

Mappeeksamen IDR4000

Kandidatnummer: 503

2024-11-22

Table of contents

Forord	4
1 Reliabilitet og verktøy for reproduserbar vitenskapelig data	5
1.1 Introduksjon	5
1.2 Metode	5
1.2.1 Deltakere	5
1.2.2 Metode:	6
1.2.3 Protokoll	7
1.2.4 Standardisering:	8
1.2.5 Mulige avvik	8
1.3 Resultater	9
1.4 Diskusjon	10
2 Regresjonsmodeller og prediksjonsdata	12
2.1 Introduksjon / Bakgrunn	12
2.2 Del 1: Laktat terskler	12
2.2.1 Introduksjon	12
2.2.2 Metode	12
2.3 Del 2: Forutsi størrelser på DNA fragmenter eller stiningene i en qPCR-kalibreringskurve	17
2.3.1 Introduksjon	17
2.3.2 Metode	17
2.3.3 Diskusjon	20
2.4 Del 3: Tolkning av regresjonsmodell	20
2.4.1 Introduksjon	20
2.4.2 Metode	21
2.4.3 Diskusjon	22
3 Slutninger fra statistiske modeller og statistisk styrke	25
3.1 Introduksjon	25
3.2 Simulasjon	25
3.2.1 Oppgave 1.	26
3.2.2 Oppgave 2.	27
3.2.3 Oppgave 3.	27

3.3	Many studies	28
3.3.1	Oppgave 4.	29
3.3.2	Oppgave 5.	30
3.3.3	Oppgave 6.	31
3.3.4	Oppgave 7.	31
3.4	Many studies without population effect	32
3.4.1	Oppgave 8.	32
4	Studiedesign	35
4.1	Introduksjon	35
4.2	Hoveddel	36
4.2.1	Grunnlag for studiene	36
4.2.2	Studiedesign og forskningsmetoder	36
4.2.3	Resultat og konklusjon	38
4.3	Avslutning	38
5	Analyse av gjentatte målte eksperiment	40
5.1	Effekten av treningsvolum på hypertrofi og styrke	40
5.1.1	Introduksjon	40
5.1.2	Metode	41
5.1.3	Resultat	43
5.1.4	Diskusjon	44
6	Labrapport; Ekstraksjon og analyse av muskelprotein	47
6.1	Introduksjon og bakgrunn	47
6.2	Metode	48
6.2.1	Homogenisering av muskelprøver	48
6.2.2	Western blot protokoll	49
6.3	Resultater	53
6.4	Diskusjon	54
6.5	Konklusjon	54
7	Filosofihistore	55
7.1	Oppgave 1	55
7.2	Oppgave 2	56
	Kilder	58

Forord

<https://github.com/Nybakk/mappeeksamen-idr4000-anders-nybakk>

1 Reliabilitet og verktøy for reproduserbar vitenskapelig data

1.1 Introduksjon

I vår studie har vi gjennomført VO₂max-tester over fire forskjellige dager for å gjennomføre VO₂max-tester. Formålet med disse testene var å undersøke hvor konsistente resultatene er under kontrollerte forhold. For å oppnå dette har vi forsøkt å standardisere flere variabler, inkludert treningsnivå og matinntak dagen før testene. Ifølge Halperin *et al.* (2015) er det avgjørende å bruke standardiserte protokoller for å oppnå pålitelige resultater i fysiologiske tester.

I tillegg har vi målt laktatnivåer umiddelbart etter avslutningen av hver test for å vurdere den fysiologiske responsen. Laktatmålinger tas etter VO₂max-tester for å kontrollere at utøveren har presset seg til det ytterste. På den måten kan vi sikre oss reliabilitet under testen. Det er laktatnivået og Respiratory exchange ratio (RER) vi vil fokusere på i dette forsøket. Vi vil sammenlikne disse målingene, regne ut standard avvik (SD) for å teste reliabiliteten, og så presentere resultatene ved hjelp av tabeller og grafer. Tidligere forskning har vist at laktatnivåer kan gi verdifull informasjon om anaerob kapasitet og treningsintensitet, noe som er viktig for å forstå atletisk ytelse (Hopkins (2000); Tanner & Gore (2012)).

Målet med forsøket er å lære seg å lede VO₂max-tester på en reliabel og valid måte, føre og tolke data samt behandle disse i R og github for å lage reproduserbare rapporter.

1.2 Metode

1.2.1 Deltakere

Forsøket hadde i alt 14 deltakere mellom 20 og 40 år. Gruppen bestod av 3 kvinner og 11 menn. Treningsgrunnlaget til deltakerne var variert og selv om samtlige deltakere meddelte at de driver idrettsaktivitet jevnlig, er det noen som primært trener styrketrening, ikke utholdenhet.

Oppsummering av Demografiske Data
Gjennomsnitt og standardavvik per tidsperiode

Tidspunkt ¹	Gj.sn.	Alder ¹	SD Alder ¹	Gj.sn.	Høyde ¹	SD Høyde ¹	Gj.sn.	Vekt ¹	SD Vekt ¹	Anta
t1		25.66	3.64		179.79	7.63		79.16	11.85	
t2		25.66	3.64		179.79	7.63		79.27	11.82	
t3		26.79	3.69		178.73	7.91		80.45	11.92	
t4		29.06	7.22		174.62	6.16		74.74	12.32	

¹Alle verdier er beregnet per tidsperiode.

1.2.2 Metode:

Vi gjennomførte fire testdager. De to første var påfølgende dager, og de to siste var med en dag mellom. Hensikten med disse testdagene var å gjennomføre fysiologiske tester med høy grad av reliabilitet. Det er mange faktorer som kan påvirke validitet og reliabilitet, og det er veldig viktig å ta høyde for dette under fysiologisk testing. Vi tok flere forhåndsregler for å sikre så like testforhold som mulig.

En VO2max-test går ut på å måle hvor mange ml en person evner å ta opp og forbruke per minutt. Oksygenkravet øker lineært med økende belastning helt til personen når sin maksimale aerobe kapasitet, da kurven enten flater ut eller synker.

Maksimalt oksygenopptak beskrives enten i absolutte tall (ml/min) eller som relative tall i forhold til kroppsvekt (ml/kg/min).

VO2max-testen ble gjennomført som en trappetrinnstest der motstanden økte med 20W hvert minutt til utmattelse, eller når RPM < 60. Det registreres målinger av VO2max ved hvert 30 sek. Deltakerne startet testen på en watt tilsvarende fysisk form og erfaring med sykkel. Etter testene var ferdige ble informasjonen innhentet og plottet i et ferdig formatert Excel-dokument.

Da en tydelig protokoll ble fulgt for å etterstrebe så sikre og reliable tester som mulig, er det flere forhold som må tas underveis. Matinntak og koffeininntak fra samme dag og kvelden før ble registrert ved første test, og skulle være likt ved de påfølgende testene. Trening dagen før test ble også registrert, men lyktes ikke i å reprodusere dette da trening dagen før test 2 ble gårldagens VO2max-test. Vi hadde også føringer om at man skulle etterstrebe lik søvn på dagene før test. For å sikre lik grad av formuleringer og verbal instruks samt grad av engasjement og heing, hadde hver deltaker samme testleder ved hver test.

Underveis i testen måles også RER som en del av VO2max-testen. Laktatmålingene ble gjort et minutt etter testen ble avsluttet.

1.2.3 Protokoll

1.2.3.1 Før test:

Vi startet forberedelsene til testing med å gjøre klar laktatmaskinen ved å starte den, sjekke standard og evt, bytte den hvis det er tomt. Videre tok vi på hansker og monterte munnstykket til testingen, festet papir rundt og festet med neseclipe for å unngå tilskitning, fant fram slangen og monterte disse til miksekammeret. Så slo vi på Lode-sykkelen, vyntus og Lode-PC

Kalibrere oksygenanalysator

Kjøre gasskalibrering, åpne gassflaske (lukkes når kalibrering er ferdig), sjekke at sensor er koblet i maskinen og starte kalibrering. Vi godtar 2% feilmargin, hvis den er høyere må man rekalkibrere. Så kjører vi volumkalibrering - feilmargin på 0,2% eller under godtas. Kalibrere kammer, flytte sensor til kammer og skru av gass til slutt.

Forberede utstyr

Lage ny pasient på Vyntus og Lode; "Idr4000_h24_g1_idx". Veie personen i så lite klær som mulig (trekke fra 300g for klær). Lage plotteark, Stille inn krankarm (172,5) og kalibrere sykkel på Lode PC. Stille inn sykkel til deltaker og fullføre plotteark. Klargjøre laktatrør, papir og teip, samt teip til neseclipe. Velge protokoll på Lode PC - dersom personen ikke har erfaring med sykkel må man bli enige om en Watt man tror kan passe.

1.2.3.2 Under test:

Informasjon til deltaker

Forklare til deltakeren hva en VO2max-test er og måler, samt forklare hvordan denne skal gjennomføres. Det gjøres Målinger per 30 sek med watt-økning per 1 min. Deltakeren skal prøve å sykle med 80-100 rpm, og vi stopper test hvis under 60. VO2max-test er en prestasjonstest hvor hvert sekund teller. For å kontrollere resultatet blir det gjort pulsmåling og laktatmåling 1 min etter test med et stikk i fingeren. Etter testen spør vi også hvor deltakeren er på Borg skala (6-20). Til slutt spør vi deltakeren om den vil ha vifte på.

Informasjon underveis Deltakeren kan følge med på tiden og RPM under testen. Testleder gir oppdateringer og heier fram deltakeren. Testlederen er den samme hver gang og prøver å gjøre dette så likt som mulig på hver test for å få best mulig reliabilitet.

1.2.3.3 Etter test:

Rett etter testen er ferdig spør testleder om Borg skala for å teste validiteten på testen. Testen blir avsluttet på begge PC'ene og etter ett 1 minutt tas laktat-prøve. Tok 2 prøver for å sikre god reliabilitet.

Hente rapport:

Finne fram labbrapport på test-PC: Rapport: INN_TABELL:30SEK_MIX, søke opp id-nr. Lagre i rett mappe: F10 (nederste knapp) Lagre i rett mappe på minnepenn. Overføre til One Drive på labbPCn. Hente ut resultat fra laktat-prøve. Plotte inn resten av verdiene i excel.

Til slutt vaskes utstyret som skal vaskes og ryddes bort, før det evt. gjøres klar til ny test.

1.2.4 Standardisering:

I dette forsøket valgte vi å standardisere matinntak og koffeininntak i forkant av testen. Vi ønsket at testdeltakerne skulle spise det samme de tre siste måltidene før testen, og ha likt koffeininntak samme dag som testen.

Vi kunne også ha valgt å standardisere trening, men på grunn av forskjellig treningsopplegg hos deltakerne lot ikke dette seg gjøre.

Standardisering av tidspunkt ble gjort for oss med reservasjon av laboratorium 2. Helt standardisert blir det likevel ikke fordi i testuke 1 er testdagene rett etter hverandre, mens i testuke 2 er det en hviledag mellom testene.

Alle beregninger, grafer, tabeller og modeller er gjort i R (Anon, n.d.). For å sikre reproduserbarhet har alt arbeidet blitt lagret som en repository i github.

1.2.5 Mulige avvik

Deltakerne klarte ikke å standardisere kosthold, trening og søvn. På t4 tydelig lavere luftfuktighet i test-laben. Tidspunktet for t3 og t4 ble endret pga dobbeltbooket lab, slik at det ikke skjedde på samme tid på dagen som de andre testene, det ble ulik restutisjonstid fra de andre testene, og dårligere tid til å gjennomføre testene - mulig stressfaktor. Laktat-prøve tatt for sent, eller unyaktig. Unøyaktighet med synkronisert start av test i vyntus og Lode. Glemte å starte stoppeklokke til deltaker. Datainnsamlingen og testene ble utført av fire ulike grupper, og vi har ikke kontroll over utførelsen til de andre gruppene. Plottefeil - byttet RER og laktat.

Oppsummering av Laktat- og RER-målinger
Gjennomsnitt, standardavvik og antall per tidsperiode

Tidspunkt ¹	Gj.sn. Laktat ¹	SD Laktat ¹	Antall Laktat ¹	Gj.sn. RER ¹	SD RER ¹	Antall RER ¹
t1	12.20	2.76	14	1.14	0.03	14
t2	12.30	2.91	14	1.15	0.04	14
t3	13.64	2.50	11	1.18	0.05	11
t4	13.82	2.31	8	1.19	0.03	8

¹Alle verdier er beregnet per tidsperiode.

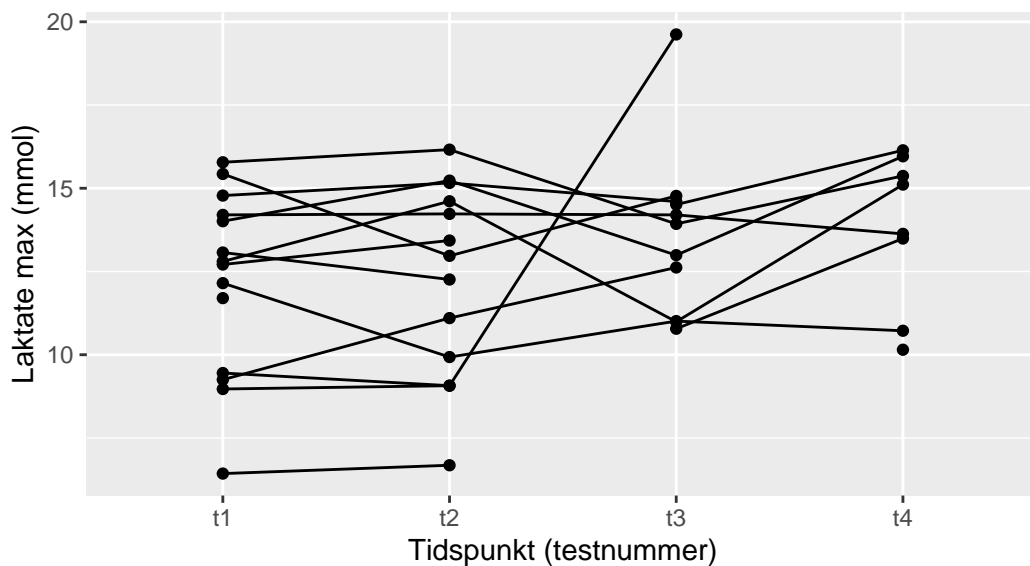
1.3 Resultater

Vi valgte å se på laktat og RER i denne oppgaven. Når RER øker over 1.0 indikere det ofte at laktatproduksjonen øker. RER samfatter ofte med den anaerobe terskelen hvor laktatnivået begynner å stige raskt. Av de 14 deltakerne var det bare 8 (2 kvinner og 6 menn) som gjennomførte alle fire testene. Dette kom delvis som følge av at tidspunktene for t3 og t4 ble endret.

I tabellen over vises gjennomsnittsverdiene for laktat og RER på hver av de fire testene samt SD for disse med variasjonskoeffisientene for laktat på: 7.6%, og 1.6% for RER.

