



Stockholms
universitet

Statistisk analys av bestämningar av maximal syreupptagningsförmåga

Tuva Andersson

Kandidatuppsats 2014:9
Matematisk statistik
September 2014

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Statistisk analys av bestämningar av maximal syreupptagningsförmåga

Tuva Andersson*

September 2014

Sammanfattning

Maximal syreupptagningsförmåga ($VO_2\max$), i dagligt tal kallad kondition, är en viktig prediktor för fysisk prestationsförmåga och hjärthälsa. I många sammanhang är det inte möjligt att mäta $VO_2\max$ eftersom det kräver avancerad laboratorietrustning och maximal ansträngning. Elin Ekblom-Bak är en forskare vid Gymnastik- och idrottshögskolan i Stockholm, som har utvecklat en ny metod för att med testcykel skatta maximal syreupptagningsförmåga. I denna uppsats så undersöker vi Ekblom-Baks regressionsmodell för att skatta $VO_2\max$ och föreslår möjliga förändringar. Ett av de tidigare mest använda testen för att skatta $VO_2\max$ är Åstrandets test. I de fall försökspersonerna som deltagit i utformningen av Ekblom-Baks test också har gjort ett Åstrandstest så jämför vi både testens prediktiva förmåga. Åstrandstestet ser ut att behöva justeras för de olika könen så vi modifierar Åstrandstestet genom att lägga till en indikatorvariabel för om deltagaren är en man eller kvinna. Efter modifiering så finner vi, för de deltagare som utfört bägge testen, att Ekblom-Baks test och Åstrandets test har liknande prediktiv förmåga.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: tuvand@gmail.com. Handledare: Rolf Sundberg.

Abstract

Maximal oxygen uptake ($VO_2\text{max}$) is an important predictor of aerobic capacity and cardiovascular health. It is in many situations not possible to measure $VO_2\text{max}$ since it requires maximal effort and advanced laboratory equipment. Elin Ekblom-Bak, a researcher at The Swedish School of Sport and Health Sciences (GIH) in Stockholm, has developed a new cycle ergometer test to estimate $VO_2\text{max}$. In this paper, we will investigate Ekblom-Bak's regression model to estimate $VO_2\text{max}$ and suggest possible changes.

One of the previously most widely used tests to estimate $VO_2\text{max}$ is Åstrand's test. In those cases where the subjects who participated in the design of Ekblom-Bak's test also did a Åstrand's test we compare both test's predictive ability. There's indications that Åstrand's test needs to be adjusted for gender. We modify Åstrand's test by adding a variable indicating if the subject is a man or a women. After modification, we find that Åstrand's test and Ekblom-Bak's test have similar predictive ability for those participants who completed both tests.

Förord

Jag vill tacka min handledare, Rolf Sundberg, professor i matematisk statistik vid Matematiska institutionen i Stockholm, som visat stort engagemang och hela tiden hjälpt mig att komma vidare i arbetet med denna uppsats. Vidare vill jag även rikta ett stort tack till Elin Ekblom-Bak, doktor i medicinsk vetenskap vid Gymnastik- och idrottshögskolan i Stockholm, för att du har gjort detta arbete möjligt genom att ta dig tid att svara på både fysiologiska och statistiska frågor under arbetets gång. Till sist vill jag också tacka min sambo för visat tålamod och stöd.

Innehåll

1	Inledning	1
1.1	Bakgrund	1
1.1.1	Åstrands test	1
1.1.2	Eklom-Baks test	3
1.2	Syfte	3
1.3	Analysmetod	3
2	Teori - Multipel linjär regression	4
2.1	Definition och skattning av modellens parametrar	4
2.2	Test av linjära hypoteser	5
2.3	Anpassningsmått och prediktionsmått	6
2.4	Residualer och inflytelserika observationer	7
2.5	Grafisk diagnostik för regressionsmodeller	8
2.5.1	Residual plot	8
2.5.2	Added variable plot	8
2.5.3	C+R plot	9
2.6	Indikatorvariabler	9
3	Beskrivning av datamaterialet	10
3.1	Responsvariabel - $VO_2\max$	10
3.2	De olika studierna	10
3.3	Förklarande variabler	11
3.4	Scatterplot och boxplot av datamaterialet	12
4	Statistisk analys	14
4.1	Studie 1 - Eklom-Baks regressionsmodell	14
4.1.1	Jämförelse med Åstrands test	17
4.2	Studie 2 - Logaritmering av responsen	19
4.3	Studie 3 - Transformation av regressor?	20
4.4	Studie 4 - Inkludering av alla försökspersoner	22
4.4.1	Korrigerig för annan mätmetod av $VO_2\max$	22
4.4.2	Vikt, längd och andra utförda test	25
5	Diskussion	26
6	Appendix	29

