

Mappeeksamen IDR4000

Student Studentson

2024-09-10

Table of contents

Introduksjon	4
1 Assignment 1: Reliability and tools for reproducible data science	5
2 Introduksjon	6
3 Metode	7
4 Resultat	9
4.1 Gjennomsnitt og standardavvik	9
4.2 Reliabilitet	11
4.3 Korrelasjon mellom Vo2maks og Wattmaks per kg	12
5 Referanser	14
6 Assignment 2: Regression models, predicting from data	15
7 Assignment 3: Drawing inference from statistical models, and statistical power	16
8 Assignment 4: Study designs	17
8.1 Overview	17
9 Assignment 5: Analyzing repeated measures experiments	18
9.1 Assignment overview	18
9.2 Introduction	18
9.3 Methods	18
9.3.1 Participants and study overview	18
9.3.2 Muscle strength and hypertrophy	18
9.3.3 Data analysis and statistics	18
9.4 Results	18
9.5 Discussion	22
9.6 Conclusion	22
10 Philosophy of science	23
11 Molecular Laboratory report	24

Introduksjon

Ny rad

Mappeeksamen består av følgende deler:

- Rapport: “Deskriptiv statistikk, reliabilitet og validitet og verktøy for reproduserbar vitenskap”.
- Laborasjonsrapport fra molekylærlabb
- Arbeidskrav i vitenskapsteori
- Rapport: “Statistisk inferens, statistiske modeller og statistisk styrke”
- Rapport: “Studiedesign”
- Rapport: “Analyse av eksperimenter med repeterte målinger”

1 Assignment 1: Reliability and tools for reproducible data science

2 Introduksjon

Det ble gjennomført fire testdager 28.08.2024, 29.08.2024, 9.09.2024 og 11.09.2024 for å teste VO₂maks. Formålet med disse testene var å øve på å kunne gjennomføre fysiologiske tester med høy reliabilitet. Reliabilitet refererer til graden av konsistens eller pålitelighet i målinger evnen til å kunne reproducere (Hopkins,2001), et eksempel på dette er ved fysiologisk testing som repeteres i forskningsprosjekter, der bedre reliabilitet vil indikere hvor god presisjonen er og måling av endring over tid (Hopkins,2001). Det er mange begreper som er relevante for å kunne si noe om reliabilitet, men standardavviket er et av disse. Standardavviket sier noe om hvor langt unna verdiens gjennomsnittlige avstand er fra gjennomsnittet (Spiegelhalter 2020)

Kroppens maksimale oksygenopptak (VO₂maks) sier noe om kroppens maksimale evne til å ta opp og omsette oksygen (Bassett and Howley 2000) . VO₂maks kan beskrives ved hjelp av Ficks likning: $VO_2\text{maks} = \dot{V}O_2\text{maks} \times a-vO_2\text{differansemaks}$. VO₂maks måles ved at man måler hvor mye oksygen kroppen klarer å omsette pr minutt (Bassett and Howley 2000). Det finnes ulike måter og fremstille VO₂maks på de to av disse er absolutt VO₂maks beskrevet som (ml/min) eller relative tall relatert til kroppsvekt (ml/kg/min).

Vi har i resultat delen valgt å fremstille effekt maks (Wmaks) som er et mål på snitteffekt det siste minuttet av VO₂maks testen basert på kroppsvekt. Wmaks/kg sett opp i sammenheng med den relative VO₂maks (ml/kg/min). Forskning viser at at høy VO maks sammen med god mekanisk effektivitet og høy laktatterskel gir bedre utholdenhetsprestasjoner, noe som reflekteres i høyere Wmaks/kg (Joyner and Coyle 2008).

