

# Mappeeksamen IDR4000

508

2024-11-22

# Table of contents

<b>Forord</b>	<b>4</b>
<b>1 Reliabilitet og verktøy for reproduserbar vitenskapelig data</b>	<b>5</b>
1.1 Introduksjon . . . . .	5
1.1.1 Protokoll . . . . .	5
1.2 Metode: . . . . .	6
1.2.1 Deltakere . . . . .	6
1.2.2 Protokoll . . . . .	6
1.2.3 Før test: . . . . .	6
1.2.4 Etter test: . . . . .	7
1.3 Standardisering: . . . . .	7
1.4 Resultater . . . . .	7
1.4.1 Graf over laktatmålingene . . . . .	10
1.4.2 Graf over RER-målingene . . . . .	11
1.4.3 Utregning av variasjonskoeffisient . . . . .	11
1.5 Avvik og feilkilder . . . . .	11
1.5.1 Standardisering . . . . .	11
1.6 Resultat . . . . .	13
<b>2 Regression Models</b>	<b>14</b>
2.1 Introduksjon . . . . .	14
2.2 Del 1: Laktatterskler . . . . .	14
2.2.1 Introduksjon . . . . .	14
2.2.2 Metode . . . . .	14
2.3 Del 2: Predikere størrelse på DNA-fragmenter . . . . .	16
2.3.1 Introduksjon . . . . .	16
2.3.2 Metode . . . . .	17
2.3.3 Målinger . . . . .	17
2.4 Diskusjon . . . . .	18
<b>3 Statistical inference</b>	<b>20</b>
3.1 Oppgave 1 . . . . .	20
3.2 Oppgave 2: . . . . .	21
3.3 Oppgave 3: . . . . .	21
3.4 Oppgave 4: . . . . .	22

3.5	Oppgave 5: . . . . .	23
3.6	Oppgave 6: . . . . .	23
3.7	Oppgave 7: . . . . .	23
3.8	Oppgave 8: . . . . .	24
<b>4</b>	<b>Study Design</b>	<b>26</b>
4.1	Study design . . . . .	26
<b>5</b>	<b>Philosophy of Science</b>	<b>29</b>
5.1	Introduksjon . . . . .	29
5.2	Metode . . . . .	30
5.2.1	Deltakerne . . . . .	30
5.2.2	Styrketrening . . . . .	30
5.2.3	Tester . . . . .	30
5.2.4	Data . . . . .	31
5.3	Diskusjon . . . . .	32
5.4	Konklusjon . . . . .	33
<b>6</b>	<b>Philosophy of Science</b>	<b>34</b>
6.1	Oppgave 1 . . . . .	34
6.2	Oppgave 2 . . . . .	35
<b>7</b>	<b>Labratory report</b>	<b>37</b>
7.1	Hensikt: . . . . .	37
7.2	Metode: . . . . .	37
7.3	Resultater: . . . . .	39
7.4	Diskusjon . . . . .	39
<b>8</b>	<b>Referanser</b>	<b>41</b>

# Forord

Dette er mappeeksamen i IDR4000.

Hver del er produsert av meg eller som gruppearbeid i undervisningssammenheng.

Link til repository på github: <https://github.com/ajuvodden/-mappeeksamen.git>

Mappeeksamen består av følgende deler:

- Rapport: "Reliabilitet"
- Rapport: "Regresjonsmodeller"
- Rapport: "Statistisk inferens, statistiske modeller og statistisk styrke"
- Rapport: "Studiedesign"
- Rapport: "Analyse av eksperimenter med repeterte målinger"
- Arbeidskrav i vitenskapsteori
- Labrapport fra molekylærlabb

# 1 Reliabilitet og verktøy for reproduserbar vitenskapelig data

## 1.1 Introduksjon

I vår studie har vi gjennomført VO<sub>2</sub> max-tester over fire forskjellige dager for å vurdere påliteligheten av de innsamlede dataene. Formålet med disse testene var å undersøke hvor konsistente resultatene er under kontrollerte forhold, altså estimere reliabiliteten. For å oppnå dette har vi forsøkt å standardisere flere variabler, inkludert treningsnivå og matinntak dagen før testene. Ifølge Halperin, Pyne, and Martin (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 reliabiliteten i testresultatene. Laktatnivåene kan være en nyttig indikator på hvor konsistente og stabile testene er ved gjentakelser, ettersom tidligere forskning har vist at denne parameteren kan variere med intensiteten og varigheten av belastningen (Hopkins (2000); Tanner (2012)). Ved å overvåke laktatnivåene kan vi sikre at de fysiologiske responsene er reproduserbare, og dermed at testene gir pålitelige målinger som reflekterer ekte treningsbelastning.

I denne rapporten velger vi å fokusere på laktatmaks og RER. Vi vil sammenligne disse, beregne standardavviket og analysere påliteligheten ved å bruke variasjonskoeffisienter (CV).

### 1.1.1 Protokoll

#### Protokoll VO<sub>2</sub>maks-test

Vi gjennomførte fire testdager, hvor de to første testene ble gjennomført på påfølgende dager, og de to siste var med én hviledag mellom. Hensikten med disse testene var å gjennomføre fysiologiske tester med høy grad av reliabilitet. Det er mange faktorer som kan påvirke validitet og reliabilitet, og det er svært viktig å ta høyde for disse under fysiologisk testing. Vi tok flere forhåndsregler for å sikre så like testforhold som mulig.

En VO<sub>2</sub> max-test går ut på å måle hvor mange ml oksygen 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).

## 1.2 Metode:

VO2 max-testen ble gjennomført som en trappetrinnstest, hvor motstanden økte med 20W hvert minutt til utmattelse, eller når RPM ble mindre enn 60. Det ble registrert målinger av VO2 max hver 30. sekund. Deltakerne startet testen på en watt tilsvarende deres fysiske form og erfaring med sykkel. Etter testene var ferdige ble dataene samlet og plottet i et ferdig formatert Excel-dokument.

For å oppnå pålitelige og reproduerbare resultater ble flere faktorer kontrollert. Matinntak og koffeininntak ble standardisert, og trening dagen før testen ble registrert. Det ble også tatt høyde for at deltakerne skulle få tilstrekkelig søvn før testene. Alle deltakere hadde samme testleder under hver test for å sikre konsistens i instruksjoner og motivasjon.

### 1.2.1 Deltakere

Det deltok 14 forsøkspersoner under forsøket, der 3 av de var kvinner og 11 av de var menn med en gjennomsnittsalder på 25,7 år. Gjennomsnittshøyde var 180 og gjennomsnittsvekt var 79,3.

### 1.2.2 Protokoll

#### 1.2.3 Før test:

Starte laktatmaskin. Bytte standard hvis den er tom (finnes i kjøleskap). Finne frem munnstykke og sette sammen dette, husk hansker, Papir med neseklype rundt når den er klar. Ta med ny slange som festes i miksekammer. Skru på Lode-sykkel. Skru på Vyntus og Lode PC

#### Kalibrere oksygenanalysator

Gasskalibrering. Åpne gassflaske (lukkes når kalibrering er ferdig). Sjekk at sensor er koblet i maskinen. Start kalibrering. Godtar 2% feilmargin, hvis den er høyere må man rekalkibrere. Volumkalibrering, Feilmargin på 0,2% eller under godtas. Kalibrere kammer, dermed flytte sensor til kammer og skru av gass.

#### 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 på pc. Stille inn krankarm (172,5) og kalibrere sykkel på Lode PC. Stille inn sykkel til deltaker. Fullføre plotteark med ny informasjon. Klargjøre laktatrør, papir og teip, samt teip til neseklype. Velge protokoll på Lode PC. Dersom personen ikke har erfaring med sykkel må man bli enige om en Watt man tror kan passe.

**Forklare forsøksperson hva testen innebærer, opplyse om borg skala og hvor på skala man vil ha den.**

VO2max – Kondisjonstest. Målinger hvert 30 sek. Prestasjonstest – hvert sekund teller. Watt-økning hvert 1 min. 80-100 rpm anbefales, stopper test hvis under 60rpm. Forklare Borg skala – 6 = sofa, 20 = maks. Laktatmåling 1 min etter test – stikk i fingeren.

#### 1.2.4 Etter test:

Avslutte test på begge PC. Spørre om Borg skala. 1 minutt etter endt test skal det tas laktat-prøve. Tok 2 prøver for å sikre god reliabilitet.

#### Hente rapport:

Rapport: INN\_TABELL:30SEK\_MIX. Søk opp id-nr. Lagre i rett mappe: F10 (nederste knapp). Lagre i rett mappe på minnepenn. Overfør til One Drive på labbPCnFør test:

### 1.3 Standardisering:

I forsøket vårt valgte vi å standardisere matinntak og koffeininntak før testene. Vi ønsket at deltakerne skulle spise de samme måltidene de tre siste timene før testen, og ha likt koffeininntak på testdagen.

Standardisering av trening var utfordrende, da deltakerne hadde ulike treningsopplegg, og dette kunne ikke helt reproduseres. Vi forsøkte imidlertid å kontrollere for andre faktorer som søvn og testtidspunkt. Selv om vi forsøkte å standardisere søvnmengden, kunne vi ikke kontrollere for søvnkvalitet, og vi fokuserte derfor mindre på dette.

### 1.4 Resultater

id	timepoint	sex	age	
Length:47	Length:47	Length:47	Min.	:22.00
Class :character	Class :character	Class :character	1st Qu.:	23.97
Mode :character	Mode :character	Mode :character	Median	:24.58
			Mean	:26.50
			3rd Qu.:	27.00
			Max.	:43.72
height	weight	w.max	vo2.max	vco2.max
Min. :162.0	Min. : 58.50	Min. : -20.0	Min. :2464	Min. :2945
1st Qu.:176.0	1st Qu.: 71.50	1st Qu.:250.7	1st Qu.:3180	1st Qu.:3774

Median :178.0	Median : 79.10	Median :318.3	Median :4235	Median :4785
Mean :178.7	Mean : 78.74	Mean :314.6	Mean :3967	Mean :4590
3rd Qu.:180.0	3rd Qu.: 83.40	3rd Qu.:384.5	3rd Qu.:4584	3rd Qu.:5294
Max. :191.0	Max. :103.90	Max. :455.7	Max. :5294	Max. :6462

rer.max	ve.max	bf.max	hr.max
Min. :1.085	Min. : 83.9	Min. : 30.35	Min. : 0.0
1st Qu.:1.125	1st Qu.:127.8	1st Qu.: 42.17	1st Qu.:180.8
Median :1.160	Median :161.5	Median : 48.35	Median :186.5
Mean :1.160	Mean :155.8	Mean : 57.38	Mean :181.6
3rd Qu.:1.192	3rd Qu.:185.5	3rd Qu.: 65.42	3rd Qu.:190.0
Max. :1.275	Max. :214.5	Max. :191.00	Max. :203.0
			NA's :3

la.max	borg.max	group
Min. : 6.43	Min. :16.00	Length:47
1st Qu.:10.99	1st Qu.:18.00	Class :character
Median :13.25	Median :19.00	Mode :character
Mean :12.85	Mean :18.62	
3rd Qu.:14.73	3rd Qu.:19.50	
Max. :19.62	Max. :20.00	
NA's :1		

#### 1.4.0.1 Standardavvik:

##### Laktat

```
# A tibble: 4 x 4
  timepoint m s n
  <chr>    <dbl> <dbl> <int>
1 t1      12.2  2.76  14
2 t2      12.3  2.91  14
3 t3      13.6  2.50  11
4 t4      13.8  2.31   8
```

##### Laktat etter kjønn

```
# A tibble: 8 x 5
  timepoint sex m s n
  <chr>    <chr> <dbl> <dbl> <int>
1 t1      f 13.0 0.954 3
2 t1      m 12.0 3.08 11
```



3	t2	f	12.9	2.70	3
4	t2	m	12.1	3.09	11
5	t3	f	12	1.40	2
6	t3	m	14.0	2.60	9
7	t4	f	12.3	3.20	3
8	t4	m	14.7	1.15	5