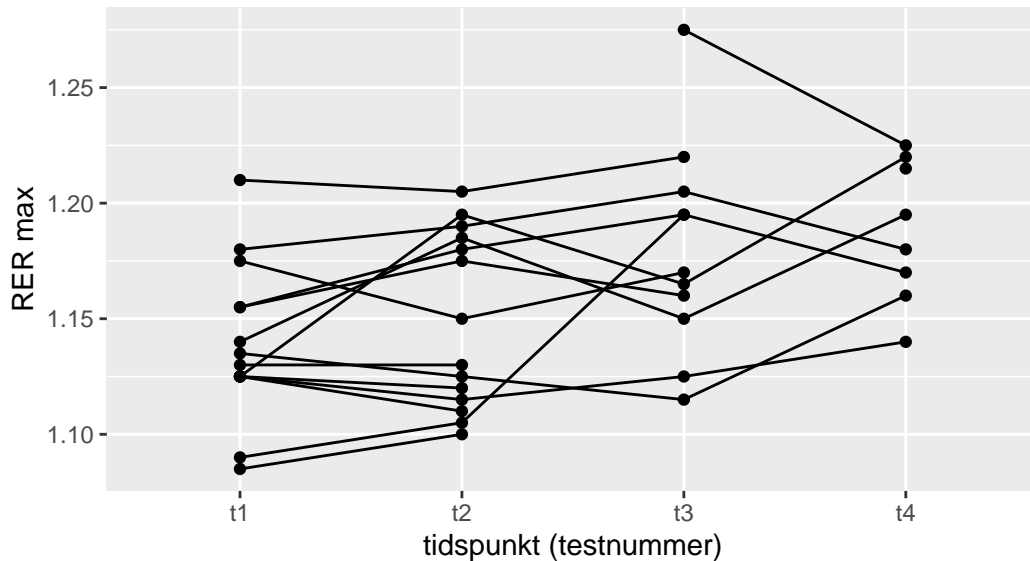
Laktatmålinger

Grafen viser de forskjellige laktatmålingene til hver enkelt deltaker, og foran



RER-målinger

Grafen viser de forskjellige RER-målingene til hver enkelt deltaker, og for



Grafene viser testresultatene for hver deltaker på hver enkelt test. Her kan vi se at noen få har relativt like målinger, men de fleste har svært varierende verdier. Noe av dette kan forklares med at noen av deltakerne er aktive utøvere som ikke ville gjennomføre flere maksimale tester på så kort tid i fare for å forstyrre treningsopplegget sitt. Noen har derfor bevisst testet submaksimalt på en eller flere tester.

1.4 Diskusjon

Variasjonskoeffisientene for laktat på: 7.6%, og 1.6% for RER indikerer at RER-målingene er betydelig mer konsistente enn laktatmålingene. Hvor RER har høy reliabilitet og laktat har betraktelig lavere reliabilitet. Dette er også noe som gjenspeiles på standardavvikene til de respektive målingene.

Med tanke på de demografiske variablene ser vi at det er noe forskjell i alder, høyde og vekt mellom tidspunktene og kjønnene. Det er også få kvinner med. Det er en særlig større variasjon i alder på t4 (7.22 år.). Dette har mest sannsynlig sammenheng med at de to siste og spesielt t4 hadde færre deltagere. Dette vil påvirke reliabiliteten negativt.

Samlet sett viser dataen at RER-målingene er høyst reliable og at laktatmålingene er mindre konsistente. Dette gir mening ettersom det er liten sjanse for feil i målingene av RER som gjøres av vintus. Måling av laktat er det testleder som gjør, og med manglende erfaring vil det naturlig nok være større sjanse for avvik i utførelsen. Vi tror en av de mest sannsynlige årsakene er vår uerfarenhet med å ta laktatprøver. Tid etter endt test før prøven ble tatt,

forurensning av for eksempel svette i blod eller prøver med for lite eller for mye blod kan være mulige feilkilder. Det trenger likevel ikke å bety at laktatmålingene er helt ubrukelige, men at de bør tolkes med forsiktighet og i sammenheng med andre fysiologiske markører som RER.

Den store variasjonen i resultatene for enkeltdeltakere som kommer fram i grafene tyder på at deltakerne har hatt stor variasjon i hvor mye de har presset seg under testene. Som vi skriver i protokollen er VO₂max-tester prestasjonstester hvor maksimal ytelse er viktig for et reliabelt resultat. De varierende verdiene vi presenterer her tyder derfor på at reliabiliteten for VO₂max-testene også trolig er lav. Dette gjør at det er vanskelig å si hva vi kan forvente av repeterte målinger i denne gruppen. En måte å bedre reliabiliteten på er å fjerne resultatene til deltakere som svinger mye i RER og laktatmålinger, eller å fortsette forsøket med økt fokus på maksimal ytelse i framtidige tester.

2 Regresjonsmodeller og prediksjonsdata

2.1 Introduksjon / Bakgrunn

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 molekyllær 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* datapakken *exscidata*.

2.2 Del 1: Laktat terskler

2.2.1 Introduksjon

Laktat terskel er en variabel som er godt brukt for å forutsi prestasjon innenfor utholdenhets idretter, til å styre intensiteten av treningsøkter og evaluere trenings effekt (Machado *et al.*, 2012). Det finnes ulike metoder for å finne testpersonens laktat terskel. Machado *et al.* (2012) forteller oss at den “maximal-deviation method” (D_{max}) anbefalt av Cheng et al. 1992, bidrar med å kunne evaluere de ulike mekanismene som virker bestemmende for prestasjon innenfor langdistanseløping og sykling (Machado *et al.*, 2012). Videre hadde denne metoden en bedre korrelasjon med prestasjon og laktat terskel sammenliknet med andre metoder. I våres reliabilitets tester ble det ikke utført laktat målinger, på bakgrunn av dette benytter vi oss av data settet til “cyclingstudy”. De representative tersklene som blir undersøkt er 2 mmol L⁻¹ og 4 mmol L⁻¹.

2.2.2 Metode

Som en kan se i den plotta grafen under, er de forskjellige grafene ikke så forskjellige rundt 2mmol og 4mmol L⁻¹. På den andre siden ser vi at den lineære modellen er feil ved 300w, den sekundærpolynomiske modellen er feil ved 275w. Den tredje- og fjerdepolynomiske modellen derimot, varierer ikke mye fra hverandre.

```

### laste ned nødvendige packages
library(tidyr)
library(tidyverse)
library(ggplot2)
library(exscidata)

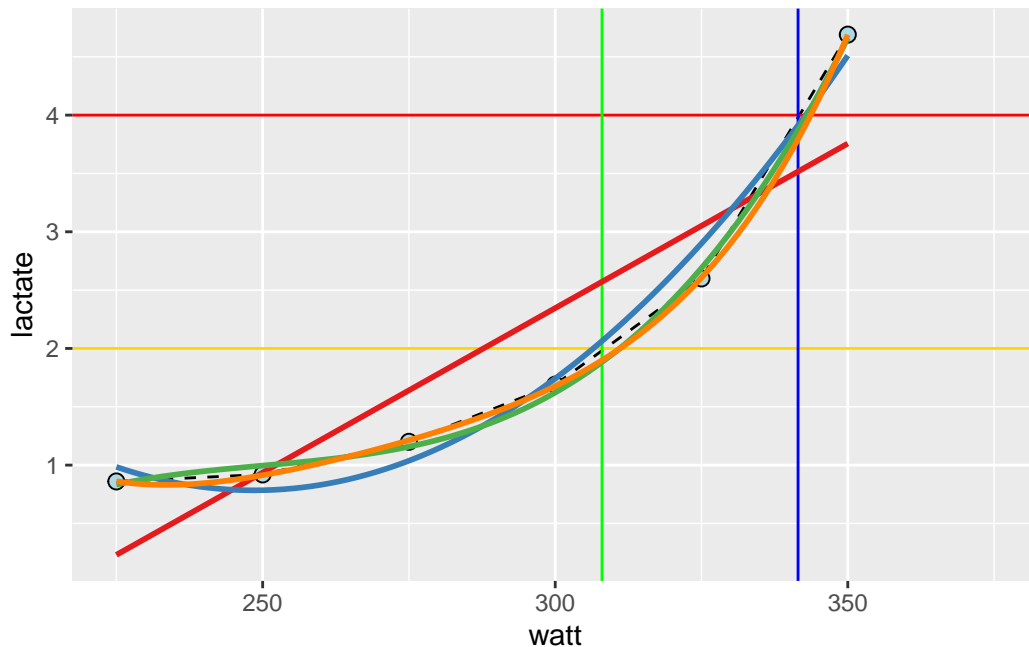
###laste inn data
data("cyclingstudy")

### 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 skjeringspunktet 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 tredjegradspynomisk 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")

```



vurdering av tilpasningen til de forskjellige lineære modellene på sammenhengen mellom t

```
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))

# Legger til en strak linje fra modellen.
m1 <- lm(lactate ~ watt, data = lactate)

# Legger til en andregradspynomisk modell.
m2 <- lm(lactate ~ poly(watt, 2, raw = TRUE), data = lactate)

# Legger til en tredjegradsynomisk modell.
```

```

m3 <- lm(lactate ~ poly(watt, 3, raw = TRUE), data = lactate)

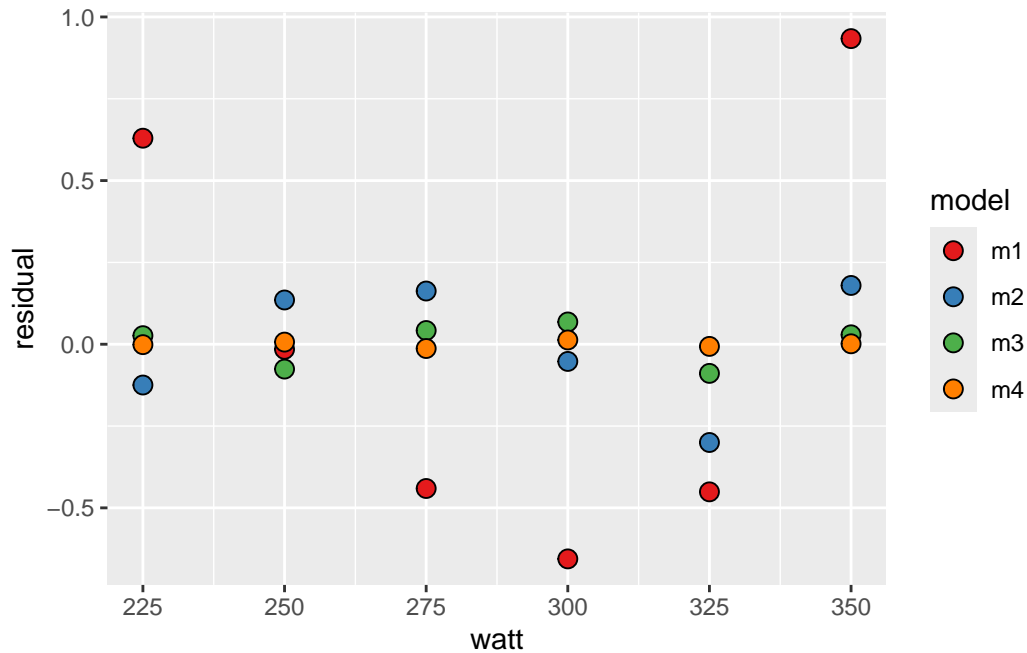
# Legger til en fjerdegradspynomisk modell.
m4 <- lm(lactate ~ poly(watt, 4, raw = TRUE), data = lactate)

# 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 modellemer.
  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 akse og restverdiene på y akse
  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"))

```



For å finne ut hva forutsatt wattverdi som er nærmest 2 og 4 mmol L-1, benytter vi koden under:

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

ndf$predictions <- predict(m3, newdata = ndf)

# for å finne ut kva forutsatt Wattverdi som er nermost 2 og 4 mmol L-1
lactate_threshold <- ndf %>%
  filter(abs(predictions - 4) == min(abs(predictions - 4)))

summary(lactate_threshold)
```

	watt	predictions
Min.	:343	Min. :4
1st Qu.:	:343	1st Qu.:4
Median :	:343	Median :4
Mean :	:343	Mean :4
3rd Qu.:	:343	3rd Qu.:4
Max.	:343	Max. :4

Her finner vi ut av at på 2 mmol får vi en wattverdi på 311 W, mens på 4 mmol får vi en wattverdi på 343 W. Verdien for 2 mmol ligg på samme dataframe som kjem fram på 4 mmol L-1.

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

2.3.1 Introduksjon

I denne delen av oppgaven tar vi utgangspunkt i forsøket vi gjorde på molekylærlabben 05. - 06. september, hvor vi ekstraherte DNA fra blod. Videre forsøkte vi å isolere genene som assosieres med hurtig muskelfibersammentrekning (R/R) og langsom muskelfibersammentrekning (X/X) ved hjelp av en PCR-maskin. Prøvene herfra ble testet videre ved hjelp av elektroforese i agarose gel sammen med en DNA-stige (ladder) som brukes som markør for å kartlegge genene. Etter elektroforesene tok vi bilde av prøven slik at vi kunne observere resultatene. Stigen markerer antall hvert fete basepar (bp) opp til 300, og hvert 100. basepar videre til 1000bp. det dominante R/R-genet har 413bp og det ressesive X/X-genet har 318bp. De små genmolekylene med få basepar vil trenge lenger in i gelen under elektroforesen, så X/X-genet vil altså trenge lenger inn i gelen ved elektroforese. Dette kan være vanskelig å observere bare med øynene, og vil ikke være særlig nøyaktig. For å få et mer reliabelt resultat har vi derfor brukt følgende metode. Fra prøvene våre var det tre brønner som gav resultat - to med et gen, og en med to (Wackerhage, 2014).

2.3.2 Metode

Først har vi brukt ImageJ Fiji for å behandle bildet vi fikk fra DNA-prøvene. Vi inverterte bildet for å få tydeligere farger, roterte det rett vei og klippet ut delen av bildet vi ville bruke - altså analysen av våre prøver. Videre brukte vi rektangelverktøyet for å markere stigen og prøvene vi ville analysere. Ut fra de markerte områdene lager ImageJ fiji grafer for hver brønn. Vi markerte toppunktene i alle grafene som indikerer gen (og trinn på stigen). Programmet registrerer plasseringen til toppunktene og disse "koordinatene" legges inn i et excell-dokument som vi bruker til beregningene.

```
library(readxl)

dat <- read_excel("data/resultat_dna_analyse.xlsx")
```

For å finne ut av molekylstørrelsen til de ukjente prøvene våre må vi først kalibrere stigen. Dette gjør vi ved å lage en data.fram som vi kaller ladder. Her er tallene omvendt proporsjonale

ettersom det er de små molekylene som trekkes lengst inn i gelen. Denne dataframen kaller vi “ladder”.

Videre må vi også lage en data.frame for de ukjente variablene som vi kaller “unknown”

```
# lage 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))

# lage ny dataramme med ukjente variabler

unknown <- data.frame(dist = c(258.5, 262.5, 265.5, 335.5))
```

For å lage en kalibreringsmodell bruker vi de samme dataene i ggplot for å vise stigen. Dette brukes videre for å estimere størrelsen på de ukjente variablene. Vi valgte å bruke en bøydd graf (poly) for å få minst mulig avvik.

