

Mappeeksamen IDR4000

Kandidatnummer: 506

Table of contents

1	Introduksjon	4
1.1	Forord	4
2	Rapport 1 - Deskriptiv statistikk, reliabilitet og validitet og verktøy for reproduserbar vitenskap	5
2.1	Introduksjon	5
2.2	Metode	5
2.3	Standardisering før test	5
2.4	Utstyr	6
2.5	Protokoll	6
2.6	Databehandling etter test	7
2.7	Resultat	8
2.8	Resultater	9
2.9	Diskusjon	9
3	Rapport 2 - Predikere data ved hjelp av regresjon	10
3.1	Introduksjon	10
3.2	Metode	10
3.2.1	Del 1: Predikert laktatterskel	10
3.2.2	Del 2: Predikert hellning på qPCR kalibreringskurve	10
3.2.3	Del 3: Tolke en regresjonstabell	10
3.3	Resultat	11
3.3.1	Del 1: Predikert laktatterskel	11
3.3.2	Del 2: Predikert hellning på qPCR kalibreringskurve	12
3.3.3	Del 3: Tolke en regresjonstabell	12
3.4	Diskusjon	13
3.4.1	Del 1	13
3.4.2	Del 2	13
3.4.3	Del 3	13
4	Rapport 3 - Statistisk inferens, statistiske modeller og statistisk styrke	14
4.1	Oppgave 1	14
4.2	Oppgave 2	15
4.3	Oppgave 3	15
4.4	Oppgave 4	16

4.5	Oppgave 5	17
4.6	Oppgave 6	17
4.7	Oppgave 7	17
4.8	Oppgave 8	18
5	Rapport 4 - Studiedesign	19
5.1	Innledning	19
5.2	Studiedesign	19
5.3	Analyse av studiene	20
5.4	Sammenligning av statistiske metoder	21
5.5	Konklusjon	22
6	Rapport 5 - Analyse av eksperimenter med repeterte målinger	23
6.1	Introduksjon	23
6.2	Metoder	23
6.2.1	Deltakere og studiedesign	23
6.2.2	Tester og målinger	24
6.2.3	Analyse	25
6.3	Resultater	25
6.4	Diskusjon	26
6.5	Konklusjon	26
7	Filosofihistorie	27
7.1	Spørsmål 1	27
7.2	Spørsmål 2	28
8	Laboratorierapport fra molekylær labb	30
8.1	qPCR	30
8.2	Utstyr	30
8.3	Metoder	30
8.4	Pipetteringsskjema	31
8.5	Diskusjon	34
	References	35

1 Introduksjon

Mappeeksamen består av følgende deler:

- Rapport 1: “Deskriptiv statistikk, reliabilitet og validitet og verktøy for reproduserbar vitenskap”.
- Rapport 2: “Predikere data ved hjelp av regresjon”
- Rapport 3: “Statistisk inferens, statistiske modeller og statistisk styrke”
- Rapport 4: “Studiedesign”
- Rapport 5: “Analyse av eksperimenter med repeterte målinger”
- Arbeidskrav i filosofihistorie
- Laboratorierapport fra molekylærlybb

1.1 Forord

Jeg vil gjerne takke Lone, Lars og Ulrich for et godt samarbeid på Rapport 1/2 og laboratorierapporten. I tillegg vil jeg takke Jens Konrad Daniel Hammarström for hjelp på veien når kodene har vært kranglete. Denne oppgaven ville ikke vært mulig uten de som er nevnt. Under ligger lenken til Mappeeksamen sin repository på GitHub med alle kodene som har blitt benyttet til å besvare de ulike delene av eksamen.

Lenke til repository: <https://github.com/DanielTEriksen/Mappeeksamen>

2 Rapport 1 - Deskriptiv statistikk, reliabilitet og validitet og verktøy for reproducerbar vitenskap

2.1 Introduksjon

Hensikten med denne rapporten er å presentere estimer for reliabilitet av data samlet i fysiologisk testlab ved Høgskolen i Innlandet. Hovedfokuset i eksperimentet er å vise til god standardisering og gjennomtenkt metode for å minske støy, noe som er viktig for å øke reliabiliteten (Halperin, Pyne, and Martin 2015). Ifølge (Hopkins 2000) er reliabilitet reproducerbarheten av verdier fra en test, analyse, eller annen måling ved gjentatte forsøk på samme individ. Vi har gjennomført $\dot{V}O_{2max}$ -tester på sykkel der 16 deltakere har gjennomført 2 eller 4 tester i løpet av tre uker. De to første testene ble gjennomført med ca. 24 timers mellomrom i uke en. De to siste testene ble gjennomført med ca. 48 timers mellomrom i uke tre. Vi har brukt Rstudio for å analysere dataen og få svar på om våre tester er reliable.

2.2 Metode

Vi gjennomførte en test-retest i vår studie hvor vi samlet inn data fra flere $\dot{V}O_{2max}$ -tester. 16 deltakere gjennomførte to til fire tester i løpet av en periode på tre uker. De to første testene ble gjennomført med 24 timers mellomrom i løpet av den første uken. De to siste testene ble gjennomført med 48 timers mellomrom i løpet av den tredje uken.

2.3 Standardisering før test

Resultatene fra en $\dot{V}O_{2max}$ -test påvirkes av ulike fysiologiske og miljømessige faktorer, så det er viktig å sette visse regler som deltakeren må følge i dagene før testen. En veiledning ble sendt til deltakerne noen dager før deres første test med informasjon om hvordan de bør forberede seg til testdagen. Denne veiledningen bør inkludere standardiseringsrutiner som hjelper til med å isolere deltakernes sanne aerobe ytelse og eliminerer eksterne faktorer som kan påvirke testresultatene. Biologiske faktorer som kontrolleres for er fysisk aktivitet, hydrering, søvn, koffein- og kaloriinntak, alkoholforbruk og tidspunktet på dagen testen gjennomføres.

Deltakeren må følge den samme rutinen før sin neste test for å sikre nøyaktighet, pålitelighet og sammenlignbarhet av testresultatene.

2.4 Utstyr

Spesifikt utstyr er nødvendig for å utføre en $\dot{V}O_{2max}$ -test. Deltakerne utførte testen på en elektromagnetisk bremset sykkelergometer (Lode Excalibur Sport, Lode. B.V., Groningen, Nederland) for å utfordre deltakernes aerobe kapasitet. Vi brukte også et system kalt Vyntus CPX mixing chamber (Vyntus CPX JaegerCareFusion, UK) for å måle $\dot{V}O_{2max}$. For å måle hjertefrekvensen brukte vi en bryststropp fra Polar. For å måle deltakernes laktatnivå brukte vi en maskin kalt Biosen c-line (EKF diagnostic GmbH, Barleben, Tyskland). Det ble også benyttet en stoppeklokke, en ekstra datamaskin for vårt Excel-ark og en vekt for å måle kroppsvekt. Hele oppsettet gir oss mulighet til å overvåke og måle nødvendige data for å bestemme deltakernes $\dot{V}O_{2max}$.

2.5 Protokoll

Testprotokollen for å utføre en $\dot{V}O_{2max}$ -test beskriver hvordan testen skal gjennomføres hver gang, da standardisering er viktig for å få nøyaktige og pålitelige resultater. En deltaker bør derfor ha samme testleder dersom en test gjentas. Under testene våre fulgte vi følgende protokoll, der et Excel-ark ble klargjort av testlederen før ankomst til laboratoriet. Ved ankomst tar testlederen på seg labfrakk og sørger for at Biosen (laktatmåling) er slått på. Maskinen må deretter kalibreres, og resultatet skal være 12 mmol/l. Ved kalibreringsfeil, som kan føre til unøyaktige data, byttes standardvæsken. Vyntus må deretter kalibreres ved hjelp av gasskalibrering og volumkalibrering. Gasskalibreringen må være innenfor 2,0 i differanse, og volumkalibreringen må være innenfor 0,2. Her må gassbeholderen åpnes, og ventilasjonsvolumtransduseren (Triple-V-turbinen) festes til Vyntus. Mens kalibreringen pågår, monterer testlederen munnstykket sammen med nese-klemmen og kobler det til miksekammeret via en slange. Deretter må ergometersykkelen kalibreres ved å posisjonere krankarmen vertikalt. Denne kalibreringen gjøres på datamaskinen koblet til ergometersykkelen, da vi bruker et annet program her enn Vyntus. Krankarmen kontrolleres for å sikre at den er 172,5 mm lang og at riktig pedaltype er montert. Til slutt samles utstyret for å måle laktat, og alt er klart for deltakerens ankomst.

En deltakerprofil må opprettes både i Vyntus og ergometersykkelens program. Profilen inkluderer deltakerens navn, fødselsdato, kjønn, høyde og vekt. Det første vi gjør når deltakeren ankommer, er å måle vekt. Vekten reduseres med 300g p.g.a. klær. Deretter justeres ergometersykkelen slik at deltakeren sitter komfortabelt. Innstillingene på sykkelen lagres deretter i sykkelprogrammet for fremtidige tester. Oppvarmingen varer i 5 minutter, der deltakeren skal ha en gradvis økning på Borgs skala for opplevd anstrengelse (RPE) fra 10 til 13. Hvis

deltakeren ønsker det, slås viften på og rettes mot dem. Testlederen gir deretter informasjon om testen og Borg-skalaen.

Vyntus settes deretter til “måling,” og testlederen sørger for at pulsmåleren er koblet til Vyntus, samt at innstillingene er satt til 30 sekunders målinger og størrelse medium for munnstykket. Gassbeholderen lukkes, og Triple-V kobles fra Vyntus og festes til miksekammeret. Oppvarmingen er ferdig, og testen starter i Vyntus. Selve testen begynner etter 1 minutt i Vyntus, og $\dot{V}O_{2max}$ -protokollen startes i sykkelprogrammet. En stoppeklokke startes også og plasseres foran deltakeren.

Under testen informerer testlederen om økninger i motstand (W), normalt 20W for kvinner og 25W for menn, og oppmuntrer deltakeren til å trække til utmattelse. Testen avsluttes når deltakeren gir seg eller RPM faller under 60. Testlederen noterer makspuls, sluttid, watt ved endt test og Borg-score på slutten av testen i excel-arket. Laktat tas ett minutt etter testslutt ved å tørke av fingeren til deltakeren, stikke et hull, tørke av den første bloddråpen og deretter fylle opp røret. Blodet analyseres deretter i Biosen. Testen er ferdig for deltakeren. Testlederen avslutter testen i både Vyntus og sykkelprogrammet. Sykkelen, munnstykket, slangen og pulsmåleren vaskes, og viften rettes mot miksekammeret for å la det tørke.