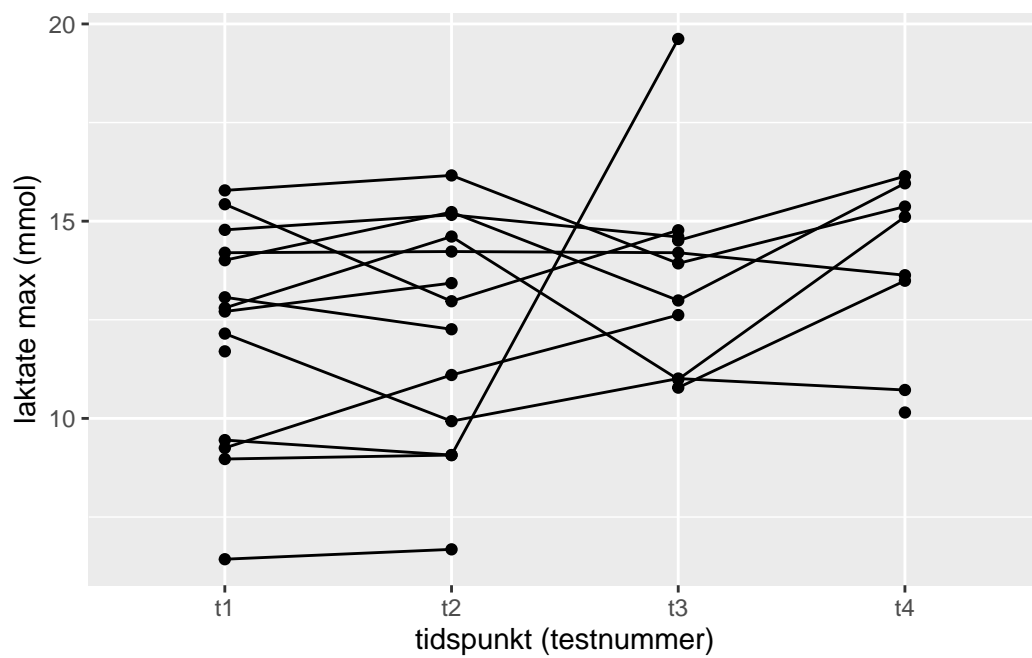
## RER

```
# A tibble: 4 x 4
  timepoint      m      s      n
  <chr>      <dbl> <dbl> <int>
1 t1         1.14 0.0337    14
2 t2         1.15 0.0379    14
3 t3         1.18 0.0454    11
4 t4         1.19 0.0308     8
```

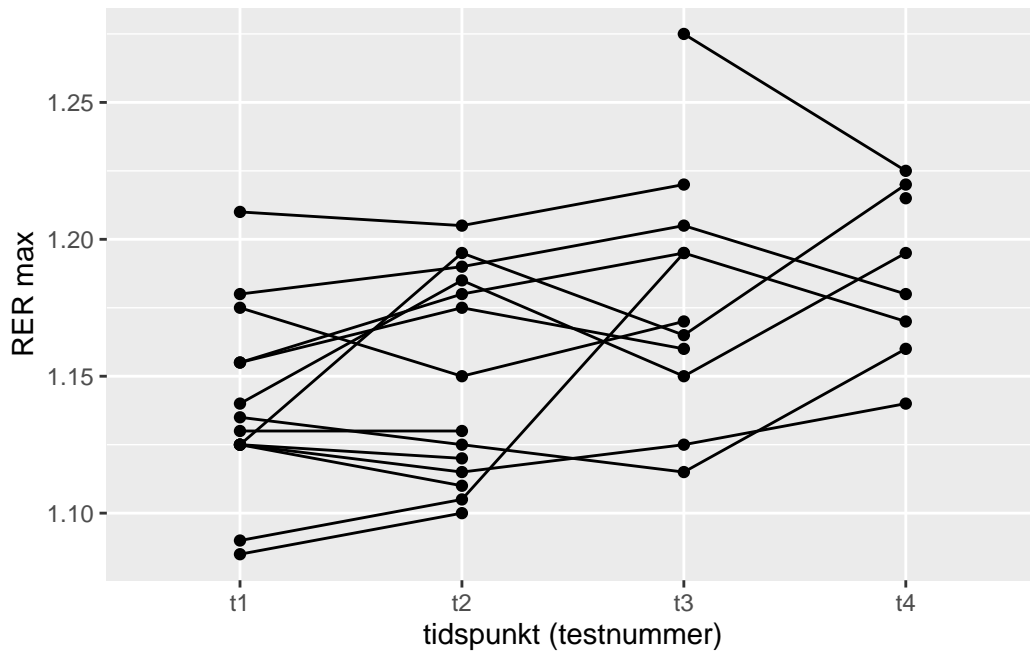
## RER etter kjønn

```
# A tibble: 8 x 5
  timepoint sex      m      s      n
  <chr>      <chr> <dbl> <dbl> <int>
1 t1        f      13.0 0.954     3
2 t1        m      12.0 3.08    11
3 t2        f      12.9 2.70     3
4 t2        m      12.1 3.09    11
5 t3        f       12  1.40     2
6 t3        m      14.0 2.60     9
7 t4        f      12.3 3.20     3
8 t4        m      14.7 1.15     5
```

### 1.4.1 Graf over laktatmålingene



### 1.4.2 Graf over RER-målingene



### 1.4.3 Utregning av variasjonskoeffisient

#### 1.4.3.1 Laktat maks

Variasjonskoeffisienten for laktat maks var 7.6%

#### 1.4.3.2 RER maks

Variasjonskoeffisienten for RER maks var 1.6%

## 1.5 Avvik og feilkilder

### 1.5.1 Standardisering

#### Livsstil

I forkant av perioden ble vi enig om å forsøke å standardisere trening og matinntak i forkant av testene. Dette visste vi med en gang kunne bli utfordrende ettersom deltakerne har egne

treningsopplegg som de følger. Noen på gruppa hadde for eksempel forskjellige treningsbolker fra t1 og t2 til t3 og t4. I tillegg til dette var t1 og t2 satt opp to dager etter hverandre, mens det var en hviledag mellom t3 og t4. Id1 gjennomførte også en test mandagen før t1, så i den første testuken ble id1 testet tre ganger.

Mat og drikke i forkant av testen var også vanskelig å standardisere da deltakerne også måtte ta hensyn til andre familimedlemmer ol. Underveis i testen var ikke dette en utfordring ettersom testen var så kort var det ikke nødvendig med inntak av mat eller drikke underveis.

Selv om vi ønsket å standardisere søvn så vi at dette ville være vanskelig å gjennomføre. Vi oppfordret deltakerne til å forsøke å få like mye søvn hver gang ved å legge seg til samme tid, men ettersom vi ikke kunne sikre lik søvnkvalitet valgte vi å ikke ha så mye fokus på dette.

Dagsform vil også påvirke resultatene våre. Travle dager og uvant livsstil i forkant av testen kan påvirke resultatet selv om trening, mat og drikke er standardisert. Dagsform vil også påvirke utøverens evne til å presse seg.

### **Unøyaktighet i framgangsmåte.**

Vi har laget et protokoll med framgangsmåte, og har forsøkt å følge den hver gang. Likevel vil det alltid skje unntak og menneskelige feil som kan påvirke resultatet.

I protokollen står det at vi skal ta laktatprøve 1 min etter at testen er ferdig. Dette har vi forsøkt å gjøre, men i noen tilfeller måtte vi forkaste den første testen fordi røret kom borti huden, ikke ble fylt opp helt og liknende. I slike tilfeller vil testen bli tatt senere enn protokollen tilsier, og vi vil risikere at laktat-nivået har endret seg. En annen typisk feil som kan skje er at man ikke redder laktaten med en gang man har fylt røret. Hvis man tar seg tid til å tørke blod og gi utøveren plaster først vil disse sekundene med forsinkelse også kunne påvirke laktatnivået i prøven.

Vi hadde de samme testlederne for hver enkelt deltaker hver gang, slik at heining, oppmuntring og oppdatering underveid i testen skulle være så likt som mulig. Dette er fordi det mentale spiller en stor rolle i deltakernes evne til å presse seg og få ut alt de har på testen. Ideelt sett skal man som testleder gjøre nøyaktig det samme hver gang, samme informasjon, samme beskjeder og lik heining. Det vil selvfølgelig være umulig å få til dette helt nøyaktig, men vi prøvde på dette også ved å følge protokollen. Har testleder derimot en god eller dårlig dag vil dette lett påvirke intensiteten i heining, som igjen vil påvirke utøverens prestasjon.

En annen feil som skjedde var at vi en eller to ganger var litt unøyaktig med å starte testen i vyntus og lode samtidig. Dette gjør at watt-økning for eksempel kunne komme litt før eller litt etter målingen i vyntus. Dette kan både være problem for testlederen som skal gi tilbakemeldinger og for utøveren som kanskje kjenner at wattøkningen kommer tidligere enn testleder gir beskjed om. Start og stopp av stoppeklokke ble også forsinket ved et par anledninger.

På t4 manglet det også stoppeklokke på laboratoriet. Vi løste dette med å bruke en telefon som stoppeklokke i stedet. Det ble litt vanskeligere for utøveren å følge med på tiden på grunn av dette.

### Andre grupper

I denne rapporten har vi også brukt data fra de andre gruppene i klassen vår. Dette gjør at vi har et bredere datautvalg som vi kan benytte oss av. Problemet med dette er at vi ikke har kontroll på andre gruppers protokoll, avvik, nøyaktighet osv. I datasettet var det for eksempel en plottfeil mellom rer.max og ve.max som vi ikke fikk endret på. Dette førte igjen til en feil i utregning av standard avvik, og at en eventuell tabell ikke ble brukbar. Det var også plottet ulikt på kjønn, da noen skrev “f”, “F”, “k” eller “kvinne” for kvinner, og “m”, “M” eller “mann” for menn. Dette fikset vi ved å gå inn i excelarket og erstatte alle med “f” og “m” på de korresponderende kjønnene.

## 1.6 Resultat

Vi valgte å se på laktat og RER i denne oppgaven. Når RER øker over 1.0 indikerer det ofte at laktatproduksjonen øker. RER samfatter ofte med den anaerobe terskelen hvor laktatnivået begynner å stige raskt.

Variasjonskoeffisientene var:

laktat 7.6%

RER 1.6%

Disse målingene 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 rentabiliteten negativt.

Samlet sett viser dataen at RER-målingene er høyst reliable og at laktatmålingene er mindre konsistente. Dette betyr ikke nødvendigvis at laktatmålingene er helt ubrukelige, men at de bør tolkes med forsiktighet og i sammenheng med andre fysiologiske markører som RER.

Det kan være flere grunner til dette. 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. Derimot er RER-målinger automatisert og mindre påvirket av faktorer som prøvetakingsteknikk.

## 2 Regression Models

### 2.1 Introduksjon

Denne oppgaven er delt inn i tre deler, som hver fokuserer på ulike konsepter innen dataanalyse og regresjon. I del 1 beregner vi laktatterskler og undersøker reliabiliteten mellom to ulike terskelnivåer. I del 2 benytter vi molekulære data for å forutsi størrelsen på DNA-fragmenter. Til slutt, i del 3, undersøker vi om det er en lineær sammenheng mellom valgte variabler fra datasettet hypertrophy i datapakken exscidata.