```
# lage en kalibreringsmodell ved hjelp av stigen

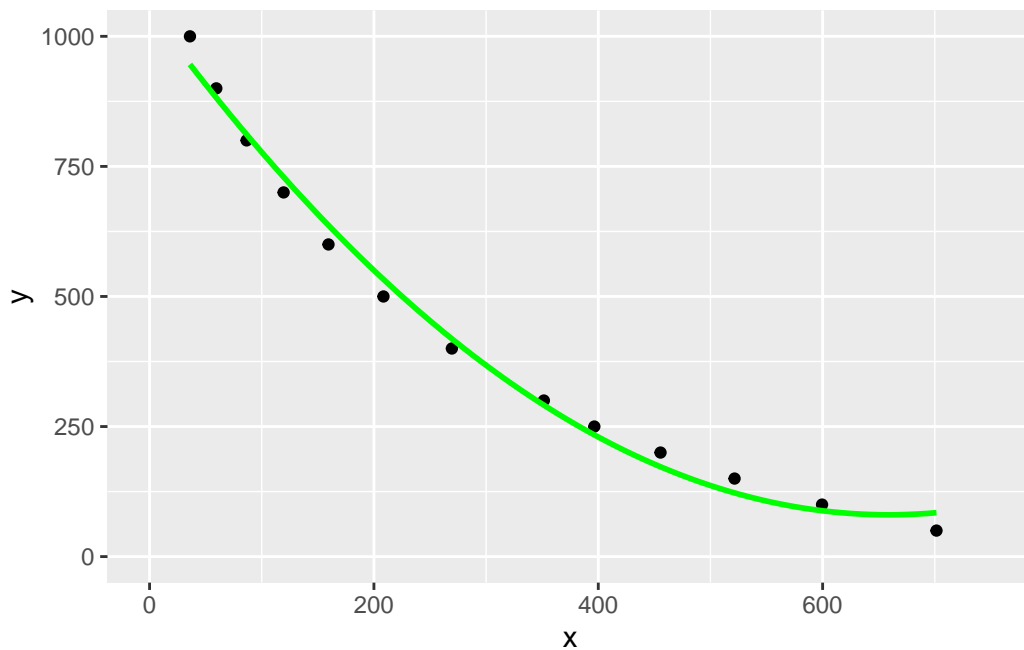
library(tidyverse)

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))
```



Til slutt brukte vi følgende koder for å estimere molekylstørrelsen på genene i prøven vår.

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

# Check model performance, R^2 should be ~ 1.
summary(cal)
```

Call:

```
lm(formula = log(mw) ~ dist, data = ladder)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.244363	-0.040218	-0.004565	0.082943	0.112630

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.0915695	0.0480419	147.61	< 2e-16 ***
dist	-0.0041842	0.0001298	-32.23	3.06e-12 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.09807 on 11 degrees of freedom

Multiple R-squared: 0.9895, Adjusted R-squared: 0.9886

F-statistic: 1039 on 1 and 11 DF, p-value: 3.059e-12

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

Brønn 1: 407 bp Brønn 2: 401 bp Brønn 3: 396 bp og 296 bp

2.3.3 Diskusjon

Denne analysen viser at ingen av genene våre har helt riktig størrelse i forhold til genene vi testet for - R/R (413bp) og X/X (318bp). Vi er likevel i nærheten som kan tyde på at allelene for brønn 1 og 2 er R/R og brønn 3 er R/X. Avviket kan forklares med unøyaktighet under DNA-testen (sansynlig ettersom validitetskontrollen i prøveresultatet ikke kom fram) og med dårlig kvalitet på bildet som vi brukte i denne oppgaven. I rapporten fra forsøket tolket vi prøvene annerledes og trodde at brønn 1, 2 og 3 alle hadde alleler litt over 300bp og at brønn 3 i tillegg hadde en feil med en ukjent allel som var på 250bp. Dette viser at det er mye unøyaktighet ved å bruke kun øynene til å tolke resultatet.

2.4 Del 3: Tolkning av regresjonsmodell

2.4.1 Introduksjon

I denne delen av oppgaven har vi valgt å se nærmere på variablene FAST_NUCLEI_T1 og TRAINING_AGE i datasettet `hypertrofi`, som er en del av `exscidata` pakken. For utforming av tabeller, figurer og grafer bruker vi `tidyverse`, `broom` og `gt`.

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

I `hypertrofi` er FAST_NUCLEI_T1 beskrevet som antall myonuclei per type-II muskelfiber, mens TRAINING_AGE viser til antall år med tidligere treningserfaring. Antall myonuclei per type-II muskelfiber, kan ha noko å si om muskelens egenskap til å utvikle kraft og personers

styrke (McArdle *et al.*, 2014, kap 22). Det er også diskutert om trening kan føre til endringer i muskelfiber type eller om de genetiske faktorene er det som er avgjørende for muskelfiber type fordelingen til den enkelte (McArdle *et al.*, 2014 p.s.535). Vi ønsker derfor å se nærmere om det er en lineær sammenheng mellom FAST_NUCLEI_T1 og TRAINING_AGE i datasettet hypertrofi.

Spørsmålet: Er det et lineært forhold mellom myonuclei per fiber CSA i type 2 og treningsalder?

2.4.2 Metode

Under i Figure 2.1 er FAST_NUCLEI_T1 satt som den avhengige variabelen på y-aksen, mens TRAINING_AGE er valgt som den uavhengige variabelen på x-aksen. Grafen er ment for å gi oss et raskt overblikk av dataene.

```
# Laster inn data
data("hypertrophy")

# 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()
```

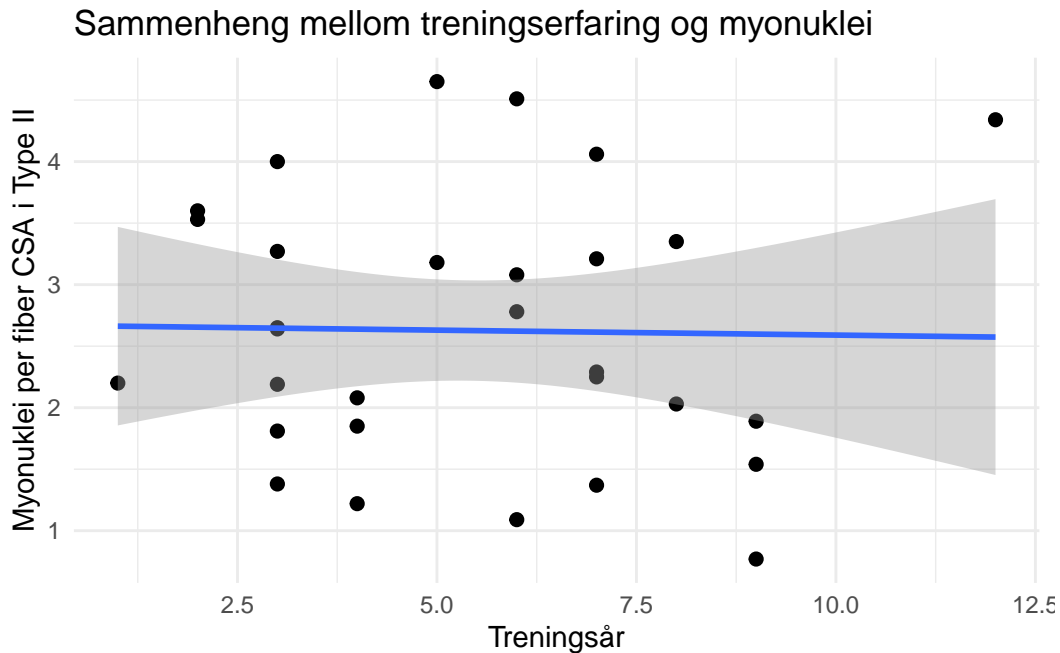


Figure 2.1: Sammenheng mellom treningsalder og myonuclei per fiber CSA i Type-II

ved hjelp av `geom_smooth` har vi lagt inn den best tilpassede linjen til datapunktene, også kalt lineær regresjonslinje (Spiegelhalter, 2019b p.s.128–129). Det gråe området omkring regresjonslinjen, visualiserer konfidensintervallet til linjen. Et bredt konfidensintervall som fremstilt her, indikerer større usikkerhet i hvordan variablene relaterer til hverandre (Spiegelhalter, 2019b p.s.240–244).

For å presentere regresjonslinjen, har vi laget en lineær statistisk modell for hjelp til videre tolkning mellom forholdet av dei to variablene. Oppsummering av de statistiske parametrene som vi har valgt å fokusere på i diskusjonen vår er listet opp i Table 2.1 under.

2.4.3 Diskusjon

I tabellen kan vi lese av verdiene for estimert koeffisient (regresjonskoeffisient), standardfeil, t-verdi og p-verdi. Den estimerte koeffisienten til “Treningserfaring (år)” forteller oss hvor mye `FAST_NUCLEI_T1` endres per enhet økning i `TRAINING_AGE`. I vårt tilfelle ser man en antall nukleikjerner per fiber reduseres med 0.008 per år med treningserfaring.

Standardfeilen måler hvor mye koeffisientene er forventet å variere fra utvalg til utvalg. Standardfeilen som vi har fått er liten i tallverdi, og man kan da fort konkludere at estimeringen er presis grunnet lav standardfeil. Samtidig er det viktig å se standardfeilen i lys av den estimerte

Table 2.1: Sammenheng mellom treningserfaring og myonuklei per muskelfiber type-II

```
# 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 ikke er aktuelt når vi kun skal se om
  # det er en lineær sammenheng mellom dei to variablene

# Lager regresjonstabell med forklarende radnavn
regression_table <- model_summary %>%
  select(term, estimate, std.error, statistic, p.value) %>%
  gt() %>%
  fmt_auto() %>%
  cols_label(
    term = "Term",
    estimate = "Estimert koeffisient",
    std.error = "Standardfeil",
    statistic = md("*t*-verdi"),
    p.value = md("*p*-verdi")
  ) %>%
  tab_source_note(
    source_note = "**Notat**: *p*-verdier mindre enn 0.05 anses som statistisk signifikante."
  )

# Vis resultatene
print(regression_table)
```

Term
Estimert koeffisient
Standardfeil
t-verdi
p-verdi
Treningserfaring (år)
−0.008
0.077
−0.104
0.918

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

koefisienten. I forhold til koefisienten selv, er standardfeilen stor, og betyr at man burde være usikker på nøyaktigheten til estimatet (Spiegelhalter, 2019b p.s.230–232)

t-verdien sier hvor mange standardavvik den estimerte koefisienten er fra 0, der jo høyere *t*-verdien (enten negativ eller positiv), dess mer signifikant er koefisienten (Spiegelhalter, 2019b p.s.275–276). Hos oss er *t*-verdien -0.104, noe som indikerer at det ikke er noe signifikant lineær sammenheng mellom `FAST_NUCLEI_T1` og `TRAINING_AGE`.

Nært knyttet til *t*-verdien, har man *p-verdien* som hjelper oss å si om *t*-verdien er statistisk signifikant. *P*-verdi er sannsynligheten for å observere en så ekstrem teststatistikk som den *t*-verdien vi har fått, gitt antagelsen at det ikke er en sammenheng mellom variablene våre (Spiegelhalter, 2019b p.s.264–265). Basert på at *p*-verdien i vår modell er 0.918, er det 91,8 % sannsynlighet at man vil observere en *t*-verdi på -0.008. Vi har derfor ikke tilstrekkelig bevis for å kunne si at den uavhengige variabelen `TRAINING_AGE` har en effekt på den avhengige variabelen `FAST_NUCLEI_T1`, og at det er en statistisk lineær sammenheng mellom variablene (Spiegelhalter, 2019b p.s.265–268).

Selv om *p*-verdi er et nyttig verktøy for å hjelpe oss å trekke slutninger om koefisientenes statistiske signifikans, sier den oss ikke noe om størrelsen på en effekt eller hva praktisk betydning den kan ha. Størrelsen på datasettet har også en betydning på *p*-verdien, der små datasett, som det vi har, kan gi høye *p*-verdier selv om det er en betydelig effekt (Spiegelhalter, 2019b p.s.285)

3 Slutninger fra statistiske modeller og statistisk styrke

3.1 Introduksjon

I denne oppgaven skal se på statistisk forskning. Vi skal simulere to forskningsprosjekt med forskjellig størrelse på utvalg. Den første gruppen (m1) har et utvalg på 8 målinger, og den andre gruppen har et utvalg på 40 målinger. Vi skal se hva forskjellig størrelse på utvalg gjør med resultatene.

3.2 Simulasjon

```
library(tidyverse)

set.seed(1)
population <- rnorm(1000000, mean = 1.5, sd = 3)

samp1 <- data.frame(y = sample(population, 8, replace = FALSE))
samp2 <- data.frame(y = sample(population, 40, replace = FALSE))

m1 <- lm(y ~ 1, data = samp1)
m2 <- lm(y ~ 1, data = samp2)

summary(m1)
```

Call:
lm(formula = y ~ 1, data = samp1)

Residuals:

Min	1Q	Median	3Q	Max
-6.5322	-1.2523	-0.0883	1.3540	4.8692

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.840	1.251	1.47	0.185

Residual standard error: 3.539 on 7 degrees of freedom

`summary(m2)`

Call:

`lm(formula = y ~ 1, data = samp2)`

Residuals:

Min	1Q	Median	3Q	Max
-5.6557	-2.2883	0.2636	2.2549	6.4212

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.5642	0.4774	3.276	0.00221 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.019 on 39 degrees of freedom

3.2.1 Oppgave 1.

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

Over kan vi se resultatene av simuleringen. og vi får følgende resultater:

	m1	m2
Estimat	1.84	1.5642
Standard feil	1.251	0.4774
t-verdi	1.47	3.276
p-verdi	0.185	0.00221

3.2.2 Oppgave 2.

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

Når vi øker størelsen på utvalget ser vi at dette påvirker resultatene. I en faktisk forskning vil vi jo ikke vite det faktiske gjennomsnittet i en populasjon. I denne simuleringen derimot har vi bestemt at gjennomsnittet (mean) er 1.5. Når vi har et utvalg på 8 observasjoner får vi resultatet 1.84, og med 40 observasjoner får vi 1.56 som estimat. Ved å øke antall faktiske observasjoner kan vi altså gjøre et mer presist estimat av hva som er gjennomsnittet i en populasjon.

På samme måte vil størelsen på utvalget påvirke standard feilen i forsøket. Standard feil beregnes ved å dele standard avvik, som i denne simuleringen er 3 på kvadratroten av antall observasjoner. I denne simuleringen blir det altså tre delt på kvadratroten av 8 i m1, og kvadratroten av 40 i m2. Som vi kan se får vi da lavere standarders feil i m2 hvor utvalget er større. Standard feil handler nemlig om hvor sannsynlig det er at utvalget er representativt for populasjonene. Lav standard feil betyr at det er stor sansynlighet for at utvalget er representativt.

T-verdien beregnes ved å dele estimatet med standard feil og forteller oss om forskjellen i gruppene er signifikant (Spiegelhalter, 2019a). Videre sier Spiegelhalter at en t-verdi over 2 tilsvarer en p-verdi, som er et mål på forskjellen mellom innsamlet data og null-hypotesen, under 0,05 som igjen vil bety at statistikken er signifikant. I vår simulering kan vi se at dette blir avgjørende. I simuleringen med 8 i utvalget får vi en t-verdi på 1.47 som tilsvarer en p-verdi på 0.185 altså ikke signifikant. I utvalget med 40 derimot er t-verdien 3,276 og p-verdien 0.00221 som vil si at resultatet er signifikant. I forsøket med 8 observasjoner ville det altså ha blitt gjort en type II feil, altså å avvise en korrekt alternativ hypotese fordi testresultatet støtter null-hypotesen (Spiegelhalter, 2019a).

3.2.3 Oppgave 3.

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

Grafen viser en tosidig p-verdi for m1. Midt i grafen ser vi det estimerte gjennomsnittet. En tosidig p-verdi sier noe om hvor mange observasjoner vi kan regne med å få fra populasjonen som er like ekstreme eller mer ekstreme enn den observerte t-verdien (Spiegelhalter, 2019a). De blå feltene viser altså hvor mange observasjoner vi kan ha utenfor null-hypotesen uten at null-hypotesen blir motbevist.

3.3 Many studies