2.6 Databehandling etter test

Korrekt innsamling av data fra testen er avgjørende, da det gir oss innsikt i deltakerens aerobe kapasitet. Rapporten generert av Vyntus lagres på en minnepinne for senere opplasting til OneDrive. Excel-arket fylles deretter med verdier fra rapporten. Disse verdiene inkluderer absolutt $\dot{V}O_{2max}$, uttrykt i milliliter oksygen per minutt (ml/min), som gir oss det totale oksygenforbruket til deltakeren og er nødvendig for å vurdere den aerobe kapasiteten. Maksimal mengde karbondioksid produsert ($vco2.max$) viser hvor mye CO₂ som utåndes under testen. Respiratoriskutveklingsforhold (RER) estimerer forholdet mellom produsert karbondioksid (CO₂) og forbrukt oksygen (O₂) under metabolisme. Maksimal minuttventilasjon (VE.max) er det høyeste volumet av luft deltakeren kan bevege inn og ut av lungene per minutt. Maksimal pustefrekvens (BF.max) er det maksimale antallet pust per minutt. Til slutt samles laktatmålingene fra Biosen. Deretter fikk vi data fra de andre gruppene og samlet alt i ett excel-ark. Dataene ble deretter eksportert til RStudio for videre analyse av resultatene.

2.7 Resultat

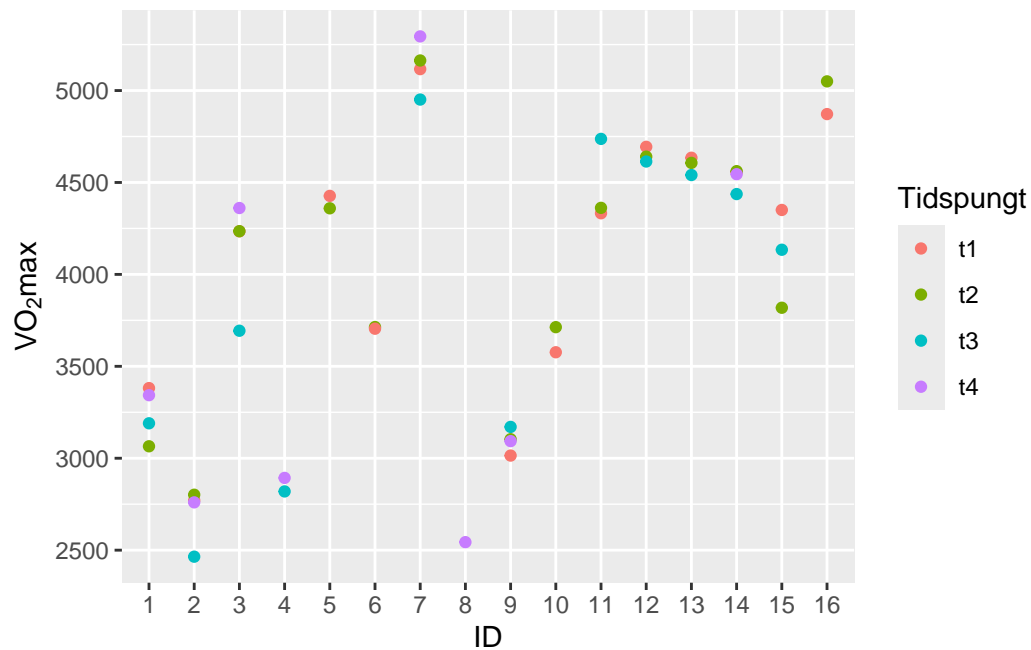


Figure 2.1: Deltakere og testresultat

Table 2.1: Deltakerenes VO₂max-verdier i ml/min

ID	Test 1	Test 2	Test 3	Test 4
ID1	3381.5	3065.0	3190.0	3343.0
ID2	2771.0	2801.5	2464.5	2760.0
ID3	4234.5	4235.0	3693.5	4361.0
ID4	NA	NA	2819.5	2893.0
ID5	4427.0	4359.5	NA	NA
ID6	3704.5	3713.5	NA	NA
ID7	5116.5	5163.5	4951.0	5294.5
ID8	NA	NA	NA	2543.5
ID12	4694.0	4640.5	4614.0	NA
ID9	3014.5	3103.5	3170.5	3093.0
ID10	3576.5	3713.0	NA	NA
ID11	4332.5	4362.0	4737.0	NA
ID13	4634.5	4606.5	4540.5	NA
ID14	4556.5	4561.5	4437.0	4545.0
ID15	4350.5	3818.5	4134.0	NA
ID16	4872.0	5050.0	NA	NA

2.8 Resultater

Resultatene fra vår utrekning av målefeil fra $\dot{V}O_{2max}$ -testene viste at mellom t1 og t2 var r cv_percent1 % variasjonskoeffesienten og mellom t3 og t4 var r cv_percent2 % variasjonskoeffesienten. Dette tilsvarer en målefeil på r te_1ml mellom t1 og t2 og r te_2ml mellom t3 og t4.

2.9 Diskusjon

Variasjonskoeffesientene r cv_percent1% og r cv_percent2% er et mål på den typiske prosentvise feilen i målingene. Det er både svakheter og styrker ved våre gjennomføringer av testene. Dette var for mange, både som deltaker og testleder, det første møte med fysiologisk testlab og $\dot{V}O_{2max}$ -test. Selv om formålet med rapporten var å estimere reliabilitet, var hensikten med testgjennomføringene at deltakerne skulle lære seg å gjennomføre en $\dot{V}O_{2max}$ -protokoll. Dette førte til svakheter som at det ble brukt 16 forskjellige testledere og assistenter, dataen ble samlet inn av fire forskjellige grupper, antall observatører varierte og innsatsen var ikke alltid maksimal da motivasjonen ikke var på hos alle deltakerne. På en annen side var utstyret det som regnes som gullstandarden ved test av $\dot{V}O_{2max}$. Læreren som underviste hadde mye erfaring og høy kompetanse.

I følge (Hopkins 2000) er observert endring reliabel om den er større enn den typiske feilen. I vårt tilfelle vil det tilsvare en større endring enn r te_1ml og r te_2ml noe vi synes å være i overkant mye. De nevnte svakhetene ved gjennomføringene vil operere som forstyrrende variabler og kan dermed virke negativt på reliabiliteten (Halperin, Pyne, and Martin 2015).

Basert på vår estimering av reliabilitet fra testene synes vi at variasjonskoeffesienter på r cv_percent1% og r cv_percent2% er for høye. Det vil da kreve i overkant mye å skulle bekrefte endring ved ny test. Dette kommer mest sannsynlig som følge av en for dårlig og variert gjennomføring av metoden.

3 Rapport 2 - Predikere data ved hjelp av regresjon

3.1 Introduksjon

Hensikten med denne rapporten er å predikere data ved hjelp av regresjon gjort i Rstudio, samt tolke en regresjonstabell. Rapporten inneholder tre deler. Ved å bestemme laktatterskel ved blodlaktatverdiene 2 og 4 mmol L⁻¹ analyserer vi forholdet mellom prestasjon i watt og treningsintensitet. Vi analyserte hellningen til en qPCR kalibreringskurve, og tolket en regresjonstabell om forholdet mellom 3RM squat og tverrsnittsareal til type II muskelfibrer.

3.2 Metode

3.2.1 Del 1: Predikert laktatterskel

I del 1 av rapporten ble datasettet cyclingstudy fra (SYLTA et al. 2016) brukt til å predikere to bestemte blodlaktatterskler ved 2 og 4 mmol L⁻¹.

3.2.2 Del 2: Predikert hellning på qPCR kalibreringskurve

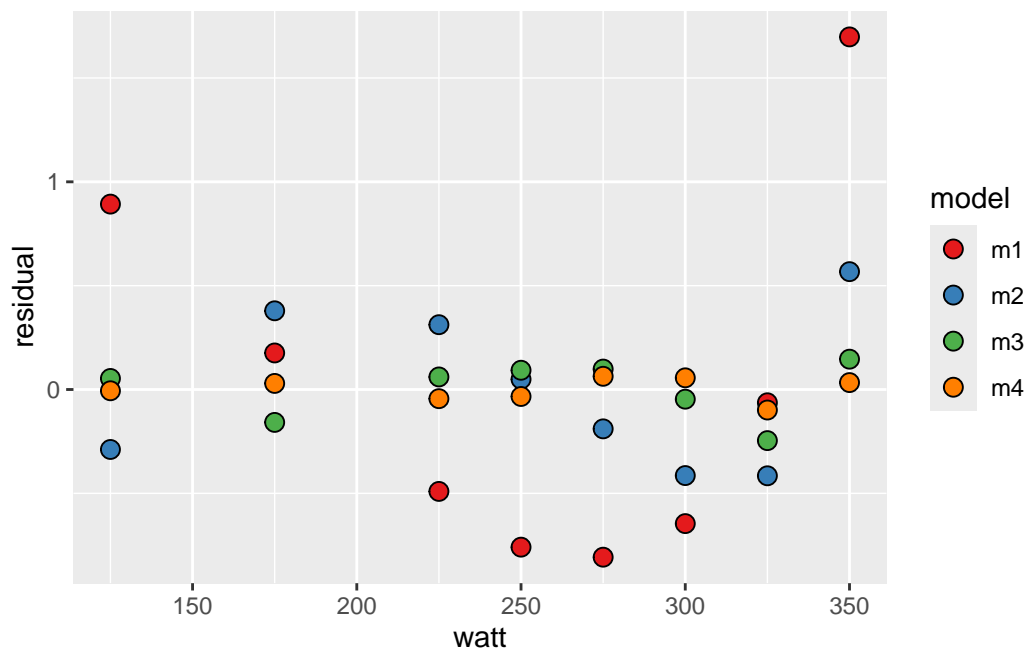
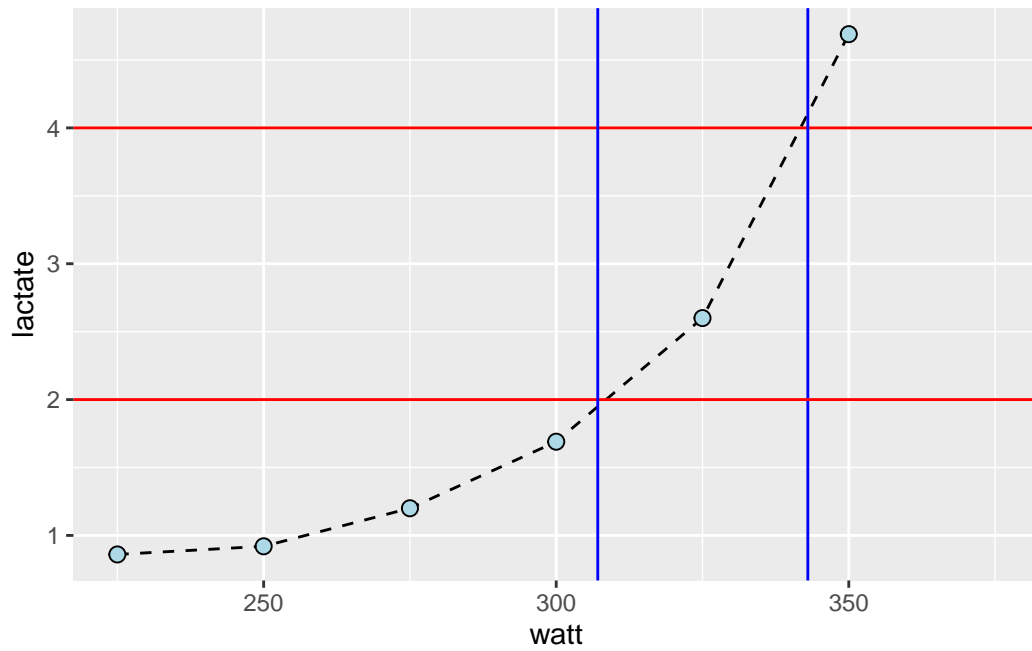
I del 2 av rapporten har vi ved hjelp av (Schindelin et al. 2012) analysert et bilde av qPCR som ble hentet fram fra forsøket DNA. Analysen av bildet gav oss data som vi anvendte for å predikere hellningen til qPCR kalibreringskurven.