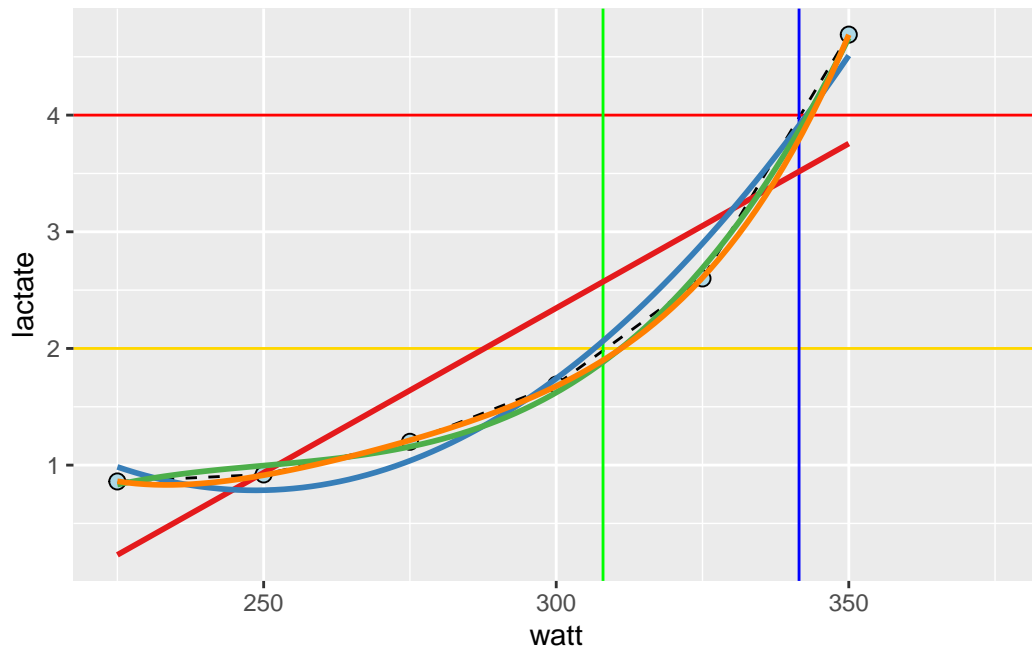
### 2.2 Del 1: Laktatterskler

#### 2.2.1 Introduksjon

Laktatterskel er en viktig variabel for å predikere prestasjoner i utholdenhetsidretter, og brukes både til å justere treningsintensiteten og evaluere effekten av trening (Machado, Nakamura, and Moraes (2012)). Det finnes flere metoder for å bestemme en utøvers laktatterskel, og ifølge Machado, Nakamura, and Moraes (2012) er “maximal-deviation method” (Dmax), som ble anbefalt av Cheng et al. i 1992, en av de mest pålitelige. Denne metoden er spesielt nyttig for å vurdere mekanismene som påvirker prestasjon i langdistanseidretter, og har vist seg å ha en sterkere korrelasjon med både prestasjon og laktatterskel sammenlignet med andre metoder. De representative tersklene som blir undersøkt er ( $\dot{V}_{O_{2max}}$ ). Dette kan oppnås ved å bruke en regresjonsmodell for å beskrive forholdet, og deretter utføre en “invers prediksjon” av treningsintensiteten Tanner (2012).

#### 2.2.2 Metode

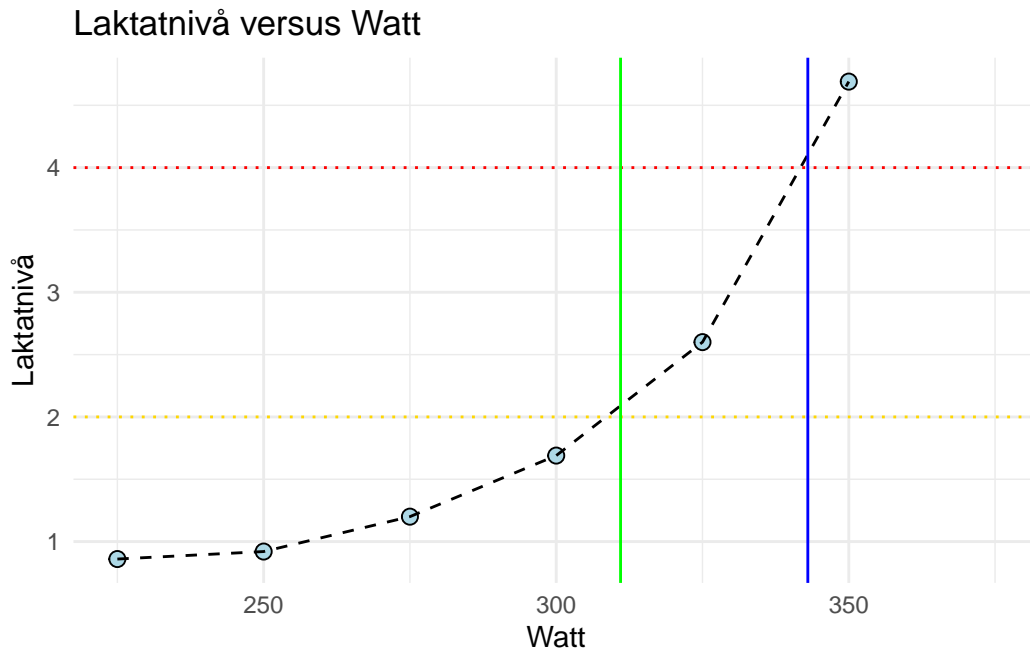
Som vi kan se i den plotta grafen under, er de forskjellige grafene ikke så forskjellige rundt 2mmol og 4mmol  $L^{-1}$ . På den andre siden ser vi at den lineære modellen er feil ved 300w, den sekundærpolynomiske modellen er feil ved 275w. Den tredje- og fjerdepolynomiske modellen derimot, varierer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



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

```
watt predictions
1 311 1.999278
```

```
watt predictions
1 343 4.000333
```



Regresjonsanalysen viser at wattverdien som gir et laktatnivå på 2 mmol L-1 er omtrent 311 watt. Når intensiteten når 311 watt, forventes laktatnivået å være nær 2 mmol L-1. laktatnivået på 4 mmol L-1, ligger predikerte wattverdien på 343 watt. Disse resultatene kan brukes til å vurdere intensiteten på treningsøktene i forhold til laktatterskelene.

For å sammenligne reliabiliteten (SD som en prosent av gjennomsnittet) mellom lac.225 og lac.375, kan vi formulere koden slik:

Resultatene viser at den typiske feilen som prosentandel av gjennomsnittet er 5.95% for lac.225 og 2.26% for lac.375. Dette betyr at for lac.225 er det en betydelig større variasjon i dataene sammenlignet med lac. 375, som har lavere SD. Samtidig indikerer begge resultatene at det er en viss usikkerhet i målingene.

## 2.3 Del 2: Predikere størrelse på DNA-fragmenter

### 2.3.1 Introduksjon

I denne delen av oppgaven ser vi nærmere på forsøket vi gjennomførte på molekylærlabben, hvor vi ekstraherte DNA fra blodprøver. Her ønsket vi å isolere gener relatert til hurtig (R/R) og langsom (X/X) muskelfibre ved hjelp av en PCR maskin. Deretter ble prøvene analysert gjennom elektroforese i agarosegel, sammen med en DNA-ladder som fungerer som markør for å kartlegge genene. DNA-ladder markerer hvert tiende basepar (bp) opp til 300bp og hvert hundrede basepar videre til 1000bp. De dominante R/R-genene har størrelse på 413bp,



mens X/X-genene har en størrelse på 318bp. På grunn av forskjellen i størrelse, vil de mindre X/X-genene trekke lenger inn i gelen under selve elektroforesen. Dette er imidlertid vanskelig å vurdere nøyaktig med bare øynene. For å få mer pålitelig resultat, har vi benyttet en metode for å analysere prøvene for mer presise målinger.

### 2.3.2 Metode

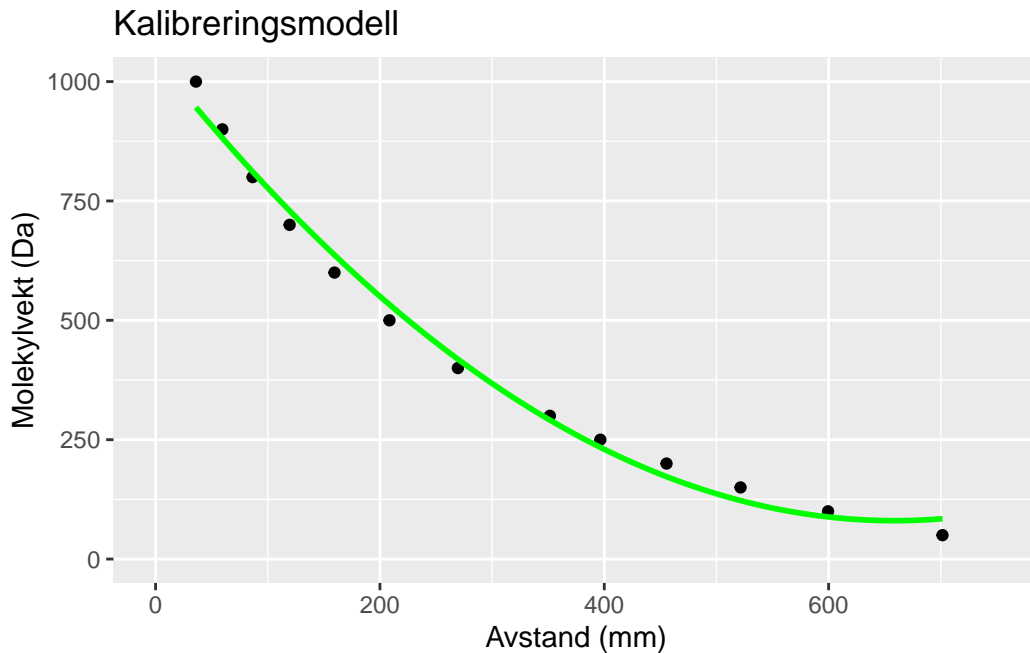
“For å analysere DNA-prøvene, brukte vi ImageJ til å behandle bildene som ble tatt etter elektroforesen. Først inverterte vi bildene for å forbedre kontrasten, deretter justerte vi dem slik at de var riktig justert, og klippet ut den relevante delen av bildet for videre analyse. Videre markerte vi stigen og prøvene vi ønsket å analysere, og ut fra de markerte områdene lagde vi grafer for hver brønn.”

For å analysere DNA-fragmentene målte vi migrasjonsavtandene i gelen ved hjelp av ImageJ. Vi markerte toppunktene i grafene som representerer både genene i prøvene og trinnene i DNA-ladder. ImageJ registrerte nøyaktig plassering av disse punktene, og koordinatene ble flyttet til et Excel-ark for videre beregning i R. Her brukte vi den kjente DNA-ladder til å lage en standardkurve ved å relatere migrasjonsavstanden til størrelsen på fragmentene i bp. Tilslutt benyttet vi en lineær regresjonsmodell for å estimere størrelsen på DNA-fragmentene i våre ukjente prøver.

### 2.3.3 Målinger

Vi starter med å kalibrere DNA-stigen for å bestemme molekylstørrelsen på de ukjente prøvene. Dette kan vi gjøre ved å lage en data.frame kalt “ladder”, der vi bruker verdiene som er omvendt proporsjonale, ettersom små molekyler beveger seg lengre i gelen. Deretter kan vi lage en annen data.frame for de ukjente prøvene, som vi kaller “unknown”, for videre analyse.

For å lage en kalibreringsmodell bruker vi dataene fra stigen til å lage en ggplot som visualiserer stigningsforhold. Deretter bruker vi denne grafen til å estimere størrelsen på de ukjente prøvene. Vi benytter en polynomfunksjon (poly) for å tilpasse en kurve som gir minst mulig avvik fra dataene.



Videre kan vi bruke følgende koder for å estimere molekylstørrelsen på genene i prøven vår.

Resultatene viser at kalibreringsmodellen har høy prediktiv nøyaktighet, med et R-squared på 0.9895, som tilsier at modellen forklarer nesten 99% av variasjonen i molekylvektene basert på migrasjonsavstanden.

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

Resultatene viser de estimerte molekylvektene (bp) for de ukjente DNA-fragmentene basert på deres migrasjonsavstand i gelen. Noe som gav oss følgende størrelser:

- Prøve 1: 407.5 bp
- Prøve 2: 400.7 bp
- prøve 3: 395.7 bp
- prøve 4: 295.2 bp

## 2.4 Diskusjon

Resultatene fra kalibreringsmodellen viste høy nøyaktighet med en R-squared på 0.9895, noe som tyder på at vi kan stole på de estimerte molekylvektene for de ukjente prøvene. Ser vi på de estimerte molekylvektene (407.5, 400.7, 395.7 bp og 295.2 bp) stemmer godt overens med de kjente størrelsene for R/R-genet (413) og X/X-genet (318 bp). De fleste prøvene er nær

400 bp, som indikerer at de inneholder R/R-genene, mens den lavere verdien (295.2 bp) kan tilhøre X/X-genet.

