

# **Eksamen IDR4000**

Kandidatnummer: 403

2023-11-17

# Innholdsfortegnelse

<b>Forord</b>	<b>4</b>
<b>Reliabilitet</b>	<b>5</b>
Introduksjon . . . . .	5
Metode . . . . .	6
Praktisk gjennomføring/tiltak for å sikre reliabilitet . . . . .	7
Resultater . . . . .	7
<b>Laborasjonsrapport</b>	<b>10</b>
Introduksjon . . . . .	10
Metode . . . . .	11
Resultat . . . . .	12
Diskusjon . . . . .	13
Konklusjon . . . . .	14
<b>Vitenskapsfilosofi</b>	<b>15</b>
Falsifikasjonisme . . . . .	15
Baysianisme . . . . .	17
<b>Studiedesign</b>	<b>19</b>
Introduksjon . . . . .	19
Sammenligning av studiedesign og metoder . . . . .	19
Statistiske tester og resultat . . . . .	21
Konklusjon . . . . .	22
<b>Analyse av repeterte målinger</b>	<b>23</b>
Introduksjon . . . . .	23
Metode . . . . .	24
Deltakere og studiedesign . . . . .	24
Måling av maksimal styrke i beinpress . . . . .	25
DXA . . . . .	25
Analyse og statistikk . . . . .	25
Resultater . . . . .	26
Diskusjon . . . . .	29
Konklusjon . . . . .	29



# Forord

Se hele repositorien på GitHub: <https://github.com/andreasohrn/eksamen-idr4000-qmd.git>.

# Reliabilitet

## Introduksjon

Vi gjennomførte to testdager 26.09.2023 og 28.09.2023 i tiden 08:00-16:00. Hensikten med disse to dagene var å gjennomføre fysiologiske tester med høy grad av reliabilitet. Det er flere faktorer som påvirker både validitet og reliabilitet, og det er viktig å ta høyde for dette under fysiologisk testing. Vi gjennomførte testdag 1 og testdag 2 med kun én dag mellom for å sikre at deltakerne var på tilsvarende fysiologisk nivå ved begge testene. Vi tok derfor en rekke forhåndsregler for å sikre så like testforhold som mulig.

Reliabilitet refererer til reproduserbarheten til en f.eks. en fysiologisk test som gjennomføres flere ganger i en repetert studie, der bedre reliabilitet indikerer bedre presisjon og måling av endring over tid (Hopkins, 2000). Innenfor reliabilitet er det en rekke relevante begreper. Standardavvik (SD) forteller hvor langt unna dataene er fra gjennomsnittet (Spiegelhalter, 2020), typical error (TE) beskrives av Hopkins (2000) som variabiliteten hos hver enkelt verdi og tenkes å kunne visualisere feilmarginen av et estimat. For å få nøyaktige målinger som kan sammenlignes er det viktig med presise måleinstrumenter som kalibreres nøye. Det er også faktorer som læringseffekt, motivasjon, restitusjon og ernæringstilstand som kan påvirke resultatene, og det er viktig å ta høyde for dette ved fysiologisk testing. Hopkins (2000) hevder at det kreves om lag 50 deltakere og 3 repeterte målinger for å kunne estimere reliabiliteten. Dette er for å utelukke de ovennevnte faktorene.

Kroppens maksimale oksygenopptak ( $VO_{2maks}$ ) gir informasjon om en persons maksimale aerobe kapasitet. Oksygenopptaket bestemmes av både sentrale- og perifere faktorer og kan illustreres ved Flick's likning:

$$VO_2 = (HR \times SV) \times (aO_2 - vO_2)$$

En  $VO_{2maks}$ -test går ut på at man måler hvor mange ml oksygen en person evner å ta opp og omsette per minutt. Oksygenkravet øker lineært med belastningen helt til personen når sin maksimale aerobe kapasitet, da vil kurven flate ut eller eventuelt synke. En persons maksimale oksygenopptak kan beskrives både i form av absolutte tall (ml/min) eller som relative tall i forhold til kroppsvekt (ml/kg/min).

## Metode

$VO_{2maks}$  testen gjennomføres som en trappetrinnstest der motstanden øker med 25W hvert minutt til utmattelse/når  $RPM < 60$ .  $VO_2$  målinger registreres hvert 30 sek. Deltakerne startet testen på enten 150W, 200W eller 250W avhengig av fysisk form og erfaring med sykkel. Hvert minutt øker watten med 25 helt til utmattelse. Etter endt test ble informasjon innhentet og plottet i ferdigstilt Excel-dokument.

## Praktisk gjennomføring/tiltak for å sikre reliabilitet

Selv om en tydelig protokoll er essensielt for å sikre reliable tester på en fysiologilab, er det flere hensyn som må tas underveis. Vi begynte hver test med å kalibrere utstyret slik at det var oppdatert etter forholdene til hver deltaker hver klokke time, for å minimere risikoen for at utstyret skal måle feil (Tanner et al., 2013). For å sikre lik grad av verbal motivasjon og formulering av instruks, valgte vi å bruke samme testleder på hver person (Halperin et al., 2015). Vi ha også instruks om at kosthold og søvn skulle være likt før og på testdagen, samt ingen trening mellom testene da dette er faktorer som kan påvirke metabolisme og prestasjon, dette ble dog ikke kontrollert (Tanner et al., 2013).

## Resultater

Etter testdag 1 (T1) fikk vi en oversikt hvilket fysiologisk nivå deltakerne i prosjektet var på. Vi valgte å undersøke variablene  $VO_{2max}$  (ml/min),  $VO_{2max}$  (ml/min/kg) og kroppsvekt. Se Tabell 1 for oversikt over T1.

Tabell 1: Oversikt over testresultater etter T1. Gjennomsnitt (Mean), laveste observasjon (Min), høyeste observasjon (Max), og standardavvik (SD)

Variable	Mean	Min	Max	SD
VO2max (ml/min)	3,990.12	2,504.00	5,893.00	999.95
VO2max (ml/min/kg)	49.83	31.74	77.54	15.28
Weight (kg)	81.75	66.40	105.90	12.32

Formålet med dette prosjektet var å teste reliabiliteten på utvalgte fysiologiske mål mellom T1 og T2. Figur 1 gir en oversikt over resultatet på T1 og T2 for hver ID. Oversikt over aktuelle reliabilitetstall finner man i Tabell 2.

Standardavvik (SD) kan forklares ved et mål på hvor stor spredningen er i forhold til datapunktens middelvei/gjennomsnitt, og definerer hvert enkelt datapunkts avvik fra gjennomsnittet (Spiegelhalter, 2020).

Variasjonskoeffisienten (CV) angir et spredningsmål for verdiene i et datasett. CV uttrykker ofte variasjon i forhold til gjennomsnittsverdien og angis i prosent (Spiegelhalter, 2020).

