

# Mappeeksamen

Håvard Crantz Lorentzen, kandidatnummer 107

2021-12-03



# Contents

<b>1</b>	<b>Realabilitet</b>	<b>5</b>
1.1	Introduksjon . . . . .	5
1.2	Metode . . . . .	5
1.3	Resultater . . . . .	6
1.4	Diskusjon . . . . .	7
<b>2</b>	<b>Labrapport - cDNA syntesering ved hjelp av Superscript IV og generell qPCR</b>	<b>9</b>
2.1	Formål . . . . .	9
2.2	Metode . . . . .	9
2.3	Resultater . . . . .	10
2.4	Diskusjon . . . . .	11
<b>3</b>	<b>Vitenskapsteori</b>	<b>13</b>
3.1	1. Falsifikasjonisme . . . . .	13
3.2	2. HD-metoden og abduksjon . . . . .	14
3.3	3. Replikasjonskrisen . . . . .	15
<b>4</b>	<b>Studiedesign</b>	<b>17</b>
4.1	Introduksjon . . . . .	17
4.2	Diskusjon av metode . . . . .	18
4.3	Resultat og anbefaling . . . . .	21
4.4	Videre forskning . . . . .	21

<b>5</b>	<b>1 vs 3 sett</b>	<b>27</b>
5.1	Introduksjon . . . . .	27
5.2	Metode . . . . .	28
5.3	Resultat . . . . .	30
5.4	Diskusjon . . . . .	32

All data og analyse som brukt i mappeeksamen finnes her: [https://github.com/Havlor/Mappeeksamen\\_2021](https://github.com/Havlor/Mappeeksamen_2021)

# 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 (eller 1.05 om laktatprofiltest er gjennomført i forkant) og blodlaktat over 8 [Tønnessen et al., 2017].

### 1.2 Metode

11 deltagere (tabell 1) gjennomførte en 10 minutter lang oppvarmingsprotokoll på tredemøllen (Woodway 4Front, Wisconsin, USA) testen skulle fåregå på. Denne oppvarmingsprotokollen bestod av fem minutter på 11-13 i Borg skala (6-20) [Borg, 1982]. Etterfulgt var det 2x1 minutter på starthastighet og startstigning med 30 sekund pausene mellom de to dragene. Siste tre minutt av oppvarmingen var også 11-13 i borg, men på valgfri stigning og hastighet. Etter oppvarming var det to minutter pause før testen begynte. Starthastighet var satt til 8km/t, med stigning på 10.5% og 5.5% for henholdsvis menn og kvinner.

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

( $\text{ml kg}^{-1} \text{ min}^{-1}$ ) . 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.

VO<sub>2</sub>max ble målt ved hjelp av en metabolsk analysator med miksekammer (vyntus CPX, mixingchamber (Vyntus CPX, Jaeger-CareFusion, UK)). Forut for alle tester ble analysatoren kalibrert for gass og volum. Analysatoren ble stilt inn til å gjøre målinger hvert 30. sekund, og VO<sub>2</sub>max ble kalkulert gjennom å bruke snittet av de to høyeste påfølgende målingene av O<sub>2</sub>. Underveis i testen mottok alle deltakerne 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 støy [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: Tabellen viser relevant informasjon om deltagerne

	Kvinner	Menn
Antall	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

I figur 1 kan man se forskjellen mellom test 1 og 2 fordelt på kjønn. Typefeilen (typical error, [Hopkins, 2000]) fra test 1 til test 2 er utregnet til å være 4.04%.

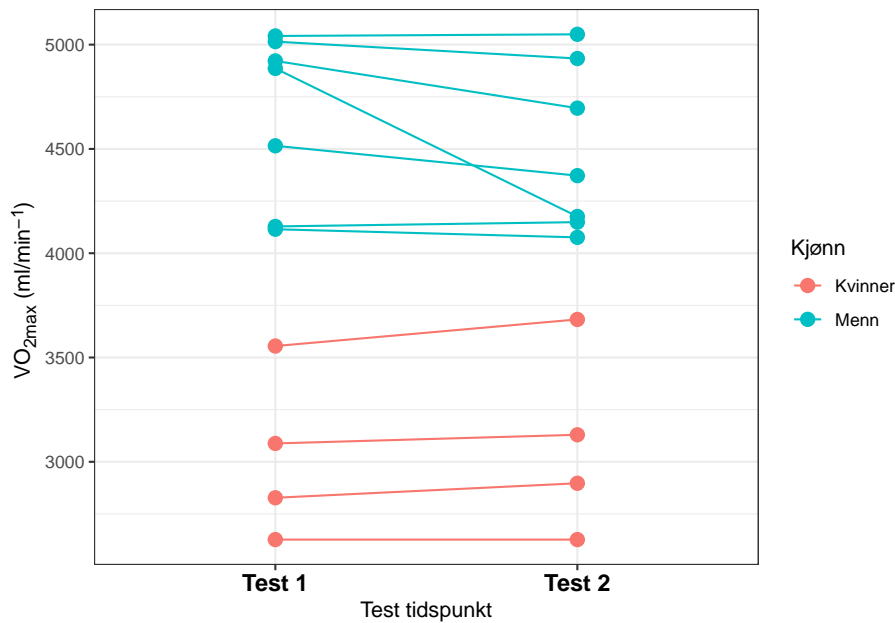


Figure 1.1: Figuren viser forskjell i Vo2max (ml/min) mellom test 1 og 2

## 1.4 Diskusjon

Resultatet i vår realibilitetstudie er funnet av en typefeil på 4.04. Typefeilen kan også tyde på at enkelte av disse resultatene kan være utsatt for støy av ulik sort [Hopkins, 2000]. 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 testingen. 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 [Halperin et al., 2015]. Faktorer som muligens påvirket testene våre var deltakernes kjennskap til testen, verbal oppmuntring og personer som var tilstede under testen [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 testdag 1. Dette fordi kjennskapen og kunnskapen de tilegnet seg på den første testen, trolig spiller inn på testresultatene.

Grunnen til at vi snakker om typefeil på en test er at vi ønsker å måle forskjellen av to tester. Når utøvere ønsker å ta VO2maks-test for å se effekt av trening, er det viktig å vite hva som er effekt av støy og hva som er effekt av trening. Desto mindre støy en test innebærer jo bedre er målingen. Hva som danner denne variasjonen som representeres ved typefeil er multifaktorelt, men hoveddelen er som oftest biologisk [Hopkins, 2000].

For å måle typefeil har vi brukt “within subject deviation metoden” [Hopkins, 2000]. 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 standardfeil [Halperin et al., 2015]. Data fra det Australske institutt for sport har også fastslått at en standardfeil 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 på 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 syntesering ved hjelp av Superscript IV og generell 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 fem folds fortynningsserie 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 mastermixer 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/referansegen, 2 µl DEPC-behandlet vann, 2 µl fortynnet cDNA. Deretter ble det fylt 71 brønner i en

qPCR-reaksjonsplate med henholdsvis 2  $\mu$ l prøve, og 8  $\mu$ l med mastermix. Reaksjonsplaten med brønner ble dekt med plastfilm, og ble sentrifugert i 1 minutt på 1200 omdreininger, 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 (figur 1). 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) (figur 1). 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. Tabell 1 viser resultatet av både referansegener og målgener.

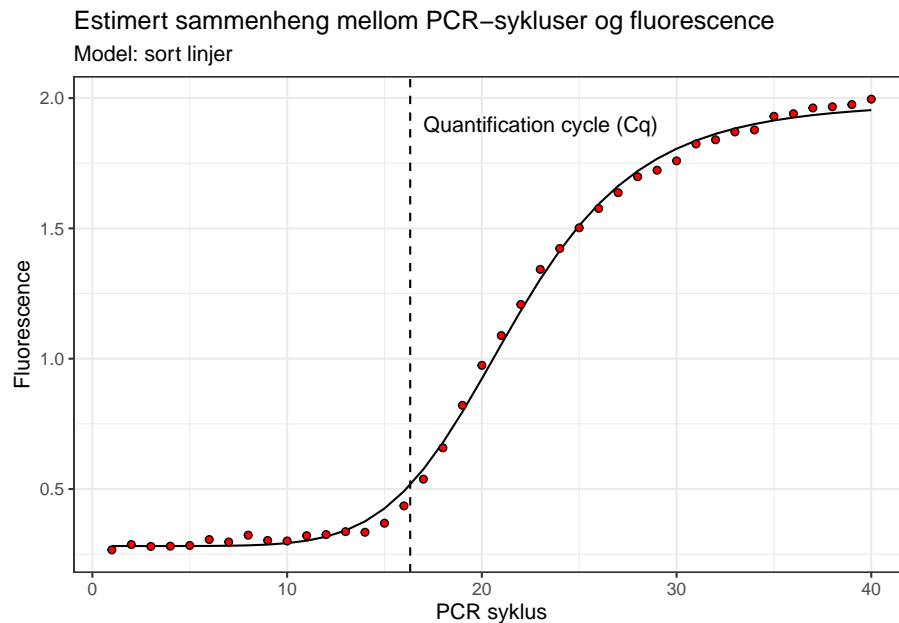


Table 2.1: Tabellen viser cq-verider per gen

sample	time	MYHC1	MYHC2A	MYHC2X	RRNA47S	CHMP2A	REEP5	B
FP1	w0	19.53	20.03	22.87	24.88	26.68	26.68	23
FP2	w0	19.70	20.04	29.80	27.54	26.78	26.59	25
FP3	w0	20.33	18.25	22.87	25.27	25.97	26.55	24
FP1	w2	19.96	17.59	26.01	32.36	26.73	26.87	24
FP2	w2	20.20	14.64	26.23	27.00	26.45	26.40	24
FP3	w2	21.57	23.34	25.95	25.67	23.79	26.17	25

