

# Mappeeksamen IDR4000

Anders Nybakk

2024-11-22

## **Table of contents**

# 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 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. 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. 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)). Ved å kontrollere for disse variablene håper vi å oppnå en høy grad av reproduserbarhet i våre resultater.

I denne rapporten velger vi å fokusere på laktat-max og RER. Vi vil sammenlikne disse, regne ut standard avvik.

## 1.2 Protokoll

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

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 VO<sub>2</sub>maks-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).

### **1.2.1 Metode:**

VO2maks-testen ble gjennomført som en trappetrinnstest der motstanden økte med 20W hvert minutt til utmattelse, eller når RPM < 60. Det registreres målinger av VO2maks 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 gårsdagens VO2maks-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 heining, hadde hver deltaker samme testleder ved hver test.

### **1.2.2 Protokoll**

#### **1.2.2.1 Før test:**

1. Starte laktatmaskin
2. Bytte standard hvis den er tom (finnes i kjøleskap)
3. Finne frem munnstykke og sette sammen dette
4. Husk hansker
5. Papir med neseclype rundt når den er klar
6. Ta med ny slange som festes i miksekammer
7. Skru på Lode-sykkel.
8. Skru på Vyntus og Lode PC

#### **Kalibrere oksygenanalysator**

1. Gasskalibrering
2. Åpne gassflaske (lukkes når kalibrering er ferdig)
3. Sjekke at sensor er koblet i maskinen
4. Start kalibrering
5. Godtar 2% feilmargin, hvis den er høyere må man recalibrere
6. Volumkalibrering-
7. Feilmargin på 0,2% eller under godtas
8. Kalibrere kammer
9. Flytte sensor til kammer
10. Skru av gass

#### **Forberede utstyr**

1. Lage ny pasient på Vyntus og Lode
2. Idr4000\_h24\_g1\_idx
3. Veie personen i så lite klær som mulig (trekke fra 300g for klær)
4. Lage plotteark
5. Stille inn krankarm (172,5) og kalibrere sykkel på Lode PC
6. Stille inn sykkel til deltaker
7. Fullføre plotteark
8. Klargjøre laktatrør, papir og teip, samt teip til neseeklype
9. 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.**

1. VO<sub>2</sub>max – Kondisjon
2. Målinger per 30 sek.
3. Prestasjonstest – hvert sekund teller
4. Watt-økning per 1 min.
5. 80-100 rpm, stopper test hvis under 60
6. Pulsmåling
7. Borg skala – 6 = sofa, 20 = maks
8. Laktatmåling 1 min etter test – stikk i fingeren
9. Vifte?

#### **1.2.2.2 Etter test:**

1. Avslutte test på begge PC
2. Spørre om Borg skala
3. 1 minutt etter endt test skal det tas laktat-prøve. Tok 2 prøver for å sikre god reliabilitet.

#### **Hente rapport:**

1. Rapport: INN\_TABELL:30SEK\_MIX
2. Søk opp id-nr.
3. Lagre i rett mappe: F10 (nederste knapp)
4. Lagre i rett mappe på minnepenn
5. Overføre til One Drive på labbPCn

### 1.2.3 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 to er det en hviledag mellom testene.

## 1.3 Resultater

```
# Last inn nødvendige biblioteker
library(readxl)
library(dplyr)
library(tidyverse)

# Definerer variablene som skal brukes
vars <- c("id", "timepoint", "temperature", "humidity",
          "sex", "age", "height", "weight", "w.max",
          "vo2.max", "vco2.max", "rer.max", "ve.max",
          "bf.max", "hr.max", "la.max",
          "borg.max")

# Leser inn dataene fra Excel-filer og kombinerer dem
dat <- bind_rows(
  read_excel("data/g1.xlsx", sheet = "data_excel", na = "na") %>%
    select(all_of(vars)) %>%
    mutate(group = "G1",
           id = paste0(group, "_", id)) ,

  read_excel("data/g2.xlsx", na = "na") %>%
    select(all_of(vars)) %>%
    mutate(group = "G2",
           id = paste0(group, "_", id)) ,

  read_excel("data/g3.xlsx") %>%
```

```

    select(all_of(vars)) %>%
    mutate(timepoint = paste0("t", timepoint),
           group = "G3",
           id = paste0(group, "_", id)) ,

read_excel("data/g4.xlsx") %>%
  select(all_of(vars)) %>%
  mutate(group = "G4",
         id = paste0(group, "_", id)) )

# Viser de første radene av det samlede datasettet

```

### 1.3.1 Tabeller

( $n$  = antall deltagere).

