

Eksamen IDR4000 Ole

Ole-K.

2024-11-05

Table of contents

Forord	4
1 Reliabilitet	5
1.1 Introduksjon	5
1.2 Metode	5
1.2.1 Gjennomføring av $\dot{V}O_{2max}$ test	5
1.2.2 Standardisering	6
1.2.3 Intern validering	6
1.3 Resultater	7
1.4 Diskusjon	8
1.5 Referanser	8
2 Regresjonsmodeller	9
2.1 Introduksjon Regresjonsmodeller	9
2.2 Laktatterskel	9
2.2.1 Introduksjon	9
2.2.2 Metode	10
2.2.3 Resultat	10
2.2.4 Diskusjon	11
2.3 Estimering av størrelse på DNA-fragmenter	11
2.3.1 Introduksjon	11
2.3.2 Metode	11
2.3.3 Resultat	12
2.3.4 Diskusjon	13
2.4 Tolke en regresjonstabell	13
2.4.1 Introduksjon	13
2.4.2 Metode	13
2.4.3 Resultat	14
2.4.4 Diskusjon	14
2.5 Referanser	15
3 Statistisk inferens	16
3.1 Introduksjon	16
3.2 Forklaring av T -verdi, P -verdi, Standard Error (SE) og Estimat	16
3.3 Hva bidrar til de ulike resultatene i de to studiene (m1 og m2)	19

3.4	Hvorfor benytter en begge “halene” i t-fordelingen (two-tailed test)	19
3.5	Kalkuler standardavvik for estimat-variabelen og gjennomsnittet av se- variabelen for begge utvalgene.	19
3.6	Lag et histogram av p -verdiene for de to uvalgene. Hvordan kan histogrammene tolkes og hva forteller de om effekten av størrelsen av utvalget for statistisk styrke?	20
3.7	Antall observasjoner fra hvert utvalg som er under signifikansnivået på 0.05 . .	21
3.8	Estimerer effekten ved one-sample t-test	21
3.9	Ved å benytte det nye datasettet: hvor mange tilfeller av “falsk positive”-tester vil vi komme over ved å repetere studien mange ganger?	22
3.10	Referanser	22
4	Studiedesign	23
4.1	Introduksjon	23
4.2	Metode	23
4.3	Resultater	24
4.4	Diskusjon	24
4.5	Konklusjon	25
4.6	Referanser	26
5	Repeterte målinger	27
5.1	Introduksjon	27
5.2	Metode	27
5.2.1	Utvalg og studieoversikt	27
5.2.2	Spesifikke beskrivelser av testmetode	28
5.2.3	Styrketreningsprotokoll	28
5.2.4	Muskulær Styrkevurdering	29
5.2.5	Måling av Kroppssammensetning og Mager Muskelmasse	29
5.2.6	Dataanalyse og Statistikk	29
5.3	Resultater	29
5.3.1	Treningsvolums effekt på mager muskelmasse i underekstremitetene . .	29
5.3.2	Treningsvolums effekt på kraftutvikling i underekstremitetene	30
5.4	Diskusjon	30
5.5	Konklusjon	32
6	Vitenskapsfilosofi	33
7	RNA ekstraksjon og analyse	34
	Referanser	35

Forord

1 Reliabilitet

1.1 Introduksjon

Reliabilitet er et viktig begrep innen treningsfysiologi og medisin, da det omhandler påliteligheten av målinger og tester. I forskning og klinisk praksis er det avgjørende at instrumenter og prosedyrer gir stabile resultater over tid, under samme forhold og for ulike individer. Hopkins (2000) definerer reliabilitet som graden av overensstemmelse mellom gjentatte målinger, og understreker at høy reliabilitet er nødvendig for å trekke pålitelige konklusjoner om effekten av trening eller behandling (Hopkins, 2000). Innen treningsfysiologi kan lav reliabilitet i målingene eksempelvis føre til unøyaktige vurderinger av treningsprogresjon, noe som kan resultere i at treningsprogrammet ikke er tilpasset det faktiske ferdighetsnivået man innehar på gitt tidspunkt. Dette kan føre til manglende fremgang, ettersom treningen ikke er optimalt justert for den enkeltes kapasitet.

1.2 Metode

1.2.1 Gjennomføring av $\dot{V}O_{2max}$ test

Testen som skal gjennomføres er en $\dot{V}O_{2max}$ -test, der testdeltakeren starter på en forhåndsbestemt wattverdi (motstand). Denne wattverdien økes gradvis hvert minutt, avhengig av den valgte protokollen, og testdeltakeren sykler til en ikke lenger klarer å fortsette. Testen avsluttes dersom deltakerens tråkkfrekvens synker under 60 rpm (rotasjoner per minutt). Under testen skal deltaker forbli sittende på sykkelen.

Testdeltakeren vil få informasjon om hvor lenge de har syklet underveis i testen. Derimot vil de ikke ha tilgang til annen informasjon som puls, $\dot{V}O_2$ eller lignende data underveis. Testlederen vil informere om hvor lang tid det er igjen til neste wattøkning og hvilken watt testdeltakeren sykler på for øyeblikket.

Testlederen vil mot slutten av testen bidra til å pushe testdeltakeren for å få ut det siste av deres kapasitet. Hvordan denne “pushingen” gjennomføres, kan variere mellom testledere, men det er viktig at tilnærmingen til hver deltaker er så konsekvent som mulig. Borg-skalaen (6-20) vil benyttes for å vurdere utmattelsen hos testdeltakeren. Ved testslutt vil deltaker bli bedt om å gi et tall på skalaen, der 6 tilsvarer “ingen anstrengelse” og 20 “maksimal anstrengelse”.

Ett minutt etter testslutt tas en laktatprøve fra testdeltakeren for å estimere maksimal laktatverdi (La_{max}). Etter at prøven er tatt, noterer testlederen ned hvor lenge testdeltakeren har syklet, og henter deretter ut relevant data fra testen via en rapport fra Vyntus-systemet.

1.2.2 Standardisering

Testdeltakeren skal spise sitt siste store måltid senest 2 timer før teststart. Det er tillatt å innta en mindre karbohydratkilde (eks: banan, energibar, gel) senest 30 minutter før teststart. Koffeininntak skal gjøres som normalt og testene gjennomføres på samme tidspunkt på dagen så godt det lar seg gjøre. Deltakeren anbefales å forsøke å opprettholde sin vanlige døgnrytme gjennom testperioden. For å forberede kroppen på testen, er det viktig at deltakerne ikke gjennomfører hard trening på underekstremitetene dagen før testen.

Oppvarmingen før testen ble satt til 5 minutter, hvor deltaker sykler på økende intensitet i intervaller på 2-2-1 minutt. Eksempelvis 2 minutter på 150W, 2 minutter på 175W og 1 minutt på 220W.

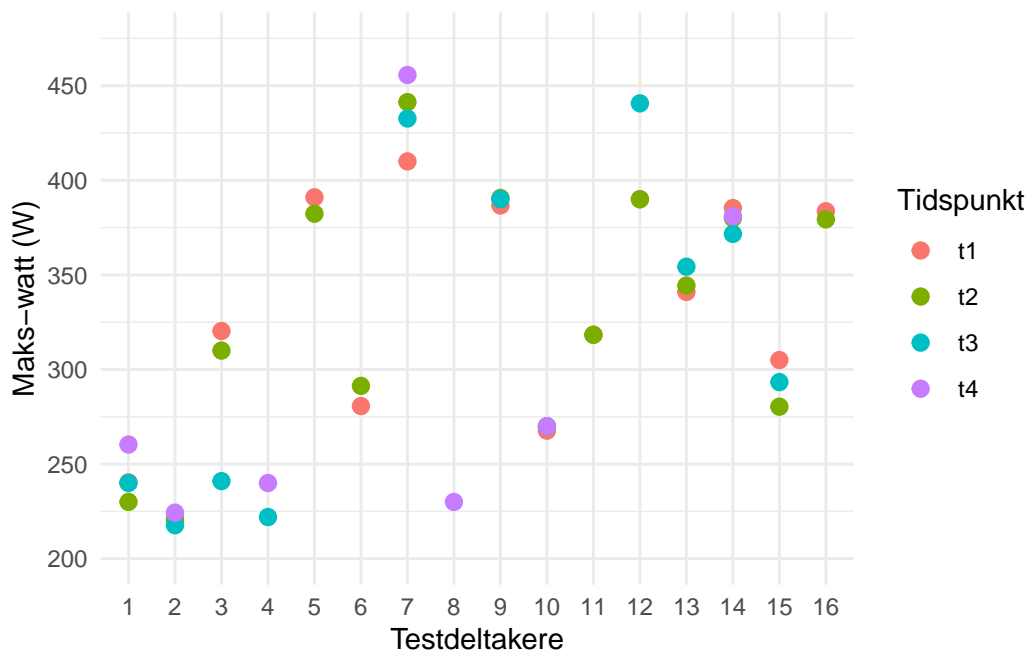
Testprotokollen er den samme for alle tester. Et eksempel på testprotokolle var start på 200W med en økning på 20W hvert minutt til utmattelse. Individuelle tilpasning mtp belastning ble gjort for hver deltaker, både ved oppvarming og testprotokoll.