## 2.4 Diskusjon

Cq-verdien sier noe om hvor mange PCR-sykluser som trengs for å detektere ulike gen [Kuang et al., 2018]. En høyere Cq-verdi indikerer altså at mengden RNA må dobles flere ganger for å detektere en terskelverdi av et gen. En lavere Cq-verdi indikerer at terskelverdien oppnås ved færre PCR-sykluser, altså at konsentrasjonen av målgen er høyere [Kuang et al., 2018]. En lavere Cq-verdi ved uke 2, sammenlignet med uke 0, 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.

I forsøket vårt har vi valg tre referansegen, dette gjør det lettere å se at resultatet stemmer. Referansegene er er gen som ikke skal være påvirkbare av den gitte intervensjoenen og skal være være en sikkerhet på at Cq-verdiene på målgenene stemmer.



## Kapittel 3

# Vitenskapsteori

### 3.1 1. Falsifikasjonisme

Når man skal diskutere falsifikasjon, er det flere spørsmål som er aktuelle, hva er vitenskap? er alt som blir publisert vitenskap, og hvor går i så fall skillet mellom vitenskap og «ikke-vitenskap»? Sistnevnte er kjent som demarkasjonsproblemet.

Spørsmålene som jeg stiller over er spørsmål som har vært diskutert av vitenskapsfilosofer lenge og spesielt de siste hundre årene [Okasha, 2016]. Karl Raimond Popper har hatt og har stor innflytelse i disse spørsmålene etter at han kom med sin teori om falsifikasjonskriteriet [Okasha, 2016]. Denne teorien sier at en vitenskapelig teori må kunne være mulig å falsifisere (motbevise) [Popper, 2002]. Dermed kommer Popper med en løsning på Demarkasjonsproblemet, vitenskap er teorier som teoretisk sett skal være mulig å motbevise [Popper, 2002]. En teori som ikke kan motbevises er da ikke-vitenskap, eller som Popper ville sagt, pseudovitenskap [Popper, 2002]. Popper løfter fram Astrologi som et eksempel på pseudovitenskap, da det ikke er mulig å falsifisere [Popper, 2002]. Astrologer baserer seg på empiriske observasjoner og danner horoskoper med vage teorier sånn at de kan bortforklare et hvert angrep mot læren [Popper, 2002]. Dette blir på mange måter det motsatte av det Popper ønsker [Popper, 2002]. En forsker skal heller sette seg teorier for å så ha som mål å falsifisere de, på denne måten vil man sikre god vitenskap [Popper, 2002]. Når man prøver å falsifisere en teori vil det være gjennom deduksjon, teorien har et premiss som empirisk blir testet og man får svaret sant eller usant. Utfallet blir da enten at teorien blir falsifisert og da lagt fram som motbevist eller at teorien blir stående. Ifølge Popper kan ikke teorier bevises, og da er argumentasjon basert på induksjon heller aldri et alternativ [Okasha, 2016]. Samtidig mener Popper at teorier som har blitt testet på riktige måter og som ikke har blitt falsifisert kan bli «corroborated», altså nesten bevist [Popper, 2002]. Poppers tanker om disse

teoriene som flere ganger ikke er blitt falsifisert ligner da veldig på en induktiv bekreftelse av en teori. Dette er altså en teori som mest sannsynlig er sann, men ifølge Popper er det ikke det samme som induktiv bekreftelse [Okasha, 2016, Popper, 2002].

Problemet med Poppers teori om vitenskap kommer tydelig fram når han ikke kan bekrefte vitenskapelige teorier. Samir Okasha får dette fram i sin bok om vitenskapsfilosofi [Okasha, 2016]. Okasha løfter fram et eksempel om personer med Downs Syndrom: Genforskere har funnet ut at alle som har Down Syndrom har tre kopier av kromosomnummer 21 istedenfor to [Okasha, 2016, s. 18]. Genforskere har da basert på et stort utvalg med personer som har Down Syndrom gjort en konklusjon som gjelder alle med Down Syndrom [Okasha, 2016]. Dette er da en induktivt bekreftet teori som er godt dokumentert og som gjør at man kan bygge videre kunnskap basert på teorien. Popper kunne ikke gjort dette da det ikke er mulig å gjøre så sterke påstander [Okasha, 2016]. På bakgrunn av dette skulle man ikke trenge å svare på demarkasjonsproblemet og heller skille mellom styrket og mindre styrket teorier. Okasha mener at en forskers rolle ikke bare handler om å finne ut om en teori er usann, men også å finne hvilke teorier som er sanne eller mest sannsynlig sanne, og da trenger man induksjon [Okasha, 2016, s. 19-20].

## 3.2 2. HD-metoden og abduksjon

I den første oppgaven så jeg på hva en vitenskapelig teori er, nå skal jeg se på hvordan teorier kan testes. Carl Gustav Hempel mente at det ikke fantes noen god metode teorier kunne testes og var uenig med Popper [Hempel, 1966]. Hempel mente at man trengte induktiv argumentasjon i vitenskapene [Hempel, 1966]. Derfor løftet Hempel fram en metode kaldt hypotetisk deduktiv metode (HD-metoden), som et bedre alternativ [Hempel, 1966]. Denne metoden sier ikke noe om hvordan teorier kommer fram, men den har klare retningslinjer på hvordan man skal teste en teori og gi den styrke [Hempel, 1966].

For å forklare HD-metoden bruker jeg følgende eksempel som teori: Høyintensiv intervalltrening på løp øker det maksimale oksygenopptaket ( $VO_{2maks}$ ). Ifølge HD-metoden skal man etter at en har utarbeidet en teori tenke ut empiriske konsekvenser til teorien [Hempel, 1966]. I vårt tilfelle er det da å finne hvilke empiriske konsekvenser det er av å få et høyere  $VO_{2maks}$ . Man skulle tro at empiriske konsekvenser av høyere  $VO_{2maks}$  er et høyere blodvolum og bedre prestasjon på fem minutter løpstest. Neste steg er å teste om de empiriske konsekvensene er sanne, dette skjer gjennom deduktiv testing [Hempel, 1966]. Dette betyr at økt blodvolum og bedre prestasjon på fem minutters løpstest blir satt som premiss for at teorien. Om blodvolumet da har økt, og prestasjonen på fem minutters løpstest har blitt bedre, gir det en styrke til teorien om at høyintensiv intervalltrening gir høyere  $VO_{2maks}$ . I vanlig deduktiv argumentasjon vil et rett premiss bety at teorien er sann, men i HD-metoden vil teorien bare bli en

grad av induktivt bekreftet [Hempel, 1966]. HD-metoden vil aldri si at en teori er bekreftet, da det ikke er mulig å vite om alle argument som kan forklarer teorien [Hempel, 1966]. Man kan også tenke seg at det er observasjoner eller resultat som kommer i fremtiden og som motsier det en har trodd før [Hempel, 1966]. I vårt eksempel kan en se for seg at høyintensitets intervalltrening under noen forhold ikke gir høyere blodvolum eller bedre prestasjon på fem minutters løpsted. HD-metoden er alltid åpen for at det er andre forklaringer og sier derfor bare noe om en teoris styrke eller svakhet ut fra det man har testet empirisk [Hempel, 1966].