#### 1.3.1.1 Tabell over demografiske variabler .

```

dat %>%
  select(timepoint, age, height, weight) %>%
  group_by(timepoint) %>%
  summarise(
    m.age = mean(age, na.rm = TRUE),
    sd.age = sd(age, na.rm = TRUE),
    m.height = mean(height, na.rm = TRUE),
    sd.height = sd(height, na.rm = TRUE),
    m.weight = mean(weight, na.rm = TRUE),
    sd.weight = sd(weight, na.rm = TRUE),
    n = n(),
    .groups = "drop"
  )

```

```

# A tibble: 4 x 8
  timepoint m.age sd.age m.height sd.height m.weight sd.weight    n
  <chr>     <dbl> <dbl>   <dbl>   <dbl>   <dbl>   <dbl> <int>
1 t1       25.7  3.64   180.    7.63    79.2    11.9   14
2 t2       25.7  3.64   180.    7.63    79.3    11.8   14
3 t3       26.8  3.69   179.    7.91    80.4    11.9   11
4 t4       29.1  7.22   175.    6.16    74.7    12.3    8

```



```
# velger etter variablene, grupperer de etter timepoint(testdagene), summerer de forskjellig
```

### 1.3.1.2 Tabell over demografiske variabler fordelt etter kjønn.

```
dat %>%
  select(timepoint, sex, age, height, weight) %>%
  group_by(timepoint, sex) %>%
  summarise(
    m.age = mean(age, na.rm = TRUE),
    sd.age = sd(age, na.rm = TRUE),
    m.height = mean(height, na.rm = TRUE),
    sd.height = sd(height, na.rm = TRUE),
    m.weight = mean(weight, na.rm = TRUE),
    sd.weight = sd(weight, na.rm = TRUE),
    n = n(),
    .groups = "drop"
  )
```

```
# A tibble: 8 x 9
  timepoint sex   m.age sd.age m.height sd.height m.weight sd.weight    n
  <chr>      <chr> <dbl> <dbl>   <dbl>   <dbl>   <dbl>   <dbl> <int>
1 t1        f     23.9  0.569   171     8.19   71.7    9.90     3
2 t1        m     26.1  4.01   182.    5.72   81.2   11.9    11
3 t2        f     24.0  0.566   171     8.19   71.5    9.95     3
4 t2        m     26.1  4.01   182.    5.72   81.4   11.8    11
5 t3        f     24.2  0.389   168.    7.78   66.7    6.79     2
6 t3        m     27.4  3.88   181.    5.67   83.5   10.7     9
7 t4        f     30.7 11.2    168.    5.69   64.0    6.85     3
8 t4        m     28.1  4.97   178.    1.67   81.2   10.2     5
```

variasjon eller reliabilitetstallet til den la.max, så den rer.max. sammenlikne til slutt

### 1.3.1.3 Standard avvik:

laktat

```

dat %>%
  group_by(timepoint) %>%
  summarise(
    m = mean(la.max, na.rm = TRUE),
    s = sd(la.max, na.rm = TRUE),
    n = n(),
    .groups = "drop")

```

```

# 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**

```

dat %>%
  group_by(timepoint, sex) %>%
  summarise(
    m = mean(la.max, na.rm = TRUE),
    s = sd(la.max, na.rm = TRUE),
    n = n(),
    .groups = "drop")

```

```

# 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**

```

dat %>%
  group_by(timepoint) %>%
  summarise(
    m = mean(rer.max, na.rm = TRUE),
    s = sd(rer.max, na.rm = TRUE),
    n = n(),
    .groups = "drop"
  )