1.2.3 Intern validering

Både tekniske og biologiske faktorer kan potensielt påvirke resultatene og true den interne validiteten av testen (Halperin *et al.*, 2015). Ved å beskrive og etablere testprosedyre samt standardisere testen, ønsker vi å skape rammer som sikrer at vi kontrollerer for mulige konfunderende faktorer som kan påvirke testprestasjon. Dette vil bidra til å redusere variasjoner og sikre at resultatene reflekterer faktiske prestasjoner, snarere enn påvirkninger fra ytre faktorer.

Selv om vi har jobbet for å sikre god intern validitet, er det faktorer som vil kunne påvirke resultatene grunnet individuelle forskjeller i måten å opptre som testleder, bruk av musikk, samt humør og mental status til testdeltager (Halperin *et al.*, 2015).

1.3 Resultater



Figur 1.1: Watten testdeltaker holdt de siste 30 sek av testen ved de ulike tidspunktene

Tabell 1.1: Gj.snitt og st.avvik for watt-maks

Tidspunkt	Watt-maks
t1	332 (62)
t2	331 (66.6)
t3	316 (84.7)
t4	294 (88.8)

Tabell 1.2: Summerende statistikk

Gj.snitt	St.avvik	Typisk feil	Variasjonskoeffisient
331.07	12.59	8.90	2.69

Figur 1.2

Den gjennomsnittlige watt-verdien som testdeltakerne holdt de siste 30 sekundene av $\dot{V}O_{2max}$ -

testene var 331.1 (± 12.6) og den typiske feilen ble målt til 8.9. Variasjonskoeffisienten ble beregnet til 2.69 % (se Figure 1.2).

1.4 Diskusjon

Den gjennomsnittlige watt-verdien som testdeltakerne holdt de siste 30 sekundene av $\dot{V}O_{2max}$ -testen var 331.1 (± 12.6) og den typiske feilen ble målt til 8.9. Ifølge Hopkins (2000) er den typisk feilen, den gjennomsnittlige variasjonen ved en test. Den typiske feilen gir en indikasjon på den forventede variasjonen ved en retest (Hopkins, 2000). I dette tilfellet kan en forvente en variasjon på 2.69 % ved en eventuell re-test.

1.5 Referanser

2 Regresjonsmodeller

2.1 Introduksjon Regresjonsmodeller

En regresjonsmodell er en modell som kvantifiserer forholdet mellom en eller flere uavhengige variabler og en avhengig variabel. Innen medisin er regresjon den analysemetoden som er hyppigst anvendt. Det finnes forskjellige regresjonsmodeller. De vanligste er lineær regresjon, polynominal regresjon og logistisk regresjon. Hva man har av datasett vil bestemme hvilken regresjonsmodell som egner seg best å benytte (Pisică *et al.*, 2022).

En lineær regresjonsmodell er en modell der en kan estimere verdien av en avhengig variabel basert på verdien av andre kjente uavhengige variabler (Pisică *et al.*, 2022). I en slik modell benyttes en rett linje for å lage en modell som beskriver dataen. Følgende funksjon benyttes for å skape det lineære plottet:

$$y_i = b_0 + b_1x_i + e_i$$

der y_i er den avhengige variabelen som kan estimeres ved å benytte de uavhengige variablene b_1x_i og b_0 . b_0 er skjæringspunktet til grafen og b_1 er stigningstallet til grafen.

I denne rapporten vil ulike tema bli presentert, henholdsvis laktatterskel, estimering av størrelse på DNA fragment og tolkning av regresjonsmodeller. De ulike temaene presenteres med egne delkapitler. Sammenhengen mellom oppgavene er at de omhandler bruker av regresjon og hvordan vi kan tolke denne typen dataanalyse.

2.2 Laktatterskel

2.2.1 Introduksjon

Laktatterskel refererer ofte til den grensen der kroppen produserer mer laktat enn den evner å omdanne. Laktatterskel brukes ofte som et mål på intensitet, samtidig sier det noe om treningsstatus og evne til å prestere. (Antonutto & Di Prampero, 1995)

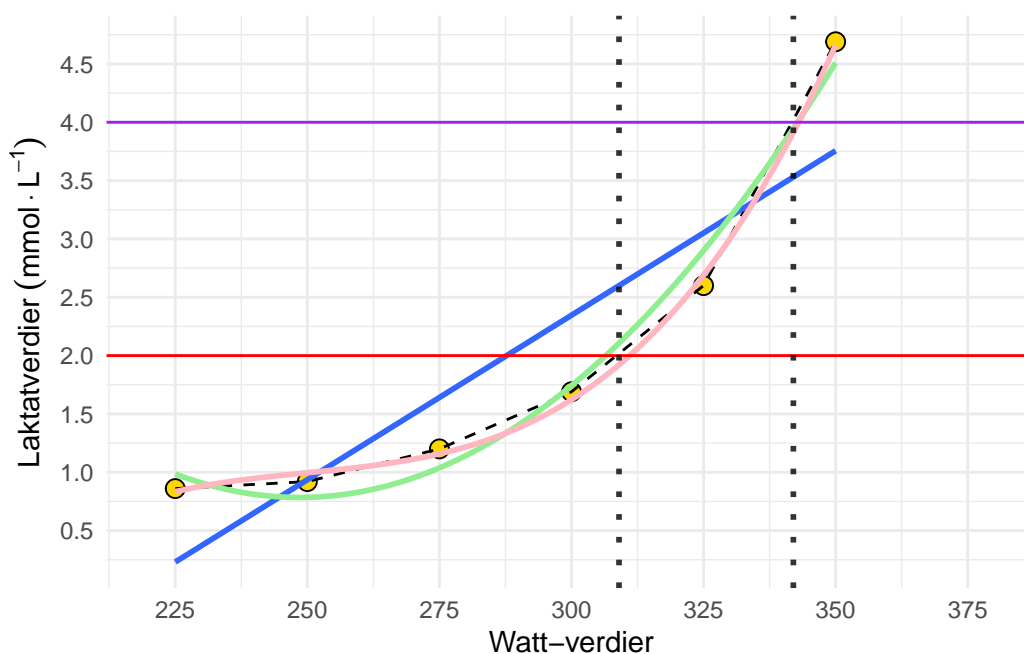
I denne delen undersøkes forholdet mellom laktatnivå og watt, samt tilpasningen av regresjonsmodeller for å beskrive dette forholdet.

2.2.2 Metode

I denne analysen ble det hentet data fra *cyclingstudy*-datasettet som en finner i pakken *exscidata* i R. Fra settet ble deltaker 10 benyttet som eksempel. Datasettet ble omstrukturert til et mer passende format (*tidy data*) for deretter å lage et hensiktsmessig plot. Det ble utført ulike regresjonsmodeller for å fremstille forholdet mellom laktat og watt (se Figure 2.1). Laktatverdiene og wattverdiene ble plottet som gule punkter i grafen. Deretter ble en stiplet svart linje tegnet for å forbinde punktene. En ønsket å se på watt-verdier ved hhv. 2 og 4 mmol/L laktat. Derfor ble de lagt til hjelpende horisontale (rød og lilla) og vertikale (stiplede svarte) linjer for å synliggjøre skjæringspunkter på grafen.

For å analysere sammenhengen mellom laktat og watt ble det utført regresjonsanalyser ved hjelp av tre forskjellige modeller: først en lineær modell (blå linje), deretter en andregradsmodell (grønn linje) og til slutt en tredjegradsmodell (rosa linje). Formålet med analysen var å vurdere hvilken av de tre modellene som best predikerer forholdet mellom laktat og watt.

2.2.3 Resultat



Figur 2.1: Relasjon mellom laktatverdier og watt. Gule punkter = laktat/watt, blå linje = lineær regresjon, grønn linje = andregradsligning, rosa = tredjegradsligning, rød linje = 2 mmol/L, lilla linje = 4 mmol/L, stiplet svarte linjer = skjæringspunkter mot rød og lilla linje

Den lineære regresjonsmodellen estimerer watt-verdier for henholdsvis 2 og 4 mmol/L intensitet til å være 288 W og 359 W. Andregrads- og tredjegradsmodellen kan man se av Figure 2.1 er tilnærmet lik den stiplede svarte linjen som knytte de gule punktene sammen.

2.2.4 Diskusjon

Av Figure 2.1 kan man se at en enkel lineær regresjonsmodell ikke er tilstrekkelig for å estimere sannsynlige verdier for watt. Derimot vil en andre- og tredjegradsmodell passe betydelig bedre opp mot relasjonen watt og laktat. I dette tilfellet predikerer den lineære modellen en watt på 288 ved 2 mmol \times L⁻¹. Den faktiske watt-verdien ligger dog nærmere 310 W når man ser på krysningspunktet mellom den røde horisontale linjen og den stiplede svarte vertikale linjen.