3 Metode

VO2maks testen ble gjennomført på en ergometersykkel med bukkestyre (Lode Excalibur Sport; Lode B.V., Groningen, Nederland). Kranken kalibreres på Lode sykkel før hver teststart, og sykkel stilles inn etter utøver sitt ønske ved første test og stilles inn til den samme sittestillingen påfølgende tester. For å måle det maksimale oksygenopptaket ble det brukt Vyntus (Jaeger Vyntus CPX, Hoechberg, Tyskland). Gassanalysator kalibreres til $< 2,0\%$ differanse og luftvolum kalibreres til $< 0,2\%$ differanse. Zeroing gjøres også alltid før test starter. Syklistene veies med de klærne de skal sykle med, og 0,3kg trekkes fra. For å kunne sikre god relabilitet ble det tydeliggjort at man skulle replisere det siste måltidet før test, ha det samme koffeininntaket, avstå fra alkohol og tobakk de siste 72 timene før test og prøve å få tilnærmet lik søvn, samt trene det samme dagen før test. Da dette er faktorer som kan spille inn på prestasjon og metabolismen (Tanner & Gore, 2013) og dermed påvirke relabiliteten. Hvorvidt dette er fulgt er noe usikkert da dette ikke er fulgt opp videre annet enn at det ble informert om før første testdag.

VO2maks testen gjennomføres etter en 5min standardisert oppvarming på ergometersykkelen. Oppvarmingen starter med en 2 min oppvarming på 11-12 i Borg, deretter 2min på 15 i Borg før 1min på 11-12 BORG. Testen starter på en belastning (Watt) basert på deltagerens nivå i samråd med utøver og testleder. Det viktigste er at påfølgende VO2maks tester starter på samme watt. Wattbelastningen økte med 20W eller 25W hvert minutt frem til utøveren når maksimal utmattelse er oppnådd. Maksimal utmattelse ble i denne sammenheng ikke evne til å kunne opprettholde RPM 60. Under VO2maks var RPM valgfritt. Testleder gjør verbal oppmuntring og sekundering underveis i testen. For at verbal oppmuntring og instruksjon ved test skulle være lik etterstrebt vi å ha samme testleder til samme test person (Halperin, Pyne, and Martin 2015). Det blir målt oksygenmålinger hvert 30 sek, og de to høyeste påfølgende målingene blir definert som VO2maks. Umiddelbart etter test oppgir utøveren opplevd anstrengelse på BORG skala. Maks hjertefrekvens blir lest av fra utøverens egen pulsklokke. Blodprøve ble tatt fra utøverens fingertupp 1 min etter endt test for å måle [BLa-]. [BLa-] blir videre målt videre målt ved hjelp av en Biosen C-line (Biosen C-line Lactate Analyzer, EKF Diagnostic GmbH, Barleben, Germany) .Etter endt test ble det hentet ut data som videre ble plottet inn i Excel og videre ført statistikk på ved hjelp av Rstudio.

```
library(readxl)
library(tidyverse)
library(ggplot2)
library(dplyr)
```

```
library(magrittr)
library(gt)
df <- read_xlsx("data/o2-test.xlsx")
```


4 Resultat

4.1 Gjennomsnitt og standardavvik

```
#Last inn datasettet
df.tbl <- df |>
#Velg hvilke variabler som skal med videre
  select(id, timepoint, weight, w.max:borg.max) |>
#Fjern verdiene som har characters
  select(-hr.max) |>
  select(-la.max) |>
  select(-bf.max) |>
  select(-vco2.max) |>
  select(-rer.max) |>
  select(-ve.max) |>
#Lag nye verdier, relativ vo2maks og wattmaks/kg
  mutate(rel.vo2max = vo2.max / weight) |>
  mutate(rel.wmax = w.max / weight) |>
#Sliter fortsatt med å forklare pivot både longer og wider
  pivot_longer(names_to = "ID",
               values_to = "verdier",
               cols = w.max:rel.wmax) |>
#Grupper etter id og names
  group_by(id, ID) |>
#Summarise gjennomsnitt og standardavvik til verdiene
  summarise(m = mean(verdier),
            s = sd(verdier)) |>
  mutate(m_s = paste0(signif(m, 3), " (", signif(s, 2), ")")) %>%
  ungroup() |>
select(-m, -s) |>
  pivot_wider(names_from = "id",
             values_from = "m_s") |>
gt()
```

```

df.tbl |>
  text_replace(
    pattern = "^bf.max$",
    replacement = md("Bf<small>(maks)</small>")) |>

  text_replace(
    pattern = "^borg.max$",
    replacement = md("Borg<small>(maks)</small>")) |>

  text_replace(
    pattern = "^rel.vo2max$",
    replacement = md("VO<sub>2maks</sub><br><small>(ml/kg/min)</br></small>")) |>

  text_replace(
    pattern = "^rel.wmax$",
    replacement = md("Wattmaks/kg")) |>

  text_replace(
    pattern = "^vo2.max$",
    replacement = md("VO<sub>2maks</sub><br><small>(ml/min)</br></small>")) |>

  text_replace(
    pattern = "w.max$",
    replacement = ("Wattmaks")) |>
  tab_options(table.font.size = 10)