Problemet med HD-metoden er at den empiriske testingen er bundet til teorien. Det vil si at man ikke kan endre teorien, selv om resultatet av en test ender opp med å passe bedre til en annen teori. Charles Sanders Peirce sin teori om abduksjon i tillegg til induksjon og deduksjon kan løse dette problemet [Peirce et al., 1992]. Abduksjon handler om at man setter flere teorier, for å deretter forklare teoriene og teste de deduktivt (som i HD-metoden) [Peirce et al., 1992]. Resultatet fra testene vil da fortelle oss hvilken teori som er sann basert på hvilken teori som gir det beste svaret på testene [Peirce et al., 1992]. En teori som forklarer mest mulig vil være å foretrekke da man i abduksjon ønsker å ha en teori som er bedre enn andre tilgjengelig teorier [Peirce et al., 1992]. Abduksjon blir derav også kaldt «slutning til den beste forklaringen» fordi en teori skal gi den beste mulige forklaringen på et fenomen [Peirce et al., 1992].

På mange måter ligner HD-metoden og abduksjon på hverandre og begge vil gå under «kvalifisert gjetning» [Persson, 2019]. Det som i hovedsak skiller abduksjon fra HD-metoden er fleksibiliteten og den praktiske tilnærmingen da det er flere teorier å velge mellom [Peirce et al., 1992].

### 3.3 3. Replikasjonskrisen

De siste 20 årene har det blitt identifisert en replikasjonskrise i vitenskapen [Begley and Ellis, 2012, Ioannidis, 2005, Open Science Collaboration, 2015]. Dette handler om at mange studier som blir gjort på nytt ikke klarer å få det samme resultatet som første gangen det ble gjort. Det gir oss et stort problem, for hvordan kan man da stole på vitenskapen? Undersøkelser viser at tillitten til vitenskapen har falt både i England og USA [Funk and Kennedy, 2016]. Dette er også noe flertallet av forskere anerkjenner som en krise basert på en spørreundersøkelse gjort på 1500 forskere [Baker, 2016]. Alexander Bird mener å ha en forklaring til denne replikasjonskrisen [Bird, 2020].

Bird mener at feilen ligger i en neglisjering av basefrekvens [Bird, 2020]. Vitenskapen konkluderer rett og slett alt for fort uten å ta hensyn til basefrekvens [Bird, 2020]. Sann som vitenskapen er i dag er det typisk gjort en randomisert kontrollert undersøkelse og resultatet er enten signifikant eller usignifikant basert på P-verdien [Bird, 2020]. Denne signifikantgrensen som ofte blir satt til 0,05 (5%) har da blitt en slags pekepinn på om noe er sant eller usant basert

på om resultatet er over eller under grensen [Bird, 2020]. P-verdien sier noe om sannsynligheten for at resultatet er falskt positivt (type-I-feil) [Bird, 2020]. Om P-verdien er 0,05 er det en 5% sjanse for at resultatet er falskt positiv [Bird, 2020]. Problemet med dette er at P-verdien man får etter et forskningsprosjekt bare sier noe om utvalget som er sett på i den gitte situasjonen [Bird, 2020]. Ved å bruke basefrekvens vil man se på resultatet basert på populasjonen og hvor sannsynlig det egentlig er for at resultatet er sant [Bird, 2020]. Et eksempel på dette, som Bird [2020] tar fram, er sannsynligheten for at en tilfeldig har en sjelden sykdom om vedkommende tester positivt på en test. Sykdommen hadde en forekomst på 1 til 1000 og testen hadde en pålitelighet på 95% [Bird, 2020]. Det er lett å tenke at sannsynligheten er 95%, men i realiteten er den 2% [Bird, 2020]. Dette kommer av at man må regne inn forekomsten i regnestykke gjennom Bayes teorem [Bird, 2020]. På samme måte som at det er lett å trekke konklusjonen om at personen i eksempelet over, kan vi dra konklusjoner om at vitenskapelige hypoteser er sanne [Bird, 2020]. Dette gjør at mange studier faktisk ikke gir et rett svar, og da ikke er repliserbare [Bird, 2020].

Andre forklaringer som Bird trekker fram er den statistiske styrken, som i mange tilfeller ikke er høy nok [Bird, 2020]. Styrken blir i stor grad styrt av forskningens størrelse (antall deltagere) og gir en sannsynlighet for å ikke få et falskt negativt resultat (type-II-feil) [Bird, 2020]. Bird mener at dette ikke er en god forklaring da en høy statistisk styrke egentlig fortsatt gir relativt høy sannsynlighet for å få en type-II-feil [Bird, 2020]. Bird viser til en utregning hvor det er høyere statistisk styrke enn i «vanlig» vitenskap, men at det fortsatt er 31% sjanse for å få type-II-feil [Bird, 2020, s. 12-14]. Det vil være bra for vitenskapen å få opp den statistiske styrken, men kan ikke forklarer krisen som vitenskapen står i [Bird, 2020].

Juks og dårlig praksis i vitenskapen vil kunne forklare noe av replikasjonskrisen, men det er vanskelig å sette en konklusjon basert på det man har sett av studier så langt [Bird, 2020]. Det er derfor ikke gode nok holdepunkt til å mene at det er dårlig moral rundt vitenskapen som er hele forklaringen på krisen [Bird, 2020].

Basert på det Bird [2020] skriver, er det tydelig at det er hans egen forklaring som gir mest styrke til teorien om at vitenskapen er i en krise. Det er tydelig at folk som driver vitenskap trenger en større forståelse av statistikk og sannsynlighet [Bird, 2020]. Gjennom et større hensyn til basefrekvensen vil man utarbeide bedre teorier og være mer forsiktig med å si at noe er gjeldene for en hel populasjon [Bird, 2020].



## Kapittel 4

# Studiedesign

### 4.1 Introduksjon

Mennesker i vesten lever lenger og lenger, og eldrebølgen er et faktum. på bakgrunn av dette er det derfor et stort behov for å finne ut hvordan eldre mennesker kan bevare helsen best mulig. I denne rapporten skal jeg se på fem studier som har sett på styrketrening som en metode for å bedre helsen til eldre mennesker [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Alle fem studiene ønsker å finne ut hvordan man kan legge opp styrketrening som gir god effekt på helsen (mentalt og fysisk) og dermed gjøre det lettere å gi gode treningsanbefalinger. En oppsummering av studiene er vist i tabell 1.

I fire av studiene ble det sett på forskjellige måter man rent praktisk bør gjennomføre styrketrening [Schott et al., 2019, ; Turpela et al., 2017; Vikberg et al., 2019; Vincent et al., 2002]. I Schott et al. [2019] ble det for eksempel sett på effekten av å trene med styrketreningsapparater mot frivektstrening [Schott et al., 2019]. Studien til Geirsdottir et al. [2012] skiller seg fra de andre da den hovedsakelig så på den generelle effekten av trening [Geirsdottir et al., 2012]. Den så som de andre på styrke- og muskeladaptasjoner, men inkluderte et spørreskjema om hvordan treningen påvirket livskvaliteten [Geirsdottir et al., 2012].

På bakgrunn av studiespørsmålene og svarene de vil gi oss, vil studiene gi en styrke til den allerede etablerte anbefalingen om at eldre bør trene styrke for å opprettholde eller bedre helsen [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al. [2002]]. I tillegg vil studiene gjøre det lettere å gi gode anbefalinger da de løser noen praktiske spørsmål rundt selve gjennomføringen av styrketrening [Geirsdottir et al., 2012, ; Schott et al., 2019; Turpela et al., 2017; Vikberg et al., 2019; Vincent et al., 2002]. Vikberg et al. [2019] foreslår at en enkel form for styrketrening som man kan gjøre hjemme er nok for å styrke muskulatur og bedre helsen. Dette vil

gjøre styrketreningseffekten oppnåelig for alle, og ikke bare de som trener på treningssenter [Vikberg et al., 2019].

## 4.2 Diskusjon av metode

Fire av studiene var av studietypen Randomisert kontrollert undersøkelse (RCT) hvor det var en, to eller tre intervensjonsgrupper og en kontrollgruppe [Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Hvem som var i hvilke grupper, ble randomisert tilfeldig. Antallet i gruppene var nesten likt fordelt i intervensjon og kontroll med unntak av Vincent et al. [2002]. «Nesten likt» vil si at det stort sett var flere i de gruppene det krevde mest av (treningsgruppene). Samtidig var dette gruppene med flest frafall, sånn at det ved studieslutt nesten var utlignet [Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019]. I Vincent et al. [2002] var det større forskjell på kontroll og intervensjon (16 i kontroll, 24 og 22 i intervensjon) uten at det blir opplyst noen grunn til det. Å ha en blokkrandomisering som de tre andre nevnt over, ville gitt studien et styrket design. Når det er forskjellige størrelser vil det bli vanskeligere å sammenligne, da mindre grupper gir bredere konfidensintervall og lavere statistisk styrke [Hulley, 2013, s. 63, 146].