3.2.3 Del 3: Tolke en regresjonstabell

I del 3 av rapporten har vi gjort en statistisk analyse av forholdet mellom “Type II (FAST) fibers cross sectional area (micrometer²)” ved baseline (FAST_CSA_T1) og “Squat 3 repetition maximum load (kg)” ved baseline (SQUAT_3RM) fra datasettet til (Haun et al. 2018) og (Haun et al. 2019) for å undersøke om det var et lineært forhold mellom disse.

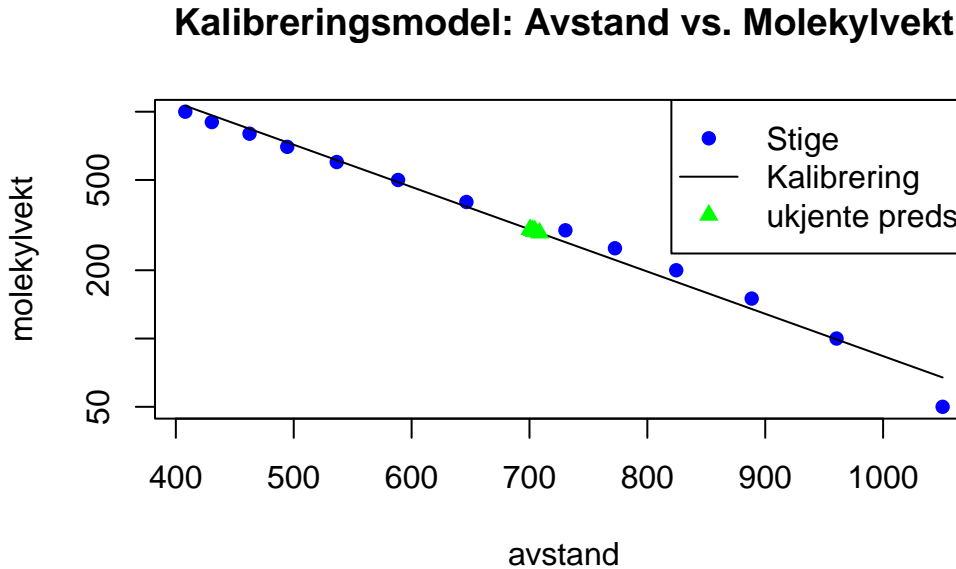
3.3 Resultat

3.3.1 Del 1: Predikert laktatterskel



Vi har kalkulert blodlaktatterskel ved 4mmol L^{-1} til å være 343 watt, og blodlaktatterskel ved 2mmol L^{-1} til å være 307 watt for subject “10” ved timepoint “pre”.

3.3.2 Del 2: Predikert hellning på qPCR kalibreringskurve



Kalibreringsmodellen beskriver forholdet mellom avstand og molekylvekt. Etter log-transformering av molekylvekten, blir forholdet mellom avstand og molekylvekt tilnærmet lineært, noe som gjør den lineære regresjonsmodellen passende. R-kvadratverdien på 0.98 indikerer hvor godt modellen passer til dataene da den er tilnærmet 1. De forutsagte molekylvektene for de ukjente prøvene er basert på den tilpassede kalibreringsmodellen. Predikasjonene for de ukjente prøvene gir et estimat av deres molekylvekt basert på deres migrasjonsavstand i gelen. QQ-plottet forteller oss om forskjellene mellom observert og predikert molekylvekt følger en normalfordeling. Ideelt sett bør punktene i QQ-plottet falle langs referanselinjen noe de tilsynelatende gjør. Det vil si at modellen er godt spesifisert, og fanger forholdet mellom migrasjonsavstand og molekylvekt.

3.3.3 Del 3: Tolke en regresjonstabell

Per økning i kilo SQUAT_3RM øker FAST_CSA_T1 med 5.483m^2 . Standardfeil (SE) = 8.032. SE forklarer gjennomsnittlig variasjon i økning i populasjonen tatt fra vårt utvalg. T-verdi = 0.683. T-verdien er relativt lav og indikerer at differansen mellom gjennomsnittet i dataen og gjennomsnittet i populasjonen sannsynligvis er liten. P-verdi = 0.50. P-verdien

her tilsier at det ikke er noe signifikant sammenheng mellom økning i kilo i SQUAT_3RM og økning i mikrometer² i FAST_CSA_T1.

3.4 Diskusjon

3.4.1 Del 1

Rapporten viser hvordan man kan kalkulere seg fram til wattbelastning på sykkel ved bestemte laktatverdier på 2 og 4 mmol L⁻¹. Dette er noe som beskriver intensiteten. Vi kom fram til at terskelwatten ved 2 mmol L⁻¹ var 307, og at terskelwatten ved 4mmol L⁻¹ var 343 for subject “10” ved timepoint “pre”.

3.4.2 Del 2

Kalibreringsmodellen viser seg å være effektiv til å beskrive forholdet mellom avstand og molekylvekt, spesielt etter at molekylvekten ble log-transformert. Denne transformasjonen bidro til å gjøre forholdet tilnærmet lineær, noe som bekrefter at en lineær regresjonsmodell er en passende tilnærming. Med en R-kvadratverdi på 0.98 ser vi at modellen gir en nesten perfekt tilpasning til dataene, som styrer modellens prediksjonsnøyaktighet. For de ukjente prøvene baseres prediksjonene av molekylvekt på modellen og gir troverdige estimater basert på migrasjonsavstand i gelen. QQ-plottet støtter også modellens robusthet, ettersom punktene faller nær referanselinjen, som indikerer at restene følger en normalfordeling. Dette antyder at modellen er godt spesifisert og gir en pålitelig beskrivelse av sammenhengen mellom migrasjonsavstand og molekylvekt. Samlet bekrefter resultatene at den utviklede modellen fanger opp de sentrale aspektene ved dataene på en god måte.

3.4.3 Del 3

Resultatene viser at det er en økning på 5.48 mikrometer² i FAST_CSA_T1 for hver kilo økning i SQUAT_3RM, men P-verdien på 0.50 indikerer at denne sammenhengen ikke er statistisk signifikant. Standardfeilen (SE) på 8.032 reflekterer variasjonen i estimatet og hvor mye det kan forventes å avvike fra populasjonsgjennomsnittet i utvalget vårt. Den lave t-verdien på 0.683 understøtter også at forskjellen mellom gjennomsnittet i utvalget og populasjonen sannsynligvis er liten. Oppsummert tyder disse funnene på at det ikke er en sterk eller signifikant sammenheng mellom økning i SQUAT_3RM og økning i FAST_CSA_T1. Samtidig som den observerte økningen i FAST_CSA_T1 heller kan være påvirket av tilfeldige variasjoner enn en systematisk effekt av økt styrke.

4 Rapport 3 - Statistisk inferens, statistiske modeller og statistisk styrke

4.1 Oppgave 1

Estimate, SE, t-verdi og p-verdi gir oss innsikt i gjennomsnittsforskjellen mellom behandlingene, hvor presist estimatet er, hvor ekstrem t-verdien er, og sannsynligheten for at resultatene er tilfeldige (p-verdi).

Estimate (Estimert verdi): m1 Estimate = 1.839727 m2 Estimate = 1.564161 Estimate (eller estimeringen) er den gjennomsnittlige verdien for den avhengige variabelen (y) i hver av de to modellene. Dette kan tolkes som forskjellen mellom de to behandlingene i studien. For modell m1 (med et mindre utvalg, $n = 8$), er den estimerte verdien for forskjellen mellom behandlingene 1.84. I modell m2 (med et større utvalg, $n = 40$) er den estimerte forskjellen litt lavere, på 1.56. Begge estimatene er positive, noe som indikerer at den ene behandlingen kan være bedre enn den andre i begge studiene.

SE (Standard Error/Standardfeil): m1 SE = 1.251293 m2 SE = 0.4774117 SE (standardfeilen) er et mål på variasjonen av estimatet, det vil si hvor presist vi kan estimere den gjennomsnittlige forskjellen mellom behandlingene. Standardfeilen for modell m1 er mye høyere (1.25) sammenlignet med modell m2 (0.47). Dette skyldes at modell m1 har et mye mindre utvalg ($n = 8$), som fører til større usikkerhet i estimatet. I modell m2, der vi har et større utvalg ($n = 40$), er standardfeilen lavere, noe som betyr at estimatet er mer presist.

T-verdi (t-statistikk): m1 T-verdi = 1.470261 m2 T-verdi = 3.276336 T-verdien er forholdet mellom estimatet (Estimate) og standardfeilen (SE), og den brukes til å avgjøre om det er statistisk signifikant forskjell mellom behandlingene. For modell m1 er t-verdien 1.47, som er relativt lav og viser at forskjellen mellom behandlingene ikke er veldig stor sammenlignet med den underliggende variasjonen. For modell m2 er t-verdien 3.27, som er mye høyere, og dette indikerer at vi har en sterkere indikasjon på at det er en faktisk forskjell mellom behandlingene når vi har et større utvalg.

P-verdi: m1 P-verdi = 0.1849546 m2 P-verdi = 0.002212965 P-verdien representerer sannsynligheten for å observere en t-verdi som er like ekstrem som den beregnede, gitt at nullhypotesen (ingen forskjell mellom behandlingene) er sann. For modell m1 er p-verdien 0.185, som betyr at det er en 18.5% sjanse for å observere en forskjell like stor eller større enn den vi har sett, selv om det ikke er noen reell forskjell. Siden p-verdien er mye større enn 0.05, kan vi ikke avvise

nullhypotesen, og vi kan ikke si med stor sikkerhet at det er en forskjell mellom behandlingene. For modell m2 er p-verdien derimot 0.0022, som er mye lavere enn 0.05. Dette betyr at det er en veldig liten sjanse (0.22%) for å observere en slik forskjell hvis nullhypotesen er sann. Her kan vi avvise nullhypotesen og konkludere med at det sannsynligvis er en reell forskjell mellom de to behandlingene.