```

Table 4.1: Tabellen viser hver deltaker

ID	1	2	3	4	5	
Borg<small>(maks)</small>	19.2 (0.96)	19 (0.82)	18 (1.2)	19 (0)	19.5 (0.71)	19
VO_{2maks} <small>(ml/kg/min)</br></small>	33.5 (1.5)	43.7 (2.6)	51.6 (4.1)	37.1 (1.1)	58.9 (0.64)	45.5 (0)
Wattmaks/kg	2.5 (0.14)	3.58 (0.044)	3.6 (0.46)	3 (0.2)	5.18 (0.082)	3.51 (0)
VO_{2maks} <small>(ml/min)</br></small>	3240 (150)	2700 (160)	4130 (300)	2860 (52)	4390 (48)	3710 (0)
Wattmaks	243 (13)	221 (2.8)	288 (36)	231 (13)	387 (6.1)	286 (0)

```

df.rel <- df |>
  select(id, timepoint, weight, vo2.max, w.max) |>
  mutate(rel.vo2 = vo2.max / weight) |>
  mutate(rel.wmax = w.max / weight)

```

4.2 Reliabilitet

```
df.rel |>

#velger hvilke variabler vi er interessert i å se på

select(id, timepoint, rel.vo2) |>

#benytter oss av pivot_wider for å lage tidy_data

pivot_wider(names_from = timepoint,
             values_from = rel.vo2) |>
mutate(diff = t2 - t1) |>
summarise(MEAN = mean(c(t1, t2), na.rm = TRUE),
          SD = sd(diff, na.rm = TRUE),
          TE = SD / sqrt(2),
          CV = TE/MEAN * 100) |>
gt() |>
fmt_number(decimals = 2)
```

Table 4.2: Reliability relativ vo2maks T1&T2

MEAN	SD	TE	CV
52.44	1.83	1.30	2.47

```
cv_perct1_t2 <- 2.47
```

Reliabiliteten mellom t1 og t2 er 2.47%.

```
df.rel |>

#velger hvilke variabler vi er interessert i å se på

select(id, timepoint, rel.vo2) |>
```

```
#benytter oss av pivot_wider for å lage tidy_data

pivot_wider(names_from = timepoint,
             values_from = rel.vo2) |>
mutate(diff = t4 - t3) |>
summarise(MEAN = mean(c(t3, t4), na.rm = TRUE),
          SD = sd(diff, na.rm = TRUE),
          TE = SD / sqrt(2),
          CV = TE/MEAN * 100) |>
gt() |>
fmt_number(decimals = 2)
```

Table 4.3: Reliability relativ vo2maks T3&T4

MEAN	SD	TE	CV
48.64	3.29	2.32	4.78

```
cv_perct3_t4 <- 4.78
```

Reliabiliteten mellom t3 og t4 er 4.78%.

4.3 Korrelasjon mellom Vo2maks og Wattmaks per kg

```
df %>%
  mutate(rel.vo2max = vo2.max / weight) %>%
  mutate(w.max.kg = w.max / weight) %>%
  ggplot(aes(x = rel.vo2max,
            y = w.max.kg,
            color = id)) +
  geom_point() +
  labs(x = ("Relativ VO2maks"),
       y = ("Wattmaks/kg"),
       title = "Korrelasjon mellom Vo2maks og Wattmaks",
       subtitle = "Verdier er delt på vekt")
```