Da dette er studier som ser på trening er «blinding» vanskelig, dermed nevnte studiene heller ikke noe om dette [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. For å unngå forskningsfusk er det likevel gunstig at testpersonell ikke vet hvem som er i hvilken gruppe, da testpersonell skal gjennomføre alle tester like godt, selv om man ofte ønsker et resultat i favør hypotesen [Hulley, 2013, s. 148-149].

Geirsdottir et al. [2012] var en klinisk studie uten kontrollgruppe, de hadde altså en stor gruppe (265 kvinner og menn) som trente samme styrketreningssprogrammet. At de ikke hadde en kontrollgruppe innrømmet studiegruppen selv at var en svakhet [Geirsdottir et al., 2012]. Samtidig argumenterer de for at eldre mennesker ikke kan forvente en økning i styrke, muskelmasse og muskelfunksjon uten å trene [Geirsdottir et al., 2012]. På bakgrunn av dette mente de at det var nok å sammenligne med testresultat ved start [Geirsdottir et al., 2012]. Det er mulig å forstå argumentasjonen til Geirsdottir et al. [2012] når de i tillegg har såpass mange deltagere som de har. Samtidig hadde de definert kontrollgruppe og et RCT-design gjort studien sterkere som en studie som ønsker å finne ut av noe [Hulley, 2013, s. 87]. Studien vil gå under studiekategorien Case-serie som er mer egnet til å se på karakteristika, og i denne sammenhengen av styrketrening på eldre [Hulley, 2013, s. 87]. Dette kommer frem da resultatet av spørreskjema om livskvalitet var noe av det viktigste, men burde kommet bedre frem når de diskuterer de andre variablene [Geirsdottir et al., 2012].

Populasjonen som er ønsket å treffe er relativt friske eldre kvinner og menn [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg

et al., 2019, Vincent et al., 2002]. Ingen av studiene har en eksplisitt definisjon av «eldre», men totalt i studiene var det et aldersspenn på 60-92 og gjennomsnittsalderen i de forskjellige studiene var rundt 70år (flest rett under) [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Den eldre befolkningen er en svært heterogen gruppe, noe som gjør det vanskelig med ekskluderingskriterier. Studiene godtok da det som ble definert som milde sykdommer som ikke påvirker treningen [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vincent et al., 2002]. Unntaket er til en viss grad Vikberg et al. [2019] hvor det den eneste eksklusjonskriteriet var sarkopeni. Samtidig kan det tyde på at det var relativt god helse blant deltagerene da de alle var 70 år, ikke hadde sarkopeni og klarete å gjennomføre treningen [Vikberg et al., 2019]. Totalt sett treffer studiene godt på populasjonen da alle har et høyt antall deltagere og har kvalifikasjonskriterier som gjør det mulig for «vanlig» eldre folk å være med [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002].

Rekrutteringen av deltagere ble gjort ved hjelp av brev og annonser med unntak av Schott et al. [2019; Geirsdottir et al., 2012; Turpela et al., 2017; Vikberg et al., 2019; Vincent et al., 2002]. Han hentet deltagere fra en treningsgruppe som var etablert på studiestedet [Schott et al., 2019]. Felles for alle studiene var at det var en frivillig deltagelse [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Schott et al. [2019] og Vincent et al. [2002] uttrykte et krav om at det måtte være statistisk styrke på 80%. Schott et al. [2019] gjorde også et estimat hvor de så at det var behov for 26 deltagere for å se en signifikant forskjell ved signifikantgrense på 5% når de godtok en sannsynlighet på 80% (de inkluderte 32 personer). De andre studiene nevnte ikke noe om statistisk styrke, og kunne med fordel vist til en estimering av deltagere som krevdes for en gitt statistisk styrke. Turpela et al. [2017] innrømmet at det kunne bli vanskelig å se forskjell mellom gruppene da utrente ofte får god effekt uansett. Dette problemet kunne blitt kontrollert for om man hadde gjort en powerberging og inkludert tilstrekkelig deltagere [Hulley, 2013, s. 48-49] .

Alle fem studiene hadde pre-post-design hvor Scott et al. også i tillegg hadde tester i uke 10/26 og 6 uker etter post-testen i uke 26 [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. I alle studiene ble det gjort idrettsfysiologiske tester av styrke, men også funksjonelle tester som hadde som mål å se hvordan styrken kom til uttrykk i dagligdagse gjøremål (gå test, gripe test osv.) [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Geirsdottir et al. [2012] la som nevnt også en del vekt på spørreskjema som ble gitt før og etter intervensjon. Schott et al. [2019] hadde også et spørreskjema, men dette ble bare gitt etter og var en evaluering av treningen. Treningen som ble gjennomført av intervensjonsgruppene ble gjort på treningssenter eller treningslokale på studiestedet [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. I studien til Vincent

et al. [2002] sendte deltagerne inn treningsdagbøker for hver økt som det så det ble gitt tilbakemeldinger på. I de fire andre studiene var det kvalifiserte trenere som veiledet deltagerne igjennom øktene [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019]. Kontrollgruppen ble instruert i å leve som vanlig og ikke trene [Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Treningsperiodene var fra 10 uker til 26 uker [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002].

Studiene hadde som nevnt som mål å finne ut hvordan eldre bør trene og hvilke effekter det er av styrketrening [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. For å finne ut dette ble det sett på forskjellige variabler for styrke, kroppssammensetning (spesielt fettfri masse), muskelstørrelse, men også variabler som skulle reflektere dagligdagse gjøremål [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Vincent et al. [2002] la for eksempel en del vekt på en trappetes hvor deltagerne skulle gå en trapp så fort som mulig. Dette er viktig da noe av hovedgrunnen for å trene styrke for eldre nettopp er å mestre dagligdagse gjøre mål [Vincent et al., 2002].

I analysen av studiene var det ønskelig å finne forskjeller i resultat mellom grupper, her ble det brukt forskjellige statistiske tester. I tre av studiene hvor det var en eller to grupper ble brukt parett t-test får så se forskjell innad i gruppene [Geirsdottir et al., 2012, Vikberg et al., 2019, Schott et al., 2019]. Vincent et al. [2002] brukte 3\*2 ANOVA for å se forskjell innad og mellom gruppene og såg etter en interaksjon mellom gruppe og tid. Vincent et al. [2002] fant ikke forandring mellom gruppene på pretest, men gjorde likevel en ANCOVA for å kontrollere for forskjeller ved pre-test når de analyserte forandringen mellom gruppene. To av de andre studier gjorde også ANCOVA for å se forandring mellom gruppene og effekten av tid ved gjentatte målinger [Schott et al., 2019, Vikberg et al., 2019]. Turpela et al. [2017] brukte her en-veis-ANOVA for å kontrollere for ulikheter ved pre-test. Denne kontrollen for pre-verider gir styrke til resultatet i studiene da den kontrollerer for eventuelle forskjeller i gruppene ved pre-test. For Geirsdottir et al. [2012] er ikke denne analysemetoden mulig da det bare er en gruppe. Schott et al. [2019] brukte også MANCOVA for å gjøre statistikk på variablene som ble målt flere enn 2 ganger og på spørreskjemaet han brukte. Geirsdottir et al. [2012] brukte Wilcoxon's test for analyse av spørreskjemaet sitt. Tre av studiene brukte også post hoc tester som vil gi et mer troverdig resultat da den kontrollerer resultatet for falsk positiv (type I feil) [Hulley, 2013, s. 50-52][Schott et al., 2019, Turpela et al., 2017, Vincent et al., 2002]. Alle hadde signifikantgrense ved  $P = 5\%$  med unntak av en ANCOVA-test Vikberg et al. [2019] gjorde som gjorde en interaksjon mellom kjønn og intervensjon, hvor  $P = 10\%$  [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vincent et al., 2002]. Denne interaksjonen til Vikberg et al. [2019] vil med dette lettere godta at det er en forskjell mellom kjønn.

### 4.3 Resultat og anbefaling

Tre av studeiene studiene, som sammenlignet intervensjon mot kontroll, ble det ikke sett forskjell på muskelmasse, men Vikberg et al. [2019] fant forskjell i muskelmasse [Schott et al., 2019, Turpela et al., 2017, Vincent et al., 2002]. Vikberg et al. [2019] fant ikke forskjell i styrke fra kontrollgruppen, men de gjorde de tre andre [Schott et al., 2019, Turpela et al., 2017, Vincent et al., 2002]. På funksjonell styrke gav alle studiene positiv effekt på funksjonelle tester [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. I de studiene som sammenlignet flere treningsprotokoller var det liten forskjell mellom det forskjellige treningsprotokollene, men alle skilte seg fra kontrollgruppen [Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Geirsdottir et al. [2012] fant en liten forbedring av livskvalitet. Alt i alt greide de delvis å gi et positivt svar på hypotesen [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002].

