

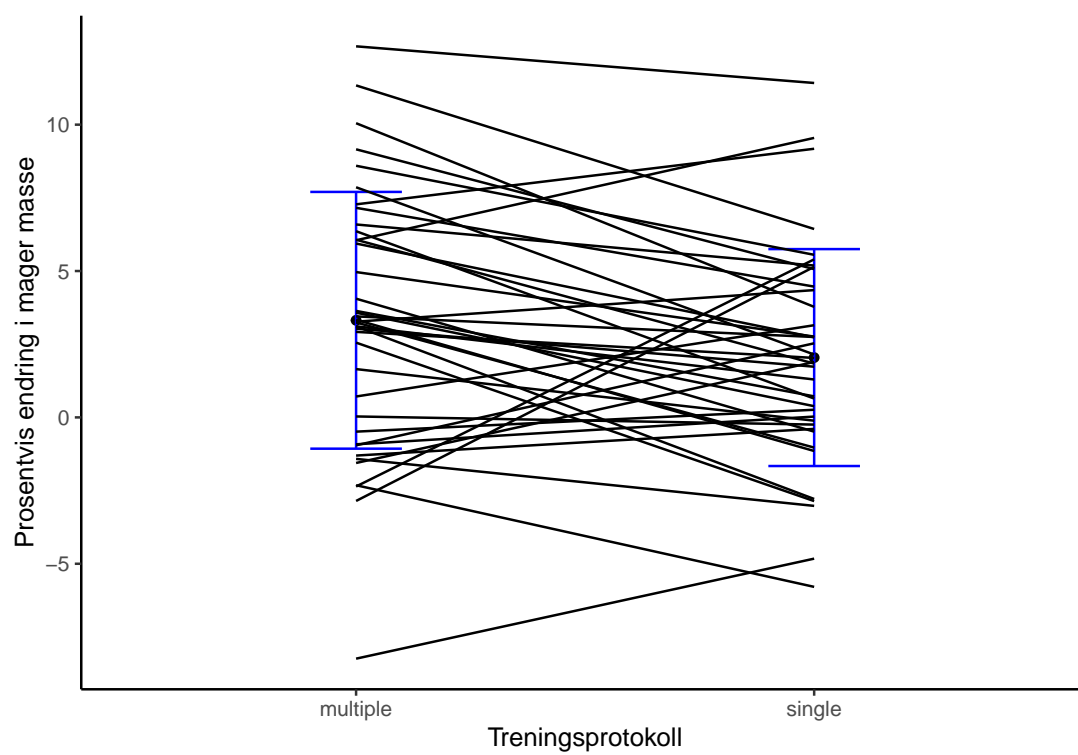


Figure 5.1: Figuren viser prosentvis økning i mager masse fra pre til post mellom multiple- sett og single- sett
(#fig:Prosentvis økning figur dexa)