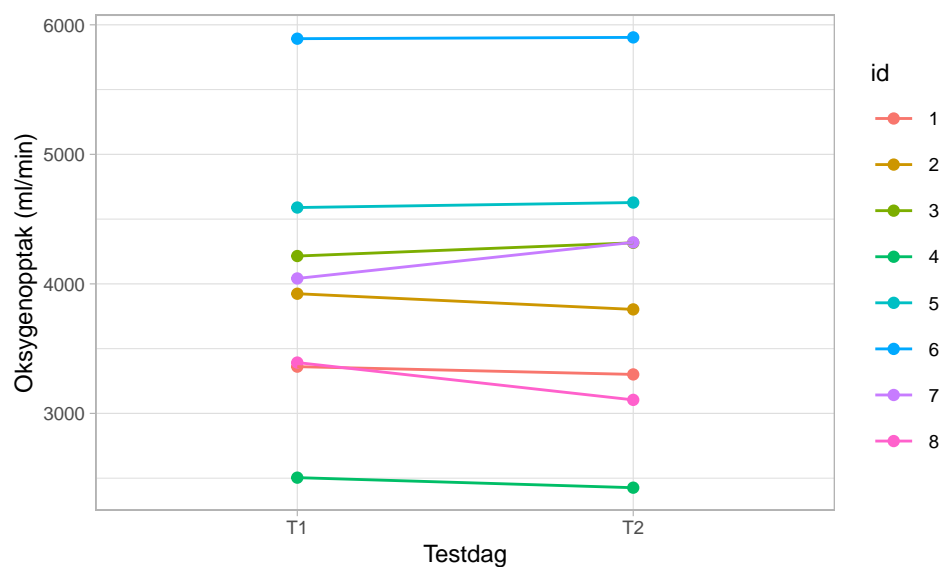
Typical Error (TE) eller standardfeil er variabiliteten hos hver enkelt verdi og tenkes å kunne visualisere feilmarginen av et estimat. Eksempelvis vil en standardfeil kunne forklares gjennom biologiske prosesser som påvirker for eksempel kraftutvikling, som følge av mentale eller fysiske faktorer (Hopkins, 2000).

Limits of agreement (LoA) viser til 1,96 standardavvik fra den gjennomsnittlige differansen mellom T1 og T2. Dette er illustrert i Figur 2 med et Bland-Altman-Plot.

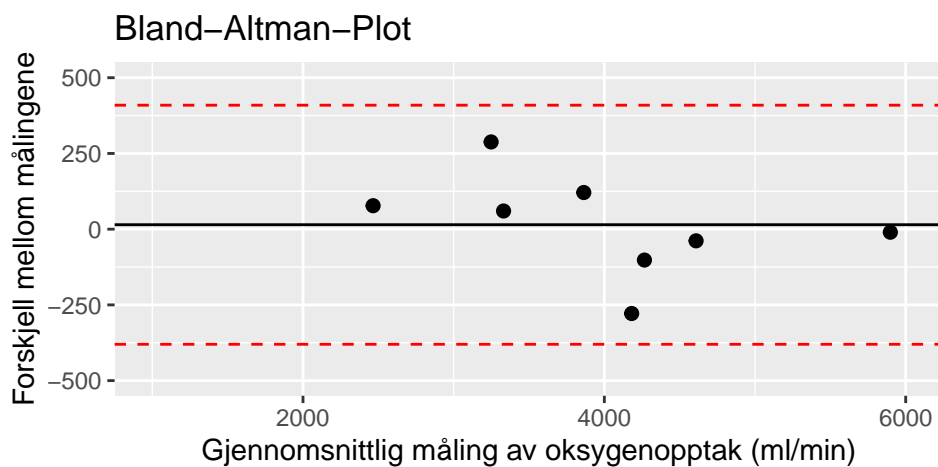
Tabell 2: Oversikt over reliabilitetstall for utvalgte tester. Tabellen viser gjennomsnitt av T1 og T2 (Mean), standardavvik (SD), typical error (TE), variasjonskoeffisient (CV), og limits of agreement (LoA)

Variable	Mean	SD	TE	CV (%)	LoA
VO2max (ml/min)	3,982.78	166.91	118.02	2.96	394.68
VO2max (ml/min/kg)	49.50	2.48	1.75	3.54	5.87
Weight (kg)	82.07	1.73	1.22	1.49	4.09





Figur 1: Sammenligning av oksygenopptak målt i ml/min mellom testdag 1 og testdag 2 for hver forsøksperson



Figur 2: Bland-Altman-Plot for maksimalt oksygenopptak (ml/min). Stiplede linjer tilsvare øvre og nedre limits of agreement. Heltrukken linje tilsvare gjennomsnittet av differansen mellom T1 og T2

# Laborasjonsrapport

## Introduksjon

Analyser av genuttrykk gjennom fluoresens-basert sanntids kvantitativ polymerase kjedereaksjon (qPCR) er fast praksis i mange medisinske treningsstudier (Kuang et al., 2018). qPCR benyttes til å måle uttrykket av et målgen i prøver fra blant annet blod og muskelvev. Et vanlig bruksområde innenfor treningsfysiologi er for eksempel å måle treningsinduserte endringer i genuttrykk for ulike muskelfibertyper. Selv om denne typen analyse er svært utbredt, finnes det mange ulike protokoller og måter å gjennomføre analysen på. Reproduserbarheten og reliabiliteten til dataen avhenger i stor grad av hvordan eksperimentene gjennomføres og tolkes. For å sikre så reliable tester som mulig, er det viktig med en detaljert og nøye protokoll (Kuang et al., 2018).

Prinsippet bak en qPCR-analyse er å følge PCR i sanntid. For å gjennomføre PCR må vi først ekstrahere RNA fra en biologisk celleprøve. Deretter blir RNAet gjort om til cDNA i en prosess kalt reversert transkripsjon. Med PCR blir dette cDNAet deretter amplifisert opp til milliarder av kopier (Kuang et al., 2018). Hver syklus i en PCR består av tre steg for å kopiere opp det aktuelle DNAet. I første steg (denaturering) blir DNAet utsatt for høy temperatur for å dele DNAet fra dobbeltrådet til enkelttrådet. I det andre steget (annealing) blir temperaturen senket og primere fester seg til templat-trådene. I det siste steget (elongering) øker temperaturen igjen og DNA-polymeraser fester seg til primerne og syntetiserer et nytt dobbeltrådet DNA, likt det vi startet med (Kuang et al., 2018). Deretter vil denne syklusen repetere seg, slik at DNAet doubles eksponentielt for hver syklus. I en

qPCR som benytter SYBR green-metoden blir det aktuelle DNAet bundet til fluoresens under hver syklus slik at man kan følge PCR-prosessen i sanntid, ved å ta et bilde som detekterer fluorescensen etter hver syklus (Kuang et al., 2018). Mengden cDNA og fluorescens doubles etter hver syklus, og signalet av fluorescens øker dermed eksponentielt. Hvor raskt signalet av fluorescens når en satt grense kalt syklisk terskel (CT), bedømmer hvor stort uttrykk av et gen vi har. Jo færre sykluser som må gjennomføres for å nå CT, jo større uttrykk av målgenet hadde vi i celleprøven vår (Livak & Schmittgen, 2001).