1 Inledning

Maximal syreupptagningsförmåga ($VO_2\text{max}$), i dagligt tal kallad kondition, är viktig för prestationen för såväl idrottare som motionärer, men också för att klara av vardagliga sysslor. En låg kondition har visat sig vara en riskfaktor för olika sjukdomar, till exempel diabetes och hjärt-kärlsjukdomar, medan en bra kondition verkar skyddande. Därför är det i olika sammanhang bra att kunna mäta en persons $VO_2\text{max}$.

Arbetande muskler behöver syre för att kunna arbeta under längre tidsperioder. När arbetsbelastningen ökar, ökar syreupptagningen genom att hjärtat slår snabbare, mer blod pumpas ut vid varje slag, samt att syreupptaget i de arbetande musklerna blir större och mer effektivt. $VO_2\text{max}$ är framför allt ett mått på hjärtats maximala kapacitet att via kärlen och blodet skicka ut syre till de arbetande musklerna, och på hur bra musklerna sedan är att tillgodogöra sig det syre de får. $VO_2\text{max}$ mäts vid maximalt arbete med avancerad utrustning i ett arbetsfysiologiskt laboratorium. [2, s. 33]

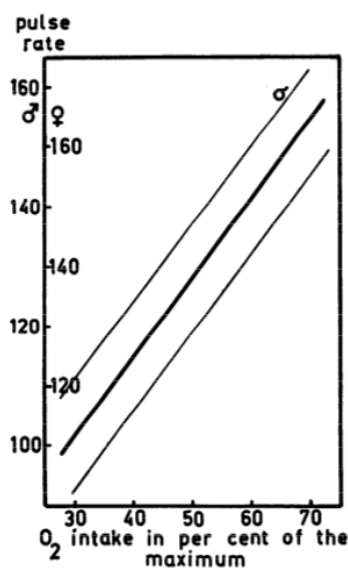
1.1 Bakgrund

I många sammanhang så är det inte möjligt att mäta en persons $VO_2\text{max}$. Att utföra maximalt arbete kan innebära en hälsorisk och/eller så finns inte avancerad laboratorieutrustning att tillgå. Därför har flera test utvecklats där man utifrån submaximalt arbete försöker skatta individens $VO_2\text{max}$. Ett av de mest använda, både i Sverige och utomlands, är Åstrands test. Det baseras liksom Ekblom-Baks test, på submaximalt arbete utfört på cykel. [2, s. 33] Cykling har visat sig vara en lämplig arbetsform därför att en viss belastning (ej maximal) kräver ungefär samma energiutveckling hos cyklisten oavsett om det är en kvinna eller man ung, gammal, tränad, otränad, elitcyklist eller ovan vid cykling [11].

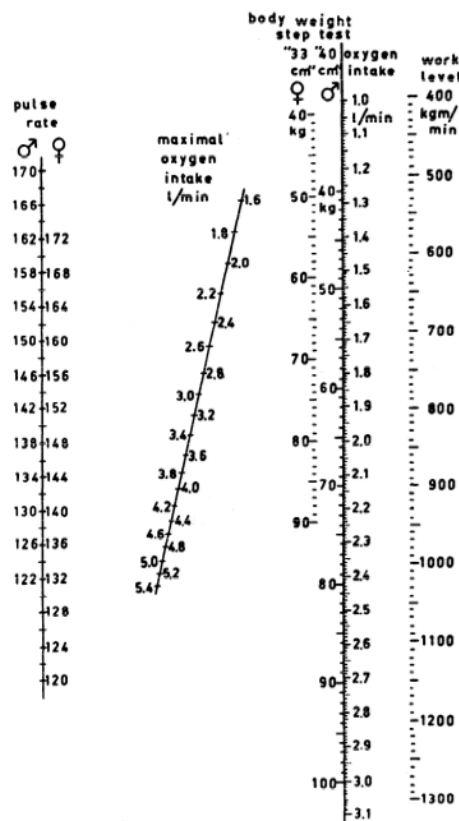
1.1.1 Åstrands test

Åstrands nomogram (Figur 1) utarbetades av Irma och P-O Åstrand i slutet av 1950-talet på Gymnastiska centralinstitutet (nuvarande Gymnastik- och idrottshögskolan) i Stockholm och baseras bland annat på studier som visat att