Noen faktorer som kan påvirke resultatene er bildekvalitet og variasjon i elektroforesen, som kan føre til små avvik i migrasjonsavstandene. Likevel, de estimerte størrelsene er konsistente med forventningene, og metoden gir en pålitelig tilnærming for å bestemme molekylvektene i DNA-prøvene våre.

## 3 Statistical inference

### 3.1 Oppgave 1

#### 3.1.0.1 Estimerer

Estimatene i regresjonsmodellene `m1` og `m2` representerer gjennomsnittene for hver gruppe. Her estimerer vi populasjonsgjennomsnittet basert på to forskjellige utvalgsstørrelser. Nøgglelen er at estimatet i begge modellene er gjennomsnittet av observerte verdier i utvalget. For `m1` (utvalgsstørrelse 8) er estimert gjennomsnitt 1.840, noe som er litt annerledes enn det sanne populasjonsgjennomsnittet på 1.5. Utvalgsgjennomsnittet er en tilnærming av populasjonsgjennomsnittet, og nøyaktigheten avhenger av utvalgsstørrelsen. For `m2` (utvalgsstørrelse 40) er estimert gjennomsnitt 1.5642, som er mye nærmere populasjonsgjennomsnittet på 1.5. Større utvalg reduserer vanligvis variasjonen i estimatet og gir en mer pålitelig tilnærming av populasjonsgjennomsnittet.

The estimates in the regression models `m1` and `m2` represent the sample means for each group. In this context, we're estimating the mean of a population based on two different sample sizes. The key point is that the estimate in both models is the average of the observed values in the sample.

For `m1` (sample size of 8), the estimated mean is 1.840, which is slightly different from the true population mean of 1.5. The sample mean is an approximation of the population mean, and the accuracy depends on sample size.

For `m2` (sample size of 40), the estimated mean is 1.5642, which is much closer to the population mean of 1.5. Larger sample sizes typically reduce the variability in the estimate, making it a more reliable approximation of the population mean.

#### 3.1.0.2 Standardfeil

Standardfeilen (SE) er et mål på usikkerheten i estimatet. Den viser hvor mye gjennomsnittet fra et nytt tilfeldig utvalg forventes å avvike fra det vi har. For `m1` er SE 1.251. En høy SE indikerer mer usikkerhet rundt estimatet på grunn av liten utvalgsstørrelse. For `m2` er SE 0.4774, som er betydelig mindre enn for `m1`. Dette skyldes at et større utvalg reduserer variasjonen i estimatet, noe som gir et mer presist resultat.

### 3.1.0.3 T-verdi og P-verdi

For m1 er t-verdien 1.47, beregnet som estimatet delt på SE. Dette indikerer hvor langt utvalgsgjennomsnittet er fra 0 målt i standardfeil.

For m2 er t-verdien 3.276, som er større på grunn av lavere SE og er et mer presist estimat i det større utvalget. En høyere t-verdi indikerer sterkere bevis for at gjennomsnittet er signifikant forskjellig fra 0.

For m1 er p-verdien 0.185, som betyr at det er 18.5 % sjanse for å observere en t-verdi så ekstrem som 1.47 hvis nullhypotesen (gjennomsnitt = 0) er sann. Siden dette er over den vanlige terskelen på 0.05, forkaster vi ikke nullhypotesen.

For m2 er p-verdien 0.00221, som er mye mindre og indikerer sterke bevis mot nullhypotesen. Dette antyder at det observerte gjennomsnittet (1.5642) er signifikant forskjellig fra 0.

## 3.2 Oppgave 2:

Den primære forskjellen mellom de to modellene er utvalgsstørrelsene:

Lite utvalg (m1): Med bare 8 observasjoner er estimatet mindre presist, noe som fører til en høyere standardfeil og lavere t-verdi. Som et resultat forkaster vi ikke nullhypotesen, siden estimatet er for upresist til å kunne fastslå en forskjell fra 0 med sikkerhet.

Stort utvalg (m2): Med 40 observasjoner er estimatet mer presist, noe som gir lavere standardfeil, høyere t-verdi og lavere p-verdi. Den økte utvalgsstørrelsen gir større statistisk styrke, noe som gjør det lettere å oppdage en signifikant forskjell fra 0. Forskjellen i resultater mellom m1 og m2 skyldes altså forskjellen i utvalgsstørrelse. Det større utvalget gir mer pålitelige estimater og sterkere statistiske bevis.

## 3.3 Oppgave 3:

I hypotesetesting bruker vi skyggeområdet i de nedre og øvre halene av t-fordelingen til å beregne p-verdien, som representerer sannsynligheten for å observere en t-verdi like ekstrem som den beregnet fra dataene, gitt at nullhypotesen er sann (gjennomsnitt = 0).

For tosidige tester ser vi på avvik fra 0 i begge retninger (positive og negative). Derfor beregner vi arealet under kurven i begge haler av fordelingen.

Skyggeområdet tilsvarer andelen t-verdier som er like ekstreme eller mer ekstreme enn den observerte t-verdien. Hvis dette arealet (p-verdi) er lite, konkluderer vi med at slike ekstreme verdier sannsynligvis ikke oppstår tilfeldig, og vi kan kanskje forkaste nullhypotesen.

For m1 er skyggeområdet (p-verdi = 0.185) stort, noe som betyr at det er en høy sannsynlighet for å observere en t-verdi like ekstrem som 1.47. Derfor forkaster vi ikke nullhypotesen.

For m2 er skyggeområdet (p-verdi = 0.00221) lite, så vi forkaster nullhypotesen og konkluderer med at gjennomsnittet er signifikant forskjellig fra 0.

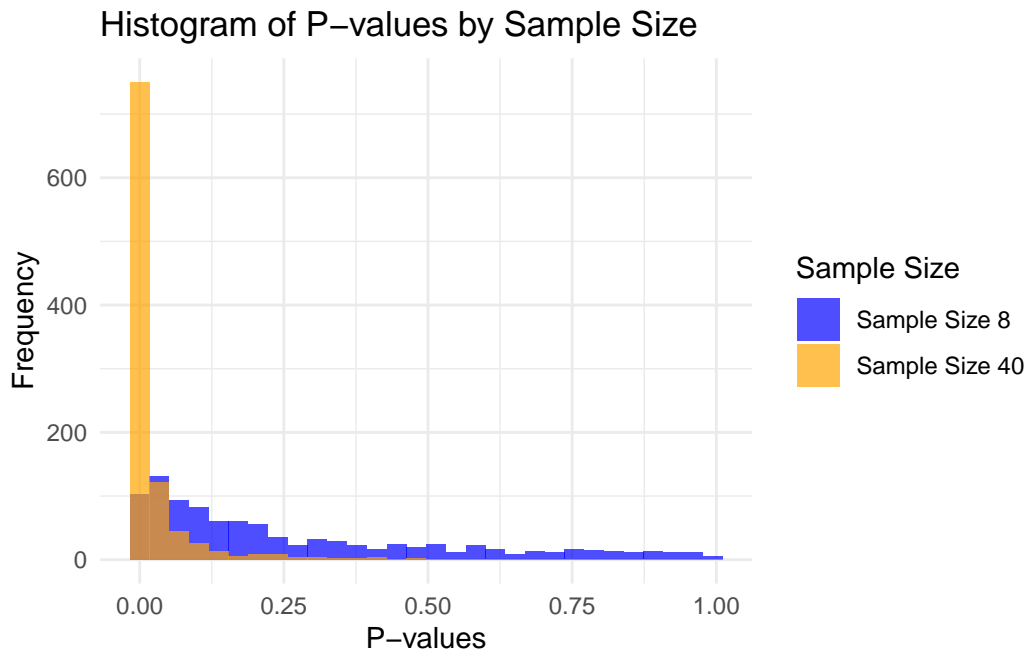
Ved å bruke både estimer og p-verdier fra m1 og m2 kan vi se hvordan utvalgsstørrelsen påvirker påliteligheten av statistiske slutninger.

### 3.4 Oppgave 4:

```
# A tibble: 2 x 3
      n sd_estimate avg_se
  <dbl>      <dbl> <dbl>
1     8        1.07  1.02
2    40        0.484  0.470
```

Standardavviket til estimatene (sd\_estimate) indikerer hvor mye estimatene varierer rundt gjennomsnittsestimatet for hver utvalgsstørrelse. Gjennomsnittlig standardfeil (avg\_se) gir innsikt i variasjonen i selve estimatet. Hvis standardavviket til estimatene er nært gjennomsnittlig standardfeil, antyder det at estimatene er tett samlet rundt gjennomsnittet, og at SE er en pålitelig indikator på denne variasjonen.

### 3.5 Oppgave 5:



Tolkning

Histogrammet lar oss visualisere fordelingen av p-verdier for hver utvalgsstørrelse. Et større utvalg resulterer vanligvis i en smalere p-verdi-fordeling, sentrert nærmere 0, noe som indikerer økt styrke til å påvise en effekt. Mindre utvalg kan derimot føre til en bredere fordeling med flere p-verdier som klynger seg nær 1.

### 3.6 Oppgave 6:

```
# A tibble: 2 x 2
  n significant_count
<dbl>          <int>
1     8             227
2    40             865
```

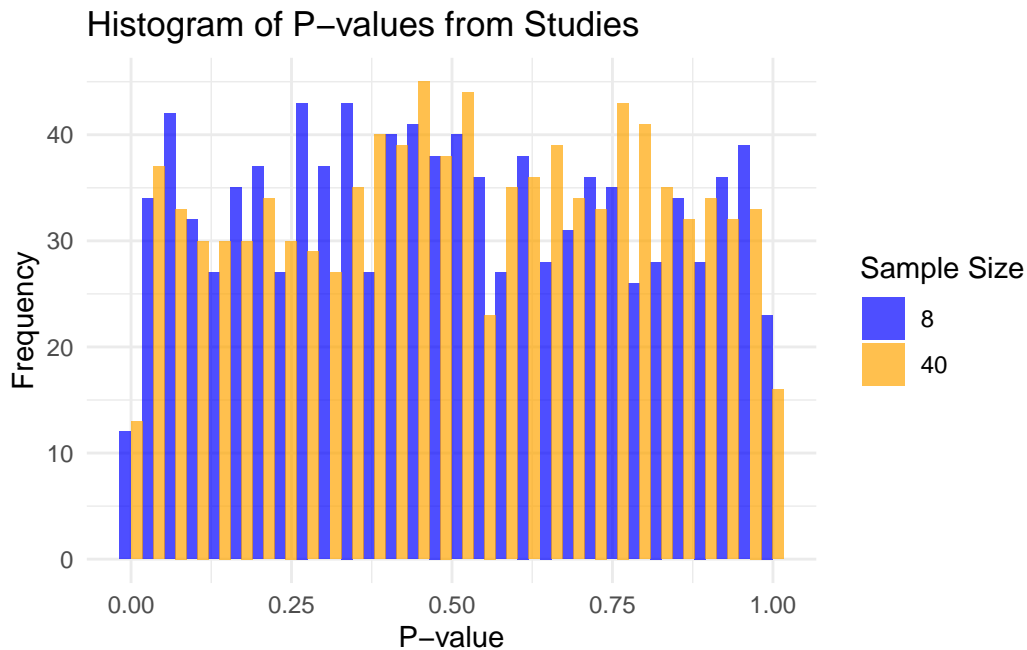
### 3.7 Oppgave 7:

```
sample_size power
1          8 0.2320770
```

### Forklaring av resultater

Styrken til en test indikerer sannsynligheten for korrekt å forkaste nullhypotesen når den faktisk er feil. Høyere styrke betyr større sjanse for å oppdage en effekt dersom den finnes. Typisk gir større utvalgsstørrelser større styrke, noe som er indikert av de beregnede styrkene for de to utvalgsstørrelsene.

## 3.8 Oppgave 8:



Number of false positives for sample size 8: 44

Number of false positives for sample size 40: 49

I hypotesetesting oppstår falske positive når vi feilaktig forkaster nullhypotesen. Ved et signifikansnivå på 0.05 forventer vi at omtrent 5 % av testene gir falske positive dersom nullhypotesen er sann.

Jeg kjørte 1000 simuleringer med utvalgsstørrelser på 8 og 40, med en populasjonsgjennomsnitt på 0 (nullhypotesen). Antall p-verdier mindre enn 0.05 representerer falske positive.