## Metode

I forkant av forsøket hadde labansvarlig forberedt cDNA. qPCR-analyse ble gjort ved bruk av cDNA og en Master mix. Master mixen besto av 5 l Cybr-green, 1 l primer mix (MCH1, MHC2a, MCH2x eller MCHb2m) og 2 l H<sub>2</sub>O. På en plate med brønner ble det tilført 8 l Master mix i brønnene sammen med 2 l cDNA. I tillegg til dette lagde vi en fortynningsserie for å teste primerne. Fortynningene vi brukte var 1/1, 1/10, 1/100, 1/1000, 1/10000, 1/100000 og 1/1000000. I 1/1 prøven var det 2 l cDNA og 8 l cmyc. Fortynningsserien tok utgangspunkt i denne prøven og ble fortynnet med H<sub>2</sub>O. Platen ble så dekket med plast og sentrifugert på 1200rpm i 1 minutt. PCR-prøvene ble analysert ved sanntids PCR (Applied Biosystems 7500 fast Real-TimePCR Systems, Life Technologies AS) og ved bruk av Quant Studio programvare (Applied Biosystems, Waltham, MA USA). PCR-prosessen besto av tre deler, en “Hold stage” en “PCR stage” og en “Melt curve stage”. Det første steget gikk ut på at temperaturen økte med 1,99°C/s opp til 50°C hvor temperaturen forble konstant i 2 min. Videre økte temperaturen med 1,99°C/s opp til 95°C hvor temperaturen forble konstant i 2 min. Deretter startet selve PCR-prosessen som besto av 40 sykluser. Én syklus besto av 1 sek på 95°C før temperaturen sank med 1,77°C/s ned til 60°C hvor temperaturen ble holdt konstant i 30 sek. Etter hver syklus ble det tatt bilde av brønnenes fluorescens. Avslutningsvis økte temperaturen med 1,99°C/s opp til 95°C hvor temperaturen holdes konstant i 15 sek. Deretter synker temperaturen med 1,77°C/s ned til 60°C hvor temperaturen holdes konstant i 1 min.

Temperaturen økte deretter med 0,15°C/s opp til 95°C hvor temperaturen ble holdt konstant i 15 sek. Etter PCR-prosessen var ferdig kunne vi hente ut resultatene i form av CT-verdier.

## Resultat

Vi ser at mengden genuttrykk for de ulike muskelfibertypene har endret seg fra uke 0 til uke 2 Tabell 1. Genuttrykket for Muskelfibertype 1 (MCH1) har sunket fra 35-27%, muskelfibertype 2a (MHC2a) har økt fra 0,6-72% og muskelfibertype 2x (MHC2x) har sunket fra 63-4%.

Tabell 1: Prosentvis fordeling av genuttrykk etter uke 0 og uke 2

Gentype	Uke 0	Uke 2
MHC1	35.43	27.33
MHC-2a	0.61	72.21
MHC-2x	63.96	4.61

Antall sykluser for å nå syklisk terskel (CT) har endret seg fra uke 0 til uke 2 Tabell 2. Antall sykluser har sunket for MCH1 fra 22-17 sykluser. For MHC2a har antall sykluser sunket 27-15 sykluser, og for MCH2x har antall sykluser til CT økt fra 21-23 sykluser. Lavere antall sykluser viser til større genuttrykk.

Tabell 2: CT-verdier

Gentype	Uke 0	Uke 2
MHC1	22.04	17.29
MHC-2a	27.91	15.89
MHC-2x	21.19	23.19

Tallene er antall sykluser før syklisk terskel er nådd

## Diskusjon

Ved bruk av fluoresens-basert sanntids kvantitativ polymerase kjedereaksjon (qPCR) er det essensielt med både egnede gener og en detaljert protokoll. Målet med undersøkelsen var å kunne vurdere, presentere og tolke genuttrykk. I denne oppgaven valgte vi å undersøke endringen i mengden genuttrykk av MHC1, MHC2a og MHC2x før trening og etter 2 uker med styrketrening.

Vi har undersøkt antall sykluser som trengs for å nå CT for de bestemte genene, der færre sykluser indikerer et større genuttrykk. Genuttrykket for MCH2x var det som trengte flest sykluser for å nå CT ved uke 2, mens for både MCH1 Og MCH2a har antall sykluser til CT sunket. Dette tyder på at genuttrykket til MCH1 og MHC2a har økt mens MHC2x har sunket.

Den reduserte mengden genuttrykk for MCH1 og MHC2x og den økte mengden genuttrykk av MHC2a virker å være fornuftige. Tabell 1 i Terzis et al. (2007) viser blant annet at etter 14 uker med styrketrening har prosentandelen av muskelfibertype 2a økt. Andersen & Aagaard (2000) viser også at styrektrening reduserer muskelfibertype MCH2x og øker MHC2a. Dette ser ut til å stemme over ens med våre resultater som vises i tabell 1.

Endringen i genuttrykk av MHC2a og MHC2x er stor. Dette kan tyde på at vi analyserte en prøve fra en veldig utrent person. Siden vi på forhånd ikke fikk noe informasjon om forsøkspersonen vi analyserte prøven fra, blir dette likevel vanskelig å fastslå. I Andersen & Aagaard (2000) beskrives det at ved å øke muskelaktiviteten, enten gjennom styrke- eller utholdenhetstrening, slås MHC2x-genet av noe som resulterer i en forhøyet andel MHC2a-gen. Dette stemmer godt med våre observasjoner. Ved inaktivitet reverseres denne prosessen, noe som gjør forekomsten av MHC2x mer vanlig hos inaktive mennesker.

## Konklusjon

Denne analysen av genuttrykk for ulike myosin tungkjeder indikerer at styrketrening bidrar til endringer som et forhøyet uttrykk av MHC2a i forhold til MHC1 og MHC2x.

# Vitenskapsfilosofi

## Falsifikasjonisme

Falsifikasjonisme er et vitenskapelig prinsipp gjort kjent av den østeriske filosofen Karl Popper. Popper mente at innen vitenskapelig metode kan en teori kun betraktes som vitenskap hvis den er falsifiserbar. Det vil si at teorien eller hypotesen er formulert på en slik måte at man kun med bruk av deduktiv logikk og empirisk testing potensielt kan falsifisere den (Popper, 1959). Det er viktig å presisere at selv om et utsagn må være falsifiserbart for å regnes som vitenskap, er den ikke nødvendigvis usann. Popper mente også at en teori med flere og mer presise prediksjoner er bedre, siden den da vil være mer falsifiserbar. Eksempelvis vil en teori som sier at «alle objekter faller mot bakken med en hastighet på  $9,8\text{m/s}^2$  i et vakuum», være bedre enn å bare si at «alle objekter faller mot bakken når de slippes i et vakuum». Det finnes altså ulike grader av falsifiserbarhet (Popper, 1959).

Det er flere grunner til at Popper mente at falsifikasjon var den eneste riktige måten å drive ordentlig vitenskap på. Popper mente blant annet at man utelukkende må basere vitenskapen på deduktiv logikk. Bakgrunnen dette baserer seg på det David Hume beskrev som «induksjonsproblemet». Induksjonsproblemet går kort sagt ut på at man ikke rasjonelt kan begrunne eller rettferdiggjøre induktive beslutninger, siden induktive resonnementer baserer seg på antagelsen om at noe som har skjedd før i tiden også vil skje i framtiden, også kalt uniformitetsprinsippet. Popper var opptatt av at uansett hvor mange ganger et fenomen har skjedd før, kan man aldri med 100% garanti si at det vil skje på akkurat lik måte også i framtiden. Verifisering av teorier og hypoteser må derfor basere seg på antagelser, noe som

Popper trakk linjer mot pseudovitenskapen. Med bakgrunn av dette burde fokuset være på å heller forsøke å falsifisere teorier ved hjelp av deduktiv logikk og empiri. Hvis en teori står lenge uten at noen klarer å falsifisere den gjennom gjentatte forsøk, regnes den ofte som en sterkere teori, uten at man noen gang kan si at den er bekreftet.