4.2 Oppgave 2

Forskjellene mellom resultatene fra m1 og m2 skyldes hovedsakelig forskjellen i utvalgsstørrelse. Modell m1 har en liten utvalgsstørrelse ($n = 8$), noe som fører til større usikkerhet (høyere SE) og en svakere t-verdi. Med større usikkerhet blir det vanskeligere å trekke sikre konklusjoner, noe som gjenspeiles i den høyere p-verdien (0.185), som tyder på at det ikke er statistisk signifikante forskjeller mellom behandlingene.

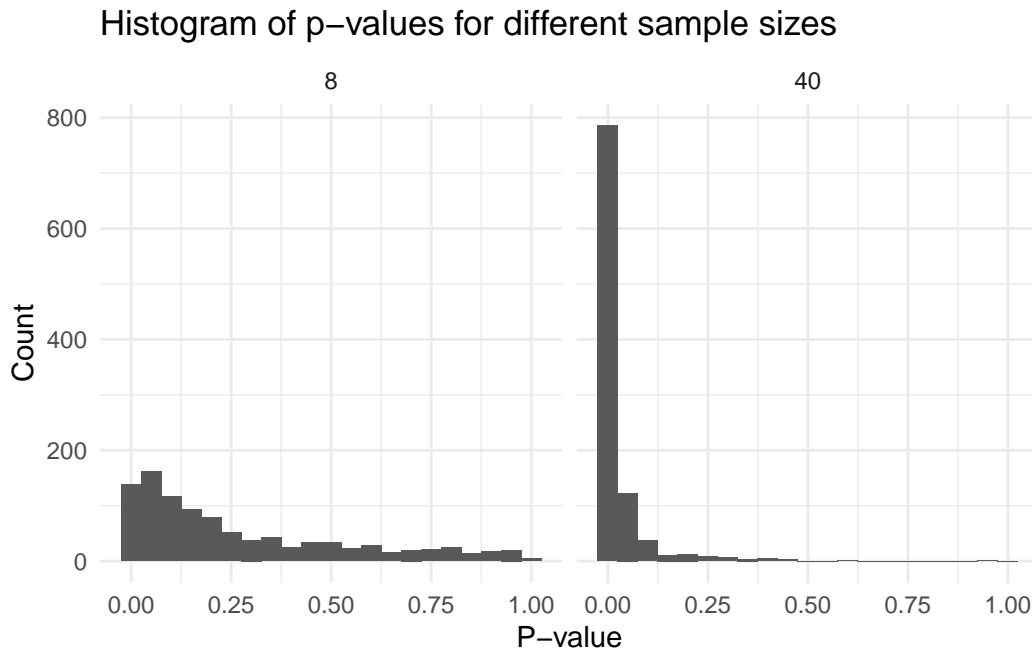
Modell m2 har en større utvalgsstørrelse ($n = 40$), noe som reduserer standardfeilen, øker t-verdien, og gir en lavere p-verdi (0.0022). Dette betyr at vi med større sikkerhet kan si at forskjellen mellom behandlingene er reell og statistisk signifikant. Med et større utvalg blir estimatene mer presise, og resultatene er mer pålitelige.

4.3 Oppgave 3

Det skyggelagte området i t-fordelingen representerer de mest ekstreme verdiene for t-statistikken, gitt at nullhypotesen er sann. Når vi utfører en t-test, er vi interessert i sannsynligheten for å få en t-verdi like ekstrem som den observerte (eller mer ekstrem) under antakelsen om at det ikke er noen reell forskjell mellom behandlingene (nullhypotesen).

Vi bruker både den nedre og øvre halvdelen av fordelingen (to-halet test), fordi vi ønsker å teste om forskjellen mellom behandlingene er signifikant i begge retninger – enten behandlingen har en positiv eller negativ effekt. Hvis t-verdien faller innenfor det skyggelagte området (dvs. at p-verdien er liten), kan vi konkludere med at forskjellen er statistisk signifikant, og at vi bør avvise nullhypotesen.

Med m1 har vi en t-verdi på 1.47 og en tilhørende p-verdi på 0.185, som betyr at vi ikke kan avvise nullhypotesen (ingen reell forskjell). I m2 derimot, gir en t-verdi på 3.28 en p-verdi på 0.0022, noe som gir en signifikant forskjell.



4.4 Oppgave 4

$N=8$ så er sd 1.07. Gjennomsnittlig SE er på 1.01. $N=40$ så er sd 0.474. Gjennomsnittlig SE er på 0.470. Standardavviket av estimatene og gjennomsnittet av standardfeilene (SE) er svært like fordi de begge måler hvor mye gjennomsnittsestimater varierer over mange studier. Standardavviket av estimatene fra 1000 simuleringer representerer den faktiske variasjonen mellom gjennomsnittene fra forskjellige studier, mens SE er den teoretiske forventningen om denne variasjonen. I praksis, når vi gjentar studien mange ganger (som vi gjør i simuleringene), vil den faktiske variasjonen i estimatene (standardavviket) tilnærme seg SE , som er en beregnet verdi for forventet variasjon.

Disse to verdiene er derfor svært like fordi SE er konstruert for å gi et anslag på hvor mye gjennomsnittsestimatene vil variere mellom forskjellige studier. SE beregnes som standardavviket i dataene delt på kvadratroten av gruppestørrelsen ($SE=SD/\sqrt{n}$), og som et mål på presisjonen i et gjennomsnittsestimat, reduseres SE når testgruppestørrelsen øker.

Når vi ser på forskjellen mellom gruppestørrelsen ($n = 8$ og $n = 40$), ser vi at både standardavviket til estimatene og SE reduseres når gruppestørrelsen øker. Dette skyldes at større utvalg gir mer nøyaktige estimater av gjennomsnittet og dermed mindre variasjon mellom ulike studier

4.5 Oppgave 5

$N = 8$: Histogrammet viser en bred fordeling av p-verdier, med en høyere konsentrasjon av p-verdier i den øvre delen av skalaen (nær 1). Dette tyder på at de fleste av testene som ble utført med denne gruppestørrelsen ikke er statistisk signifikante, med mange p-verdier som overstiger 0.05. Det er også en liten topp på venstre side av histogrammet, men det er relativt få p-verdier som er betydelig lave (mindre enn 0.05).

$N = 40$: Histogrammet viser en klar konsentrasjon av p-verdier som ligger mye nærmere null, noe som indikerer at flere av testene i denne gruppen er statistisk signifikante. Det er en høy topp ved 0.00, som tyder på at mange av testene har avvist nullhypotesen

Dette histogrammet demonstrerer tydelig effekten på statistisk styrke for den større utvalgsstørrelsen. I utvalgsstørrelsen ($n = 40$) ser vi en mye høyere andel p-verdier som indikerer statistisk signifikans (mindre enn 0.05), noe som betyr at vi har en større sjanse til å oppdage reelle effekter i dataene. I motsetning til dette, den mindre utvalgsstørrelsen ($n = 8$) resulterer i et lavere antall signifikante p-verdier, som kan indikere at vi har større usikkerhet og mindre presisjon i estimatene.

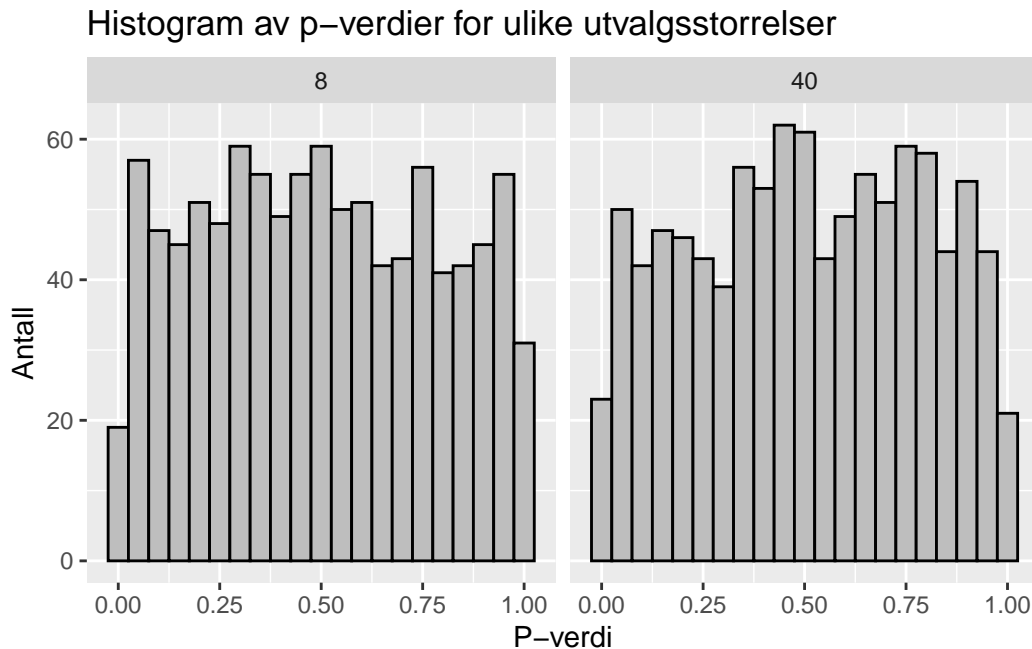
4.6 Oppgave 6

Basert på analysen av dataene, har vi beregnet antall studier fra hver utvalgsstørrelse ($n = 8$ og $n = 40$) som erklærer en statistisk signifikant effekt ved et signifikansnivå på 0.05. Resultatene viser at omtrent 24.3% av studiene med en utvalgssgruppe på 8 oppnådde signifikante p-verdier, mens hele 87.5% av studiene med utvalgssgruppe på 40 var signifikante. Dette indikerer at større testgrupper resulterer i høyere andel signifikante funn. Dette er i samsvar med forventningen om at større utvalg gir mer pålitelige estimater og øker den statistiske styrken, noe som gjør det lettere å oppdage reelle effekter i dataene. Dermed bekrefter disse resultatene viktigheten av tilstrekkelig med deltakere i statistiske analyser for å oppnå meningsfulle og signifikante funn.

4.7 Oppgave 7

Med en utvalgsstørrelse på 8, en effektstørrelse på 0.5 og et signifikansnivå på 0.05, viser beregningene at styrken er 23.2%. Dette indikerer at det er en lav sannsynlighet (23.2%) for å avdekke en faktisk effekt i studien, dersom effekten eksisterer. Denne lave styrken forklarer hvorfor så få av studiene med $n = 8$ oppnådde signifikante resultater i simuleringene (24.3%). Det illustrerer også risikoen for type II-feil ved små utvalgsstørrelser, der man ofte ikke klarer å fange opp reelle effekter.

