

Mappeeksamen IDR4000: Kvantitativ metode og statistikk

Kandidatsnummer: 516. Antall ord: 16 113

2024-11-22

Table of contents

1	Innhold	3
2	Reliabilitet og verktøy for reproduserbar vitenskap	4
2.1	Metode	4
2.1.1	Utstyr og kalibrering	5
2.1.2	Forberedelser av testdeltakeren	6
2.1.3	Testgjennomføring	6
2.1.4	Datahåndtering	6
2.2	Resultat	7
2.3	Diskusjon	10
2.4	Konklusjon	11
2.5	Vedlegg: Koder brukt	12
3	Regresjonsmodeller, predikasjon og tolkning	16
3.1	Del 1: Laktat terskler	16
3.1.1	Metode	17
3.1.2	Resultat	17
3.1.3	Diskusjon	19
3.2	Del 2: Forutsi størrelser på DNA fragmenter eller stiningene i en qPCR-kalibreringskurve	21
3.2.1	Metode	22
3.2.2	Resultater	23
3.2.3	Diskusjon	24
3.3	Del 3: Tolkning av regresjonsmodell	25
3.3.1	Metode	25
3.3.2	Resultat	25
3.3.3	Diskusjon	27
3.4	Vedlegg: Koder brukt	28
4	Statistisk inferens, statistiske modeller og statistisk styrke	37
4.1	Simulering til oppgave 1-4	37
4.1.1	Spørsmål 1 - Forklaring av regresjonsresultater: Estimat, SE, t-verdi og p-verdi	37
4.1.2	Spørsmål 2 - Årsaker til forskjeller mellom studieresultatene i modellene m1 og m2.	39

4.1.3	Spørsmål 3 - Betydningen av skyggeområdene i halområdene til t-fordelingen i Figure 4.1.	40
4.2	Simulering til spørsmål 4-7	40
4.2.1	Spørsmål 4 - Sammenligning av standardavvik og SE for utvalgsstørrelser ($n = 8$ og $n = 40$): Forklaring og definisjon.	40
4.2.2	Spørsmål 5 - Histogram av p-verdier og effekten av utvalgsstørrelse på statistisk styrke	41
4.2.3	Spørsmål 6 - Antall studier med signifikante effekter for hver utvalgsstørrelse	42
4.2.4	Spørsmål 7 - Beregning av statistisk styrke for en t-test med ulike utvalgsstørrelser ved bruk av pwr -pakken.	42
4.2.5	Spørsmål 8 - Antall falske positive resultater ved 5 % signifikansnivå ved gjentatte studier.	43
4.3	Vedlegg: Koder brukt	44
5	Studiedesign	49
5.1	Analyse	49
5.1.1	Problemstilling og hypoteser	50
5.2	Metode	51
5.2.1	Valg av studiedesign	51
5.2.2	Utvalg	52
5.2.3	Tester og sentrale variabler	53
5.2.4	Statistiske tester	53
5.3	Resultat	54
5.4	Konklusjonene fra studiene	55
5.5	Anbefaling til fremtidige studier	55
6	Analyse av eksperimenter med repeterte målinger	57
6.1	Metode	58
6.1.1	Deltagere og studieoversikt	58
6.1.2	Styrketreningsprotokoll	59
6.1.3	Målinger av maksimal muskelstyrke og muskelens tverrsnittsareal	60
6.1.4	Data analyse og statistikk	60
6.2	Resultater	61
6.3	Endring i muskelstyrke	62
6.4	Endring i muskeltverrsnitt	63
6.5	Diskusjon	64
6.6	Begrensninger	65
6.7	Implikasjoner for fremtidig forskning	65
6.8	Konklusjon	65

7	Arbeidskrav i vitenskapsfilosofi	75
7.1	Spørsmål 1	75
7.1.1	Innledning	75
7.1.2	Humes argument mot induksjon	75
7.1.3	Innvending mot Hume sitt argument	76
7.1.4	Humes respons på innvendinga	76
7.2	Spørsmål 2	76
7.2.1	Innledning	76
7.2.2	Beskrivelse av falsifikasjonisme og Poppers motivasjon	77
7.2.3	Utfordring med falsifikasjonisme	77
7.2.4	Poppers løsning	78
7.2.5	Konklusjon	78
8	Protein ekstraksjon og analyse	79
8.1	Metode	80
8.1.1	Homogenisering av muskelprøver	80
8.1.2	Western blot protokoll	81
8.2	Resultater	83
8.3	Diskusjon	84
8.4	Konklusjon	85
8.5	Oppløsninger	85
8.6	Koder brukt	85
	Referanser	88

1 Innhold

Mappen består av flere oppgaver som dekker sentrale temaer i kurset. Oppgavene er utviklet gjennom semesteret med tilbakemeldinger fra medstudenter og lærere, og de er videre bearbeidet og forbedret for inkludering i denne mappen. Mappeeksamen er strukturert som en Quarto-bok, der hver oppgave utgjør et eget kapittel. Dette gir en helhetlig presentasjon av arbeidet som er gjennomført.

- Rapport: “Reliabilitet og verktøy for reproduserbar vitenskap”
- Rapport: “Regresjonsmodeller, predikasjon og tolkning”
- Rapport: “Statistisk inferens, statistiske modeller og statistisk styrke”
- Rapport: “Studiedesign”
- Rapport: “Analyse av eksperimenter med repeterte målinger”
- Arbeidskrav i vitenskapsteori
- Laborasjonsrapport fra molekylærlab: “Protein ekstraksjon og analyse”

Gjennom disse oppgavene demonstreres forståelse og anvendelse av metodiske og teoretiske konsepter som er sentrale i kurset, med fokus på både praktiske ferdigheter og kritisk refleksjon. De mest sentrale kodene er lagt ved til slutten av hvert del kapittel. Kodene, tabeller og referanselist er ekskludert fra ordtellingen.

2 Reliabilitet og verktøy for reproducerbar vitenskap

I 2005 satte John Ioannidis forskningsverden på prøve med utsagnet om at de fleste publiserte forskningsfunn er feilaktig (Ioannidis, 2005). Han fremhevet mangel på reliabilitet som en sentral utfordring, der forskeres fleksibilitet i studiedesign og analyser bidrar til resultater som ikke kan reproduseres. Reliabilitet referer her til muligheten for å gjenskape studier og oppnå lignende resultater (Spiegelhalter, 2019a, pp. s.342–348).

Reliabilitet av en test kan kvantifiseres på flere måter, blant annet som variasjonen innen samme person, endring i gjennomsnitt og retest-korrelasjon (korrelasjonen mellom målinger ved gjentatte tester) (Hopkins, 2000, p. s.2). Endring i gjennomsnitt viser hvordan tilfeldige eller systematiske faktorer, som læringseffekter, påvirker forskjellen mellom to tester. Retest-korrelasjon vurderer hvor godt en test reproducerer rangeringen av deltakere ved gjentatte forsøk. (Hopkins, 2000, p. s.2).

Den tredje måten, variasjon innen samme person, er særlig viktig for å forstå reliabiliteten til målinger. Denne variasjonen fanger opp tilfeldige forskjeller mellom målinger på samme individ og kvantifiseres som standardavviket til de individuelle verdiene, også kjent som “typical error” eller “standard error of measurement” (Hopkins, 2000, p. s.2). “Typical error”, på norsk typisk feil, kan dermed forstås som variasjonen man vil forvente å observere fra forsøk x til forsøk y ved repeterte forsøk hos samme testdeltakere. En nyttig måte å presentere typisk feil på, er å uttrykke det som prosentdel av gjennomsnittsverdien (Hopkins, 2000, p. s.2). Dette åpner opp for sammenligning av en heterogen gruppe der prestasjonsnivået er ulikt, da den absolutte verdien av typisk feil vil gjerne øke sammen med måleverdien.

I denne rapporten vil VO₂max-maks data fra fysiologi-laben bli analysert for å kvantifisere reliabilitet i form av typisk feil av maksimal watt oppnådd under testen. Gjennomføring og standardisering av testene vil bli forklart under metode, mens resultatene vil bli presentert med hovedvekt på betydningen av “typical error” for reliabiliteten til målingene.

2.1 Metode

Dataene som skal brukes for å måle reliabilitet er hentet fra 16 ulike testpersoner som gjennomførte gjentatte VO₂max-makstest på ergospirometri på fysiologilaben. Det ble satt opp slik at alle skulle gjennomføre testen fire ganger spredd utover to uker, men det var kun åtte

stk som gjennomførte fire tester. Det ble gjort antropometeriske mål av testpersonene før hver test, men i Table 2.1 er det kun tatt utgangspunkt i målingene før første test. Totalt var det 14 stk som hadde baseline målinger som ble inkludert i Table 2.1.

Tabell 2.1: Karakteristikkene på testpersonene

	Kvinne	Mann
N	3	11
Alder	23.3 (0.6)	25.9 (4.0)
Vekt	71.7 (9.9)	81.2 (11.9)
Stature	171 (8)	182 (6)

Deltakerne ble delt opp i grupper, der en var testdeltaker, en skulle være testansvarlig, og de resterende var observatører som også hjalp testansvarlig med praktiske oppgaver. Det ble rult på hvem som var testperson, testansvarlig og observatører, slik at alle fikk være testansvarlig.

2.1.1 Utstyr og kalibrering

For å måle den metabolske responsen under testen ble miksekammeret fra Vyntus CPX benyttet. Nødvendig volum- og gasskalibrert ble utført før hver test i henhold til produsenten sine retningslinjer. Ved differanser på over 2% under volumkalibrering eller over 0.2% under gasskalibrering ble nødvendig feilsøking igangsatt. Dette inkluderte kontroll for skader på turbinen eller sampleslangen, sjekk av gasskoblinger, samt inspeksjon for fukt i sampleslange og andre relevante faktorer.

I tillegg til måling av metabolsk respons ble laktatnivå målt etter ett minutt etter testslutt. Blodprøver ble samlet fra pekefingeren ved hjelp av en lansett, og laktat ble analysert med Biosen fra EKF Diagnostics. Apparatet hadde en innebygd timer som varslet behov for ny kalibrering hver time.

Testpersonene syklet på en ergometersykkel fra Lode. Tilpasning av sykkelen, som høyde på sete eller avstand mellom sete og styre, ble gjort før den første testen. Disse innstillingene ble lagret i Lode sin programvare, slik at hver testperson hadde standardisere sykkelinnstillinger ved alle tester. Alle sammen skulle benytte seg av MTB-pedaler (“Mountain Bike-pedaler”) og ha 172,5 mm lengde på kranarmene. En vifte ble plassert foran sykkelen for å gi testpersonene mulighet til avkjøling underveis i testen. Samme innstilling på viften ble brukt ved hver test for å sikre standardisering. Puls ble registrert ved hjelp av et Garmin pulselte, som var tilkoblet programvaren SentrySuite fra Vyntusen.

2.1.2 Forberedelser av testdeltakeren

Før testingen ble det etablert retningslinjer for mat- og drikkeinntak, døgnrytme, trening og oppvarming. Testpersonene ble oppfordret til å opprettholde normal døgnrytme gjennom de to ukene testene pågikk. Det ble spesifisert at siste hovedmåltid skulle inntas senest to timer før teststart, men det var tillat å spise en mindre karbohydratkilde som en banan, en bar eller en gel, opptill 30 min før testen. Deltakerne ble også bedt om å opprettholde sitt vanlige koffeininntak og sørge for å være i væskebalanse. oppvarmingen besto av 5 min på ergometersykel, hvor intensiteten gradvis ble økt: de første 2 minuttene på en bestemt watt (for eksempel 150W), de neste 2 minuttene på høyere intensitet (175W), og det siste minuttet, eksempelvis på endå høyere intensitet (220W). Selv om intensiteten ble tilpasset individuelt, skulle den samme oppvarmingsprotokoll følges ved hver test.

2.1.3 Testgjennomføring

Testen startet på en gitt watt (motstand), som for hvert minutt økte gradvis helt til testpersonen ikke mestret å sykle lengre (ramp protokoll). Hva watt man startet på og hvor mye den økte per minutt, varierte fra person til person på bakgrunn av treningsbakgrunn. Hele testen skulle gjennomføres sittende, og ved en tråkkfrekvens på under 60 rpm (rate per minute) eller om testpersonen reiste seg fra setet, ble testen avsluttet. Testansvarlig skulle kun informere om hvor lenge det var til neste økning og hvilken watt det ble syklet på. Informasjon som puls, V2max eller andre parametre fikk testpersonen ikke tilgang til underveis i testen. Mot slutten av testen skulle testansvarlig bidra til å motivere testpersonen slik at han/hun tok ut sitt ytterste. Hvordan testansvarlige valgte å løse dette varierte, men det viktigste var at han/hun oppførte seg tilnærmet lik ved hver test ovenfor den som syklet.

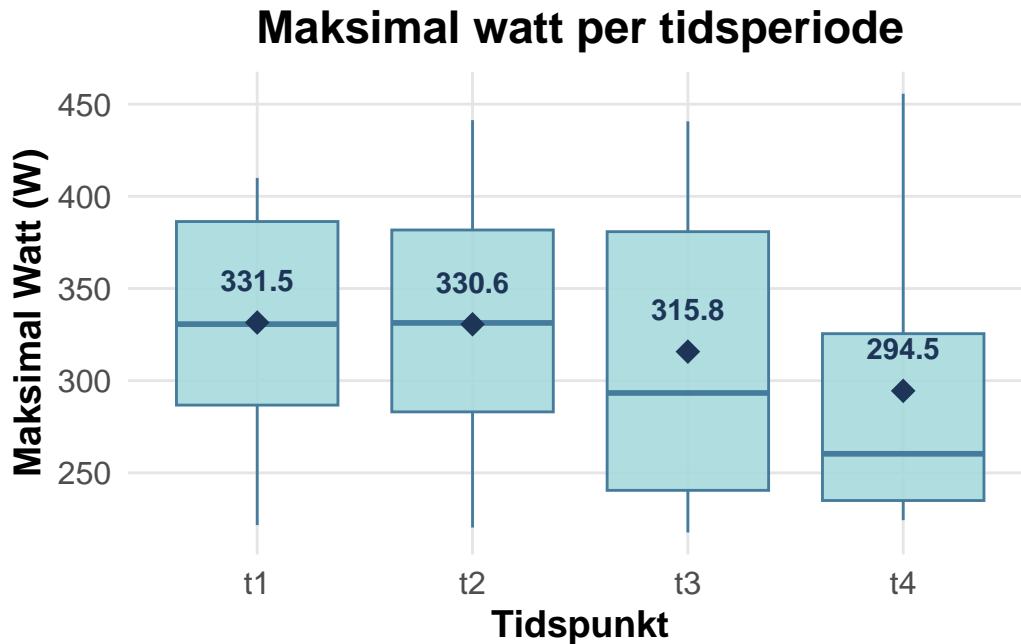
Rett etter testslutt ble testpersonen spurt om Borg skala for å beskrive grad av anstrengelse. Han/hun ble spurt på å oppgi et tall mellom 6-20 på hvor sliten han/hun er, der 6 er “ingen anstrengelse” og 20 er “maksimal anstrengelse”. Laktatprøve ble tatt etter ett minutt, og plassert i beger med løsning før den ble vendt og satt i Biosen for måling. Testansvarlig noterte ned hvor lenge testpersonen syklet og hva watt han/hun endte på.

2.1.4 Datahåndtering

Rådata om metabolsk respons ble samlet inn med programvare fra Vyntus og eksportert til Excel for en første organisering. Testens varighet og sluttwatt ble manuelt registrert av testansvarlig i samme excel-fil. Videre datahåndtering, som filtrering, visualisering og beregninger av gjennomsnitt, standardavvik og “typical error”, ble utført i R (versjon 4.4.1). Manglende data ble håndtert ved ekskludering av ufullstendige rader for variabler som var sentrale for analysene.

2.2 Resultat

Gjennomføringen av de fire VO2-maks-testene viste en nedadgående trend i maksimal watt-verdier fra t1 til t4. Som vist i Figure 2.1, falt gruppens gjennomsnittlige og median watt-verdier fra første test til fjerde test. Samtidig er det en betydelig større spredning av målingene i test tre og fire sammenlignet med første og andre test. Antall deltakere som fullførte tester ved de ulike tidspunktene avtok gradvis fra t1 (n = 14) og t2 (n = 14) til t3 (n = 11) og t4 (8 deltakere).



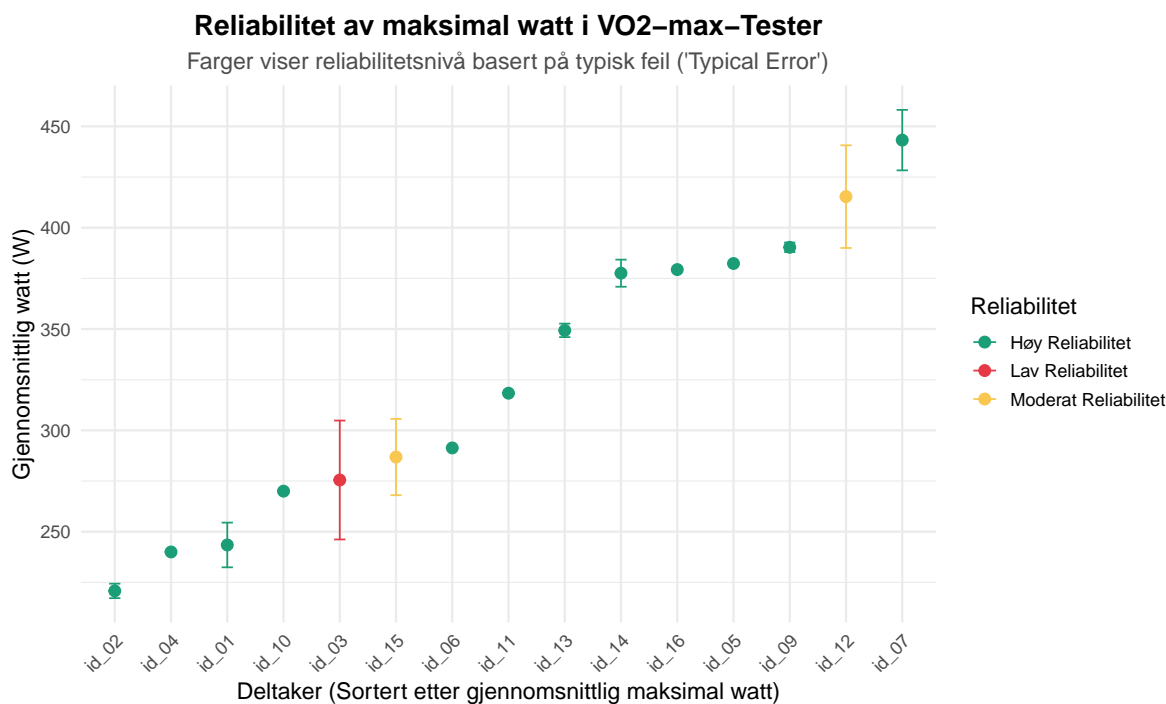
Figur 2.1: Gjennomsnittlige watt-verdier for hele gruppen for hver testforsøk.

For å forstå variasjonene på individnivå, ble maksimal watt analysert for hver deltaker. Figure 2.2 viser gjennomsnittlig watt for hver deltaker, sortert fra lavest til høyest, med feilstenger som representerer typisk feil for å visualisere variasjonen mellom målinger for hver deltaker. Typisk feil gir en indikasjon på den forventede variasjonen i watt ved en ny test. Fargekodingen angir enten høy (CV ≤ 3%), moderat (CV > 3% og ≤ 5%) og lav (CV > 5%) reliabilitet (Hopkins, 2000). Generelt viste de fleste testpersonene høy reliabilitet, mens noen hadde større variasjon og lavere reliabilitet mellom målingene. Det er viktig å påpeke at Figure 2.2 kun inkluderer testpersoner som har gjennomført to eller flere tester, ettersom dette er nødvendig for å kunne beregne typisk feil

$$\text{Typical Error} = \frac{\text{SD of differences}}{\sqrt{2}}$$

og CV

$$CV (\%) = \frac{\text{Typical Error}}{\text{Mean Watt}} \times 100$$



Figur 2.2: Gjennomsnittlige wattverdier med typisk feil fra alle testforsøk for samtlige testpersoner.

Table 2.2 gir en detaljert oversikt over typisk feil og CV for hver deltaker, beregnet basert på alle tester gjennomført av den enkelte. De er rangert etter gjennomsnittlig watt, og fargene følger samme rangering mellom høy (grønn, CV $\leq 3\%$), moderat (gul, CV $> 3\%$ og $\leq 5\%$) og lav (rød) (CV $> 5\%$) reliabilitet. Som i Figure 2.2, viser Table 2.2 generelt høy reliabilitet blant deltakerne, med koeffisient av variasjon (CV) $\leq 3\%$ for flertallet. Id_09, id_11, id_16, og id_10 skiller seg ut med svært lav CV ($< 1\%$), noe som indikerer konsistente målinger mellom testene. Imidlertid ble betydelig variasjon observert for id_03 (CV = 10.50 %) og id_12 (CV = 5.08 %), som begge har lav reliabilitet. Id_01 og id_04 ligger i kategorien moderat reliabilitet med CV mellom 3 % og 5 %. Det er verdt å merke seg at id_11 ikke viser noen variasjon mellom målingene (CV = 0 %), som kan indikere perfekt samsvar mellom testene, men bør vurderes for mulig feilregistrering. Mens id_08 mangler tilstrekkelige data for beregning av typisk feil og CV.”

For å evaluere reliabiliteten og konsistensen i målingene over tid, har det blitt gjort parvise sammenligninger av målinger på to spesifikke tidspunkter. I Table 2.3 så referer testpar til

Tabell 2.2: Testpersonenes typiske feil (typical error) og koeffisient av variasjon (CV).

Sammenligning av typisk feil og CV

Hver rad representerer en deltaker

Deltaker-ID	Gj.snitt watt	Typisk feil	CV (%)
id_07	434.90	13.53	3.11
id_12	406.90	20.68	5.08
id_09	389.10	1.52	0.39
id_05	386.70	4.33	1.12
id_16	381.50	2.17	0.57
id_14	379.50	4.04	1.06
id_13	346.60	4.91	1.42
id_11	318.30	0.00	0.00
id_15	292.90	8.73	2.98
id_03	290.40	30.50	10.50
id_06	286.00	5.33	1.86
id_10	269.40	0.82	0.31
id_01	242.70	8.99	3.71
id_04	231.00	9.00	3.90
id_08	230.00	NA	NA
id_02	221.00	1.96	0.89

Tabell 2.3: Gjennomsnittlig typisk feil og CV av alle testpersonen for hvert testpar.

Typisk feil og koeffisient av variasjon (CV) per testpar
Kun basert på w.max for tidspunktene t1, t2, t3 og t4

Testpar	Typisk Feil	Gj.snitt CV (%)
t1 -> t2	12.59	2.53
t2 -> t3	29.53	5.45
t3 -> t4	8.96	4.43

denne parvise sammenligningen mellom målinger utført på to påfølgende tidspunkt for samme deltaker. I denne analysen har man sett på første test og andre test (t1 -> t2), andre test til tredje test (t2 -> t3) og tredje test til fjerde test (t3 -> t4). Typisk feil og koeffisient av variasjon (CV) har blitt beregnet basert på forskjellen mellom målingene i hvert par i tråd med anbefalingene til Hopkins et al. (2000) (Hopkins, 2000, p. s.11). Samme fargekode som er blitt brukt i de andre tabellene og figurene for å illustrere grad av reliabilitet, er også brukt i tabellen under.

Basert på funnene i Table 2.3 viste testparet t1 -> t2 lavest CV (2.53 %), noe som indikerer høy reliabilitet mellom disse målingene, mens t2 -> t3 hadde høyest CV (5.45 %), som tilsier lav reliabilitet og større variasjon. Antall testpersoner i testparene varierte, med 14 personer for testpar t1 -> t2, 10 personer for t2 -> t3, og 6 personer for t3 -> t4. Denne variasjonen kan påvirke beregningene av typisk feil og CV, spesielt i testpar med færre deltakere. Det er derfor mulig at den lavere reliabiliteten observert for t3 -> t4 delvis skyldes usikkerhet forårsaket av den mindre utvalgsstørrelsen.

2.3 Diskusjon

Denne rapporten undersøkte reliabiliteten i målinger av maksimal watt hentet fra VO2max-maks-tester, med fokus på typisk feil og koeffisienten av variasjon (CV) som mål på reliabilitet. Resultatene viste generelt høy reliabilitet, men med variasjon mellom testparene. Dette indikerer at reliabiliteten kan påvirkes av både biologiske faktorer, som fatigue og fokus, og teknologiske faktorer, som variasjoner i utstyr eller testledere.

Typisk feil reflekterer variasjon fra flere kilder, inkludert biologiske faktorer som tretthet og motivasjon, samt teknologiske faktorer som kalibrering av utstyr og operatørens tilnærming (Hopkins et al., 2000) (Halperin et al., 2015; Hopkins, 2000, p. s.2). Selv om standardisering av oppvarming, protokoller og testforhold ble gjennomført, er det vanskelig å eliminere alle kilder til variasjon. Typisk feil gir derfor et mål på denne variasjonen ved gjentatte målinger. I denne analysen ble typisk feil beregnet som standardavviket av differansen mellom påfølgende målinger, justert med faktoren $\sqrt{2}$:

$$TE = \frac{SD_{\text{diff}}}{\sqrt{2}}$$

Denne tilnærmingen ble valgt for å fokusere på variasjonen mellom målinger, uavhengig av gjennomsnittsverdien for maksimal watt (Hopkins, 2000, p. s.3). Ved å fjerne innflytelsen fra prestasjonsnivået, gir denne metoden et mer rettferdig grunnlag for å sammenligne reliabiliteten mellom testparene, særlig når deltakernes prestasjonsnivå varierer.

Selv om typisk feil er nyttig, har den en begrensning: Den øker med måleverdien, noe som gjør det vanskelig å sammenligne reliabiliteten mellom deltakere med ulik prestasjonsevne. For eksempel vil en utøver med høy maksimal watt ha en høyere absolutt typisk feil enn en utøver med lavere watt (Hopkins, 2000). For å håndtere dette ble koeffisienten av variasjon (CV) brukt som et supplerende mål:

$$CV(\%) = \frac{\text{Typical Error}}{\text{Mean}} \times 100$$

CV er dimensjonsløs og gir et mål på variasjonen i prosent av gjennomsnittet, noe som gjør det mulig å sammenligne reliabiliteten mellom deltakere, testforhold og utstyr (Hopkins, 2000, p. 3). Dette var særlig viktig i denne analysen for å evaluere systematiske forskjeller i reliabilitet mellom testparene og over tid.

En begrensning er variasjon i antall deltakere mellom testparene, noe som kan ha påvirket presisjonen i utregninge av typisk feil og CV. For eksempel var det bare 6 personer som ble brukt i analysen av testpar t3 -> t4, noe som fører til økt usikkerhet i estimatene. I tillegg vil biologiske og teknologiske faktorer som dagsform, tretthet, variasjon i kalibrering av utstyr og testleders tilnærming, ha bidratt til variasjon, til tross for standardisering.

Samtidig som det er flere ting som kan ha påvirket estimatene, vil standardiseringen som ble gjort minimere for ytre påvirkninger. I tillegg vil kombinasjonen av typisk feil og CV gi et mer nyansert bilde av variasjonen i målingene. Bruken av $TE = \frac{SD_{\text{diff}}}{\sqrt{2}}$ gir et bedre sammenligningsgrunnlag for å analysere reliabilitet over tid mellom ulike testpar, ved å fjerne innflytelsen fra gjennomsnittsverdien.