```

```

# 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

```

dat %>%
  group_by(timepoint, sex) %>%
  summarise(
    m = mean(rer.max, na.rm = TRUE),
    s = sd(rer.max, na.rm = TRUE),
    n = n(),
    .groups = "drop"
  )

```

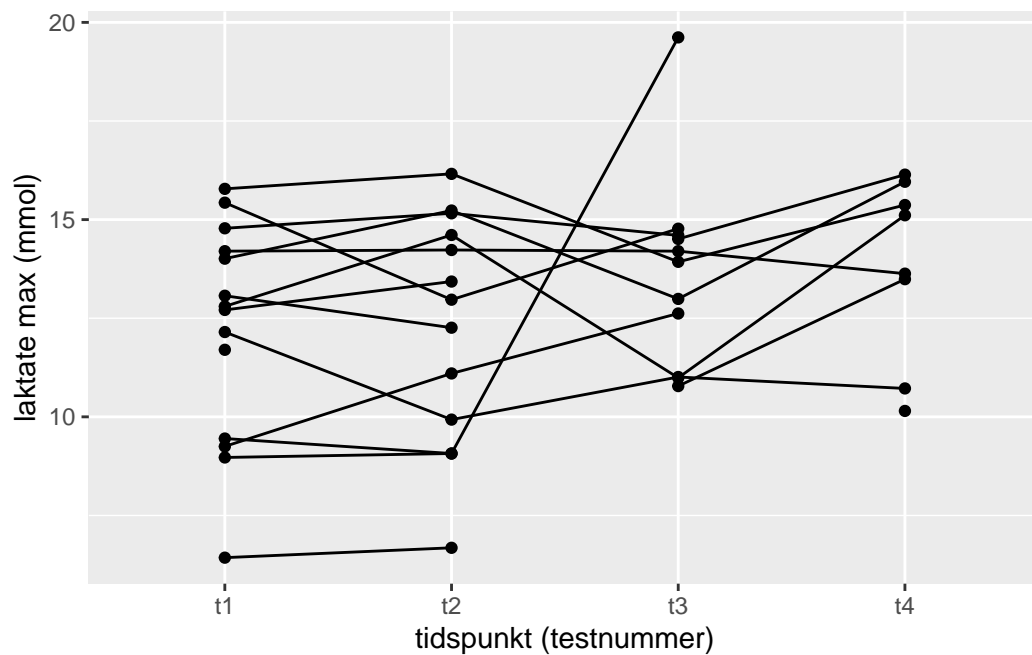
```

# A tibble: 8 x 5
  timepoint sex      m      s      n
  <chr>    <chr> <dbl> <dbl> <int>
1 t1      f      1.14 0.0126      3
2 t1      m      1.14 0.0379     11
3 t2      f      1.16 0.0304      3
4 t2      m      1.14 0.0398     11
5 t3      f      1.17 0.0318      2
6 t3      m      1.18 0.0493      9
7 t4      f      1.19 0.0225      3
8 t4      m      1.18 0.0371      5

```

#### 1.3.1.4 Graf over laktatmålingene

```
ggplot(data = dat,  
  aes(y = la.max,  
    x = timepoint,  
    group = id)) +  
  
  geom_point() +  
  
  geom_line() +  
  
  labs(y = "laktate max (mmol)",  
    x = "tidspunkt (testnummer)")
```



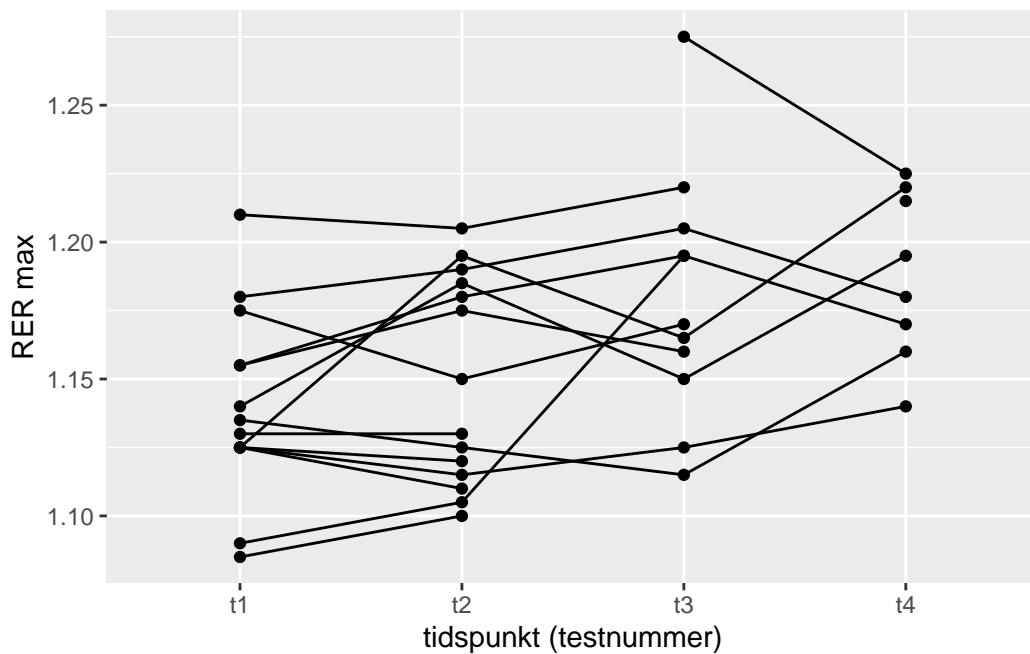
#### 1.3.1.5 Graf over RER-målingene