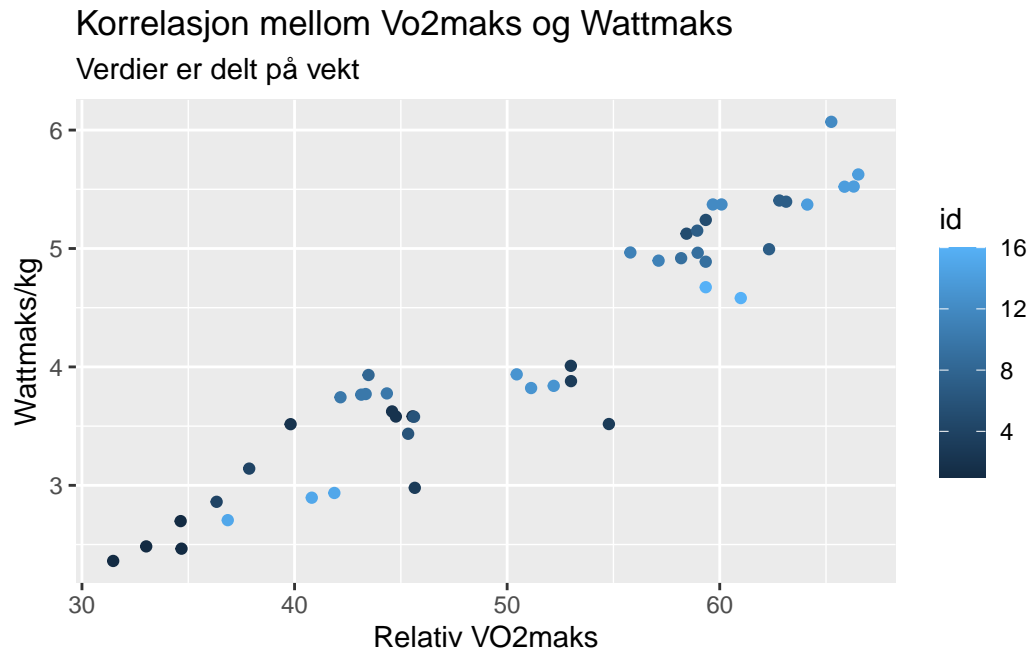


Figure 4.1: Figur 1: Hvert punkt = én observasjon

Jeg fikk dessverre ikke til å gjøre om på hver enkelt farge, for å lettere kunne skille ulike id'er fra hverandre.

5 Referanser

6 Assignment 2: Regression models, predicting from data

The assignment has three parts:

- Part 1: Lactate thresholds
- Part 2: Predicting sizes of DNA fragments
- Part 3: Interpreting a regression table

7 Assignment 3: Drawing inference from statistical models, and statistical power

This assignment is set up as a statistical laboratory, we will perform simulations and your assignment is to interpret and explain the results. Create a report based on the code used in the lab and make sure you answer the specified questions (1-8). You can be as creative as you want and explore the results further.

8 Assignment 4: Study designs

8.1 Overview

Choose an area of interest (e.g. protein supplementation for muscle hypertrophy or the effect of block periodization on VO2max). Find at least five *original research studies*¹ in your selected area and describe strength and weakness of these studies. The report should focus on the design of the studies and selection of statistical tests to answer study aims. Conclude your report with a recommendation, how should future studies in your area be designed to best answer similar questions?

¹Avoid using review articles or meta-analyses

9 Assignment 5: Analyzing repeated measures experiments

9.1 Assignment overview

In this assignment you will analyse and report on trial investigating the effect of resistance training volume on lean mass and muscle strength. The data are part of the `exscidata` package and can be accessed as `data("strengthvolume")` and `data("dxadata")`. Read the [instructions carefully!](#)

Below you will find a basic outline of the report and example code that we worked on in class.

9.2 Introduction

9.3 Methods

9.3.1 Participants and study overview

9.3.2 Muscle strength and hypertrophy

9.3.3 Data analysis and statistics

9.4 Results

The average difference in lean mass changes between sets were 122.8, 95% CI: [8.6, 237], $p = 0.036$.

```
## Time points in strength data set

strengthvolume %>%
  distinct(exercise)
```

```
# A tibble: 6 x 1
  exercise
  <chr>
1 legpress
2 legext
3 isok.60
4 isok.120
5 isok.240
6 isom
```

```
## Exploratory plot of strength data
```

```
str <- strengthvolume %>%
  filter(include == "incl") %>%
  mutate(time = factor(time, levels = c("pre", "session1",
                                         "week2", "week5",
                                         "week9", "post"))) %>%
  print()
```