2.4 Konklusjon

Denne rapporten evaluerte reliabiliteten i målinger av maksimal watt fra VO2max-maks-tester ved bruk av typisk feil og CV. Analysen viste generelt høy reliabilitet, men med variasjon mellom testparene som kan tyde på biologiske og teknologiske faktorer. For å muliggjøre sammenligning på tvers av testpersonene og over ulike tidspunkter av testing, ble CV brukt som et dimensjonsløst mål sammen med typisk feil. Standardisering av protokoller bidro til å redusere variasjon og øke reproduserbarheten.

2.5 Vedlegg: Koder brukt

```
# Laste opp nødvendige pakker

library(readxl)
library(tidyverse)
library(knitr)
library(gt)

# Innlesing av datasett
df_all <- read_xlsx("data/plot_data_all_gr.xlsx") %>%
  mutate(w.max = as.numeric(w.max),
         la.max = as.numeric(la.max),
         hr.max = as.numeric(hr.max),
         rer.max = as.numeric(rer.max))

# Opprydding av datasett til tabell 2.1
baseline_df_all <- df_all %>%
  filter(timepoint == "t1") %>% # Filtrer for timepoint == t1
  select(sex, age, weight, height) %>% # Velg kolonner for tabellen
  group_by(sex) %>% # Gruppér etter kjønn
  summarise(
    N = n(), # Antall deltakere (unik for t1)
    Alder = sprintf("%.1f (%.1f)", mean(age, na.rm = TRUE),
                      sd(age, na.rm = TRUE)),
    Vekt = sprintf("%.1f (%.1f)", mean(weight, na.rm = TRUE),
                      sd(weight, na.rm = TRUE)),
    Stature = sprintf("%.0f (%.0f)", mean(height, na.rm = TRUE),
                      sd(height, na.rm = TRUE)),
    .groups = "drop"
  ) %>%
  # Konverter alle kolonnene til karakter for å unngå datatypeproblemer i pivot_longer
  mutate(across(everything(), as.character)) %>%
  pivot_longer(cols = c(N, Alder, Vekt, Stature), names_to = "Variable",
              values_to = "Value") %>%
  pivot_wider(names_from = sex, values_from = Value)

# Endre kolonnenavn fra k og m til Kvinne og Mann
baseline_df_all <- baseline_df_all %>%
  rename(Kvinne = k, Mann = m)
```

```

# Utregninger for figur 2.1

# Beregn differanser mellom tidspunkter
df_typical <- df_all %>%
  arrange(id, timepoint) %>%
  group_by(id) %>%
  mutate(
    diff_watt = w.max - lag(w.max), # Beregn differanser
    # Parvise tidspunkter
    time_pair = paste(lag(timepoint), timepoint, sep = " -> ")
  ) %>%
  filter(!is.na(diff_watt)) %>% # Fjern NA i differanser
  ungroup()

# Beregn typisk feil og CV basert på parvise tidsperioder
typical_error_summary <- df_typical %>%
  group_by(id) %>%
  summarise(
    mean_watt = mean(w.max, na.rm = TRUE), # Gjennomsnittlig watt
    # Typisk feil basert på differanser
    typical_error = sd(diff_watt, na.rm = TRUE) / sqrt(2),
    cv = (typical_error / mean_watt) * 100, # CV i prosent
    .groups = "drop"
  ) %>%
  mutate(
    # Sett typisk feil til 0 hvis NA
    typical_error = ifelse(is.na(typical_error), 0, typical_error),
    cv = ifelse(is.na(cv), 0, cv) # Sett CV til 0 hvis NA
  ) %>%
  arrange(mean_watt) %>%
  mutate(ID = factor(id, levels = id)) # Bevar sorteringsrekkefølgen

# Klassifiser reliabilitet basert på CV
typical_error_summary <- typical_error_summary %>%
  mutate(reliability = case_when(
    cv <= 5 ~ "Høy Reliabilitet", # CV 5%
    cv > 5 & cv <= 10 ~ "Moderat Reliabilitet", # CV > 5% og 10%
    cv > 10 ~ "Lav Reliabilitet" # CV > 10%
  ))

# Filtrer kun deltakere med gyldige verdier
typical_error_summary <- typical_error_summary %>%

```

```

filter(!is.na(mean_watt))

# Beregn gjennomsnitt, typisk feil og CV for hver deltaker
summary_tf_cv <- df_all %>%
  group_by(id) %>%
  summarise(
    mean_watt = mean(w.max, na.rm = TRUE),
    typical_error = sd(w.max, na.rm = TRUE) / sqrt(2), # Typisk feil
    cv = (typical_error / mean_watt) * 100, # Typisk prosentfeil (CV)
    .groups = "drop"
  ) %>%
  mutate(
    mean_watt = round(mean_watt, 1),
    typical_error = round(typical_error, 2),
    cv = round(cv, 2)
  ) %>%
  arrange(desc(mean_watt)) # Sorter etter CV i synkende rekkefølge

# Nødvendige utreninger og opprydding for tabelle 2.3

# Les inn data og filtrer for relevante variabler og målepunkter
df_filtered <- read_xlsx("data/plot_data_all_gr.xlsx") %>%
  filter(timepoint %in% c("t1", "t2", "t3", "t4")) %>%
  select(id, timepoint, w.max) %>%
  mutate(w.max = as.numeric(w.max))

# Fyll inn manglende tidspunkter for hver deltaker
df_complete <- df_filtered %>%
  # Fyller inn NA for manglende tidspunkter
  complete(id, timepoint, fill = list(w.max = NA))

# Beregn typisk feil og CV mellom testpar
df_pairwise <- df_complete %>%
  arrange(id, timepoint) %>%
  group_by(id) %>%
  mutate(
    diff_watt = w.max - lag(w.max), # Differanse mellom tester
    mean_watt_pair = (w.max + lag(w.max)) / 2, # Gjennomsnitt for testparet
    test_pair = paste(lag(timepoint), timepoint, sep = " -> ")
  )

```



```

) %>%
filter(!is.na(diff_watt)) %>% # Fjern rader uten gyldige verdier
ungroup()

# Beregn typisk feil og CV på gruppenivå
summary_gruppe <- df_pairwise %>%
  group_by(test_pair) %>%
  summarise(
    typical_error = sd(diff_watt, na.rm = TRUE), # SD av differansene
    # Gjennomsnittlig CV
    mean_cv = mean((abs(diff_watt) / mean_watt_pair) * 100, na.rm = TRUE),
    .groups = "drop"
  )

```

3 Regresjonsmodeller, predikasjon og tolkning

Denne oppgaven er delt inn tre separate deler som tar for seg konsepter innenfor analyse av data og regresjon. I del 1 kalkulerer vi laktat terskler, og ser nærmere på reliabiliteten mellom to ulike terskelnivåer. Del 2 bruker vi molekylær data til å predikere størrelsen på DNA-fragment ved hjelp av en veileder. I del 3 skal vi se nærmere på om det finnes en lineær sammenheng mellom to valgte variabler fra datasettet **hypertrophy** i datapakken **exscidata**. Hver del vil ha sin egen introduksjon, metode, resultatdel og diskusjon.

3.1 Del 1: Laktat terskler

Laktatterskel (LT) er en sentral variabel innen treningsfysiologi, spesielt i utholdenhetsidretter, hvor den brukes til å forutsi prestasjon, styre intensiteten av treningsøkter og evaluere effekten av trening (Machado et al., 2012; Poole et al., 2021). LT representerer den arbeidsintensiteten hvor produksjonen av laktat i blodet overstiger kroppens evne til å fjerne det, noe som fører til en akkumulering av laktat i blodet (Poole et al., 2021, p. s.738). Dette markerer overgangen fra en stabil til en progressivt økende metabolsk belastning.

Det finnes mange ulike metoder for å bestemme LT, og en av de mest brukte er å måle intensitetene ved faste blodlaktatnivåer, som 2 og 4 mmol/L, ved hjelp av regresjonsmodeller som predikerer intensiteten ved disse verdiene (Kindermann et al., 1979; Tanner & Gore, 2012). Andre tilnærminger, som “maximal-deviation method” (Dmax), beskrevet av Machado et al. (2012), tilbyr analyser som kan bedre reflektere de underliggende metabolske mekanismene som er bestemmende for prestasjon (Machado et al., 2012).

For å evaluere påliteligheten av disse målemetodene, er det viktig å vurdere testens reliabilitet. Dette kan kvantifiseres gjennom variasjon innen samme person, også kjent som “typical error”, på norsk typisk feil (Hopkins, 2000, p. s.2). Typisk feil forteller noe om den forventede variasjonen mellom målinger på samme individ og kan uttrykkes som en prosentandel av gjennomsnittsverdien, kjent som koeffisienten av variasjon (CV) (Hopkins, 2000, p. 3). CV er dimensjonsløs og åpner opp for muligheten å sammenligne reliabiliteten mellom personer med ulik prestasjonsevne, og er derfor et relevant mål i denne rapporten.

Laktatterskelen er en viktig fysiologisk parameter for å forstå sammenhengen mellom treningsintensitet og metabolsk respons. Selv om det finnes flere metoder for å beregne laktat-terskel, er det vanlig å bruke faste blodlaktatnivåer som referansepunkter, som 2 mmol/L og 4 mmol/L, for å forutsi treningsintensitet (Wickham, 2014, kap. 6). Formålet med denne

deloppgave er å sammenligne reliabiliteten til disse tersklene, målt som typisk feil i prosent av gjennomsnittet.

På grunn av begrenset datainnsamling fra vårt reliabilitetsforsøk, benyttes datasettet fra `cyclingstudy` som grunnlag for analysen. Dette gir mulighet til å beregne og evaluere reliabiliteten av de to tersklene og utforske deres egnethet i praksis.

3.1.1 Metode

For å ha laktatterskler til å undersøke, ble det benyttet et åpent datasett `cyclingstudy`, som inneholdt en rekke fysiologiske variabler fra en sykkelstudie (Sylta et al., 2016). Fra dette datasettet ble det hentet ut informasjon om treningsintensitet, målt i watt, og blodlaktatkonsentrasjoner. Ved hjelp av lineære og polynomiske modeller ble det beregnet to laktatterskler, 2 mmol/L og 4 mmol/L, for å evaluere forholdet mellom treningsintensitet og laktatakkumulering. Prediksjoner fra hver modell ble brukt til å identifisere wattverdier nærmest tersklene 2 mmol/L og 4 mmol/L. Den tredjegradspolynomiske modellen ble valgt for å illustrere resultater i detalj.

Data ble filtrert for en deltaker (subject = 10) ved et bestemt tidspunkt i studien (pre). Laktatkonsentrasjoner mellom 225 og 375 watt ble omformet til et langt format for enkel visualisering og analyse. Laktatnivåene ble analysert for å finne treningsintensiteter nærmest tersklene 2 mmol/L og 4 mmol/L.

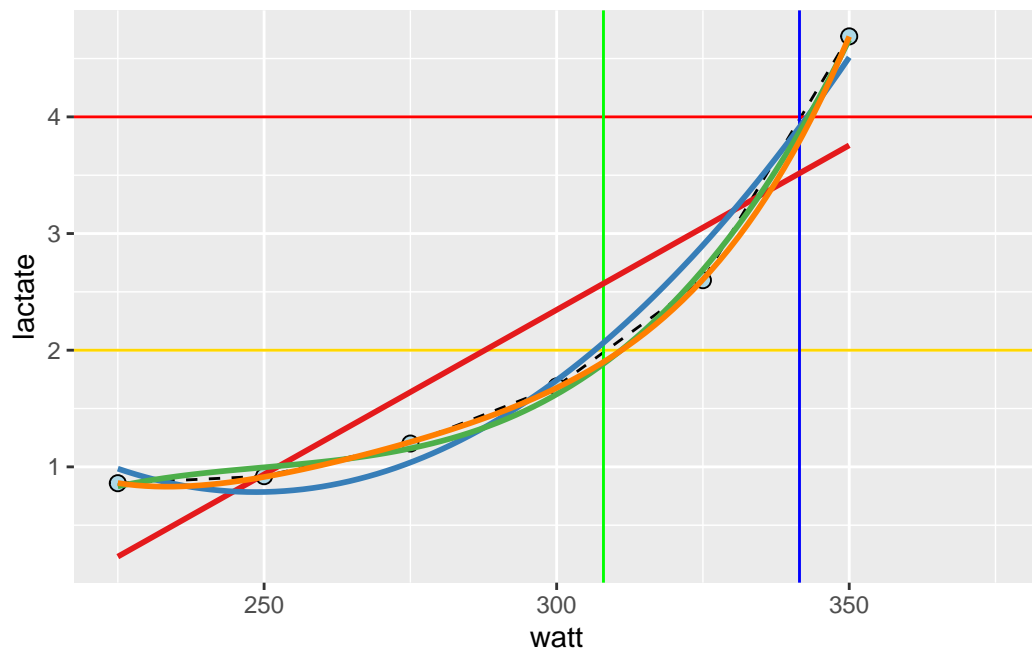
Restverdiene (residuals) fra hver modell ble beregnet for å vurdere hvor godt modellene beskrev dataene. Disse restverdiene ble visualisert i Figure 3.2 for å illustrere modellenes avvik fra de observerte verdiene ved ulike treningsintensiteter. Viderre ble de ulike modellene, inkludert lineær, andre-, tredje- og fjerdegradspolynomiske tilpasninger, brukt for å sammenligne hvordan hver modelle beskriver sammenhengen mellom watt og blodlaktatkonsentrasjon, grafisk fremstilt i Figure 3.1.

Reliabilitet til tersklene ble analysert ved å beregne typisk feil $TE = \frac{SD_{diff}}{\sqrt{2}}$, som prosentandel av gjennomsnittsverdien, altså koeffisienten av variasjon (CV) $CV(\%) = \frac{\text{Typical Error}}{\text{Mean}} \times 100$.

All datahåndtering, analyse og grafisk fremstilling ble utført i R (versjon 4.1.1).

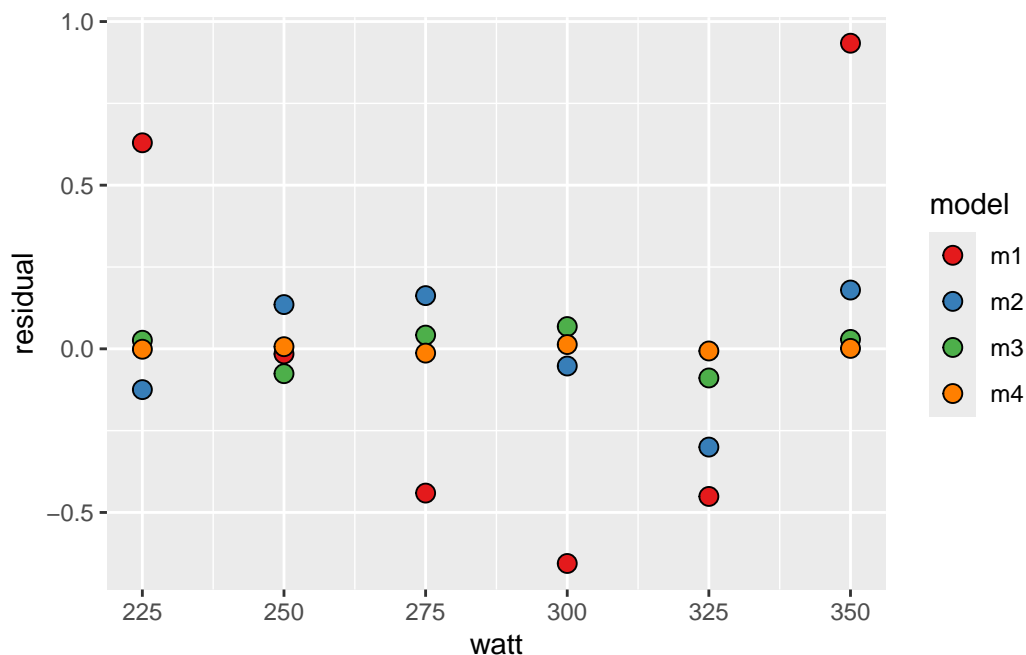
3.1.2 Resultat

Figure 3.1 viser sammenhengen mellom treningsintensitet (watt) og blodlaktatkonsentrasjon, sammen med tilpasninger basert på ulike modeller. De observerte dataene er koblet med rette linjer, her synlig som stiplet linje, som interpolerer mellom punktene. Tersklene 2 mmol/L (gul) og 4 mmol/L (rød) er markert med horisontale linjer, mens de vertikale linjene (blå og grønn) markerer hvor tersklene møter den stiplede linjen på x-aksen, basert på visuell estimering.



Figur 3.1: Tilpasning av ulike modeller til sammenhengen mellom treningsintensitet (watt) og blodlaktatkonsentrasjon. Grafen viser lineær modell (rød linje), andregradspolynomisk modell (blå linje), tredjegradspolynomisk modell (grønn linje), og fjerdegradspolynomisk modell (oransje linje), sammen med de observerte dataene (punkter). Tersklene ved 2 mmol/L (gul linje) og 4 mmol/L (rød linje) er indikert.

De ulike modellene ble sammenlignet med hensyn til hvordan de beskriver dataene, og evaluert ved restverdier (Figure 3.2). Samlet sett gir de tredje- og fjerdegradspolynomiske modellene de beste tilpasningene, særlig nær tersklene ved 2 mmol/L og 4 mmol/L, mens den lineære modellen viser større avvik ved høyere wattverdier. Andregradspolynomiske modellen ligger mellom disse to ytterpunktene.



Figur 3.2: Restverdier for de ulike modellene (lineær, andregradspolynomisk, tredjegradspolynomisk og fjerdegradspolynomisk). Grafen viser hvordan modellene avviker fra de observerte verdiene over treningsintensiteter (watt), med hver modell representert av forskjellige farger.

Table 3.1 og Table 3.2 viser estimerte watt verdier ved laktatttersklene 2mmol/L og 4 mmol/L, basert på de ulike modellene. Gjennomsnittet av intensitetene for hver terskel er også inkludert. Disse predikasjonene ble beregnet ved interpolasjon mellom de observerte datapunktene for blodlaktatkonsentrasjon.

3.1.3 Diskusjon

Estimatene variere noe mellom modellene for begge terskler. For 2 mmol/L varierer de fra 287.7 W (lineær modell) til 311.1 W (fjerdegradspolynomisk), mens snittet ble på 304.1 W. For 4 mmol/L varierte de fra 342.8 W (andregradspolynomisk modell) til 350 W (lineær modell), med et snitt på 344.9 W. Dette illustrerer hvordan valg av modell, spesielt ved høyere intensiteter kan påvirke predikasjonene.

Tabell 3.1: De estimerte treningsintensitetene ved laktatterksel 2mmol/L basert på de fire ulike modellene: lineær (m1), andregradspolynomisk (m2), tredjegradspolynomisk (m3), og fjerdegradspolynomisk (m4).

Laktatterskel ved 2 mmol/L
Treningsintensitet per modell og gjennomsnitt

Modell	Treningsintensitet (Watt)
m1	287.7
m2	306.5
m3	311.0
m4	311.1
Gjennomsnitt	304.1

Tabellen viser modellene og gjennomsnittet for terskelen ved 2 mmol/L.

Tabell 3.2: De estimerte treningsintensitetene ved laktatterksel 2mmol/L basert på de fire ulike modellene: lineær (m1), andregradspolynomisk (m2), tredjegradspolynomisk (m3), og fjerdegradspolynomisk (m4).

Laktatterskel ved 4 mmol/L
Treningsintensitet per modell og gjennomsnitt

Modell	Treningsintensitet (Watt)
m1	350.0
m2	342.8
m3	343.0
m4	343.7
Gjennomsnitt	344.9

Tabellen viser modellene og gjennomsnittet for terskelen ved 4 mmol/L.

Tabell 3.3: Beregnede verdier av gjennomsnittlig treningsintensitet (watt), standardavvik for differanser (watt), typisk feil (watt), og koeffisient av variasjon (CV %), ved 2 mmol/L og 4 mmol/L

Reliabilitet ved ulike terskler
Typisk feil og koeffisient av variasjon

Terskel (mmol/L)	Gjennomsnitt (Watt)	SD Diff (Watt)	Typisk Feil (Watt)	CV (%)
2 mmol/L	304.07	11.13	7.87	2.59
4 mmol/L	344.88	3.44	2.43	0.70

Tabellen viser beregnet reliabilitet ved tersklene 2 mmol/L og 4 mmol/L.

I Table 3.3, presenteres mål på reliabilitet ved tersklene 2 mmol/L og 4 mmol/L. Ved terskel 2 mmol/L har treningsintensiteten en høyere typisk feil (7.87 W) og koeffisient av variasjon (2.59%) sammenlignet med terskel ved 4 mmol/L, hvor typisk feil og CV% er henholdsvis 2.43 W og 0.70%. Man kan derfor antyde at reliabiliteten er bedre ved høyere terskler (4 mmol/L) enn ved lavere terskler (2 mmol/L).

3.2 Del 2: Forutsi størrelser på DNA fragmenter eller stiningene i en qPCR-kalibreringskurve

Sportslige prestasjoner påvirkes av både miljømessige og genetiske faktorer (Tucker & Collins, 2012). Et viktig gen i denne sammenhengen er *ACTN3*, som koder for proteinet alpha-actinin-3. Dette proteinet finnes nesten utelukkende i hurtige muskelfibre og er kjent for sine rolle i kraftbaserte aktiviteter som sprint og vektløfting (Mikami et al., 2014; North & Beggs, 1996; Schadock et al., 2015). Mutasjoner i genet kan føre til en ikke-funksjonell variant, kjent som R577X-polymorfismen, som resulterer i manglende produksjon av proteinet (North et al., 1999). Genotyper som inneholder R allelet, er assosiert med bedre ytelse i kraftfulle og eksplosive idretter, mens X-allelet kan være gunstig for utholdenhet (Mikami et al., 2014; Yang et al., 2003, pp. s.629–630).

For å analysere genotypene til *ACTN3*-genet, vil molekylær teknikker som PCR (polymerasekjedereaksjon) og elektroforese være nyttig. PCR muliggjør spesifikk amplifisering av DNA-sekvenser, slik at man kan identifisere genetiske variasjoner (Schadock et al., 2015). Elektroforese i agarosegel brukes deretter til å separere DNA-fragmentene basert på størrelse, noe som vil gi en visuell representasjon av genotypene (Schadock et al., 2015).

I denne delen av oppgaven ble *ACTN3*-genet undersøkt gjennom DNA-analyse, som ble gjennomført som en del av et forsøk på molekylærelaboratoriet. Ved hjelp av PCR og elektroforese

forsøkte man å separere og analyser fragmentstørrelsene til *ACTN3*-genet for å kartlegge genotypen i de ulike prøvene, for å innsikt i genetiske bidrag til fysisk ytelse og idrettsprestasjoner.

3.2.1 Metode

3.2.1.1 DNA-ekstraksjon

DNA ble ekstrahert fra blodprøver samlet i prøverør med EDTA (etylendiamintetraeddiksyre) ved hjelp av en modifisert protokoll basert på Bartlett og Stirling (Bartlett & Stirling, 2003, kap 6). Etter overføringen av 3 mL blod til et 15 mL rør, ble cellene lysert med Reagens A og sentrifugert (3000 g i 5 min) for å isolere en cellepellet. Pelleten ble resuspendert i Reagens B, og DNA ble frigjort ved tilsetning av natriumperklorat (250 µL av 5 M sodium perchlorate) og inkubasjon ved 65 °C. Etter avkjøling i romtemperatur, ble iskald kloroform (2 mL) tilsatt for å skille DNA fra andre cellekomponenter, og mikset i en roterende misker i mellom 30 til 60 min. Etterpå ble den sentrifugert etterfulgt av sentrifugering (2400 g i 2 min) for å hente ut den øvre delen av prøven. DNA ble uthentet med kald 100% etanol (2-3 mL), tørket og resuspendert i TE-buffer (200 µL). Konsentrasjonen ble målt med et spektrofotometer, og verdiene låg rundt 200 og 500 ng/µL.

3.2.1.2 Bestemmelse av *ACTN3*-genotypen

ACTN3-genotypen ble bestemt ved bruk av en fire-primer PCR-protokoll tilpasset fra Schadock et al. (2015) (Schadock et al., 2015). PCR-reaksjonen ble satt opp i et totalvolum på 20 µL, bestående av 10 µL 2X master mix, 5 µL primermiks (inneholdt hsACTN3_F1, hsACTN3_R1, hsACTN3Tif_F2, og hsACTN3Cir_R2) og 5 µL DNA-prøve (se over). PCR-syklusen inkluderte initial denaturering ved 95 °C i 2 minutter, etterfulgt av 35 sykluser med 95 °C i 10 sekunder, og 72 °C i 45 sekunder, og til slutt ved 72 °C i 2 minutter.

3.2.1.3 Elektroforese for analyse av PCR-produkter

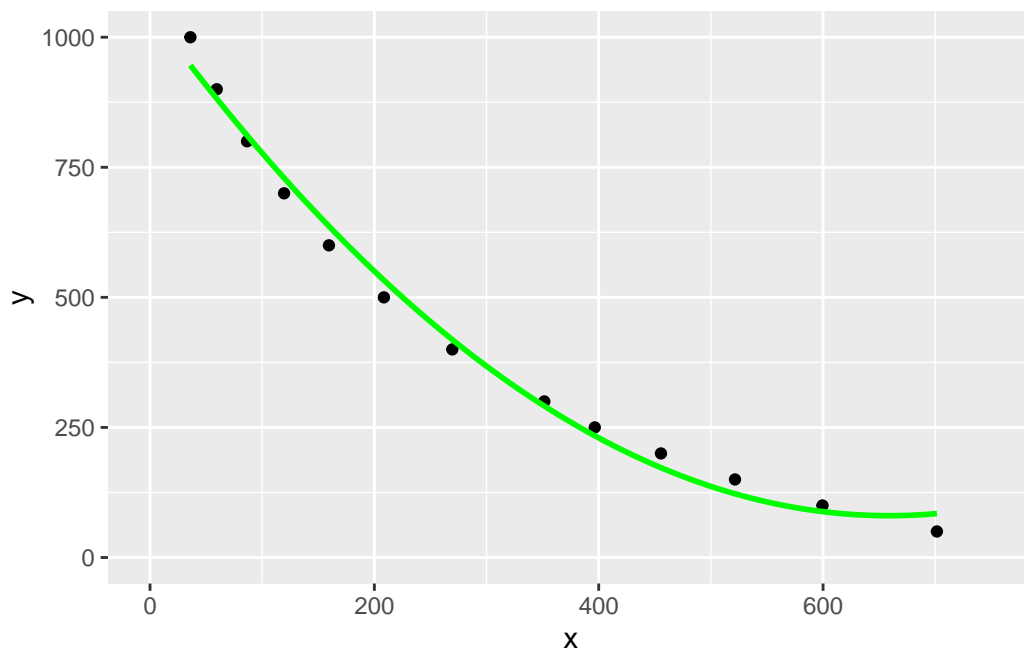
PCR-produktene ble analysert ved hjelp av agarosegelelektroforese i en 2 % agarosegel. Gelen ble fremstilt ved å løse 2 g agarose i 100 mL 1X TBE-buffer, med tilsetning av 10 µL Sybr Safe for visualisering av DNA. Løsningen ble oppvarmet til klarhet, avkjølt til cirka 60 grader, og deretter helt i en støpeform med gelkammer. Etter cirka 1 time hadde gelet blitt fast, og ble plassert i en horisontal elektroforeseenhet fylt med 1X TBE-buffer

DNA-prøvene ble blandet med 6X farge (1 µL per 5 µL DNA-prøve), og 2-5 µL av hver prøve ble lastet i egne brønner sammen med en DNA-stige som referanse. Elektroforesen ble utført ved 150 V i cirka 1 time, inntil fargeindikatoren hadde vandret rundt 80% av gelens lengde. Gelen med prøvene ble visualisert i en G ved bruk av UV-lys og Sybr Green-instillinger.