```
ggplot(data = dat,  
  aes(y = rer.max,  
    x = timepoint,  
    group = id)) +
```

```
geom_point() +

geom_line() +

labs(y = "RER max",
      x = "tidspunkt (testnummer)")
```



## 1.4 Utregning av variasjonskoeffisient

### 1.4.1 Laktat maks

```
cv <- dat %>%
  select(id, timepoint, la.max) %>%
  pivot_wider(names_from = timepoint,
              values_from = la.max) %>%

  mutate(diff = t2 - t1) %>% # Change/difference score
  summarise(m = mean(c(t1, t2), na.rm = TRUE),
            s = sd(diff, na.rm = TRUE), # Summarize to calculate sd, and...
```

```

te = s / sqrt(2),
cv = 100 * (te/m))

cv_percent <- round(cv$cv,1)

```

Variasjonskoeffisienten for laktat maks var 7.6%

## 1.4.2 RER maks

```

cv <- dat %>%
  select(id, timepoint, rer.max) %>%
  pivot_wider(names_from = timepoint,
              values_from = rer.max) %>%

  mutate(diff = t2 - t1) %>% # Change/difference score
  summarise(m = mean(c(t1, t2), na.rm = TRUE),
            s = sd(diff, na.rm = TRUE), # Summarize to calculate sd, and...
            te = s / sqrt(2),
            cv = 100 * (te/m))

cv_percentRER <- round(cv$cv,1)

```

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.

### **Temperatur og luftfuktighet<sup>^</sup>**

Alle testene vi på gruppe 1 utførte var på det samme laboratoriet - Laboratorium 2. Her finnes det flere tiltak som skal sikre at temperaturen og luftfuktigheten er relativ lik hele tiden, som klimaanlegg og luftfukter. Likevel ble det noe variasjon i temperatur og luftfuktighet. Temperaturen varierte fra nn-nn grader, og luftfuktigheten varierte fra nn-nn %. Det skal nevnes at luftfuktigheten var ganske stabil på t1-t3, og at det var på t4 at luftfuktigheten var veldig lav.

### **tidspunkt**

Tidspunkt skulle være standardisert. Vi skulle ha laboratoriet fra kl 09:15-12:00 på alle testene, og vi fulgte samme rekkefølge hver gang. Det oppstod likevel noen problemer. På t3 var laboratoriet dobbeltbooket så det ble kortere tid til å gjennomføre testene, som førte til noe stress. På t4 var laboratoriet også dobbeltbooket så testen ble utsatt til kl 13.00. Også her hadde vi fått mindre tid enn det som først var oppsatt, men klarte å gjennomføre alle testene med god margin.

Vi klarta altså bare å standardisere tidspunkt skikkelig på t1 og t2. Resultatene på t3 og t4 kan derfor bli noe påvirket av dette.

### **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 at vi en eller to ganger var litt unøyaktig med å starte testen i vintus og lode samtidig. Dette gjør at watt-økning for eksempel kunne komme litt før eller litt etter målingen i vintus. 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 exelarket og erstatte alle med “f” og “m” på de korresponderende kjønnene.

## 1.6 Resultat

```
sd_age_t4 <- dat %>%
  filter(timepoint == "t4") %>%
  summarise(sd_age = sd(age, na.rm = TRUE)) %>%
  pull(sd_age)