Med en utvalgsstørrelse på 40 øker styrken til 86.9%, noe som gir en langt høyere sannsynlighet for å oppdage en faktisk effekt. Dette er i tråd med simuleringene, hvor 87.5% av studiene med $n = 40$ oppnådde signifikante resultater. Den betydelige økningen i statistisk styrke ved større utvalgsstørrelser viser at man med større utvalg får mer pålitelige resultater og reduserer risikoen for å overse ekte effekter (type II-feil)



4.8 Oppgave 8

Med et signifikansnivå på 5 % vil omtrent 5 % av studiene kunne gi et “falsk positivt” resultat, noe som betyr at vi avviser nullhypotesen selv om den faktisk er sann. Basert på resultatene fra simuleringen, ser vi at for studier med en utvalgsstørrelse på 8, var det 57 studier av 1000 som ga et falskt positivt resultat, noe som tilsvarer en andel på 5,7 %. For studier med en utvalgsstørrelse på 40, var det 56 falske positive resultater, noe som tilsvarer en andel på 5,6 %. Dette er svært nært det forventede nivået på 5 %, noe som bekrefter at selv uten en reell effekt vil omtrent 5 % av studiene vise statistisk signifikante resultater, kun på grunn av tilfeldigheter.

5 Rapport 4 - Studiedesign

5.1 Innledning

Treningsperiodisering er en systematisk treningsstrategi som blir benyttet for å optimalisere prestasjon og fysiologisk tilpasning. Dette skjer gjennom variasjon i intensitet, volum og belastning (B. R. Rønnestad, Hansen, and Ellefsen 2012). Periodisering er svært relevant i utholdenhetsidretter, der VO max ofte brukes som en indikator på aerob kapasitet og prestasjonsnivået til utøveren. Det blir benyttet ulike modeller som blokkbasert, tradisjonell og høyintensiv periodisering. Disse metodene har blitt utviklet for å maksimere treningsutbyttet.

De fem originale forskningsstudiene som undersøker effektene av periodisering er valgt for å representere et bredere spekter av ulike design og tilnærminger. Rapporten fokuserer på studiedesign og valgene av statistiske metoder. Dette er avgjørende for å tolke funnene på en valid og pålitelig måte (Hulley et al. 2013). Resultatene fra denne analysen er relevante for både forskere og trenere som ønsker å utvikle effektive treningsprogrammer for utholdenhetsutøvere.

5.2 Studiedesign

Valg av studiedesign er avgjørende for å besvare forskningsspørsmål på en pålitelig måte, og trekke gyldige konklusjoner. Generelt sett kan forskningsdesign deles inn i eksperimentelle og observasjonelle tilnærminger, hvor randomiserte kontrollerte studier (RCT) er gullstandarden for å etablere kausalitet. Dette er noe som gir en høy intern validitet, men slike studier kan ofte kreve betydelige ressurser og ha begrenset ekstern validitet. Longitudinelle studier gir en verdifull innsikt i hvordan variabler som VO max kan utvikle seg over tid, men de er sårbare for små utvalg og fraværet av randomisering.

Observasjonelle studier og case-studier gir dyp innsikt i spesifikke mønstre og fysiologiske tilpasninger. Selv om disse studiedesignene mangler kontrollgrupper og ikke kan etablere kausalitet, kan de fortsatt være svært nyttige for å lage hypoteser og skape forståelse for detaljerte mekanismer. Avanserte metoder, som for eksempel nær-infrarød spektroskopi (NIRS) og blandede modeller illustrerer hvordan ulike tilnærminger kan tilpasses til spesifikke forskningsspørsmål. (Hulley et al. 2013) understreker det at valg av studiedesign bør settes opp mot forskningsspørsmålets krav til intern validitet, generaliserbarhet og ressursbruk. Ved å kombinere styrkene fra ulike typer studiedesign kan forskningen gi en mer helhetlig forståelse

av forskningsfeltet. De fem studiene i denne rapporten inkluderer et spekter av design hvor hver tilnærming har sine styrker og svakheter. Valget av studiedesign påvirker både kvaliteten på dataene som bringes frem og hvordan resultatene kan generaliseres (Hulley et al. 2013).

5.3 Analyse av studiene

(Sellés-Pérez et al. 2024) er en longitudinell studie av triatleter hvor de fulgte tre kvinnelige U23-triatleter. Utøverne ble fulgt gjennom en sesong for å evaluere hvordan treningsbelastningen påvirket VO max og kroppssammensetning. Det ble benyttet lineær regresjon for å analysere datamaterialet. Ettersom at studien inneholder et lite utvalg og har en mangel på randomisering, så blir genrelaiserbarheten svekket. Til tross for dette gir studien en verdifull innsikt i disse type endringer blant unge utøvere.

(B. R. Rønnestad, Hansen, and Ellefsen 2012) er en RCT-studie av blokkbasert periodisering. Denne studien sammenlignet effekten av blokkbasert og tradisjonell periodisering på VO max og anaerob terskel. Studien besto av 21 godt trente syklister, og varte over 12 uker. Det ble brukt t-tester og ANOVA ble for å analysere forskjellene mellom grupper. Studiens validitet blir styrket av randomiseringen, men ettersom at utvalget fortsatt et mindre vil det begrense overførbarheten til andre populasjoner.

(Bent R. Rønnestad and Hansen 2018) er en langtids case-studie hvor de fulgte en elitesyklist over 58 uker for å nærmere undersøke effekten av blokkbasert trening. Det ble benyttet deskriptiv statistikk og trendanalyser for å evaluere endringene i VO max og laktatterskel. Styrken til studien er at den er såpass langsiktig, men mangelen på kontrollgrupper reduserer muligheten for å trekke kausale konklusjoner.

(Vaccari et al. 2022) sin studie sammenlignet ulike høyintensitetsprotokoller og hvilken innvirkning dette hadde på VO max. Det ble benyttet avanserte statistiske metoder, og blandede modeller som skapte en mer robust analyse av datasettet. Studien viser frem hvilken betydning individuelle tilpasninger har til høyintensiv trening. Mangelen på kontrollgrupper svekker de kausale konklusjonene.

(Hopker, O'Grady, and Pageaux 2016) gjennomførte en observasjonsstudie for å undersøke hvordan langvarig konstant belastning påvirker VO max og muskeleffektivitet. Studien brukte nær-infrarød spektroskopi (NIRS) for å måle muskeloksygenopptak under ulike treningsprotokoller. Forskningsspørsmålet fokuserte på hvordan muskulære og kardiovaskulære tilpasninger skjer over tid ved utholdenhetstrening. (NIRS) målingene for muskeloksygenopptak ga en detaljert innsikt i de fysiologiske tilpasningene til utholdenhetstrening. Fraværet av kontrollgrupper og en mindre utvalgsstørrelse begrenser den kasuale validiteten til studien.

5.4 Sammenligning av statistiske metoder

Studiene benyttet ulike statistiske metoder for å svare på forskningsspørsmålene. Dette var tilpasset studienes design og datagrunnlag. T-tester og ANOVA som ble brukt i (B. R. Rønnestad, Hansen, and Ellefsen 2012) er standardmetoder som benyttes for å evaluere forskjeller mellom ulike grupper. Disse metodene passer godt til å analysere effektene av intervensjonene som sammenligner de ulike treningsprotokollene. Dette er noe som forutsetter et normalfordelt datamateriale og homogene variansforhold. Det kan være utfordrende med mindre utvalg, men randomiseringen i studien styrket derimot validiteten. Metodene alene gir en begrenset innsikt i de individuelle variasjonene.

Lineær regresjon ble benyttet i (Sellés-Pérez et al. 2024). Dette er en effektiv metode for å analysere sammenhengene mellom treningsvolum og VO max. Denne metoden er derimot noe sårbar når det gjelder effekten av små utvalg. Dette kan påvirke presisjonen til estimatene og redusere generaliserbarheten av funnene i studien. I tillegg manglet studien randomisering, som kan øke risikoen for bias.

Deskriptiv statistikk og trendanalyser ble brukt i (Bent R. Rønnestad and Hansen 2018) og (Hopker, O’Grady, and Pageaux 2016), men med ulike formål. Disse metodene gir enkle og visuelle tilnærminger når det gjelder å identifisere langsiktige mønstre og fysiologiske tilpasninger. Selv om dette gir en verdifull innsikt i individuelle og overordnede mønstre, mangler metodene robusthet for å etablere kausale sammenhenger. Særlig for de studiene som mangler kontrollgrupper. Der (Bent R. Rønnestad and Hansen 2018) fokuserte på trendanalysene over 58 uker i en enkelt case-studie, kombinerte (Hopker, O’Grady, and Pageaux 2016) deskriptiv statistikk med avanserte teknologiske målinger. Begge studiene har et fravær av randomisering og kontrollgrupper, som svekker kausaliteten til begge studiene.

(Vaccari et al. 2022) benyttet blandede modeller. Dette gir en mer avansert statistisk tilnærming, hvor det gis større mulighet til å kontrollere variasjon mellom deltakere. Det gir grunnlag for å kunne analysere komplekse datasett med gjentatte målinger, som øker robustheten og validiteten av funnene i studien. Ettersom at det er et fravær av kontrollgrupper i studien svekker dette muligheten til å trekke sterke kausale konklusjoner.

5.5 Konklusjon

Studiene demonstrerer at forskjellige design og statistiske metoder har ulike styrker og svakheter. Longitudinelle design gir gjerne mer innsikt i endringer over tid, men kan lide av at uvalget er for lite. RCT-studier gir høy validitet, men kan ha begrenset generaliserbarhet. Case-studier og observasjonsstudier gir spesifikke detaljer, men kan ofte mangle kausale forklaringer. For fremtidig forskning anbefales det å kombinere styrkene i disse metodene. Dette kan være ved å bruke større utvalg, crossover-design og benytte avanserte statistiske og blandede modeller. Dette kan forbedre generaliserbarheten og kausaliteten til studiene, og samtidig gi en mer helhetlig forståelse av periodiseringens effekt på VO max.

6 Rapport 5 - Analyse av eksperimenter med repeterte målinger

6.1 Introduksjon

Styrketrening er en veletablert treningsform for å øke maksimal styrke og muskelmasse (Krieger 2009);(Schoenfeld, Ogborn, and Krieger 2016). Tidligere studier har vist at treningsvolum er en viktig faktor for å optimalisere resultatene for maksimal styrke og muskelmasse. En meta-analyse av (Krieger 2009) fant ut at det å gjennomføre flere sett gir bedre resultater enn det å gjennomføre ett sett, mens en studie av (Mitchell et al. 2012) viste det at ulik treningsbelastning ikke nødvendigvis førte til signifikante forskjeller i muskelhypertrofi.

Hvilken effekt man får av ulikt treningsvolum kan imidlertid variere mellom ulike individer. (Grgic et al. 2022) fant ut at det å trene til utmattelse opp mot ikke-utmattelse kan påvirke styrke og hypertrofi ulikt. Dette viser ytterligere kompleksiteten til treningsdesign og og resultater. Derfor er det viktig å ha god forståelse for hvilke treningsvolumer og treningsprogrammer som er optimale for ulike målgrupper.

Formålet til denne rapporten er å analysere hvordan lavt volum (1-sett) kontra moderat volum (3-sett) påvirket den maksimale styrken og muskelmassen til deltakerne som gjennomførte et 12 ukers treningsprogram. Basert på tidligere forskning som har blitt gjort på fagfeltet og en foreløpig analyse av den tilgjengelige dataen, så er hypotesen at moderat volum gir en større økning i både styrke og muskelmasse, sammenlignet med lavt volum.