```
library(tidyverse)

set.seed(1)
population <- rnorm(1000000, mean = 1.5, sd = 3)

# Create data frames to store the model estimates
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)

# A for loop used to sample 1000 studies, each iteration (i) will draw a new sample
# from the population.

for(i in 1:1000) {

  # Draw a sample
  samp1 <- data.frame(y = sample(population, 8, replace = FALSE))
  samp2 <- data.frame(y = sample(population, 40, replace = FALSE))

  # Model the data
  m1 <- lm(y ~ 1, data = samp1)
  m2 <- lm(y ~ 1, data = samp2)

  # Extract values from the models
  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]
```

```

}

# Save the results in a combined data frame

results <- bind_rows(results_8, results_40)

```

3.3.1 Oppgave 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?

m1: 8

```

library(dplyr)

# Beregn standardavviket til estimate og gjennomsnittet av se for hver utvalgsstørrelse
results_summary <- results %>%
  group_by(n) %>%
  summarise(
    mean = mean(estimate),
    std_estimate = sd(estimate), # Standardavviket til estimate
    avg_se = mean(se)           # Gjennomsnittet av standardfeilen
  )

# Vis sammendraget av resultatene
print(results_summary)

```

```

# A tibble: 2 x 4
   n   mean std_estimate avg_se
<dbl> <dbl>         <dbl> <dbl>
1     8  1.52          1.07  1.02
2    40  1.51          0.484 0.470

```

Som vi kan se over er standard avvik og gjennomsnittlig standard feil svært like. Forskjellen er 0,05 i m1 og 0,014 i m2. Grunnen til at standard avvik og standard feil er så like er på grunn av de henger sammen. Standard feil finner vi som tidligere nevnt ved å dele standard avvik på kvadratroten av utvalget. Begge disse verdiene blir brukt for å lage kurvemodeller som viser p-verdi og t-verdi. Når vi får lavere standard avvik og standard feil vil kurven bli smalere og

spissere, fordi estimatet kommer nærmere det faktiske gjennomsnittet. hvis vi skulle laget en kurve ut fra utvalgene m1 og m2, ville altså m1 være bredere og rundere enn m2 som ville være smal og spiss.

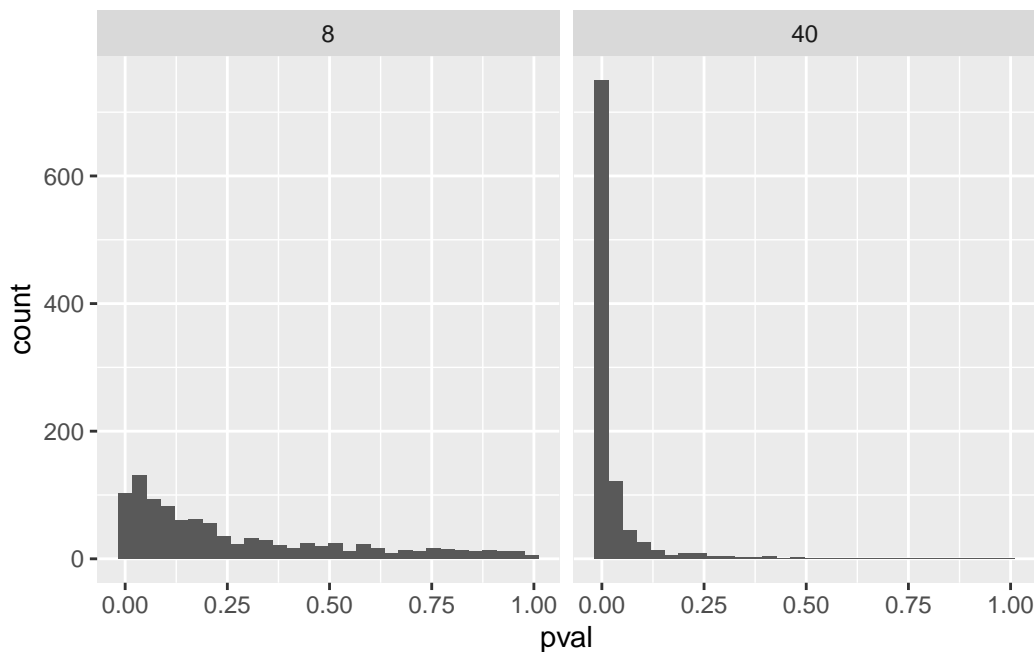
beregne kurven

3.3.2 Oppgave 5.

Create a histogram (see example code below) 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?

```
# Example code for copy and paste

# A two facets histogram can be created with ggplot2
results %>%
  ggplot(aes(pval)) +
  geom_histogram() +
  facet_wrap(~ n)
```



I denne oppgaven kan vi se hva jeg snakket om i forrige oppgave. Et større utvalg vil gjøre at histogrammet blir smalere og spissere. Vi samler estimatene mot gjennomsnittet. Som vi kan se fra histogrammene vil et større utvalg gi lavere p-verdi, som i dette tilfelle vil gjøre resultatet

signifikant som jeg har vist tidligere i oppgaven. Statistisk styrke er sansynligheten for å korrekt forkaste nullhypotesen gitt at den nye hypotesen stemmer, har en klar sammenheng med utvalgstørrelsen (Spiegelhalter, 2019a). Med større utvalg vil altså styrken øke, som vi også kan se i utregningen over.

3.3.3 Oppgave 6.

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

```
# Count the proportion of tests below a certain p-value for each
results %>%
  filter(pval < 0.05) %>%
  group_by(n) %>%
  summarise(sig_results = n()/1000)
```

```
# A tibble: 2 x 2
      n sig_results
<dbl>   <dbl>
1     8     0.227
2    40     0.865
```

I denne beregningen skal jeg finne ut hvor mange studier med utvalgsstørrelser 8 (lik m1) og 40 (lik m2) som vil få statistisk signifikante resultat med en p-verdi $< 0,05$. Her får jeg resultatene 0,227 for m1, og 0,865 for m2. Dette kan vi gjøre om til prosentverier: m1: 22,7 % og m2: 86,5 %. Det er altså stor forskjell på den betydelige effekten på de forskjellige studiene. I studier med 8 deltakere vil det bare være 22,7 % sjanse for å få et statistisk signifikant resultat, mens det i studier med 40 vil være 86,5 % sjanse for å få et statistisk signifikant resultat.

3.3.4 Oppgave 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.

```
# Using the pwr package
library(pwr)

pwr.t.test(n = 8, sig.level = 0.05, d = 1.5/3, type = "one.sample")
```

```
One-sample t test power calculation
```

```
      n = 8  
      d = 0.5  
sig.level = 0.05  
power = 0.232077  
alternative = two.sided
```

```
pwr.t.test(n = 40, sig.level = 0.05, d = 1.5/3, type = "one.sample")
```

```
One-sample t test power calculation
```

```
      n = 40  
      d = 0.5  
sig.level = 0.05  
power = 0.8693981  
alternative = two.sided
```

Statistisk styrke er sannsynligheten for å korrekt avvise null-hypotesen, gitt at den alternative hypotesen er sann (Spiegelhalter, 2019a). Lav statistisk styrke øker sannsynligheten for å begå type I feil som vil si at man feilaktiv avviser en korrekt null-hypotese (Spiegelhalter, 2019a). Her får m1 en statistisk styrke på 0,232. Dette stemmer bra med simuleringen i oppgave 6. I m2 derimot får vi en statistisk styrke på 0,869 som er en ganske høy statistisk styrke. I studien m2 er det altså lav sannsynlighet for å begå en type I feil. Dette støtter igjen det jeg har kommet fram til i tidligere oppgaver om at større utvalg øker sannsynligheten for et korrekt resultat, og at hvis vi skulle stole på resultatene fra m1 ville vi begå en type II feil.

3.4 Many studies without population effect

3.4.1 Oppgave 8.

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

```
population <- rnorm(1000000, mean = 0, sd = 3)  
  
# Create data frames to store the model estimates
```



```

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)

# A for loop used to sample 1000 studies, each iteration (i) will draw a new sample
# from the population.

for(i in 1:1000) {

  # Draw a sample
  samp1 <- data.frame(y = sample(population, 8, replace = FALSE))
  samp2 <- data.frame(y = sample(population, 40, replace = FALSE))

  # Model the data
  m1 <- lm(y ~ 1, data = samp1)
  m2 <- lm(y ~ 1, data = samp2)

  # Extract values from the models
  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]

}

# Save the results in a combined data frame

results_null <- bind_rows(results_8, results_40)

# Calculate number of false positives

```

```

false_positives <- results_null %>%
  filter(pval < 0.05) %>%
  group_by(n) %>%
  summarise(sig_results = n() / 1000)