Utvalgsstørrelse 8: 58 falske positive.

Utvalgsstørrelse 40: 46 falske positive.

Resultatene viser at med et større utvalg reduseres antall falske positive noe, men ikke like mye som forventet. Selv om teorien antyder at større utvalg gir mer nøyaktige estimater og reduserer utvalgsvariasjonen, eliminerer de ikke falske positive helt. Forskjellen i falske positive mellom de to utvalgsstørrelsene (58 mot 46) er liten, men samsvarer fortsatt med den forventede raten for falske positive (rundt 5 %) over tid.

## 4 Study Design

### 4.1 Study design

Temaet jeg har valgt å ta for meg er intervalltrenings påvirkning på det maksimale oksygenopptaket ( $VO_{2\text{maks}}$ ).  $VO_{2\text{maks}}$  kan defineres som kroppens maksimale evne til å ta opp og omsette oksygen, altså den maksimale hastigheten på den aerobe energiomsetningen.

Fick's ligning illustrerer de bestemmende faktorene for  $VO_{2\text{maks}}$ , som er minuttvolum (MV), arteriell oksygen og venøs oksygen.

$VO_{2\text{maks}}$  har en sterk assosiasjon til prestasjon og toppidrett, men sees også på som en sterk indikator for generell helse. Kunnskap knyttet til dette temaet omhandler derfor ikke kun prestasjon innen toppidretten, men også den generelle helsen i befolkningen. En lav  $VO_{2\text{maks}}$  sees på som en sterk prediktor for tidlig død (Kodama (2009); Bassett and Howley (1997); Sparks et al. (2024); Lee et al. (2010)).

Intervalltrening er en effektiv og tidsbesparende måte å øke sin  $VO_{2\text{maks}}$  på, og kan derfor være enklere og smartere å prioritere for de som ønsker en bedre utholdenhet og sterkere helse. Det er nyttig å vite hvilken intervalltrening som er mest effektiv, enten det er kortere intervalldrag, eller lengre. Det er dette jeg skal undersøke i denne oppgaven.

Forskerne i studiene jeg har valgt ut prøver alle å finne ut hvilken intervallform som best øker  $VO_{2\text{maks}}$  og utholdenhetsprestasjonen. Studiene har noe ulik tilnærming til temaet, og ulike hypoteser og problemstillinger. Duffield, Edge, and Bishop (2006) undersøkte hvordan høyintensiv intervalltrening (HIT) påvirket  $VO_{2\text{maks}}$  under kraftig, konstant belastning. Montero et al. (2015) testet hypotesen om at økningen i  $VO_{2\text{maks}}$  etter 6 uker med moderat utholdenhetstrening skyldtes hematologiske tilpasninger, ikke muskeltilpasninger. Sindiani et al. (2017) sammenliknet effekten av to ulike intervalltreningsprogrammer. Formålet var å finne ut om rekkefølgen på intervallene, til tross for totalt likt arbeid, ville gi ulike fysiologiske effekter. Helgerud et al. (2007) sammenliknet effekten av utholdenhetstrening med ulik intensitet og metode, men med likt totalt arbeid. Hypotesen deres var at trening ved lav intensitet (LIT) og moderat intensitet (MIT) ville gi mindre effekt på  $VO_{2\text{maks}}$  enn HIT. Rønnestad et al. (2020) sammenliknet effektene av en ukes HIT-mikrosyklus med enten korte eller lange intervalldrag. Hypotesen deres var at korte intervaller ville gi bedre effekt på utholdenhetsprestasjon enn lange intervaller. Denne studien sammenliknet effekten av en ukes HIT-mikrosyklus med enten korte eller lange intervaller. Hypotesen var at korte intervaller ville gi bedre forbedringer i indikatorer på utholdenhetsprestasjon enn lange intervaller.

Likt for alle studiene er at de har en hypotese om at HIT trening gir bedre fysiologiske tilpasninger enn trening med lavere intensitet, da kortere intervaller ofte vil gi et høyere treningsindusert stress og større fysiologiske adaptasjoner.

Alle de fem studiene er intervensjonsstudier som undersøker effekten av ulik intensitet på intervalltreningen på VO<sub>2</sub>maks. Studiene til Duffield, Edge, and Bishop (2006) og Rønnestad et al. (2020) har benyttet seg av et crossover-design, mens de andre benyttet seg av et randomisert kontrollstudie-design.

Utvalgsstørrelse i studiene varierer stort, fra 10 til 55 deltakere, og det er lite informasjon å oppdrive om hvordan de innhentet deltakere til studiene sine. Ingen av studiene opplyser om de har gjort statistiske analyser for å bestemme nødvendig utvalgsstørrelse. Dette er en viktig begrensning, fordi små utvalg ofte kan føre til lav statistisk styrke og gir økt risiko for type II feil. Både lite informasjon om rekruttering og hvordan de beregnet utvalgsstørrelse er svakheter i studiene. Sannsynligvis er det blitt gjort, men det er ikke opplyst om i artiklene. Det gjør at en kan stille spørsmålsteget om hvorvidt dataene deres er generaliserbare for en større del av populasjonen. Når en studie har et lite eller ikke-representativt utvalg, kan det være utfordrende å overføre resultatene til den større befolkningen. For eksempel, hvis deltakerne er valgt på en måte som ikke reflekterer eller representerer den totale populasjonen (som f.eks idrettsstudenter i forskningsprosjekter på trening), kan resultatene være ikke-representative for befolkningen ellers med et lavere fysisk nivå eller utgangspunkt. Det er viktig med informasjon om hvordan deltakerne ble rekruttert for å vite om utvalget er representativt for den gruppen man ønsker å generalisere til.

En utilstrekkelig beregnet utvalgsstørrelse kan føre til at studien mangler statistisk styrke, som er evnen til å oppdage en faktisk effekt hvis den eksisterer. Uten en presis beregning av nødvendig utvalgsstørrelse risikerer man å ha for få deltakere til å oppdage små og betydningsfulle effekter, og øker igjen risikoen for type II feil.

Samlet sett er informasjon om rekruttering og beregning av utvalgsstørrelse avgjørende for å vurdere studienes vitenskapelige styrke, troverdighet og hvorvidt resultatene kan overføres til en bredere befolkning.

Rønnestad et al. (2020) skriver at 18 mannlige deltakere på nasjonalt nivå enten i terrengsykling eller landeveissykling meldte seg frivillig til å delta i studiet. Duffield, Edge, and Bishop (2006) opplyser kun om at de rekrutterte 10 aktive og friske kvinner frivillig. Montero et al. (2015) opplyser om at de rekrutterte 16 utrente, men friske menn til studien sin. Sindiani et al. (2017) og kolleger beskriver at de rekrutterte 40 friske, unge menn i alderen 22-25 som var idrettsstudenter. Helgerud et al. (2007) har også rekruttert 45 mannlige studenter, som alle var aktive i idrett fra før av.

Studiene er designet på ulike måter, og varierer i lengde på 3 til 8 uker. Samtlige studier er lagt opp med en testperiode før treningsintervensjonen, og en testperiode etter treningsintervensjonen. Testene ble gjennomført pre- og post-intervensjon, deriblant VO<sub>2</sub>maks-test, laktatprofil, arbeidsøkonomi, flebotomi, hemoglobinmasse og konsentrasjon, m.m. VO<sub>2</sub>maks er den mest

relevante variabelen å se på i henhold til mitt tema, og responsen en får på ulike typer intervalltrening. Alle de andre variablene som er målt er høyst aktuelle og knytter seg sterkt opp til VO2maks responsen. Parede og u-parede t-tester og ANOVA er den vanligste statistiske analysen og ble brukt i samtlige studier for å sammenlikne effekten mellom grupper. Det ble også brukt Cohen´s d for å finne effektstørrelse, som brukes for å beskrive størrelsen på en forskjell mellom to grupper, uavhengig av utvalgsstørrelse. Effektstørrelsen gir en indikasjon på hvor stor eller liten en effekt er, og brukes for å vurdere betydningen av forskningsfunn i tillegg til statistisk signifikans.

Samtlige studier konkluderte med at trening med høyere intensitet er bedre for å øke VO2maks enn trening med lavere intensitet eller kontinuerlig arbeid. Det er brukt både parede og uavhengige t-tester for å sammenlikne forskjeller mellom to grupper, og ANOVA for å studere forskjellene mellom tre eller flere grupper, for å finne ut om det er en signifikant forskjell mellom de, og om det er statistisk betydningsfullt. Dette er gjort i tråd med oppbygningen av studiedesignene i disse studiene.

Helgerud et al. (2007) sammenliknet effektene av 4 ulike intervaller sin påvirkning på VO2maks, og fant at 4x4 min intervaller og 15x15 sekund økte VO2maks signifikant i forhold til gruppen som hadde lav intensitet eller trente på laktatterskel med henholdsvis 7,3% og 5,5%. Rønnestad et al. (2007) fant data som støttet hypotesen deres om at kortere intervaller gir bedre adaptasjoner til utholdenhet og prestasjon enn lengre intervaller. Dog økte begge gruppene i VO2maks, men gruppen med kortere intervaller økte mest innad i gruppen. De resterende studiene er enig i at intervaller gjennomført på en høy intensitet gir bedre utholdenhetsadaptasjoner enn LIT eller kontinuerlig arbeid, for å forbedre VO2maks hos både moderat godt trente og toppidrettsutøvere.

# 5 Philosophy of Science

## 5.1 Introduksjon

Styrketrening er ment som all trening som utvikler eller vedlikeholder vår evne til å skape størst mulig kraft ved en spesifikk eller forutbestemt hastighet, eller enkelt sagt øke eller vedlikeholde muskelmasse eller -styrke. Treningsformen består av sett bestående av x antall repetisjoner, som sammen med belastning og treningsfrekvens utgjør det totale treningsvolumet. I tillegg defineres treningsintensitet i styrketrening som en prosent av en repetisjon maksimum (%RM) eller repetisjoner i reserve. Resultatet man får av systematisk styrketrening er muskler som både blir større og sterkere. Årsaken til dette er at muskelfiberne øker i tverrsnitt (hypertrofi) grunnet fysiologiske adaptasjoner som skjer i muskulaturen knyttet til styrketrening. Hos utrente ser man typisk en økning i muskelfibertverrsnitt på  $\sim 0,1-0,5\%$  per treningsøkt og en økning i muskelstyrke på  $\sim 1\%$  per treningsøkt i de første 8-12 ukene av en styrketreningsintervensjon (Wernbom, Augustsson, and Thome?? (2007)).

Hvilken treningsmetode som er mest optimal for å øke muskelmasse og -styrke er et omdiskutert og interessant tema. På lik linje med andre treningsformer vil den mest optimale metoden være individuelt og avhengig av den enkeltes utgangspunkt og mål med treningen. Mange opplever å ha stillesittende arbeid kombinert med lite tid til overs til trening, er ikke bra for helsa, og det er viktig at fysisk aktivitet i større grad prioriteres. Det er for mange med en travel hverdag viktig å finne den treningsformen som gir god effekt, samtidig som den er tidseffektiv. Studier viser at både muskelmasse og -styrke forfaller betrakkelig etter fylte 50 år (Deschenes (2004); Janssen et al. (2000)), som utvikler seg til sarkopeni dersom man ikke påvirker prosessen ved å øke muskelproteinsyntesen. Sarkopeni er tap av muskelmasse som er aldersrelatert, som ofte begynner i 50-årene. Dette reduserer livskvaliteten og funksjonsevnen, samt øker risiko for sykdommer og tidlig død (Beaudart et al. (2017); Deschenes (2004)). Styrketrening er vist å ha god utvikling på både muskelstyrke og muskelmasse hos både menn og kvinner (Cureton et al. (1988)).

Det er flere studier som har undersøkt hvilken forskjell ett sett og tre sett utgjør. (Rhea-2002?) fant at tre sett ga signifikant større økning i styrke enn ett sett. Rønnestad et al. (2007) fant tilsvarende resultater, men også at tre sett ga bedre effekter på hypertrofi enn ett sett.

Med vissheten om denne kunnskapen er hensikten med denne studien å undersøke forskjellen, eventuelt hvor stor forskjellen er, på styrketrening med ett sett eller tre sett, på voksne kvinner og menns utvikling av styrke og muskelmasse.