Med falsifikasjonismen ønsket også Popper å svare på «demarksjonsproblemet». Altså det å skille mellom hva som er vitenskap og ikke-vitenskap. Ved å sette klare kriterier på hva som kan regnes som vitenskap, vil også vitenskapen ha høyere status. Hvis en teori eller hypotese ikke kan falsifiseres, kan man ikke lenger kalle det vitenskap, men heller konspirasjoner, myter eller ideologier. En korrekt måte å drive forskningsfelt videre på var ifølge Popper å komme opp med en hypotese   prediksjon   falsifisering   ny hypotese osv. osv. Denne vitenskapssyklusen fungerer bra hvis hypotesen er falsifiserbar og gir prediksjoner som kan testes empirisk. Popper kritiserte sterkt store åpne teorier som kunne gi svar til veldig mange ulike spørsmål. Eksempelvis kan evolusjonær psykologi kritiseres med bakgrunn i dette. Der er fokuset å forklare menneskelig atferd ved hjelp av evolusjonsteori. Problemet er det at man kun er på leting etter å bekrefte en teori eller hypotese man har, og man tolker bevisene på en måte som best mulig passer sitt eget narrativ. En slik framgangsmåte er ifølge popper intet bedre enn pseudovitenskap, og blir altfor lettvint.

Poppers radikale syn på hvordan man bør drive vitenskap har også blitt kritisert fra flere hold. Et aktuelt motsvar til falsifikasjon ble presentert av den franske fysikeren Pierre Duhem. Han mente at spesielt innen fysikk så kan ikke en teori i seg selv gi en prediksjon, uten hjelp av flere støttehypoteser. Bruker vi gravitasjonsloven som eksempel så sier den at alle objekter faller mot bakken med lik hastighet i et vakuum. Skal man teste dette utendørs, er det likevel veldig mange andre faktorer som spiller inn. Dette kan blant annet være luftmotstand, vind, temperatur og målefeil. Det kreves dermed støttehypoteser om hvordan disse faktorene påvirker eksperimentet for å kunne undersøke den faktiske teorien Duham mente at når eksperimenter ikke stemmer overens med prediksjoner er det flere potensielle årsaker, inkludert feil i eksper-



imentoppsettet, feil i forståelsen av teorien eller andre faktorer som kan spille inn. Dermed blir det nesten svært vanskelig å vite om det er støttehypotesene som er feil eller selve teorien. Dette betyr også man er avhengig av å verifisere støttehypoteser for å kunne falsifisere en teori. Problemet her er at denne verifiseringen vanligvis krever induksjon, og induksjon kan ikke rasjonelt begrunnes i følge falsifikasjonismen.

Duhams argument mot deduktiv logikk i falsifikasjonisme står enda sterkt i dag, og det virker ikke som det er mulig å unngå delvis induktiv logikk. Spørsmålet da er hvor stort problem det er at vitenskapen delvis bygger på induksjon. Det finnes flere syn på hvordan vitenskapsmetodologien bør være, og en fellesnevner er at de forsøker å sette induktive argumenter i system. Ved å undersøke hva gode induktive argumenter har til felles kan man forsøke å lage system for å optimalisere forskning og vitenskap. Ett forsøk på en slik systematisering er Bayesianisme som jeg skal gå nærmere inn på i neste oppgave.

## Bayesianisme

Baysiansime er en vitenskapsfilosofisk retning som baserer seg på bruken av matematisk sannsynlighetsregning og statistikk til å svare på vitenskapelige spørsmål (Godfrey-Smith, 2003). Et veldig sentralt konsept i bayesianismen er at troen rundt en teori eller tema skal oppdateres kontinuerlig som ny informasjon blir tilgjengelig. Hele rammeverket rundt baysiansime er basert på «Bayes Teorem», oppkalt etter statistikeren og filosofen Thomas Bayes, og kan skrives som:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

Bayes Teorem beskriver sjansen for at en hendelse skal skje basert på tidligere observasjoner, også kalt prior-sannsynlighet (Godfrey-Smith, 2003). I Bayesiansk analyse brukes denne prior-sannsynligheten til å representere oppfatningen om en hendelse før man observerer ny data. Etter ny data blir tilgjengelig, oppdateres «posterior-sannsynligheter» ved hjelp av Bayes' teorem. Slik vil man kontinuerlig regne seg fram til sannsynligheten til at noe skal forekomme.

Hvis den nye informasjonen bekrefter hypotesen, vil sannsynligheten for at det forekommer igjen øke, og omvendt.

Med bakgrunn i dette vil en Bayesianer svært sjeldent påstå at noe er en absolutt sannhet, men heller se på hvor sannsynlig det er å tro at noe er sant basert på matematiske utregninger. Sann sett kan man for eksempel tenke seg til at et utfall er 60% sannsynlig i et felt hvor tidligere litteratur er ambivalent. Hvis nyere og bedre forskning på dette feltet kommer med data som støtter hypotesen, kan denne sannsynligheten økes til eks. 80% over tid. Slik leter man ikke etter verifisering eller falsifisering av hypoteser, men heller oppdaterer kunnskapen sin gradvis over tid.

Denne holistiske tilnærming til vitenskapen blir ofte sett på som en styrke med Bayesianismen sammenlignet med blant annet HD-metoden. HD-metoden baserer seg i større grad på en binær tilnærming til hypoteser. Enten er en hypotese verifisert, eller så er den falsifisert. Ved en slik metode kan det være vanskelig å reflektere usikkerheter i forskningsfelt. I tillegg vil en slik metode i større grad slite med å ta hensyn til en gradvis endring innenfor feltet.

Selv om rammeverket rundt Bayesianismen kan gi et godt og realistisk bilde på ulike forskningsspørsmål er det også noen svakheter. Ett problem som kan dukke opp er hvor man skal få all de tidligere sannsynlighetene fra. Forskningsfelt eller tema som er forsket lite på fra før er særlig sårbare. I disse situasjonene kan data fra studier med lav kvalitet eller mangler påvirke posterior-resultate i for stor grad. Man kan tenke seg til at dette vil jevnes ut når man får mer data og informasjon på feltet, men også da må man være obs på eventuelle bias. På svært store forskningsfelt kan det også være en stor utfordring å få oversikt over og strukturere en representasjon av prior sannsynlighet. I tillegg til dette er det ingen veldefinert metode å definere prior sannsynlighet. Dette kan medføre at forskjellige personer kan bruke ulike prior sannsynligheter for de samme eksperimentene (K. Hackenberger, 2019). Hvor store konsekvenser dette medfører er det vanskelig og si, og det kommer nok helt an på situasjonen. På en side kan det være interessant å utforske samme eksperiment med ulike prior sannsynligheter, for å få et større overblikk. På den andre siden er det sårbart mot at personer kan velge å benytte seg av prior som er fordelaktig for dem.