2.3 Estimering av størrelse på DNA-fragmenter

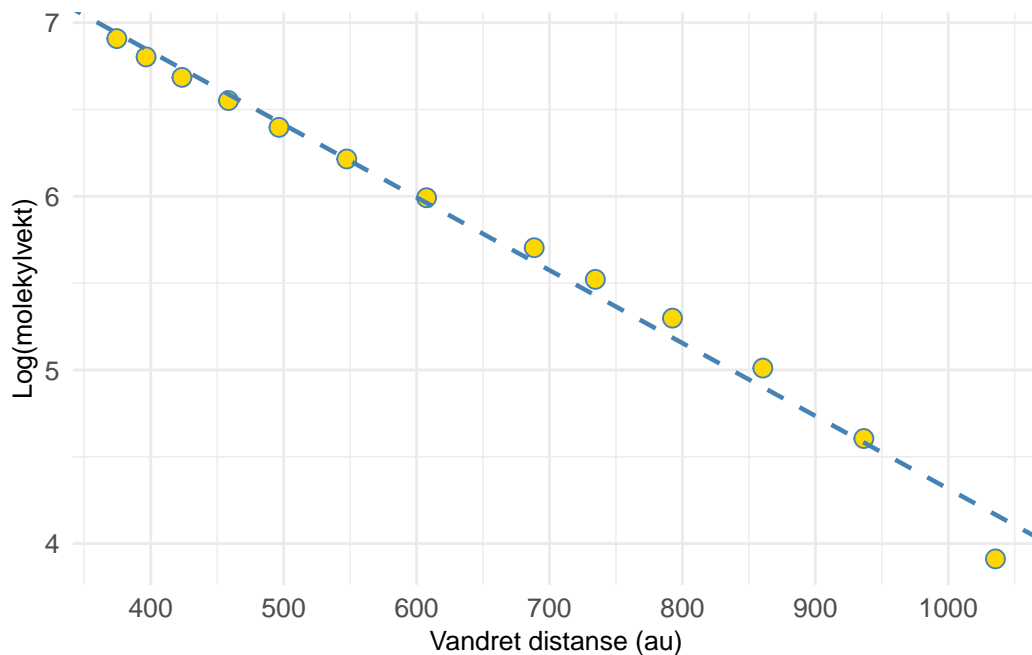
2.3.1 Introduksjon

Beregning av størrelse på DNA-fragmenter er en viktig prosess i lab for å kunne identifisere og analysere DNA-prøver. Gelelektroforese er en av metodene som en hyppigst anvendt for dette formålet. DNA-prøvene tilsettes da en agrose gel og separeres etter størrelse ved hjelp av elektrisitet. Store fragment beveger seg langsommere enn små fragment og en vil få en spredning på gelen. Ved å benytte en kjent referansemarkører, kan man estimere størrelsen på DNA-fragmentene. (Schadock *et al.*, 2015)

2.3.2 Metode

Først ble det tatt helblod fra en forsøksperson, og DNA ble ekstrahert ved hjelp av en standard protokoll for blod-DNA-isolering. Ekstraheringen involverte flere trinn, blant dem tilsetning av bufferløsninger og enzymatiske behandlinger for å bryte ned cellemembraner og frigjøre DNA-et. Det ble videre kjørt en elektroforese for å kunne separere DNA fragmentene og bestemme lengden på DNA-fragmentene i basepar. Når elektroforesen var ferdig, ble det tatt et bilde av agrose gel 2%.

Videre har bildet fra elektroforesen blitt analysert ved hjelp av ImageJ/Fiji og videre data-analyser ble gjort i R. Logaritmen av molekylvekten ble plottet mot vandret distanse på gel for å estimere størrelse av DNA-fragment med høyere presisjon.



Figur 2.2

2.3.3 Resultat

Figure 2.2 viser sammenhengen mellom logaritmen av molekylvekten og distanse vandret på gel. Predikert molekylvekt er oppgitt i basepar (bp).

Tabell 2.1: Predikert molekylvekt (bp) for gitte distanser (au)

Ukjente distanser (au)	Predikert molekylvekt (bp)
18	4,596
384	994
408	895
436	796
470	690
508	588
560	475
600	400
618	371
696	267
742	221
798	174
862	133
936	98
994	77
1,208	31

Av Table 2.1 ses estimerte verdier for molekylvekten (bp) etter gitt distanse de har vandret

nedover gelen. Testresultatene indikerer at for hver enhet man vandrer nedover gelen vil DNA-fragmentets vekt minke med -0.00419. Med modellen hvor molekylvekten predikerers, estimeres en minkning i molekylvekt på -0.00126.

2.3.4 Diskusjon

Årsaken til forskjellen i de faktisk målte resultatene fra labben og de predikerte resultatene kan være målefeil ved bruk av Fiji-programmet, der vi estimerte vandret distanse etter bildene som ble tatt.

2.4 Tolke en regresjonstabell

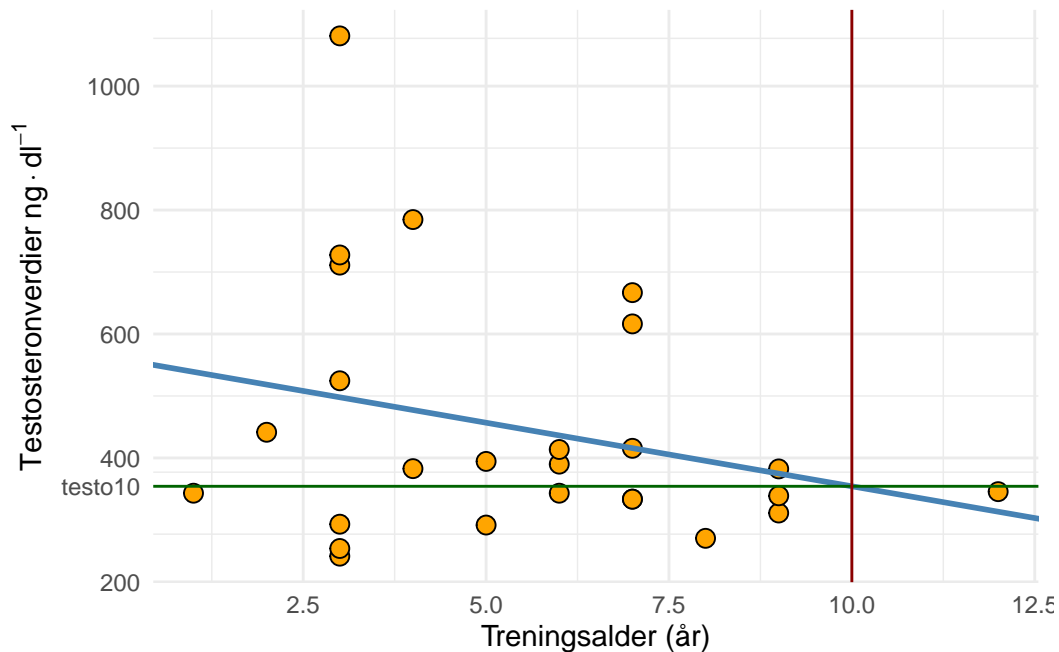
2.4.1 Introduksjon

Mennesker reponderer forskjellig på styrketrening og det er ønskelig å undersøke hvorfor det er slik. Ulike biomarkører ble undersøkt for å forsøke å finne ut av hva som er årsak til den ulike responsen mellom personer som responderer godt på styrketrening, kontra de som ikke responderer så godt. (Haun *et al.*, 2019)

2.4.2 Metode

Denne studien undersøker biomarkører for muskelhypertrofi for lav (LOW) og høy (HIGH) respons på motstandstrening. 30 mannlige universitetsstudenter gjennomførte et 6 ukers høyvolum treningsprogram, og ble delt inn i LOW og HIGH grupper basert på endringer i muskelstørrelse og sammensetning. Det ble tatt blodprøver der testosteronivå i blodet var en av biomarkørene av interesse. (Haun *et al.*, 2019)

2.4.3 Resultat



Figur 2.3: Figur 3: Sammenheng mellom treningsalder og testosteronverdier i blodet

Den lineære modellen forteller i dette tilfellet at for hvert år man trener, så vil nivå av testostertone synke med $20.51 \text{ ng} \cdot \text{dl}^{-1}$.

2.4.4 Diskusjon

Fra datasettet hypertrophy valgte vi å se på sammenhengen mellom testosteronkonsentrasjon i blodet ($\text{ng} \times \text{dl}^{-1}$) og treningsalder (antall år med trening). Den lineære modellen forteller at testosteronkonsentrasjonen i blodet synker med $20.51 \text{ ng} \times \text{dl}^{-1}$ for hvert treningsår. Etter 10 år med trening, estimerer den lineære modellen et testosteronnivå på $354.26 \text{ ng} \times \text{dl}^{-1}$.

Analysen av dataene viser en t -verdi på -1.39 og en p -verdi lik 0.178 , noe som indikerer at det ikke er statistisk signifikant bevis for en sammenheng mellom treningsalder og nivået av testosteron i blodet. Den negative t -verdien indikerer at testostertonnivå synker i tråd med økende treningsalder. Den lave t -verdien tyder dog på at denne effekten er relativt liten.

Siden p -verdien er høyere enn det tradisjonelle signifikansnivået på 0.05 , kan vi ikke avvise nullhypotesen, som antyder at det ikke er noen betydelig effekt eller sammenheng mellom de to variablene i dette datasettet. Dette betyr at variasjonen i testosteronnivåer ikke ser ut til å være relatert til hvor lenge individene har trent.

2.5 Referanser

3 Statistisk inferens

3.1 Introduksjon

Statistisk inferens er en type statistikk som omhandler metoder for å trekke konklusjoner om en populasjon basert på et utvalg. Målet er å generalisere funnene fra utvalget og bruke det til å si noe om hele populasjonen. En viktig del av statistisk inferens er å vurdere usikkerheten i estimatene, typisk ved hjelp av konfidensintervaller og p-verdier (Cremers *et al.*, 2017).