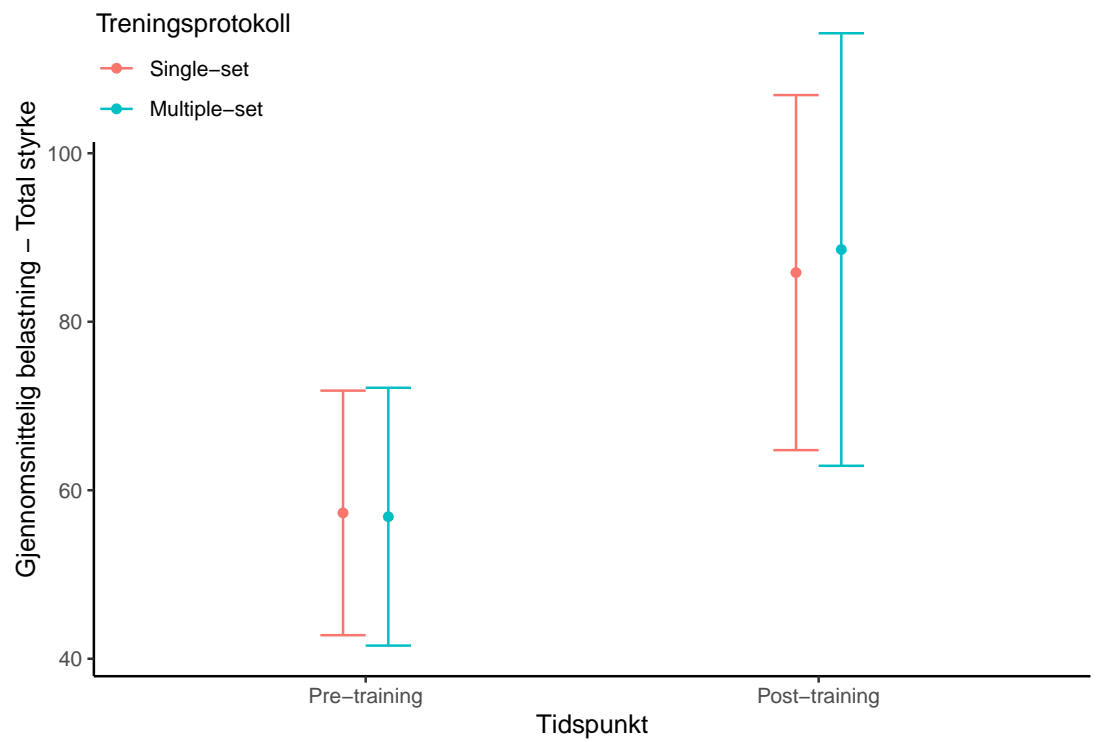


Figure 5.2: Figuren viser økning i muskelstyrke mellom multiple- sett og singel-sett fra pre til post test (diff= -0.029AU)

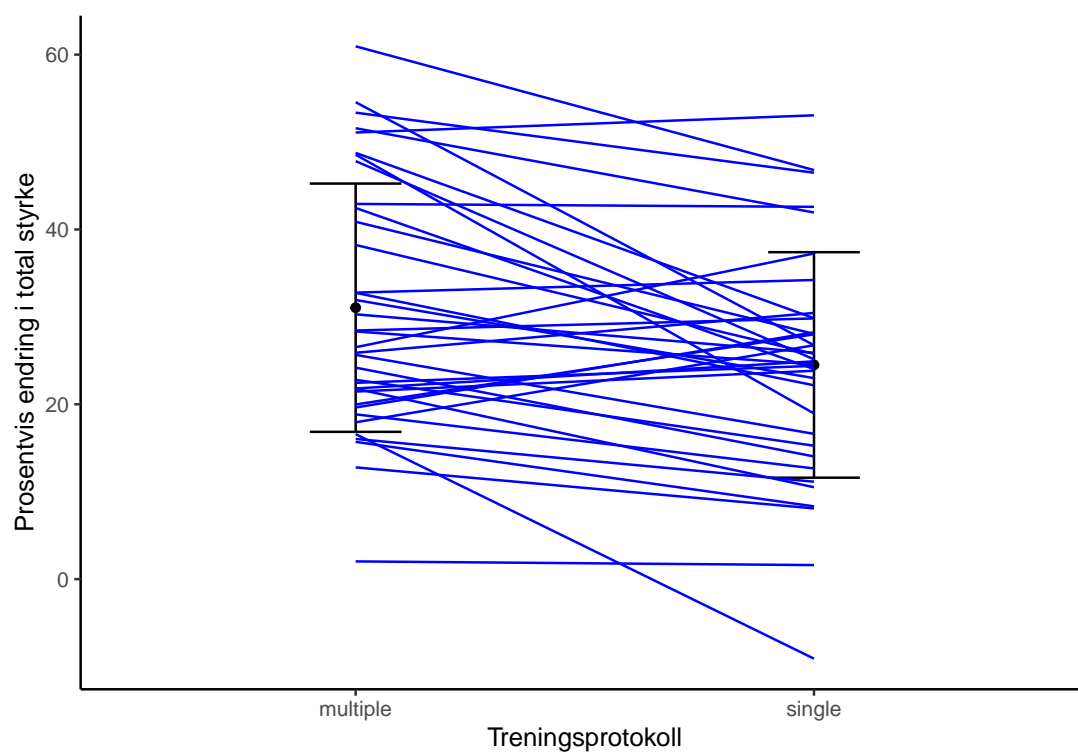


Figure 5.3: Figuren viser prosentvis endring i muskelstyrke mellom multiple-sett og singe- sett fra pre til post test
(#fig:Prosentvis økning figur styrke)

kan tenkes at for å oprettholde denne utviklingen, må etterhvert treningsvolumet økes fra singel-sett protokollen til noe liknende flere.-sett protokollen følger, basert på det graderte doserespons forholdet.