3.2.1.4 Analyse av PCR-produkter med Image J og R

For å bestemme størrelsen på PCR-produktene ble bildeanalyse utført med ImageJ Fiji. Gel-bildet ble invertert, rotert og trimmet for å isolere prøvene og DNA-stigen. Ved bruk av rektangelverktøyet ble stigen og prøvene markert, og toppunktene i intensitetsgrafene ble registrert. Disse punktene representerte DNA-fragmentenes migrasjonsavstand, og dataene ble eksportert til Excel for videre analyse.

I R ble en kalibreringskurve laget basert på DNA-stigen (Figure 3.3), hvor logaritmen av molekylvekten ble plottet mot migrasjonsavstanden. Kalibrering ble utført ved hjelp av en polynommodell for å sikre høy presisjon. Denne modellen ble deretter brukt til å estimere molekylstørrelsen for de ukjente prøvene. Modellen ble vurdert basert på R^2 -verdien fra lineær regresjon, og verdien lå nær 1, noe som indikerte høy modellpresisjon. Den justerte modellen ble deretter brukt til å estimere molekylstørrelsene for de ukjente prøvene, som ble beregnet på migrasjonsavstandene fra gelelektroforesen.



Figur 3.3: Kalibreringskurve basert på DNA-stigen.

3.2.2 Resultater

Analyse av PCR-produktene viste at DNA-fragmentet i brønn 1 (prøve 1) hadde en estimert båndstørrelse på 407 bp, mens brønn 2 (prøve 2) viste et fragment på 401 bp. I brønn 3 (prøve 3) ble det identifisert to fragmenter, med størrelser på henholdsvis 396 bp og 296 (Table 3.4).

Tabell 3.4: Resultater fra PCR-analyse: Estimert båndstørrelse for DNA-fragmenter basert på agarosegelelektroforese.

Resultater fra PCR-analyse

Brønn	Båndstørrelse (bp)
1	407
2	401
3	396, 296

Disse resultatene ble beregnet basert på kalibreringskurven som ble laget ved hjelp av DNA-stigen, og fragmentstørrelsene reflekterer migrasjonsmønsteret observert i gelanalysen.

3.2.3 Diskusjon

Denne analysen viste at ingen av DNA-fragmentene hadde nøyaktig den forventede størrelsen på 413 bp (R/R) eller 318 bp (X/X) (Schadock et al., 2015). Fragmentene fra brønn 1 (417bp) og brønn 2 (401 bp) ligger imidlertid nær den forventede størrelse for R/R-genotypen, mens brønn 3 (396 bp og 296 bp) indikerer en mulig heterozygot genotype (R/X) (Schadock et al., 2015). Avvikene kan forklares med flere faktorer, inkludert tekniske og menneskelige feil under eksperimentet.

Blant de tekniske feilene er kvaliteten på gelbildet noe som kan ha bidratt til usikkerhet i målingene, ettersom dårlig oppløsning eller utilstrekkelig kontrast gjør det vanskelig å nøyaktig identifisere båndenes plassering. Videre kan kalibreringsmodellen ha blitt påvirket av små feil i dataregistreringen, noe som kan ha påvirket nøyaktigheten i estimeringen av fragmentstørrelsene.

Menneskelige feil er også en viktig faktor å vurdere. Feil pipetering kan ha ført til variasjon i mengden DNA eller reagenser, noe som kan påvirke amplifiseringen. Under elektroforese kan små variasjoner i prøvelasting, som ulike mengde DNA-prøve i brønnene, ha forårsaket skjevheter i båndenes intensitet og plassering. I tillegg kan subjektiv tolkning av gelbilder uten digitale analyseverktøy føre til feiltolkninger.

Denne oppgaven viste at PCR-analysen ga fragmenter som lå nær forventede størrelser, men med mindre avvik, noe som kan tilskrives både tekniske og menneskelige feil. Dette understreker viktigheten av presisjon i laboratoriearbeid og bruk av objektive metoder for dataanalyse.

3.3 Del 3: Tolkning av regresjonsmodell

Muskeltilpasninger til trening avhenger av en kompleks kombinasjon av genetiske og miljømessige faktorer. Myonuclei (cellekjerner), som finnes i muskelfibrene, spiller en nøkkelrolle i å regulere musklens kapasitet for proteinsyntese og dermed også evnen til å utvikle styrke og kraft (McArdle et al., 2014, kap 22). Antall myonuclei i type-II muskelfibre er spesielt relevant, da disse fibrene er avgjørende for eksplosive bevegelser og maksimal styrke (McArdle et al., 2014, kap 22). Samtidig er det uklart i hvilken grad treningserfaring, mål som antall år med systematisk trening, kan påvirke antallet myonuclei (McArdle et al., 2014, p. s.535).

Ved å undersøke sammenhengen mellom antall myonuclei i type II muskelfibre og treningsalder, kan man belyse om langvarig trening har en målbar effekt på muskelfiberens egenskaper. Dette kan gi innsikt som er relevant både for praktisk trening og for å forstå mekanismene bak muskeltilpasning.

Spørsmålet: Finnes det en lineær sammenheng mellom antall myonuclei per type-II muskelfiber og treningsalder?

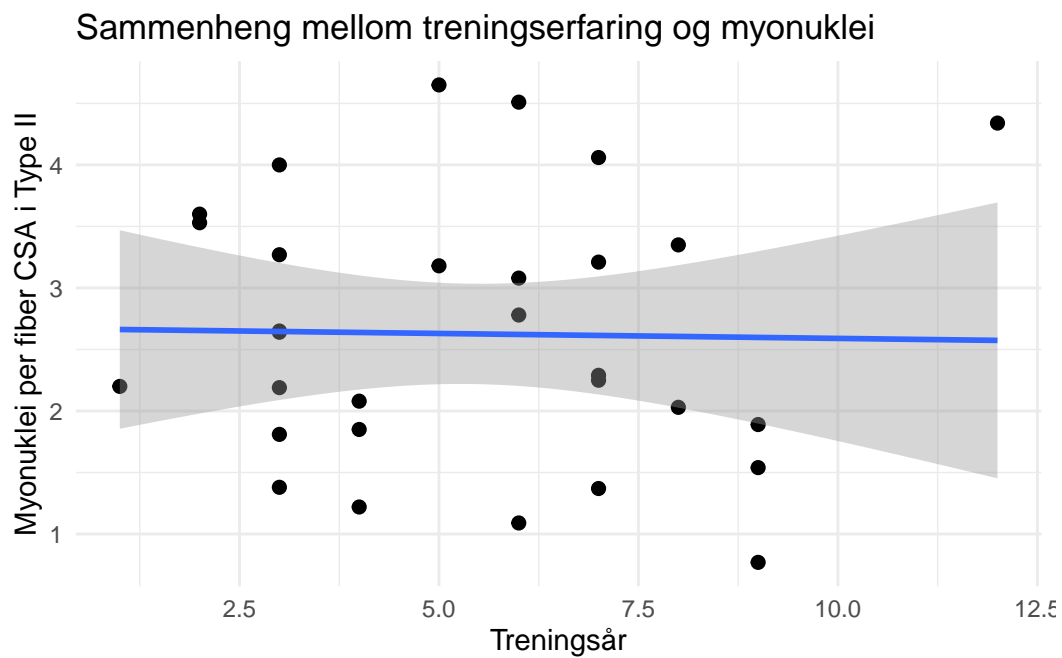
3.3.1 Metode

For å undersøke sammenhengen mellom antall myonuclei per type-II muskelfiber og treningsalder, brukte såg man på variablene `FAST_NUCLEI_T1` og `TRAINING_AGE` i datasettet `hypertrofi` fra `exscidata`-pakken. Lineær regresjon ble benyttet, da denne er velegnet for å evaluere en potensiell lineær relasjon mellom en avhengig variabel (`FAST_NUCLEI_T1`) og en uavhengig variabel (`TRAINING_AGE`) (Spiegelhalter, 2019a, pp. s.128–129).

Datasettet ble først filtrert for å ekskludere observasjoner med manglende verdier, og vi valgte kun de relevante variablene for analysen. Dataene ble deretter visualisert med scatterplott og en tilhørende regresjonslinje generert av `geom_smooth`, se Figure 3.4. Regresjonslinjen gir en indikasjon på hvordan variablene relaterer seg til hverandre, mens det gråe konfidensintervallet rundt linjen reflekterer usikkerheten i modellen. Et bredt konfidensintervall, som sett her, indikerer større usikkerhet i forholdet mellom variablene (Spiegelhalter, 2019a, pp. s.240–244).

3.3.2 Resultat

Tabellen i Table 3.5 oppsummerer de statistiske parametrene fra den lineære modellen, inkludert den estimerte koeffisienten, standardfeilen, t -verdien og p -verdien. Disse komponentene gir innsikt i styrken og usikkerheten rundt sammenhengen mellom variablene. Intercept ble ekskludert fra tabellen, da det representerer verdien av den avhengige variabelen når den uavhengige variabelen er null, noe som ikke er relevant for denne analysen.



Figur 3.4: Sammenheng mellom treningalder og myonuclei per fiber CSA i Type-II

Tabell 3.5: Sammenheng mellom treningserfaring og myonuklei per muskelfiber type-II

Term	Estimert koeffisient	Standardfeil	<i>t</i> -verdi	<i>p</i> -verdi
Treningserfaring (år)	-0.008	0.077	-0.104	0.918

Notat: *p*-verdier mindre enn 0.05 anses som statistisk signifikante.

3.3.3 Diskusjon

I tabellen kan vi lese av verdiene for estimert koeffisient (regresjonskoeffisient), standardfeil, t-verdi og p-verdi. Den estimerte koeffisienten til “Treningserfaring (år)” indikerer at FAST_NUCLEI_T1 reduseres med 0.008 per år med treningserfaring. Denne negative endringen er imidlertid svak, og standardfeilen er relativt stor i forhold til koeffisienten, noe som tyder på at estimatet er usikkert (Spiegelhalter, 2019a, pp. s.230–232)

Standardfeilen gir en indikasjon på hvor mye koeffisienten kan forventes å mellom ulike utvalg (Spiegelhalter, 2019a, pp. s.230–232). Selv om standardfeilen numerisk er lav, er dens forhold til koeffisienten avgjørende for tolkningen. I dette tilfellet bør man være forsiktig med å trekke konklusjoner om nøyaktigheten til estimatet.

t-verdien sier hvor mange standardavvik den estimerte koeffisienten er fra null. En større t-verdi, enten negativ eller positiv, tyder på at man med større sikkerhet kan si at koeffisienten er signifikant (Spiegelhalter, 2019a, pp. s.275–276). Med en t-verdi på -0.104, i vårt tilfelle, kan man ikke med sikkerhet si at det er en signifikant lineær sammenheng mellom FAST_NUCLEI_T1 og TRAINING_AGE.

p-verdien er nært knyttet til t-verdien, og hjelper oss å vurdere om t-verdien er statistisk signifikant. P-verdien representerer sannsynligheten for å observere en teststatistikk som er like ekstrem, eller like ekstrem, som den t-verdien vi har fått, gitt antagelsen at det ikke er en sammenheng mellom variablene våre (Spiegelhalter, 2019a, pp. s.264–265). I vår modell er p-verdien 0.918, noe som tilsier at det er 91,8 % sannsynlighet at man vil observere en t-verdi på -0.008, gitt nullhypotesen. Dermed kan man ikke konkludere med at den uavhengige variabelen TRAINING_AGE har en effekt på den avhengige variabelen FAST_NUCLEI_T1 eller at det finnes en signifikant lineær sammenheng mellom disse variablene (Spiegelhalter, 2019a, pp. s.265–268).

Det er viktig å understreke at p-verdien kun sier noe om den statistiske signifikansen av resultatene, og ikke om størrelsen på effekten eller dens praktiske relevans (Spiegelhalter, 2019a, p. s.285). I små datasett, som i vårt tilfelle, kan p-verdien være høy selv om det er en reel effekt, fordi små utvalg har lavere statistisk styrke til å påvise sammenhenger (Spiegelhalter, 2019a, p. s.285).

Denne analysen har undersøkt sammenhengen mellom treningserfaring og antall myonuclei per type II-muskelfiber, men resultatene viser ingen signifikant lineær sammenheng mellom variablene. Selv om p-verdien indikerer manglende statistisk signifikans, er det viktig å merke seg begrensningene ved små datasett og mulig usikkerhet knyttet til estimatene.

3.4 Vedlegg: Koder brukt

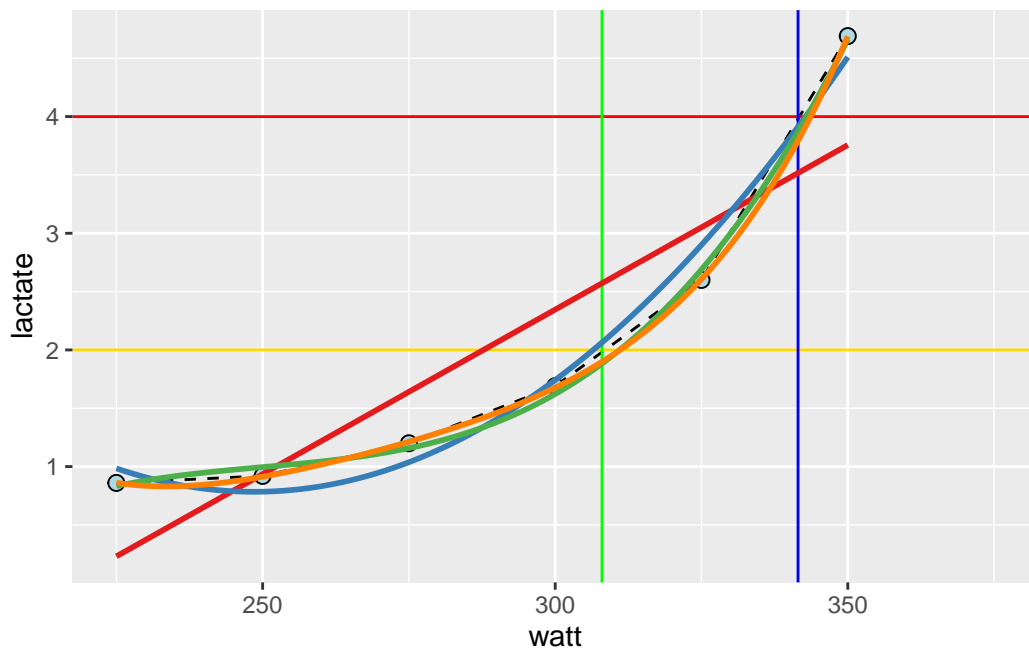
```
# Laster inn nødvendige biblioteker
library(exscidata)
library(readxl)
library(tidyverse)
library(ggplot2)
library(gt)
library(broom)
library(dplyr)

# Laste inn data
data("cyclingstudy")

## Estimering av laktatterskelen og treningsintensiteten ved 4mmol L-1

cyclingstudy %>%
  # utvalg av nødvendige kolonner i analysen.
  select(subject, group, timepoint, lac.225:lac.375) %>%
  # Kun ein deltaker og ett tidspunkt.
  filter(timepoint == "pre", subject == 10) %>%
  # lang format ved å bruke laktatkolonnene.
  pivot_longer(names_to = "watt",
               values_to = "lactate",
               names_prefix = "lac.",
               names_transform = list(watt = as.numeric),
               cols = lac.225:lac.375) %>%
  # Plotte data, group = subject nødvendig for å sammenkoble punktene.
  ggplot(aes(watt, lactate, group = subject)) +
  geom_line(lty = 2) +
  geom_point(shape = 21, fill = "lightblue", size = 2.5) +
  # Linjer på spesifikke punktene for 2mmol og 4mmol, samt skjæringspunktet
  # mellom linjene.
  geom_hline(yintercept = 4, color = "red") +
  geom_hline(yintercept = 2, color = "gold") +
  geom_vline(xintercept = 341.5, color = "blue") +
  geom_vline(xintercept = 308, color = "green") +
  # legge til en strak linje fra den lineære modellen.
  geom_smooth(method = "lm", se = FALSE, formula = y ~ x, color = "#e41a1c") +
```

```
# poly(x, 2) Legger til en andregradspynomisk modell.
geom_smooth(method = "lm", se = FALSE, formula = y ~ poly(x, 2),
            color = "#377eb8") +
# poly(x, 3) Legger til en tredjegradsplnomisk modell.
geom_smooth(method = "lm", se = FALSE, formula = y ~ poly(x, 3),
            color = "#4daf4a") +
# poly(x, 4) Legger til en fjerdegradspynomisk modell.
geom_smooth(method = "lm", se = FALSE, formula = y ~ poly(x, 4),
            color = "#ff7f00")
```



```
lactate <- cyclingstudy %>%
  # utvalg av nødvendige kolonner i analysen.
  select(subject, group, timepoint, lac.225:lac.375) %>%
  # Kun ein deltaker og ett tidspunkt.
  filter(timepoint == "pre", subject == 10) %>%
  # lang format ved å bruke laktatkolonnene.
  pivot_longer(names_to = "watt",
               values_to = "lactate",
               names_prefix = "lac.",
               names_transform = list(watt = as.numeric),
               cols = lac.225:lac.375) %>%
  # Fjerne dei ugyldige veriene NA for å hindre feilmeldinger.
```

```

filter(!is.na(lactate))

# Opprettelse av modellene

## Lagrer den lineære modellen
m1 <- lm(lactate ~ watt, data = lactate)

## Lagrer den polynomoniske modellen
m2 <- lm(lactate ~ poly(watt, 2, raw = TRUE), data = lactate)

## Lagrer den tredjegradsplynomisk modellen
m3 <- lm(lactate ~ poly(watt, 3, raw = TRUE), data = lactate)

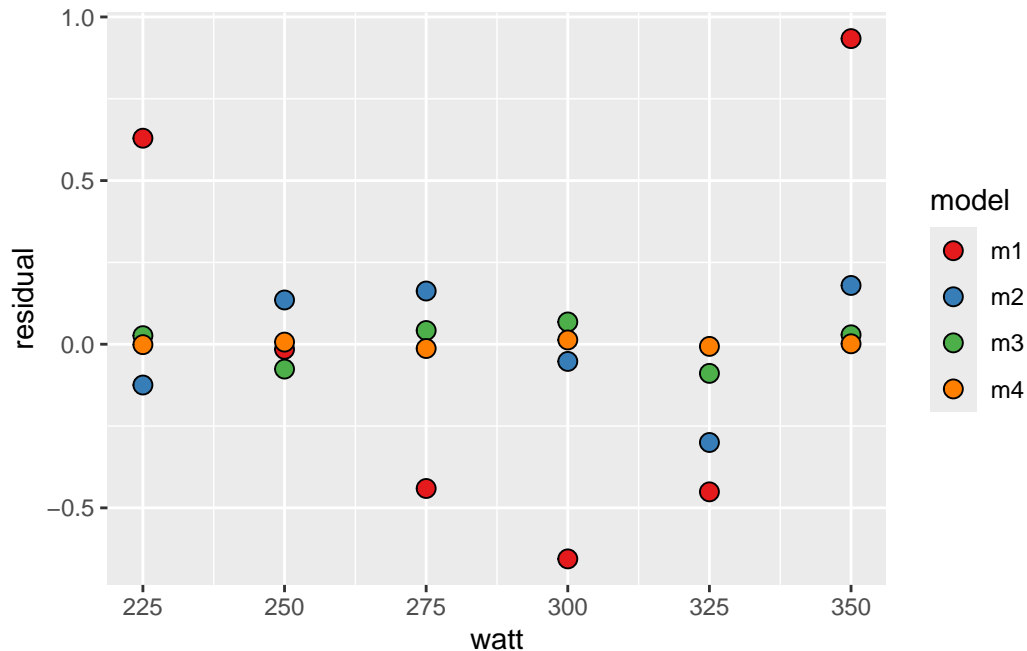
## Lagrer den fjerdegradsplynomisk modell.
m4 <- lm(lactate ~ poly(watt, 4, raw = TRUE), data = lactate)

# Lagre alle restverdiene som nye variabler.
lactate$resid.m1 <- resid(m1)
lactate$resid.m2 <- resid(m2)
lactate$resid.m3 <- resid(m3)
lactate$resid.m4 <- resid(m4)

lactate %>%
  # Samle all data fra modellele.
  pivot_longer(names_to = "model",
               values_to = "residual",
               names_prefix = "resid.",
               names_transform = list(residual = as.numeric),
               cols = resid.m1:resid.m4) %>%
  # Plotte verdiene fra den observerte watten på x aksen og restverdiene
  # på y aksen
  ggplot(aes(watt, residual, fill = model)) + geom_point(shape = 21, size = 3) +

  # For å ha samme farger som over, bruker me scale fill manual.
  scale_fill_manual(values = c("#e41a1c", "#377eb8", "#4daf4a", "#ff7f00"))

```

```
# Predikasjoner

## Ny dataramme for predikasjoner
ndf <- data.frame(watt = seq(from = 225, to = 350, by = 0.1))

## Predikasjoner fra modellene
ndf$predictions_m1 <- predict(m1, newdata = ndf) # Lineær modell
ndf$predictions_m2 <- predict(m2, newdata = ndf) # Andregradspolynomisk modell
ndf$predictions_m3 <- predict(m3, newdata = ndf) # Tredjegradspolynomisk modell
ndf$predictions_m4 <- predict(m4, newdata = ndf) # Fjerdegradspolynomisk modell

# Beregn terskel 2 mmol/L
lactate_threshold_2mmol <- data.frame(
  Model = c("m1", "m2", "m3", "m4"),
  Watt = c(
    ndf %>% filter(abs(predictions_m1 - 2) == min(abs(predictions_m1 - 2))) %>%
      pull(watt),
    ndf %>% filter(abs(predictions_m2 - 2) == min(abs(predictions_m2 - 2))) %>%
      pull(watt),
    ndf %>% filter(abs(predictions_m3 - 2) == min(abs(predictions_m3 - 2))) %>%
      pull(watt),
    ndf %>% filter(abs(predictions_m4 - 2) == min(abs(predictions_m4 - 2))) %>%
      pull(watt)
  )
)
```

```

)
)

# Beregn terskel 4 mmol/L
lactate_threshold_4mmol <- data.frame(
  Model = c("m1", "m2", "m3", "m4"),
  Watt = c(
    ndf %>% filter(abs(predictions_m1 - 4) == min(abs(predictions_m1 - 4))) %>%
      pull(watt),
    ndf %>% filter(abs(predictions_m2 - 4) == min(abs(predictions_m2 - 4))) %>%
      pull(watt),
    ndf %>% filter(abs(predictions_m3 - 4) == min(abs(predictions_m3 - 4))) %>%
      pull(watt),
    ndf %>% filter(abs(predictions_m4 - 4) == min(abs(predictions_m4 - 4))) %>%
      pull(watt)
  )
)

# Kombiner tersklene i en tabell
lactate_thresholds <- data.frame(
  Terskel = c(rep("2 mmol/L", 4), rep("4 mmol/L", 4)),
  Modell = c(lactate_threshold_2mmol$Model, lactate_threshold_4mmol$Model),
  Watt = c(lactate_threshold_2mmol$Watt, lactate_threshold_4mmol$Watt)
)

# Beregn gjennomsnitt direkte for 2 mmol/L og 4 mmol/L
mean_2mmol <- mean(lactate_threshold_2mmol$Watt)
mean_4mmol <- mean(lactate_threshold_4mmol$Watt)

# Beregn standardavviket (SD_diff) for hver terskel
sd_diff_2mmol <- sd(lactate_threshold_2mmol$Watt)
sd_diff_4mmol <- sd(lactate_threshold_4mmol$Watt)

# Beregn typisk feil (TE)
te_2mmol <- sd_diff_2mmol / sqrt(2)
te_4mmol <- sd_diff_4mmol / sqrt(2)

# Beregn koeffisient av variasjon (CV %)
cv_2mmol <- (te_2mmol / mean_2mmol) * 100
cv_4mmol <- (te_4mmol / mean_4mmol) * 100