- Förhållandet mellan puls och en persons syreupptagning i % av $VO_2\text{max}$ under submaximalt arbete kan antas vara linjärt.
- Det går att skatta en persons syreupptagning under arbete utifrån arbetets belastning (effekt).



(a) Relation mellan puls, under arbete, och faktiskt syreupptagning i % av VO_2 max. Den tunna linjen representerar standardavvikelse. (Baserad på resultat från en studie med 16 kvinnor och 17 män i åldern 20-30 år.)



(b) Åstrands nomogram

Figur 1: Bilder tagna från Irma och P-O Åstrands artikel; *A Nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work*, 1954. [12]

Värt att notera är att nomogrammet är baserat på studier av friska, unga, vältränade personer, eller som skaparna själva skriver: *It should be emphasized that the nomogram is based from experiments with healthy subjects 18-30 years of age.* [12, s. 221]

Personen som genomför testet cyklar i sex minuter, på en individuellt vald belastning, med en tramphastighet på 50 varv per minut. Arbetet ska upplevas något ansträngande, ca 50-70 procent av maximal kapacitet. Utifrån kön, ålder, hur hög pulsen är under den sista minuten samt vilken belastning som deltagarna cyklar på så skattas VO_2 max sedan utifrån nomogrammet. Otaliga studier har undersökt noggrannheten och precisionen av Åstrands nomogram och på gruppnivå stämmer medelvärden av beräknat VO_2 max bra med uppmätt VO_2 max, medan det på individnivå finns stora variationer. Testet ska också gärna utföras vid upprepade tillfällen då hjärtfrekvensen vid första testtillfället kan vara påverkad av yttre faktorer såsom stress och nervositet. Men i många sammanhang, till exempel vid hälsokontroller och

screening, är det bara möjligt att genomföra testet vid ett enskilt testtillfälle. [2, s. 33-34]

1.1.2 Ekblom-Baks test

Elin Ekblom-Bak är en forskare på Gymnastik- och idrottshögskolan som har utvecklat ett nytt test för att med träningscykel bestämma $VO_2\max$. Ekblom-Baks test bygger på samma grundtanke som Åstrandstestet, det vill säga att var enkelt tidseffektivt, enbart innebära en låg risk för individen samt att det bara behövs en vanlig stationär cykel och ett pulsband. Skillnaden är att testpersonen cyklar med en trampfrekvens på 60 varv per minut på två olika belastningar, istället för bara en. Först fyra minuter på en lägre standardiserad belastning, som alla genomför, följt av fyra minuter på en högre belastning som bestämts i samspråk med testledaren för att uppnå en nivå som testpersonen upplever som något ansträngande (cirka 50-70 % av maximal kapacitet eller cirka 14 på den så kallade Borgskalan). Fyra minuter anses vara tillräckligt för att nå *steady-state*, dvs att syreupptagningen i lungorna motsvarar vävnadernas behov. Det nya testet bygger på den arbetsfysiologiska förutsättningen att pulsökningen mellan två givna belastningar är högre hos en person med lägre $VO_2\max$ jämfört med en person med högre $VO_2\max$. Och genom att använda pulsökning mellan två nivåer, istället för en absolut puls på en belastning, hoppas man kunna eliminera felkällor som kan påverka pulsen såsom temperatur, nervositet eller känslor, likväl som individuella variationer i VO_2 på en given arbetsbelastning. En regressionsmodell har anpassat där man utifrån pulsökning, belastning, ålder och kön kan skatta $VO_2\max$. [1, s. 320] [2, s. 34]

1.2 Syfte

Syftet med detta kandidatarbete är att undersöka EBs (Ekblom-Baks) redan anpassade regressionsmodell samt att jämföra mot Åstrands test i de fall där försökspersoner har utfört bägge testen. Under arbetets gång erhöles mer data av kompletterande typ, både data över personer som uteslöts då EBs modell utformades samt kompletterande data i form av nya försökspersoner som utfört både EBs test och en mätning av $VO_2\max$. Det nya datat inkluderas successivt i regressionsmodellen.