6.2 Metoder

6.2.1 Deltakere og studiedesign

Studien til (Hammarström et al. 2020) inkluderte opprinnelig 40 deltakere i alderen 18–40 år. Inklusjonskriteriene var at deltakerne skulle være ikke-røykere og mellom 18-40 år. Eksklusjonskriteriene inkluderte intoleranse mot lokalbedøvelse, at deltakerne ikke skulle ha trent mer enn 1 styrkeøkt i uka de siste 12 månedene før interensjonsstart, nedsatt muskelstyrke på grunn av skader, og medikamenter som kunne påvirke tilpassninger som var assosiert med trening.

Av de 40 deltakerne, så ble 7 deltakere ekskludert underveis i dataanalysen fordi de ikke fullførte minst 85 % av de planlagte treningsøktene. Årsakene til at disse deltakerne ble ekskludert var: ubehag eller smerte i underekstremitetene under trening ($n = 5$), skade som ikke var relatert til studien ($n = 1$), og avvik fra studieprotokollen ($n = 1$). Dette resulterte i at den endelige analyse ble av 33 deltakere.

Alle deltakerne gjennomførte et 12-ukers styrketreningsprogram hvor hvert bein ble tildelt enten en 1-setts- eller 3-setts-protokoll. Muskelmassen til deltakerne ble målt ved DXA-skanning før og etter intervensjonen, og den maksimale styrken ble målt ved 1RM-tester for beinpress og leg extension.

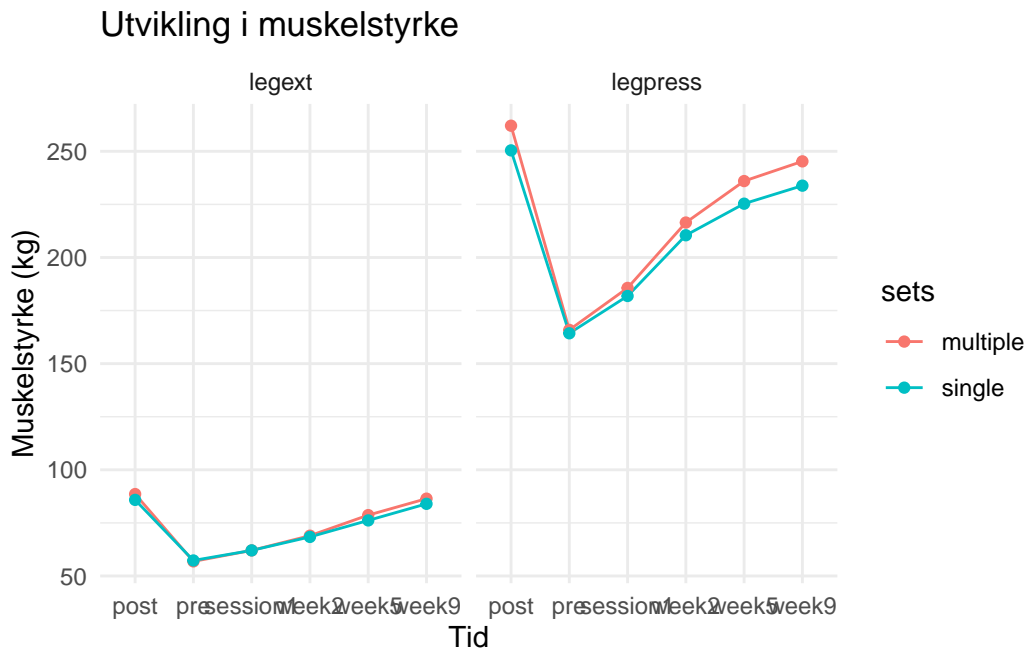
6.2.2 Tester og målinger

Muskelmassen til deltakerne ble målt ved bruk av Dual-energy X-ray Absorptiometry, også kjent som DXA-skanning. Under disse skanningene lå deltakerne på ryggen med armene langs siden. Målingene ble gjennomført før og etter intervensjonen. DXA-analysen estimerte “lean mass” for hvert bein. Endringene i muskelmassen til deltakerne ble beregnet ved å sammenligne resultatene fra de målingene før og etter intervensjonen.

Den maksimale styrken til deltakerne ble målt ved 1RM for beinpress og leg extension. Deltakerne gjennomførte en oppvarming som inneholdt 5–10 repetisjoner på en lav belastning. Deretter ble belastningen på øvelsen gradvis økt med 5–10 % av den estimerte 1RM til deltakeren. Deltakerne utførte deretter én repetisjon per belastning de fikk fram til de ikke klarte å gjennomføre en korrekt eller godkjent repetisjon. Mellom hvert forsøk fikk deltakeren 2–3 minutters hvile.

Hvert bein ble tilfeldig tildelt enten en 1-setts- eller 3-setts-protokoll: 1-sett-protokollen inneholdt én serie med 8–12 repetisjoner til utmattelse. 3-sett-protokollen inneholdt tre serier med 8–12 repetisjoner til utmattelse med 1–2 minutters hvile mellom seriene.

6.2.3 Analyse



6.3 Resultater

Resultatene fra undersøkelsen viste at 3-setts-protokollen viste en signifikant større økning i muskelmasse sammenlignet med 1-setts-protokollen. Den gjennomsnittlige forskjellen var på 122.8 g, 95 % CI: 8.6 til 237.0, $p = 0.0359$. Dette indikerer at et høyere treningsvolum vil gi bedre resultater i muskelmasse over en 12-ukers treningsperiode.

Når det gjaldt den maksimale styrken viste analysene at 3-setts-protokollen resulterte i større økninger i styrke sammenlignet med 1-sett. (Estimate = 289.06 g, $p < 0.05$). Økningene var mer markante i beinpress, og viste jevn progresjon for 3-setts-protokollen gjennom treningsperioden. Dette understøtter hypotesen om at et høyere treningsvolum vil gi bedre resultater for den maksimale styrken.

En korrelasjonsanalyse mellom muskelmasse og maksimal styrke viste en moderat, positiv sammenheng ($r = 0.45$, $p < 0.05$). Dette viser at økninger i muskelmasse vil bidra til økt styrke, men også at det er andre faktorer som vil påvirke. Disse funnene viser godt fram kompleksiteten i hvordan ulikt treningsvolum påvirker resultater for både muskelmasse og styrke. Det fremhever også betydningen av å vurdere flere faktorer når det gjelder utvikling innefor styrketrening.

6.4 Diskusjon

Resultatene som kommer frem fra denne analysen bekrefter hypotesen om at moderat volum (3-sett) gir større effekt på muskelmasse og maksimal styrke enn et lavt volum (1-sett). Dette samsvarer med funnene til (Krieger 2009), som viser at flere sett gir bedre resultater for styrke og hypertrofi. Dette er noe som også blir støttet av (Schoenfeld, Ogborn, and Krieger 2016), som viser en sammenheng mellom økt treningsvolum og effekt. (Grgic et al. 2022) forteller om at det kan være individuelle forskjeller, og at evnen til å trene til utmattelse er noe som kan påvirke resultatene.

Fremtidige studier bør gå nærmere inn på hvordan treningsvolum kan tilpasses ulike til populasjoner. For eksempel hvordan trening kan tilpasses for utrente eller eldre personer, slik at treningseffekten kan optimaliseres. Dette vil bidra til en bedre forståelse av hvordan treningsvolum kan påvirke ulike grupper og deretter hjelpe til med å utvikle skreddersydde treningsprogrammer for best mulig effekt.

6.5 Konklusjon

Denne studien bekrefter at et moderat volum (3-sett) er mer effektivt enn lavt volum (1-sett), både for å forbedre muskelmasse og maksimal styrke. Funnene tydeliggjør viktigheten av å inkludere flere sett i styrketreningen for å oppnå optimale treningsresultater. Det vil nok være nødvendig med mer forskning når det gjelder de individuelle forskjellene i respons ettersom at effekten av styrketrening kan være forskjellig.

7 Filosofihistorie

7.1 Spørsmål 1

Ifølge Hume er det umulig å rasjonelt begrunne bruken av induksjon. Hva er argumentet for denne konklusjonen? Gi en innvending mot ett av premissene i Humes argument og prøv å svare på denne innvendingen på Humes vegne.

Induksjon er en metode for å trekke generelle konklusjoner basert på tidligere observasjoner. Dette kan benyttes for å fortelle eller forutse noe som ikke har skjedd enda (Holmen 2024). Det kan gi induktiv sterke argumenter, som tilsier at hvis premissene er sanne, så vi konklusjonen sannsynligvis være sann, men ikke garantert være sann. Et klassisk eksempel på et slikt type argument er f.eks.: «Solen sto opp i morges, solen sto opp i går, solen sto opp dagen før i går, og derfor vil også solen stå opp i morgen» (Vassend 2024). Induksjon gir sannsynlige, men ikke sikre konklusjoner, fordi konklusjonene er basert på erfaring og ikke nødvendigvis tar hensyn til andre mulige tilfeller.

Dette er noe David Hume som en av flere vitenskapsfilosofer har stilt seg kritisk til. Hume argumenterer for at det ikke er mulig å rasjonelt eller logisk forsvare bruken av induksjon. På bakgrunn av at et induktivt argument forutsetter uniformitetsprinsippet. En påstand som kalles for uniformitetsprinsippet, er at «fremtiden vil være som fortiden». Man kan ikke bevise at UP stemmer kun gjennom logisk tankegang, ettersom at fremtiden lett kan være annerledes fra fortiden. At fremtiden vil være som fortiden er et sirkulært argument, og noe som Hume mener det ikke ligger noe rasjonelt bak (Vassend 2024).

Induksjon kan ikke begrunnes rasjonelt, og man kan ikke bevise argumentene kun gjennom ren logikk. Derfor kan man ikke empirisk sjekke at en påstand er sann. Ethvert forsøk på å bevise UP ender opp med å forutsette UP, noe som gjør argumentet sirkulært (Vassend 2024).

Innvending mot ett av premissene til Humes argument: UP vil være en forutsetning for erfaring. Hvis man ikke tror på at for eksempel naturlovene er forutsigbare, så kan man ikke tolke erfaringer eller gjøre vitenskapelige observasjoner. På denne innvendingen ville Hume svart at selv om UP ved ulike tilfeller kan være nødvendig for hvordan vi tolker verden rundt oss, vil det ikke si at det kan rettferdiggjøres logisk. Grunnen til at vi tror på UP er fordi vi pleier å se slike mønstre, ikke fordi vi har bevis på at dette er noe som alltid gjelder.

7.2 Spørsmål 2

Gi en kort beskrivelse av falsifikasjonisme og si litt om hvorfor Popper var motivert til å utvikle denne teorien. Presenter så ett problem med teorien og vurder hvorvidt problemet kan løses.