# Studiedesign

## Introduksjon

I denne oppgaven skal jeg sammenligne studier som ser på effekten av samtidig løps- og styrketrening på arbeidsøkonomi i løping. Løpsøkonomi (LØ) regnes vanligvis som en av tre bestemmende faktorer for utholdenhetsprestasjons, sammen med  $VO_{2max}$  og unntyttingsgrad (Tønnessen & Rønnestad, 2018). Dermed vil det være fordelaktig for løpere å finne treningsmetoder som optimaliserer sin egen løpsøkonomi. LØ beskrives som energibehovet for en gitt hastighet, og blir som regel målt som oksygenopptaket på denne hastigheten. Jo bedre løpsøkonomi en utover har, jo lavere oksygenopptak vil personen ha på en gitt hastighet (Tønnessen & Rønnestad, 2018). LØ bestemmes av flere faktorer, og utenom teknikk, ser økt muskelstyrke, fjærstivhet og nevralt adaptasjoner ut til å være mest sentralt (Tønnessen & Rønnestad, 2018). Det er derfor interessant å undersøke hvordan samtidig løps- og styrketrening påvirker løpsøkonomien og generell prestasjon. Det har blitt observert bedring i RE allerede etter 4 uker hos godt trente utøvere (Guglielmo et al., 2009). Tilsvarende observasjoner har også blitt gjort i ulike studier som undersøker effektene fra 8-40 uker med styrketrening (Beattie et al. (2017); Støren et al. (2008); Taipale et al. (2010); Millet et al. (2002).

## Sammenligning av studiedesign og metoder

De fem studiene som jeg har valgt ønsket alle å undersøke hvordan samtidig løps- og styrketrening påvirker LØ og andre fysiologiske faktorer. Jeg har valgt og kun fokusere på LØ i denne

oppgaven. Alle fem studiene er intervensjonsstudier som ser effektene av tung styrketrening og/eller eksplosiv styrketrening. Måten dette har blitt undersøkt på er i alle tilfellene ved å sammenligne resultater fra pre-tester med posttester. Alle studiene er gjort på godt trente utøvere, og vil dermed være begrenset til den populasjonen. LØ ble i alle studiene målt som oksygenopptak på en gitt submaksimal belastning på tredemølle.

For å svare på problemstillingen benyttet Millet et al. (2002) og Støren et al. (2008) et randomisert kontrollert studiedesign (Hulley, 2013). Det blir oppgitt i begge artiklene at de ulike gruppene ble matchet for blant annet nivå, slik at gruppene skulle bli så lik som mulig. Det ble ikke oppgitt hvordan denne randomiseringen ble gjort, men det kan virke som at det ble brukt en “matched pair” randomisering. En slik randomisering brukes for å fordele forsøkspersoner med viktige karakteristikk (Hulley, 2013). I disse studiene ble gruppene matchet med blant annet tid på 5000m, alder og konkurransenivå for å gjøre gruppene så lik som mulig. I Millet et al. (2002) var 7 av deltakere på internasjonalt nivå, og disse ble dermed satt i par og fordelt på de to gruppene. I Støren et al. (2008) ble gruppene matchet for alder og tid på 5000m. Begge studiene hadde mellom 15-19 deltakere, der halvparten av forsøkspersonene fungerte som kontrollgruppe. Med under 10 forsøkspersoner i hver gruppe, vil også en matchet randomisering være viktig for å unngå for store ulikheter mellom gruppene (Hulley, 2013). I begge studiene var det en gruppe som trente tung styrketrening + løping, og en som kun trente løping. Utøverne gjennomførte en pre- og posttest i laboratorium, hvor blant annet løpsøkonomien ble testet.

Taipale et al. (2010) og Guglielmo et al. (2009) undersøkte hvilken effekt kombinert løping med tung styrketrening og eksplosiv styrketrening har på løpsøkonomi. I begge studiene ble forsøkspersoner fordelt enten i en tung styrketrenings- eller eksplosiv styrketreningsgruppe. Begge studiene var randomiserte mixed model designs, der gruppene både ble sammenlignet mot hverandre, og med seg selv pre/post. I et slikt design kan forskere undersøke både innenfor- og mellomgruppeeffekter samtidig (Hulley, 2013). Guglielmo et al. (2009) hadde ikke en egen kontrollgruppe, men utøverne kan fungere som sin egen kontroll når effektene av styrketreningen ble sammenlignet med baseline-testene, som i et within subject design (Hulley, 2013). En

mulig svakhet med dette er at utøverne kan forbedre seg over tid bare på grunn av repetisjon eller kjennskap til testene, uavhengig av intervensjonen. Dette kan føre til falske positive resultater grunnet læringseffekt (Hulley, 2013).

Den siste studien Beattie et al. (2017) var en “two group repeated measures design”, som gikk over 40 uker. 20 deltakere ble delt inn i en styrke+løping-gruppe og en kontrollgruppe. En svakhet med denne studien er mangel på randomisering (Hulley, 2013). Utøverne ble fordelt i en av gruppen basert på blant annet hvor godt styrketrening passet i timeplanen deres, og hvor nære de bodde styrketreningfasiliteter. En slik inndeling kan føre til at utøvere som i utgangspunktet er svært motivert for styrke, og har troen på god effekt ender opp i intervensjonsgruppen. En slik fordeling kan ha innvirkning på hvor pålitelige resultatene i studien er (Hulley, 2013).

## Statistiske tester og resultat

For å undersøke effektene styrketrening har på løpsøkonomi har det blitt brukt flere ulike statistiske tester, avhengig av formålet. For å undersøke endring i LØ før og etter treningsintervensjone innad i gruppene brukte Millet et al. (2002) og Støren et al. (2008) en paret t-test. Millet et al. (2002) “benyttet også en repeatead measures ANOVA” for å teste forskjellene mellom gruppene. De fant at gruppen som trente styrke + utholdenhet bedret LØ mer enn gruppen som bare trente utholdenhet ( $P < 0.05$ ). Begge studiene brukte Pearson korrelasjonstest for å undersøke sammenhengen mellom LØ og ulike styrkerelaterte variabler. Guglielmo et al. (2009) brukte en tosidet ANOVA (group  $\times$  time) for å teste hvordan gruppene “tung styrketrening” og “eksplosiv styrketrening” endret seg fra pre til post. De fant at “tung styrketrening” bedret LØ signifikant fra fra pre til post ( $P < 0.05$ ), men fant ingen signifikant effekt i “eksplosiv styrke”-gruppen. En svakhet her er at det også hadde vært interessant å sammenligne gruppene seg imellom. Taipale et al. (2010) undersøkte forskjell mellom gruppene med en “one way ANOVA” og forskjellene innad i gruppene med en “repeated measures ANOVA”. Disse fant også at både tung styrketrening ser ut til å være mer effektivt enn eksplosiv styrke-

trening når sammenlignet pre/post. Gruppen med tung styrketrening økte LØ fra pre til post signifikant ( $P < 0.01$ ). Til slutt Beattie et al. (2017) gjennomførte både, pre - mid - posttesting. De brukte “one way repeated measures anova” for hver gruppe, og “two-way repeated measures anova” for å sammenligne mellom gruppene. De fant en signifikant bedring i LØ fra uke 0-20 ( $P < 0.01$ ), men ingen signifikant bedring fra uke 0-40. Det skal sies at utøverne trente styrke 2x ganger i uken fra uke 0-20, og 1x i uken fra uke 20-40. Dette kan være en årsak til at vi så en større effekt de første 20 ukene.