```

```

# Opprett separate datarammer for tersklene
threshold_2mmol <- lactate_thresholds %>%
  filter(Terskel == "2 mmol/L") %>%
  bind_rows(data.frame(
    Terskel = "Gjennomsnitt",
    Modell = "Gjennomsnitt",
    Watt = mean(lactate_threshold_2mmol$Watt)
  ))

threshold_4mmol <- lactate_thresholds %>%
  filter(Terskel == "4 mmol/L") %>%
  bind_rows(data.frame(
    Terskel = "Gjennomsnitt",
    Modell = "Gjennomsnitt",
    Watt = mean(lactate_threshold_4mmol$Watt)
  ))

# Kombiner alt i en dataramme
reliability_data <- data.frame(
  Terskel = c("2 mmol/L", "4 mmol/L"),
  Mean_Watt = c(mean_2mmol, mean_4mmol),
  SD_diff = c(sd_diff_2mmol, sd_diff_4mmol),
  Typical_Error = c(te_2mmol, te_4mmol),
  CV_Percent = c(cv_2mmol, cv_4mmol)
)

# Lager dataramme for å finne avstand og molekylærvekt
ladder <- data.frame(dist = c(36, 59.5, 86.5,
                             119.5, 159.5, 208.5,
                             269.5, 351.5, 396.5,
                             455.5, 521.5, 599.5, 701.5),
  mw = c(1000, 900, 800,
         700, 600, 500,
         400, 300, 250,
         200, 150, 100, 50))

# Lager ny dataramme med ukjente variabler
unknown <- data.frame(dist = c(258.5, 262.5, 265.5, 335.5))

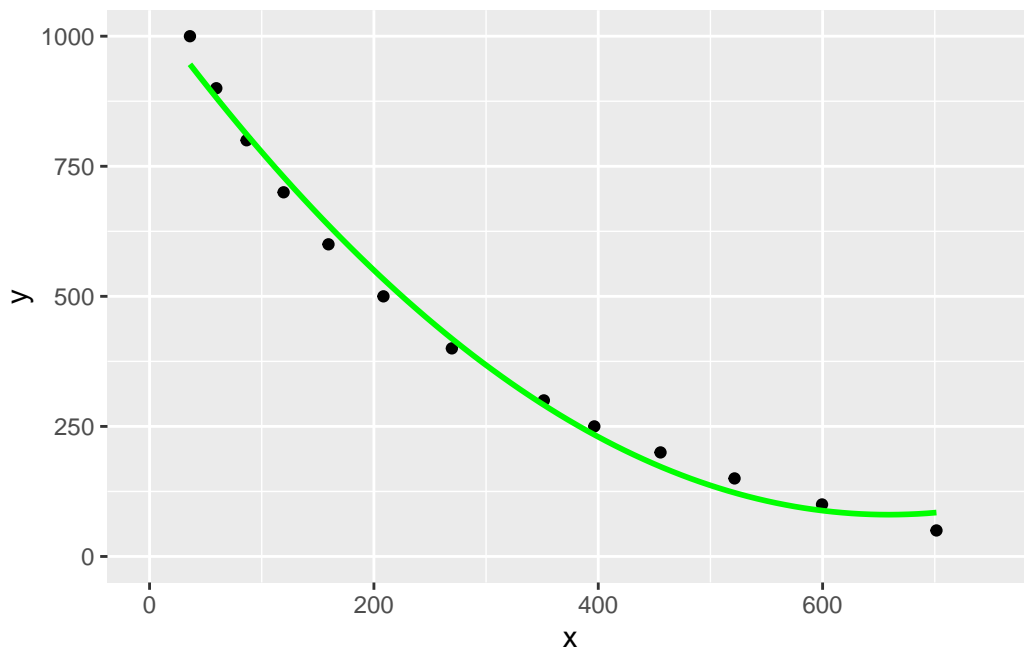
```

```
# Lager en kalibreringsmodell ved hjelp av stigen
ggplot(data = data.frame(x = c(36, 59.5, 86.5,
                              119.5, 159.5, 208.5,
                              269.5, 351.5, 396.5,
                              455.5, 521.5, 599.5, 701.5),
                          y = c(1000, 900, 800,
                              700, 600, 500,
                              400, 300, 250,
                              200, 150, 100, 50)),
        aes(x, y)) + geom_point() +

geom_smooth(method = "lm", formula = y ~ poly(x, 2),
            color = "green", se = FALSE)+

scale_y_continuous(limits = c(0, 1000)) +

scale_x_continuous(limits = c(0, 750))
```



```
# Fit the model
cal <- lm(log(mw) ~ dist, data = ladder)

# Check model performance, R^2 should be ~ 1.
```

```

summary(cal)

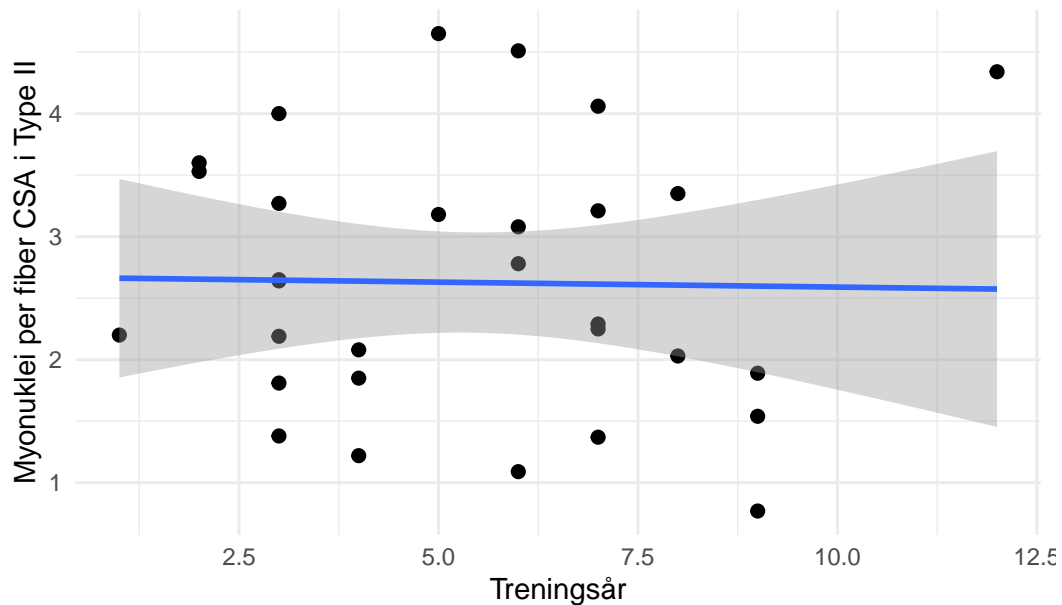
# Estimate molecular weights from migration distances
preds <- exp(predict(cal, newdata = unknown))

# Filtrerer ut NA-verdier før du velger variabler
ds <- hypertrophy %>%
  filter(!is.na(TRAINING_AGE) & !is.na(FAST_NUCLEI_T1)) %>%
  select(PARTICIPANT, GROUP, TRAINING_AGE, FAST_NUCLEI_T1)

# Plotter data uten NA-verdier
ds %>%
  ggplot(aes(TRAINING_AGE, FAST_NUCLEI_T1)) +
  geom_point(size = 2, fill = "red") +
  geom_smooth(method = "lm", se = TRUE) +
  labs(
    title = "Sammenheng mellom treningserfaring og myonuklei",
    x = "Treningsår",
    y = "Myonuklei per fiber CSA i Type II") +
  theme_minimal()

```

Sammenheng mellom treningserfaring og myonuklei



```
# Lager lineær modell med ds uten NA-verdier
mod1 <- lm(FAST_NUCLEI_T1 ~ TRAINING_AGE, data = ds)

# Henter ut koeffisienter og deres statistikker
model_summary <- tidy(mod1)

# Tilpasser p-verdier og runder av, og fjerner interceptet
model_summary <- model_summary %>%
  mutate(
    term = ifelse(term == "(Intercept)", "Intercept (Konstantledd)",
                  "Treningserfaring (år)"),
    p.value = ifelse(p.value < 0.001, "< 0.001", round(p.value, 3)),
    estimate = round(estimate, 3),
    std.error = round(std.error, 3),
    statistic = round(statistic, 3)
  ) %>%
  # Filtrer ut interceptet
  filter(term != "Intercept (Konstantledd)")
# Velger å filtrere ut intercept da det ikkje er aktuelt når vi kun skal se om
# det er en lineær sammenheng mellom dei to variablene
```

4 Statistisk inferens, statistiske modeller og statistisk styrke

Denne rapporten er basert på simuleringer utført i R, der resultatene blir analysert gjennom svar på de spesifiserte spørsmåla (1–8). Gjennom grundig analyse og tolking av resultata gir rapporten ei forklaring som belyser de viktigste funna fra simuleringene.

4.1 Simulering til oppgave 1-4

For å kunne svare på spørsmålene 1 til 4 så er det blitt gjort mange simuleringer med to ulike utvalgsstørrelser ($n = 8$ og $n = 40$), som danner grunnlaget for to lineære modeller, `m1` og `m2`. De viktigste parametrene for modellene er vist i Table 6.2.

Under i Figure 4.1 vises t-fordelingen av de ulike lineære modellene, der det skyggelagte området visualiserer sannsynligheten for å observere en t-verdi som er like ekstrem eller mer ekstrem enn den observerte t-verdien i `samp1`, gitt at nullhypotesen om ingen endring er sann.

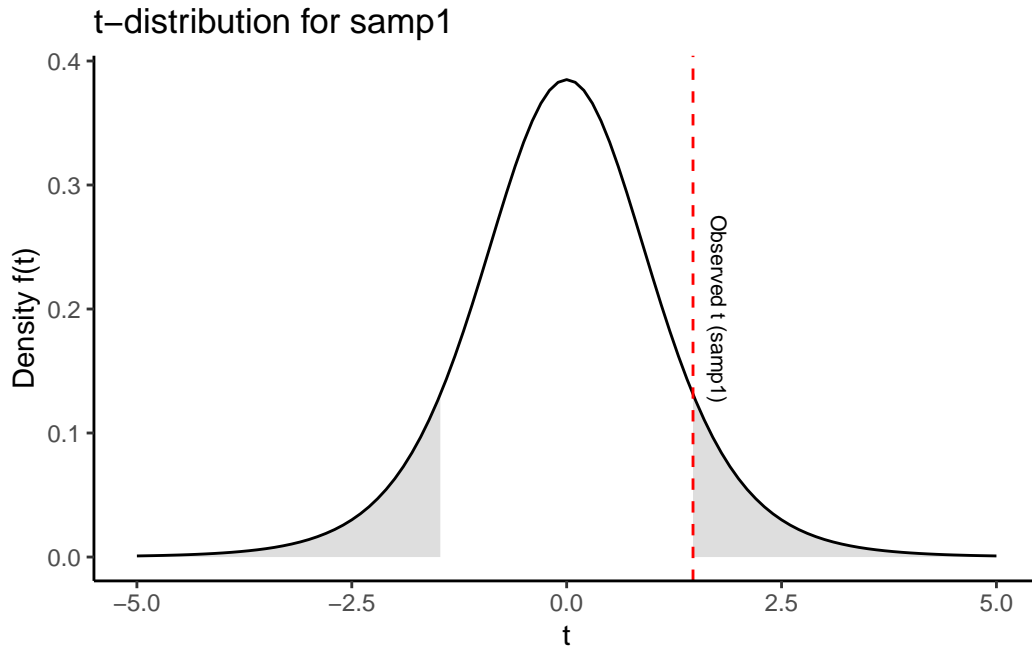
4.1.1 Spørsmål 1 - Forklaring av regresjonsresultater: Estimat, SE, t-verdi og p-verdi

1. Explain the estimate, SE, t-value, and p-value from the regression models that we created previously (`m1` and `m2`).

Tabell 4.1: Oversikt over de parametrene som vil bli diskutert videre i oppgaven.

Parametere for modellene `m1` og `m2`

Parameter	<code>m1</code>	<code>m2</code>
Estimat	1.8397275	1.564160975
Standardfeil	1.2512930	0.477411701
t-verdi	1.4702611	3.276335647
p-verdi	0.1849546	0.002212965



Figur 4.1: t-fordeling av `samp1`.

Estimate: Den estimerte koeffisienten representerer i regresjonsanalyser med to variabler hvor mykje vi kan forvente at den avhengige variabelen endrer seg per endring i den uavhengige variabelen (Spiegelhalter (2019a), s.275). I våre “intercept-only” modeller, vil *estimate* derimot korrespondere til gjennomsnittet av alle verdiene i de ulike utvalgene, og gi oss et estimat på gjennomsnittet til populasjonen.

I `samp1` er utvalgets gjennomsnitt 1.84 noe som er litt ifra populasjonens faktiske gjennomsnitt på 1.5. Når vi øker størrelsen på utvalget, slik som i `samp2`, ser vi at gjennomsnittet nærmer seg populasjonen: `samp2` = 1.564 vs $\mu = 1.5$.

Standardfeil (SE): Standardfeilen måler hvor mykje *estimate* forventes å variere fra utvalg til utvalg (Spiegelhalter, 2019a, pp. s.231, 403–404). Standardfeilen sier dermed hvor mye gjennomsnittet vil kunne variere fra utvalg til utvalg grunnet tilfeldig variasjon i dataene. For eksempel sier standardfeilen i `m2` at gjennomsnittet vil variere med 0.4774117 for hvert utvalg. SE beregnes slik:

$$SE = \frac{s}{\sqrt{n}}$$

hvor:

- (s) er standardavviket til utvalget

- (n) er antall observasjoner

t-verdi: t -verdien, eller t -statistikk, beregnes som $t = \frac{\text{estimate}}{\text{standard error}}$, og kan forstås som hvor mange standardfeil estimatet er fra 0 (Spiegelhalter, 2019a, pp. s.276–277). Den vil hjelpe oss med å avgjøre om det er fornuftig å anta at vårt gjennomsnitt er forskjellig fra 0, og om null-hypotesen burde forkastes eller godtas (Spiegelhalter, 2019a, pp. s.276–277). Dess høyere t -verdien er, dess sikrere kan vi være på at estimatet vårt, her y , oppsto ved tilfeldig variasjon gitt at nullhypotesen var sann.

p-verdi: Gitt at vi veit t -verdien og størrelsen på utvalget, kan vi benytte oss av R til å regne ut p -verdien, (se `summary()` eller `p1/p2` i koden over): henholdsvis 0.185 for `samp1` og 0.002 for `samp2`.

P -verdi defineres som sannsynligheten for å få et resultat så ekstremt som det observert, viss null hypotesen er virkelig sann (Spiegelhalter, 2019a, s. 264). Hva terskel man setter for p -verdien vil variere basert på type studie og ønsket effekt, men historisk sett så er den $p < 0.05$ sett på som statistisk signifikant. I vårt tilfelle er det kun `m2` som gir oss en p -verdi under 0.05, og er da statistisk signifikant. Samtidig er det viktig å understreke at p -verdien ikke sier noe om sannheten til en hypotese eller at statistisk signifikans betyr at det har praktisk eller klinisk betydning (Spiegelhalter, 2019a, pp. s.297–303).

4.1.2 Spørsmål 2 - Årsaker til forskjeller mellom studieresultatene i modellene `m1` og `m2`.

2. Discuss what contributes to the different results in the two studies (`m1` and `m2`).

Det første som burde påpekes er ulik utvalgsstørrelse på de to modellene. Ved større utvalg blir estimatene gjerne mer presise fordi tilfeldige variasjoner gattes ut. For eksempel så vil standardfeilen bli påvirket av størrelsen på utvalget, ved at et større utvalg vil føre til en lavere standardfeil (ref formelen over), som igjen kan gi en mer signifikant t -verdi og lavere p -verdi. Dette er tilfelle i modellene våre over, der `m2` har et større utvalg som fører til en lavere standardfeil, mer signifikant t -verdi og lavere p -verdi sammenlignet med `m1` (Spiegelhalter, 2019a, pp. s.191–192).

Den lave utvalgsstørrelsen i `m1` vil også føre til at tilfeldige variasjoner i datane vil kunne føre til at resultatene avviker mer fra den sanne populasjons gjennomsnittet enn `m2`. Forskjellen i standardfeilen mellom de to modellene, sier at det er forskjellig spredning i observasjonene. Den økte spredningen i `m1` vil bidra til mindre presise estimer, høyere p -verdi og mindre signifikante resultater (Spiegelhalter, 2019a, p. s.195).

4.1.3 Spørsmål 3 - Betydningen av skyggeområdene i halområdene til t-fordelingen i Figure 4.1.

3. Why do we use the shaded area in the lower and upper tail of the t-distribution (See Figure 4.1).

De skyggelagte områdene hjelper å visualisere hvor sannsynlig eller usannsynlig den observerte t-verdien, gitt at nullhypotesen er sann (dvs. at $\mu = 1.5$). Ved å skyggelegge begge halene visualiseres man sannsynligheten for å få en t-verdi enten på den positive eller negative siden. Dess mindre det skyggelagte området er, desto mindre er det at vi vil observere en så ekstrem t-verdi, og desto mer sannsynlig er det at vi kan forkaste nullhypotesen.

4.2 Simulering til spørsmål 4-7

Med hjelp av R har det blitt laget et datasett med resultat fra 1000 utførte studier, for å kunne utvikle en faktisk utvalgsdistribusjon. Ved å bruke `results` datasettet, skal eg svare på spørsmålene 4-7.

4.2.1 Spørsmål 4 - Sammenligning av standardavvik og SE for utvalgsstørrelser ($n = 8$ og $n = 40$): Forklaring og definisjon.

4. Calculate the standard deviation of the estimate variable, and the average of the se variable for each of the study sample sizes (8 and 40). Explain why these numbers are very similar. How can you define the Standard Error (SE) in light of these calculations?

I Table 4.2 er det en oppsummering av standardavviket til `estimate` og gjennomsnittet av `se`:

Tabell 4.2: Oppsummerende statistikk for standardavviket til `estimate` og gjennomsnittet av `se`.

Utvalgsstørrelse	Standardavvik	SE
8	1.0708432	1.0213745
40	0.4838475	0.4696954

Standardavviket av estimatene vil fortelle hvor mye estimatene vil variere fra prøve til prøve når vi trekker et tilfeldig utvalg av populasjonen - med andre ord vil det være den naturlige variasjonen i gjennomsnittet i utvalgene. Standardfeilen, har vi tidligere forklart, måler hvor mye vi forventer et estimat fra et enkelt utvalg vil avvike fra det sanne gjennomsnittet av

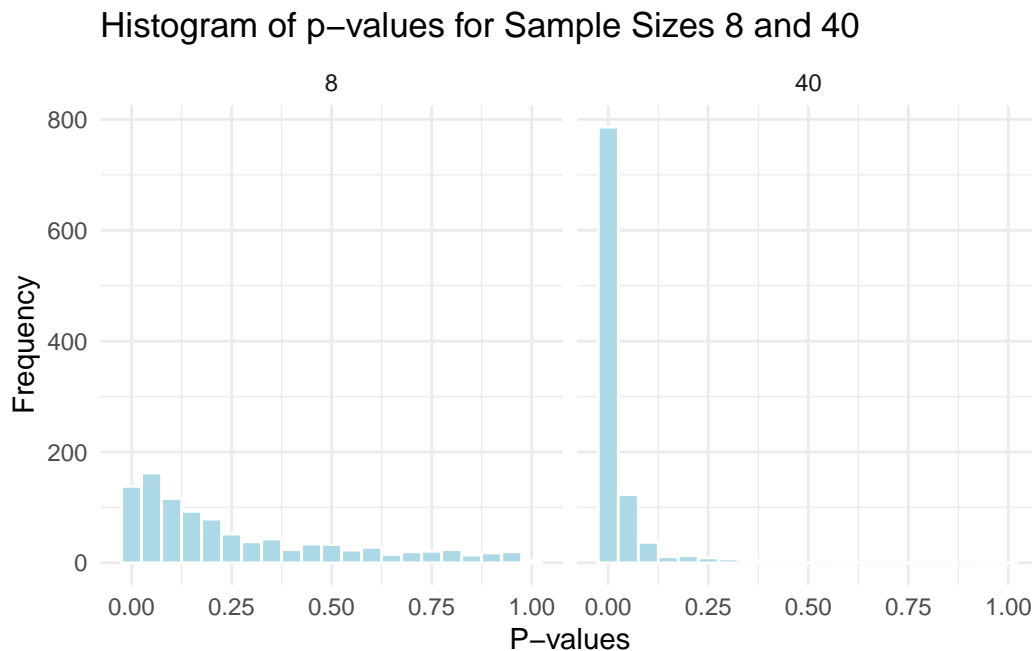
populasjonen. Begge disse målene er relatert til den variasjonen vil vi kunne se i dataene, og sier oss noe om usikkerheten rundt gjennomsnittet i utvalget.

Standardfeilen kan på mange måter ses på som et mål for presisjonen på estimatet (Spiegelhalter (2019a), s.231). I vårt tilfelle vil det si hvor presist utvalgets gjennomsnitt estimerer populasjonens gjennomsnitt. Dess flere prøver man tar, vil standardfeilen nærme seg standardavviket av estimatene, fordi begge måler hvor mye estimatene varierer rundt det sanne gjennomsnittet.

4.2.2 Spørsmål 5 - Histogram av p-verdier og effekten av utvalgsstørrelse på statistisk styrke

5. Create a histogram of the p-values from each study sample-size. How do you interpret these histograms, what do they tell you about the effect of sample size on statistical power?

Figure 4.2 under viser fordelingen av p-verdier for $n = 8$ og $n = 40$, basert på de 1000 prøvene som har blitt generert.



Figur 4.2: Histogram of p-values for Sample Sizes 8 and 40

Basert på tidligere diskusjon rundt usikkerheten og variasjon rundt estimatene, vil vi forvente å se en fordeling med flere p-verdier som er høyere enn 0.05 og dermed ikke statistisk signifikant

for $n = 8$. Noe som kommer tydelig frem i Figure 4.2. For $n = 40$ vil vi motsetning forvente at flere p-verdier som er lavere en 0.05, noe som vi også er tydelig i Figure 4.2.

Statistisk power, er beskrevet som sannsynligheten for å oppdage en sann effekt om den er tilstede (Spiegelhalter (2019a), s.285). I praksis kan vi tenke på at høyere statistisk power øker sjansen til å finne signifikante resulateter når det finnes en reel effekt.

En faktor som vi så fint får visualisert i Figure 4.2, er at statistisk power avhenger av blant annet utvalgsstørrelsen. Basert på det vi observerer i histogrammet, vil vi med større sikkerhet si at vi står bedre egnet til å oppdage en sann effekt (her det sanne gjennomsnittet) ved $n = 40$.

4.2.3 Spørsmål 6 - Antall studier med signifikante effekter for hver utvalgsstørrelse

6. Calculate the number of studies from each sample size that declare a statistical significant effect (specify a threshold α for your significance level)

For at studiene fra hvert utvalg skal bli med i summeringen, må de ha eit signifikansnivå på $\alpha < 0.05$. Under i Table 4.3 ser man oversikten av hvor mange studier fra hvert utvalgsmål som er statistiske signifikante:

Tabell 4.3: Proportion of Significant Results by Sample Size

n	
8	0.227
40	0.865

4.2.4 Spørsmål 7 - Beregning av statistisk styrke for en t-test med ulike utvalgsstørrelser ved bruk av pwr-pakken.

7. Using the pwr package, calculate the power of a one-sample t-test, with a effect size of 1.5/3, your specified significance level and sample sizes 8 and 40. Explain the results in the light of your simulations.

I Table 4.4 er det vist den statistiske poweren for de ulike utvalgsstørrelsene. Som det ble nevnt tidligere, vil en høyere power styrke sannsynlighetene for å finne en signifikant forskjell viss den eksisterer, f.eks: en statistisk power på 0.8 tilsvarer at det er 80% sannsynlighet for at vi vil oppdage en statistisk forskjell viss den eksisterer.

Tabell 4.4: Statistical Power for Different Sample Sizes

Sample.Size	Power
8	0.2320770
40	0.8693981

Den lave styrken for $n = 8$ sammenlignet med $n = 40$ som vi ser i Table 4.4, henger sammen med at det var færre ikke-signifikante tester for $n = 8$ versus $n = 40$ i Table 4.3. Gjennomgående for når vi har sammenlignet de statistiske parametrene mellom de ulike utvalgene, er at eit større utvalg har gitt meir presisjon. Dette samsvarer godt med den statistiske poweren også: større utvalg vil øke sannsynligheten for signifikante resultater og økt statistisk power.

4.2.5 Spørsmål 8 - Antall falske positive resultater ved 5 % signifikansnivå ved gjentatte studier.

8. With a significance level of 5%, how many studies would give you a “false positive” result if you did many repeated studies?

I koden under har eg laget et nytt datasett med resultater fra flere tilfeldige utvalg med $n = 8$ og $n = 40$, fra en populasjon med en gjennomsnittlig effekt på 0.

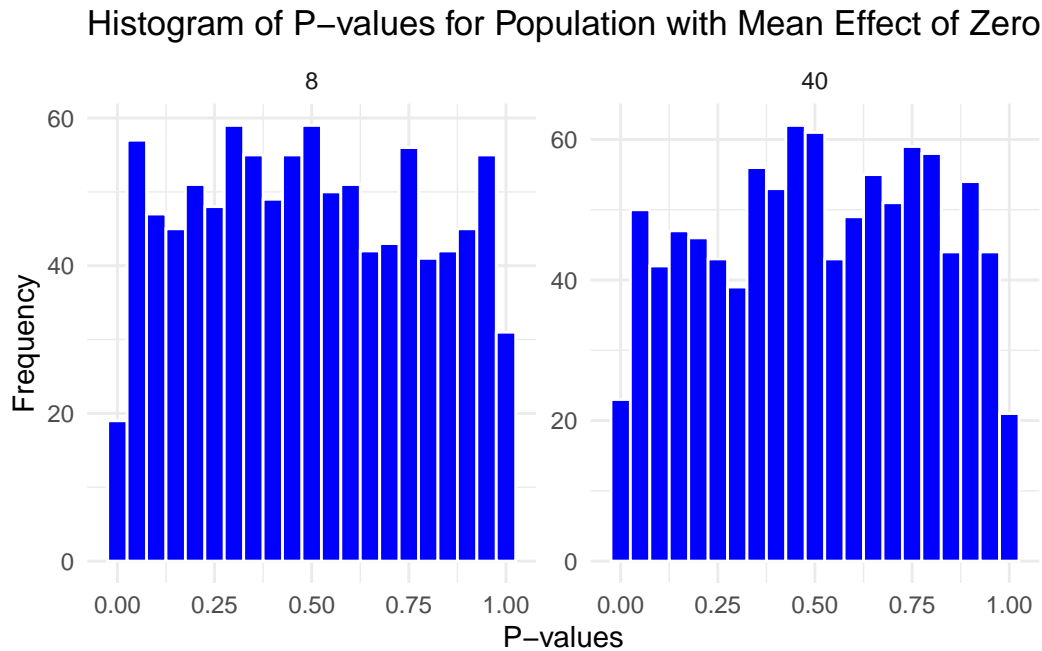
Videre har eg laget et histogram, Figure 4.3, for å visualisere andelen av p-verdier for de forskjellige utvalgsstørrelsene.

Et “falskt positivt” resultat oppstår når p-verdien er mindre enn signifikansnivået (her $\alpha < 0.05$), selv om den sanne effekten i populasjonen er null (Spiegelhalter (2019a), s.278). I Table 4.5 ser vi at selv om effekten på populasjonen er lik null, er det omtrent 5% av studiene for begge utvalgsstørrelsene som feilaktig viser en statistisk signifikant effekt. Dette demonstrer i praksis konseptet med et signifikansnivå ($\alpha < 0.05$), som betyr at vi tillater en 5% sjanse for å forkaste nullhypotesen når den er sann, bedre kjent som type I feil (Spiegelhalter (2019a), s.283).