1.3 Analysmetod

Analysmetoden är multipel linjär regressionsanalys, med undersökningar om transformation av respons och förklarande variabler kan vara aktuellt samt korsvalidering i form av *leave-one-out* då modellen ska användas för prediktion. För statistisk modellering används programvaran SAS 9.4.

2 Teori - Multipel linjär regression

2.1 Definition och skattning av modellens parametrar

Om vi har en responsvariabel som vi tror väl kan beskrivas linjärt av några förklarande variabler¹, låt säga p stycken, så kan vi använda modellen multipel linjär regression,

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \varepsilon_i \quad i = 1, \dots, n \quad (1)$$

där vi antar att

- i) De förklarande variablerna kan betraktas som fixa
- ii) De ej observerbara slumpmässiga felen är oberoende
- iii) Konstant varians hos försöksfelen, $Var(\varepsilon_i) = \sigma^2$, samt $E(\varepsilon_i) = 0$ för $i = 1, \dots, n$
- iv) $\varepsilon_i \sim N(0, \sigma^2)$.

En ekvivalent formulering av (1) kan fås genom att dra bort de aritmetiska medelvärden från var och en av regressorerna,

$$y_i = \alpha + \beta_1(x_{1i} - \bar{x}_1) + \beta_2(x_{2i} - \bar{x}_2) + \cdots + \beta_p(x_{pi} - \bar{x}_p) + \varepsilon_i, \quad (2)$$

där "nya α är relaterad till "gamla α genom sambandet

$$\text{"gamla } \alpha = \text{"nya } \alpha - (\beta_1 \bar{x}_1 + \cdots + \beta_p \bar{x}_p), \quad \beta_1, \dots, \beta_p \text{ oförändrad.}$$

Fördelen med (2) är att räkningarna blir enklare och risken för avrundningsfel minskar. Vi kommer att använda formuleringen (2) när vi presenterar teorin.

Låt nu

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & (x_{11} - \bar{x}_1) & \cdots & (x_{p1} - \bar{x}_p) \\ 1 & (x_{12} - \bar{x}_1) & \cdots & (x_{p2} - \bar{x}_p) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (x_{1n} - \bar{x}_1) & \cdots & (x_{pn} - \bar{x}_p) \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

så kan vi skriva (2) på formen

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

¹Kallas också regressorer.

En skattning av modellens parametrar fås nu med minsta-kvadratmetoden, där vi minimerar $\|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2$, till

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}.$$

Under förutsättningarna *i)–iii)* så ger minsta-kvadratmetoden en väntevärdesriktig skattning och ingen annan linjär väntevärdesriktig skattare är effektivare (*Gauss-Markovs sats*). En väntevärdesriktig skattare av σ^2 ges av residualkvadratsumman dividerad med antalet frihetsgrader,

$$\hat{\sigma}^2 = \frac{Kvs(residual)}{n - (p + 1)},$$

där $Kvs(residual) = \sum_{i=1}^n (y_i - (\hat{\alpha} + \hat{\beta}_1(x_{1i} - \bar{x}_1) + \dots + \hat{\beta}_p(x_{ip} - \bar{x}_p)))^2$. [8]

2.2 Test av linjära hypoteser

Att de slumpmässiga felen ska vara normalfördelade, (*iv*), är ej ett krav för att modellen ska vara tillförlitlig, men en förutsättning för vissa test och beräkning av p-värden i modellen om antalet observationer är få, annars blir dessa test approximativa [7, s. 106].

För test av om modellens regressorer förklarar något av variationen i \mathbf{Y} , dvs för test av hypotesen

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0 \quad (3)$$

så används teststatistikan

$$\frac{Kvs(regression)}{p\hat{\sigma}^2} = \frac{\sum_{i=1}^n (\sum_{j=1}^p \hat{\beta}_j (x_{ji} - \bar{x}_j))^2}{p\hat{\sigma}^2}$$

Under H_0 och normalfördelningsantagandet så är denna teststatistika $F(p, n - p - 1)$ -fördelad och vi förkastar för stora värden (om variationen som modellens regressorer förklarar är större än vad som kan anses bara vara slump). För alla redovisade modeller nedan kan (3) förkastas på 1% signifikansnivå.