Effektstørrelser gir en kvantitativ vurdering av hvor betydelig en observert effekt er, og er ikke avhengig av utvalgsstørrelse. Dette er nyttig i forskning, da det hjelper med å forstå de praktiske betydningene av funnene man har gjort. Ved å kombinere statistisk inferens med effektstørrelser kan man bedre vurdere resultatene sine og sette det i en bred kontekst (Cremers *et al.*, 2017).

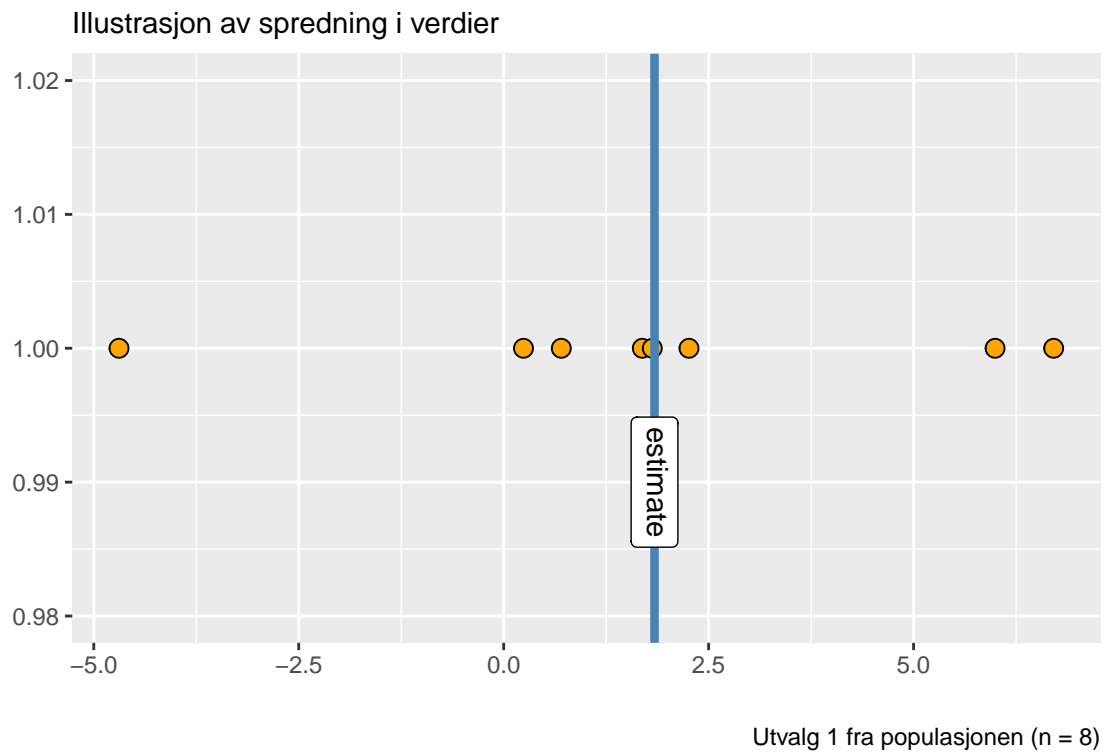
For å illustrer

3.2 Forklaring av *T*-verdi, *P*-verdi, Standard Error (SE) og Estimat

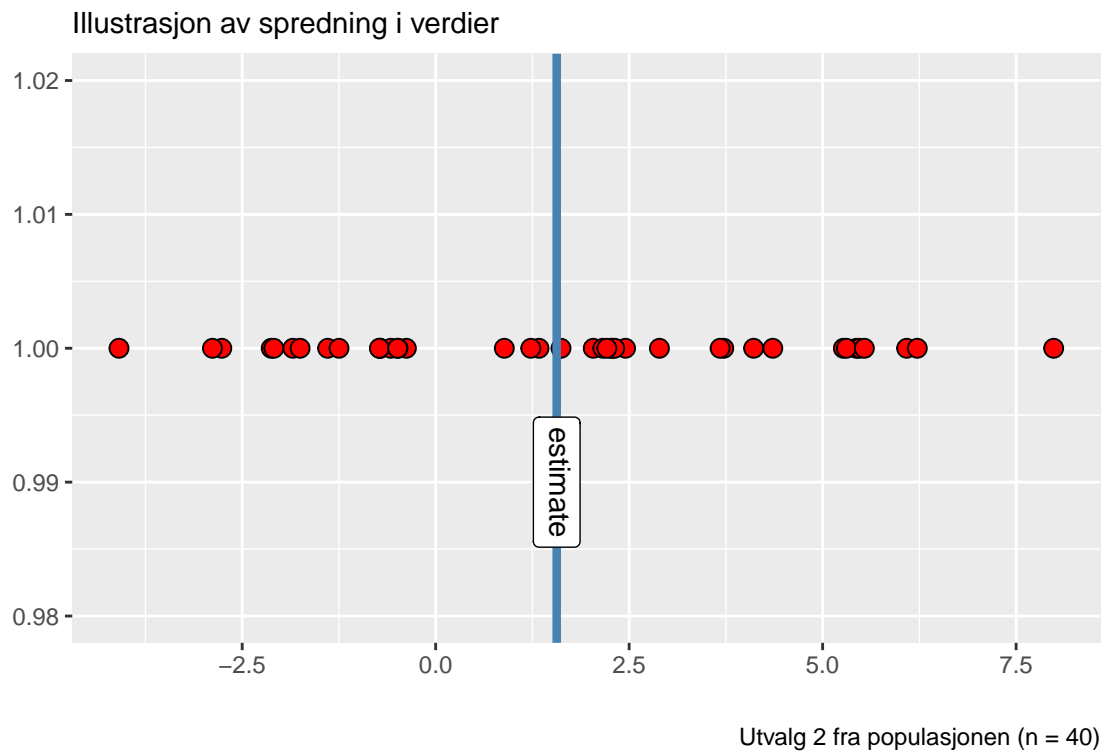
T-verdien er et mål som brukes i hypotesetesting for å vurdere om det er en signifikant forskjell mellom gjennomsnittet av et utvalg og et kjent eller hypotetisk populasjonsgjennomsnitt. *T*-verdiens størrelse er en indikasjon på hvor stor forskjell det er mellom utvalget og populasjonen. Jo større *t*-verdi, desto mer signifikant er forskjellen.

$t = \frac{(\bar{x} - \mu_0)}{\frac{s}{\sqrt{n}}}$, der \bar{x} = gjennomsnittet av utvalget, μ_0 = gjennomsnittet av populasjonen, s = standardavviket, n = størrelsen av utvalget

P-verdien sier noe om usikkerheten i trekningen av utvalget. Kanskje trakk man i sample 1 bare ekstreme verdier. Utvalget vil neppe bli en perfekt representasjon av populasjonen uansett hvordan man trekker ett utvalg. Dette er en usikkerhet man må ta høyde for i ulike studier og dette kan *p*-verdien hjelpe oss å si noe om. Om vi tar høyde for at nullhypotesen stemmer, hvor mange verdier kan vi forvente er mer ekstreme enn våre observasjoner. *p*-verdien måler altså observasjonene våre opp mot den spesifikke nullhypotesen. Lave *p*-verdier indikerer at det er lite sannsynlig at den observerte forskjellen skyldes tilfeldigheter, og en kan forkaste nullhypotesen. Motsatt indikerer høye *p*-verdier at den observerte forskjellen trolig skyldes tilfeldigheter og en har ikke godt nok grunnlag til å si at det er en forskjell mellom gruppene.



Figur 3.1: Spredning i verdier for $n = 8$



Figur 3.2: Spredning i verdier for $n = 40$

Standard error (SE) er en måte å angi feilmarginen av et estimat eller måling. Sagt på en annen måte, beskriver SE variasjonen ved repeterte gjennomsnitt for et utvalg knyttet til en populasjon. Man benytter ofte sentralgrenseteoremet i denne sammenhengen, som sier at hvis man gjør en undersøkelse mange ganger, vil resultatene samle seg omkring den sanne verdien. Jo større utvalg, jo mindre feilmargin. Manuelt kan man beregne SE ved å ta standardavviket og dele på kvadratroten av antall observasjoner. Dette gir oss altså et svar på hvor langt unna det sanne gjennomsnittet vi kan forvente å komme. I dette tilfellet ser en at SE for sample 1 er betydelig større enn i sample 2, nettopp pga et mindre utvalg i sample 1. En kan altså forvente å få verdier lenger unna den sanne gjennomsnittsverdien ved et mindre utvalg.

Estimatet er en verdi som blir kalkulert basert på et utvalg og blir brukt til å estimere noe om en populasjon. Det kan eksempelvis være et estimat på det sanne gjennomsnittet i en befolkning.

3.3 Hva bidrar til de ulike resultatene i de to studiene (m1 og m2)

Forskjellen mellom sample 1 og 2 er først og fremst antall som trekkes ut (utvalget). Med et større utvalg blir SE lavere, t -verdien høyere og p -verdien lavere, sammenlinket med et utvalg som er mindre. En får mindre variabilitet med større utvalg og større usikkerhet ved mindre utvalg.