	Kvinner		Menn	
	Inkludert	Ekskludert	Inkludert	Ekskludert
Antall	18	4	16	3
Alder	22 (1.3)	22.9 (1.6)	23.6 (4.1)	24.3 (1.5)
Vekt (kg)	64.4 (10)	64.6 (9.7)	75.8 (11)	88.2 (22)
Høyde (cm)	168 (6.9)	166 (7.6)	183 (5.9)	189 (4.6)

Verdiene er presentert som gjennomsnitt og standardavvik

## 5.2 Metode

### 5.2.1 Deltakerne

Det ble rekruttert 41 kvinner og menn med alder 18-40 år til å delta i studien. Deltakerne måtte være ikke-røykere, ikke ha trent mer enn en økt i uka de siste 12 månedene, ikke ha fysiske skader som hindret de i å gjennomføre treningen, ikke være allergisk mot bedøvelse, og heller ikke bruke medisiner som kunne påvirke treningsadapsjonene. Grunnet at deltakerne ikke hadde gjennomført mer enn 85% av øktene av ulike årsaker, ble syv deltakere ekskludert fra dataanalysene.

### 5.2.2 Styrketrening

Treningsintervensjonen bestod av 12 uker styrketrening for hele kroppen med en belastning på 7-10 RM. Øvelsene på underkroppen ble utført unilateralt, altså at ett ben jobber om gangen, for å kunne gjøre within subject differensiering på treningsvolum. Hvilket ben som skulle ha ett sett, og hvilket som skulle ha tre sett, ble bestemt tilfeldig. Pre og post intervensjon ble muskelstyrke målt, samt ved uke 3, 5, og 9. Pre og post ble også kroppsammensetning målt.

### 5.2.3 Tester

Muskelstyrke ble målt som maksimal styrke, derav en repetisjon maksimum (1RM) i unilateral benpress. Testen hadde en standardisert oppvarming, som bestod av (estimert) 50, 75 og 85% av 1RM. Motstanden økte gradvis helt til deltakeren ikke klarte å gjennomføre en repetisjon, og fikk 4-6 forsøk. Den høyeste vekten deltakeren klarte ble deres 1RM. Kroppsammensetning ble målt med Dual Energy x-ray Absorptiometry (DXA) ved hjelp av en standardisert protokoll. Deltakerne kom fastende og hadde ikke gjennomført noe anstrengende aktivitet siste 48 timene.

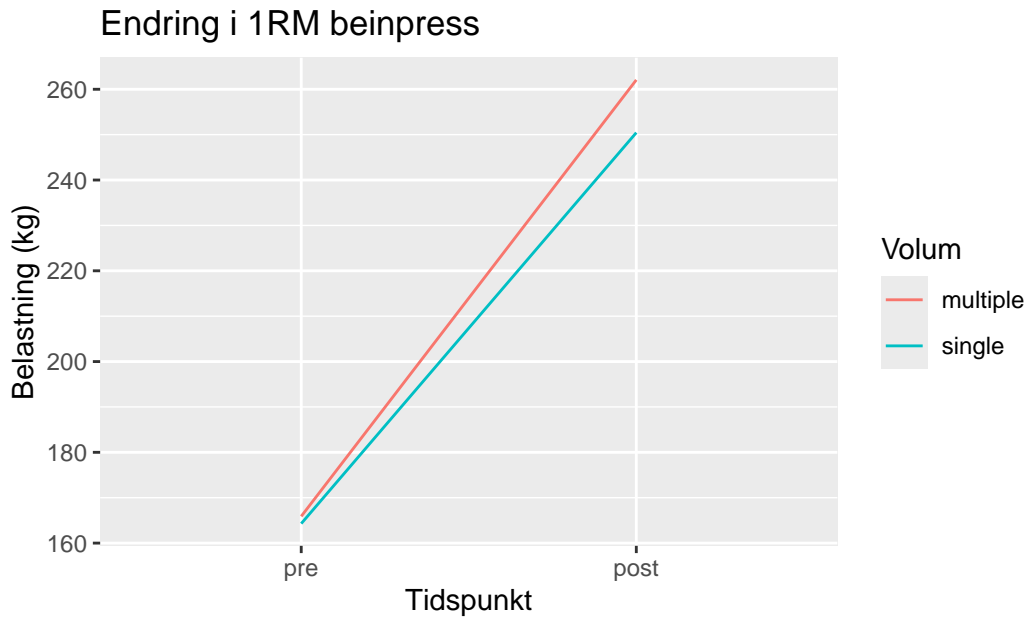
## 5.2.4 Data

Det ble gjennomført en parret t-test på differansen mellom ett og tre sett på lean mass i beina, og på differansen mellom ett og tre sett på 1RM unilateral benpress. Figurene viser den gjennomsnittlige endringen fra pre til post i absolutte verdier for både ett og tre sett i lean mass og 1 RM i unilateralt beinpress. Data er analysert ved bruk av RStudio 2024.04.2.



Figuren viser gjennomsnittlig endring i lean mass i beina fra pre til post

Resultatene fra den parrede t-testen viser at det er en signifikant forskjell mellom ett og tre sett på lean mass i beina med en p-verdi på 0.036. Figuren viser at begge beina hadde tilnærmet likt utgangspunkt ved pre-test. Begge beina har hatt fremgang, men tre sett har en større økning enn ett sett.



Figuren viser gjennomsnittlig endring i 1RM unilateral beinpress fra pre til post

Resultatene fra den parrede t-testen viser at det er en signifikant forskjell mellom ett og tre sett på maksimal styrke i unilateral benpress med p-verdi på 0,025. Figuren viste at begge bena var tilnærmet like sterke ved pre-test. Begge bena hadde fremgang, men tre sett hadde en større økning i maksimal styrke enn ett sett.

### 5.3 Diskusjon

I studien min ble det undersøkt forskjellen mellom ett sett og tre sett styrketrening utgjorde på maksimal styrke og muskelmasse i bena hos kvinner og menn. Dette ble besvart med at tre sett ga en større økning i maksimal styrke og muskelmasse enn det ett sett gjorde. Resultatene viste at benet som trente tre sett hadde en signifikant større endring i maksimal styrke i unilateral benpress og lean mass enn det benet som trente ett sett. Resultatene våre stemmer overens med Krieger (2009) og Krieger (2010) som også fant at styrketrening med moderat treningsvolum ga bedre effekt på muskelmasse og -styrke enn styrketrening med lavt treningsvolum. De fysiologiske prosessene som skjer i kroppen når man trener styrke ser ut til å bli påvirket av mer enn bare styrketreningen man gjør, men også av treningsvolumet, der det foretrekkes et høyere volum fremfor et lavere volum Schoenfeld, Ogborn, and Krieger (2016).



## 5.4 Konklusjon

Til tross for dette hadde også benet som kun trente ett sett, god fremgang i både maksimal styrke og muskelmasse. Det er altså ikke bortkastet å kun trene ett sett, men det er bedre å trene tre sett enn ett sett.

## 6 Philosophy of Science

### 6.1 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.**

David Hume var en filosof på 1700-tallet som blant annet er kjent for meningene sine om bruken av induksjon. Han mente det er umulig å rasjonelt begrunne bruken av induksjon og argumenterer for dette i all hovedsak ved å bruke to premisser: - All begrunnelse må enten være induktiv eller deduktiv - Hverken deduktiv eller induktiv begrunnelse kan rettferdiggjøre induksjon.

Det første premisset begrenser kunnskap til å stamme fra erfaring (induksjon) eller fornuften (deduksjon). Den induktive begrunnelsen gir kunnskap om verden basert på en antagelse om at fremtiden vil likne på fortiden. Den deduktive begrunnelsen gir sikker kunnskap, men begrenses til logiske sannheter.

Det andre premisset hevder verken induksjon eller deduksjon kan gi et klart grunnlag for å tro at induksjon er pålitelig. Induktivt blir et argument som er sirkulært, og deduktivt er det ikke mulig å bevise at fremtiden vil likne på fortiden.

Jeg har valgt abduksjon, som også er kalt «slutning til beste forklaring», som en innvending mot Humes argument. Abduksjon handler om å velge hypotesen som på best måte forklarer de observerte dataene. Hume kunne likevel argumentere for at abduksjon fortsatt er avhengig av induksjon. Skal vi vurdere hvilken hypotese som vil gi den beste forklaringen må vi gå ut ifra at fortidens mønstre vil fortsette i fremtiden.

Hvis vi for eksempel observerer at en plante trives i sollyset, kan vi abduktivt komme frem til at sollys er en nødvendighet for plantevekst. Imidlertid forutsetter denne konklusjonen at fremtidige planter vil reagere på samme måte på sollys, som de har gjort i fortiden, og dette er en antagelse som er induktiv.

Argumentet til Hume påpeker et grunnleggende problem i epistemologien. Hvordan kan vår tro på at fremtiden vil likne fortiden rettferdiggjøres? Selv om det ikke finnes et logisk bevis for dette, kan det argumenteres for at induksjon i praksis har vist seg å være en pålitelig metode. Evnen vår til å forutsi og kontrollere verden, fra hverdagslige handlinger til vitenskapelige oppdagelser, kan det tyde på at induksjon har en viss grad av gyldighet.

Poenget til Hume er ikke nødvendigvis at vi bør kutte bruken av induksjon. Han argumenterer for at vi må godta at induksjon ikke har et absolutt rasjonelt grunnlag. Kunnskapen vår om verden, basert på induksjon, er derfor alltid preget av en viss grad med usikkerhet.

## 6.2 Oppgave 2

**Forklar de grunnleggende ideene i Bayesianisme, samt hvordan Bayesianske sannsynligheter kan tolkes. Presenter så ett problem med Bayesianisme og evaluér hvor alvorlig problemet er.**

Bayesianisme har blitt en svært populær tilnærming til statistikk og matematisk sannsynlighetsteori, basert på Bayes' teorem (Godfrey-Smith (2003)). Grunnlaget for Bayes Teorem og bayesianismen er filosof og statistiker, Thomas Bayes. Bayes teorem skrives som følger: Dette teoremet beskriver en metode for å oppdatere en «prior sannsynlighet» for en hypotese, gitt ny informasjon og data, for å få en «posterior sannsynlighet». Det handler altså om sannsynligheten for om en hendelse skal skje basert på tidligere observasjoner, det er dette som er kalt prior-sannsynlighet (Godfrey-Smith (2003)). I slike analyser i Bayesianismen brukes nettopp denne tilnærmingen til å sannsynliggjøre utfallet av en hendelse før ny data observeres. Etter hvert som ny informasjon blir tilgjengelig oppdateres den såkalte posterior-sannsynligheten, ved hjelp av Bayes Teorem. Dette betyr at man kan kontinuerlig regne seg frem til sannsynligheten for at noe skal skje. Hvis den nye informasjonen støtter den opprinnelige hypotesen, vil sannsynligheten for at dette skjer igjen øke, og vice versa.

Som følge av denne prosessen, vil en bayesianer sjelden hevde at noe er en absolutt sannhet, men heller gi uttrykk for hvor sannsynlig det er å kunne tro på at noe er en sannhet, kun basert på matematiske utregninger. For eksempel kan dette bety at dersom tidligere forskning på et område er noe ambivalent, er et utfall 60% sannsynlig. Dersom ny forskning på området kommer på banen med data som støtter hypotesen, kan sannsynligheten for dette økes til eks. 80%. På denne måten oppdaterer man kunnskapen sin kontinuerlig og naturlig over tid.

En slik holistisk tilnærming til vitenskapen er sett på som en styrke med Bayesianismen, sammenlignet med for eksempel HD-metoden. Denne metoden baserer seg i stor grad på en binær tilnærming til hypotese, som krever at en hypotese enten falsifiseres eller verifiseres gjennom testing. En utfordring ved HD-metoden er at den ikke tar noen høyder for usikkerheter eller andre gradvise endringer i forskningsfeltet. Derfor er Bayesianismen bedre egnet til å tolke slike endringer ved å tillate fleksibilitet i form av oppdaterte sannsynlighetsberegninger.