Tabell 4.5: Resultater av falske positive og ikke-signifikante tester per utvalgsstørrelse

Tabell 4.5: Resultater av falske positive og ikke-signifikante tester per utvalgsstørrelse

n	Antall falske positive	Falske positive (%)	Antall ikke-signifikante
8	44	4.4	956
40	49	4.9	951



Figur 4.3: Histogram of P-values for Population with Mean Effect of Zero

4.3 Vedlegg: Koder brukt

```
# Laster inn nødvendige pakker
library(tidyverse)
library(gridExtra)
library(gt)
library(knitr)
library(pwr)

set.seed(1) # Sikrer at alle de tilfeldige tallene som generes er de
# sammen hver gang koden kjøres
population <- rnorm(1000000, mean = 1.5, sd = 3) # Generer 1 million
# tilfeldige verdier som er normalfordelt

# Sample 1
samp1 <- data.frame(y = sample(population, 8, replace = FALSE))

# Sample 2
```

```

samp2 <- data.frame(y = sample(population, 40, replace = FALSE))

# Lager lineære modeller for hver sample med kun en konstant
m1 <- lm(y ~ 1, data = samp1)
m2 <- lm(y ~ 1, data = samp2)

# Henter ut nødvendige parametere fra modellene
m1_summary <- summary(m1)$coefficients
m2_summary <- summary(m2)$coefficients

# Oppretter data for tabellen
parameter_table <- data.frame(
  Parameter = c("Estimat", "Standardfeil", "t-verdi", "p-verdi"),
  m1 = c(m1_summary[1, 1], m1_summary[1, 2], m1_summary[1, 3],
        m1_summary[1, 4]),
  m2 = c(m2_summary[1, 1], m2_summary[1, 2], m2_summary[1, 3],
        m2_summary[1, 4])
)

## Definer t-verdier for begge utvalg
t_value1 <- 1.47 # t-verdi for samp1
t_value2 <- 3.276 # t-verdi for samp2
df1 <- 7 # Antall frihetsgrader for samp1 (n-1, her 8-1=7)
df2 <- 39 # Antall frihetsgrader for samp2 (n-1, her 40-1=39)

# Lag dataramme for x-verdier
x_vals <- seq(-5, 5, length.out = 1000)

# Utrekne p-verdier for m1 og m2
p1 <- 2 * pt(t_value1, df = df1, lower.tail = FALSE)
p2 <- 2 * pt(t_value2, df = df2, lower.tail = FALSE)

# Lager data frames for å lagre modell estimator
results_8 <- data.frame(estimate = rep(NA, 1000),
                        se = rep(NA, 1000),
                        pval = rep(NA, 1000),
                        n = 8)

results_40 <- data.frame(estimate = rep(NA, 1000),
                          se = rep(NA, 1000),
                          pval = rep(NA, 1000),
                          n = 40)

```

```

# En løkke som for å hente ut 1000 prøver(samples), hver iterasjon (i) vil
# hente ut en ny prøve fra populasjonen

for(i in 1:1000) {

  # Henter ut en prøve
  samp1 <- data.frame(y = sample(population, 8, replace = FALSE))
  samp2 <- data.frame(y = sample(population, 40, replace = FALSE))

  # Lager en lineær modell
  m1 <- lm(y ~ 1, data = samp1)
  m2 <- lm(y ~ 1, data = samp2)

  # Henter ut verdier fra modellen
  results_8[i, 1] <- coef(summary(m1))[1, 1]
  results_8[i, 2] <- coef(summary(m1))[1, 2]
  results_8[i, 3] <- coef(summary(m1))[1, 4]

  results_40[i, 1] <- coef(summary(m2))[1, 1]
  results_40[i, 2] <- coef(summary(m2))[1, 2]
  results_40[i, 3] <- coef(summary(m2))[1, 4]

}

# Lagrer resultatene i en data frame
results <- bind_rows(results_8, results_40)

# Filtrer ut resultatene for n = 8 og n = 40
results_n8 <- results %>% filter(n == 8)
results_n40 <- results %>% filter(n == 40)

# Beregn standardavvik for estimate og gjennomsnitt for SE
summary_table <- data.frame(
  "Utvalgsstørrelse" = c(8, 40),
  "Standardavvik" = c(sd(results_n8$estimate), sd(results_n40$estimate)),
  "SE" = c(mean(results_n8$se), mean(results_n40$se)))

# Beregn andelen signifikante resultater

```



```

sig_table <- results %>%
  filter(pval < 0.05) %>%
  group_by(n) %>%
  summarise(` ` = n()/1000)

# Beregn statistisk styrke for n = 8 og n = 40
power_n8 <- pwr.t.test(n = 8, sig.level = 0.05, d = 1.5/3, type = "one.sample")
power_n40 <- pwr.t.test(n = 40, sig.level = 0.05, d = 1.5/3, type = "one.sample")

# Lag en dataframe med resultatene
power_results <- data.frame(
  "Sample Size" = c(8, 40),
  "Power" = c(power_n8$power, power_n40$power)
)

population <- rnorm(1000000, mean = 0, sd = 3)

# Lag data frames for å lagre modellene sine estimater
results_8 <- data.frame(estimate = rep(NA, 1000),
  se = rep(NA, 1000),
  pval = rep(NA, 1000),
  n = 8)

results_40 <- data.frame(estimate = rep(NA, 1000),
  se = rep(NA, 1000),
  pval = rep(NA, 1000),
  n = 40)

# Lagre resultatene i en data frame

results_null <- bind_rows(results_8, results_40)

# Beregn totalt antall tester og falske positive
total_tests <- 1000 # Antall tester
false_positives <- results_null %>%
  filter(pval < 0.05) %>%
  group_by(n) %>%
  summarise("Antall falske positive" = n(),
    "Falske positive (%)" = (n() / total_tests) * 100)

```

```
# Legg til totalt antall tester for å beregne ikke-signifikante resultater
false_positives <- false_positives %>%
  mutate("Antall ikke-signifikante" = total_tests - `Antall falske positive`)
```

5 Studiedesign

Tradisjonelt sett er plyometrisk trening den vanligste treningsmetoden for å forbedre en persons evne til rask kraftutvikling, og er derfor ofte brukt innen sprint- og hoppidretter (Markovic & Mikulic, 2010). Selv om evnen til rask kraftutvikling er viktig for prestasjon, er utvikling av maksimal kraft også viktig. Et individs evne til å produsere impuls, som er et produkt av rate of force development (RFD) og maksimal kraft, er en avgjørende egenskapen for prestasjon i eksplosive og kraftfulle idretter (Cleather, 2021, pp. s.20–22). For å bedre forstå hvordan plyometrisk trening påvirker prestasjon i slike idretter, har jeg valgt ut fem studier som ser nærmere på effekten av plyometrisk trening på kraftutvikling.

De fem inkluderte studiene vil bli analysert ved hjelp av QALMRI-metoden (“Question”, “Alternatives”, “Methods”, “Results” og “Inferences”), en evidensbasert tilnærming som sikrer en grundig forståelse av nøkkelpunktene og kritisk vurdering av empiriske artikler (Brosowsky et al., 2020). I stedet for å gi en detaljert gjennomgang av hver enkelt studie, vil jeg vektlegge og sammenligne studienes design og valg av statistiske analyser som er brukt for å besvare forskningsspørsmålene. Til slutt vil jeg komme med min anbefaling for hvordan fremtidige studier kan designes for best mulig å undersøke effekten av plyometrisk trening på muskelstyrke.

5.1 Analyse

De studiene som jeg har valgt er:

- Harput et al. (2023): “Effects of Plyometric Training on Sonographic Characteristics of Quadriceps Muscle and Patellar Tendon, Quadriceps Strength, and Jump Height in Adolescent Female Volleyball Players”
- Fatouros et al. (2000): “Evaluation of Plyometric Exercise Training, Weight Training, and Their Combination on Vertical Jumping Performance and Leg Strength”
- McKinlay et al. (2018): “Effects of Plyometric and Resistance Training on Muscle Strength, Explosiveness, and Neuromuscular Function in Young Adolescent Soccer Players”
- Vissing et al. (2008): “Muscle Adaptations to Plyometric vs. Resistance Training in Untrained Young Men”

- Whitehead et al. (2018): “A Comparison of the Effects of Short-Term Plyometric and Resistance Training on Lower-Body Muscular Performance”

5.1.1 Problemstilling og hypoteser

De utvalgte studiene har som overordnet mål å undersøke hvordan plyometrisk trening påvirker sentrale aspekter av fysisk prestasjon, inkludert maksimal styrke, eksplosivitet og muskeltykkelse, som er avgjørende i eksplosive og kraftfulle idretter. Hver studie adresserer hver sin spesifikke problemstilling knyttet til dette målet. Fatourous et al. (2000) ønsker å sammenligne effekten plyometrisk trening, tradisjonell styrketrening og kombinasjonstrening har på vertikal hopp prestasjon (Fatouros et al., 2000). Vissing et al. (2008) undersøker hvilke adaptive endringene i maksimal styrke og power som skjer som følge av enten tradisjonell styrketrening eller plyometrisk styrketrening (Vissing et al., 2008). McKinlay et al. (2018) adresserer hvordan plyometrisk trening påvirker nevro-muskulære funksjoner og muskulære tilpasninger hos unge fotballutøvere (McKinlay et al., 2018). Whitehead et al. (2018) fokuserer på kortsiktige effekter av plyometrisk og styrketrening på muskelstyrke og eksplosiv prestasjon (Whitehead et al., 2018). Mens Harput et al. (2023) utforsker effekten av plyometrisk trening på hopptøyle, muskeltykkelse og muskelstyrke i quadriceps hos unge kvinnevennlige volleyballspillere (Harput et al., 2023).

Ingen av studiene definerer et klart forskningsspørsmål knytt til problemstillingen sin. De fleste av de aktuelle studiene presenterer en klar hypotese om at plyometrisk trening vil gi lik eller bedre resultater i eksplosivitet, muskelstyrke eller nevro-muskulære tilpasninger sammenlignet med andre treningsmetoder, som Harput et al. (2023), Vissing et al. (2008), McKinlay et al. (2018) og Whitehead et al. (2018) (Harput et al., 2023; McKinlay et al., 2018; Vissing et al., 2008; Whitehead et al., 2018). Fatouros et al. (2000) derimot, beskriver heller formålet med studien uten å formulere et forskningsspørsmål eller en hypotese (Fatouros et al., 2000).

Hypotesene i studiene deler en antakelse om at plyometrisk trening vil kunne forbedre muskelens evne til å generere kraft hurtig, altså power (Raastad et al., 2010, p. s.225). Det finnes imidlertid flere alternative forklaringer på resultatene som kan påvirke dataene. En mulig årsak kan være at intensiteten i den plyometriske treningen var høyere enn kontrolltreningen, noe som førte til forbedringer uavhengig av treningsmetode (Raastad et al., 2010, pp. s.231–233). Deltagernes treningsbakgrunn kan også påvirke responsen på treningen, i tillegg til faktorer som dagsformen på testdagen, som kan bidra til ytterligere variasjon i prestasjonene. Det er også mulig at forbedringene i eksplosivitet skyldes nevro-muskulære tilpasninger enn spesifikke effekter av plyometrisk trening (Raastad et al., 2010, p. s.64). Videre kan placeboeffekten spille en rolle, da forbedringene kan være påvirket av deltakernes forventninger om framgang ved å delta i et treningsopplegg.

Plyometrisk trening beskrives som en sentral metode for å forbedre prestasjon der det stilles krav til hurtig kraftutvikling, og antas å kunne gi lignende eller bedre muskulære og nevro-muskulære tilpasninger sammenlignet med tradisjonell styrketrening (Markovic & Mikulic,

2010; Whitehead et al., 2018, pp. s.2743–2744). Samtidig påpekes det at det finnes begrenset forskning på hvorfor plyometrisk trening forbedrer prestasjon, og hvilke spesifikke tilpasninger som ligger bak disse forbedringene (Fatouros et al., 2000, p. s.471; Harput et al., 2023, p. s.89; McKinlay et al., 2018, pp. s.3039–3040; Vissing et al., 2008, p. s.1800). Basert på disse kunnskapshullene har forfatterne av de utvalgte studiene formulert hypoteser om at plyometrisk trening vil kunne gi lik eller større forbedring i prestasjon i bevegelser som krever rask kraftutvikling, sammenlignet med tradisjonell styrketrening, og samtidig undersøke hvilke underliggende mekanismer som fører til disse tilpasningene.

5.2 Metode

5.2.1 Valg av studiedesign

Valg av studiedesign er avgjørende for å kunne svare på forskningsspørsmålet eller teste hypoteser. Studiedesignet strukturerer studien på en måte som gjør det mulig å trekke holdbare slutninger (Browner et al., 2022, p. s.3). Valget av design avhenger blant annet om man planlegger å igangsette en intervensjon for å undersøke dens effekter. Slike studier, kjent som analytiske studier, prøver å evaluere sammenhenger og trekke slutninger om årsak-virkning-forholdet (Browner et al., 2022, p. s.5). En typisk analytisk studiedesign er en randomisert kontrollert studie (RCT). Før man gjennomfører en slik studie, er det vanlig at deskriptive studier har blitt utført for å undersøke karakteristikker ved en populasjonen, som for eksempel muskelmasse hos eliteutøvere (Browner et al., 2022, p. s.4). En tydelig beskrivelse av studiedesignet og begrunnelse av valget er avgjørende for å gi leserne instinkt i forskerens vurderinger.

Som nevnt tidligere, er det overordnet målet med de utvalgte studiene å undersøke hvordan plyometrisk trening påvirker fysisk prestasjon, inkludert maksimal styrke, eksplosivitet og muskelmasse. Fire av studiene har valgt å tilfeldig fordele deltakerne til enten å gjennomføre trening (intervensjon) eller ikke (kontroll) (Fatouros et al., 2000; McKinlay et al., 2018; Vissing et al., 2008; Whitehead et al., 2018), slik det gjøres i en RCT-studie. Den femte studien valgte imidlertid et prospektivt kohortdesign, der deltakerne ble fulgt opp over tid uten randomisering (Harput et al., 2023). Ved å randomisere deltagerne mellom gruppene i RCT-studiene minimeres påvirkningen av konfunderende variabler som kan forstyrre resultatene, mens kohortstudier gir mulighet til å observere naturlige endringer i en populasjon over tid uten intervensjon i gruppefordelingen (Browner et al., 2022, pp. s.4, 6, 116, 196).

I RCT-studiene som Fatouros et al. (2000), McKinlay et al. (2018), Vissing et al. (2008) og Whitehead et al. (2018) gjennomførte, vil valget om randomisering bidra til å redusere risikoen for at andre faktorer enn treningen selv påvirker resultatene, som treningsbakgrunn, noe som vil styrke validiteten av konklusjonene om årsakssammenheng mellom intervensjon og fysisk prestasjon. Ingen av studiene benyttet derimot blinding i sitt design, noe som kan ha introdusert rapporteringsbias, særlig siden både deltakerne og forskerne visste hvilken gruppe

deltakerne tilhørte (Browner et al., 2022, pp. s.196, 397). Dette er ikke noe nytt for treningsintervensjoner, der det er praktisk vanskelig å skjule gruppetilhørighet.

I kohortstudien av Harput et al. (2023), var det derimot ikke mulig å bruke randomisering. Forskerne fulgte deltakerne over tid for å observere endringer etter en allerede igangsatt treningsintervensjon, noe som tillot forskerne å se naturlige variasjoner i prestasjonene uten inngrep i gruppefordelingen (Browner et al., 2022, pp. s.4, 116). Samtidig gjør dette at kohortstudier er mer utsatt for konfunderende variabler, siden deltakerne ikke blir tilfeldig fordelt i grupper. En konfunderende variabel er en faktor som kan påvirke både intervensjonen og utfallet, og dermed skape en skjev sammenheng (Spiegelhalter, 2019a, p. 110). Dette kan redusere validiteten, men slike design er likevel nyttige for å undersøke langtidseffekter i realistiske settinger.

5.2.2 Utvalg

Ingen av studiene definerer eller beskriver en bestemt populasjon i detalj, og det er varierende grad av informasjon om deltakerne. Generelt tar studiene for seg unge, aktive idrettsutøvere eller utrente menn/ungdommer, men ikke alle går i dybden på demografiske faktorer som kan påvirke generaliserbarheten av resultatene (Browner et al., 2022, pp. s.26–27). For eksempel bestod utvalget i Vissing et al. (2008) av en relativt homogen gruppe med utrente menn, mens Harput et al. (2023) fokuserte på en mer spesifikk, men potensielt heterogen gruppe av unge volleyballspillere (Harput et al., 2023; Vissing et al., 2008). Selv om informasjon om gruppene kan være mangelfull, ser de fleste studiene ut til å ha relativt homogene utvalg med hensyn til treningsstatus og idrettsbakgrunn.

Det er blitt brukt varierende metoder for å rekruttere deltakere i studiene, og graden av detaljer informasjon om rekruttering og begrunnelse for utvalgsstørrelse varierer. Både Vissing et al. (2008) og Fatouros et al. (2000) rekrutterte frivillige (Fatouros et al., 2000; Vissing et al., 2008), mens Harput et al. (2023) brukte en kohort av volleyballspillere (Harput et al. (2023)). De to siste studiene gir derimot lite informasjon om rekrutteringsprosessen, men det kan virke som deltagerne er rekruttert fra lokale ungdomslag.

Kun Vissing et al. (2008) rapporterte en power-beregning for å sikre tilstrekkelig statistisk styrke (0.8), mens de andre ikke oppga dette (Vissing et al., 2008, p. s.1830). Power referer til sannsynligheten for å oppdage en statistisk signifikant forskjell i utfall mellom gruppene, gitt at det eksisterer en slik forskjell, og det er basert på antakelsen om repeterte studier under like betingelser (Browner et al., 2022, p. s.6). En power-beregning, som gjøres i planleggingsfasen, estimerer hvor mange deltakere som trengs i hver gruppe av studien for å sikre tilstrekkelig sannsynlighet for å oppdage en effekt dersom den finnes (Browner et al., 2022, p. s.6). Dette er særlig viktig i hypotesedrevne studier for å redusere risikoen for å overse reelle forskjeller (type 2-feil) (Spiegelhalter, 2019a, pp. s.283–286).

Når Vissing et al. (2008) oppgir en power på 0.8, betyr det at de med 80% sannsynlighet vil kunne oppdage en forskjell dersom den eksisterer. Selv om power-beregning rapporteres

eller ikke, gir en slik beregning en vurdering av hvor robust designet er til å trekke pålitelige konklusjoner. Mangelen på power-beregning i de andre studiene kan indikere at de ikke har vurdert dette i designfasen, noe som kan svekke deres evne til å oppdage eksisterende effekter og sikre sikre slutninger.

5.2.3 Tester og sentrale variabler

I de inkluderte studiene ble ulike tester gjennomført før og etter treningsintervensjonene for å måle endringer i fysisk prestasjon, muskelstyrke og eksplosivitet. Harput et al. (2023) benyttet isokinetiske styrkemålinger og ultralyd for å vurdere muskeltykkelse, styrke og hopphøyde før og etter en seks ukers periode med plyometrisk trening (Harput et al., 2023). Vissing et al. (2008) og McKinlay et al. (2018) brukte på sin side dynamiske tester som vertikale hopp og sprint før og etter treningsperioden (McKinlay et al., 2018; Vissing et al., 2008). Selv om alle studiene hadde som mål å undersøke effekten av plyometrisk trening, varierer testdesignene avhengig av spesifikke forskningsspørsmål eller hypoteser. Dette fører til ulike valg av variabler som hopphøyde, muskelstyrke og eksplosivitet, som alle er sentrale faktorer for prestasjon i kraftbaserte idretter. Valget av disse variablene reflekterer hver studies formål og hypotese, og illustrerer variasjonene i design og metodevalg mellom studiene.

5.2.4 Statistiske tester

Som tidligere nevnt, har de inkluderte studiene som mål å undersøke hvordan plyometrisk trening påvirker viktige utfallsmål på fysisk prestasjon, som muskelstyrke og hopphøyde. Selv om studiene benytter ulike studiedesign og variabler, tilpasset sitt spesifikke forskningsspørsmål eller hypoteser, har de alle til felles at de måler endringer i sine definerte utfallsmål over tid og mellom treningsintervensjoner. For å kunne trekke konklusjoner om effekten av intervensjonene har alle studiene valgt Analysis of Variance (ANOVA) som sin primære statistiske metode.

ANOVA er en statistisk test som gjør det mulig å sammenligne gjennomsnittet mellom flere grupper for å avgjøre om variasjonen mellom gruppene er så stor at det er usannsynlig å anta at det kun skyldes tilfeldigheter (Diez et al., 2022, pp. s.322, 326). I denne analysen beregnes mean square between groups (MSG), som representerer variabiliteten mellom gruppene. Dette sammenlignes med mean square error (MSE), som reflekterer variasjonene innenfor hver gruppe, og gir et mål på naturlig variasjon når nullhypotesen er sann (Diez et al., 2022, p. s.326). Forholdet mellom MSG og MSE, utgjør F-statistikken:

$$F = \frac{MSG}{MSE}$$

En høy F-verdi indikerer at forskjellen mellom gruppene er stor sammenlignet med variasjonene innad i gruppene, noe som kan tyde på en signifikant effekt av treningen (Diez et al., 2022, pp.

s.326–327). Det er viktigå påpeke at for at ANOVA skal gi sterke bevis mot nullhypotesen om at utvalgsgjennomsnittene (μ_i) er like, må flere forutsetninger være oppfylt. For det første må observasjonene være uavhengige både innad og mellom gruppene. Videre må dataene innenfor hver gruppe være tilnærmet normalfordelt, og variasjonen mellom gruppene er lik (Diez et al., 2022, p. s.322).

I de inkluderte studiene ble både en-veis og to-veis ANOVA brukt for å analysere dataene. En-veis ANOVA sammenligner gjennomsnittene mellom flere grupper ved å bruke en avhengig variable for å avgjøre om det finnes en statistisk signifikant forskjell mellom gruppene (Diez et al., 2022, pp. s.323–325). To-veis ANOVA evaluerer effekten av to forskjellige faktorer på resultatet, noe som gjør det mulig å undersøke både hoved- og samspilleffekter mellom de to faktorene (Diez et al., 2022, pp. s.329–331).

For eksempel brukte Whitehead et al. (2018) en-veis ANOVA for å sammenligne baseline data mellom gruppene (Whitehead et al., 2018, p. s.2746). Denne analysen adresserte hvorvidt gruppene var likeverdige ved starten av studien, en nødvendig forutsetning for å evaluere effekten av treningen. Harput et al. (2023) og McKinlay et al. (2018) brukte to-veis ANOVA for å undersøke interaksjonen mellom tid (før og etter intervensjon) og gruppetilhørighet (treningsintervensjon versus kontroll) (Harput et al., 2023, pp. s.92–93; McKinlay et al., 2018, p. s.3046). Disse analysene var sentrale for å svare på spørsmål om hvorvidt treningsintervensjonen førte til større forbedringer over tid i forhold til kontrollgruppen.

Mens ANOVA forteller at det finnes forskjeller mellom gruppene, sier den ikke noe om hvilke grupper som faktisk skiller seg fra hverandre, noe som gjør det vanskelig å sammenligne treningsintervensjonene opp mot hverandre. For å identifisere spesifikke gruppedifferanser når ANOVA viste signifikante resultater, benyttet flere av studiene post hoc-tester som Bonferroni (Spiegelhalter, 2019b, p. s.280) med signifikansnivå på $p \leq 0.05$. Disse tilleggsanalysene tillot en mer detaljert undersøkelse av hvor forskjellene mellom gruppene oppstod, noe som var avgjørende for å vurdere treningsinduserte effekter og besvare forskningsspørsmålene om effekten av ulike treningsintervensjoner.

5.3 Resultat

Studiene dokumenterer samlet sett en positiv effekt av plyometrisk trening på sentrale prestasjonsmål som muskelstyrke, eksplosivitet og hopphøyde sammenlignet med kontrollgruppene (Fatouros et al., 2000, p. s.474; Harput et al., 2023, p. s.93; McKinlay et al., 2018, p. s.3046; Vissing et al., 2008, p. s.1804; Whitehead et al., 2018, p. s.2747). For eksempel rapporterte Harput et al. (2023) signifikante forbedringer i hopphøyde og quadriceps-styrke etter seks uker med strukturert plyometrisk trening (Harput et al., 2023, p. s.93). Vissing et al. (2008) viste at plyometrisk trening førte til målbare økninger i eksplosiv styrke, inkludert forbedrede resultater i vertikale hopp og sprinttid (Vissing et al., 2008, p. s.1804). På tilsvarende måte dokumenterte McKinlay et al. (2018) at regelmessig plyometrisk trening

hos unge fotballspillere førte til en markant forbedring i nevromuskulær funksjon og styrke (McKinlay et al., 2018, p. s.3046).

Fatouros et al. (2000) fant også at plyometrisk trening kombinert med styrketrening resulterte i ytterligere forbedringer, sammenlignet med plyometrisk trening alene (Fatouros et al., 2000, p. s.474). Totalt sett belyser disse funnene forskningsspørsmålene og hypotesene i studiene, og dokumenterer målbare forbedringer i prestasjon som et resultat av plyometrisk trening, både alene og i kombinasjon med andre treningsformer.

5.4 Konklusjonene fra studiene

Hovedfunnene fra studiene viser en signifikant sammenheng mellom plyometrisk trening og bedring i sentrale prestasjonsmål innen kraftfulle og eksplosive idretter, da særlig muskelstyrke, hopp høyde og hopp høyde. Studien rapporterte målbare forbedringer blant deltakerne sammenlignet med kontrollgruppene, noe som var et sentralt formål i hver studie. Fatouros et al. (2000) fant også ut at en kombinasjon av plyometrisk trening og styrketrening ga større prestasjonsforbedringer enn plyometrisk trening alene, noe som indikerer at sammensatte treningsformer kan gi ytterligere fordeler sammenlignet med isolerte plyometriske øvelser (Fatouros et al., 2000, p. s.474). Likevel så understreket Fatouros et al. (2000), at man skal være forsiktig å trekke konkrete slutninger, da gruppen som mottok den kombinerte treningsintervensjon også ble eksponert for et høyere samlet treningsvolum enn gruppen som kun gjennomførte plyometriske trening (Fatouros et al., 2000, p. s.474).

Utvalgene i studiene bestod hovedsaklig av unge, aktive individer, inkludert utrente menn og ungdomsidrettsutøvere som volleyballspillere og fotballspillere. Forfatterne bemerket at den observerte effekten kan være påvirket av spesifikke egenskaper ved utvalget, som alder og treningsbakgrunn (Harput et al., 2023, pp. s.94–95; McKinlay et al., 2018, pp. s.3048–3049). Dette indikerer at resultatene hovedsaklig er relevante for lignende grupper og kanskje ikke kan overføres til andre grupper, som eldre voksne eller personer med mer treningserfaring.

5.5 Anbefaling til fremtidige studier

Ved å bygge videre på de inkluderte studiene om plyometrisk trening, bør fremtidige studier fokusere på å optimalisere forskningsdesignet for å oppnå sterkere og mer generaliserbare konklusjoner. En anbefaling er å inkludere flere aldersgrupper, treningsnivåer og kjønn. Dette vil gi bedre innsikt i hvordan plyometrisk trening påvirker ulike populasjoner, inkludert eldre voksne og personer med mer omfattende treningserfaring. Videre bør fremtidige studier inkludere flere målepunkter over tid for å undersøke de langsiktige effektene av plyometrisk trening.

For å kunne trekke kausale slutninger om effekten av plyometrisk på ulike populasjoner er RCT-er det foretrukne studiedesign, ettersom de reduserer påvirkningen av konfunderende faktorer. Selv om blinding er ideelt, er det vanskelig å få til praktisk i treningsintervensjoner. For å få innsikt i langsiktige endringer og vedvarende effekter anbefales prospektive kohortstudier over lengre tid. Kombinasjonen av disse designene vil gi et mer helhetlig bilde av både kortsiktige og langsiktige effekter av plyometrisk trening.

Med tanke på statistiske analyser er variasjonsanalyser velegnede for de anbefalte studiedesign. I RCT-er der kontroll over baselineforskjeller er ønskelig, som alder eller treningsbakgrunn, vil **Analysis of Covariance (ANCOVA)** være en passende metode. ANCOVA justerer for varians knyttet til konfunderende faktorer (kovariater, f.eks alder eller kjønn) for å sikre at treningseffektene kan observeres med større presisjon (Keselman et al., 1998, p. s.373). I prospektive kohortstudier kan **mixed effects modeller** være ideelle for å håndtere data med repeterte målinger og variasjon både mellom og innen individer (Baayen et al., 2008, pp. s.409–410). Dette vil muliggjøre justeringer for både individuelle forskjeller og gruppeforskjeller, slik at man får en dypere forståelse av hvordan treningseffekten utvikler seg over tid.

Ved å kombinere slike statistiske metodene med nøye gjennomtenkte forskningsdesign kan fremtidige studier gi en mer presis og helhetlig forståelse av plyometrisk trening sine effekter på fysisk prestasjon i ulike populasjoner.

6 Analyse av eksperimenter med repeterte målinger

Treningsadaptasjon er en av de mest grunnleggende søylene i treningsvitenskapen, som sier at kroppen vår kan tilpasse seg den belastningen den utsettes for, gitt at den får nok tid til å tilpasse seg den nye belastningen (Bompa & Buzzichelli, 2019, s. 8). Spesielt skjelettmuskulaturen har vist seg å være et høgt tilpasningsdyktig vev som kan endre sin egen arkitektur som respons på en gitt stimuli (Coffey & Hawley, 2007). Resultatene på disse tilpasningene er påvirket av volumet, intensiteten og frekvensen av treningen, i tillegg vil tilpasningene avhengig av den spesifikke belastningen kroppen og muskulaturen utsettes for, for eksempel treningsmetode (Bompa & Buzzichelli, 2019, p. s.3; Coffey & Hawley, 2007; Izquierdo et al., 2004; Raastad et al., 2010, p. s.17).

Styrketrening har vist seg å være avgjørende, ikke bare for idrettsprestasjoner (Washif & Kok, 2022), men også for folkehelse og redusert dødelighet (Ruiz et al., 2008). Denne treningsformen omfatter aktiviteter som utvikler eller opprettholder vår evne til å generer maksimal kraft eller dreiemoment ved en spesifikk hastighet (Raastad et al., 2010, p. s.17). Trening for maksimal muskelstyrke, der målet er å øke den høyeste kraften en muskel kan produsere (målt som en repetisjon maksimum) (Raastad et al., 2010, p. s.13), skiller seg fra trening med hensikt om muskelvekst (hypertrofi), hvor fokus ligger på å øke muskelens størrelse (Schoenfeld, 2016). Flere studier har undersøkt hvilket treningsvolum som best stimulerer slike tilpasninger og har identifisert et dose-respons-forhold, hvor økt treningsvolum er assosiert med større muskelvekst (Currier et al., 2023; Schoenfeld, 2016). Et lignende forhold er også blitt observert for muskelstyrke (Ralston et al., 2017), selv om det er også foreslått at et høgt treningsvolum kan bremse utviklingen av maksimal muskelstyrke (Zhang et al., 2023). Muskelens tverrsnitt er en sentral faktor å ta med, da et større tverrsnitt vil gi et større potensiale for kraftutvikling (Raastad et al., 2010, p. s.20). Dette understreker behovet for videre undersøkelser av hvordan treningsvolum påvirker både muskelhypertrofi og styrkeutvikling, samt i hvilken grad endringer i muskelmasse er sammenfallende med økninger i styrke.