Bibliography

- Progression models in resistance training for healthy adults. *Medicine & Science in Sports & Exercise*, 41(3):687–708, 03 2009. doi: 10.1249/mss.0b013e3181915670. URL <http://dx.doi.org/10.1249/MSS.0b013e3181915670>.
- Monya Baker. 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604): 452–454, 05 2016. doi: 10.1038/533452a. URL <http://dx.doi.org/10.1038/533452a>.
- D. R. Bassett and E. T. Howley. Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Medicine and Science in Sports and Exercise*, 32(1):70–84, 01 2000. doi: 10.1097/00005768-200001000-00012. PMID: 10647532.
- Alexander Bird. Understanding the replication crisis as a base rate fallacy. *The British Journal for the Philosophy of Science*, pages 000–000, 12 2020. doi: 10.1093/bjps/axy051. URL <http://dx.doi.org/10.1093/bjps/axy051>.
- Stephen P Bird, Kyle M Tarpenning, and Frank E Marino. Designing resistance training programmes to enhance muscular fitness. *Sports Medicine*, 35(10): 841–851, 2005. doi: 10.2165/00007256-200535100-00002. URL <http://dx.doi.org/10.2165/00007256-200535100-00002>.
- G. A. Borg. Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise*, 14(5):377–381, 1982. PMID: 7154893.
- M. Bottaro, J. Veloso, D. Wagner, and P. Gentil. Resistance training for strength and muscle thickness: Effect of number of sets and muscle group trained. *Science & Sports*, 26(5):259–264, 11 2011. doi: 10.1016/j.scispo.2010.09.009. URL <http://dx.doi.org/10.1016/j.scispo.2010.09.009>.
- Felipe A. Brigatto, Tiago V. Braz, Thamires Cristina da Costa Zanini, Moisés D. Germano, Marcelo S. Aoki, Brad J. Schoenfeld, Paulo H. Marchetti, and Charles R. Lopes. Effect of resistance training frequency on neuromuscular performance and muscle morphology after 8 weeks in trained men. *Journal of Strength and Conditioning Research*, 33(8):2104–2116, 08 2019. doi: 10.1519/jsc.0000000000002563. URL <http://dx.doi.org/10.1519/jsc.0000000000002563>.