Alle studiene konkludere med at styrketrening fungerte [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. For å oppsummere anbefalingene, er det nok å trene styrke 1-2 ganger i uken, så lenge det er helkroppssøkt [Turpela et al., 2017, Vikberg et al., 2019]. En ytterlig effekt vil man få om man trener med frivekter, spesielt på ben og triceps, som er viktige muskelgrupper for eldre [Schott et al., 2019]. Det er også nok å gjøre 13 rep på 50% 1RM [Vincent et al., 2002]. Antall serier er ikke sett på, men basert på studiene burde 1-2 være nok [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017]. Det kommer også fram at all trening som ble sett på gav en effekt [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Treningen vil vedlikeholde muskelmassen din, øke styrken din, gjøre praktiske gjøremål lettere og det vil kunne øke livskvaliteten din [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002]. Denne konklusjonen og anbefalingen gjelder for relativt friske eldre mennesker som ikke er vant til å trene og som har en lav muskelmasse, dog ikke kritisk lav [Geirsdottir et al., 2012, Schott et al., 2019, Turpela et al., 2017, Vikberg et al., 2019, Vincent et al., 2002].

### 4.4 Videre forskning

Det er enda ikke veldig tydelig hvilke treningsprogram som gir god effekt på bedret funksjonsevne [Turpela et al., 2017]. Schott et al. [2019] foreslår at studeier med frivekter kan være løsningen. I tillegg burde man gjøre studier hvor man kontrollerer proteininntaket hos eldre da dette kan være så lavt at det påvirker styrkefremgangen [Turpela et al., 2017].

Table 4.1: Tabellen viser en oppsummering av studiene

Studie	Spørsmål	Hypoteser	Logikk	Metode	Resultat	Interferes
Turpela et al. 2017	Hvordan påvirker treningsfrekvensen muskelstyrke, muskelmasse og funksjonell kapasitet hos utrente eldre mennesker?	To til tre treninger i uken er bedre enn en trening.	Det vil bli enklere å gi klare råd om hvor ofte eldre bør trene styrke for å oppnå treningseffekt.	4 grupper: ingen økter, 1 økt, 2 økter og 3 økter. Intervensjonsperioden på 6 måneder hvor tester ble gjort før og etter. Treningsøktene var tradisjonell tung og eksplosiv trening. 106 friske kvinner og menn (64-75 år). Ble brukt en-veis-ANOVA og en-veis-ANOVA med gjentatte målinger.	Ikke sammenheng mellom treningsfrekvens og ganghastighet eller kroppssammenheng. Alle skilte seg fra kontrollgruppen Dose-respons i forhold til 1RM og dynamisk styrke (i treningsapparat)	Sunne eldre vil ha en gunstig effekt av å trene lavfrekvens helkroppsstyrketrening (1-2 ganger i uken).

Studie	Spørsmål	Hypoteser	Logikk	Metode	Resultat	Interferes
Schott et al. 2019	Er det forskjell på treningseffekt mellom frivektstrening og apparatstrening hos eldre mennesker?	Begge gruppene vil ha en fremgang i styrke, men gruppen som trener med frivekter vil få et bedre resultat sammenlignet med gruppen som trente med treningsapparat.	Vil gjøre det lettere   å gi treingsanbefalinger for godt fungerende eldre	32 friske kvinner og menn mellom 60 og 86 år ble fordelt i to grupper (frivekter eller apparat). To treninger i uken med 3 sett og 10-12 repetisjoner (10RM). Testing av styrke ved start, 10 uker, 26 uker og 6 uker etter intervensjon. Spørreskjema om treningen etter intervensjon. Statistikk ble gjort ved uavhengig t-test, parett-test, MANCOVA, Cohens d og Pearsons korrelasjon.	Begge gruppene økte i alle øvelser og gjorde det bedre på dynamiske styrketester. For triceps og ben var det større effekt av å trene med frivekter. Frivektstrening var mer motiverende.	Trening med frivekter og maskin gir god styrkeeffekt, men frivektstrening gir ytterligere effekt på ben og triceps. Disse to muskelgruppene er viktige for å forebygge fall og i dagliglivet, frivektstrening er derfor å anbefale for eldre som er godt fungerende.

Studie	Spørsmål	Hypoteser	Logikk	Metode	Resultat	Interferes
Vikberg et al. 2019	Gir trening med enkle øvelser effekt på funksjonell styrke og kroppssammensetning?	Styrketrening vil gi positiv effekt på funksjonell styrke	Eldre kan på en enkel måte få god effekt av styrketrening.	70 kvinner og menn med pre-sarkopeni ble fordelt i kontroll (36) og intervensjonsgruppe (34). Intervensjonsgruppen trente tre gruppetreninger i uken i 10 uker. Helkroppsprogram med RPE mellom 6 og 7 (skala 1-10). Det ble gjort et tester før og etter. Paret T-test ble brukt for å se forskjell innad i gruppene. ANCOVA ble brukt for å se forskjell mellom gruppene.	Det var økning i styrke fra post til pre hos intervensjonsgruppen, men de skilte seg ikke fra kontroll. Kroppssammensetning var bedret innad i gruppen og i forhold til kontroll.	Styrketrening med enkelt utstyr som man kan gjøre nesten hvor som helst er effektivt for å hindre tap av muskelstyrke og øke muskelmasse for eldre med pre-sarkopeni



Studie	Spørsmål	Hypoteser	Logikk	Metode	Resultat	Interferes
Geirsdottir et al. 2012	Hva er effekten av trening på eldre mennesker? Og er det relatert til bedret livskvalitet, helse kroppssam- ning og funksjonell styrke?	Styrketrening vil gi god effekt på muskel- massen og styrken som igjen vil øke livskvalitet.	Styrketrening vil gi høyere livskvalitet hos eldre	265 friske kvinner og menn mellom 65 og 92 år trente i 12 ukers med tre økter i uken. Øktene var helkropp- sprogram (3 sett, 6-8 rep på 75-80% av 1RM per øvelse). Mot- standen ble økt 5-10% hver uke. Tester og spør- reskjema ble gjort før og etter. Paret t-test ble brukt for å se på forskjell mellom pre og post. Wilcoxon test ble brukt for å se på pre og post spør- reskjema. Persons og Spear- mans r ble brukt for å se på kor- relasjoner.	Det var en økning i muskel- masse, styrke og funksjonell kapasitet. Det var en liten, men signifikant økning i livskvalitet.	12 uker med styr- ketrening vil gi god effekt på styrke, kroppssam- ning og muskel- masse. Dette vil igjen gi bedre livskvalitet. Ved styr- ketrening vil en også hindre fallet i muskel- styrke og muskel- masse.

Studie	Spørsmål	Hypoteser	Logikk	Metode	Resultat	Interferes
Vincent et al. 2002	Er det forskjell på treningseffekten av å trene høyere vekt og lavere antall repetisjoner og lavere vekt og flere repetisjoner hos eldre mennesker?	Styrketrening med lavere vekt og flere repetisjoner fungerer like bra som høy vekt og få repetisjoner.	Man kan anbefale trening som oppleves lettere og mindre skadeutsett.	62 friske kvinner og menn mellom 60 og 83 år fordelt i to grupper (trening eller ikke trening) i 26 uker. De som trente, trente enten 50% av 1RM *13 rep eller 80% av 1RM 8 rep i en serie per treningsapparat (volum blir ca. det samme for begge). Ble gjort tester før og etter. ANOVA ble brukt for å se forskjell innad og mellom gruppene over tid. Pre og post-testene ble gjort i ANCOVA.	Begge treningsgruppene fikk økning i muskelstyrke og muskelutholdenhet, samt i trappetesten i forhold til kontroll, men ingen forskjell mellom treningsgruppene.	Eldre friske mennesker kan få like godt resultat med 13 rep på 50% av 1 RM, 8 rep på 80% av 1RM. Dette gjør styrketrening enklere i tillegg til lavere skaderisiko.