På bakgrunn av tidligere forskning og den noe usikre dose-respons-sammenhengen mellom optimal treningsvolum og muskeltilpasninger, er formålet med denne studien å undersøke hvordan ulike nivåer av treningsvolum påvirker både muskelvekst og muskelstyrke, ettersom skjelettmuskulaturen er svært tilpasningsdyktig på den stimulien den utsettes for. Ved å utforske denne sammenhengen håper man å kunne bidra med økt innsikt som kan være nyttig ved anbefalinger om optimalt treningsvolum.

6.1 Metode

6.1.1 Deltagere og studieoversikt

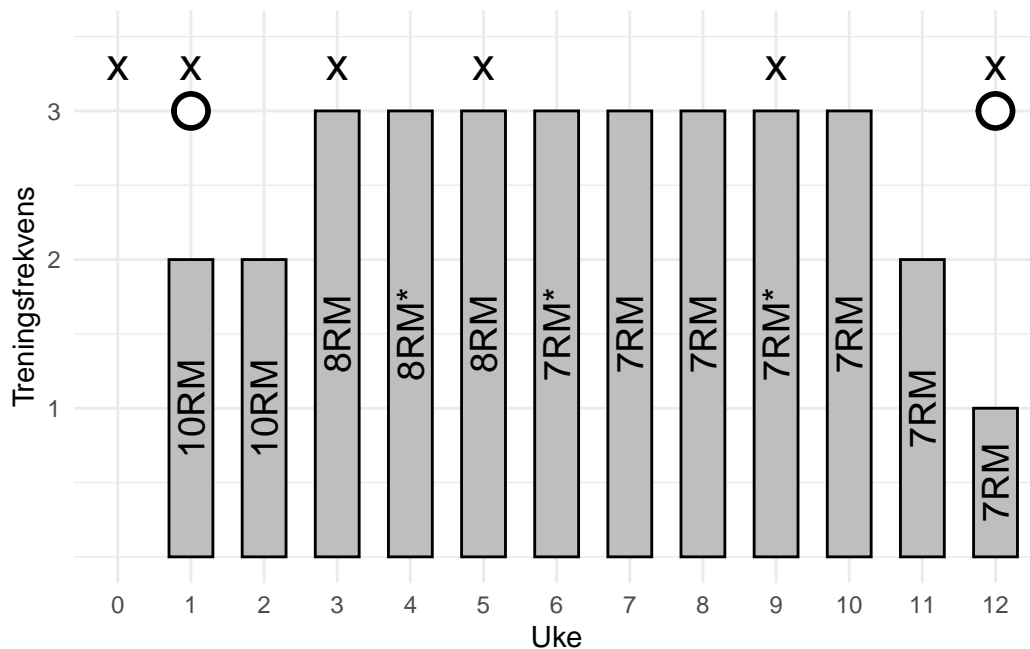
Totalt ble førti-en menn og kvinner inkludert i studien, der inklusjonskriteriene var at de ikke var røykere og var mellom 18 og 40 år. Klare eksklusjonskriterer som redusert muskelstyrke grunnet tidligere eller nåværende skader, har hatt mer en en ukentlig styrketrenings økt de siste 12 månedene fra inklusjonsdato, intoleranse for lokal bedøvelse, samt bruk av medikamenter som kan påvirke treningsadaptasjoner. Syv deltakere ble ekskludert fra data analysen, fordi de ikke hadde gjennomført minimum 85% av de oppsatte treningene grunnet årsaker som: ubehag eller smerter i underekstremitetene under trening ($n = 5$), skade påført utenom studieprotokollen ($n = 1$), eller manglende overholdelse av studiedesignet ($n = 1$). Selv om ingen av de inkluderte deltakerne utførte mer en 1 ukentlig styrkeøkt, var det tjue deltakere som rapporterte at de var fysisk aktive ved inklusjon (median på 2 økter i uken, med en range på 0.5-4). Alle de inkluderte hadde tidligere treningserfaringer med ulike idretter, som for eksempel lagidretter, skiskytting etc...).

Tabell 6.1: Deltakerkarakteristikker: Verdiene er presentert som gjennomsnitt og standardavvik (SD)

	Kvinne		Mann	
	Ekskludert	Inkludert	Ekskludert	Inkludert
N	4	18	3	16
Alder	22.9 (1.6)	22.0 (1.3)	24.3 (1.5)	23.6 (4.1)
Vekt	64.6 (9.7)	64.4 (10.4)	88.2 (22.4)	75.8 (10.7)
Stature	166 (8)	168 (7)	189 (5)	183 (6)

Deltakerne gjennomgikk en treningsintervensjon med 12 uker styrketrening for hele kroppen i perioden september til november. For å muliggjøre for differensiering av treningsvolum for hver deltaker, ble beinøvelsene utført unilateralt. Hver deltaker fikk tilfeldig utdelt enten ett sett eller tre sett til enten høyre eller venstre bein. Deltagerne utførte dermed begge volumprotokollene. Det ble gjennomført måling av maksimal styrke ved baseline, under (uke 3,5 og 9) og rett etter intervensjonen, mens muskeltverrsnitt ble målt før og etter, se Figure 6.1.

De ulike søylene illustrer treningsfrekvens per uke med en treningsintensitet lik x repetisjon maksimum (RM). De intensitetne som er markert med (*), referer til at en av de øktene den uken ble utført med 90% av 1RM. Sirkelsymbolet viser til når det ble utført målinger av muskeltverrsnitt med hjelp av fullkroppss DXA og MR av kneekstensjons muskelen. X markerer de ulike styrkemålingene: før intervensjonen ($n = 34$), under ($n = 18$) og etter ($n = 34$). Den maksimale styrken før intervensjonen ble satt som den høyeste verdien deltagerne oppnådde under to ulike testsekvenser før intervensjonen startet.



Figur 6.1: Studieoversikt

6.1.2 Styrketreningsprotokoll

Det ble laget en standardisert oppvarming som deltakerne skulle gjennomføre før de begynte med styrketreningen. Oppvarmingen begynte med 5 minutter på ergometer sykkel, der de skulle holde 12-14 på gjennomgått grad av anstrengelse (RPE). Etter syklingen skulle de gjennomføre 10 repetisjoner av push-ups (tilpasset vanskelighetsgrad basert på vinkel), sit-ups, knebøy og rygghev, etterfulgt av et sett med 10 repetisjoner av hver styrkeøvelse med en motstand tilsvarende rundt 50% av deres repetisjons maksimum (1RM).

Styrketreningen inneholdt en del for beina og en for overkroppen. For beina skulle det gjennomføres unilateral beinpress, knefleksjon og knestrekk, i den nevnte rekkefølgen. Avhengig av gruppen deltakerne ble tildelt til, ble hver øvelse enten utført med ett sett eller tre sett. Beinets som ble tildelt ett-setts protokollen gjennomførte arbeidet sitt mellom det andre og tredje settet til beinet som skulle gjennomføre tre sett. Pausene mellom settene skulle vare mellom 90-180 sekunder. Etter beinøvelsene hadde blitt fullført, skulle deltakerne gjøre to sett av bilateral benkpress, nedtrekk og enten skulderpress eller sittende roinger som ble gjort i hver sin økt. Som man kan se fra Figure 6.1, så ble treningsintensiteten gradvis økt fra 10RM to første ukene hadde 10RM, til 8RM de påfølgende tre ukene og 7RM de siste syv ukene. Etter den niende treningsøkten hadde hver uke som inneholdt tre økter i uken, en økt med lavere motstand, men der man beholdt samme antall repetisjoner (motstanden tilsvarte 90% av økten før). Øktene med maksimal innsats skulle ikke gjentas før det var gått 48 timer,

mens det trengte kun å gå 24 timer mellom ny økt og øktene med redusert motstand. For å fasiletere til restitusjon, ble det gitt en standardisert drikk med 0.15 g kg^{-1} protein, 11.2 g kg^{-1} karbohydrater og 0.5 g kg^{-1} fett, etter hver treningsøkt. For å sikre bærekraftighet av treningsprotokollen, ble det lagt opp til at noen av øktene kunne gjennomføres uten tilsyn. For å sikre progresjon og etterlevelse, ble deltakerne instruert til å føre detaljerte loggbøker for de øktene som ble utført uten tilsyn, slik at forskerteamet kunne i samarbeid med deltakerne gå gjennom øktene.

6.1.3 Målinger av maksimal muskelstyrke og muskelens tverrsnittsareal

Deltakerne sin maksimale styrke ble satt som deres en repetisjon maksimum (1RM) i unilateral beinpress og knestrekk. Selve testen inneholdt en standardisert oppvarming med 10, 6 og 3 repetisjoner på henholdsvis 50,75 og 85% av deres antatte repetisjon maksimum. Etter dette, ble motstanden gradvis økt helt til deltagerne ikke mestret å løfte vekten gjennom hele bevegelsesbanen, for å finne deres 1RM. Hver deltaker fikk mellom fire og seks forsøk, der den høyeste vekten de mestret å løfte i hver øvelse, ble satt som 1RM. Ved baseline, ble 1RM målt to ganger med minst fire dager mellom hver måling, der den høyeste i hver test ble brukt i videre analyser. I tillegg til nye målinger etter endt intervensjon, gjennomførte en del av deltakerne ($n = 18$) styrkemålinger underveis i studien (i uke 2, 5 og 9). Det skulle ha gått minst 48 timer fra forrige treningsøkt før styrketest. De resterende som ikke utførte tester underveis, ble de ordinære treningsøktene prioritert ved sykdom eller tidsutfordringer som medførte at de gikk glipp av trening eller test.

Ved hjelp av magnetisk resonans bildefremstilling (MR), målte man deltagerens quadriceps (vastus lateralis, medialis, intermedius og rectus femoris) muskeltverrsnittsareal. Dette ble gjort i henhold til produsentens protokoll (S-Scan, Esaote Europe B.V., Maastricht, Nederland), og personen som skulle analysere MR-bildene var blindet ved hjelp av OsiriX (v.5.6, Pixmeo Sarl, Bernex, Sveits). Målingen av muskeltverrsnittet både før og etter intervensjon, ble gjort på samme sted på låret, cirka midt på, med samme avstand fra kneleddet. Resultatet av målingen måtte inneholde minst fire påfølgende bilder med 5 mm tykkelse og 10mm avstand.

6.1.4 Data analyse og statistikk

Med mindre noe annet er spesifisert, er alle deskriptive data presentert som gjennomsnitt og standardavvik. Før studien ble det gjort en forhåndsberegning av utvalgsstørrelse, som viste at 40 deltakere ville være tilstrekkelig for å kunne oppdage forskjeller på omtrent 3 og 5 prosentpoeng for henholdsvis muskeltverrsnittsareal og maksimal styrke mellom de ulike volumforholdene, med en ønsket statistisk styrke på 80%. Denne beregningen er basert på data fra tidligere studier, som antar at forskjellene mellom volumforholdene tilsvarer en effektstørrelse på mellom 0,19 og 0,24 (Ralston et al., 2017; Schoenfeld, 2016).

Tabell 6.2

Modell	Parameter	Estimert Koeffisient	Standard Feil	t-verdi	p-verdi
M1.1: Legext	(Intercept)	43.479	2.836	15.334	0.05*
M1.1: Legext	timepost	33.089	1.539	21.499	0.05*
M1.1: Legext	timepost:setssingle	-4.278	2.188	-1.956	0.0513
M1.2: Legpress	(Intercept)	138.164	12.091	11.427	0.05*
M1.2: Legpress	timepost	95.326	4.158	22.926	0.05*
M1.2: Legpress	timepost:setssingle	-8.541	5.865	-1.456	0.1462
M2: Muskel Tverrsnitt	(Intercept)	7,126.800	254.701	27.981	0.05*
M2: Muskel Tverrsnitt	timepost	289.876	55.046	5.266	0.05*
M2: Muskel Tverrsnitt	timepost:setssingle	-118.167	77.375	-1.527	0.1294
Oppsummering av Modellene for Pre og Post					

For å undersøke hvordan de ulike volumforholdene påvirket muskelvekst og styrke, ble det benyttet lineære blandede modeller (LLM). De relative endringene fra baseline ble satt som den avhengige variabelen, med antall sett som den uavhengige variabelen. For å vurdere om større muskelvekst også gir tilsvarende endringer i muskelstyrke, ble en interaksjonseffekt mellom treningsvolum og endring i muskeltverrsnitt inkludert. Ved å inkludere denne interaksjonseffekten kan modellen fange opp eventuelle forskjeller i effekten av muskelvekst på muskelstyrke under ulike treningsvolum. Baseline-verdier og kjønn ble brukt som kovariater for å kontrollere for deres potensielle effekt på muskelvekst og styrke.

For å evaluere om de statistiske modellene oppfylte forutsetningene for lineære blandede modeller, ble diagnostiske plot undersøkt. Q-Q plot ble brukt for å vurdere normaliteten til residualene, residualer mot predikerte verdier for å sjekke for homoskedastisitet, og histogram av residualer for å identifisere eventuelle skjevheter eller avvik.

Resultater med et signifikansnivå på $\alpha = 0.05$ ble ansett som statistisk signifikante. Analysen av dataene ble gjort i R (R Core Team (2018)).

6.2 Resultater

I Table 6.2, er det en oversikt over de sentrale parametrene for de tre LMM-ene som ble brukt i denne studien.

Tabell 6.3: Deskriptiv data for styrke pre og post

Exercise	Sex	Mean Load (kg)	Standard Deviation (kg)
post - multiple			
Leg Extension	female	71.45	16.32
Leg Extension	male	108.91	19.02
Leg Press	female	231.18	45.04
Leg Press	male	292.94	70.65
post - single			
Leg Extension	female	71.32	14.82
Leg Extension	male	101.25	14.86
Leg Press	female	223.75	48.73
Leg Press	male	280.47	65.58
pre - multiple			
Leg Extension	female	46.75	10.79
Leg Extension	male	67.50	11.76
Leg Press	female	136.00	38.72
Leg Press	male	197.37	61.62
pre - single			
Leg Extension	female	47.50	9.97
Leg Extension	male	67.63	10.98
Leg Press	female	137.50	35.41
Leg Press	male	192.50	58.45

6.3 Endring i muskelstyrke

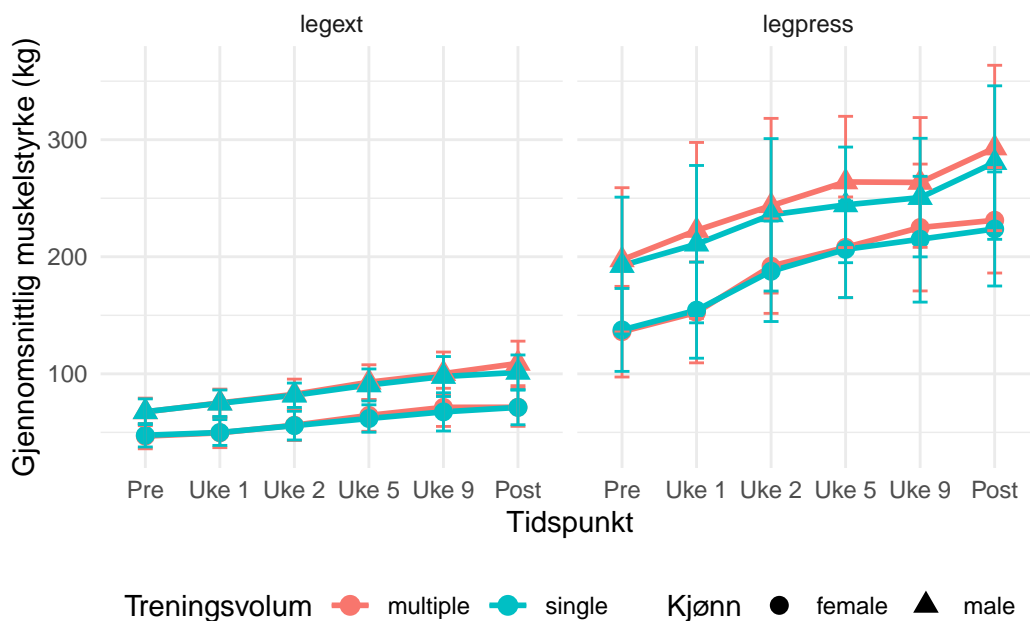
Generelt viser resultatene en økning i muskelstyrke for både leg press og leg extension for alle grupper, og at flere sett gir større økning i muskelstyrke (Table 6.3). For leg extension var den prosentvise økningen i muskelstyrke større for menn enn for kvinner, uavhengig av om de trente med ett eller flere sett. Kvinner som utførte flere sett økte gjennomsnittlig muskelstyrke med cirka 53 %, mens menn som utførte flere sett økte med omtrent 61 %. Tilsvarende var økningen for ett sett 50 % for kvinner og 49 % for menn.

For leg press var den prosentvise økningen også større for menn enn for kvinner, og forskjellene mellom treningsvolumene var tydeligere. Kvinner som trente med flere sett økte gjennomsnittlig muskelstyrke med omtrent 70 %, mens menn økte med omtrent 48 %. For ett sett var den prosentvise økningen for kvinner cirka 63 %, mens menn økte med omtrent 46 %.

Gjennomsnittlig muskelstyrke økte over tid for begge øvelsene (Figure 6.2). Modellresultatene

(lme-modellen) viser at økningen fra baseline (pre) til hver påfølgende tid (økt 1, uke 2, uke 5, uke 9 og post) var statistisk signifikant ($p < 0.05$). Når det gjelder endringene fra pre til post for de to øvelsene, ble det estimert at legpress økte med 33.09 kg (SE = 1.54 kg, 95% konfidensintervall (KI): [30.07, 36.11]). Mens med leg press, ble det estimert at muskelstyrken økte med 95.33 kg (SE 4.16 , 95% konfidensintervall (KI): [87.18, 103.48]).

Når vi ser på om volumforholdene (ett sett vs. flere sett) hadde noen effekt på muskelstyrken, viser konfidensintervallene for leg extension at forskjellen mellom ett sett og flere sett var mellom 30.07 kg og 36.11 kg. Tilsvarende var konfidensintervallet for leg press mellom 87.18 kg og 103.48 kg. Begge disse konfidensintervallene inkluderer null, noe som indikerer at det ikke var noen statistisk signifikant forskjell mellom ett sett og flere sett for endringen i muskelstyrke, verken for leg extension eller leg press. Dette tyder på at begge treningsvolumene hadde lignende effekt på muskelstyrken.



Figur 6.2: Utvikling i muskelstyrke for leg press og leg extension.

6.4 Endring i muskeltversnitt

Resultatene viste en signifikant økning i muskelens tverrsnittsareal etter treningsperioden (Table 6.4). Den estimerte koeffisienten for post-testen var 289.88 g (SE = 55.05), med et 95 % konfidensintervall fra 181.99 til 397.76. Dette indikerer at deltakerne i gjennomsnitt økte muskelens tverrsnittsareal med 289,88 gram etter intervensjonen.

Analysen av effekten av treningsvolum på muskelens tverrsnittareal viste at interaksjonen mellom tid og antall sett ikke var signifikant (estimat = -118.17, SE = 77.37, 95 % KI [-269.82, 33.48], $p = 0.13$). Dette indikerer at det ikke var en statistisk signifikant forskjell ($p > 0.05$) i økningen av muskelens tverrsnittareal mellom deltakerne som utførte ett sett og de som utførte tre sett. Selv om estimatet var negativt, tyder det på at økt treningsvolum ikke nødvendigvis førte til en større økning i muskelens tverrsnittareal i denne studien.

Tabell 6.4: Oppsummering av muskeltverrsnitt før og etter intervensjon (kombinert for kjønn, kun pre og post)

Tidspunkt	Treningsvolum	Gjennomsnittlig muskeltverrsnitt (g)	Standardavvik (SD)
post	multiple	9093.368	1297.290
post	single	8983.975	1219.944
pre	multiple	8835.974	1189.980
pre	single	8845.317	1175.207

6.5 Diskusjon

Denne studien undersøkte effekten av to ulike treningsvolum, ett sett versus tre sett, på muskelstyrke og muskeltverrsnittareal. Resultatene viste en signifikant økning i både muskelstyrke og muskeltverrsnittareal etter treningsperioden for begge gruppene. Imidlertid var det ingen statistisk signifikante forskjeller mellom gruppene som utførte ett sett og de som utførte tre sett, verken for muskelstyrke eller muskelhypertrofi. De diagnostiske analysene bekreftet at modellforutsetningene var rimelig oppfylt, noe som gir tillit til validiteten av funnene våre. Ingen betydelige avvik ble observert i residualene, og modellene ble derfor ansett som passende for dataene.

Disse funnene utfordrer den etablerte oppfatningen om at et høyere treningsvolum alltid fører til større muskeltilpasninger. Tidligere forskning har ofte indikert et dose-respons-forhold mellom treningsvolum og muskelvekst, hvor økt volum er assosiert med større hypertrofi og styrkeøkninger (Currier et al., 2023; Schoenfeld, 2016). For eksempel fant Schoenfeld et al. (2016) at flere sett per øvelse var mer effektive for muskelhypertrofi enn ett sett. På den annen side har noen studier ikke funnet signifikante forskjeller mellom ulike treningsvolum, spesielt når treningsintensiteten er kontrollert (Ralston et al., 2017).

En mulig forklaring på våre funn er at begge treningsvolumene var tilstrekkelige for å fremkalle maksimale adaptasjoner innenfor den gitte treningsperioden. Det kan være en terskelverdi for hvor mye volum som er nødvendig for å stimulere muskelvekst og styrke, og at denne terskelen ble nådd med ett sett i vår studie. Dette er også sett i forskning som antyder at selv lave til moderate treningsvolumer kan være effektive for muskelhypertrofi hos utrente individer (Krieger, 2010; Schoenfeld et al., 2019).

Når det gjelder muskelstyrke, viste begge gruppene betydelige forbedringer uten signifikante forskjeller mellom volumene. Dette kan skyldes at styrkeøkninger i stor grad påvirkes av nevro-muskulære tilpasninger, spesielt i de tidlige stadiene av treningsintervensjoner (Folland & Williams, 2007). Det er mulig at den relativt korte treningsperioden i vår studie ikke var tilstrekkelig for å avdekke forskjeller i styrkeøkninger basert på treningsvolum.

6.6 Begrensninger

Flere begrensninger bør tas i betraktning. For det første var utvalgsstørrelsen begrenset, noe som kan ha redusert studiens statistiske styrke til å oppdage mindre forskjeller mellom gruppene. Selv om forhåndsberegningen av utvalgsstørrelse indikerte tilstrekkelig kraft, kan individuelle variasjoner ha påvirket resultatene. For det andre var treningsperioden relativt kort, og lengre intervensjoner kan være nødvendig for å observere de fulle effektene av ulikt treningsvolum på muskelhypertrofi og styrke. I tillegg ble ikke faktorer som kosthold, søvn og daglige aktiviteter kontrollert, noe som kan ha påvirket muskeltilpasningene.

6.7 Implikasjoner for fremtidig forskning

Fremtidige studier bør vurdere å inkludere større og mer varierte utvalg, samt lengre treningsperioder, for å bedre forstå effekten av treningsvolum på muskeltilpasninger. Det kan også være verdifullt å undersøke hvordan ulike populasjoner, som erfarne idrettsutøvere versus nybegynnere, responderer på forskjellige treningsvolum. Videre bør fremtidig forskning ta hensyn til andre treningsvariabler som intensitet, frekvens og type øvelser for å gi en mer helhetlig forståelse av optimal treningsdesign.

6.8 Konklusjon

Resultatene fra denne studien indikerer at både ett sett og tre sett med styrketrening fører til signifikante økninger i muskelstyrke og muskeltverrsnittareal hos voksne individer. Mangelen på signifikante forskjeller mellom de to treningsvolumene antyder at et lavere volum kan være like effektivt som et høyere volum for å fremme muskelstyrke og hypertrofi i en kortere treningsperiode. Disse funnene har praktiske implikasjoner for treningsprogrammering, spesielt for individer med begrenset tid til trening. Likevel er det behov for ytterligere forskning for å bekrefte disse resultatene og for å utforske de langsiktige effektene av treningsvolum på muskeltilpasninger.