# Print results
print(false_positives)

```

```

# A tibble: 2 x 2
      n sig_results
  <dbl>     <dbl>
1     8     0.053
2    40     0.053

```

Fra resultatene o simuleringen over kan vi se at vi får det samme resultatet for begge utvalgsg størrelsene, nemlig 0,053. Dette vil si at ved å gjennomføre forsøkene gjennntatte ganger i den samme populasjonen vil 5,3 % av forsøkene gi falsk positivt resultat. Dette virker ved første øyekast rart fordi et større utvalg skal i utgangspunkte redusere sjansen for å få falsk positivt svar fordi man tester en større andel av populajsonen. Men hvis vi gjennomfører det samme forsøket gjennntatte ganger innenfor samme populajson vil vi ende opp med samme resultat gitt at hypotesen stemmer, fordi det også på denne måten vil teste en større andel av populasjonen.

Likevel vil få få falsk positivt resultat så mye som 5,3 % av tilfellene. Dette kan forklares med at utvalgene er tilfeldig, og i noen utvalg vil vi da ha grupper som ikke stemmer med gjennomsnittet av populasjonen. Dette vil alltid skje med mindre vi faktisk tester hele populasjonen. Dette er grunnen til at vi ikke bare kan lene oss på testresultatene, men at vi også må støtte opp med statistisk analyse.

4 Studiedesign

4.1 Indroduksjon

I denne oppgaven skal jeg analysere og sammenlikne fem forskningsartikler som undersøker samme tema. Tema jeg har valgt er *“Blockperiodisering og VO2max”*. Jeg skal forsøke å bruke QALMRI-metoden for å analysere og sammenlikne artiklene på en systematisk og god måte. QALMRI står for “Questions”, “Alternatives”, “Logic”, “Method”, “Results”, “Inferences”, og handler om å systematisk analysere hva forfatterne ønsker å finne ut av, hva de tror vil skje, logikken bak hypotesen, metodene de har brukt i forskningen, resultatene og hva de betyr (Brosowsky *et al.*, 2020).

Artiklene jeg har valgt er:

- *“Block training periodization in alpine skiing: effects of 11-day HIT on VO2max and performance”* - (Breil *et al.*, 2010)
- *“Block periodization of high-intensity aerobic intervals provides superior training effects in trained cyclists”* - (Rønnestad *et al.*, 2012b)
- *“Effects of 12 weeks of block periodization on performance and performance indices in well-trained cyclists”* - (Rønnestad *et al.*, 2012a)
- *“Training Prescription Guided by Heart Rate Variability Vs. Block Periodization in Well-Trained Cyclists”* - (Javaloyes *et al.*, 2020)
- *“No Differences Between 12 Weeks of Block- vs. Traditional-Periodized Training in Performance Adaptations in Trained Cyclists”* - (Almquist *et al.*, 2022)

Det skal sies at ikke alle artiklene jeg har valgt i denne oppgaven ser spesifikt på sammenhengen mellom bolkperiodisering og VO2max. De fleste av studiene ser på den generelle effekten bolkperiodisering har på prestasjonsevnen til utøvere. Ettersom VO2max er svært viktig for utholdenhetsidretter har jeg likevel valgt å fokusere på dette. Tester som går på andre egenskaper kommer jeg derfor ikke til å fokusere så mye på her.

4.2 Hoveddel

4.2.1 Grunnlag for studiene

Hovedfokuset til studiene jeg har valgt er å se på effekten av blokkperiodisert trening, med fokus på utholdenhet. Fire av artiklene ser spesifikt på hvordan effekten er for syklister, mens den siste tar for seg alpinister. (Rønnestad *et al.*, 2012b), (Rønnestad *et al.*, 2012a) og (Almquist *et al.*, 2022) sammenlikner i sine artikler effekten av blokkbasert trening (BP) med tradisjonell periodisering (TP). (Javaloyes *et al.*, 2020) sammenlikner effekten av BP med “*heart rate variability guided training*” (HRV). (Breil *et al.*, 2010) som forskjer på alpinister skiller seg også ut fra dette fordi de sammenlikner ikke BP med en annen form for periodisering. De forskjer utelukkende på hva en blokk med høyintensitetsintervaller (HIT) gjør med den aerobe kapasiteten til junior-alpinister.

(Breil *et al.*, 2010), (Rønnestad *et al.*, 2012b), (Rønnestad *et al.*, 2012a) har alle en hypotese om at BP vil gi en større forbedring i aerob kapasitet enn TP. Dette gi mening fordi det blir referert til studiene i artiklene. (Breil *et al.*, 2010) sine funn blir referet til i (Rønnestad *et al.*, 2012b), og (Rønnestad *et al.*, 2012a) referer igjen til sitt eget studie tidligere det samme året, og handler mye om å utvide forskningen de selv har gjort. Tanken bak BP er å lure kroppen til og styrkes raskere som følge av en relativt rask belastningsøkning i form av en høyintensitets treningsbolk på for eksempel en uke. Det er da viktig at man eterfølger dette med en rolig bolk slik at kroppen får tid til å restituere seg. Tanken er at en slik periodisering skal gi en raskere fremgang en TP hvor man fordeler de harde øktene jevnt utover treningsperioden.

(Almquist *et al.*, 2022) presenterer ingen klar hypotese, men viser til nyere forskning hvor resultatene viser at BP ikke gir noe særlig bedre resultat enn andre former for periodisering som TP. Motivasjonen bak denne forskningen og artikkelen virker å være å undersøke hva som faktisk stemmer av tidligere forskning - altså om BP er bedre enn TP, eller ikke.

(Javaloyes *et al.*, 2020) har ingen klar hypotese i sin artikkel men viser til annen forskning som viser at HRV kan gi en mer pålitelig formening om en utøvers evne til å trene. Ut fra utøverens HRV-score før trening kan man bestemme hva slags intensitet treningsøkten skal ha. Høy HRV-score = høy intensitet og motsatt. Dette kan altså være en måte å få optimalisert treningen for hver utøver individuelt ut fra dagsfor slik at treningen også blir optimal (Walker, 2024). Det blir ikke skrevet eksplisitt, men det virker som om de har en hypotese om at HRV kan være en bedre måte å periodisere treningen etter enn BP.

4.2.2 Sudiedesign og forskningsmetoder

Jeg synes det er vanskelig å si hva slags studiedesign som blir brukt i studiene. Alle studiene har som mål å finne ut om en form for periodisering er mer effektiv enn en annen. I de fleste studiene jeg har valgt er det BP som undersøkes med en kontrollgruppe som har TP, men unntak av studien til (Javaloyes *et al.*, 2020) som sammenlikner HRV med BP. Dette kan likne

på det (Hulley *et al.*, 2013) kaller “*Case Controll Studies*”, men passer samtidig ikke helt fordi det beskrives som en forskningsmetode som brukes for å finne ut hvordan en bestemt diagnose påvirker en gruppe, men “diagnosen” kan kanskje byttes ut med treningsmetode.

Studiene kan også minne om “*Randomized blinded trial*” som brukes for å finne ut om en behandling faktisk virker. Disse studiene blir blindet som vil si at deltakerne ikke vet om de er med i intervensjonsgruppen eller kontrollgruppen - derfor passer dette heller ikke helt. Kanskje er det designformen “*Randomiserte kontrollerte studier*” som passer best hvor man deler utvalget tilfeldig, slik at eventuelle ulikheter i gruppene også blir tilfeldig. Alle studiene deler i hvert fall utvalget i to grupper: intervensjonsgruppe (BP eller HRV i (Javaloyes *et al.*, 2020) sitt stude) og en kontrollgruppe (TP eller BP i (Javaloyes *et al.*, 2020) sitt studie).

Alle studiene har utvalg med godt trente utøvere. I studien til (Breil *et al.*, 2010) består utvalget av eliteutøvere på juniornivå. I de fire sykkelstudiene er utøverne alle godt trente syklistere som enten konkurrerer aktivt eller som har bakgrunn fra konkurransesykling. Ut ifra tallene som blir presentert for VO₂max er de fleste utøverne i studiene er på prestasjonsnivå 3 etter modellen til (Pauw *et al.*, 2013), og noen på nivå 2 (men dette er hovedsaklig alpinister hvor VO₂max ikke er like viktig som for syklistere). Noen utøvere er også på nivå 4 og 5 i følge (Almquist *et al.*, 2022).

I studiene til (Rønnestad *et al.*, 2012b) og (Rønnestad *et al.*, 2012a) meldte deltakerne seg frivillig for å delta på prosjektet, men alle var som sagt godt trente eller aktive syklistere. (Breil *et al.*, 2010) hentet deltakerne fra et nasjonalt treningssenter og (Almquist *et al.*, 2022) fra en lokal sykkelklubb. (Javaloyes *et al.*, 2020) skriver ikke hvor de hentet sine deltakere fra, men at de ble plukket ut på bakgrunn av kvalifikasjoner. Ingen av studiene oppgir noen utregning for statistisk styrke, ut fra utvalgene kan man anta at det ikke er meningen at de skal være representative for hele populasjonen. Forskningen i disse artiklene handler nok hovedsaklig om prestasjonsutvikling for aktive utøvere.

Alle studiene har forsøkt å standardisere så mye som mulig. Utøverne har fått instruksjoner i forkant av pretest om å unngå harde økter rett før testen, hva de kan spise og drikke på testdagen og liknende. I alle prosjektene blir det også gjort flere tester som VO₂max, sprint, laktatprofil, spenst og så videre, og i noen av prosjektene foregår testene over to dager. I alle prosjektene er derfor rekkefølgen på testingen også standardisert. Treningsopplegget mellom testene er selvfølgelig laget og standardisert av prosjektlederene ettersom det er effekten av disse de skal undersøke. Varigheten på treningsperiodene varierer en del. I studiene til (Breil *et al.*, 2010) og (Rønnestad *et al.*, 2012b) trener utøverne i 4 uker mellom testene, og i gruppen med BP har de først en hard uke med mye HIT etterfulgt av tre rolige uker med færre HIT-økter, mens TP hadde jevn treningsbelastning hele perioden. (Rønnestad *et al.*, 2012a) og (Almquist *et al.*, 2022) la opp treningen likt, men over 12 uker i stedet for fire og da med tre treningsblokker i BP-gruppen. (Javaloyes *et al.*, 2020) hadde et treningsopplegg som gikk over åtte uker hvor HRV-gruppen trente ut fra HRV-scoren de fikk hver morgen, og BP-gruppen fulgte et forhåndsbestemt blokk-opplegg.

Det har blitt brukt veldig like former for statistisk analyse i studiene. Utregning av standard avvik blir gjort i alle studiene og blir oppgitt som en del av resultatene. De fleste studiene med unntak av (Almquist *et al.*, 2022) understreker at det blir gjort en ANOVA-analyse for å sammenlikne gruppene. T-test blir gjennomført av (Breil *et al.*, 2010), (Rønnestad *et al.*, 2012b) og (Rønnestad *et al.*, 2012a), mens (Almquist *et al.*, 2022) og (Javaloyes *et al.*, 2020) bruker “*Shapiro-Wilks-test*”. Sistnevnte bruker også en “*Levene’s-test*”. Alle studeiene regner en p-verdi $< 0,05$ som statistisk signifikant.

4.2.3 Resultat og konklusjon

I samtlige studier var BP den formen for periodisering som gav mest framgang på VO2max. Spesielt i studiene til (Breil *et al.*, 2010), (Rønnestad *et al.*, 2012b) og (Rønnestad *et al.*, 2012a) kommer BP godt ut, og har signifikant resultat - noe TD ikke har. I studien til (Almquist *et al.*, 2022) er forskjellen vesentlig mindre. Her har BP fortsatt en fordel fremfor TP, men mye mindre enn resultatene fra de andre studiene sier. En mulig forklaring på dette i følge forfatterne kan være den store spredningen i prestasjonsnivå. I denne studien varierer deltakerne fra nivå 2-5 (Pauw *et al.*, 2013), så individuell framgang som følge av økning i treningsbelastning kan skjule eventuelle forskjeller mellom BP og TP. BP har også bedre effekt på VO2max. enn HRV i følge resultatene i studien til (Javaloyes *et al.*, 2020), men også her er forskjellene mindre enn de tre første studiene.

Studiene til (Breil *et al.*, 2010), (Rønnestad *et al.*, 2012b) og (Rønnestad *et al.*, 2012a) konkluderer med at BP *kan* være fordelaktig for å forbedre aerob kapasitet. Spesielt (Breil *et al.*, 2010) er positive til BP, men dette er også en av de korteste studiene, og også den studien som tydeligst hadde fokus på VO2max. Selv om de er positive til BP stiller de fortsatt et spørsmål: stegn til hvor viktig VO2max er for alpinister - selv om tidligere studier viser at alpinister har høyere VO2max enn gjennomsnittet i populasjonen. (Rønnestad *et al.*, 2012b) og (Rønnestad *et al.*, 2012a) viser som sagt også til at disse studiene tyder på at BP kan være fordelaktig i forhold til TP, men er tydelige på at det er vanskelig å si hva som er den beste måten å legge opp treningen på for syklistene. (Almquist *et al.*, 2022) konkluderer med at deres data ikke støtter hypotesen om at BP er en bedre metode for trente syklistene enn TD. Resultatene viser at det er litt forskjell på treningsutbyttet, men at den ikke er signifikant. Studien til (Javaloyes *et al.*, 2020) skiller seg litt ut fra resten i og med at den har som mål å finne ut om HRV er et valid og reliabelt verktøy for å periodisere treningen. Derfor er det ikke fokus på hva som er mest effektivt av HRV og BP, men de konkluderer med at HRV er et godt hjelpemiddel og at man kan stole på det i treningsplanleggingen.

4.3 Avslutning

I denne oppgaven har jeg sett på fire forskjellige studier som handler om blokkperiodisering og VO2max. Jeg har forsøkt å analysere artiklene til disse studiene og se på hva slags studiedesign

de har brukt, hva slags statistiske metoder de har brukt og hva forskerne har funnet ut og konkludert med i studiene. Det siste punktet var litt vanskelig ettersom det bare var et av studiene jeg fant som forsket spesifikt på BP sin effekt på VO2max. Resten av studiene hadde mer generelt fokus på prestasjonsforbedring innenfor sykling hvor VO2max bare var en av mange variabler de forsket på. Jeg synes likevel det var spennende å se hvordan (Rønnestad *et al.*, 2012a) og (Almquist *et al.*, 2022) fikk forskjellige resultat. Disse studiene var veldig like, og Rønnestad var med i begge studiene, så det var interessant at resultatene ble såpass ulike. Alt i alt synes jeg alle artiklene var gode, og de virker svært seriøse, så det at resultatene ble ulike tyder bare på at det trengs mer forskning på feltet.

5 Analyse av gjentatte målte eksperiment

5.1 Effekten av treningsvolum på hypertrofi og styrke

5.1.1 Introduksjon