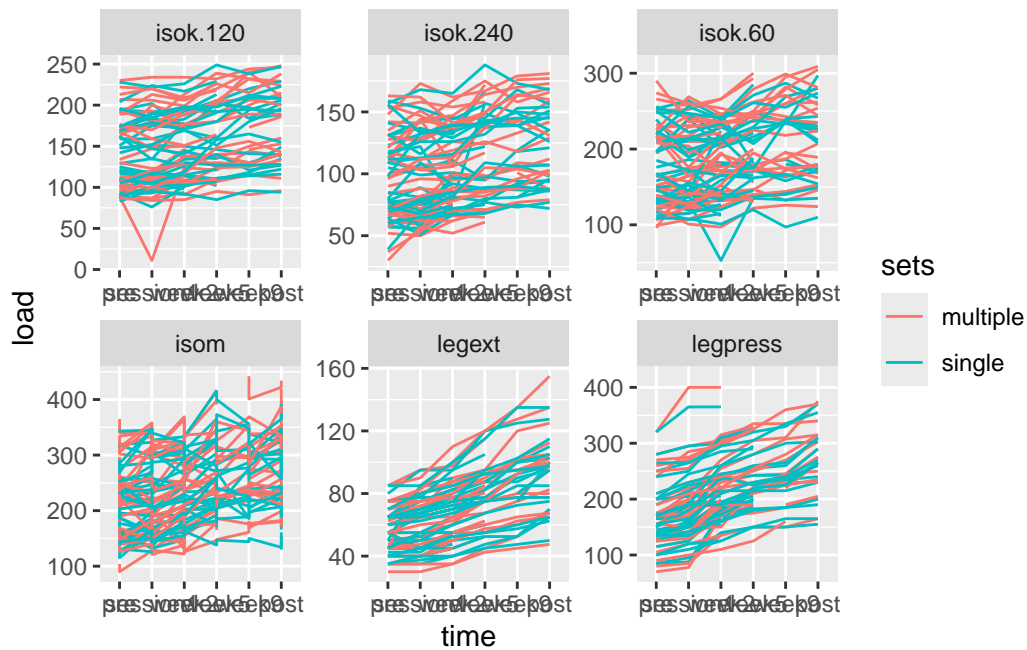
```
# A tibble: 2,856 x 8
```

	participant	sex	include	time	sets	leg	exercise	load
	<chr>	<chr>	<chr>	<fct>	<chr>	<chr>	<chr>	<dbl>
1	FP13	male	incl	pre	single	R	legpress	115
2	FP13	male	incl	pre	multiple	L	legpress	115
3	FP13	male	incl	pre	single	R	legext	55
4	FP13	male	incl	pre	multiple	L	legext	55
5	FP13	male	incl	session1	single	R	legpress	125
6	FP13	male	incl	session1	multiple	L	legpress	125
7	FP13	male	incl	session1	single	R	legext	55
8	FP13	male	incl	session1	multiple	L	legext	55
9	FP13	male	incl	week2	single	R	legpress	185
10	FP13	male	incl	week2	multiple	L	legpress	175

```
# i 2,846 more rows
```

```
str %>%
  ggplot(aes(time,
              load,
              group = paste(participant, sets),
              color = sets)) +
  geom_line() +
  facet_wrap(~ exercise, scales = "free")
```