- Nicholaus Brosowsky, Olga Parshina, Anthony Locicero, and Matthew John Charles Crump. Teaching undergraduate students to read empirical articles: An evaluation and revision of the qalmri method, 06 2020. URL <http://dx.doi.org/10.31234/osf.io/p39sc>.
- Nicholas A. Burd, Andrew M. Holwerda, Keegan C. Selby, Daniel W. D. West, Aaron W. Staples, Nathan E. Cain, Joshua G. A. Cashaback, James R. Potvin, Steven K. Baker, and Stuart M. Phillips. Resistance exercise volume affects myofibrillar protein synthesis and anabolic signalling molecule phosphorylation in young men. *The Journal of Physiology*, 588(16):3119–3130, 08 2010. doi: 10.1113/jphysiol.2010.192856. URL <http://dx.doi.org/10.1113/jphysiol.2010.192856>.
- Sally A. Clark, P. C. Bourdon, W. Schmidt, B. Singh, G. Cable, K. J. Onus, S. M. Woolford, T. Stanef, C. J. Gore, and R. J. Aughey. The effect of acute simulated moderate altitude on power, performance and pacing strategies in well-trained cyclists. *European Journal of Applied Physiology*, 102(1):45–55, 12 2007. doi: 10.1007/s00421-007-0554-0. PMID: 17882451.
- Paulo Gentil, James Fisher, James Steele, Mario H. Campos, Marcelo H. Silva, Antonio Paoli, Jorgen Giessing, and Martin Bottaro. Effects of equal-volume resistance training with different training frequencies in muscle size and strength in trained men. *PeerJ*, 6:e5020, 06 2018. doi: 10.7717/peerj.5020. URL <http://dx.doi.org/10.7717/peerj.5020>.
- Israel Halperin, David B. Pyne, and David T. Martin. Threats to internal validity in exercise science: a review of overlooked confounding variables. *International Journal of Sports Physiology and Performance*, 10(7):823–829, 10 2015. doi: 10.1123/ijsp.2014-0566. PMID: 25756869.
- Carl G. Hempel. *Philosophy Of Natural Science*. Prentice Hall, 1966.
- A. V. Hill and H. Lupton. Muscular exercise, lactic acid, and the supply and utilization of oxygen. *QJM*, os-16(62):135–171, 01 1923. doi: 10.1093/qjmed/os-16.62.135. URL <https://academic.oup.com/qjmed/article-lookup/doi/10.1093/qjmed/os-16.62.135>.
- W. G. Hopkins. Measures of reliability in sports medicine and science. *Sports Medicine (Auckland, N.Z.)*, 30(1):1–15, 07 2000. doi: 10.2165/00007256-200030010-00001. PMID: 10907753.
- Stephen B. Hulley. *Designing clinical research*. Wolters Kluwer/Lippincott Williams and Wilkins, 4 edition, 2013.
- Emil Johnsen and Roland van den Tillaar. Effects of training frequency on muscular strength for trained men under volume matched conditions. *PeerJ*, 9:e10781, 02 2021. doi: 10.7717/peerj.10781. URL <http://dx.doi.org/10.7717/peerj.10781>.

- WILLIAM J. KRAEMER and NICHOLAS A. RATAMESS. Fundamentals of resistance training: Progression and exercise prescription. *Medicine and Science in Sports and Exercise*, 36(4):674–688, 04 2004. doi: 10.1249/01.mss.0000121945.36635.61. URL <http://dx.doi.org/10.1249/01.mss.0000121945.36635.61>.
- William J. Kraemer, Nicholas A. Ratamess, and Duncan N. French. Resistance training for health and performance. *Current Sports Medicine Reports*, 1(3):165–171, 06 2002. doi: 10.1249/00149619-200206000-00007. URL <http://dx.doi.org/10.1249/00149619-200206000-00007>.
- Jujiao Kuang, Xu Yan, Amanda J. Genders, Cesare Granata, and David J. Bishop. An overview of technical considerations when using quantitative real-time pcr analysis of gene expression in human exercise research. *PLOS ONE*, 13(5):e0196438, 05 2018. doi: 10.1371/journal.pone.0196438. URL <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0196438>. Publisher: Public Library of Science.
- Thiago Lasevicius, Brad Jon Schoenfeld, Jozo Grgic, Gilberto Laurentino, Lucas Duarte Tavares, and Valmor Tricoli. Similar muscular adaptations in resistance training performed two versus three days per week. *Journal of Human Kinetics*, 68(1):135–143, 08 2019. doi: 10.2478/hukin-2019-0062. URL <http://dx.doi.org/10.2478/hukin-2019-0062>.
- Jeffrey M. McBride, John B. Blaak, and Travis Triplett-McBride. Effect of resistance exercise volume and complexity on emg, strength, and regional body composition. *European Journal of Applied Physiology*, 90(5-6):626–632, 11 2003. doi: 10.1007/s00421-003-0930-3. URL <http://dx.doi.org/10.1007/s00421-003-0930-3>.
- Karl R. Popper. *Conjectures and refutations: the growth of scientific knowledge*. Routledge classics. Routledge, London ; New York, 2002.
- Regis Radaelli, Steven J. Fleck, Thalita Leite, Richard D. Leite, Ronei S. Pinto, Liliam Fernandes, and Roberto Simão. Dose-response of 1, 3, and 5 sets of resistance exercise on strength, local muscular endurance, and hypertrophy. *Journal of Strength and Conditioning Research*, 29(5):1349–1358, 05 2015. doi: 10.1519/jsc.0000000000000758. URL <http://dx.doi.org/10.1519/JSC.0000000000000758>.
- Grant W. Ralston, Lon Kilgore, Frank B. Wyatt, and Julien S. Baker. The effect of weekly set volume on strength gain: A meta-analysis. *Sports Medicine*, 47(12):2585–2601, 07 2017. doi: 10.1007/s40279-017-0762-7. URL <http://dx.doi.org/10.1007/s40279-017-0762-7>.
- Eileen Y. Robertson, Philo U. Saunders, David B. Pyne, Robert J. Aughey, Judith M. Anson, and Christopher J. Gore. Reproducibility of performance changes to simulated live high/train low altitude. *Medicine and Science in Sports and Exercise*, 42(2):394–401, 02 2010. doi: 10.1249/MSS.0b013e3181b34b57. PMID: 19927018.