3.4 Hvorfor benytter en begge “halene” i t -fordelingen (two-tailed test)

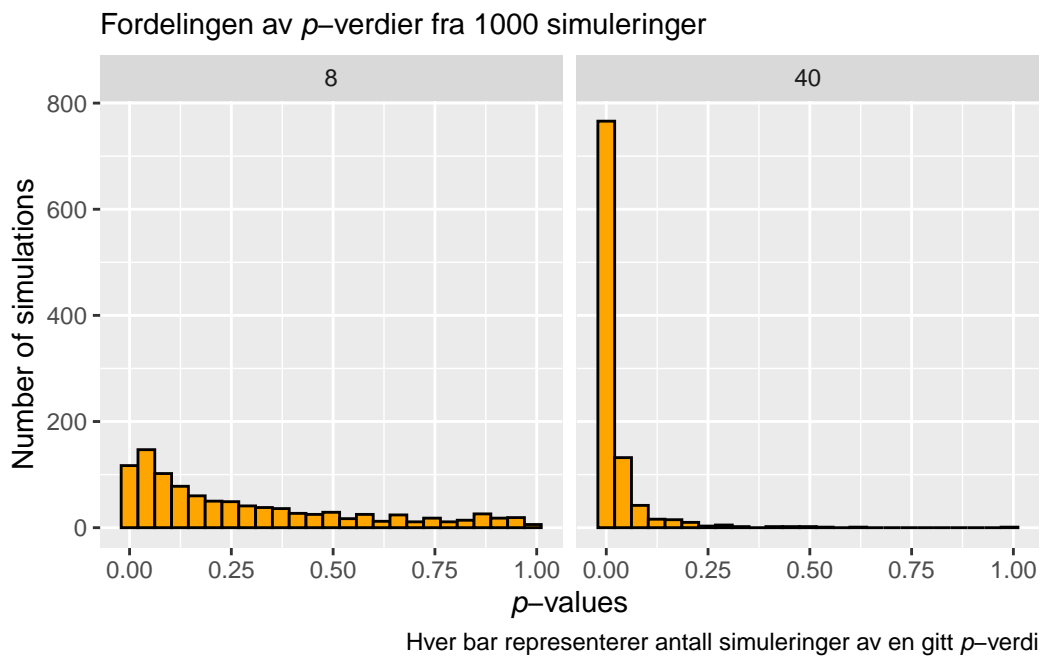
En two-tailed test tillater at man ser på mulighet for endring i begge retninger, altså positiv og negativ retning. Når man undersøker ulike behandlinger, for eksempel for blodtrykk, vil man være interessert i å vurdere både positive og negative endringer mellom gruppene. En ser etter et avvik i begge retninger av fordelingen og ser om avviket er stort nok til å kunne forkaste nullhypotesen (ref. avsnitt om t -verdi over). Høye t -verdier referer ofte til lave p -verdier, som vil si at det er lite sannsynlig at den observerte forskjellen skyldes tilfeldigheter, gitt at nullhypotesen er sann. Ettersom en ser på forskjeller i begge retninger av fordelingen, må en doble p -verdien ved bruk av two-tailed test.

3.5 Kalkuler standardavvik for estimat-variabelen og gjennomsnittet av se-variabelen for begge utvalgene.

Standard error: $SE = \frac{\sigma}{\sqrt{n}}$, der σ = standardavviket for utvalget og n = antall i utvalget.

For utvalget med $n=8$ er standardavviket for snittet av utvalget 1.039 og gjennomsnittet av standardfeilen 1.012. Altså tilnærmet likt. Det samme gjelder for utvalget med $n=40$ (bare med mindre avvik) der standardavviket for snittet av utvalget er 0.474 og gjennomsnittet av standardfeilen er 0.472. Hvordan kan man definere standardfeilen, sett i lys av disse variablene? På generell basis kan man si at SE for et utvalg er et estimat på standardavviket til den teoretiske fordelingen av gjennomsnitt. Observerer at når “ n ” (utvalget) øker, vil standardfeilen synke (ref. formel for beregning av SE).

3.6 Lag et histogram av p -verdiene for de to uvalgene. Hvordan kan histogrammene tolkes og hva forteller de om effekten av størrelsen av utvalget for statistisk styrke?



Utvalgsstr.	Signif. res. / 1000
8	0.222
40	0.877

I histogrammet med $n=8$ kan man se stor variasjon innad i utvalgene. Ved de 1000 simuleringene som er gjort er det sannsynlig at det har vært stor spredning i de variablene som er trekt i samme utvalg (og motsatt). Dermed vil disse enkelt-utvalgene gjøre utslag på p -verdien begge veier. Ved å plote det slik fremheves også tendensen til falsk-positive simuleringer ved $p =$

0.05. En falsk positiv oppstår når en test indikerer en signifikant effekt, slik at man forkaster nullhypotesen (H_0), men det egentlig ikke er en effekt der (H_0 er sann).

3.7 Antall observasjoner fra hvert utvalg som er under signifikansnivået på 0.05

Setter signifikansnivå til $p < 0.05$ Det er 877 tilfeller der $p < 0.05$ for utvalget med 40. Det er 222 tilfeller der $p < 0.05$ for utvalget med 8.

3.8 Estmerer effekten ved one-sample t-test

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

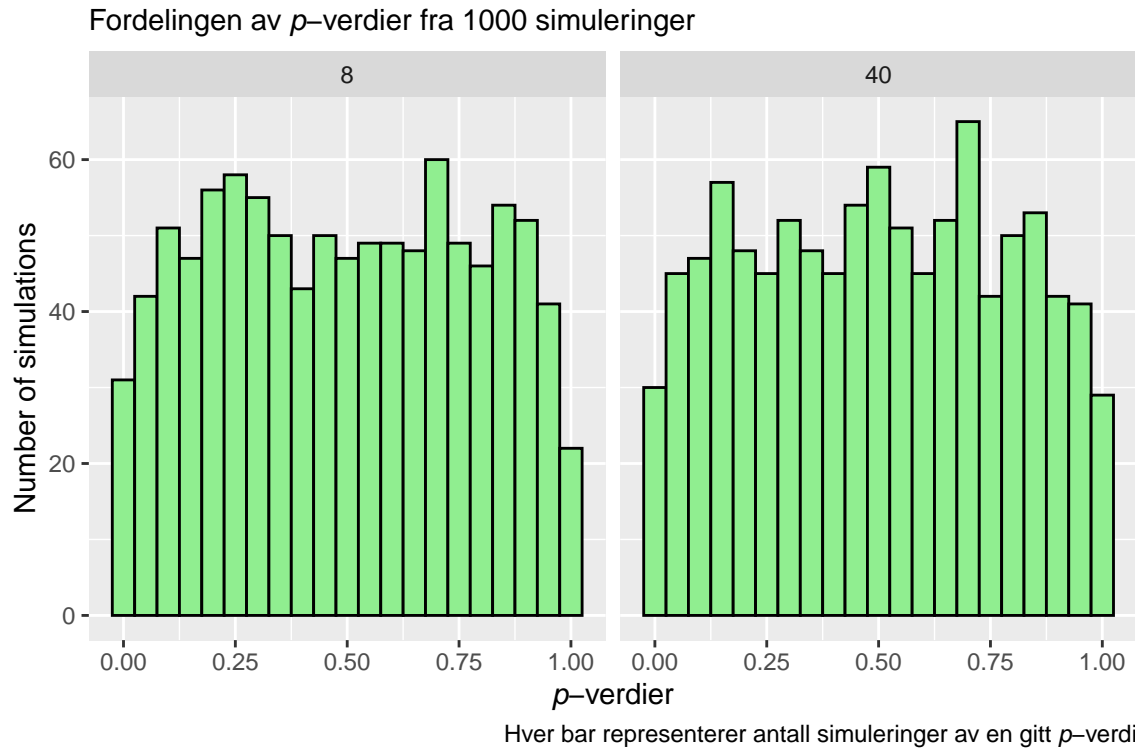
```
      n = 8
      d = 0.5
sig.level = 0.05
  power = 0.232077
alternative = two.sided
```

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

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

Effekt ved utvalget på 8 stk: 0.232. Effekt ved utvalget på 40 stk: 0.869. En kan forvente høyere effekt med et større utvalg. Dette observeres i vårt tilfelle. Med høyere effekt, vil en også forvente å se større grad av signifikans, altså en lavere p -verdi (som også er tilfellet i denne simuleringen).

3.9 Ved å benytte det nye datasettet: hvor mange tilfeller av “falsk positive”-tester vil vi komme over ved å repetere studien mange ganger?



Ved å repetere studien 1000 ganger, vil vi komme over 53 falsk-positive tester for utvalget med 8, og 57 for utvalget med 40. Totalt antall falsk-positive tester blir 110.

3.10 Referanser

4 Studiedesign

4.1 Introduksjon

Studiedesign er en viktig komponent i forskning da det vil være avgjørende for hvordan data samles inn, tolkes og analyseres. Et godt studiedesign vil ofte resultere i mer pålitelige resultater og større grad av reproduserbarhet. Studier som ser seg mer anvendbare i praksis er å foretrekke. (Zabor *et al.*, 2020)

Struktur er en viktig komponent når man planlegger en studie. QALMRI-metoden kan bidra til å sikre en ryddig og god struktur. QALMRI adresserer spørsmål som en ønsker svar på, for så å oppgi hvilke tilnærminger som skal benyttes for å besvare spørsmålet. Videre ser man på sammenhengen mellom metodene som benyttes i studien og koblingen til spørsmålet, eller hypotesten, man har stilt i oppgaven. Til slutt presenteres resultatene og man drar konklusjoner til den aktuelle populasjonen.