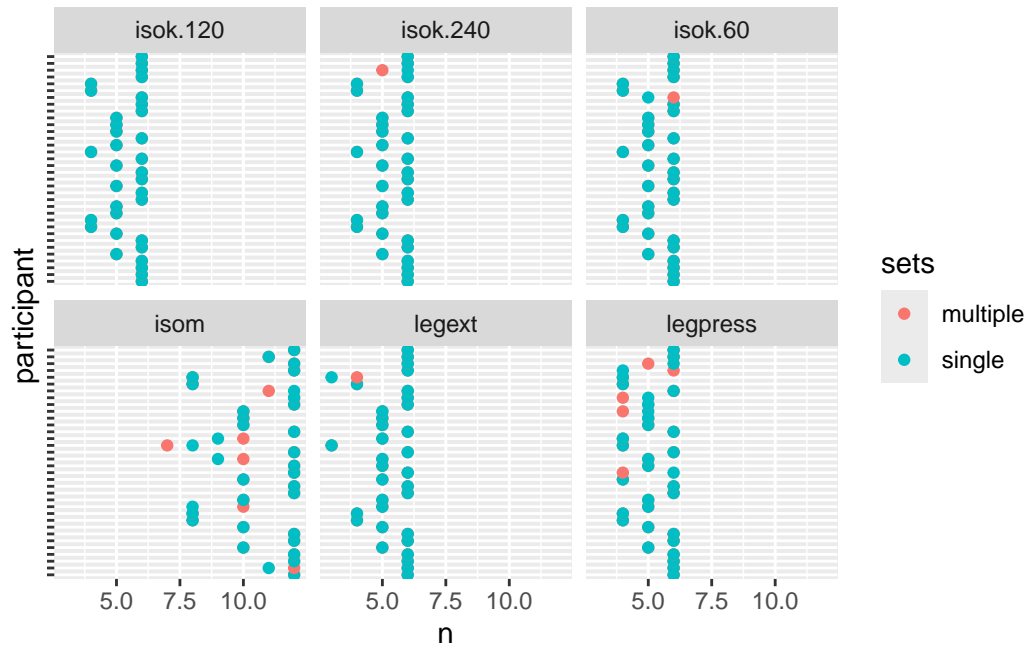
Warning: Removed 5 rows containing missing values or values outside the scale range (``geom_line()``).



```
## How many measurements per participant
```

```
str %>%
  filter(!is.na(load)) %>%
  group_by(participant, exercise, sets) %>%
  summarise(n = n() ) %>%
  ggplot(aes(n, participant, color = sets)) +
  geom_point() +
  facet_wrap(~ exercise) +
  theme(axis.text.y = element_blank())
```

``summarise()`` has grouped output by 'participant', 'exercise'. You can override using the ``.groups`` argument.



```
## Use pre and post data
# Combine pre data prior to data analysis
# per exercise, leg, participant, and sets

str %>%
  mutate(time = if_else(time %in% c("pre", "session1"), "pre", time)) %>%

  filter(time %in% c("pre", "post")) %>%

  summarise(load = max(load, na.rm = TRUE),
            .by = c(participant,
                    sex,
                    time,
                    sets,
                    exercise,
                    leg)) %>%

  print()
```

Warning: There were 7 warnings in `summarise()`.
 The first warning was:
 i In argument: `load = max(load, na.rm = TRUE)`.

```
i In group 62: `participant = "FP6"`, `sex = "female"`, `time = "post"`, `sets
  = "multiple"`, `exercise = "legpress"`, `leg = "L"`.
Caused by warning in `max()``:
! no non-missing arguments to max; returning -Inf
i Run `dplyr::last_dplyr_warnings()` to see the 6 remaining warnings.
```

```
# A tibble: 816 x 7
  participant sex    time sets    exercise leg    load
  <chr>      <chr> <chr> <chr>    <chr>    <chr> <dbl>
1 FP13      male   pre   single  legpress R     125
2 FP13      male   pre   multiple legpress L     125
3 FP13      male   pre   single  legext   R      55
4 FP13      male   pre   multiple legext   L      55
5 FP13      male   post  single  legpress R     230
6 FP13      male   post  multiple legpress L     235
7 FP13      male   post  single  legext   R     97.5
8 FP13      male   post  multiple legext   L     100
9 FP16      female  pre   single  legpress R      95
10 FP16     female  pre   multiple legpress L      85
# i 806 more rows
```

9.5 Discussion

9.6 Conclusion

10 Philosophy of science

See instructions on canvas.

11 Molecular Laboratory report

Select one laboratory assignment and write a detailed report.

References

- Bassett, D R, Jr, and E T Howley. 2000. "Limiting Factors for Maximum Oxygen Uptake and Determinants of Endurance Performance." *Med. Sci. Sports Exerc.* 32 (1): 70–84.
- Halperin, Israel, David B Pyne, and David T Martin. 2015. "Threats to Internal Validity in Exercise Science: A Review of Overlooked Confounding Variables." *Int. J. Sports Physiol. Perform.* 10 (7): 823–29.
- Joyner, Michael J, and Edward F Coyle. 2008. "Endurance Exercise Performance: The Physiology of Champions." *J. Physiol.* 586 (1): 35–44.
- Spiegelhalter, David. 2020. "Introducing the Art of Statistics: How to Learn from Data." *Numeracy* 13 (1).