Selv om Bayesianisme gir en dynamisk og oppdatert forståelse på ulike forskningsspørsmål, har tilnærmingen også svakheter. Et problem er hvor man skal hente all den tidligere sannsynligheten fra, særlig når temaet er svakt representert i forskningsfeltet og har lite tidligere data å vise til. Temaer eller forskningsfelt det er svært lite data fra, som det er forsket lite på, er svært sårbare. Forskning av lav kvalitet eller begrenset omfang kan påvirke resultatene i stor grad, altså posterior-resultatene. Det kan tenkes at med mer informasjon og ny data vil

dette utjevnes, men man skal være obs på at dette kan føre til skjevheter, bias. I tilfeller det det er gjort en del forskning på et tema kan det være en utfordring å få et klart bilde av prior-sannsynligheten, som kan føre til misvisende eller ufullstendige konklusjoner, men også at ulike personer kan bruke ulik prior-sannsynlighet for de samme eksperimentene (K. Hackenberger (2019)) . Det er vanskelig å si noe om hvor store konsekvenser dette kan få, og det er nok avhengig av situasjonen.

Bayesianismen gir samlet sett en fleksibel tilnærming til forskningsspørsmål, men krever en nøye vurdering av hvilke data og tidligere sannsynligheter som brukes og legges til grunn for arbeidet. På en side gir metoden interessante muligheter til å utforske samme eksperiment med ulike prior-sannsynligheter og hypoteser justeres i takt med nye funn. På en annen side kan det være en utfordring å unngå subjektive valg som påvirker analysen.

## 7 Laboratory report

### 7.1 Hensikt:

Ekstrahere DNA fra hel-blod og analysere ACTN3 genotypen ved hjelp av PCR og elektroforese-separasjon av DNA-fragmenter.

Hensikten med dette er å undersøke genetiske variasjoner som kan påvirke fysisk prestasjon, spesielt muskelfunksjon. ACT3N-genet koder for et protein kalt alfa-aktinin-3 som er mye uttrykt i type II muskelfibre. Variasjoner i dette genet kan ha betydning for et individs atletiske prestasjon, spesielt aktiviteter som krever eksplosiv styrke eller sprint. Testen brukes ofte innen idrettsforskning for å forstå den genetiske bakgrunnen for atletisk kapasitet, da spesielt innen styrkeøvelser og sprint. Ved å analysere ACTN3-genotypen kan man få en innsikt i hvordan genetiske faktorer påvirker muskelprestasjon og om visse genetiske varianter er assosiert med bedre prestasjon i spesifikke idretter eller øvelser.

### 7.2 Metode:

**Protokollen for DNA-ekstraksjon fra fullblod kan beskrives som følgende:**

Blod samles i heparin- eller EDTA-rør og må behandles samme dag hvis det oppbevares ved romtemperatur. Tre milliliter fullblod overføres til et 15 mL rør, og 12 mL av Reagens A tilsettes. Blandingen roteres forsiktig i fire minutter ved romtemperatur. Prøvene fraktes ned fra molekylab til testlab for sentrifugering.

Prøven sentrifugeres deretter ved 3000 g i fem minutter, og supernatanten fjernes forsiktig uten å forstyrre cellepellet. Eventuelle rester av væske fjernes ved blotting eller pipettering. Avvik-10min + 5 min 1400g Prøvene fraktes opp fra testlab til molekylab for videre behandling. Avvik ca 35 grader i 20 min, + 10 min 70 grader

Cellepelletet resuspenderes i 1 mL Reagens B, før 250  $\mu$ L 5 M natriumperklorat tilsettes. Røret vendes flere ganger for å blande, og prøven inkuberes i vannbad på 65 °C i 15–20 minutter før den avkjøles til romtemperatur. Avvik Ingrid- tok over løsning i 1,5ml rør- sentrifugerte på nytt hastighet 6 temperatur 21 grader =for å at DNA som er tygst skulle legge seg på bunnen igjen.

Deretter tilsettes 2 mL iskald kloroform, og blandingen roteres i 30–60 minutter. Dette trinnet må utføres i et avtrekksskap, da kloroform er giftig. Etter sentrifugering ved 2400 g i to minutter overføres den øvre vandige fasen, som inneholder DNA, til et nytt rør.

For å felle ut DNA tilsettes 2–3 mL avkjølt 100 % etanol, og røret vendes. Det utfelte DNA-et overføres til et 1,5 mL rør og lufttørkes før det resuspenderes i 200 L TE-buffer.

Til slutt måles DNA-konsentrasjonen med et spektrofotometer. Forventet utbytte er mellom 200 og 500 ng/ L.

### **Bestemmelse av ACTN3-genotypen involverer følgende trinn:**

DNA ble utvunnet fra prøver, enten fra fullblod eller muskelvev. Master mix og primer mix ble forberedt på forhånd og gjort klart for videre bruk. For hver reaksjon blandes prøven med master mix og primer mix i et PCR-reaksjonsrør for å gjennomføre PCR (polymerasekjedereaksjon).

Resultatene analyseres videre ved hjelp av generell protokoll for agarosegelelektroforese for DNA, hvor prøvene separeres basert på størrelse.

### **Generell protokoll for DNA Agarosegelelektroforese**

#### **Forberedelse av gel**

Start med å fortynne 10X TBE-buffer til en 1X-løsning ved å blande 100 ml konsentrert buffer med 900 ml vann. En av gruppene gjennomførte dette trinnet. Deretter tilsettes 100 ml 1X TBE i en konisk kolbe, sammen med agarose for å lage en gel med ønsket konsentrasjon, for eksempel 2 g for en 2 % gel. Dette steget ble utført av oss.

Dersom du bruker Sybr Safe fargestoff, tilsettes det i en konsentrasjon på 1:10 000 (10 µl per 100 ml). Blandingen varmes opp til løsningen blir klar, for eksempel i mikrobølgeovn eller på varmeplate. Vi varmet den opp i omtrent 3 minutter på middels styrke i mikrobølgeovnen. La blandingen avkjøles til rundt 60 grader – den skal kjennes varm, men ikke varm nok til å brenne.

Hell blandingen i en støpeform for gel, og plasser kammen på sin dedikerte plass. Gelen polymeriseres i løpet av én time, og kammen fjernes etterpå.

#### **Kjøring av gelen**

Plasser gelen i elektroforeseapparatet og hell 1X TBE-buffer i tanken slik at brønnene dekkes. Bland prøvene med ladningsfarge, for eksempel 1 µl av 6X farge per 5 µl prøve. I vårt tilfelle brukte vi 2 µl 6X farge per 10 µl prøve.

Tilsett DNA-ladder og prøver i brønnene. Dette steget tok litt tid, og det kan hende en prøve havnet utenfor brønnen. Sett spenningen til 150 V og kjør gelen i omtrent 1 time, eller til fargen har beveget seg rundt 80 % gjennom gelen. Pass på at prøvene beveger seg mot den positive elektroden! Vår kjøring tok 54 minutter.

#### **Visualisering**

Gelen visualiseres i et G:Box-system ved bruk av UV-lys og Sybr green-innstillinger.

## 7.3 Resultater:

(Resultatene hadde manglende data, ved at det manglet et bånd. Dette resulterte i at vi ikke klarte å identifisere ACTN3 genotype, hvor genotype R/R produserer to bånd på 690 og 413 bp, mens en X/X genotype produserer to bånd på 690 og 318 bp. Feilkilder kan være en årsak til hvorfor vi ikke fikk ønsket resultat).

Resultatene til 2 av medlemmene av gruppa viser 318 bp altså X/X , mens 1 viser ingenting og 1 viser 413 bp eller R/R. Simon var den eneste som målte R/R ettersom han hadde 413 bp. Ingrid var den eneste på gruppen som ikke fikk registrert noe resultat, dette kan ha en sammenheng med trinn 5 i protokollen hvor hun hadde et avvik. De forskjellige avvikene som er nevnt kan ha ført til at resultatene ble som de ble.

## 7.4 Diskusjon

### Mulige feilkilder:

Flere faktorer kan ha bidratt til feil eller uventede resultater under DNA-ekstraksjonen og analysen. Effektiv lysering av cellene er avgjørende for å frigjøre DNA, og dersom lysereagensene ikke ble tilsatt korrekt, eller inkubasjonstiden var for kort, kan det ha resultert i ufullstendig lysering og lavt DNA-utbytte. I tillegg kan DNA ha blitt degradert under ekstraksjonsprosessen, spesielt ved overdreven vortexing eller eksponering for høye temperaturer, noe som kan skade DNA-strengene.

Videre kan problemer ha oppstått under etanolutfellingen, som feil mengde etanol, upassende saltbuffer eller utilstrekkelig sentrifugering. Disse feilene kan forhindre korrekt utfelling av DNA. Under påvisningsprosessen, for eksempel ved bruk av gelelektroforese, kan feil i teknikken ha skjedd, som feil mengde DNA i brønnene, feil gelkonsentrasjon eller problemer med fargestoffet som brukes for visualisering. Dersom DNA er tilstede, men bare ett bånd er synlig i stedet for to eller tre som forventet, kan dette skyldes spesifikke problemer med DNA-ekstraksjonen eller analysen.

Ulike trinn i analysen kan ha bidratt til avvikene. Ineffektiv enzymatisk behandling, for eksempel lav aktivitet av restriksjonsenzymer, kan resultere i ukuttet DNA som kun viser ett bånd. Problemer under PCR-amplifikasjon, som suboptimale betingelser eller manglende spesifisitet i primerbindingen, kan også føre til at kun deler av DNA blir forsterket. Feil under gelelektroforese, som feil konsentrasjon av gelen eller utilstrekkelig separasjon av fragmenter, kan også påvirke resultatene.

Til slutt kan kontaminasjon eller feil pipettering føre til unøyaktige resultater. Dette kan forklare hvorfor bare ett bånd er synlig, selv om flere var forventet. Disse mulige feilene

understreker viktigheten av nøye håndtering gjennom hele prosessen for å sikre pålitelige resultater.



## 8 Referanser

- Bassett, D. R., & Howley, E. T. (1997). Maximal oxygen uptake: Classical versus contemporary viewpoints: *Medicine & Science in Sports & Exercise*, 29(5), 591–603. <https://doi.org/10.1097/00005768-199705000-00002>
- Beaudart, C., Zaaria, M., Pasleau, F., Reginster, J.-Y., & Bruyère, O. (2017). Health Outcomes of Sarcopenia: A Systematic Review and Meta-Analysis. *PLOS ONE*, 12(1), e0169548. <https://doi.org/10.1371/journal.pone.0169548>
- Cureton, K. J., Collins, M. A., Hill, D. W., & McElhannon, F. M. (1988). Muscle hypertrophy in men and women: *Medicine & Science in Sports & Exercise*, 20(4), 338–344. <https://doi.org/10.1249/00005768-198808000-00003>
- Deschenes, M. R. (2004a). Effects of Aging on Muscle Fibre Type and Size: *Sports Medicine*, 34(12), 809–824. <https://doi.org/10.2165/00007256-200434120-00002>
- Duffield, R., Edge, J., & Bishop, D. (2006). Effects of high-intensity interval training on the response during severe exercise. *Journal of Science and Medicine in Sport*, 9(3), 249–255. <https://doi.org/10.1016/j.jsams.2006.03.014>
- Godfrey-Smith, P. (2003). *Theory and reality: An introduction to the philosophy of science*. University of Chicago Press.
- Halperin, I., Pyne, D. B., & Martin, D. T. (2015). Threats to Internal Validity in Exercise Science: A Review of Overlooked Confounding Variables. *International Journal of Sports Physiology and Performance*, 10(7), 823–829. <https://doi.org/10.1123/ijsp.2014-0566>
- Helgerud, J., Høydal, K., Wang, E., Karlsen, T., Berg, P., Bjerkaas, M., Simonsen, T., Helgesen, C., Hjorth, N., Bach, R., & Hoff, J. (2007). Aerobic High-Intensity Intervals Improve V̇O<sub>2</sub>max More Than Moderate Training. *Medicine & Science in Sports & Exercise*, 39(4), 665–671. <https://doi.org/10.1249/mss.0b013e3180304570>
- Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine (Auckland, N.Z.)*, 30(1), 1–15. <https://doi.org/10.2165/00007256-200030010-00001>
- Janssen, I., Heymsfield, S. B., Wang, Z., & Ross, R. (2000). Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr. *Journal of Applied Physiology*, 89(1), 81–88. <https://doi.org/10.1152/jappl.2000.89.1.81>
- K. Hackenberger, B. (2019). Bayes or not Bayes, is this the question? *Croatian Medical Journal*, 60(1), 50–52. <https://doi.org/10.3325/cmj.2019.60.50>