Menneskekroppen er relativt god til å tilpasse seg miljøet den lever i eller blir utsatt for. Ved økt treningsbelastning vil kroppen vår, særlig muskelen våre over tid tilpasse seg den økte belastningen ved å øke styrke og volum. Hvordan vi kan gjøre dette på mest mulig effektiv måte er ofte målet innenfor treningsfysiologisk forskning. Det er mange faktorer som spiller inn på hvordan og hvor raskt musklene våre tilpasser. Treningsvolum, antall sett, antall repetisjoner og pause mellom sett vil påvirke effekten av treningen. Det som vanligvis anbefales for hypertrofi er for eksempel moderat til høy motstand, høyt volum og korte pauser mellom sett. For å øke styrke er det vanlig å øke motstand, redusere volum og ha lengre pauser (Anon, 2009) (Schoenfeld *et al.*, 2016). Dette har trolig sammenheng med at trening med få repetisjoner til utmattelse kan bidra til bedre muskelaktivering (Ruple *et al.*, 2023).

Når det gjelder treningsvolum kan det se ut som at dette har en større betydning for hypertrofi enn for styrke. Ved trening av maksimal styrke vil man få stor effekt av relativt lavt treningsvolum. Økt treningsvolum vil føre til enda større økning i styrke, men ikke proporsjonalt i forhold til treningsvolum. For hypertrofi er det derimot en mer proporsjonal utvikling i forhold til treningsvolum (SCHOENFELD *et al.*, 2019). Forkning gjort på godt trente idrettsutøvere tyder også på at større treningsvolum i styrketrening gir større økning i styrke, selv om dette ikke alltid er det lureste for konkurrerende utøvere på grunn av totalbelastningen (Naclerio *et al.*, 2013).

Som sagt gir også lavere treningsvolum gode resultater, og selv om det antageligvis er størst effekt med høyt treningsvolum er det i noen tilfeller er det vanskelig å bestemme hva som gir best resultat av høyt og lavt treningsvolum (Mitchell *et al.*, 2012). En metastudie fra 2019 foreslår at en enkelt set med 6-12 repetisjoner med motstand fra 70-85 % 1RM, 2-3 ganger per uke med høy intensitet (til utmattelse) kan over en periode på 8-12 uker gi en signifikant økning av maksimal styrke hos trente personer (Androulakis-Korakakis *et al.*, 2019).

Målet med denne studien er å se hvordan forskjellig treningsvolum påvirker muskelhypertrofi og muskelstyrke hos relativt lite trente personer. Ut fra tidligere forskning som jeg har presentert her vil det være naturlig å anta at trening med høyt volum vil gi bedre resultat enn trening med lavt volum. Det vil også være naturlig å anta at vi ser en større effekt på hypertrofi enn på muskulær styrke.

Table 5.1: Deltakeroversikt

	Age (years)	Stature (cm)	Body mass (kg)
Female n = 18	22 (1.3)	167.7 (6.9)	64.4 (10.4)
Male n = 16	23.6 (4.1)	182.9 (5.9)	75.8 (10.7)

5.1.2 Metode

5.1.2.1 Etisk godkjenning

Alle deltakerne ble informert om potensiell risiko og ubehag studien kunne medføre og gav informert bekreftelse og godkjenning av dette i forkant av studieopptaket. Studiedesign var forhåndsregistrert (ClinicalTrials.gov Identifier: NCT02179307), og godkjent av den lokale etikk-komiteen ved Høgskolen i innlandet Lillehammer, avdeling for idrettsvitenskap (no. 2013-11-22:2) og alle prosedyrende ble gjennomført i henhold til *Helsinki-erklæringen*

5.1.2.2 Deltakere

41 menn og kvinner var med i denne studien. Alle deltakerne måtte være ikke-røykere mellom 18 og 40 år. For å kunne observere effekten av treningsintervensjonen best mulig kunne ikke deltakerne ha en treningshistorie med mer enn én økt med styrketrening per uke de siste 12 månedene opp til intervensjonen. Deltakerne kunne ikke ha redusert muskelstyrke som følge av tidligere eller nåværende skade av samme årsak, og de kunne ikke gå på faste medisiner da også disse kunne påvirke trenignseffekten.

7 av deltakerne ble ekskludert fra data-analysen på grunn av at de ikke gjennomførte 85 % eller mer av de planlagte treningsøktene under intervensjonen. Årsakene var; ubehag eller smerte i underekstremitetene under trening (n = 5), skade urelatert til studiet (n = 1) og mislyktes i å følge protokollen (n = 1).

Alle deltakerne meddelte tidligere erfaring med idrettsaktivitet (f.eks. lagsport, langrenn og turn). 20 av deltakerne meddelte at de jevnlig drev med fysisk aktivitet eller trening når de ble med i studien (ca. to ganger i uken), og 10 av disse utførte sporadiske styrketreningsøkter, men ikke mer enn en økt per uke.

Deltakerne er beskrevet i Table [5.1](#)

5.1.2.3 Studie design

5.1.2.3.1 Mål av muskelmasse (mål av regional lean mass)

Muskeltverrsnitt av knestrekkerne (MTK; vastus lateralis, medialis, intermedius og rectus femoris) ble målt før og etter treningsintervensjonen ved hjelp av “*magnetic resonance imaging*” (MRI) i samsvar med produsentens protocol (S-Scan, Esaote Europe B.V., Maastricht, the Netherlands). Bildene ble analysert blindt av den samme teknikeren ved hjelp av OsiriX (v.5.6, Pixmeo Sarl, Bernex, Switzerland). Tverrsnittet ble tatt med samme avstand til kneledet og ved bruk av minst fire sammenhengende bilder (5 mm tykkelse, 10 mm atskillelse) hos alle deltakerne. Kroppssammensetning hos deltakerne ble målt før og etter intervensjonen ved hjelp av “*Dual-energy x-ray absorptiometry*” (DXA) (Lunar Prodigy, GE Healthcare, Oslo, Norway), også dette i henhold til protokoll. I forkant av MRI- og DXA-målingen ble deltakerne bedt om å faste i minimum to timer før testing, og unngå hard fysisk aktivitet 48 timer før testing. Etter siste treningsøkt under intervensjonen var det to dager til post-målingene med MRI og DXA.

5.1.2.3.2 Mål av maksimal styrke

Testing av maksimal styrke i knestrekkerne ble gjort med en repetisjon maks (1RM) med etbeins beinpress. Test-økta ble standardisert for alle deltakerne med et spesifikt oppvarmingsprogram bestående av 10, 6 og 3 repetisjoner med belastning på 50, 75 og 85 % av forventet maksimal styrke. Etter dette ble 1RM funnet ved å gradvis øke motstanden i beinpress til deltakeren ikke klarte å fullføre bevegelsen i øvelsen på grunn av for høy vekt. Den høyeste vekten registrert med fullført bevegelse ble registrert som 1RM. Alle deltakerne fikk seks forsøk.

5.1.2.3.3 Treningsprotokoll

Intervensjonen bestod av 12 ukers standradisert styrketrening for hele kroppen. Alle øvelsene for bein ble utført med et bein om gangen for å få best tilpasset treningsvolum. Alle deltakerne utførte styrketrening som bestod av et enkelt sett (single set protokoll) og tre sett (multiple set protokoll). Hvilket av beina til deltakerne som skule brukes til hvilken protokoll ble tilfeldig trukket på forhånd.

Oppvarmingen før alle treningsøktene bestod av følgende; 5 min sykling på ergometersykel, etterfulgt av 10 repetisjoner av hver av kroppsvektøvelser (push-ups med tilpasset motstand for deltakernes nivå, sit-ups, rygg-hev og knebøy), og til slutt et sett med 10 repetisjoner av hver øvelse i økta med 50 % av 1RM.

Styrkeøvelsene for bein ble utført i følgende rekkefølge; etbeins beinpress, beincurl og knestrek. Øvelsene ble utført med et sett (single sets) eller tre sett (multiple sets) per øvelse. Single sets ble utført mellom andre og tredje sett i multiple sets protokollen. Deltakerne gjorde det samme med overkroppsøvelsene en hånds benkpress, pull-down og enten skulder-press eller sittende roing (annenhver gang) som en del av et utvidet forskningsprosjekt. Pauser mellom sett var på 90-180 s. Treningsintensiteten økte gradvis gjennom intervensjonen og startet med 10RM i uke 1-2, 8RM i uke 3-6 og 7RM uke 7-12. Treningsvolum økte også i løpet av intervensjonen ved at det ble flere økter per uke i snitt. Uke 1,3 og 5 hadde to økter og uke 2 og 4 hadde tre

økter. Uke 6, 7 og 8 hadde alle tre økter, to økter i uke 9, tre økter i uke 10 og 11 og den 12 og siste uka hadde to økter. I ukene med tre økter ble en av øktene utført med litt redusert vekt (90 % av forrige økt med samme antall repetisjoner). Øktene med maksimal innsats ble atskilt med minimum 48 timer, og de submaksimale øktene (90 %) ble atskilt med minimum 24 timer. For å bidra til best mulig restitusjon fikk utøverne en standardisert drikk som inneholdt 0.15 g kg⁻¹ protein, 11.2 g kg⁻¹ karbohydrater and 0.5 g kg⁻¹ fett.

For å tilpasse studien til deltakernes hverdag kunne noen økter gjennomføres uten oppsyn, men i gjennomsnitt ble 91 % av øktene gjennomført med oppsyn. for å holde oversikt over øktene uten oppsyn måtte deltakerne føre detaljerte logg som gjennom intervju ble sjekket av forsker-teamet sammen med deltakeren.

5.1.2.4 Dataanalyse og statistikk

Alle data som blir presentert i denne rapporten er oppgitt i gjennomsnitt og standard avvik (mean (SD)) med mindre noe annet er spesifisert. For å se på hvordan forskjellig volum påvirker hypertrofi og styrke har vi regnet ut den gjennomsnittlige forskjellen i endring mellom singel-set-protokoll (SSP) og multiple-set-protokoll (MSP). Disse dataene blir også visualisert i modeller i resultatdelen. Modellene viser hva slags utvikling deltakerne hadde fra SSP og MSP.

Testing mot null-hypotesen om at det ikke vil være noen forskjell på treningseffekten fra SSP og MSP ble gjort i R ved hjelp av lme4-pakken (Bates *et al.*, 2015). Grensen for statistisk signifikans ble satt til $\alpha = 0.05$. All data-analyse ble gjort i R (Anon, n.d.).

5.1.3 Resultat

Av 41 deltakere i studiet gjennomførte 34 hele treningsopplegget og hvis resultater er med i følgende beregninger

5.1.3.1 Hypertrofi

Etter 12 uker med styrketrening har det alt i alt vært en økning av muskelmasse og som forventet har det vært en større økning i muskelmasse ved bruk av flere sett under treningen (MSP) enn et enkelt sett (SSP). Den gjennomsnittlige forskjellen i muskelmasseendring i lår mellom sett-protokollene var 122.895g CI: [8.6,237.0], p -value =0.036, $t_{33} = 2.19$) (95% konfidensintervall).

Tabellen Figure 5.1 viser hva slags utvikling hver enkelt deltaker har hatt i løpet av intervensjonen og forskjellen på økning av muskelmasse mellom SSP og MSP. Her kan vi se at de fleste har fått en klart større økning i muskelmasse ved MSP enn SSP. Vi kan også se at for noen

deltakere er det relativt liten forskjell i protokollene og for noen er det faktisk en klart mindre økning i muskelmasse ved MSP enn SSP.

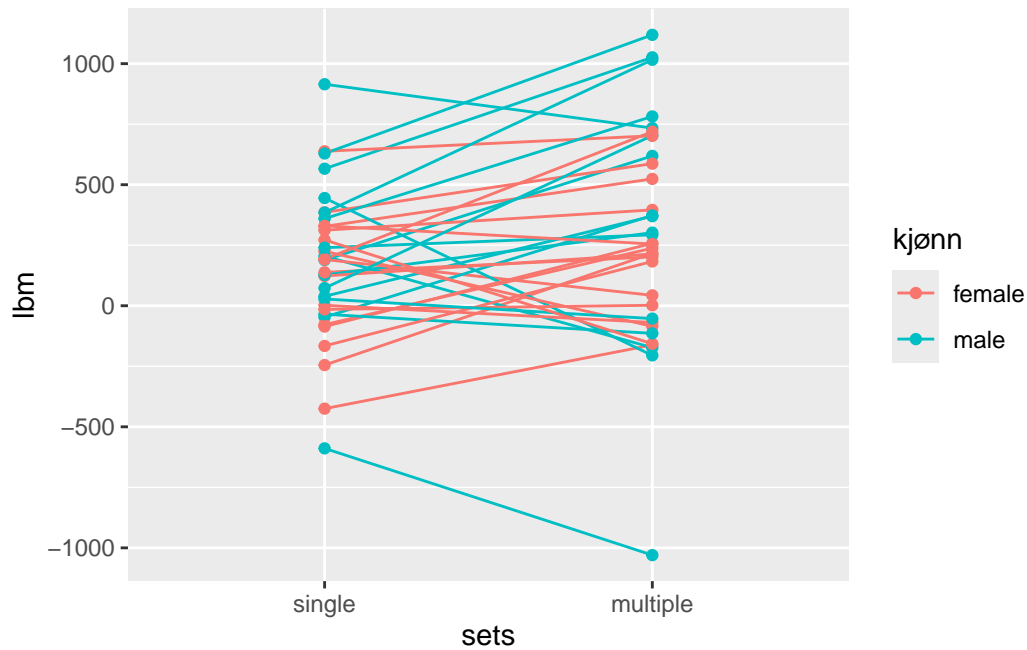


Figure 5.1: Forskjell på endring av muskelmasse under intervensjonen mellom SSP og MSP hos mannlige og kvinnelige deltakere

5.1.3.2 Maksimal styrke

Vi kan se en liknende tendens da det gjelder muskelstyrke. Den gjennomsnittlige forskjellen i muskelstyrkeendring i beinpress mellom sett-protokollene var 7.895kg CI: [1.1,14.6], p -value = 0.025, $t_{30} = 2.36$ (95 % konfidensintervall). Også her er det i gjennomsnitt en større økning av 1RM ved MSP enn SSP. Men da vi ser på Figure 5.2 kan vi også her se at resultatene varierer hos deltakerne. Her er det færre som har en klar fordel av MSP selv om det er en tendens til det. Også her kan vi se at det er noen som opplever en større økning ved SSP enn ved MSP.

5.1.4 Diskusjon

I dette studiet fikk deltakerne en større økning i muskelmasse ved å følge treningsprotokoll med flere sett (MSP) enn med et enkelt sett (SSP). Dette stemmer over ens med anbefalingene fra (Anon, 2009) og funnene til (Schoenfeld *et al.*, 2016). Det som er overraskende i denne studien er at det ikke var stor nok forskjell til å si at funnet er statistisk signifikant med en p -verdi på

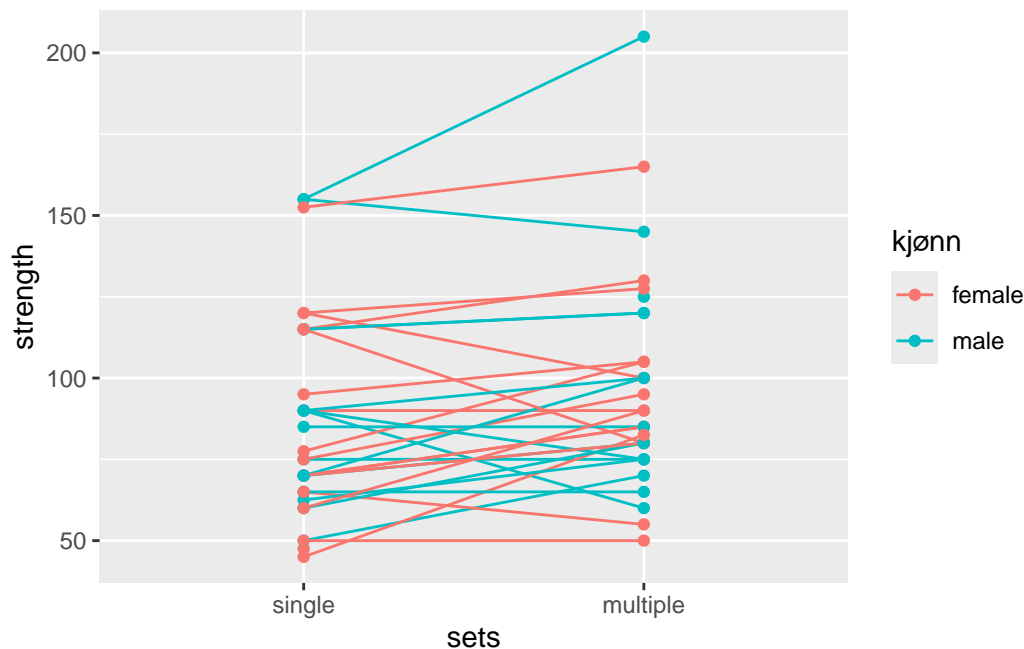


Figure 5.2: Forskjell på endring i styrke i legpress under intervensjon mellom SSP og MSP hos mannlige og kvinnelige deltakere

0.036. Fra tabellen kunne vi også se at noen deltakere ikke bare responderte like godt på SSP som MSP, men faktisk bedre. Selv om vi ser en tendens til at det i gjennomsnitt gir bedre effekt av MSP enn SSP kan vi altså ikke konkludere med dette ut fra denne studien.

Vi får litt samme resultat i resultane for muskelstyrkeøkning (1RM). Også her viser er det i gjennomsnitt en større økning i styrke ved MSP enn ved SSP, men p -verdien er enda lavere på 0,025, Så igjen er ikke funnene signifikante. På styrke kan vi se at det er enda mindre forskjell på effekten av SSP og MSP. Dette stemmer overens med studien til (Mitchell *et al.*, 2012) om at det kan være vanskelig å si hva som gir best effekt av høyt og lavt treningsvolum.

Selv om vi ser en tendens i denne studien til at det totale utbyttet av MSP er større enn SSP, er ikke forskjellen stor nok til å motbevise null-hypotesen om at det ikke er noen forskjell. Den tendensen vi kan se stemmer likevel med studien til (SCHOENFELD *et al.*, 2019), hvor de sier at treningsvolum er viktigere for hypertrofi enn for muskelstyrke. Så kan vi likevel bruke funnene? (Naclerio *et al.*, 2013) sier i sin studie at selv om økt treningsvolum gir noe bedre effekt er ikke dette alltid anbefalt for konkurrerende utøvere på grunn av det høye totalvolumet. Kanskje kan en liknende anbefaling gjelde for utrente personer slik som deltakerne i denne studien. Denne studien tyder på at lavt treningsvolum ikke har noen signifikant dårligere effekt en høyt treningsvolum. At et sett per øvelse gir nesten like godt resultat som tre vil være godt nytt for personer som sliter med motivasjon eller som ikke liker styrketrening. Dette kan være svært positivt i et folkehelseperspektiv for å motivere flere til

å trene og holde seg i form med litt trening. Dette stemmer også med meta-studien gjort av (Androulakis-Korakakis *et al.*, 2019).

Vi kan altså ikke komme med noen klar konklusjon i denne studien. Det var noe overraskende at hypotesen ikke stemte, spesielt med tanke på hypertrofi. Det kan tenkes at det er en forskjell på trente og godt trente i en slik situasjon så mer forskning innenfor dette feltet er nok nødvendig. Selv om hypotesen ikke stemte kan det som sagt likevel være positivt med tanke på folkehelse.

6 Labrapport; Ekstraksjon og analyse av muskelprotein

6.1 Introduksjon og bakgrunn