I denne oppgaven studeres fem originalartikler som tar for seg effekten av blokkperiodisering på $\dot{V}O_{2max}$ (Rønnestad *et al.*, 2019), (Breil *et al.*, 2010), (Rønnestad *et al.*, 2016), (Rønnestad *et al.*, 2014b), (Rønnestad *et al.*, 2014a). I oppgaven blir det lagt mest vekt på studiedesignene og de statistiske analysemetodene fra de utvalgte studiene, samt en vurdering av styrker og svakheter knyttet til disse.

4.2 Metode

Alle studiene i oppgaven er randomiserte kontrollerte forsøk (RCT) der man var interessert i å undersøke effekten av periodisering av høyintensiv trening. Utvalget ble delt i to grupper, en blokkperiodiseringsgruppe (BP) og en kontrollgruppe (CT). I studiene trente utvalget i BP og CT det samme volumet totalt over intervensjonsperioden. Forskjellen var at BP periodiserte sine høyintensive økter i bolker/mikrosykluser (periodiserignen varierte mellom studiene). Det ble gjennomført målinger både pre og post treningsintervensjonen og intervensjonene hadde varighet fra 4 til 12 uker.

I studiene lå utvalget et sted mellom 16 og 22 personer. I noen studier så en på både kvinner og menn, i andre studier kun menn. Gjengangeren var at utvalget ble definert som godt trente utøvere, eller at de hadde drevet kompetetivt i sin idrett de siste årene.

Det ble gjennomført ulike statistiske analyser i de ulike forsøkene. Noen benyttet ANOVA, mens andre benyttet t-tester, eller en kombinasjon av begge. Flere av studiene benyttet Cohen's d for å beregne effektstørrelse (ES) da dette er en gunstig metode når man vil vurdere effekten av en intervensjon og sammenlikne to grupper (Hopkins *et al.*, 2009). I studiene ble resultater presentert som gjennomsnittsøkning i $\% \pm$ standardavvik. Resultater ble ansett som statistisk signifikante dersom p -verdier ≤ 0.05 .

4.3 Resultater

Etter å ha sammenliknet kontrollgruppen med blokkperiodiseringsgruppen post intervensjon, kom samtlige studier frem til at blokkperiodisering gav signifikant økning i prestasjonsparametere. Endringen kan ses i tabell 1.

P -verdier < 0.05 i samtlige studier. Den gjennomsnittlige effekten blir målt til å ligge en plass mellom 0.76 og 1.34 i 4/5 studier. En av studiene gjorde i tillegg beregninger på kun de mannlige utøverne i sin studie og fant mer signifikante resultater. Denne studien presenterte ingen effektstørrelser.

Studie	Relativ endring i $\dot{V}O_{2max}$	p -verdier	Effektstørrelse
1	$5.1 \pm 3.3\%$	$=0.05$	1.04
2	$5.1 \pm na$	<0.01	na
3	$2.0 \pm 2.5\%$	<0.05	0.76
4	$4.6 \pm 3.7\%$ (relative)	<0.05	1.34
5	$8.8 \pm 5.9\%$	<0.05	1.08

Tabell 1: Alle verdier er fra BP-gruppen.

4.4 Diskusjon

Styrken ved å benytte RCT som studiedesign er at det minimerer bias og skaper godt sammenlikningsgrunnlag blant liknende intervensjoner. Ved å benytte RCT blir også resultatene mer pålitelige. Ved å benytte objektive data i sine resultater, skapes en anvendbarhet i form av at studiene i større grad er sammenliknbare på tvers av studier og intervensjoner. (Zabor *et al.*, 2020)

En svakhet ved studiene er de små utvalgene. Dette øker sjansen for å gjøre type II-feil (godtakelse av nullhypotesen, tross at den er falsk). Gjennomsnitt i små utvalg påvirkes i relativt stor grad av verdier som ligger langt unna gjennomsnittet (outliers) (Mishra *et al.*, 2019). For å objektivt fjerne outliers fra datasettet kan en bestemme seg for grenseverdier man ønsker at sine verdier skal ligge innenfor (Hopkins *et al.*, 2009).

Det kan tenkes at det er gunstig for fremtidige studier å ha større grad av mangfold, dvs. menn og kvinner, ulike alder og treningsbakgrunn. Dette for å kunne si noe om en bredere populasjon. Utvalget bør dog stå i stil til populasjonen man ønsker å si noe om. Ønsker man eksempelvis å forskrive gode treningsmetoder for toppidrettsutøvere innen sykkelporten, bør utvalget bestå av godt trente syklistere. Da utvalgene i disse studiene bestod av godt trente utøvere, kan det antyde at blokkperiodisering av høyintensiv trening kan være gunstig for utøvere som allerede er på et høyt nivå.

Tid og varighet av intervensjoner er ofte en svakhet i studier. Disse studiene varte i hhv. 5-, 4-, 5-, 6- og 12-uker, noe som gjør det vanskelig å si noe om de langsiktige effektene av blokkperiodisering vs tradisjonell trening. Funnene i studiene kan styrkes dersom man studiene repeteres og/eller følges opp over lengre perioder.

Målet med studiene var å se hvorvidt mikrosykluser med høyintensiv trening øker prestasjonssparameteren $\dot{V}O_{2max}$ mer enn tradisjonell periodisering av trening. For å undersøke potensielle endringer mellom BP og CT i studiene ble det som nevnt benyttet ulike statistiske tester. Grunnet små utvalg og en forventet liten endring i prestasjonsparametere blant de godt trente utøverne, ble det benyttet t-tester da det viser seg hensiktsmessig ved små utvalg og benyttes for å sammenlikne gjennomsnittet mellom to grupper (Mishra *et al.*, 2019).

Det at én av studiene ikke presenterte effektstørrelse kan man stille spørsmål ved. En av grunnene kan være at små utvalg kan skape store variasjoner i resultater og skape lite pålitelige effekter. Et annet punkt å diskutere er hvorvidt resultatene er av klinisk relevans. Dersom forskerne mener at en effektsørrelsen kan gi et feilaktig inntrykk av resultatenes styrke, kan de la være å presentere effekten.

4.5 Konklusjon

Blokkperiodisering viser å gi større økning i $\dot{V}O_{2max}$ i samtlige studier, sammenliknet med tradisjonell trening. Langrennsutøvere, ishockeyspillere, alpinister og syklistere har alle hatt økning i $\dot{V}O_{2max}$ ved å periodisere hard trening. Godt trente utøvere, som potensielt har stagnert med tradisjonell trening, kan ha god effekt av blokkperiodisering. Funnene vil særlig være relevant for godt trente utøvere, da dette var utvalget som ble sett på i samtlige studier. Dette er dog et snevert utvalg, som trolig har egenskaper som skiller seg fra resten av befolkningen. Derfor vil det være utfordrende å generalisere funn og anbefalinger opp mot den generelle befolkningen. Repetisjon av studiene der utvalget innehar større grad av spredning når det kommer til treningsbakgrunn, kjønn og alder kan derav være av interesse.

Alle studiene er randomiserte kontrollerte studier, noe som gir tyngde til grad av validitet og grunnlag for direkte sammenlikning av ulike treningsmetoder. Det parallelle gruppedesignet som ble benyttet isolerer effekten av de ulike treningsintervensjonene.

ANOVA (analysis of variance) og t -tester ble benyttet slik at man kan sammenlikne flere grupper over ulike tidpunkter i intervensjonen. På den måten kan man få innblikk i hvordan ulik form for trening påvirker prestasjonparametere over tid. Valget av testene bidrar å styrke funnene om at blokkperiodisering gir økning i $\dot{V}O_{2max}$, sammenliknet med tradisjonell trening.

4.6 Referanser

5 Repeterte målinger

5.1 Introduksjon

I dagens samfunn er det i stor grad enighet om at styrketrening er viktig for god helse. Befolkningen trender mot en mindre aktiv livsstil, samtidig som man blir eldre og eldre. Etter fylte 50 år, viser studier at man taper et sted mellom 5 og 10 % muskelmasse årlig, som tilsvarer ~ 0.4 kg (Westcott, 2012). Tilstander som aldersrelatert muskeltap, diabetes type II, hjertesykdom, redusert beinmasse og økende grad av fettmasse ses hyppigere og hyppigere. Ved sykdom ser en høy grad av inaktivitet som igjen leder til tap av muskelmasse, redusert metabolsk respons, økt fettmasse og andre liknende tilstander. Aerob aktivitet er for tiden veldig populært blir i stor grad anbefalt for å motvirke utvikling av disse tilstandene, men hva gjelder så styrketrening? (Westcott, 2012)

Styrketrening kan defineres som trening med mål om å bli sterkere, forebygge skader, fremme prestasjon eller trening med mål om å øke muskelmasse (Stone *et al.*, 2007). Noen definerer volum som den viktigste faktoren for å se endringen på cellenivå, mens andre mener at høyintensiv trening kan gi samme effekten (Santana *et al.*, 2021). Radaelli *et al.* (2024) ser at lavintensiv trening også kan ha god effekt for å øke muskelmasse og styrke i underekstremiteter blant eldre individer (≥ 60 år) (Radaelli *et al.*, 2024). Antall sett, antall repetisjoner og motstanden det trenes på benyttes for definere *volumet* av trening, mens *intensiteten* på treningen ofte defineres som en prosentandel av én maksimal repetisjon (1RM). For mange er tidsklemma en problem i hverdagen og flere er trolig interessert i å vite hvor lite trening som skal til for å oppnå ønsket stimuli av treningen man gjennomfører. (Santana *et al.*, 2021)

I denne studien er det ønskelig å se nærmere på hvorvidt volumet av styrketrening er av betydning for økning av muskelmasse og muskelstyrke.