## Konklusjon

Overordnet viser resultatene fra studiene at tung styrketrening har en positiv effekt på arbeidsøkonomi i løping. Disse effektene ser vi både når styrketreningsgruppene har blitt sammenlignet med seg selv pre/post, og når sammenlignet med kontrollgrupper. Ut ifra disse studiene ser det også ut til at tung styrketrening har bedre effekt på LØ enn eksplosiv styrketrening. En svakhet er at ingen av disse studiene inkluderte kvinnelige forsøkspersoner. I nyere studier som undersøker dette tema burde kvinner også inkluderes, for å undersøke om effektene styrketrening har på LØ også er representativt for kvinner. Basert på disse studiene ser det også ut til at to økter i uken er mer optimalt enn en økt i uken.

# Analyse av repeterte målinger

## Introduksjon

Det er stor enighet om at de største adaptasjonene vi får av å trene tung styrketrening er økt styrke og muskelvekst. For utrente personer kan man forvente å øke maksimal muskelstyrke med 1% i uka, men med variasjoner mellom 0,1 - 3 % (McDonagh & Davies, 1984). Tilsvarende ser det ut til at muskeltverrsnittet kan øke med 0,1 - 0,5 % per økt (Wernbom et al., 2007). Det er mange faktorer som kan påvirke hvor store adaptasjonene blir. Noen grunner til at vi finner så stor spredning i respons kan være bruk av ulike treningsformer, ulikt antall serier og repetisjoner på øktene, og hvilke muskler som blir målt (Tønnessen & Rønnestad, 2018).

Det er mye som tyder på at det er en sammenheng mellom styrketreningsvolum og treningsadaptasjoner (Raastad, 2010). Det er likevel stor variasjon i funnene, og det ser ikke ut til at det er noe klar fasit på hvor mye man bør trene for å oppnå gode resultater (Raastad, 2010). For utrente personer er det av interesse å undersøke hvor små treningsmengder man kan gjennomføre, og likevel oppnå gunstige resultater. Dette siden tid ofte er mangelvare hos mange mennesker, og trening fort kan bli en nedprioritert aktivitet (Choi et al., 2017). De lærde strides når det kommer til hvor mange serier utrente personer burde trene for mest gunstig effekt. Cannon & Marino (2010) og Mitchell et al. (2012) viser til at økter med lavt volum gir like gode adaptasjoner i styrke og masse som økter med moderat volum. Likevel er det andre større meta-analyser som favoriserer 3 serier over 1 serie (Krieger, 2009). Denne diskrepansen kan tyde på at det er store individuelle forskjeller på styrketreningsadaptasjoner.

Et forskningsdesign med fokus på intra-individe forskjeller på ulikt treningsvolum, vil kunne være med på å forstå dette tema bedre.

Hensikten med denne studien er å undersøke hvordan effektene en og tre serier påvirker maksimal styrke målt som 1RM beinpress og muskelvekst. Ved at deltakerne trener en serie på det ene beinet og tre serier på det andre kan vi undersøke forskjeller innad i hvert individ.

## Metode

### Deltakere og studiedesign

41 kvinner og menn mellom ble rekrutert til studien, med kriterie om å være mellom 18 og 40, samt ikke-røykende se Tabell 1. Eksklusjonskriterier var intoleranse til lokal bedøvelse, at man hadde trent mer enn en styrkeøkt i uka i snitt de siste 12 månedene før intervensjonen, reduksjon i muskelstyrke grunnet skader eller sykdom, og hvis man går på medisin som kan påvirke adaptasjoner til styrketrening. Alle deltakerne som fullførte minimum 85% av styrkeøkter ble inkludert i analysen.

Intervensjonen besto av 12 uker med fullkroppss styrketrening som ble gjennomført mellom september og november. For å kunne undersøke forskjellene av ulikt treningsvolum innen hver deltaker, ble beinøvelsene gjennomført unilateralt. Beinene til deltakerne ble dermed tilfeldig fordelt til å gjennomføre en serie på det ene beinet, og tre serier på det andre beinet på på hver øvelse. Slik kunne alle deltakerne gjennomføre både 1 og 3 serier. Deltakerne målte muskelstyrke ved baseline, underveis (uke; 3, 5 og 9) og etter intervensjonen. Kroppssammensetningen til deltakerne ble målt rett før og etter treningsintervensjonen.

Tabell 1: Oversikt over de inkluderte deltakerne ved pre-test

Variabler	Gj.snitt $\pm$ SD
N	34
Alder	22.79 $\pm$ 3.02



Høyde	$174.84 \pm 9.98$
Vekt	$69.72 \pm 11.87$

---

## Måling av maksimal styrke i beinpress

Maksimal styrke i unilateral beinpress ble definert som en repetisjon-maks (1RM). Før testen ble det gjennomført en spesifikk oppvarming som besto av ti, seks og tre repetisjoner på 50, 75 og 85% av predikert 1RM. Videre ble 1RM målt med å gradvis øke belastningen til forsøkspersonen ikke lenger klarte å gjennomføre øvelsen med fullt bevegelsesutslag. Den tyngste vekten med godkjent gjennomføring ble definert som 1RM. Hver forsøksperson ble gitt fire til seks forsøk.

## DXA

For å måle mager kroppsmasse og generelle kroppssammensetning ble det benyttet dual-energy X-ray absorptiometry (DXA) (Lunar Prodigy, GE Healthcare, Oslo, Norge). Alle deltakerne skulle være fastende i minimum to timer før DXA-målinger for å så nøyaktige målinger som mulig på mager kroppsmasse. Deltakerne skulle også avstå fysisk anstrengende aktivitet den siste tiden før testen.

## Analyse og statistikk

For å undersøke effektene av ulikt treningsvolum på maksimal styrke og økning i muskelmasse benyttet jeg en linear mixed model (LMM). Ved å bruke LMMs, kan vi sammenligne effektene av ulikt treningsvolum på styrke og muskelmasse, samt undersøke hvordan disse effektene varierer mellom deltakerne og over tid. LMMs ble spesifisert med load/lean mass som avhengig variabel, og med tid og tid til sets-interaksjonen som faste effekter. I tillegg ble modellene spesifisert med random intercepts for hver deltaker. Dette muliggjør estimering av individuelle effekter, slik at vi også kan undersøke hvordan effekten av treningsvolum varierer mellom