#for å referere til standard avik til t4 i teksten
```

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.



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 prøvetakingteknikk.

## 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<sup>-1</sup> og 4 mmol L<sup>-1</sup>.

#### 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<sup>-1</sup>. 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)

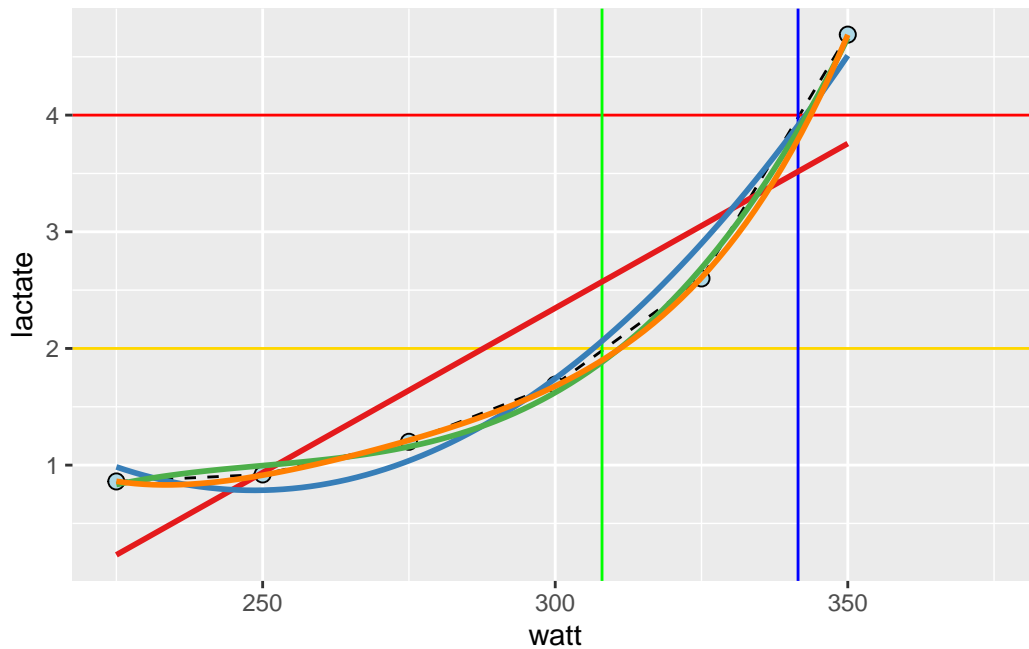
###laste inn data
data("cyclingstudy")

### 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 verdiene 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)

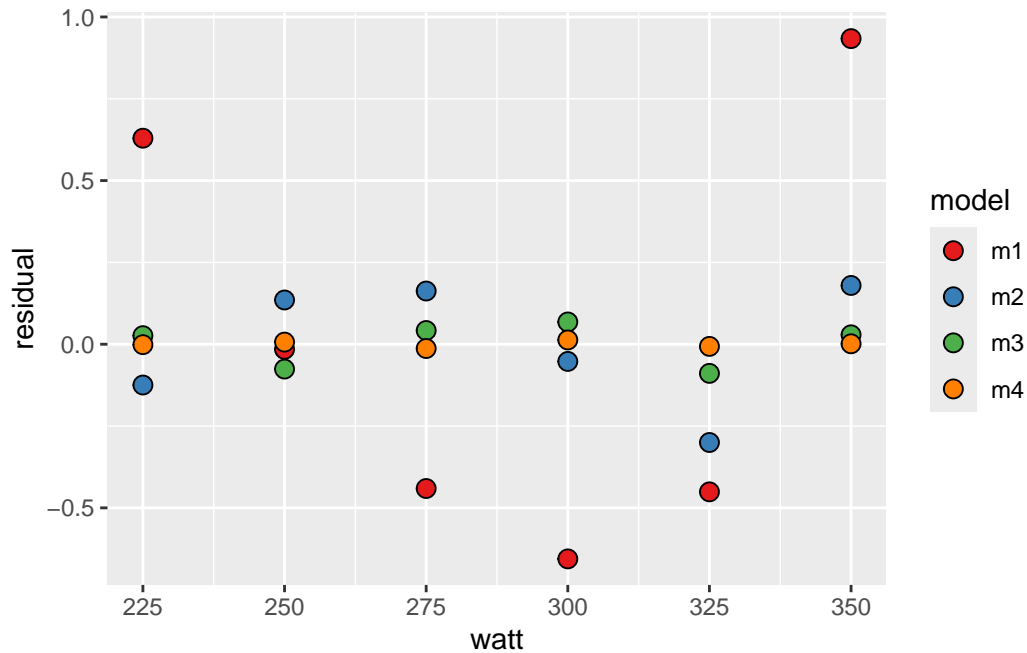
# Legger til en fjerdegradspynomisk modell.
m4 <- lm(lactate ~ poly(watt, 4, raw = TRUE), data = lactate)

# 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