I jakten på å bedre forstå de molekylære reaksjonene som er involvert i reguleringen av transkripsjon og translasjon ved trening, ernæring, sykdom og aldring, har også interessen og bruken av Western blot (WB) innenfor treningsfysiologi økt. WB har flere bruksområder for undersøkelse av regulatoriske molekylære reaksjoner, som for eksempel kvantifisere proteinmengde og protein-protein-interaksjoner, som er avgjørende for å kunne konstatere fysiologiske adaptasjoner til trening (Bass *et al.*, 2016). Prosessen i WB deles inn i følgende steg:

- 1) Ekstraksjon av protein fra en sammensatt mikstur av intracellulær og ekstracellulær protein
- 2) Kvantifisering av proteinkonsentrasjon og elektroforese av protein
- 3) Overføring til en stabil membran med høy binding av proteiner
- 4) “Blokking” av membranen for å redusere bindingen av ikke-ønskede molekyler
- 5) Binding av antistoffer til de spesifikke proteinene man ønsker å undersøke
- 6) Binding av et sekundært antistoff, med en markør, til det første antistoffet
- 7) Utvikling og gjenkjenning av markøren
- 8) Kvantifisering av de resulterende båndene med hjelp av densitometriprogramvare

Tidligere har WB kun blitt assosiert med selve prosessen med å overføre proteinet fra en gel til en mer stabil membran, mens nå referer man til WB når man skal forklare hele prosessen (Bass *et al.*, 2016).

Formålet med vårt eksperimentet er å benytte seg av WB for å se på den totale mengden av p70S6K (t-p70) og mengden fosforisert p70S6K fra muskelvev. Med dette kan vi si noe om fosforyleringsnivået av p70S6K i muskulaturen, som reflekterer hvor stor andel av p70S6K som er aktiv i forhold til den totale proteinmengden i vevet. Muskelprøvene er hentet fra høyre og venstre ben til ulike testpersoner, der alle har gjennomgått styrketrening av de ene beinet ved ulike tidspunkt før selve biopsien:

- Hennie (HE) = 10 min før biopsi
- Lars (LØ) = 30 min før biopsi
- Trond (T) = 60 min før biopsi

Prøven ble homogenisert forut for WB-analysen, som vil si at vi har frigjort proteinene gjennom mekanisk eller kjemisk nedbryting av cellemembranen. Selve protokollen og hvordan vi har utført homogeniseringen og WB er beskrevet under, samt presentasjon og diskusjon av resultatene.

6.2 Metode

6.2.1 Homogenisering av muskelprøver

Vi benyttet oss av fryst vev av m.vastus lateralis som var blitt ekstrahert fra både det høyre og venstre låret til en av prøvedeltakerne. Vekten på de ulike prøvene var:

Høgre bein (HE-R): 14,0 g Venstre bein (HE-L): 14,5 g Til homogeniseringen benyttet vi oss av en plastpistill, og tilsatte 600 μ l med lysat-buffer som var satt sammen av 594 μ l t-per og 6 μ l pink for å visualisere om vevet var fullstendig homogenisert.

For å måle proteinkonsentrasjonen i prøven vår, brukte vi Bradford Assay-metoden. Denne metoden bruker et fargestoff, Coomassie Brilliant Blue G-250, som vil binde seg til proteiner og skape en reaksjon der prøven vil skifte farge til blå i proporsjon med proteinkonsentrasjonen mengden protein. For å kvantifisere proteinkonsentrasjonen laget vi en standardkurve med bovint serumalbumin som referanse, slik at vi kunne sammenligne blåfargen i vår prøve med kjente proteinkonsentrasjoner. Vi måler deretter, med hjelp av et spektrofotometer, hvor mye av prøvene som blir absorbert ved 595 nm for å kunne bestemme om prøvene våre inneholder lave eller høye konsentrasjoner av proteiner (Noble & Bailey, 2009). Protokollen under er en generell fremgangsmåte for homogenisering av vev, og avvik og sentrale anmodninger er kommentert under de ulike punktene.

Definisjoner: - BSA = Bovint serumalbumin, et protein som blir utvunne fra blodplasmaet til storfe. Ofte brukt som standard for proteinkonsentrasjon.

6.2.1.1 Utstysliste

- Microfuge tubes (1.5 ml Eppendorf)
- Plastic pestle
- Centrifuge (capable of 10 000 g and 4°C)
- Ice
- Lysis buffer (e.g. Hepes-buffer, Ripa-buffer)

6.2.1.2 Protokoll

Kursiv skrift: Sentrale kommentarer fra selve utføringen.

Fet skrift: Mulige avvik

1. Prepare the tissue by dissection away any connective tissue and blood. If you are working with wet tissue, this step should be done prior to freezing. Freeze dried tissue can be dissected in room temperature under the microscope. Use 10-50 mg wet-weight tissue or 2-10 mg freeze-dried tissue
2. Add protease/phosphatase inhibitors to the ice-cold lysis buffer (10 l/ml). **Pipetering mulig feilkilde i t-per, og i ene prøven (høyre)**
3. Keep the sample on dry ice or cooling block until you add the ice cold lysis buffer of choice. Use 20 l/mg wet-weight or 80 l/mg dry weight. *Dry: 2 mg → 160 l, 5 mg → 400 l, 10 mg → 800 l*
4. Quickly disrupt the tissue by hand using the plastic pestle until there are no visible pieces.
5. Keep the sample on ice, vortex or rotate the sample for 20-60 min. *Varighet av roteringen var 33min.*
6. Spin the sample 10 min, 10 000 g, 4°C
7. Carefully remove the supernatant to a new tube without disrupting the pellet.
8. Aliquot the sample to a tube for protein concentration determination (1:10 dilution, 4 l to 36 ul ddH₂O. Keep a known volume for later normalization of the protein concentration (e.g. to 3 ug/ul) keep on ice. *Myo homogenat ofte utenfor standard. Bruker 1:6 fortynner for å komme innenfor standard. 12 l prøve og 60 l vann. Pipetering gikk nokså bra også i dette leddet, liten sannsynlighet for feilkilde, men skal aldri utelukkes.*
9. Determine protein concentration with Bradford Assay. 10 l sample + 250 l reagent. Make BSA (bovint serumalbumin) standards according to Thermo Sci guidelines (ref. 23209). *Igjen ble pipeteringen godt utført, liten sannsynlighet for feilkilder, men skal ikke utelukkes.*

6.2.1.3 Mulige avvik:

- 7.2.1.2, Punkt 2: Feil ved pipetering

6.2.2 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 mTOR-signalveien.

Kursiv skrift: Kommentarer fra utføringen **Fet skrift:** Mulige avvik

6.2.2.1 Løsninger

Table 6.1: **Løsninger**

Buffer.Solution	Komponenter	Konsentrasjoner..mM.	Masser..g.
TBS 1L 10x, pH 7.6 (HCl adjust)	Tris BaseNaCl	20 mM x 10 = 200 mM 137 mM x 10 = 1.37 M	0.2 x 121.1 = 24.22 g 1.37 x 58.44 = 80.06 g
Running buffer 1L 10x, pH 8.3	Tris BaseG- lycinSDS	25 mM x 10 = 250 mM 192 mM x 10 = 1.92 M 3.5 mM x 10 = 35 mM	0.25 x 121.1 = 30.28 g 1.92 x 75.07 = 144.1 g 0.035 x 288.38 = 10.09 g
Transfer buffer* 1L 10x	Tris BaseGlycin	25 mM x 10 = 250 mM 192 mM x 10 = 1.92 M	0.25 x 121.1 = 30.28 g 1.92 x 75.07 = 144.1 g

- 1 x Transfer buffer = 100 ml TB stock x 10 + 100 ml methanol + 800 ml dH₂O

6.2.2.2 Sample preparation

- Muscle samples should be prepared in an appropriate sample buffer
- Dilute samples to a final concentration of 1.5-2 g/l in Laemmli buffer (Bio-Rad)
- Boil samples for 5 min, 95°C (denaturation). Cool to room temp. prior to loading. Spin down condensate

6.2.2.3 Sample loading, gel electrophoresis

- Assemble electrophoresis chamber on ice, add buffer to limit-mark (*Det ble brukt Running buffer*)
- Add protein ladder/standard (5 l) (*Proteininstige: Precision Plus Protein Dual Color Standards, 500 µl #1610374*), (*Pipetering gikk fint*)
- Load samples according to pre-written load scheme (max. Loading capacity in GTX precast gels: 30 l) (**Pipeteringsfeil i brønn 5: trakk prøven tilbake i pipetten en gang. Fikk alt ut igjen, men kan ha blitt blandet.**), (*Brønn 2-6 er p-p70 og brønn 9-13 er t-p70.*)

1	2	3	4	5	6	7	8	9	10	11	12	13	14
M	H-L	H-R	T-L	T-R	L-R	M	M	H-L	H-R	T-L	T-R	L-R	M

- M = Markør
- H-L = Hennie - Venstre ben
- H-R = Hennie - Høyre ben
- T-L = Trond - Venstre ben
- T-R = Trond - Høyre ben
- L-R = Lars - Høyre ben

- Move electrophoresis chamber into 4 C. Top up buffer. Put on top lid.
- Run gels: constant voltage, 30 min, 300 volt

6.2.2.4 Protein transfer (blotting)

- Disassemble gels and place them in transferbuffer 15-30 min
- Prepare sandwich:
 - Put sponges in dH₂O to get rid of bubbles
 - Put membranes (cut in upper left corner) in methanol to activate (5-10 min on shaker) *Klipte hjørnet på membran ved brønn 18.*
 - Place sandwich in assembly tray – black side down. Squeeze dH₂O from sponges and place them in transferbuffer – remove all bubbles
 - Wet 2 filter papers in buffer, place them on top of gel and gently remove them together. Place filterpapers with gel on top of sponge. Remove bubbles
 - Put membrane on top of gel. Careful with direction, membranes should be marked upper left corner. Remove bubbles.
 - Place the last sponge on top, close the sandwich
 - Overføring til sandwich gikk fint, lav risiko for feilkilde.
- Place the sandwich in transfer tray
- Run transfer: Constant ampere/current 300 mA, 3 h. **Brukte hurtigprogram: 100V i 30 min.**

6.2.2.5 Stain for total protein and cut membrane

- Mix destain solution with methanol 1:1 (*30:30 ml*)
- Mix Eraser solution with methanol 1:1 (*30:30 ml*)
- Rinse the membrane quickly in dH₂O
- Add MemCode Sensitizer to the the membrane, place on shaker for 2 min. Decant solution,

reuse 1 time

- Add MemCode Reversible stain, on shaker 1 min. Decant. Reuse 1 time.
- Destain by adding MemCode destain to membrane, rinse quickly 3 times
- Add methanol/destain solution, on shaker, 5 min- Rinse with dH₂O 4 times#
- Image membrane
- Cut the membrane. If multiple membranes, number each
- Proceed with Eraser → Add eraser/methanol solution to the membrane on shaker, 10 min*
- Rinse the membrane in dH₂O 4 times. Store in TBS until cut
- Cut membranes according to protein weight and/or wells used

6.2.2.6 Blocking and primary antibody

- Block the membrane in 2,5 % milk in TBS-T for 1 h at room temp
- Decant blocking solution, rinse in TBS
- Add primary antibody (AB) in BSA/milk solution (5% in TBS-T), incubate over night in 4 C

Brukte 5 % milk i TBS-T

6.2.2.7 Antibody wash and secondary antibody

Denne delen ble gjort av Vilde, og vi har derfor ikke oversikt over mulige avvik her, men har spesifisert muligheter for avvik under

- Wash with TBS-T 2x1 min + 3x5 min (**Unøyaktig vasking vil kunne føre til bakgrunnsstøy.**)
- Add secondary AB in 2,5% milk with TBS-T for 1 h at room temp
- Wash with TBS 4x5 min (**Unøyaktig vasking vil kunne føre til bakgrunnsstøy.**)
- Perform ECL
 - Mix working solution (Thermo super signal ...)
 - Incubate membrane for 5 min at room temp
 - Cover blot in clear plastic wrap
- Collect image on Gel Doc

6.2.2.8 Mulige avvik:

- Punkt 7.2.2.3: Pipeteringsfeil i brønn 5: trakk prøven tilbake i pipetten en gang. Fikk alt ut igjen, men kan ha blitt blandet.
- Punkt 7.2.2.4: Brukte hurtigprogram: 100V i 30 min.
- Punkt 7.2.2.7: En annen person utførte vaskingen - derfor ikke kontroll over dette.

6.3 Resultater

Vi benyttet oss av WB-analyse for å kvantifisere total p70S6K og fosforylert 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 Tabell 1 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 høyre bein hos testperson T (1.550) og lavest i venstre bein hos testperson LØ (0.920).

Table 6.3: Oppsummering av absolutte verdier til de individuelle testpersonene

Sub.	Bein	M.blot..p.p70	M.blot..t.p70	p.p70..raw	t.p70..raw	p.p70..norm	t.p70..norm	Fosf.nivå
HE	Venstre	16530.472	15516.16	4276.447	4276.447	0.2587	0.2756	0.939
HE	Høyre	10135.430	13009.67	4238.276	4238.276	0.4182	0.3257	1.284
T	Venstre	11187.966	12917.48	4457.276	4457.276	0.3985	0.3449	1.155
T	Høyre	7012.388	10868.74	4148.447	4148.447	0.5916	0.3816	1.550
L	Høyre	14791.037	13617.43	4249.861	4249.861	0.2872	0.3122	0.920

Den gjennomsnittlige aktiviteten til p70S6K på tvers av gruppen var 1.168 ± 0.248 , med et høyere nivå i høyre bein (1.426 ± 0.147) sammenlignet med venstre bein (0.986 ± 0.090). Variasjonen i fosforyleringsnivået var høyest på gruppenivå med en CV på 21.23%, og lavere når høyre (CV = 10.31%) og venstre bein ble analysert separat (CV = 9.13%).

Table 6.4: **Gjennomsnittlige nivåer av total og fosforylert p70S6K, samt forskjellen mellom disse, i høyre og venstre bein. Presenteres som gjennomsnitt (gj.snitt), standardavvik (SD) og coefficient of variation (CV)**

Parameter	Hele.gruppen..n...20.	Høyre.bein..n...10.	Venstre.bein..n...10.
t-p70 (norm.)	Gj.sn: 0.3280SD: 0.0366CV = 11.16%	Gj.sn: 0.3537SD: 0.0396CV = 11.19%	Gj.sn: 0.3127SD: 0.0298CV = 9.53%
p-p70 (norm.)	Gj.sn: 0.3908SD: 0.1174CV = 30.04%	Gj.sn: 0.5049SD: 0.1227CV = 24.30%	Gj.sn: 0.3148SD: 0.0704CV = 22.37%
Fosforyleringsnivå	Gj.sn: 1.168SD: 0.248CV = 21.23%	Gj.sn: 1.426SD: 0.147CV = 10.31%	Gj.sn: 0.986SD: 0.090CV = 9.13%

6.4 Diskusjon

Intensitetsverdiene gir oss et bilde av p70S6K-aktiviteten i prøvene, og selv om absolutt kvantifisering ikke ble utført, indikerer et høyere fosforyleringsnivå en økt aktivitet av mTOR-signalveien, som spiller en viktig rolle for proteinsyntesen. Det høyere fosforyleringsnivået i høyre bein (1.4261) sammenlignet med venstre (0.986) kan forklares ved at kun det ene beinet ble trent før biopsien. Dette viser til den akutte aktiveringen av mTOR-signalveien som respons av styrketrening, der den største aktiveringen ser ut til å forekomme kort tid etter trening, med en gradvis reduksjon over tid. Et høyere fosforyleringsnivå reflekterer potensielt et større kapasitet for proteinsyntese i det trente beinet. Nivå over 1.0 i flere prøver kan tyde på høy aktivering av p70S6K, men kan også være et resultat av tekniske faktorer.

Høyere variasjon i fosforylert p70S6K (CV = 30.04 %) sammenlignet med total p70S6K (CV = 11.16 %) kan skyldes biologisk variasjon i mTOR-signalveien mellom testpersoner. Samtidig kan tekniske faktorer som variasjoner i prøvehåndtering ha bidratt til dette. Lavere CV i høyre (10.31%) og venstre bein (9.13%) indikerer mer konsistens når prøvene analyseres separat.

6.5 Konklusjon