5.2 Metode

5.2.1 Utvalg og studieoversikt

I studien ble det rekruttert 41 individer, menn ($n = 19$) og kvinner ($n = 22$). For å bli inkludert i studien måtte man kunne få lokal anestesi, ikke ha trent mer enn én styrkeøkt i uken siste året, ikke ha nedsatt eller påvirket muskelstyrke grunnet tidligere eller nåværende skade og ikke innta medisin som kan påvirke treningsadaptasjon. Etter gjennomført treningsintervensjon

ble 7 deltakere ekskludert fra dataanalysen da de ikke oppfylte kravene om gjennomføring av intervensjon. Table 5.1 viser en oversikt over antall kvinner og menn som er inkludert i analysering av data og deres respektive gjennomsnittlige karakteristikk. Deltakere som ble inkludert i analysen hadde alle tidligere erfaring med idrett.

Tabell 5.1: Deltakerkarakteristikk

Kjønn	Alder (År)	Høyde (cm)	Vekt (kg)
Kvinne n = 18	22 (1.3)	167.7 (6.9)	64.4 (10.4)
Mann n = 16	23.6 (4.1)	182.9 (5.9)	75.8 (10.7)

Alle verdier er gjennomsnitt og standardavvik.

5.2.2 Spesifikke beskrivelser av testmetode

Intervensjonen varte i 12 uker og bestod av helkroppss styrketrening. Beinøvelser ble gjennomført unilateralt. De ble tilfeldig trukket hvilket bein som skulle gjennomføre ett sett (single set) og hvilket som skulle gjøre tre sett (multiple sets).

5.2.3 Styrketreningsprotokoll

Treningsøktene startet med en standardisert oppvarming som besto av (i) ergometer-sykling i 5 minutter (RPE 12–14), (ii) øvelser med kroppsvekt med 10 repetisjoner av push-ups, sit-ups, ryggekestensjon og knebøy, og til slutt (iii) et sett med 10 repetisjoner på ~ 50 % av 1RM for hver styrkeøvelse i intervensjonen.

Beinøvelsene ble gjennomført i følgende rekkefølge: ettbeins beinpress, leg curl og kneekstensjon. Single settene ble gjennomført mellom andre og tredje sett for multiple settene. Treningsintensiteten økte gjennom intervensjonen. Det ble trent 10RM første to ukene, 8RM de neste tre ukene og 7RM de siste syv ukene. For å gjøre treningen mer gjennomførbar for deltakerne, ble noen av øktene gjennomført uten tilsyn. Deltakerne ble da bedt om å føre detaljert logg. Fra den niende treningsøkten, ble én av de tre ukentlige øktene gjennomført på 90 % av intensitet, mens tiltenkt volum skulle opprettholdes. Det skulle være minimum 48 timer mellom hver fullintensitetsøkt og minimum 24 timer mellom submaksimale økter (90 %). For å sikre restitusjon, ble hver deltaker tildelt en drikk som inneholdt $0.15 \frac{g}{kg}$ protein, $11.2 \frac{g}{kg}$ karbohydrat og $0.5 \frac{g}{kg}$ fett.

5.2.4 Muskulær Styrkevurdering

Muskelstyrke ble vurdert ved hjelp av en dynamometer for å måle isokinetisk og isometrisk styrke. Deltakerne ble kjent med testprosedyrer og fikk flere forsøk på hvert nivå for å sikre nøyaktighet. En repetisjon maksimum (1RM) ble testet for unilateral beinpress og kneekstension, der testene ble forberedt med spesifikke oppvarmingssett.

5.2.5 Måling av Kroppssammensetning og Mager Muskelmasse

Kroppssammensetning ble vurdert med dual-energy X-ray absorptiometri (DXA). Deltakerne ble instruert om å faste og unngå intensiv trening før målingene.

5.2.6 Dataanalyse og Statistikk

Utvalgsstørrelsesberegninger indikerte at 40 deltakere var tilstrekkelig for å oppdage meningsfulle forskjeller. Lineære blandede modeller (LMM) ble benyttet for å analysere endringer i muskelhypertrofi og styrke. Alle analyser ble utført i R, med signifikansnivå satt til $\alpha = 0,05$.

De statistiske analysene hadde som mål å avdekke effekten av forskjellige motstandstreningsvolumer på muskelstyrke og hypertrofi.

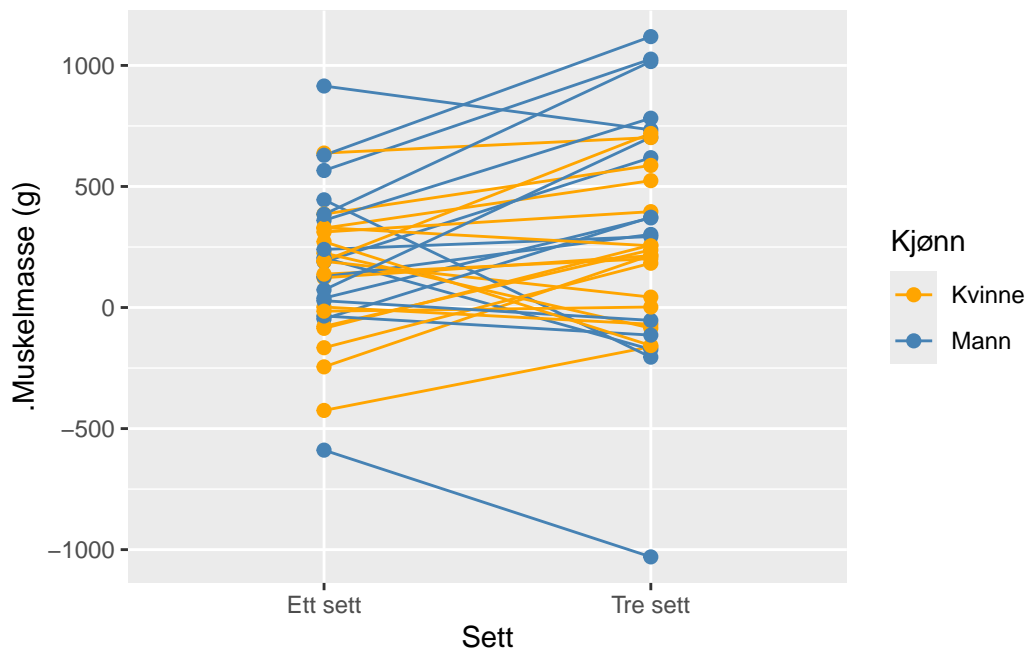
5.3 Resultater

5.3.1 Treningsvolums effekt på mager muskelmasse i underekstremitetene

Etter den 12 uker lange styrketreningsintervensjonen er den prosentvise økningen i mager muskelmasse for benet som trente tre sett 3.25% og for det som trente ett sett 1.9%. Den gjennomsnittlige forskjellen i mager muskelmasse i beina mellom de to treningsvolumene (ett sett vs. tre sett) var 122.8 g (95% CI: [8.6, 237.0], p -value = 0.036, $t_{33} = 2.19$).

Benytter man en blandet modell som tar man høyde for hver deltaker sine pre-verdier, og ser på hvordan antall sett påvirker mager muskelmasse etter treningsintervensjonen, blir den gjennomsnittlige endringen i mager muskelmasse mellom ett og tre sett 122.7 g (95% CI: [10.9, 234.5], p -value = 0.037, $t_{32.8} = 2.18$).

Figure 5.1 viser endringen i mager muskelmasse målt i gram for de ulike deltakerenes bein. Linjen som er trukket mellom punkt viser hver deltakers resultat og gir et inntrykk av forskjellen mellom effekten av treningsvolum på muskelmasse. Tre sett virker av figuren å ha større effekt for å øke muskelmasse, sammenliknet med lavere treningsvolum.



Figur 5.1: Endring i mager muskelmasse fra pre- til post-dexamåling (målt i gram)

5.3.2 Treningsvolums effekt på kraftutvikling i underekstremitetene

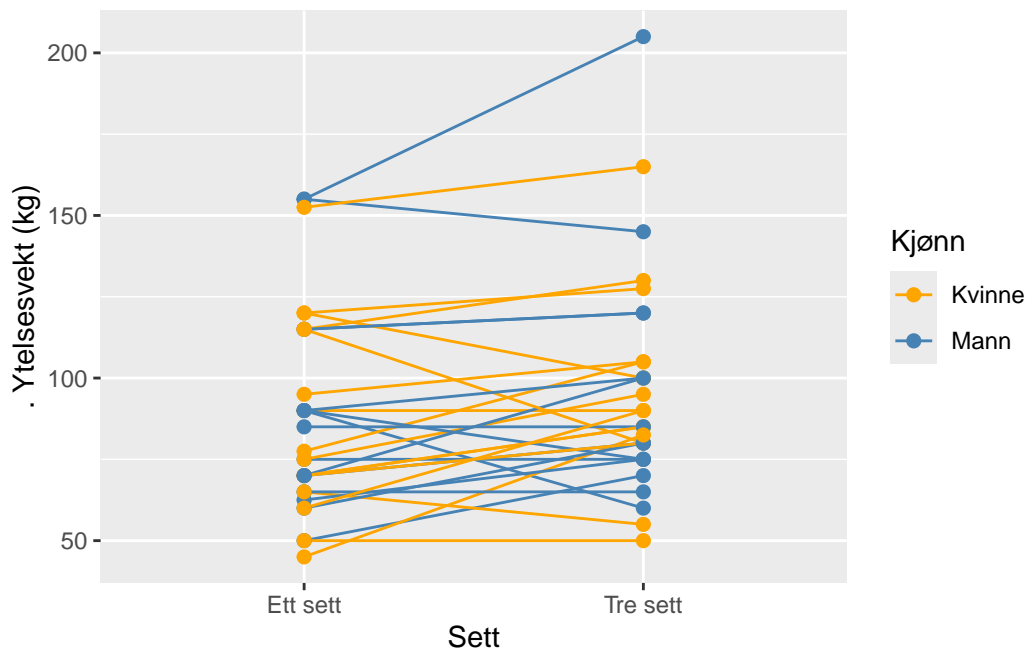
Den gjennomsnittlige prosentvise økningen i ytelsesvekt ved 1RM (målt i kg) for henholdsvis tre og ett sett var 36.89 og 34.15. Den gjennomsnittlige forskjellen i ytelsesvekt mellom beina ved de ulike treningsvolumene (ett sett vs. tre sett) var 7.8 kg (95% CI: [1.1, 14.6], p -value = 0.025, $t_{30} = 2.36$).