Falsifikasjonisme er et kjent begrep innenfor vitenskapsfilosofien, og noe som var meget sentralt hos synspunktene til Karl Popper (Alnes 2021). Ifølge Popper er en teori bare vitenskapelig dersom den er falsifiserbar, det vil si at den kan motbevises gjennom empirisk testing og deduktiv logikk. En vitenskapelig teori må derfor være formulert på en måte som gjør det mulig å teste den og potensielt vise at den er feil. Dette prinsippet skiller vitenskapelige teorier fra påstander som ikke kan testes eller utfordres (Popper 1959). Det er viktig å understreke at selv om et utsagn må være falsifiserbart for å regnes som vitenskapelig, betyr ikke dette at det nødvendigvis er usant.

Popper mente at falsifikasjon er den beste måten å drive vitenskap på. Popper mente at man utelukkende måtte basere vitenskap på deduktiv logikk. Bakgrunnen for at Popper var motivert til å utvikle denne teorien var på bakgrunn av «Hume's problem of induction» (Popper 1959). Induksjonsproblemet handler kort sagt om at man ikke kan begrunne induktive slutninger rasjonelt, fordi de bygger på antakelsen om at det som har skjedd tidligere, også vil skje i fremtiden. Denne antakelsen er kjent som uniformitetsprinsippet.

Popper mente at problemet med induksjon kunne løses ved å vektlegge falsifisering fremfor verifikasjon. Han hevdet at en teori er vitenskapelig dersom den kan motbevises gjennom systematiske tester. Denne tilnærmingen unngår behovet for induktive slutninger og baserer seg i stedet på deduktiv logikk, som ikke har de samme utfordringene (Popper 1959).

Teorien kan by på flere problemer, og et av disse problemene er at tilnærmingen til teorien er ganske direkte: «en teori bare vitenskapelig dersom den er falsifiserbar». Dette vil bli problematisk ettersom at noen sannsynlighetspåstander ikke er falsifiserbare.

Dette skaper utfordringer, ettersom noen sannsynlighetspåstander ikke kan falsifiseres. Dette skyldes hvordan de er formulert og hvilken kontekst de befinner seg i. Påstander som er generelle og ikke knyttet til spesifikke empiriske observasjoner, er ofte vanskeligere å falsifisere. For eksempel kan en påstand som “det er sannsynlig at det vil regne i morgen” være vanskelig å motbevise, fordi den baserer seg på en sannsynlighet snarere enn konkrete observasjoner. Derimot, når sannsynlighetspåstander knyttes til spesifikke hendelser eller observasjoner, blir de ofte enklere å teste og dermed mer falsifiserbare.

Et problem med falsifikasjonisme er at det i praksis kan være vanskelig å konkludere med at en teori er helt falsifisert. Dette skjer på bakgrunn av at forskere ofte kan justere teorier ved å legge til hjelpehypoteser eller andre forutsetninger, som gjør at det blir mulig å unngå falsifikasjon. Popper anerkjenner dette som en utfordring, men argumenterer for at vitenskapens styrke ligger metodene for systematisk testing av teorier med streng kritikk.

Men kan problemet løses? Det viser seg at både induksjon og deduksjon ikke er helt feilfrie metoder innenfor vitenskapen. For å løse dette problemet kan man for eksempel ha en bayesiansk tilnærming. Bayesianisme er et forsøk innenfor vitenskapsfilosofien på å benytte den matematiske sannsynlighetsteorien til å forstå den vitenskapelige prosessen, med inkludering av induksjon, falsifisering, bekreftelse, osv. (Vassend 2024). En bayesiansk tilnærming kan bidra til å adressere problemet ved å introdusere en ramme for hvordan vi oppdaterer sannsynligheten basert på nye data, i stedet for å søke etter absolutte sannheter. En slik tilnærming kan gi oss en mer fleksibel vei til kunnskap, hvor sannsynligheter justeres over tid.

8 Laboratorierapport fra molekylærlabb

8.1 qPCR

Formålet med denne rapporten er å sjekke primereffektivitet og gjennomføre målrettet amplifisering av cDNA ved bruk av spesifikke primere. Å analysere genuttrykk ved bruk av qPCR er en metode som blir benyttet mye for å undersøke hvordan gener responderer på fysiologiske stimuli som trening. I denne undersøkelsen så vi nærmere å hvordan styrketrening påvirker genuttrykket i ulike muskelfibertyper.

8.2 Utstyr

En sanntids PCR-maskin (vi bruker QuantStudio 5), en qPCR-reaksjonsplate, nukleasefritt vann og pipettespisser SYBR-grønn Master mix

8.3 Metoder

Fikk ferdig cDNA prøver fra fryser av personer som har gjennomført systematisk styrketrening i 2 uker.

Vi fortynnet prøvene med qPCR:

Warning: package 'kableExtra' was built under R version 4.4.2

1. Flytter 2 μ l fra rør 1 til 2a, og 10 μ l fra 1 til 2b, vortex rør 2a+2b.
2. Flytter 2 μ l fra 2a til 3a og 2 μ l fra 2b til 3b, vortex rør 3a+3b.
3. Flytter 2 μ l fra 3a til 4a og 2 μ l fra 3b til 4b, vortex rør 4a + 4b.

Mulig avvik: første: litt lite innhold i 3b ifh til de andre

1. Kombiner en mesterblanding:

Table 8.1: Fortynningsdata

Ror	Fortynning	Prøve_volum	H2O_volum
1	1	30µl	0µl
2a	1/10	2µl	18µl
3a	1/100	2µl	18µl
4a	1/1000	2µl	18µl
2b	1/2	10µl	10µl
3b	1/20	2µl	18µl
4b	1/200	2µl	18µl

Table 8.2: Komponentdata for qPCR

Komponent	Volum_per_reaksjon	Total_volum	cmyc
Sybr-green 2X master-mix	5 l	50 l	125 l
Primermix (Forward and Reverse, 5 M each)	1 l	10 l	25 l
H2O	2 l	20 l	50 l

Mulig avvik i første steg: sybr-grønn: tok 200 + 50, hang igjen dråper i pipettespiss på begge to

Andre steg: vond pipette for sybr-grønn 50ml

2. Fyll platen med 8 l primerspesifikk masterblanding
3. Tilsett 2 l cDNA-prøve.

8.4 Pipetteringsskjema

Table 8.3: Pipetteringsskjema

Rad	Kol_13	Kol_14	Kol_15	Kol_16	Kol_17	Kol_18	Kol_19	Kol_20	Kol_21	Kol_22	Kol_23
EN	Fp1	Fp2			cmyc 1	cmyc 2a	cmyc 3a	cmyc 4a	cmyc 2b	cmyc 3b	cmyc 4b
B	myhc 1	myhc 1			cmyc 1	cmyc 2a	cmyc 3a	cmyc 4a	cmyc 2b	cmyc 3b	cmyc 4b
C	myhc 1	myhc 1			cmyc 1	cmyc 2a	cmyc 3a	cmyc 4a	cmyc 2b	cmyc 3b	cmyc 4b
D	myhc 2a	myhc 2a									
E	myhc 2a	myhc 2a									
F	myhc 2a	myhc 2a									
G	myhc 2x	myhc 2x									
H	myhc 2x	myhc 2x									
I	myhc 2x	myhc 2x									
J	myhc	myhc			cmyc 1	cmyc 2a	cmyc 3a	cmyc 4a	cmyc 2b	cmyc 3b	cmyc 4b
K	myhc	myhc			cmyc 1	cmyc 2a	cmyc 3a	cmyc 4a	cmyc 2b	cmyc 3b	cmyc 4b
L	myhc	myhc			cmyc 1	cmyc 2a	cmyc 3a	cmyc 4a	cmyc 2b	cmyc 3b	cmyc 4b

MHC Delta-delta-CT method: relative quantification

Gene fam normalization 2^{-Ct}

Table 8.4: qPCR Resultater

Sample_name	Target_name	Ct1	Ct2	Ct3	Average	Reference_gene	Ct1_ref	Ct2_ref	Ct3_ref	Average_ref	Delta_Ct	Delta_Delta_Ct	Two_Delta_Delta_Ct
FP6 R w0	myhc1	19.798	19.901	19.621	19.77333	b2m	24.670	24.513	24.691	24.625	-4.851	NA	28.867
FP6 R w2pre	myhc1	18.944	19.240	19.861	19.34833	b2m	22.913	23.950	23.819	23.561	-4.212	NA	18.536
FP6 R w0	mhc2a	21.029	21.247	20.627	20.96800	b2m	24.670	24.513	24.691	24.625	-3.657	NA	12.613
FP6 R w2pre	mhc2a	19.549	19.304	19.580	19.47800	b2m	22.913	23.950	23.819	23.561	-4.083	NA	16.944
FP6 R w0	myhc2x	27.019	26.898	25.907	26.60795	b2m	24.670	24.513	24.691	24.625	1.983	NA	0.253
FP6 R w2pre	myhc2x	24.871	24.105	24.256	24.41062	b2m	22.913	23.950	23.819	23.561	0.850	NA	0.555

Table 8.5: Prosentfordeling for myhc1, myhc2a og mhc2x

X100..	myhc1	myhc2a	mhc2x
41.733	69,17 %	30,22 %	0,61 %
36.034	51,44 %	47,02 %	1,54 %

Etter en to uker lang styrketreningsintervensjon observeres en endring i genuttrykket og fordeling for myhc1, myhc2a og myhc2x. Ved pre-testen var fordelingen henholdsvis 69,17 % for myhc1, 30,22 % for myhc2a og 0,61 % for myhc2x. Etter intervensjonen, ved post-testen, endret fordelingen seg til 51,44 % for myhc1, 47,02 % for myhc2a og 1,54 % for myhc2x. Ct verdiene indikerer en økning av genuttrykket av myhc1, myhc2a og myhc2x.