# Kapittel 5

## 1 vs 3 sett

### 5.1 Introduksjon

Det finnes utallige metoder å trene styrketrening på, men hvilken metode som er den mest effektive er enda uklart. Tradisjonelt gjøres styrketrening i  $X$  antall ganger (repetisjoner) i  $X$  antall serier (sett). Men hvor mye har det å si at man gjør ett eller flere sett? Flere studier har sett på effekten av å gjøre ett sett og tre sett [Galvão and Taaffe, 2005, Hass et al., 2000, Krieger, 2009, Radaelli et al., 2014, Schoenfeld et al., 2019]. Resultatet i studiene er sprikende. Noen studier finner ikke forskjell mellom grupper som trener ett og tre sett [Hass et al., 2000, Radaelli et al., 2014], mens andre studier ser at begge gruppene øker, men tre sett gir en ytterlige effekt på styrke [Krieger, 2009, Galvão and Taaffe, 2005]. Det er også sett at man kan oppnå like styrkeeffekter, men at økningen av muskelmasse trenger større volum enn ett sett for å gi effekt [Schoenfeld et al., 2019].

Basert på statistikk fra Statistisk sentralbyrå [2019] er styrketrening blitt svært populært, men bare ett fåtall trener mer enn en gang i uken. Studier viser derimot at den ideelle effekten av styrketrening kommer først når man trener to til tre ganger i uken [Schoenfeld et al., 2016]. På bakgrunn av dette vil det være gunstig å finne ut av spørsmålet om ett sett er nok, da en treningsøkt tar mye kortere tid og lettere å prioritere. Noe av det som gjør at det er vanskelig å dra en tydelig konklusjon av studiene på området er at det vil være individuelle variasjoner mellom to treningsgrupper. I vår studie ser vi derfor på personer som trener ett sett på ene benet og tre sett på det andre benet. På denne måten vil en sikre at framgangen ikke blir forskjellig på bakgrunn av gruppeforskjeller, men på bakgrunn av selve treningen.

I studien ønsker vi å se effekten av både fem uker og 12 uker. Hypotesen er at begge treningsmetodene vil ha en god framgang i både styrke og muskelstørrelse,

men at benet som trener tre sett vil få en bedre styrkefremgang ved både fem og 12 uker.

## 5.2 Metode

### 5.2.1 Etikk

Alle deltagere i studien ble informert om potensielle risikoer ved trening og testing samt eventuelle ubehagelige og anstrengende situasjoner. Dette gav de et informert samtykke på. Alle prosedyrer er i tråd med Helsinkideklarasjonen.

### 5.2.2 Deltagere

Det ble rekruttert 41 kvinner og menn til en treningsperiode på 12 uker. Kriterier for å være med var at de måtte være mellom 18-40 år og ikke røyke. De var ikke vant til å trene (ble ekskludert om de hadde mer enn en økt i uken det siste året). I analysene av styrkefremgang ble bare 29 deltagere inkludert, de resterende møtte ikke møtte opp på test ved fem uker. 34 deltagere ble analysert for muskelmasseøkning (her var det bare test ved pre og post), resterende gjennomførte ikke studien (fem stk pga. smerter, en stk. pga skade som ikke hadde med studien å gjøre og en pga. ikke fulgte protokoll).

### 5.2.3 Trening

Trening som ble gjennomført var et helkroppstreningsprogram hvor treningen av ble gjort forskjellig på høyre og venstre fot. De ene bene trente ett sett, mens det andre trente tre sett. Hvilket ben som trente hva, ble randomisert. Antall repetisjoner i uke 0-2 var 10 repetisjoner maksimum (RM), uke 2-5 var det 8RM, uke 5-12 var det 7RM. Det ble gjort tre treninger i uken, med unntak av ukene hvor det ble gjort test. Alle økter ble gjort på ved kvalifiserte trenere (91% av øktene, resterende økter ble gjort uten trener for at det skulle bli gjennomførbart). Øvelsene på bena ble gjort i denne rekkefølgen: unilateral benpress, kneleksjon og kneekstensjon med enten ett sett eller tre sett. Det ene settet ble gjort mellom andre og tredje sett i tre sett serien. Etter benøvelsene ble øvelser på overkroppen gjort i to sett: bilateral benkpress, nedtrekk og enten skulderpress eller sittende roing (gjort annenhver økt). Pauser mellom sett var 90-180 sekund og pause mellom økten var på minimum 24 timer.

## 5.2.4 Testing

### 5.2.4.1 Styrketester

Styrketestne ble gjort ved pre, etter fem uker og ved post (12 uker). Tre av testene ble gjort som unilateral kneekstensjon i dynamometer (Cybex 6000, Cybex International, Medway USA). I forkant av testingen ble en standardisert oppvarming gjennomført på ergometersykkel i fem minutter. Dynamometeret ble justert av testleder sånn at individuelle innstillinger ble gjort. Før hver test ble det gjort standardiserte oppvarmingsrepetisjoner i dynamometeret. Isokinetisk (isok) og isometrisk (isom) styrke ble målt ved maksimalt dreiemoment målt i newtonmeter, og ble notert for hver test (isok120, isok240 og isok60 grader per sekund bevegelseshastighet og isom0 (kneet i 60 graders vinkel)). Det ble gitt to forsøk ved isok60 og isom60 og tre forsøk ved isok240 og isok120 og høyeste ble registrert.

De to andre testene var 1RM-tester i unilateral kneekstensjon og benpress. Som oppvarming ble det gjort ti, seks og tre repetisjoner ved henholdsvis 50, 75 og 85% av forventet 1RM. Deretter ble 1RM funnet ved at vekten ble økt gradvis til de ikke klarte å fullføre full bevegelse i øvelsen. Den høyeste gjennomførte repetisjonen ble satt som 1RM.

Resultatet ble regnet om til en kombinert score som er et gjennomsnitt av alle styrketestene.

### 5.2.4.2 Estimering av Muskelmasse

Ved pre og post-test ble det brukt «dual-energy X-ray absorptiometry» (DXA)(Lunar Prodigy, GE Healthcare, Oslo, Norway) for å estimere muskelmasse i gram. Deltagerne fikk beskjed om å faste de to timene før test og å unngå fysisk aktivitet 48 timer før.

## 5.2.5 Statestikk

For å se forskjell mellom ett og tre sett ble den kombinerte scoren i styrke regnet om til prosentvis økning. For å finne signifikante forskjeller ble ANCOVA brukt, hvor det ble kontrollert for pre-verdier og tatt hensyn til at dataene er korrelerte. Dette er da man gjør to forskjellige protokoller på samme person. For muskelmasse ble det sett på forskjell mellom pre og post. Endring i muskelstyrke ble analysert ved pre til fem uker, fem uker til post, og pre til post. En P-verdi under 0,05 blir sett på som signifikant endring. ANCOVA-modellen ble gjort kombinert med R-pakken lmerTest [Kuznetsova et al., 2017]. Alle tall er gitt som gjennomsnitt med standardavvik. Tabeller, figurer og analyser ble gjort i RStudio (versjon 1.4.1717; R Foundation for Statistics Computing, Vienna AT).

## 5.3 Resultat

### 5.3.1 Muskelmasse

Benet som trente tre sett økte signifikant mer enn benet som trente ett sett, med henholdsvis økning på 3.37% ( $\pm 4.59\%$ ) og 2.05% ( $\pm 3.62$ ) ( $p < 0.05$ ) (figur 1). Det var ingen forskjell mellom benene ved pre.

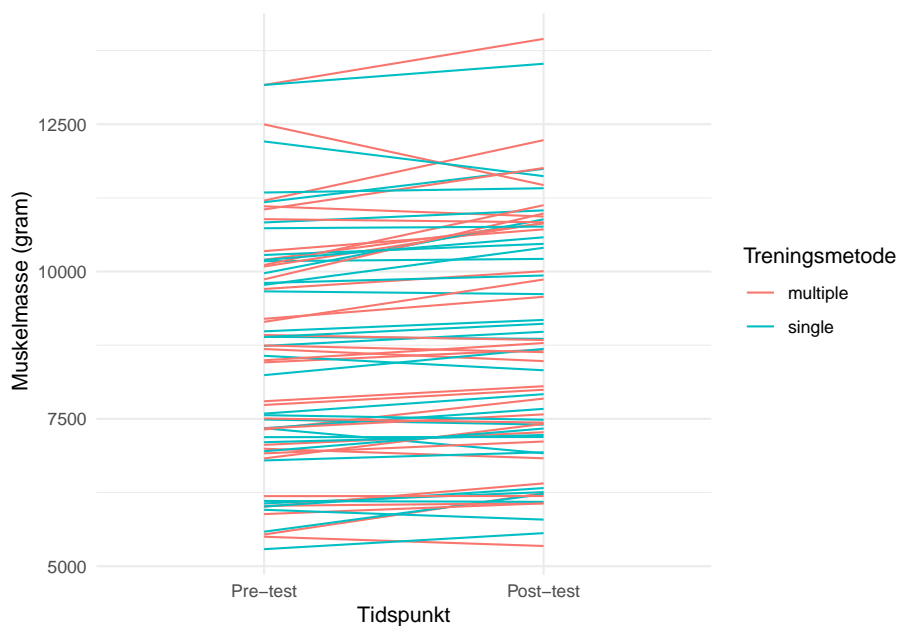


Figure 5.1: Figur 1 viser økningen i muskelvekst fra pre-test til post-test for alle forsøkspersoner skildt ved single- sett (1 sett) og multiple- sett (3 sett).