Benytter man en blandet modell der man tar høyde for hver deltaker sine pre-verdier, og ser hvordan antall sett påvirker ytelsesvekt etter treningsintervensjonen, blir den gjennomsnittlige endringen i ytelsesvekt mellom ett og tre sett 8.4 kg (95% CI: [1.9, 15], p -value = 0.017, $t_{30.5} = 2.53$).

Figure 5.2 viser endringen i den maksimale ytelsesvekten deltakerne evnet å presse ved 1RM for henholdsvis de som trente ett og tre sett. Av figuren ser man en tendens til at høyere treningsvolum gir større evne til å yte mer på øvelsen beinpress.

5.4 Diskusjon

I tråd med tidligere forskning, som har undersøkt effekten av treningsvolum på muskelmasse og styrke (Stone *et al.*, 2007), (Santana *et al.*, 2021), viser denne studien at høyere volum



Figur 5.2: Endring i ytelsesvekt fra pre- til post-test for beinpress (målt i kg)

(tre sett) har en tendens til å føre til større muskelvekst og styrkeøkning enn lavere volum (ett sett). Dette styrker ideen om at volum kan være en viktig faktor for å oppnå optimal muskelhypertrofi og styrkeutvikling.

I en større kontekst, spesielt i forhold til aldersrelatert muskeltap, er det viktig å påpeke at forskning viser at personer over 50 år mister 5–10 % av muskelmassen årlig (Westcott, 2012). Studien antyder at styrketrening kan være et effektivt tiltak for å motvirke eller bremse dette tapet av muskelmasse. Dette er særlig relevant for eldre voksne, og andre individer, som er lite aktive eller har begynt å trene styrke for første gang.

Studien er dog gjennomført på yngre individer, som tidligere har drevet med idrett (se Table 5.1). Det kan av den grunn være misvisende å trekke paralleller til en populasjon med andre forutsetninger. På en annen side hadde deltakerne i studien trent lite til ingen styrke det siste året. Av den grunn kan en anta at treningseffekten en ser vil være tilstede for andre individer med tilsvarende treningsbakgrunn, uavhengig av alder.

I denne studien med unge utøvere så en frafall underveis i intervensjonen. En kan anta at dette skyldes at treningsmengden blir for høy i forhold til hva kroppen er vant til, noe som kan føre til skader eller andre utfordringer. Dersom en står ovenfor ulike helseutfordringer, kan det være utfordrende å utføre flere sett per øvelse, og derfor kan den praktiske anvendelsen av tre sett være utfordrende. I lys av dette kan det være relevant å vurdere om ett sett også gir tilstrekkelig stimuli for muskel forbedring i populasjoner med begrensede ressurser.

Intensiteten (70–90 % 1RM) brukt i denne studien er i tråd med anbefalingene for både muskelstyrke og hypertrofi (Santana *et al.*, 2021). Imidlertid kunne det vært interessant å undersøke om lavere intensitet på treningen kombinert med høyere volum kunne ha hatt en like god effekt på muskelmasseøkning for denne gruppen.

En langtidsoppfølging av deltakerne etter at treningsintervensjonen er avsluttet, kan gi mer informasjon om hvordan treningsvolum påvirker vedlikehold av muskelmasse og styrke i lengden. Hvordan deltakerne vedlikeholder gevinster over tid kan være viktig for å vurdere praktisk anvendbarhet av treningsvolumet som en strategi for å bevare helsen.

Ved å benytte en blandet modell (ANCOVA) i tillegg til t-test får man en modell som tar hensyn til individuell variasjon. Modellen tar høyde for at hver deltaker har ulike pre-verdier og at de responderer ulikt gjennom intervensjonen. Ved å gi hver deltaker sin egen startverdi, kan man mer presist vurdere og isolere effekten av treningsintervensjonen (ett sett vs tre sett).

5.5 Konklusjon

Treningsvolum og intensitet virker å være viktige parametre for å øke muskelmasse og muskelstyrke. Tre sett med styrkeøvelser synes å gi en større tilpasning enn kun ett sett med øvelser. De langsiktige effektene av denne treningsmodellen bør undersøkes nærmere, i tillegg til om de samme effektene av treningsvolum kan observeres i andre deler av befolkningen.

6 Vitenskapsfilosofi

7 RNA ekstraksjon og analyse

Referanser

- Antonutto G & Di Prampero PE (1995). The concept of lactate threshold. A short review. *J Sports Med Phys Fitness* **35**, 6–12.
- Breil FA, Weber SN, Koller S, Hoppeler H & Vogt M (2010). Block training periodization in alpine skiing: Effects of 11-day HIT on VO₂max and performance. *Eur J Appl Physiol* **109**, 1077–1086.
- Cremers HR, Wager TD & Yarkoni T (2017). The relation between statistical power and inference in fMRI. *PLoS One* **12**, e0184923.
- Halperin I, Pyne DB & Martin DT (2015). Threats to internal validity in exercise science: A review of overlooked confounding variables. *Int J Sports Physiol Perform* **10**, 823–829.
- Haun CT, Vann CG, Mobley CB, Osburn SC, Mumford PW, Roberson PA, Romero MA, Fox CD, Parry HA, Kavazis AN, Moon JR, Young KC & Roberts MD (2019). Pre-training skeletal muscle fiber size and predominant fiber type best predict hypertrophic responses to 6 weeks of resistance training in previously trained young men. *Front Physiol* **10**, 297.
- Hopkins WG (2000). Measures of reliability in sports medicine and science. *Sports Med* **30**, 1–15.
- Hopkins WG, Marshall SW, Batterham AM & Hanin J (2009). Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc* **41**, 3–13.
- Mishra P, Singh U, Pandey CM, Mishra P & Pandey G (2019). Application of student's t-test, analysis of variance, and covariance. *Ann Card Anaesth* **22**, 407–411.
- Pisică D, Dammers R, Boersma E & Volovici V (2022). Tenets of good practice in regression analysis. A brief tutorial. *World Neurosurg* **161**, 230–239.e6.
- Radaelli R, Rech A, Molinari T, Markarian AM, Petropoulou M, Granacher U, Hortobágyi T & Lopez P (2024). Effects of resistance training volume on physical function, lean body mass and lower-body muscle hypertrophy and strength in older adults: A systematic review and network meta-analysis of 151 randomised trials. *Sports Med*.

- Rønnestad BR, Ellefsen S, Nygaard H, Zacharoff EE, Vikmoen O, Hansen J & Hallén J (2014a). Effects of 12 weeks of block periodization on performance and performance indices in well-trained cyclists. *Scand J Med Sci Sports* **24**, 327–335.
- Rønnestad BR, Hansen J & Ellefsen S (2014b). Block periodization of high-intensity aerobic intervals provides superior training effects in trained cyclists. *Scand J Med Sci Sports* **24**, 34–42.
- Rønnestad BR, Hansen J, Thyli V, Bakken TA & Sandbakk Ø (2016). 5-week block periodization increases aerobic power in elite cross-country skiers. *Scand J Med Sci Sports* **26**, 140–146.
- Rønnestad BR, Øfsteng SJ & Ellefsen S (2019). Block periodization of strength and endurance training is superior to traditional periodization in ice hockey players. *Scand J Med Sci Sports* **29**, 180–188.
- Santana DA de, Castro A & Cavaglieri CR (2021). Strength training volume to increase muscle mass responsiveness in older individuals: Weekly sets based approach. *Front Physiol* **12**, 759677.
- Schadock I, Schneider A, Silva ED, Buchweitz MRD, Correa MN, Pesquero JB, Paredes-Gamero EJ, Araujo RC & Barros CC (2015). Simple method to genotype the ACTN3 r577x polymorphism. *Genet Test Mol Biomarkers* **19**, 253–257.
- Stone MH, Stone M & Sands WA (2007). *Principles and practice of resistance training*. Human Kinetics.
- Westcott WL (2012). Resistance training is medicine: Effects of strength training on health. *Curr Sports Med Rep* **11**, 209–216.
- Zabor EC, Kaizer AM & Hobbs BP (2020). Randomized controlled trials. *Chest* **158**, S79–S87.