- Felipe Romero. Novelty versus replicability: Virtues and vices in the reward system of science. *Philosophy of Science*, 84(5):1031–1043, 12 2017. doi: 10.1086/694005. URL <http://dx.doi.org/10.1086/694005>.
- BENT R. RØNNESTAD, WILHELM EGELAND, NILS H. KVAMME, PER E. REFSNES, FAWZI KADI, and TRULS RAASTAD. Dissimilar effects of one- and three-set strength training on strength and muscle mass gains in upper and lower body in untrained subjects. *Journal of Strength and Conditioning Research*, 21(1):157–163, 02 2007. doi: 10.1519/00124278-200702000-00028. URL <http://dx.doi.org/10.1519/00124278-200702000-00028>.
- Juraj Saric, Domagoj Lisica, Ivan Orlic, Jozo Grgic, James W. Krieger, Sasa Vuk, and Brad J. Schoenfeld. Resistance training frequencies of 3 and 6 times per week produce similar muscular adaptations in resistance-trained men. *Journal of Strength and Conditioning Research*, 33(1):S122–S129, 07 2019. doi: 10.1519/jsc.0000000000002909. URL <http://dx.doi.org/10.1519/JSC.0000000000002909>.
- P. U. Saunders, R. D. Telford, D. B. Pyne, A. G. Hahn, and C. J. Gore. Improved running economy and increased hemoglobin mass in elite runners after extended moderate altitude exposure. *Journal of Science and Medicine in Sport*, 12(1):67–72, 01 2009. doi: 10.1016/j.jsams.2007.08.014. PMID: 18069063.
- Brad J. Schoenfeld, Dan Ogborn, and James W. Krieger. Effects of resistance training frequency on measures of muscle hypertrophy: A systematic review and meta-analysis. *Sports Medicine*, 46(11):1689–1697, 04 2016. doi: 10.1007/s40279-016-0543-8. URL <http://dx.doi.org/10.1007/s40279-016-0543-8>.
- B. Sibbald and M. Roland. Understanding controlled trials: Why are randomised controlled trials important? *BMJ*, 316(7126):201–201, 01 1998. doi: 10.1136/bmj.316.7126.201. URL <http://dx.doi.org/10.1136/bmj.316.7126.201>.
- DAVID B. STARKEY, MICHAEL L. POLLOCK, YOSHI ISHIDA, MICHAEL A. WELSCH, WILLIAM F. BRECHUE, JAMES E. GRAVES, and MATTHEW S. FEIGENBAUM. Effect of resistance training volume on strength and muscle thickness. *Medicine & Science in Sports & Exercise*, 28(10):1311–1320, 10 1996. doi: 10.1097/00005768-199610000-00016. URL <http://dx.doi.org/10.1097/00005768-199610000-00016>.
- Espen Tønnessen, Erlend Hem, Ida Svendsen, Eirik V. Larsen, Martin Skau-gen, and Elisabeth Solbakken. Utholdenhetstester ved olympiatoppen. 2017. URL https://www.olympiatoppen.no/fagomraader/trening/testing/testing_av_utholdenhet/media53703.media.