#” Vedlegg: Koder brukt

```

# Last inn dataene
library(exscidata)
library(tidyverse)
library(knitr)
library(kableExtra)
data("strengthvolume"); data("dxadata")

baseline_tbl <- dxadata %>%
  select(participant:time, sex:weight, age, height) %>%
  # Sørg for at disse kolonnene finnes
  filter(time == "pre") %>%
  group_by(sex, include) %>%
  summarise(
    N = n(),
    Alder = sprintf("%.1f (%.1f)", mean(age, na.rm = TRUE),
                        sd(age, na.rm = TRUE)),
    Vekt = sprintf("%.1f (%.1f)", mean(weight, na.rm = TRUE),
                        sd(weight, na.rm = TRUE)),
    Stature = sprintf("%.0f (%.0f)", mean(height, na.rm = TRUE),
                        sd(height, na.rm = TRUE)),
    .groups = "drop" # For å unngå advarsler om gruppering
  ) %>%
  # Konverter alle kolonnene til karakter for å unngå datatypeproblemer i
  # pivot_longer
  mutate(across(everything(), as.character)) %>%
  pivot_longer(cols = c(N, Alder, Vekt, Stature), names_to = "Variable",
               values_to = "Value") %>%
  unite("sex_include", sex, include, sep = "_") %>%
  pivot_wider(names_from = sex_include, values_from = Value)

# Dataframe som representerer treningsfrekvens og RM per uke
data <- data.frame(
  Uke = factor(1:12, levels = c(0, 1:12)), # Spesifiserer nivåene,
  # inkludert 0
  RM = c("10RM", "10RM", "8RM", "8RM*", "8RM", "7RM*", "7RM", "7RM",
         "7RM*", "7RM", "7RM", "7RM"),
  Frekvens = c(2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 2, 1)
)

# Ukene der styrkemålingene ble gjort (x-markeringene)
måling_uker <- data.frame(
  Uke = factor(c(0, 1, 3, 5, 9, 12), levels = c(0, 1:12)), # Inkluderer 0 foran 1

```

```

    Frekvens = 3.1,          # Plassering over søylene
    label = "x"             # Bruk "x" som markør
  )

# Ukene der muskeltverrsnit ble gjort (sirkel-markeringene)
sirkel_uker <- data.frame(
  Uke = factor(c(1, 12), levels = c(0, 1:12)), # Sikrer at 0 er med på skalaen
  Frekvens = 3.0          # Plassering over søylene
)

# Last inn nødvendige pakker
library(exscidata)
library(tidyverse)
library(nlme)
library(ggplot2)
library(gridExtra)
data("dxadata")

# Forbered dataene
muskeltverr <- dxadata %>%
  select(participant:include, lean.left_leg, lean.right_leg) %>%
  pivot_longer(
    names_to = "leg",
    values_to = "lean.mass",
    cols = lean.left_leg:lean.right_leg
  ) %>%
  mutate(
    leg = if_else(leg == "lean.left_leg", "L", "R"),
    sets = if_else(multiple == leg, "multiple", "single"),
    time = factor(time, levels = c("pre", "post"))
  ) %>%
  select(participant, time, sex, include, sets, leg, lean.mass)

# Bygg modellen
model_muskel <- lme(
  fixed = lean.mass ~ time * sets + sex,
  random = ~ 1 | participant,
  data = muskeltverr,
  na.action = na.omit
)

# Undertrykk output av modelloppsummering

```

```

summary_modmuskel <- summary(model_muskel)

# Ekstraher residualer og predikerte verdier
residuals_muskeltverr <- residuals(model_muskel)
fitted_values_muskeltverr <- fitted(model_muskel)

# Lag plottene
# 1. Q-Q plot
qq_plot_muskeltverr <- ggplot(data = data.frame(sample = residuals_muskeltverr),
                              aes(sample = sample)) +

  stat_qq() +
  stat_qq_line(col = "red") +
  ggtitle("Q-Q Plot") +
  theme_minimal()

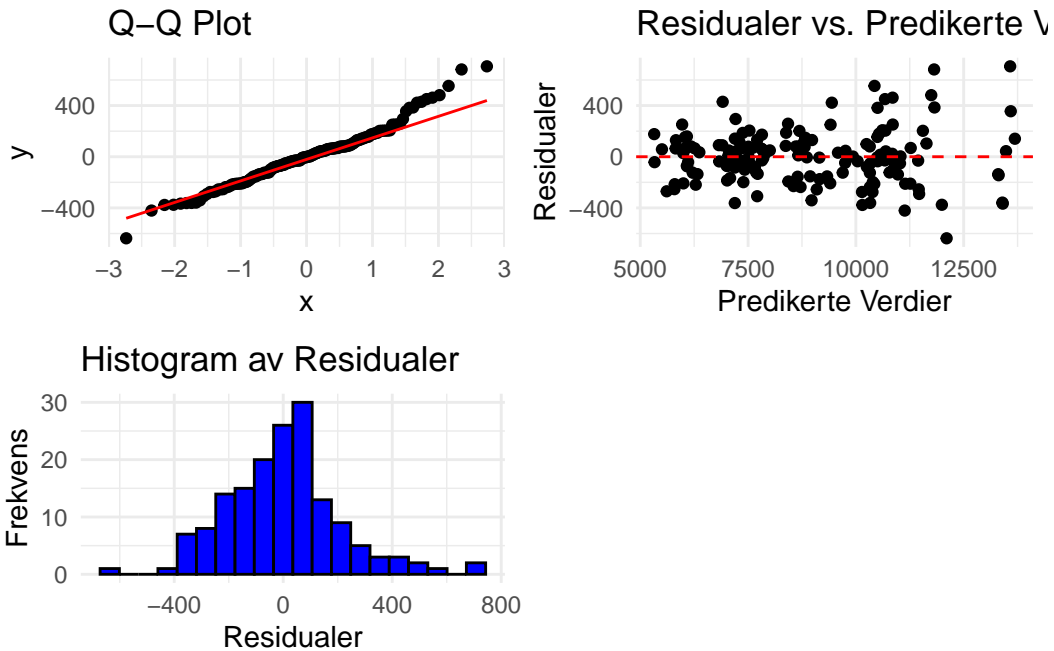
# 2. Residualer vs. Predikerte Verdier
residuals_plot_muskeltverr <- ggplot(data = data.frame(
  fitted = fitted_values_muskeltverr, residuals = residuals_muskeltverr),
  aes(x = fitted, y = residuals)) +

  geom_point() +
  geom_hline(yintercept = 0, linetype = "dashed", color = "red") +
  ggtitle("Residualer vs. Predikerte Verdier") +
  xlab("Predikerte Verdier") +
  ylab("Residualer") +
  theme_minimal()

# 3. Histogram av Residualer
hist_plot_muskeltverr <- ggplot(data = data.frame(
  residuals = residuals_muskeltverr), aes(x = residuals)) +
  geom_histogram(bins = 20, fill = "blue", color = "black") +
  ggtitle("Histogram av Residualer") +
  xlab("Residualer") +
  ylab("Frekvens") +
  theme_minimal()

# Kombiner plottene
grid.arrange(qq_plot_muskeltverr, residuals_plot_muskeltverr,
             hist_plot_muskeltverr, nrow = 2)

```



```
# Samme fremgangsmåte ble brukt for både legpress og legext for styrke.

# Filtrer datasettet for leg extension-øvelsen og sett 'time' som
# en faktor med ønsket rekkefølge
strengthvolume_legext <- strengthvolume %>%
  filter(exercise == "legext") %>%
  mutate(time = factor(time, levels = c("pre", "session1", "week2", "week5",
    "week9", "post")))

# Modell for muskelstyrke i leg extension-øvelsen uten manglende verdier
model_legext <- lme(
  fixed = load ~ time * sets + sex,
  random = ~ 1 | participant,
  data = strengthvolume_legext, # Bruker det oppdaterte datasettet
  na.action = na.omit
)

# Oppsummer modellresultatene
summary_modellegext <- summary(model_legext)

# Hent ut estimatet og standardfeilen for 'post'
```

```

post_koeff_ext <- summary_modell_legext$tTable["timepost", "Value"]
post_se_ext <- summary_modell_legext$tTable["timepost", "Std.Error"]

# Beregn konfidensintervallet for post-estimatet
alpha <- 0.05
z_value <- qnorm(1 - alpha / 2)

lower_bound_ext <- post_koeff_ext - z_value * post_se_ext
upper_bound_ext <- post_koeff_ext + z_value * post_se_ext

konfidensintervall_post_ext <- c(lower_bound_ext, upper_bound_ext)

# Hent ut estimatet og standardfeilen for interaksjonen mellom 'timepost'
# og 'setssingle'
interaksjon_koeff_ext <- summary_modell_legext$tTable["timepost:setssingle",
  "Value"]
interaksjon_se_ext <- summary_modell_legext$tTable["timepost:setssingle",
  "Std.Error"]

# Beregn konfidensintervallet for interaksjonsestimatet
lower_bound_interaksjon_ext <- interaksjon_koeff_ext - z_value *
  interaksjon_se_ext
upper_bound_interaksjon_ext <- interaksjon_koeff_ext + z_value *
  interaksjon_se_ext

konfidensintervall_interaksjon_ext <- c(lower_bound_interaksjon_ext,
  upper_bound_interaksjon_ext)

# Samme fremgangsmåte ble brukt for både legpress og legext for styrke.

# Ekstraher oppsummeringer for hver modell
tidy_legext <- tidy(model_legext)
tidy_legpress <- tidy(model_legpress)
tidy_muskel <- tidy(model_muskel)

# Legg til en kolonne som angir modellnavnet for enkel referanse
tidy_legext$model <- "Model 1.1: Leg Extension"
tidy_legpress$model <- "Model 1.2: Leg Press"
tidy_muskel$model <- "Model 2: Muskel Tverrsnitt"

```



```

# Kombiner alle dataene i én dataframe
tidy_models <- rbind(tidy_legext, tidy_legpress, tidy_muskel)

# Hvis nødvendig, tilordne unike termer til en variabel uten å skrive ut
unique_terms <- unique(tidy_models$term)

# Filtrer ut kun de radene som er relevante
tidy_models_filtered <- tidy_models %>%
  filter(term %in% c("(Intercept)", "timepost", "timepost:setssingle"))

# Velg ut de viktigste kolonnene og legg til formatering for p-verdier med stjerner
tidy_models_summary <- tidy_models_filtered %>%
  dplyr::select(model, term, estimate, std.error, statistic, p.value) %>%
  mutate(
    estimate = round(estimate, 3),
    std.error = round(std.error, 3),
    statistic = round(statistic, 3),
    p.value = ifelse(p.value < 0.05, "0.05*", format.pval(p.value, digits = 3,
                                                         nsmall = 3))
  )

# Filtrer datasettet for leg press-øvelsen
strengthvolume_legpress <- strengthvolume %>%
  filter(exercise == "legpress")

# Filtrer datasettet for leg extension-øvelsen
strengthvolume_legext <- strengthvolume %>%
  filter(exercise == "legext")

# Filtrer for pre og post, og beregn gjennomsnitt og standardavvik for `legext`
deskriptiv_legext <- strengthvolume_legext %>%
  filter(time %in% c("pre", "post")) %>%
  group_by(time, sets, sex) %>%
  summarise(
    mean_load = mean(load, na.rm = TRUE),
    sd_load = sd(load, na.rm = TRUE),
    n = n()
  ) %>%
  mutate(exercise = "Leg Extension")

# Filtrer for pre og post, og beregn gjennomsnitt og standardavvik for `legpress`
deskriptiv_legpress <- strengthvolume_legpress %>%

```

```

filter(time %in% c("pre", "post")) %>%
group_by(time, sets, sex) %>%
summarise(
  mean_load = mean(load, na.rm = TRUE),
  sd_load = sd(load, na.rm = TRUE),
  n = n()
) %>%
mutate(exercise = "Leg Press")

# Kombiner begge tabellene til én
deskriptiv_samlet <- bind_rows(deskriptiv_legext, deskriptiv_legpress)

# Formater tabellen for å inkludere deskriptiv statistikk for pre og post
tabell_deskriptiv <- deskriptiv_samlet %>%
  dplyr::select(exercise, time, sets, sex, mean_load, sd_load) %>%
  rename(
    Exercise = exercise,
    Time = time,
    Sets = sets,
    Sex = sex,
    Mean_Load = mean_load,
    SD_Load = sd_load
  )

# Utrekning av gjennomsnitt og SD for muskelstyrke

library(dplyr)

strength_mean.sd <- strengthvolume %>%
  filter(exercise %in% c("legpress", "legext")) %>%
  group_by(time, sets, sex, exercise) %>%
  summarise(mean_strength = mean(load, na.rm = TRUE),
            sd_strength = sd(load, na.rm = TRUE)) %>%
  ungroup()

# Diagram av utviklingen i muskelstyrke

library(ggplot2)

## Endre navn på tidspunktene

strength <- strength_mean.sd %>%

```

```

mutate(time = recode(time,
                      "pre" = "Pre",
                      "session1" = "Uke 1",
                      "week2" = "Uke 2",
                      "week5" = "Uke 5",
                      "week9" = "Uke 9",
                      "post" = "Post"),
       time = factor(time, levels = c("Pre", "Uke 1", "Uke 2", "Uke 5",
                                      "Uke 9", "Post")))

# Utregning av gjennomsnitt og SD for muskeltverrsnitt
library(dplyr)
library(knitr)
library(ggplot2)

# Dataprosessering - omstrukturering og kombinerer av data for høyre og
# venstre ben
muskeltverr <- dxadata %>%
  select(participant:include, lean.left_leg, lean.right_leg) %>%
  pivot_longer(
    cols = lean.left_leg:lean.right_leg,
    names_to = "leg",
    values_to = "lean.mass"
  ) %>%
  mutate(
    leg = if_else(leg == "lean.left_leg", "L", "R"),
    sets = if_else(multiple == leg, "multiple", "single")
  ) %>%
  select(participant, time, sex, include, sets, leg, lean.mass)

# Oppsummering av gjennomsnitt og standardavvik etter tid, treningsvolum og
# kjønn
muskeltverr_summary <- muskeltverr %>%
  group_by(time, sets, sex) %>%
  summarise(
    mean_lean_mass = mean(lean.mass, na.rm = TRUE),
    sd_lean_mass = sd(lean.mass, na.rm = TRUE),
    .groups = "drop" # Fjerner grupperingen automatisk etter summarise
  )

# Filtrer til pre og post tidspunkter og kombiner kjønn for en enklere oversikt

```

```

muskeltverr_summary_pre_post <- muskeltverr_summary %>%
  filter(time %in% c("pre", "post")) %>%
  group_by(time, sets) %>%
  summarise(
    mean_lean_mass = mean(mean_lean_mass, na.rm = TRUE),
    sd_lean_mass = sqrt(mean(sd_lean_mass^2, na.rm = TRUE)), # Kombinert SD
    # ved å bruke kvadratet av SD
    .groups = "drop"
  ) %>%
  arrange(time, sets) # Sortér etter tid og treningsvolum

```

7 Arbeidskrav i vitenskapsfilosofi

7.1 Spørsmål 1

“Ifølge Hume er det umulig å rasjonelt begrunne bruken av induksjon. Hva er argumentet for denne konklusjonen? Gi en innvending mot ett av premissene i Humes argument og prøv å svare på denne innvendingen på Humes vegne.”

7.1.1 Innledning

David Hume var en skotsk empirist, som gjennom sitt syn og tanker har hatt stor innflytelse på moderne vitenskapsfilosofi (Norton & Taylor, 2008, kap 1). Selv om Hume sine meninger og tanker har vekket en rekke spørsmål og debatter fra hans levetid til nå, er det nok hans skepsis og analyse av årsak og induksjon som har fått mest oppmerksomhet (Norton & Taylor, 2008, pp. s.214–221).

7.1.2 Humes argument mot induksjon

Induksjon er en vitenskapelig metode som forsøker å trekke generelle lover eller prinsipper basert på enkelte observasjoner (Tranøy, 2021), for eksempel: “Jeg brand meg når eg tok på flammene, derfor er flammer varme.”. Hume mente at det ikke er mulig å fastslå årsaksforholdet som vi erfarer gjennom sansene våre, som eksempelet med flammene over, men på grunn av vår gjentatte erfaring med at det er en konstant sammenheng mellom to typer hendinger, så antar vi at det ene er årsaka til den andre (Hume, 2019).

En forutsetning for at konklusjonen i mitt induktive argument skal holde, er at naturens gang ikke endrer seg, altså at framtida vil likne på fortida. Denne førestillinga, kjent som uniformitetsprinsippet, ligger til grunn for alle induktive slutninger. Hume hevdet at dette prinsippet ikke kan bevises uten å ende i en sirkulær argumentasjon, fordi enhver rettferdiggjøring av induksjon forutsetter det samme prinsippet man prøver å bevise – at framtida vil følge samme mønster som fortida (Hume, 2019, pp. s.38–39; Norton & Taylor, 2008, p. s.215). Når jeg konkluderer med at jeg vil brenne meg ved berøring av flammer, bygger jeg på denne førestillinga, men ifølge Hume kan jeg ikke bevise den uten å allerede forutsette at framtida vil være lik fortida. Dette utgjør kjernen i Humes kritikk av induksjon: vi kan ikke rasjonelt grunngi bruken av induksjon uten å bygge på den samme førestillinga vi forsøker å rettferdiggjøre.

7.1.3 Innvending mot Hume sitt argument

Selv om Hume påpekte at vi ikke kan observere et nødvendige årsaksforhold mellom to hendelser, bare at en hending følger en annen, kan det innvendast mot dette premisset at naturvitenskapen har vist at induktive slutninger fungerer pålitelig i praksis. For eksempel kan Newtons gravitasjonslov brukast til å forutsi at en gjenstand vil falle mot jorda når den har blitt sleppt. Selv om det kan være vanskelig å bevise nødvendige årsaksforhold direkte, kan en hevde at den praktiske suksessen til naturvitenskaplege lover, som Newtons gravitasjonslov, rettferdiggjør bruken av induksjon. Dette kommer tydelig til uttrykk i vår evne til nøyaktig å forutsi framtidige hendelser, som planetbanar eller gjenstander sin bevegelse på jorda, basert på tidligere erfaringer.

7.1.4 Humes respons på innvendinga

Hume ville sannsynligvis ha svart på innvendinga ved å fasthalde sin skepsis til at induksjon er basert på rasjonell fornuft. I steden ville han ha påpeikt at våre induktive slutningar kjem frå en naturlig tilbøyelegheit til å forvente at framtida vil likne fortida, basert på vaner og førestellingsevne (Hume, 2019, pp. s.48–49). Denne tilbøyelegheita til å trekke slutninger om framtida er det Hume kalte for customs, altså vaner (Hume, 2019, p. s.44). Selv om Hume ville ha opprettholdt sin påstand om at induksjon ikke er rasjonelt grunnlagt, ville han også ha erkjent, i likskap med min innvending, at induksjon er avgjørende for vår praktiske suksess. Hume mente at disse induktive slutningene, som bygger på vaner og erfaringer, er mer effektive for vår overleving og praktiske handlinger enn om vi skulle ha basert alle slutninger på ren rasjonalitet (Hume, 2019, pp. s.57–58).

7.2 Spørsmål 2

Gi en kort beskrivelse av falsifikasjonisme og si litt om hvorfor Popper var motivert til å utvikle denne teorien. Presenter så ett problem med teorien og vurder hvorvidt problemet kan løses.

7.2.1 Innledning

Demarksjonsproblemet er et grunnleggende tema innen vitenskapsfilosofi, og det handler om hvordan vi kan skille mellom vitenskap og ikke-vitenskap. Dette skilet er avgjørende for vår forståing av hva som reiknast som gyldig kunnskap, og har implikasjoner for områder som metafysikk og religion (Popper, 1959, pp. s.34–39). Mange filosofer har forsøkt å definere klare kriterier for denne avgrensinga, og en de mest sentrale i denne debatten er Karl Popper.

7.2.2 Beskrivelse av falsifikasjonisme og Poppers motivasjon

Popper var en markant kritiker av induksjon. Gjennom sin avvisning av induksjonsprinsippet utfordra han ikke bare empirisk vitenskap sin grunnleggende metode, men også det tradisjonelle skillet mellom vitenskap og metafysisk spekulasjon (Popper, 1959, p. s.34). For Popper var en vitenskapelig teori ikke gyldig fordi den kunne verifisert gjennom observasjonar, men fordi den kunne falsifiseres gjennom testing (Popper, 1959, p. s.40). Han hevdet at vitenskapens rolle er ikke å verifisere en teori basert på enkeltobservasjonar eller erfaring, men å utvikle teorier som kunne falsifiseres gjennom observasjoner eller eksperimenter (Popper, 1959, pp. s.41–42, 47–48).

Et eksempel på en vitenskapelig teori er påstanden: “Newton sin lov om gravitasjon seie at stenen eg slipp vil falle mot bakken på grunn av tiltrekkinga mellom to masser.” Denne påstanden er vitenskapelig fordi den kan testes og potensielt falsifiseres. Derimot vil påstanden “Stenen faller fordi universet bestemmer det.”, ikke kunne falsifiseres og er dermed ikke vitenskapelig

Popper var motivert til å utvikle falsifikasjonismen på grunn av de svakhetene han såg i induktiv logikk. Han mente at ideen om at generaliseringar basert på observasjonar kunne gi sikker kunnskap, var logisk problematisk. Som et svar på dette utvikla han falsifikasjonsprinsippet, som metode for å sikre at teorier kunne utfordrast og testast på en måte som kunne avsløre deres potensielle muligheter for feil.

7.2.3 Utfordring med falsifikasjonisme

Selv om Popper sin falsifikasjonisme gir en streng vitenskapleg metode, gjenspeiler den ikke alltid hvordan vitenskapen faktisk utvikler seg. Ifølge Popper bør en teori forkastet når den møter motstriden data, ettersom den då er falsifisert og dermed “bevist” å gjøre feil. Popper var kritisk til bruken av tilpasninger og hjelpehypotesar som bare ble introdusert for å redde teorier fra falsifikasjon, fordi dette kunne redusere falsifiserbarheita til teorien og dermed gjere den uvitenskapleg (Popper, 1959, p. s.145).

I praksis er det sjelden forskarar forkaster en teori ved første tegn til motstridene data. Ofte vel de å gjøre tilpasninger eller forbedrer teoriane med hjelp av hjelpehypotesar. Dette skaper derfor et misforhold mellom Popper sitt ideal om falsifikasjon og vitenskapleg praksis, der teorier ofte justeres i staden for å bli forkasta umiddelbart. Et eksempel på dette er Newton sine lover om bevegelse, som ble revidert flere ganger gjennom historia. Selv om Newtons lover forklarte bevegelse i de fleste praktiske situasjoner med høy presisjon, viste de seg utilstrekkelige for å beskrive bevegelse i ekstreme forhold som ved svært høge hastigheter. I staden for å forkaste lovene, ble de utvida og korrigert av Albert Einstein for å forklare avvika. Dette er et eksempel på hvordan vitenskapen i praksis ofte justerer teorier framfor å forkaste de umiddelbart.

7.2.4 Poppers løsning

Popper erkjenner denne utfordringa, men hevder at ho kan løysast ved å følge strenge metodologiske regler (Popper, 1959, pp. s.81–84, 145). Han argumenterer for at tilpasninger som bare løyser et enkelt problem i en teori, såkalla ad hoc-hypotesar, bør unngå, da dei svekker teoriens vitenskaplege verdi. Modifikasjoner som fører til nye, testbare prediksjoner er derimot akseptable, ettersom de styrker teorien ved å gjøre den mer falsifiserbar (Popper, 1959, p. s.83). Videre fremhever Popper prinsippet om sparsommelighet, kjent som Ockhams barberkniv, som tilsier at man bør unngå unødvendig antakingar for å forhindre at teorier blir unødige komplekse (Store norske leksikon, 2005--2007).

Selv om disse reglene gir en metodologiske løsning på problemet, kan de være betre egna som ideal enn som praktiske verktøy i vitenskapelig praksis. Vitenskapen i praksis er mer fleksibel og kompleks enn det Popper førestilte seg, og det er ikke like enkelt å vite når man bør forkaste en teori og når man burde revidere den.

7.2.5 Konklusjon

Selv om falsifikasjonisme i teorien seie at en teori skal forkastes ved motstridende data, viser vitenskapelig praksis at teorier ofte modifierast og tilpasset i staden. Dette reflekterer vitenskapens natur som en kontinuerlig prosess med prøving, feil, og forbedring, og at falsifikasjon i praksis viser seg å være en mer gradvis og fleksibel prosess enn det Poppers opprinnelege modell antydgar.

8 Protein ekstraksjon og analyse

I jakten på å oppnå en dypere forståelse av de molekylære reaksjonene som regulerer transkripsjon og translasjon med trening, ernæring, sykdom og aldring, har bruken av Western blot (WB) i treningsfysiologi økt. WB er en sentral metode for å undersøke regulatoriske molekylære reaksjoner, blant annet ved å kvantifisere proteinmengder og analysere protein-protein-interaksjoner. Dette er avgjørende for å kunne dokumentere fysiologiske adaptasjoner som oppstår som respons til trening (Bass et al., 2017).

Western blot har tradisjonelt vært assosiert med overføring av proteiner fra en gel til en stabil membran, men idag referer metoden til hele prosessen, fra proteinekstraksjon til kvantitativ analyse (Bass et al., 2017, p. s.2). Prosessen inkluderer ekstraksjon av proteiner, separasjon ved elektroforese, overføring til en membran, blokkering for å hindre uspesifikke bindinger, og påføring av antistoffer for å identifisere proteiner. Til slutt kvantifiseres proteinmengdene ved hjelp av densitometrisk programvare (Bass et al., 2017).

p70S6K er en nøkkelkomponent i mTORC1-signalveien, som spiller en sentral rolle i reguleringen av proteinsyntese og muskelvekst (D'Souza & al., 2014, p. s.7). Singalveien aktiveres som respons på mekanisk belastning, og er avgjørende for muskulære tilpasninger til trening (Atherton & Smith, 2012). Fosforlyering av p70S6K er en indikator på aktivering av mTORC1, som igjen reflekterer cellulære prosesser som fremmer anabole responser i muskulaturen (D'Souza & al., 2014; Holz & al., 2005, pp. s.577–579). Ved å måle både totalmengden av p70S6K og den fosforlyerte formen kan man få innsikt i balansen mellom inaktive og aktive proteiner, noe som gir et mer helhetlig bilde av signalveiens funksjon (Bass et al., 2017). Derfor er p70S6K et ideelt mål for å studere molekylære tilpasninger til trening.

Formålet med dette eksperimentet er å bruke og demonstrere Western Blot som en metode for å analysere proteinuttrykk. Dette gjøres ved å sammenligne totalmengden av p70S6K (t-p70) og mengden fosforlyert p70S6K i muskelvev fra et ben som har gjennomført en styrketreningsøkt, med et ben som ikke har trent, over tid etter treningen. Ved å sammenligne det trenede og utrenede ønsker man å evaluere hvordan trening påvirker aktiveringen av mTORC1-signalveien, og samtidig demonstrere hvordan WB kan brukes til å undersøke slike molekylære mekanismer.

8.1 Metode

Muskelprøvene som ble brukt i dette eksperimentet, ble hentet fra m. vastus lateralis på høyre og venstre lår hos testpersoner. Høyre ben gjennomførte styrketrening, mens venstre ben fungerte som kontroll. Prøvene ble samlet inn på ulike tidspunkter etter trening for å vurdere tidsavhengige endringer: én prøve ble tatt 10 minutter etter trening (HE), en annen 30 minutter etter (LØ), og den siste 60 minutter etter (T). Før Western blot (WB)-analysen ble muskelprøvene homogenisert for å frigjøre proteiner. Fremgangsmåten for homogenisering og WB er detaljert beskrevet i den følgende delen. Kodene som er blitt brukt for å lage tabeller ligger vedlagt nederst under overskriften “Koder”, som ligger etter konklusjonen.

8.1.1 Homogenisering av muskelprøver

Prøvene ble veid før homogenisering, og vekten for testperson HE var henholdsvis 14,0 g for høyre ben (HE-R) og 14,5 g for venstre ben (HE-L). Til homogeniseringen ble 600 μ l lysisbuffer tilsatt, sammensatt av 594 μ l T-PER (Thermo Fisher Scientific) og 6 μ l Pink Reagent for å indikere homogeniseringens fullstendighet. En plastpistill ble brukt til å mekanisk bryte ned vevet til en homogen løsning.

Homogenisering ble utført ved hjelp av en plastpistill for å bryte ned vevet og frigjøre proteinene. Til dette brukte vi 600 μ l lysisbuffer, bestående av 594 μ l T-PER og 6 μ l Pink Reagent for å sikre at vevet ble fullstendig homogenisert. Den homogeniserte prøven ble deretter behandlet for å måle proteinkonsentrasjonen ved hjelp av Bradford Assay-metoden.

Prøvene ble holdt på is under behandlingen og rotet i 33 minutter for å sikre tilstrekkelig homogenisering. Etter homogeniseringen ble prøvene sentrifugert ved 10 000 g i 10 minutter ved 4 °C for å separere supernatanten fra pelleten. Supernatanten ble deretter forsiktig overført til et nytt rør for å unngå forstyrrelse av pelleten. Denne prosessen gikk uten problemer og med minimal sannsynlighet for feilkilder.

For å måle proteinkonsentrasjonen ble det utført en fortynning av prøvene. En fortynning på 1:6 ble brukt for å bringe homogenatet innenfor standardkurvens område. Dette innebar at 12 μ l prøve ble blandet med 60 μ l destillert vann. Pipetteringen ble utført nøyaktig og uten merkbare feil, selv om mulige feilkilder alltid må vurderes.

Bradford Assay-metoden benytter fargestoffet Coomassie Brilliant Blue G-250, som binder seg til proteiner og gir en fargeendring som er proporsjonal med proteinkonsentrasjonen (Noble & Bailey, 2008). For å kvantifisere proteinkonsentrasjonen laget vi en standardkurve basert på bovint serumalbumin (BSA), som fungerer som en referanse. Absorbansen til prøvene ble målt ved 595 nm ved hjelp av et spektrofotometer. Dette gjorde det mulig å fastslå om prøvene inneholdt lave eller høye konsentrasjoner av proteiner, med grunnlag i sammenligning med standardkurven (Noble & Bailey, 2008, kap 8).