Table 8.6: Fortynningsrekke

Fortynning	Ct1	Ct2	Ct3	Avg_Ct	Sample_quan	Log_sample_quan	Slope_stigningstall	Primer_Efficiency
1	28,678	28,708	29,155	28,847	1.000	0.000	-2,6104	141,5901
1/2	29,414	29,620	29,264	29,433	0.500	-0.301		
1/10	31,776	31,416	32,413	31,868	0.100	-1.000		
1/20	33,241	32,653	Undetermined	32,947	0.050	-1.301		
1/200	Undetermined	Undetermined	34,574	34,574	0.005	-2.301		

Stigningstallet på -2,6 indikerer at primer effektiviteten er ikke-optimal da den er kalkulert til 142%. Vi ser at observasjonene ikke er lineære (Figur 8.1)

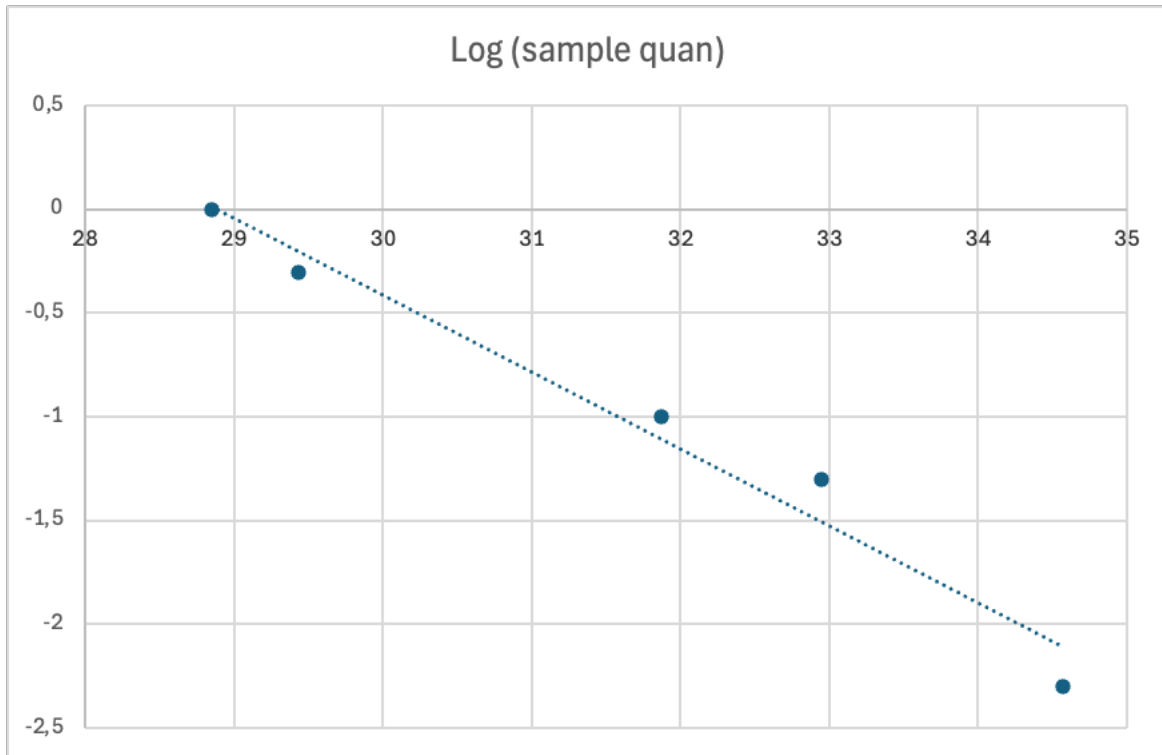


Figure 8.1: Ffiguren viser en negativ sammenheng mellom gjennomsnittlige Ct-verdier og log (sample quantity).

Table 8.7: Pooled Sample Data

Pooled.sam	Avg	SD	CV_percent
1.0	28.847	0.2670490	0,93 %
0.5	29.433	0.1788764	0,61 %
0.1	31.868	0.5050010	1,58 %

Standardavvikene er relativt lave for fortynningene i “pooled sample” som tyder på presisjon i resultatene. Standardfeilen er lavest for 1/2 fortynnet prøve ved 0,61% og høyest for 1/10 fortynnet prøve ved 1,58%. Det indikerer at det er større variabilitet mellom målingene for prøvene som er mest fortynnet. Standardfeilen til den ufortynnede prøven er høyere enn den fortynnede 1/2.

8.5 Diskusjon

(Wilborn and Willoughby 2004) sa at etter 8 uker med tung styrketrening var genuttrykket for myhc 1 og 2a økt, og mindre for mych 2x. Det er ikke studier på to uker med styrketrening som forklarer endringen i myhc genuttrykk, men vi spekulerer i at det vil være samme endringer som for (Wilborn and Willoughby 2004), men i mindre grad. Dette stemmer med våre funn da vi ser en økning i genuttrykk for både mych 1 og mych 2a, men det strider i mot at vi ser en økning i mych 2x.

Primer effektiviteten er ikke-optimal da den er kalkulert til 142%. Referanseverdiene på 90-110% indikerer at hver PCR syklus doubler mengden mål DNA nøyaktig. Vår høye verdi tyder på ulike feilkilder underveis i forsøket. Som feil pipettering eller kontaminasjon. Primere kan binde seg til flere plasser på DNA, og kan også binde seg til hverandre, noe som kan gi ut falskt signal.

CV målingene forteller at det er høyere variabilitet i målingene når prøven blir mer fortynnet. Gjennomsnittet er lavest på den ufortynnede prøven. Dette er noe som kan påvirke at CV er høyere enn den 1/2 fortynnede prøven da CV regnes ut fra standardavviket delt på gjennomsnittet, som gjør at et lavere gjennomsnitt vil gi en høyere CV. Vi mistenker at dårlig pipitering er årsak til at høyere CV for ufortynnet prøve enn 1/2 fortynnet prøve.

References

- Alnes, Jostein Hellesnes. 2021. “Falsifikasjon (vitenskapsteori).” https://snl.no/falsifikasjon_-_vitenskapsteori.
- Grgic, Jozo, Brad J. Schoenfeld, John Orazem, and Filip Sabol. 2022. “Effects of Resistance Training Performed to Repetition Failure or Non-Failure on Muscular Strength and Hypertrophy: A Systematic Review and Meta-Analysis.” *Journal of Sport and Health Science* 11 (2): 202–11. <https://doi.org/10.1016/j.jshs.2021.01.007>.
- 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/ijsp.2014-0566>.
- Hammarström, Daniel, Sjur Øfsteng, Lise Koll, Marita Hanestadhaugen, Ivana Hollan, William Apró, Jon Elling Whist, Eva Blomstrand, Bent R. Rønnestad, and Stian Ellefsen. 2020. “Benefits of Higher Resistance-Training Volume Are Related to Ribosome Biogenesis.” *The Journal of Physiology* 598 (3): 543–65. <https://doi.org/10.1113/jp278455>.
- Haun, Cody T., Christopher G. Vann, C. Brooks Mobley, Shelby C. Osburn, Petey W. Mumford, Paul A. Roberson, Matthew A. Romero, et al. 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.” *Frontiers in Physiology* 10 (March). <https://doi.org/10.3389/fphys.2019.00297>.
- Haun, Cody T., Christopher G. Vann, Christopher B. Mobley, Paul A. Roberson, Shelby C. Osburn, Hudson M. Holmes, Petey M. Mumford, et al. 2018. “Effects of Graded Whey Supplementation During Extreme-Volume Resistance Training.” *Frontiers in Nutrition* 5 (September). <https://doi.org/10.3389/fnut.2018.00084>.
- Holmen, Håvard A. 2024. “Induksjonsproblemet.” <https://snl.no/induksjonsproblemet>.
- Hopker, J. G., C. O’Grady, and B. Pageaux. 2016. “Prolonged Constant Load Cycling Exercise Is Associated with Reduced Gross Efficiency and Increased Muscle Oxygen Uptake.” *Scandinavian Journal of Medicine & Science in Sports* 27 (4): 408–17. <https://doi.org/10.1111/sms.12673>.
- Hopkins, Will G. 2000. “Measures of Reliability in Sports Medicine and Science.” *Sports Medicine* 30 (1): 1–15. <https://doi.org/10.2165/00007256-200030010-00001>.
- Hulley, Stephen B., Steven R. Cummings, Warren S. Browner, Deborah G. Grady, and Thomas B. Newman. 2013. *Designing Clinical Research*. 4th ed. Lippincott Williams & Wilkins. https://tilda.tcd.ie/epidemiology-biostatistics-course/course-material/assets/Class2/Designingclinicalresearch_4th-edition.pdf.

- 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>.
- Mitchell, Cameron J., Tyler A. Churchward-Venne, Daniel W. D. West, Nicholas A. Burd, Leigh Breen, Steven K. Baker, and Stuart M. Phillips. 2012. "Resistance Exercise Load Does Not Determine Training-Mediated Hypertrophic Gains in Young Men." *Journal of Applied Physiology* 113 (1): 71–77. <https://doi.org/10.1152/jappphysiol.00307.2012>.
- Popper, Karl. 1959. "The Logic of Scientific Discovery." 1959. https://edisciplinas.usp.br/pluginfile.php/7833283/mod_resource/content/1/%5BKarl_Popper%5D_The_logic_of_scientific_discovery%28z-lib.org%29.pdf.
- Rønnestad, B. R., J. Hansen, and S. Ellefsen. 2012. "Block Periodization of High-Intensity Aerobic Intervals Provides Superior Training Effects in Trained Cyclists." *Scandinavian Journal of Medicine & Science in Sports* 24 (1): 34–42. <https://doi.org/10.1111/j.1600-0838.2012.01485.x>.
- Rønnestad, Bent R., and Joar Hansen. 2018. "A Scientific Approach to Improve Physiological Capacity of an Elite Cyclist." *International Journal of Sports Physiology and Performance* 13 (3): 390–93. <https://doi.org/10.1123/ijsp.2017-0228>.
- Schindelin, Johannes, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, et al. 2012. "Fiji: An Open-Source Platform for Biological-Image Analysis." *Nature Methods* 9 (7): 676–82. <https://doi.org/10.1038/nmeth.2019>.
- Schoenfeld, Brad J., Dan Ogborn, and James W. Krieger. 2016. "Dose-Response Relationship Between Weekly Resistance Training Volume and Increases in Muscle Mass: A Systematic Review and Meta-Analysis." *Journal of Sports Sciences* 35 (11): 1073–82. <https://doi.org/10.1080/02640414.2016.1210197>.
- Sellés-Pérez, Sergio, Hector Arévalo-Chico, José Fernández-Sáez, and Roberto Cejuela. 2024. "Training Characteristics, Performance, and Body Composition of Three U23 Elite Female Triathletes Throughout a Season." *Sports* 12 (2): 53. <https://doi.org/10.3390/sports12020053>.
- SYLTA, ØYSTEIN, ESPEN TØNNESEN, DANIEL HAMMARSTRÖM, JØRGEN DANIELSEN, KNUT SKOVERENG, TROELS RAVN, BENT R. RØNNESTAD, ØYVIND SANDBAKK, and STEPHEN SEILER. 2016. "The Effect of Different High-Intensity Periodization Models on Endurance Adaptations." *Medicine & Science in Sports & Exercise* 48 (11): 2165–74. <https://doi.org/10.1249/mss.0000000000001007>.
- Vaccari, Filippo, Jacopo Stafuzza, Nicola Giovanelli, and Stefano Lazzer. 2022. "High-Intensity Interval Training: Optimizing Oxygen Consumption and Time to Exhaustion Taking Advantage of the Exponential Reconstitution Behaviour of D." *European Journal of Applied Physiology* 123 (1): 201–9. <https://doi.org/10.1007/s00421-022-05059-2>.
- Vassend, Olav. 2024. "Vitenskapsfilosofi dag 1-4."
- Wilborn, Colin D, and Darryn S Willoughby. 2004. "The Role of Dietary Protein Intake and Resistance Training on Myosin Heavy Chain Expression." *Journal of the International Society of Sports Nutrition* 1 (2). <https://doi.org/10.1186/1550-2783-1-2-27>.