Den asymmetriske aktiveringen av mTOR-signalveien mellom høyre og venstre bein kan forklares ved at kun det ene beinet ble trent før biopsien. Resultatene viser hvordan styrketrening kan gi en akutt økning i fosforyleringsnivået av p70S6K, med størst aktivering observert kort tid etter trening. Resultatene burde likevel tolkes med omhu, da tekniske faktorer som bakgrunnsstøy på blottet, feil pipettering, ujevn proteinoverføring, manglende kalibreringsskurve kan ha påvirket intensitetsmålingene.

7 Filosofihistore

7.1 Oppgave 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.

Induksjon handler om å predikere et resultat ut fra tidligere erfaringer. Dette er det vi kaller uniformitetsprinsippet – «fremtiden vil være som fortiden». For eksempel at de første fem eggene i pakken var gode, så da må det sjette også være bra. Uforutsette ting kan skje. Det sjette egget kan være eldre enn de andre i pakken, det kan ha fått en skade, eller komme fra en syk høne, og kan derfor være dårlig selv om de første fem var bra. David Hume var kritisk til dette, og mente at vi ikke kunne med sikkerhet predikere et utfall basert på tidligere erfaringer eller observasjoner. Han argumenterer med at selv om naturen har vært uniform fram til nå, kan vi ikke konkludere med at den vil fortsette å være uniform. Påstander om at noe vil skje fordi det har skjedd tidligere mente Hume var sirkulære argument, og derfor ikke rasjonelle (Okasha, 2016).

Hume mener altså at det vil være en irrasjonelt å konkludere at det sjette egget også er bra, fordi vi ikke lever i en uniform verden. Uforutsette ting kan skje. Det sjette egget kan være eldre enn de andre i pakken, det kan ha fått en skade, eller komme fra en syk høne, og kan derfor være dårlig selv om de første fem var bra. Hume mente altså at vi aldri med sikkerhet kan predikere framtid basert på fortid fordi fortiden ikke er et sikkert grunnlag for en slik påstand. Humes påstand om at man ikke kan rasjonelt forsvare induksjon uten å bruke et sirkulært argument er det som i dag kalles «Induksjonsproblemet» (Okasha, 2016).

Mange filosofer og forskere har argumentert mot dette problemet. Et eksempel kan være å vise til hvor langt man har kommet innenfor vitenskapen ved å bruke induksjon. Ved hjelp av induksjon har vi landet på månen, oppfunnet laseren og så videre. Personer som ikke har stolt på induksjon kan i verste fall ha dødd på grusomme måter fordi de har spist gift eller hoppet fra bygninger i tro om at de kan fly (Okasha, 2016). Et slik argument ville nok ikke overbevise Hume om at induksjonsproblemet ikke var reelt. Han ville nok kontre med at dette argumentet kun var gyldig hvis man antok at uniformitetsprinsippet stemte – noe han mente det ikke alltid gjorde. Han innrømmet på en annen side at induksjon ble brukt hele tiden, både i hverdag og vitenskap, men det vil ikke bety at det er rasjonelt. Poenget til Hume var nemlig at man ikke kan med sikkerhet predikere hva som kommer til å skje basert på det som har skjedd tidligere. Kanskje ville han gå med på at vi kan godta en antakelse om hva

som kommer til å skje basert på statistisk analyse, men en konklusjon ville han fortsatt kalle irrasjonell (Okasha, 2016).

7.2 Oppgave 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.

Karl Popper var enig med Humes filosofi om at induksjon var irrasjonelt og at man ikke kan predikere framtid basert på fortid. Han mente også at Humes induksjonsproblem var umulig å løse, og at induksjon rett og slett var et mislykket prosjekt som ikke burde ha noe med vitenskap å gjøre. Som et svar mot induksjon kom Popper med en deduktiv metode som kalles «*Falsifikasjonisme*». Kort forklart er dette en teori som sier at vi aldri kan bevise en hypotese – bare motbevise den, altså falsifisere. Popper mente at alle vitenskapelige hypoteser må være falsifiserbare. Dette vil ikke si at alle vitenskapelige teorier skal være feil, men at det skal være mulig å forsøke og motbevise de gjennom testing (Okasha, 2016).

Popper mente at flere såkalte vitenskapelige teorier ikke var tilstrekkelig falsifiserbare til å kvalifiseres som vitenskap, men i stedet var pseudo-vitenskap. Freuds psykoanalyse var en av teoriene Popper ikke anerkjente som ekte vitenskap. Grunnen til dette var at uansett hva en pasient gjorde eller led av kunne Freud forklare det med teorien sin. Det var aldri noe som ikke passet inn i teorien og den ville aldri bli motbevist. Ettersom den aldri kunne motbevises var den ikke kvalifisert til å bli kalt vitenskap ifølge Popper (Okasha, 2016). Et eksempel Popper selv bruker er at en mann dytter et barn ut i vannet for å drukne det, og at en man hopper ut i vannet for å redde barnet. Dette er motstridende handlinger som begge passer like godt i teoriene til Freud og Adler, uten at teoriene motbevises. Dette mente Popper var et bevis for svakhetene til teoriene deres, ikke styrkene (Popper, 1969).

Denne type teorier som aldri kunne motbevises, men alltid fant bekreftende bevis var grunnen til at Popper kom med teorien om falsifikasjonisme. Han mente som sagt at en hypotese bare var vitenskapelig hvis den kunne falsifiseres. En vitenskapelig hypotese skulle predikere noe – gjerne så nøyaktig som mulig, og den skulle være mulig å teste og eventuelt motbevises. Hvis et forsøk ikke klarte å motbevise hypotesen var den ikke bekreftet, bare styrket. Eksempelet Popper mente at all vitenskap burde strebe etter var Einsteins generelle relativitetsteori. Her kom Einstein med en presis prediksjon om at solens gravitasjon ville påvirke stråling fra stjerner. Dette ble senere testet under en solformørkelse og observasjonene stemte med Einstein prediksjon. Dette var akkurat det Popper mente var korrekt vitenskap – en presis detaljert prediksjon som kunne testes (Popper, 1969).

En utfordring med Poppers teori om falsifikasjonisme er at den vil utelukke, eller i alle fall gjøre det vanskelig for noen fagfelt å passe inn i hans definisjon av vitenskap. I naturvitenskapelige fagfelt som fysikk er det relativt enkelt å lage tester som skal forsøke å falsifisere en hypotese.

I psykologi for eksempel kan dette være langt vanskeligere fordi det kan handle om hypoteser som ikke er målbare. Et eksempel kan være følgende to påstander.

1. Et barn vokser opp med en kriminell far. Barnet ser opp til faren sin og ønsker å bli som han. Barnet vil ende opp som kriminell selv.
2. Et barn vokser opp med en kriminell far. Barnet er flau over faren sin og han handlinger og ønsker å ta avstand fra det. Barnet vil ende opp som politi.

Begge disse hypotesene er plausible, men det vil være umulig å lage en test som motbeviser eller beviser de. Psykologi, personlighet, samfunn og andre personer er variabler som er umulig å fjerne og som vil hindre et eventuelt forsøk i å være reliabelt. Her kan vi bare vente å se på oppveksten til barnet og så trekke en konklusjon etterpå slik som Freud og Adler gjorde i sin tid. Dette vil ikke være særlig imponerende eller vitenskapelig ifølge Popper, men det betyr ikke at påstandene er feil.

En mulig løsning på dette problemet er å kombinere deduksjon med induksjon ved hjelp av statistikk og sannsynlighetsregning. I stedet for å si med sikkerhet at 1: barnet vil bli kriminelt, eller 2: barnet vil bli politi, kan man se på statistikk for barn som har vokst opp under liknende forhold, og regne ut sannsynlighetene for hva som vil skje med dette barnet. For eksempel 1: ut fra barnets oppvekst vil det være x % sannsynlig at barnet selv vil bli kriminelt, eller 2: ut fra barnets oppvekst vil det være x % sannsynlig at barnet vil ta avstand fra det livet og bli politi. Det kan altså være mulig å kombinere deduksjon og induksjon.

Kilder

- Almquist NW, Eriksen HB, Wilhelmsen M, Hamarsland H, Ing S, Ellefsen S, Sandbakk Ø, Rønnestad BR & Skovereng K (2022). No differences between 12 weeks of block- vs. Traditional-periodized training in performance adaptations in trained cyclists. *Frontiers in Physiology*; DOI: [10.3389/fphys.2022.837634](https://doi.org/10.3389/fphys.2022.837634).
- Androulakis-Korakakis P, Fisher JP & Steele J (2019). [The Minimum Effective Training Dose Required to Increase 1RM Strength in Resistance-Trained Men: A Systematic Review and Meta-Analysis](#). *Sports Medicine* **50**, 751–765.
- Anon (2009). [Progression Models in Resistance Training for Healthy Adults](#). *Medicine & Science in Sports & Exercise* **41**, 687–708.
- Anon (n.d.). *R: A language and environment for statistical computing*.
- Bass JJ, Wilkinson DJ, Rankin D, Phillips BE, Szewczyk NJ, Smith K & Atherton PJ (2016). [An overview of technical considerations for Western blotting applications to physiological research](#). *Scandinavian Journal of Medicine & Science in Sports* **27**, 4–25.
- Bates D, Mächler M, Bolker B & Walker S (2015). Fitting linear mixed-effects models using {lme4}. ; DOI: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).
- Breil FA, Weber SN, Koller S, Hoppeler H & Vogt M (2010). [Block training periodization in alpine skiing: effects of 11-day HIT on VO2max and performance](#). *European Journal of Applied Physiology* **109**, 1077–1086.
- Brosowsky N, Parshina O, Locicero A & Crump MJC (2020). Teaching undergraduate students to read empirical articles: An evaluation and revision of the QALMRI method. Available at: <http://dx.doi.org/10.31234/osf.io/p39sc>.
- Halperin I, Pyne DB & Martin DT (2015). [Threats to internal validity in exercise science: A review of overlooked confounding variables](#). *Int J Sports Physiol Perform* **10**, 823–829.
- Hopkins WG (2000). [Measures of reliability in sports medicine and science](#). *Sports Med* **30**, 1–15.

- Hulley SB, Cummings SR, Browner WS, Grady DG & Newman TB (2013). Designing clinical research - fourth edition. *LIPPINCOTT WILLIAMS & WILKINS, a WOLTERS KLUWER business*. Available at: [file:///C:/Users/Anders%20Nybakk/OneDrive/Skrivebord/Kvantitativ%20metode%20og%20statistik/Hulley%20\(2013\)%20-%20Designingclinicalresearch_4th-edition.pdf](file:///C:/Users/Anders%20Nybakk/OneDrive/Skrivebord/Kvantitativ%20metode%20og%20statistik/Hulley%20(2013)%20-%20Designingclinicalresearch_4th-edition.pdf).
- Javaloyes A, Sarabia JM, Lamberts RP, Plews D & Moya-Ramon M (2020). [Training Prescription Guided by Heart Rate Variability Vs. Block Periodization in Well-Trained Cyclists](#). *Journal of Strength and Conditioning Research* **34**, 1511–1518.
- Machado FA, Nakamura FY & Moraes SMFD (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**, 1267–1274.
- McArdle WD, Katch FI & Katch VL (2014). *Exercise physiology: Nutrition, energy, and human performance*, 8th edn. Wolters Kluwer Health.
- Mitchell CJ, Churchward-Venne TA, West DWD, Burd NA, Breen L, Baker SK & Phillips SM (2012). [Resistance exercise load does not determine training-mediated hypertrophic gains in young men](#). *Journal of Applied Physiology* **113**, 71–77.
- Naclerio F, Faigenbaum AD, Larumbe-Zabala E, Perez-Bibao T, Kang J, Ratamess NA & Triplett NT (2013). [Effects of Different Resistance Training Volumes on Strength and Power in Team Sport Athletes](#). *Journal of Strength and Conditioning Research* **27**, 1832–1840.
- Noble JE & Bailey MJA (2009). Chapter 8 quantitation of protein. In, pp. 73–95. Elsevier. Available at: [http://dx.doi.org/10.1016/S0076-6879\(09\)63008-1](http://dx.doi.org/10.1016/S0076-6879(09)63008-1).
- Okasha S (2016). *Philosophy of science - a very short introduction*, 2nd edn. Oxford University Press.
- Pauw KD, Roelands B, Cheung SS, Geus B de, Rietjens G & Meeusen R (2013). [Guidelines to classify subject groups in sport-science research](#). *International Journal of Sports Physiology and Performance* **8**, 111–122.
- Popper K (1969). *Conjectures and refutations : The growth of scientific knowledge*, 1st edn. Routledge; Kegan Paul.
- Rønnestad BR, Ellefsen S, Nygaard H, Zacharoff EE, Vikmoen O, Hansen J & Hallén J (2012a). [Effects of 12 weeks of block periodization on performance and performance indices in well-trained cyclists](#). *Scandinavian Journal of Medicine & Science in Sports* **24**, 327–335.

- Rønnestad BR, Hansen J & Ellefsen S (2012*b*). [Block periodization of high-intensity aerobic intervals provides superior training effects in trained cyclists](#). *Scandinavian Journal of Medicine & Science in Sports* **24**, 34–42.
- Rupple BA, Plotkin DL, Smith MA, Godwin JS, Sexton CL, McIntosh MC, Kontos NJ, Beausejour JP, Pagan JJ, Rodriguez JP, Sheldon D, Knowles KS, Libardi CA, Young KC, Stock MS & Roberts MD (2023). The effects of resistance training to near failure on strength, hypertrophy, and motor unit adaptations in previously trained adults. *Physiological Reports*; DOI: [10.14814/phy2.15679](https://doi.org/10.14814/phy2.15679).
- SCHOENFELD BJ, CONTRERAS B, KRIEGER J, GRGIC J, DELCASTILLO K, BELLIARD R & ALTO A (2019). [Resistance Training Volume Enhances Muscle Hypertrophy but Not Strength in Trained Men](#). *Medicine & Science in Sports & Exercise* **51**, 94–103.
- Schoenfeld BJ, Pope ZK, Benik FM, Hester GM, Sellers J, Nooner JL, Schnaiter JA, Bond-Williams KE, Carter AS, Ross CL, Just BL, Henselmans M & Krieger JW (2016). [Longer Interset Rest Periods Enhance Muscle Strength and Hypertrophy in Resistance-Trained Men](#). *Journal of Strength and Conditioning Research* **30**, 1805–1812.
- Spiegelhalter D (2019*a*). The art of statistics: How to learn from data. *New York: Basic books, Hatchet book group*.
- Spiegelhalter DJ (2019*b*). *The art of statistics : Learning from data*, 1th edn. Pelican.
- Tanner RK & Gore CJ (2012). *Physiological tests for elite athletes*, 2nd edn. Human Kinetics. Available at: <https://books.google.no/books?id=0OPIiMks58MC>.
- Wackerhage H (2014). *Molecular exercise physiology - an introduction*, 1st edn. Routledge.
- Walker O (2024). Heart rate variability (HRV). *Science for Sport*. Available at: <https://www.scienceforsport.com/heart-rate-variability-hrv/>.