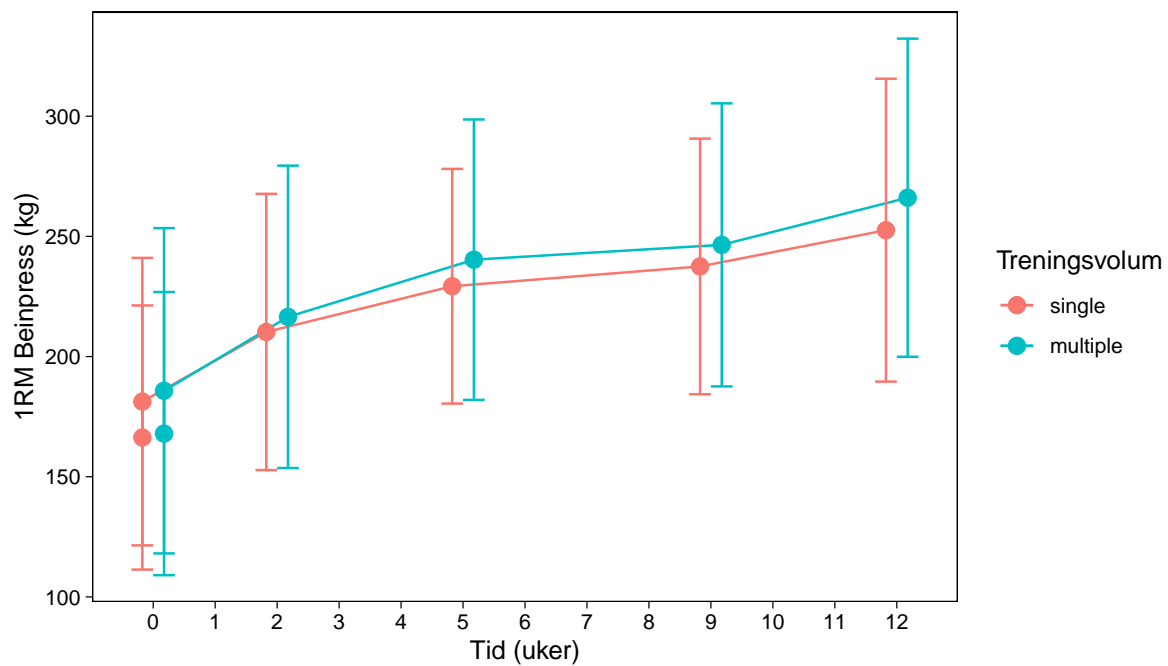
deltakerne. I modellen for muskelmasse ble det også inkludert en random slope på “time” for å ta hensyn til at det er individuell variasjon i både baseline (skjæringspunktet) og endringshastigheten over tid (stigningstallet). Alle deltakerne som har gjennomført minst 85% av øktene i intervensjonen er inkludert i analysen  $N = 34$ .

## Resultater

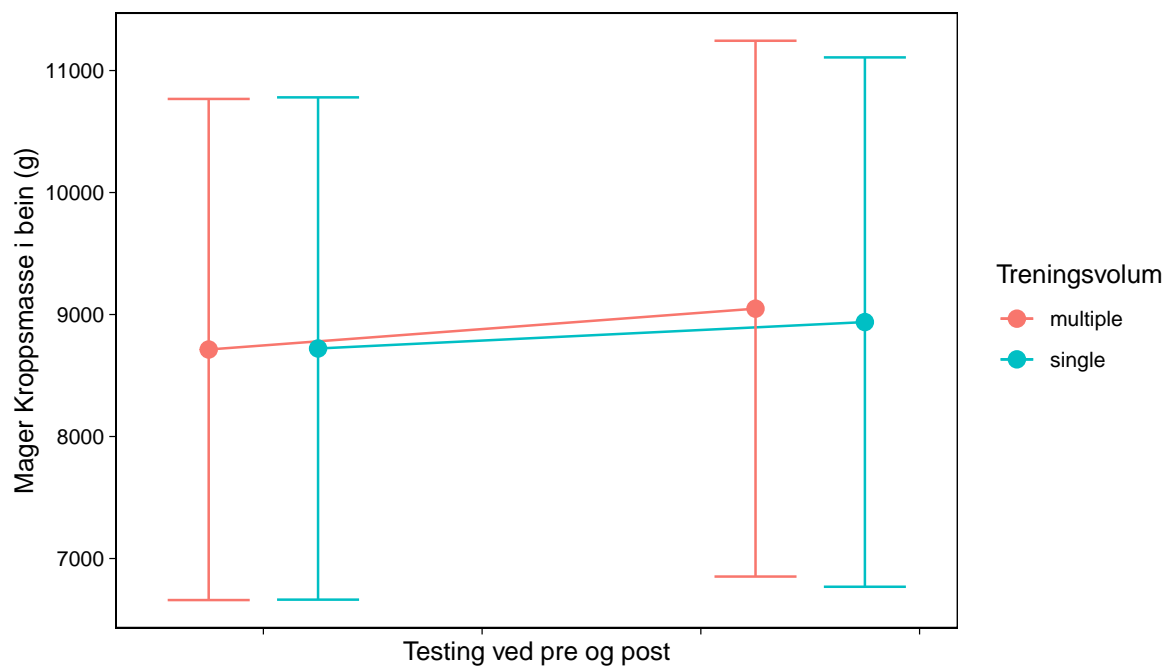
Resultatene viser at både en serie og tre serier øker 1RM i beinpress og muskelveksten i en 12-ukers treningsperiode, se Figur 1 og Figur 2. Økningen i maksimal styrke ser ikke ut til å komme av forskjeller i antall serier, da vi ikke observerte en effekt mellom de to ulike gruppene i 1RM beinpress. Når det gjelder muskelvekst, ble det observert en større økning i mager kroppsmasse hos gruppen som trente 3 serier, mot gruppen som trente 1 serie se Tabell 2.

Tabell 2: Oversikt over effektene av styrketrening med 1 og 3 serier på 1RM beinpress og mager kroppsmasse qnalsert av LMMs. Tid:Gruppemultiple viser til forskjeller av stigningstall mellom gruppene. SE = standardfeil og DF = frihetsgrader.

<b>Koeffisienter</b>	<b>Estimat</b>	<b>SE</b>	<b>DF</b>	<b>t.verdi</b>	<b>p.verdi</b>
<i>1RM Beinpress</i>					
Intercept	182.57	10.19	35.19	17.91	0.00
Tid	6.42	0.35	314.03	18.47	0.00
Gruppemultiple	2.92	3.06	313.98	0.96	0.34
Tid:Gruppemultiple	0.71	0.49	314.00	1.45	0.15
<i>Fettfri Masse</i>					
Intercept	8,589.00	347.80	33.23	24.70	0.00
Tid	166.26	64.85	50.03	2.56	0.01
Gruppemultiple	14.53	41.04	66.00	0.35	0.72
Tid:Gruppemultiple	122.79	58.04	66.00	2.12	0.04



Figur 1: Volumavhengige endringer i 1RM beinpresse i uke 0, 2, 5, 9 og 12



Figur 2: Volumavhengig endringer i mager kroppsmasse mellom pre og posttest

## Diskusjon

Resultatene viser til at over en 12 ukers treningsperiode på utrente deltakere gir 1 og 3 serier tilsvarende lik effekt på økning i 1RM beinpress. Hvis vi ser på muskelvekst, tyder det på at tre serier er fordelaktig sammenlignet med en serie. At tre serier er fordelaktig over 1 serie for muskelvekst stemmer overens med større metastudier som sammenligner ulikt treningsvolum (Krieger, 2009; Schoenfeld et al., 2016). Motstridene til de nevnte studiene, fant ikke den gjeldende studien noen forskjell mellom gruppene når det gjelder styrkeøkning. Dette samsvarer likevel med Cannon & Marino (2010) og Mitchell et al. (2012), som ikke fant en fordel med flere serier på styrke hos utrente. Det er viktig å nevne at deltakerne i denne studien er utrente person, og det kan se ut til at det er viktigere med flere serier jo bedre trent man er (Raastad, 2010). Denne studien viser likevel at både 1 og 3 serier gir god progresjon, så hvis tid er mangelvare vil 1 serie gi gunstige resultater om man er på et lavt treningsnivå

## Konklusjon

Tallene fra denne studien indikerer at en og tre serier under styrketrening gir tilnærmet lik økning i 1RM beinpress for utrente personer over en 12-ukers periode. Likevel ser vi en større effekt på økning i muskelmasse på tre serier sammenlignet med en. Dette indikerer at for utrente personer er en serie per øvelse tilstrekkelig for å oppnå god progresjon, men tre serier kan være fordelaktig om muskelvekst er målet.