- Kodama, S. (2009). Cardiorespiratory Fitness as a Quantitative Predictor of All-Cause Mortality and Cardiovascular Events in Healthy Men and Women: A Meta-analysis. *JAMA*, 301(19), 2024. <https://doi.org/10.1001/jama.2009.681>
- Krieger, J. W. (2009). Single Versus Multiple Sets of Resistance Exercise: A Meta-Regression. *Journal of Strength and Conditioning Research*, 23(6), 1890–1901. <https://doi.org/10.1519/JSC.0b013e3181b370be>
- Krieger, J. W. (2010). Single vs. Multiple Sets of Resistance Exercise for Muscle Hypertrophy: A Meta-Analysis. *Journal of Strength and Conditioning Research*, 24(4), 1150–1159. <https://doi.org/10.1519/JSC.0b013e3181d4d436>
- Lee, D., Artero, E. G., Xuemei Sui, & Blair, S. N. (2010). Review: Mortality trends in the general population: the importance of cardiorespiratory fitness. *Journal of Psychopharmacology*, 24(4\_suppl), 27–35. <https://doi.org/10.1177/1359786810382057>
- Machado, F. A., Nakamura, F. Y., & Moraes, S. M. F. D. (2012). Influence of regression model and incremental test protocol on the relationship between lactate threshold using the maximal-deviation method and performance in female runners. *Journal of Sports Sciences*, 30(12), 1267–1274. <https://doi.org/10.1080/02640414.2012.702424>
- Montero, D., Cathomen, A., Jacobs, R. A., Flück, D., De Leur, J., Keiser, S., Bonne, T., Kirk, N., Lundby, A., & Lundby, C. (2015). Haematological rather than skeletal muscle adaptations contribute to the increase in peak oxygen uptake induced by moderate endurance training. *The Journal of Physiology*, 593(20), 4677–4688. <https://doi.org/10.1113/JP270250>
- Rhea, M. R., Alvar, B. A., Ball, S. D., & Burkett, L. N. (2002). Three sets of weight training superior to 1 set with equal intensity for eliciting strength. *Journal of Strength and Conditioning Research*, 16(4), 525–529.
- Rønnestad, B. R., Egeland, W., Kvalheim, N. H., Refsnes, P. E., Kadi, F., & Raastad, T. (2007). DISSIMILAR EFFECTS OF ONE- AND THREE-SET STRENGTH TRAINING ON STRENGTH AND MUSCLE MASS GAINS IN UPPER AND LOWER BODY IN UNTRAINED SUBJECTS: *Journal of Strength and Conditioning Research*, 21(1), 157–163. <https://doi.org/10.1519/00124278-200702000-00028>
- Rønnestad, B. R., Hansen, J., Nygaard, H., & Lundby, C. (2020). Superior performance improvements in elite cyclists following short-interval vs effort-matched long-interval training. *Scandinavian Journal of Medicine & Science in Sports*, 30(5), 849–857. <https://doi.org/10.1111/sms.13627>
- Schoenfeld, B. J., Ogborn, D., & Krieger, J. W. (2016). Effects of Resistance Training Frequency on Measures of Muscle Hypertrophy: A Systematic Review and Meta-Analysis. *Sports Medicine*, 46(11), 1689–1697. <https://doi.org/10.1007/s40279-016-0543-8>
- Sindiani, M., Eliakim, A., Segev, D., & Meckel, Y. (2017). The effect of two different interval-training programmes on physiological and performance indices. *European Journal of Sport Science*, 17(7), 830–837. <https://doi.org/10.1080/17461391.2017.1321687>

- Sparks, J. R., Wang, X., Lavie, C. J., Zhang, J., & Sui, X. (2024). Cardiorespiratory Fitness as a Predictor of Non-Cardiovascular Disease and Non-Cancer Mortality in Men. *Mayo Clinic Proceedings*, 99(8), 1261–1270. <https://doi.org/10.1016/j.mayocp.2023.11.024>
- Tanner, R. (2012). *Physiological Tests for Elite Athletes 2nd Edition* (2nd utg.). Human Kinetics. <https://books.google.no/books?id=0OPIiMks58MC>
- Wernbom, M., Augustsson, J., & Thome??, R. (2007). The Influence of Frequency, Intensity, Volume and Mode of Strength Training on Whole Muscle Cross-Sectional Area in Humans: *Sports Medicine*, 37(3), 225–264. <https://doi.org/10.2165/00007256-200737030-00004>
- Bassett, David R., and Edward T. Howley. 1997. “Maximal Oxygen Uptake: ???classical??? Versus???contemporary??? Viewpoints.” *Medicine & Science in Sports & Exercise* 29 (5): 591–603. <https://doi.org/10.1097/00005768-199705000-00002>.
- Beaudart, Charlotte, Myriam Zaaria, Françoise Pasleau, Jean-Yves Reginster, and Olivier Bruyère. 2017. “Health Outcomes of Sarcopenia: A Systematic Review and Meta-Analysis.” Edited by James M. Wright. *PLOS ONE* 12 (1): e0169548. <https://doi.org/10.1371/journal.pone.0169548>.
- Cureton, Kirk J., Mitchell A. Collins, David W. Hill, and Fayette M. Mcelhannon. 1988. “Muscle Hypertrophy in Men and Women.” *Medicine & Science in Sports & Exercise* 20 (4): 338–44. <https://doi.org/10.1249/00005768-198808000-00003>.
- Deschenes, Michael R. 2004. “Effects of Aging on Muscle Fibre Type and Size.” *Sports Medicine* 34 (12): 809–24. <https://doi.org/10.2165/00007256-200434120-00002>.
- Duffield, Rob, Johann Edge, and David Bishop. 2006. “Effects of High-Intensity Interval Training on the Response During Severe Exercise.” *Journal of Science and Medicine in Sport* 9 (3): 249–55. <https://doi.org/10.1016/j.jsams.2006.03.014>.
- Godfrey-Smith, Peter. 2003. *Theory and Reality: An Introduction to the Philosophy of Science*. University of Chicago Press.
- Halperin, Israel, David B. Pyne, and David T. Martin. 2015. “Threats to Internal Validity in Exercise Science: A Review of Overlooked Confounding Variables.” *International Journal of Sports Physiology and Performance* 10 (7): 823–29. <https://doi.org/10.1123/ijsspp.2014-0566>.
- Helgerud, Jan, Kjetill Høydal, Eivind Wang, Trine Karlsen, Pål Berg, Marius Bjerkaas, Thomas Simonsen, et al. 2007. “Aerobic High-Intensity Intervals Improve V'O2max More Than Moderate Training.” *Medicine & Science in Sports & Exercise* 39 (4): 665–71. <https://doi.org/10.1249/mss.0b013e3180304570>.
- Hopkins, W. G. 2000. “Measures of Reliability in Sports Medicine and Science.” *Sports Medicine (Auckland, N.Z.)* 30 (1): 1–15. <https://doi.org/10.2165/00007256-200030010-00001>.
- Janssen, Ian, Steven B. Heymsfield, ZiMian Wang, and Robert Ross. 2000. “Skeletal Muscle Mass and Distribution in 468 Men and Women Aged 18–88 Yr.” *Journal of Applied Physiology* 89 (1): 81–88. <https://doi.org/10.1152/jappl.2000.89.1.81>.
- K. Hackenberger, Branimir. 2019. “Bayes or Not Bayes, Is This the Question?” *Croatian Medical Journal* 60 (1): 50–52. <https://doi.org/10.3325/cmj.2019.60.50>.

- Kodama, Satoru. 2009. "Cardiorespiratory Fitness as a Quantitative Predictor of All-Cause Mortality and Cardiovascular Events in Healthy Men and Women: A Meta-analysis." *JAMA* 301 (19): 2024. <https://doi.org/10.1001/jama.2009.681>.
- Krieger, James W. 2009. "Single Versus Multiple Sets of Resistance Exercise: A Meta-Regression." *Journal of Strength and Conditioning Research* 23 (6): 1890–1901. <https://doi.org/10.1519/JSC.0b013e3181b370be>.
- . 2010. "Single Vs. Multiple Sets of Resistance Exercise for Muscle Hypertrophy: A Meta-Analysis." *Journal of Strength and Conditioning Research* 24 (4): 1150–59. <https://doi.org/10.1519/JSC.0b013e3181d4d436>.
- Lee, Duck-chul, Enrique G Artero, Xuemei Sui, and Steven N Blair. 2010. "Review: Mortality Trends in the General Population: The Importance of Cardiorespiratory Fitness." *Journal of Psychopharmacology* 24 (4\_suppl): 27–35. <https://doi.org/10.1177/1359786810382057>.
- Machado, Fabiana Andrade, Fábio Yuzo Nakamura, and Solange Marta Franzói De Moraes. 2012. "Influence of Regression Model and Incremental Test Protocol on the Relationship Between Lactate Threshold Using the Maximal-Deviation Method and Performance in Female Runners." *Journal of Sports Sciences* 30 (12): 1267–74. <https://doi.org/10.1080/02640414.2012.702424>.
- Montero, David, Adrian Cathomen, Robert A. Jacobs, Daniela Flück, Jeroen De Leur, Stefanie Keiser, Thomas Bonne, Niels Kirk, Anne-Kristine Lundby, and Carsten Lundby. 2015. "Haematological Rather Than Skeletal Muscle Adaptations Contribute to the Increase in Peak Oxygen Uptake Induced by Moderate Endurance Training." *The Journal of Physiology* 593 (20): 4677–88. <https://doi.org/10.1113/JP270250>.
- Rønnestad, Bent R., Wilhelm Egeland, Nils H. Kvamme, Per E. Refsnes, Fawzi Kadi, and Truls Raastad. 2007. "DISSIMILAR EFFECTS OF ONE- AND THREE-SET STRENGTH TRAINING ON STRENGTH AND MUSCLE MASS GAINS IN UPPER AND LOWER BODY IN UNTRAINED SUBJECTS." *Journal of Strength and Conditioning Research* 21 (1): 157–63. <https://doi.org/10.1519/00124278-200702000-00028>.
- Rønnestad, Bent R., Joar Hansen, Håvard Nygaard, and Carsten Lundby. 2020. "Superior Performance Improvements in Elite Cyclists Following Short-Interval Vs Effort-Matched Long-Interval Training." *Scandinavian Journal of Medicine & Science in Sports* 30 (5): 849–57. <https://doi.org/10.1111/sms.13627>.
- Schoenfeld, Brad J., Dan Ogborn, and James W. Krieger. 2016. "Effects of Resistance Training Frequency on Measures of Muscle Hypertrophy: A Systematic Review and Meta-Analysis." *Sports Medicine* 46 (11): 1689–97. <https://doi.org/10.1007/s40279-016-0543-8>.
- Sindiani, Mahmood, Alon Eliakim, Daria Segev, and Yoav Meckel. 2017. "The Effect of Two Different Interval-Training Programmes on Physiological and Performance Indices." *European Journal of Sport Science* 17 (7): 830–37. <https://doi.org/10.1080/17461391.2017.1321687>.
- Sparks, Joshua R., Xuwen Wang, Carl J. Lavie, Jiajia Zhang, and Xuemei Sui. 2024. "Cardiorespiratory Fitness as a Predictor of Non-Cardiovascular Disease and Non-Cancer Mortality in Men." *Mayo Clinic Proceedings* 99 (8): 1261–70. <https://doi.org/10.1016/j.mayocp.2023.11.024>.
- Tanner, Rebecca. 2012. *Physiological Tests for Elite Athletes 2nd Edition*. 2nd ed. Human

Kinetics.

Wernbom, Mathias, Jesper Augustsson, and Roland Thome?? 2007. “The Influence of Frequency, Intensity, Volume and Mode of Strength Training on Whole Muscle Cross-Sectional Area in Humans:” *Sports Medicine* 37 (3): 225–64. <https://doi.org/10.2165/00007256-200737030-00004>.