8.1.2 Western blot protokoll

Selv om gruppen jobbet med en prøve, er WB-analysen blitt gjort på flere prøver fra 3 ulike testpersoner. For elektroforesen brukte vi Bio-Rad Criterion og en ferdiglaget gel. Selv om det ikke ble brukt en kalibreringskurve for å kvantifisere absolutt mengde p70S6K, ble intensitetsverdiene normalisert mot total proteinmengde og deretter brukt som en indikator for proteinaktivitet. Fosforyleringsnivået (p-p70/t-p70p) ble beregnet for å gi et relativt mål på aktivisering av mTORC1-signalveien.

De homogeniserte prøvene ble fortynnet til en endelig konsentrasjon på 1.5-2 g/l i Laemmli-bufer (Bio-Rad). Prøvene ble deretter denaturert ved oppvarming til 95 °C i 5 minutter før de ble avkjølt til romtemperatur. Proteinstandarder (Precision Plus Protein Dual Color Standards, fra Bio-Rad) ble brukt for å sikre nøyaktig beregning av proteinstørrelse. Prøvene ble lastet på en ferdiglaget Bio-Rad Criterion-gel, og elektroforese ble utført ved konstant spenning (300 V) i 30 minutter ved 4 °C, med bruk av Bio-Rad Running Buffer.

En potensiell feil ble identifisert under prøvelastingen i brønn 5, hvor en prøve utilsiktet ble trukket tilbake i pipetten. Selv om alle væske ble overført til brønnen, kan dette ha resultert i blanding. Oversikt over de ulike brønnene er vist i Table 8.1, der brønn 2-6 er p-p70 og brønn 9-13 er t-p70.

Etter elektroforese ble prøvene først lagt i metanol i 2-10 minutter, for å gjøre det mulig for proteinene å binde seg mer effektivt. Deretter ble de overført til en PVDF-membran ved hjelp av en "wet blotting"-metode, hvor proteiner ble drevet ut av gelen og overført til membranen under påvirkning av en elektrisk strøm i en bufferløsning (Bass et al., 2017, pp. s.10–12). Selve overføringen skulle etter protokollen bli utført ved konstant strøm (300 mA) i 3 timer ved romtemperatur, med en buffer bestående av 25 mM Tris, 192 mM glycin og 20% metanol, hvor metanol hjalp til med å stabilisere proteinene. Grunnet mangelen på tid, ble det heller kjørt et hurtigprogram (100 V i 30 minutter). Dette kan ha redusert effektiviteten for større proteiner. Membranet ble deretter farget med MemCode Reversible Protein Stain for å visualisere om proteinene var blitt overført og jevnt fordelt. Fargingen ga også mulighet til å visualisere total proteinmengde og justere for variasjoner mellom prøvene i den neste analysen.

Membranen ble blokkert i en 5 % melkebuffer (TBS-T) i 1 time ved romtemperatur for å redusere ikke-spesifikk binding. Primærantistoffet mot fosforylert p70S6K (p-p70S6K) ble inkubert over natten ved 4 °C, og deretter vasket grundig med TBS-T for å fjerne overskuddsantistoff. Sekundærantistoffet ble tilsatt i 2.5 % melkebuffer og inkubert i 1 time ved romtemperatur. Ufullstendig vask under dette trinnet kan ha ført til bakgrunnsstøy, da vaskingen ble utført av en annen person.

For å få frem signalene ble det benyttet ECL (Enhanced Chemiluminescence) og bildeopptak på en Gel Doc-enhet (Bass et al., 2017, pp. s.18–19). Kvantifisering av Western blot-resultatene ble utført ved hjelp av densitometrisk analyse i programmet ImageJ. Den høyeste intensiteten av proteinbåndene for p-p70S6K og t-p70S6K ble målt og normalisert mot total proteinmengde for å korrigere for variasjoner i prøvene (Bass et al., 2017, pp. s.20–24).

Tabell 8.1: Distribusjon av prøvene.

Western Blot Brønndistribusjon

Brønn	Innhold
1	M
2	HE-L
3	HE-R
4	T-L
5	T-R
6	LØ-R
7	M
8	M
9	HE-L
10	HE-R
11	T-L
12	T-R
13	LØ-R
14	M

Forklaringer: M = Markør, HE-L = Venstre ben (10 min), HE-R = Høyre ben (10 min), T-L = Venstre ben (60 min), T-R = Høyre ben (60 min), LØ-R = Høyre ben (30 min)

Tabell 8.2: Intensitetsverdiene for p-p70S6K og t-p70S6K fra densiometer analysen i ImageJ.

Western Blot analyse av fosforyleringsnivå

Variabel	HE_Venstre	HE_Høgre	T_Venstre	T_Høgre	LØ_Høgre
Membran blot (p-p70)	16 530.47	10 135.43	11 187.97	7 012.39	14 791.04
Membran blot (t-p70)	15 516.16	13 009.67	12 917.48	10 868.74	13 617.43
p-p70 (raw)	4 276.45	4 238.28	4 457.28	4 148.45	4 249.86
t-p70 (raw)	4 276.45	4 238.28	4 457.28	4 148.45	4 249.86
p-p70 (norm.)	0.26	0.42	0.40	0.59	0.29
t-p70 (norm.)	0.28	0.33	0.34	0.38	0.31
Fosforyleringsnivå	0.94	1.28	1.16	1.55	0.92

Merk: Tallene er råverdier uten en spesifikk enhet.

Forholdet mellom p-p70S6K og t-p70S6K ble beregnet som en indikator på aktiveringen av mTORC1-signalveien.

8.2 Resultater

Vi benyttet oss av WB-analyse for å kvantifisere total p70S6K og fosforlyert p70S6K, slik at vi kunne undersøke fosforyleringsnivået til proteinet i vastus lateralis til de ulike testpersonene. Intensiteten til fosforylert p70S6K (p-p70) og total p70S6K (t-p70) ble normalisert mot totalprotein fra membran blottet. Fosforyleringsnivået ble beregnet som forholdet mellom p-p70 og t-p70 ($p\text{-p70} / t\text{-p70}$). Intensitetsverdiene som presenteres i tabellene er basert på optisk tetthet fra Western blot-bildene. Verdiene representerer relative signalstyrker og har ingen absolutt måleenhet, da kalibreringskurve ikke ble benyttet. For å korrigere for variasjoner i proteinmengde mellom brønnene, f.eks grunnet pipeteringsfeil eller feil i total protein-målinger, ble intensitetene normalisert mot total proteinmengde (membran blot).

I Table 8.2 vises de normaliserte intensitetsverdiene for fosforylert (p-p70) og total (t-p70) p70S6K for høyre og venstre bein hos hver testperson, basert på membran blottet. Fosforyleringsnivået er beregnet som forholdet mellom normalisert p-p70 og t-p70. Fosforyleringsnivået ($p\text{-p70}/t\text{-p70}$) varierte mellom prøvene, med høyest nivå i trent bein hos testperson T (1.550) og lavest i utrent bein hos testperson LØ (0.920).

Den gjennomsnittlige aktiviteten til p70S6K i hele gruppen var 1.17 ± 0.25 . Det trente beinet hadde et høyere fosforyleringsnivå (1.43 ± 0.15) enn det utrente beinet (0.99 ± 0.09). Disse funnene indikerer at mekanisk belastning fra treningen førte til økt aktivering av mTORC1-signalveien.

Tabell 8.3: Deskriptiv data fra densimetri analysen.

Tabell 2: Gjennomsnittlige nivåer av total og fosforylert p70S6K

Nivåene er presentert som gjennomsnitt \pm standardavvik (SD)

Parameter	Hele gruppen (n = 20)	Høyre bein (n = 10)	Venstre bein (n = 10)
t-p70 (norm.)	0.33 ± 0.04	0.35 ± 0.04	0.31 ± 0.03
p-p70 (norm.)	0.39 ± 0.12	0.50 ± 0.12	0.31 ± 0.07
Fosforyleringsnivå	1.17 ± 0.25	1.43 ± 0.15	0.99 ± 0.09

Total p70S6K (t-p70) viste en liten forskjell mellom de to gruppene. Gjennomsnittet for hele gruppen var 0.33 ± 0.04 , mens det trenete beinet hadde et noe høyere nivå (0.35 ± 0.04) enn det utrente beinet (0.31 ± 0.03).

For fosforylert p70S6K (p-p70) var forskjellene mer markante. Gjennomsnittet for hele gruppen var 0.39 ± 0.12 , mens nivået i det trenete beinet (0.50 ± 0.12) var betydelig høyere enn i det utrente beinet (0.31 ± 0.07). Disse resultatene understøtter observasjonen om at trening spesifikt øker fosforyleringen av p70S6K og aktiverer mTORC1-signalveien.

8.3 Diskusjon

Intensitetsverdiene gir oss et bilde av p70S6K-aktiviteten i prøvene. Selv om absolutt kvantifisering ikke ble utført, indikerer et høyere fosforyleringsnivå en økt aktivitet av mTORC1-signalveien, som spiller en sentral rolle for proteinsyntese (Atherton & Smith, 2012). Resultatene viser et høyere fosforyleringsnivå av p70S6K i det trenete beinet sammenlignet med det utrente, noe som understreker treningens evne til å stimulere molekylære mekanismer som fremmer muskulær tilpasning. Aktiveringen ser ut til å være tidsavhengig, med høyest nivå kort tid etter trening og en gradvis reduksjon over tid, i tråd med tidligere forskning (Atherton & Smith, 2012; D'Souza & al., 2014).

Variasjonen i fosforyleringsnivåene kan tilskrives både biologiske og tekniske faktorer. Forskjellene mellom høyre og venstre bein var mer konsistente enn variasjonen på gruppenivå, noe som tyder på en spesifikk effekt av trening. En viktig begrensning er mangelen på en kalibreringskurve, som gjør det vanskelig å kvantifisere proteinmengden absolutt. Dette begrenser muligheten til å sammenligne resultatene direkte med tidligere studier, men demonstrerer samtidig hvordan Western blot kan brukes for å undersøke slike mekanismer.

Ved å belyse aktiveringen av mTORC1-signalveien, kan man få en dypere forståelse av molekylære tilpasninger til trening.

Tabell 8.4: Bufferoppløsninger

Oversikt over oppløsninger

Tabellen viser bufferløsninger med komponenter, konsentrasjoner og masser.

Bufferoppløsning	Komponent	Konsentrasjon (mM)	Masse (g)
TBS 1L 10x, pH 7.6 (HCl adjust)	Tris Base NaCl	200 mM 1.37 M	24.22 g 80.06
Running buffer 1L 10x, pH 8.3	Tris Base Glycin SDS	250 mM 1.92 M 35 mM	30.28 g 144.1 g 1
Transfer buffer 1L 10x	Tris Base Glycin	250 mM 1.92 M	30.28 g 144.1
1 x Transfer buffer = 100 ml TB stock x 10 + 100 ml methanol + 800 ml dH ₂ O			

8.4 Konklusjon

Resultatene indikerer en asymmetrisk aktivering av mTORC1-signalveien mellom høyre og venstre bein, noe som kan forklares ved at kun det ene beinet ble trent før biopsien. Styrketrening førte til en akutt økning i fosforyleringsnivået av p70S6K, med størst aktivering observert kort tid etter trening. Dette understreker treningens evne til å stimulere molekytlære mekanismer som fremmer proteinsyntese og muskulær tilpasning.

Selv om resultatene gir innsikt i p70S6K-aktiveringen etter trening, må de tolkes med forsiktighet på grunn av metodiske begrensninger som bakgrunnsstøy på blottet, pipetteringsfeil, ujevn proteinoverføring og mangelen på en kalibreringskurve.

8.5 Oppløsninger

8.6 Koder brukt

```
library(dplyr)
library(tidyr)
library(gt)

# Opprett datasettet med originale kolonnenavn
data <- data.frame(
  Testperson = c("HE", "HE", "T", "T", "LØ"),
  Bein = c("Venstre", "Høgre", "Venstre", "Høgre", "Høgre"),
  "Membran blot (p-p70)" = c(16530.472, 10135.430, 11187.966, 7012.388, 14791.037),
  "Membran blot (t-p70)" = c(15516.158, 13009.673, 12917.480, 10868.744, 13617.430),
  "p-p70 (raw)" = c(4276.447, 4238.276, 4457.276, 4148.447, 4249.861),
```

```

"t-p70 (raw)" = c(4276.447, 4238.276, 4457.276, 4148.447, 4249.861),
"p-p70 (norm.)" = c(0.2587, 0.4182, 0.3985, 0.5916, 0.2872),
"t-p70 (norm.)" = c(0.2756, 0.3257, 0.3449, 0.3816, 0.3122),
Fosforyleringsnivå = c(0.939, 1.284, 1.155, 1.550, 0.920),
check.names = FALSE # Hindrer R fra å endre kolonnenavnene
)

# Bytt kolonner og rader
data_long <- data %>%
  pivot_longer(
    cols = -c(Testperson, Bein),
    names_to = "Variabel",
    values_to = "Verdi"
  ) %>%
  pivot_wider(
    names_from = c(Testperson, Bein),
    values_from = Verdi
  )

# Opprett datasettet med gj.sn ± SD format
data_result <- data.frame(
  Parameter = c("t-p70 (norm.)", "p-p70 (norm.)", "Fosforyleringsnivå"),
  `Hele gruppen (n = 20)` = c(
    "0.33 ± 0.04",
    "0.39 ± 0.12",
    "1.17 ± 0.25"
  ),
  `Høyre bein (n = 10)` = c(
    "0.35 ± 0.04",
    "0.50 ± 0.12",
    "1.43 ± 0.15"
  ),
  `Venstre bein (n = 10)` = c(
    "0.31 ± 0.03",
    "0.31 ± 0.07",
    "0.99 ± 0.09"
  ),
  check.names = FALSE
)

```



```

# Opprett datasettet
buffer_data <- data.frame(
  `Buffer Solution` = c("TBS 1L 10x, pH 7.6 (HCl adjust)",
                        "Running buffer 1L 10x, pH 8.3",
                        "Transfer buffer 1L 10x"),
  `Komponent` = c(
    "Tris Base\nNaCl",
    "Tris Base\nGlycin\nSDS",
    "Tris Base\nGlycin"
  ),
  `Konsentrasjon (mM)` = c(
    "200 mM\n1.37 M",
    "250 mM\n1.92 M\n35 mM",
    "250 mM\n1.92 M"
  ),
  `Masse (g)` = c(
    "24.22 g\n80.06 g",
    "30.28 g\n144.1 g\n10.09 g",
    "30.28 g\n144.1 g"
  ),
  check.names = FALSE
)

# Lag tabellen med `gt`
buffer_table <- buffer_data %>%
  gt() %>%
  tab_header(
    title = "Oversikt over oppløsninger",
    subtitle = "Tabellen viser bufferløsninger med komponenter, konsentrasjoner og masser."
  )

```

Referanser

- Atherton, P. J., & Smith, K. (2012). Muscle protein synthesis in response to nutrition and exercise. *The Journal of Physiology*, 590(5), 1049–1057. <https://doi.org/10.1113/jphysiol.2011.225003>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Bartlett, J. M., & Stirling, D. (2003). *PCR protocols: Methods in molecular biology* (Vol. 226). Humana Press.
- Bass, J. J., Wilkinson, D. J., Rankin, D., Phillips, B. E., Szewczyk, N. J., Smith, K., & Atherton, P. J. (2017). An overview of technical considerations for western blotting applications to physiological research. *Scandinavian Journal of Medicine & Science in Sports*, 27(1), 4–25. <https://doi.org/10.1111/sms.12702>
- Bompa, T. O., & Buzzichelli, C. (2019). *Periodization: Theory and methodology of training* (6th ed.). Human Kinetics.
- Brosowsky, N. P., Parshina, O., Locicero, A., & Crump, M. J. C. (2020). *Teaching undergraduate students to read empirical articles: An evaluation and revision of the QALMRI method*. <https://doi.org/10.31234/osf.io/p39sc>
- Browner, W. S., Newman, T. B., Cummings, S. R., Grady, D. G., Huang, A. J., Kanaya, A. M., & Pletcher, M. J. (2022). *Designing clinical research* (5th ed.). Lippincott Williams & Wilkins.
- Cleather, D. (2021). *Force: The biomechanics of training*. Independently Published. <https://books.google.no/books?id=TmyjzgEACAAJ>
- Coffey, V. G., & Hawley, J. A. (2007). The molecular bases of training adaptation. *Sports Medicine*, 37(9), 737–763. <https://doi.org/10.2165/00007256-200737090-00001>
- Currier, B. S., Mcleod, J. C., Banfield, L., Beyene, J., Welton, N. J., D’Souza, A. C., Keogh, J. A. J., Lin, L., Coletta, G., Yang, A., Colenso-Semple, L., Lau, K. J., Verboom, A., & Phillips, S. M. (2023). Resistance training prescription for muscle strength and hypertrophy in healthy adults: A systematic review and bayesian network meta-analysis. *British Journal of Sports Medicine*, 57(18), 1211–1220. <https://doi.org/10.1136/bjsports-2023-106807>
- D’Souza, R. F., & al., et. (2014). Dose-dependent increases in p70S6K phosphorylation and intramuscular branched-chain amino acids in older men following resistance exercise and protein intake. *Physiological Reports*, 2(8), e12112. <https://doi.org/10.14814/phy2.12112>
- Diez, D. M., Barr, C. D., & Çetinkaya-Rundel, M. (2022). *OpenIntro statistics* (4th ed.). OpenIntro, Inc.

- Fatouros, I. G., Jamurtas, A. Z., Leontsini, D., Taxildaris, K., Aggelousis, N., Kostopoulos, N., & Buckenmeyer, P. (2000). Evaluation of plyometric exercise training, weight training, and their combination on vertical jumping performance and leg strength. *Journal of Strength and Conditioning Research*, 14(4), 470–476.
- Folland, J. P., & Williams, A. G. (2007). The adaptations to strength training: Morphological and neurological contributions to increased strength. *Sports Medicine*, 37(2), 145–168. <https://doi.org/10.2165/00007256-200737020-00004>
- Halperin, I., Pyne, D. B., & Martin, D. T. (2015). Threats to internal validity in exercise science: A review of overlooked confounding variables [Journal Article]. *Int J Sports Physiol Perform*, 10(7), 823–829. <https://doi.org/10.1123/ijsp.2014-0566>
- Harput, G., Toprak, U., Colakoglu, F. F., Temel, E., Saylisoy, S., & Baltaci, G. (2023). Effects of plyometric training on sonographic characteristics of quadriceps muscle and patellar tendon, quadriceps strength, and jump height in adolescent female volleyball players. *International Journal of Athletic Therapy and Training*, 28(2), 89–96.
- Holz, M. K., & al., et. (2005). mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. *Cell*, 123(4), 569–580. <https://doi.org/10.1016/j.cell.2005.10.024>
- Hopkins, W. G. (2000). Measures of reliability in sports medicine and science [Journal Article]. *Sports Med*, 30(1), 1–15. <http://www.ncbi.nlm.nih.gov/pubmed/10907753>
- Hume, D. (2019). *An enquiry concerning human understanding*. Electric Book. <https://ebookcentral.proquest.com/lib/hilhmr-ebooks/detail.action?docID=5302411>
- Ioannidis, J. P. A. (2005). Why most published research findings are false [Journal Article]. *PLOS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Izquierdo, M., Ibáñez, J., Häkkinen, K., Kraemer, W. J., Ruesta, M., & Gorostiaga, E. M. (2004). Maximal strength and power, muscle mass, endurance and serum hormones in weightlifters and road cyclists. *Journal of Sports Sciences*, 22(5), 465–478. <https://doi.org/10.1080/02640410410001675342>
- Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., Kowalchuk, R. K., Lowman, L. L., Petoskey, M. D., Keselman, J. C., & Levin, J. R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research*, 68(3), 350–386. <https://doi.org/10.3102/00346543068003350>
- Kindermann, W., Simon, G., & Keul, J. (1979). The significance of the aerobic-anaerobic transition for the determination of work load intensities during endurance training. *European Journal of Applied Physiology and Occupational Physiology*, 42, 25–34. <https://doi.org/10.1007/BF00421101>
- Krieger, J. W. (2010). Single vs. Multiple sets of resistance exercise for muscle hypertrophy: A meta-analysis. *Journal of Strength and Conditioning Research*, 24(4), 1150–1159. <https://doi.org/10.1519/JSC.0b013e3181d4d436>
- Machado, F. A., Nakamura, F. Y., & Moraes, S. M. F. D. (2012). Influence of regression model and incremental test protocol on the relationship between lactate threshold using the maximal-deviation method and performance in female runners. *Journal of Sports Sciences*, 30(12), 1267–1274.

- Markovic, G., & Mikulic, P. (2010). Neuro-musculoskeletal and performance adaptations to lower-extremity plyometric training. *Sports Medicine*, 40(10), 859–895. <https://doi.org/10.2165/11318370-000000000-00000>
- McArdle, W. D., Katch, F. I., & Katch, V. L. (2014). *Exercise physiology: Nutrition, energy, and human performance* (8th ed.). Wolters Kluwer Health.
- McKinlay, B. J., Wallace, P., Dotan, R., Long, D., Tokuno, C., Gabriel, D. A., & Falk, B. (2018). Effects of plyometric and resistance training on muscle strength, explosiveness, and neuromuscular function in young adolescent soccer players. *Journal of Strength and Conditioning Research*, 32(11), 3039–3050.
- Mikami, E., Fuku, N., Murakami, H., Tsuchie, H., Takahashi, H., Ohiwa, N., Tanaka, H., Pitsiladis, Y. P., Higuchi, M., Miyachi, M., Kawahara, T., & Tanaka, M. (2014). ACTN3 R577X genotype is associated with sprinting in elite japanese athletes. *International Journal of Sports Medicine*, 35(2), 172–177. <https://doi.org/10.1055/s-0033-1347171>
- Noble, J. E., & Bailey, M. J. (2008). Chapter 8 quantitation of protein. In *Methods in enzymology* (Vol. 463, pp. 73–95). Academic Press. [https://doi.org/10.1016/S0076-6879\(09\)63008-1](https://doi.org/10.1016/S0076-6879(09)63008-1)
- North, K. N., & Beggs, A. H. (1996). Deficiency of a skeletal muscle isoform of -actinin (-actinin-3) in merosin-positive congenital muscular dystrophy. *Neuromuscular Disorders*, 6(4), 229–235. [https://doi.org/10.1016/0960-8966\(96\)00361-6](https://doi.org/10.1016/0960-8966(96)00361-6)
- North, K. N., Yang, N., Wattanasirichaigoon, D., Mills, M., Easteal, S., & Beggs, A. H. (1999). A common nonsense mutation results in -actinin-3 deficiency in the general population. *Nature Genetics*, 21(4), 353–354. <https://doi.org/10.1038/7675>
- Norton, D. F., & Taylor, J. (Eds.). (2008). *The cambridge companion to hume* (2nd ed.). Cambridge University Press.
- Poole, D. C., Rossiter, H. B., Brooks, G. A., & Gladden, L. B. (2021). The anaerobic threshold: 50+ years of controversy. *The Journal of Physiology*, 599(3), 737–767. <https://doi.org/10.1113/JP279963>
- Popper, K. R. (1959). *The logic of scientific discovery* (p. 479). Hutchinson.
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Raastad, T., Paulsen, G., Refsnes, P. E., Rønnestad, B. R., & Wisnes, A. R. (2010). *Styrketrening - i teori og praksis* (2nd ed.). Gyldendal Norsk Forlag.
- Ralston, G. W., Kilgore, L., Wyatt, F. B., & Baker, J. S. (2017). The effect of weekly set volume on strength gain: A meta-analysis. *Sports Medicine*, 47, 2585–2601. <https://doi.org/10.1007/s40279-017-0762-7>
- Ruiz, J. R., Sui, X., Lobelo, F., Morrow, J. R., Jackson, A. W., Sjöström, M., & Blair, S. N. (2008). Association between muscular strength and mortality in men: Prospective cohort study. *BMJ: British Medical Journal*, 337(7661), 92–95. <http://www.jstor.org/stable/20510282>
- Schadock, I., Schneider, A., Silva, E. D., Buchweitz, M. R. D., Correa, M. N., Pesquero, J. B., Paredes-Gamero, E. J., Araujo, R. C., & Barros, C. C. (2015). Simple method to genotype the ACTN3 r577x polymorphism. *Genetic Testing and Molecular Biomarkers*, 19(5), 253–256. <https://doi.org/10.1089/gtmb.2014.0299>

- Schoenfeld, B. J. (2016). *Science and development of muscle hypertrophy*. Human Kinetics.
- Schoenfeld, B. J., Grgic, J., Ogborn, D., & Krieger, J. W. (2019). Effects of low-vs. High-load resistance training on muscle strength and hypertrophy in well-trained men. *Journal of Strength and Conditioning Research*, 33(S1), S1–S10. <https://doi.org/10.1519/JSC.0000000000002468>
- Spiegelhalter, D. J. (2019a). *The art of statistics : How to learn from data* (First US edition., pp. xvi, 426 pages) [Book]. Basic Books.
- Spiegelhalter, D. J. (2019b). *The art of statistics : Learning from data* (1th ed.). Pelican.
- Store norske leksikon. (2005–2007). *Ockhams barberkniv*. https://snl.no/Ockhams_barberkniv.
- Sylta, Ø., Tønnessen, E., Hammarstrøm, D., Danielsen, J., Skovereng, K., Ravn, T., Rønnes-tad, B. R., Sandbakk, Ø., & Seiler, S. (2016). The effect of different high-intensity peri-odization models on endurance adaptations. *Medicine and Science in Sports and Exercise*, 48(11), 2165–2174. <https://doi.org/10.1249/MSS.0000000000001007>
- Tanner, R. K., & Gore, C. J. (2012). *Physiological tests for elite athletes 2nd edition* [Book]. Human Kinetics. <https://books.google.no/books?id=0OPIiMks58MC>
- Tranøy, K. E. (2021). *Induksjon – filosofi*. https://snl.no/induksjon_-_filosofi.
- Tucker, R., & Collins, M. (2012). What makes champions? A review of the relative contribu-tion of genes and training to sporting success. *British Journal of Sports Medicine*, 46(8), 555–561. <https://doi.org/10.1136/bjsports-2011-090548>
- Vissing, K., Brink, M., Lønbro, S., Sørensen, H., Overgaard, K., Danborg, K., Mortensen, J., Elstrøm, O., Rosenhøj, N., Ringgaard, S., Andersen, J. L., & Aagaard, P. (2008). Muscle adaptations to plyometric vs. Resistance training in untrained young men. *Journal of Strength and Conditioning Research*, 22(6), 1799–1810.
- Washif, J. A., & Kok, L. Y. (2022). Relationships between vertical jump metrics and sprint performance, and qualities that distinguish between faster and slower sprinters. *Journal of Science in Sport and Exercise*, 4, 135–144. <https://doi.org/10.1007/s42978-021-00122-4>
- Whitehead, M. T., Scheett, T. P., McGuigan, M. R., & Martin, A. V. (2018). A comparison of the effects of short-term plyometric and resistance training on lower-body muscular performance. *Journal of Strength and Conditioning Research*, 32(10), 2743–2749.
- Wickham, H. (2014). Tidy data [Journal Article]. *Journal of Statistical Software; Vol 1, Issue 10* (2014). <https://www.jstatsoft.org/v059/i10>
- Yang, N., MacArthur, D. G., Gulbin, J. P., Hahn, A. G., Beggs, A. H., Easteal, S., & North, K. (2003). ACTN3 genotype is associated with human elite athletic performance. *American Journal of Human Genetics*, 73(3), 627–631. <https://doi.org/10.1086/377590>
- Zhang, X., Feng, S., & Li, H. (2023). The effect of velocity loss on strength development and related training efficiency: A dose–response meta–analysis. *Healthcare*, 11(3). <https://doi.org/10.3390/healthcare11030337>