### 5.3.2 Muskelstyrke

Begge treningsmetodene gav økning i styrke fra pre til uke fem og uke fem til post (figur 2) ( $p < 0.05$ ). Den største økningen kom for begge treningsmetodene mellom pre og uke fem (begge  $p < 0.01$ ) (tabell 1). Fra uke fem og post var det en økning på 10% for benet som trente tre sett og 6.8% for benet som trente ett sett. Benet som trente tre sett økte mer enn benet som trente ett sett ved både fem uker og ved post-test (begge  $p < 0.01$ ) (tabell 1).

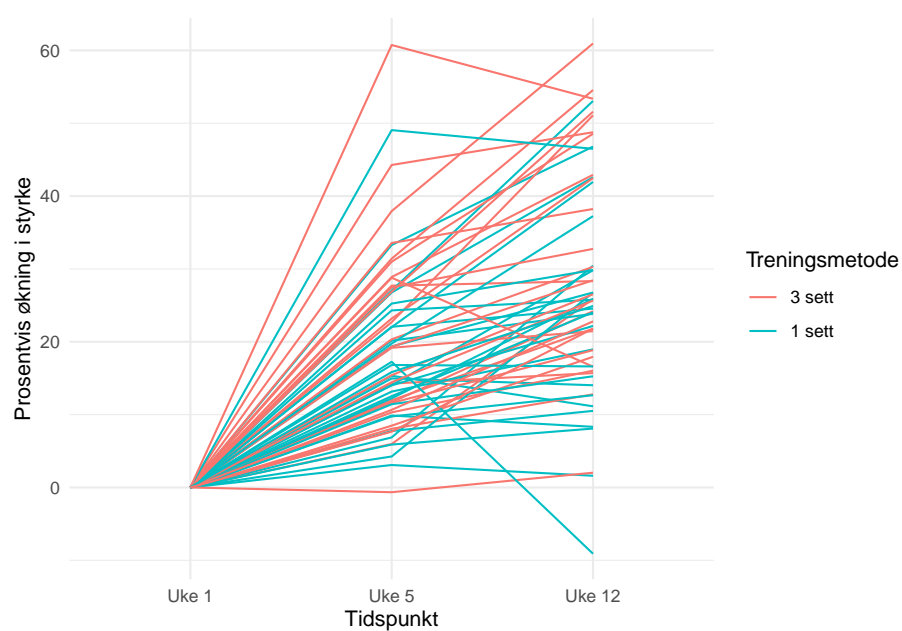


Figure 5.2: Figur 2 viser prosentvis økning i muskelstyrke, fra pre-tests til 5 uker og 5 uker til post-test for alle forsøkspersoner skildt ved single- sett (1 sett) og multiple- sett (3 sett).

Table 5.1: Tabellen viser prosentvis økning fra pre til uke 5 og uke 5 til uke 12(post)

	Uke 5	Post
3 sett	21.1(13.1)	31.1(15.1)
1 sett	17(9.58)	23.8(14.2)

Verdiene er gitt som gjennomsnitt og standardavvik (SD).  $p < 0.01$  ved pre til uke fem og uke fem til post for begge ben,  $p < 0.01$  ved begge tidspunkt for at tre sett økte mer enn fem sett

## 5.4 Diskusjon

Hovedfunnene i denne studien er at tre sett gir en større økning i både styrke og muskelmasse enn å trene ett sett etter 12 uker med styrketrening. Den største økningen i styrke kom etter de fem første ukene. Funnene våre stemmer overens med flere av studiene som har ser på grupper (ett sett mot tre sett) opp mot hverandre [Krieger, 2009, Galvão and Taaffe, 2005, Schoenfeld et al., 2016].

Etter fem uker med trening var ikke forskjellen like stor som ved 12 uker. Dette kan stemme med Radaelli et al. [2014], selv om han ikke så forskjell mellom ett og tre sett. I den studien så de på effekten av seks uker hos utrente hvor forfatterne konkluderte med at ett sett gir like god effekt som tre sett hos utrente i startfasen av styrketrening [Radaelli et al., 2014]. Selv om studien vår viser at tre sett er fordelaktig også i en startfase, kan man likevel se en overlegen effekt av tre sett fra uke fem til 12. Dette tyder på at volum blir viktigere og viktigere jo lengre treningsperioden varer. I motsetning til dette så Hass et al. [2000] i sin studie at ett sett var like bra som tre sett også hos trente personer, på både kort og lang sikt. Det som likevel skiller denne studien fra vår i tillegg til at vi ikke hadde to grupper, er at de så på styrketester i både overkropp og underkropp [Hass et al., 2000]. Når man ser på resultatet i studien er det signifikant forskjell i favør tre sett i leggfleksjon og generelt ikke overbevisende resultat for ett sett på underkroppsovelser [Hass et al., 2000]. Det kan dermed virke som at tre sett er å foretrekke på underkroppen mens ett sett kan være nok på overkroppen. Denne konklusjonen stemmer godt overens med andre studier [Rønnestad et al., 2007, Paulsen et al., 2003]. Likevel kunne det vært interessant om man også kan se denne effekten om man har samme studiedesign som i denne studien, hvor man hadde trent forskjellig på venstre og høyre arm.

I forhold til økning i muskelmasse var det en liten, men gyldig økning i vår studie. Lignende funn så Schoenfeld et al. [2019], som konkluderer med at det er



en dose-respons-effekt på økning i muskelmasse. Det samme klarete ikke Galvão and Taaffe [2005] å finne når han så på ett vs. tre sett hos eldre mennesker. Det kan potensielt forklares med at eldre trenger mer protein for å oppnå økning i muskelmasse samtidig som at de ikke er like flinke til få i seg proteiner som unge [Kraemer et al., 1999, Moore et al., 2015]. Vår studie skiller seg generelt fra andre da potensialet et menneske har for muskelmasseøkningen teoretisk sett er lik på begge ben.

#### 5.4.1 Treningsanbefaling

På bakgrunn av denne studien og andre studeier nevnt over, kan man som utrent få god effekt av å trene ett sett på hele kroppen i starten av en treningsperiode. Etter ca. en måned bør man øke antall sett på underkroppspøvelsene, men kan forsette med ett sett på overkroppen. Det er også viktig at alle repetisjoner utføres som RM for å få mest mulig effektiv av treningen.

#### 5.4.2 Konklusjon

Resultatet i studien stemte med hypotesen. Både tre sett og ett sett gav god effekt på beinstyrken hos trente personer, men tre sett gav ytterlige styrkeeffekt enn ett sett, spesielt etter fem uker. Tre sett er også å foretrekke for å oppnå økning muskelmasse.



# Bibliography

- Monya Baker. 1,500 scientists lift the lid on reproducibility. *Nature*, 533 (7604):452–454, 05 2016. doi: 10.1038/533452a. URL <https://www.nature.com/articles/533452a>. Bandiera\_abtest: a Cg\_type: Nature Research Journals Number: 7604 Primary\_atype: News Publisher: Nature Publishing Group Subject\_term: Peer review;Publishing;Research management Subject\_term\_id: peer-review;publishing;research-management.
- 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.
- C. Glenn Begley and Lee M. Ellis. Raise standards for preclinical cancer research. *Nature*, 483(7391):531–533, 03 2012. doi: 10.1038/483531a. URL <https://www.nature.com/articles/483531a>. Bandiera\_abtest: a Cg\_type: Nature Research Journals Number: 7391 Primary\_atype: Comments Opinion Publisher: Nature Publishing Group Subject\_term: Cancer;Drug development Subject\_term\_id: cancer;drug-development.
- 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 <https://www.journals.uchicago.edu/doi/full/10.1093/bjps/axy051>. Publisher: The University of Chicago Press.
- G. A. Borg. Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise*, 14(5):377–381, 1982. PMID: 7154893.
- 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.
- Cary Funk and Brian Kennedy. The politics of climate. 10 2016. Pew Research Center.