## References

- Andersen, J. L., & Aagaard, P. (2000). Myosin heavy chain IIX overshoot in human skeletal muscle. *Muscle & Nerve*, 23(7), 1095–1104. [https://doi.org/10.1002/1097-4598\(200007\)23:7%3C1095::AID-MUS13%3E3.0.CO;2-O](https://doi.org/10.1002/1097-4598(200007)23:7%3C1095::AID-MUS13%3E3.0.CO;2-O)
- Beattie, K., Carson, B. P., Lyons, M., Rossiter, A., & Kenny, I. C. (2017). The Effect of Strength Training on Performance Indicators in Distance Runners. *Journal of Strength and Conditioning Research*, 31(1), 9–23. <https://doi.org/10.1519/JSC.0000000000001464>
- Cannon, J., & Marino, F. E. (2010). Early-phase neuromuscular adaptations to high- and low-volume resistance training in untrained young and older women. *Journal of Sports Sciences*, 28(14), 1505–1514. <https://doi.org/10.1080/02640414.2010.517544>
- Choi, J., Lee, M., Lee, J., Kang, D., & Choi, J.-Y. (2017). Correlates associated with participation in physical activity among adults: A systematic review of reviews and update. *BMC Public Health*, 17(1), 356. <https://doi.org/10.1186/s12889-017-4255-2>
- Godfrey-Smith, P. (2003). *Theory and reality : An introduction to the philosophy of science*. Kapiteltittel: Bayesianism and Modern Theories of Evidence. Kapitel 14, s.202-218. University of Chicago Press.
- Guglielmo, L., Greco, C., & Denadai, B. (2009). Effects of Strength Training on Running Economy. *International Journal of Sports Medicine*, 30(01), 27–32. <https://doi.org/10.1055/s-2008-1038792>
- Halperin, I., Pyne, D. B., & Martin, D. T. (2015). Threats to internal validity in exercise science: A review of overlooked confounding variables. *International Journal of Sports Physiology and Performance*, 10(7), 823–829. <https://doi.org/10.1123/ijsp.2014-0566>

- Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science: *Sports Medicine*, 30(1), 1–15. <https://doi.org/10.2165/00007256-200030010-00001>
- Hulley, S. B. (Ed.). (2013). *Designing clinical research* (4th ed). Wolters Kluwer/Lippincott Williams & Wilkins.
- K. Hackenberger, B. (2019). Bayes or not Bayes, is this the question? *Croatian Medical Journal*, 60(1), 50–52. <https://doi.org/10.3325/cmj.2019.60.50>
- Krieger, J. W. (2009). Single Versus Multiple Sets of Resistance Exercise: A Meta-Regression. *Journal of Strength and Conditioning Research*, 23(6), 1890–1901. <https://doi.org/10.1519/JSC.0b013e3181b370be>
- Kuang, J., Yan, X., Genders, A. J., Granata, C., & Bishop, D. J. (2018). An overview of technical considerations when using quantitative real-time PCR analysis of gene expression in human exercise research. *PLOS ONE*, 13(5), e0196438. <https://doi.org/10.1371/journal.pone.0196438>
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method. *Methods*, 25(4), 402–408. <https://doi.org/10.1006/meth.2001.1262>
- McDonagh, M. J. N., & Davies, C. T. M. (1984). Adaptive response of mammalian skeletal muscle to exercise with high loads. *European Journal of Applied Physiology and Occupational Physiology*, 52(2), 139–155. <https://doi.org/10.1007/BF00433384>
- Millet, G. P., Jaouen, B., Borrani, F., & Candau, R. (2002). Effects of concurrent endurance and strength training on running economy and  $\dot{V}O_{2\max}$  kinetics: *Medicine & Science in Sports & Exercise*, 34(8), 1351–1359. <https://doi.org/10.1097/00005768-200208000-00018>
- Mitchell, C. J., Churchward-Venne, T. A., West, D. W. D., Burd, N. A., Breen, L., Baker, S. K., & Phillips, S. M. (2012). Resistance exercise load does not determine training-mediated hypertrophic gains in young men. *Journal of Applied Physiology (Bethesda, Md.: 1985)*, 113(1), 71–77. <https://doi.org/10.1152/jappphysiol.00307.2012>
- Popper, K. (1959). *The Logic of Scientific Discovery*. Hutchinson & Co.
- Raastad, T. (2010). *Styrketrening: i teori og praksis*. Gyldendal Norsk forlag.

- Schoenfeld, B. J., Ogborn, D., & Krieger, J. W. (2016). Effects of Resistance Training Frequency on Measures of Muscle Hypertrophy: A Systematic Review and Meta-Analysis. *Sports Medicine*, 46(11), 1689–1697. <https://doi.org/10.1007/s40279-016-0543-8>
- Spiegelhalter, D. J. (2020). *The art of statistics: Learning from data* (Published in paperback). Pelican, an imprint of Penguin Books.
- Støren, Ø., Helgerud, J., Støa, E. M., & Hoff, J. (2008). Maximal Strength Training Improves Running Economy in Distance Runners. *Medicine & Science in Sports & Exercise*, 40(6), 1087–1092. <https://doi.org/10.1249/MSS.0b013e318168da2f>
- Taipale, R., Mikkola, J., Nummela, A., Vesterinen, V., Capostagno, B., Walker, S., Gitonga, D., Kraemer, W., & Häkkinen, K. (2010). Strength Training in Endurance Runners. *International Journal of Sports Medicine*, 31(07), 468–476. <https://doi.org/10.1055/s-0029-1243639>
- Tanner, R. K., Gore, C. J., & Australian Institute of Sport (Eds.). (2013). *Physiological tests for elite athletes* (2nd ed). Human Kinetics.
- Terzis, G., Georgiadis, G., Stratakos, G., Vogiatzis, I., Kavouras, S., Manta, P., Mascher, H., & Blomstrand, E. (2007). Resistance exercise-induced increase in muscle mass correlates with p70S6 kinase phosphorylation in human subjects. *European Journal of Applied Physiology*, 102(2), 145–152. <https://doi.org/10.1007/s00421-007-0564-y>
- Tønnessen, E., & Rønnestad, B. (2018). *Trening fra barneidrett til toppidrett*. Gyldendal Olympiatoppen.
- Wernbom, M., Augustsson, J., & Thomee, R. (2007). The Influence of Frequency, Intensity, Volume and Mode of Strength Training on Whole Muscle Cross-Sectional Area in Humans: *Sports Medicine*, 37(3), 225–264. <https://doi.org/10.2165/00007256-200737030-00004>