- Daniel A. Galvão and Dennis R. Taaffe. Resistance exercise dosage in older adults: Single- versus multiset effects on physical performance and body composition. *Journal of the American Geriatrics Society*, 53(12): 2090–2097, 2005. doi: 10.1111/j.1532-5415.2005.00494.x. URL <http://onlinelibrary.wiley.com/doi/abs/10.1111/j.1532-5415.2005.00494.x>. \_eprint: <https://agsjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1532-5415.2005.00494.x>.
- Olof Gudny Geirsdottir, A. Arnarson, K. Briem, A. Ramel, K. Tomasson, P. V. Jonsson, and I. Thorsdottir. Physical function predicts improvement in quality of life in elderly icelanders after 12 weeks of resistance exercise. *The journal of nutrition, health and aging*, 16(1):62–66, 01 2012. doi: 10.1007/s12603-011-0076-7. URL <https://doi.org/10.1007/s12603-011-0076-7>.
- 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.
- Chris J. Hass, Linda Garzarella, Diego De Hoyos, and Michael L. Pollock. Single versus multiple sets in long-term recreational weightlifters. *Medicine Science in Sports Exercise*, 32(1):235, 01 2000. URL [https://journals.lww.com/acsm-msse/Fulltext/2000/01000/Single\\_versus\\_multiple\\_sets\\_in\\_long\\_term.35.aspx](https://journals.lww.com/acsm-msse/Fulltext/2000/01000/Single_versus_multiple_sets_in_long_term.35.aspx).
- 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, editor. *Designing clinical research*. Wolters Kluwer/Lippincott Williams and Wilkins, Philadelphia, 4th ed edition, 2013.
- John P. A. Ioannidis. Why most published research findings are false. *PLOS Medicine*, 2(8):e124, 08 2005. doi: 10.1371/journal.pmed.0020124. URL <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020124>. Publisher: Public Library of Science.
- William J. Kraemer, Keijo Häkkinen, Robert U. Newton, Bradley C. Nindl, Jeff S. Volek, Matthew McCormick, Lincoln A. Gotshalk, Scott E. Gordon, Steven J. Fleck, Wayne W. Campbell, Margot Putukian, and William J. Evans. Effects of heavy-resistance training on hormonal response patterns in younger vs. older men. *Journal of Applied Physiology*, 87(3):982–992, 09 1999.

- doi: 10.1152/jappl.1999.87.3.982. URL <https://journals.physiology.org/doi/full/10.1152/jappl.1999.87.3.982>. Publisher: American Physiological Society.
- James W. Krieger. Single versus multiple sets of resistance exercise: A meta-regression. *The Journal of Strength and Conditioning Research*, 23(6):1890–1901, 09 2009. doi: 10.1519/JSC.0b013e3181b370be. URL [https://journals.lww.com/nsca-jscr/Fulltext/2009/09000/Single\\_Versus\\_Multiple\\_Sets\\_of\\_Resistance.37.aspx](https://journals.lww.com/nsca-jscr/Fulltext/2009/09000/Single_Versus_Multiple_Sets_of_Resistance.37.aspx).
- 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.
- Alexandra Kuznetsova, Per B. Brockhoff, and Rune H. B. Christensen. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26, 2017. doi: 10.18637/jss.v082.i13.
- Daniel R. Moore, Tyler A. Churchward-Venne, Oliver Witard, Leigh Breen, Nicholas A. Burd, Kevin D. Tipton, and Stuart M. Phillips. Protein ingestion to stimulate myofibrillar protein synthesis requires greater relative protein intakes in healthy older versus younger men. *The Journals of Gerontology: Series A*, 70(1):57–62, 01 2015. doi: 10.1093/gerona/glu103. URL <https://doi.org/10.1093/gerona/glu103>.
- Samir Okasha. *Philosophy of science: a very short introduction*, volume 67 of *Very short introductions*. University Press, Oxford, second edition. edition, 2016.
- Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*, 349(6251):aac4716, 08 2015. doi: 10.1126/science.aac4716. URL <https://www.science.org/doi/10.1126/science.aac4716>. Publisher: American Association for the Advancement of Science.
- Gøran Paulsen, Dag Myklestad, and Truls Raastad. The influence of volume of exercise on early adaptations to strength training. *Journal of Strength and Conditioning Research*, 17(1):115–120, 02 2003. doi: 10.1519/1533-4287(2003)017<0115:tiovoe>2.0.co;2. PMID: 12580666.
- Charles S. Peirce, Nathan Houser, and Christian J. W. Kloesel. *The essential Peirce: selected philosophical writings*. Indiana University Press, Bloomington, 1992.
- Charlotte Price Persson. Abduksjon: Metoden for å finne den beste forklaringen, 04 2019. URL <https://forskning.no/a/1317339>. Section: kultur.
- Karl R. Popper. *Conjectures and refutations: the growth of scientific knowledge*. Routledge classics. Routledge, London ; New York, 2002.

- Regis Radaelli, Eurico N. Wilhelm, Cíntia E. Botton, Anderson Rech, Martin Bottaro, Lee E. Brown, and Ronei S. Pinto. Effects of single vs. multiple-set short-term strength training in elderly women. *AGE*, 36(6):9720, 10 2014. doi: 10.1007/s11357-014-9720-6. URL <https://doi.org/10.1007/s11357-014-9720-6>.
- 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.
- 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. PMID: 17313291.
- 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, 11 2016. doi: 10.1007/s40279-016-0543-8. URL <https://doi.org/10.1007/s40279-016-0543-8>.
- Brad J. Schoenfeld, Bret Contreras, James Krieger, Jozo Grgic, Kenneth Delcastillo, Ramon Belliard, and Andrew Alto. Resistance training volume enhances muscle hypertrophy but not strength in trained men. *Medicine Science in Sports and Exercise*, 51(1):94–103, 01 2019. doi: 10.1249/MSS.0000000000001764. URL [https://journals.lww.com/acsm-msse/Fulltext/2019/01000/Resistance\\_Training\\_Volume\\_Enhances\\_Muscle.13.aspx](https://journals.lww.com/acsm-msse/Fulltext/2019/01000/Resistance_Training_Volume_Enhances_Muscle.13.aspx).
- Nadja Schott, Bettina Johnen, and Benjamin Holfelder. Effects of free weights and machine training on muscular strength in high-functioning older adults. *Experimental Gerontology*, 122:15–24, 07 2019. doi: 10.1016/j.exger.2019.03.012. URL <https://www.sciencedirect.com/science/article/pii/S0531556518307952>.
- Statistisk sentralbyrå. Styrketrening stadig mer populært, 10 2019. URL <https://www.ssb.no/kultur-og-fritid/artikler-og-publikasjoner/styrketrening-stadig-mer-populaert>.
- Mari Turpela, Keijo Häkkinen, Guy Gregory Haff, and Simon Walker. Effects of different strength training frequencies on maximum strength, body

- composition and functional capacity in healthy older individuals. *Experimental Gerontology*, 98:13–21, 11 2017. doi: 10.1016/j.exger.2017.08.013. URL <https://www.sciencedirect.com/science/article/pii/S0531556517303315>.
- Espen Tønnessen, Erlend Hem, Ida Svendsen, Eirik V. Larsen, Martin Skaugen, and Elisabeth Solbakken. Utholdenhetstester ved olympiatoppen. 2017. URL [https://www.olympiatoppen.no/fagomraader/trening/testing/testing\\_av\\_utholdenhet/media53703.media](https://www.olympiatoppen.no/fagomraader/trening/testing/testing_av_utholdenhet/media53703.media).
- Sanna Vikberg, Niklas Sörlén, Lisa Brandén, Jonas Johansson, Anna Nordström, Andreas Hult, and Peter Nordström. Effects of resistance training on functional strength and muscle mass in 70-year-old individuals with pre-sarcopenia: A randomized controlled trial. *Journal of the American Medical Directors Association*, 20(1):28–34, 01 2019. doi: 10.1016/j.jamda.2018.09.011. URL <https://www.sciencedirect.com/science/article/pii/S1525861018305024>.
- Kevin R. Vincent, Randy W. Braith, Ross A. Feldman, Pete M. Magyari, Rachel B. Cutler, Stephanie A. Persin, Shannon L. Lennon, Abdel H. Gabr Md, and David T. Lowenthal. Resistance exercise and physical performance in adults aged 60 to 83. *Journal of the American Geriatrics Society*, 50(6):1100–1107, 2002. doi: 10.1046/j.1532-5415.2002.50267.x. URL <http://onlinelibrary.wiley.com/doi/abs/10.1046/j.1532-5415.2002.50267.x>. [\\_eprint: https://agsjournals.onlinelibrary.wiley.com/doi/pdf/10.1046/j.1532-5415.2002.50267.x](https://agsjournals.onlinelibrary.wiley.com/doi/pdf/10.1046/j.1532-5415.2002.50267.x).