För test av hypotesen att några av regressorerna, x_1, x_2, \dots, x_k , inte förklarar något av variationen i \mathbf{Y} , dvs för test av

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

så används teststatistikan

$$\frac{\sum_{i=1}^n (\sum_{j=1}^k \hat{\beta}_j (x_{ji} - \bar{x}_j))^2}{k\hat{\sigma}^2} \sim F(k, n - k - 1) \quad \text{under } H_0$$

där $\hat{\sigma}^2$ är σ^2 skattningen då alla parametrar är inkluderade i modellen. Då $k = 1$ redovisar `proc reg` i SAS p-värdet för ett tvåsidigt t-test vilket är ekvivalent med testet ovan. [8]

2.3 Anpassningsmått och prediktionsmått

Förklaringsgrad - R^2

Ett vanligt anpassningsmått är förklaringsgraden, R^2 , som anger andelen av modellen förklarad variation, och beräknas genom

$$R^2 = \frac{Kvs(regression)}{Kvs(total)} = 1 - \frac{Kvs(residual)}{Kvs(total)}, \quad (4)$$

där

$$Kvs(total) = \sum_{i=1}^n (y_i - \bar{y})^2.$$

Den andra likheten i (4) följer av att $Kvs(total) = Kvs(regression) + Kvs(residual)$. Beteckningen R^2 hänger ihop med att storheten är kvadraten på korrelationskoefficienten som mäter korrelationen mellan \mathbf{y} och $\hat{\mathbf{y}}$. När vi tillför en x -variabel till modellen så ökar alltid R^2 . Därför kan det vid variabelselektion vara bra att undersöka om $\hat{\sigma}^2$ minskar när vi lägger till en variabel. En storhet för detta är R_{adj}^2 .

Justerad förklaringsgrad - R_{adj}^2

R_{adj}^2 mäter hur mycket variansreduktion vi har i modellen jämfört med en modell med bara ett intercept och kan beräknas genom

$$R_{adj}^2 = 1 - \frac{\hat{\sigma}^2}{\hat{\sigma}_0^2},$$

där $\hat{\sigma}_0^2 = Kvs(total)/(n - 1)$ är σ^2 -skattningen när vi inte har någon x -variabel i modellen.

Root mean squared error of prediction - RMSEP

Ett sätt att undersöka en modells prediktiva förmåga är genom korsvalidering som i sin enklaste form, *leave-one-out*, går till på följande vis;

Tag tillfälligt bort observation nummer i ur datamaterialet och utför regression utan observation i . Använd regressionsmodellen för att prediktera observation nummer i med resultatet $\hat{y}_{i,(-i)}$. Indexet $(-i)$ syftar här till att observation nummer i utelämnats vid anpassningen av modellen. I kvadratisk medel får vi nu

$$MSEP = \frac{\sum_{i=1}^n (y_i - \hat{y}_{i,(-i)})^2}{n}.$$

$RMSEP = \sqrt{MSEP}$ är ett mått jämförbart med $\hat{\sigma}$ (också kallat *RMSE*) men för prediktion. Om framtida observationer kan tänkas likna de som används för att anpassa regressionsmodellen så kan *RMSEP* ses som en naturlig uppskattning av prediktionsfelet.

[8, s. 65-67]

2.4 Residualer och inflytelserika observationer

När vi anpassar modellen (2) så skattar vi β med $\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$ och predikterade värden med

$$\hat{\mathbf{Y}} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} = \mathbf{H} \mathbf{Y},$$

där \mathbf{H} kallas för *hattmatrisen*.

Residualerna definieras som

$$\hat{\varepsilon} = \mathbf{Y} - \hat{\mathbf{Y}}$$

och det går att visa att

$$E(\hat{\varepsilon}) = 0 \quad \text{och} \quad \text{Var}(\hat{\varepsilon}) = \sigma^2(\mathbf{I} - \mathbf{H}).$$

Variansen för den *ite* residualen är

$$\text{Var}(\hat{\varepsilon}_i) = \sigma^2(1 - h_{ii}), \quad (5)$$

där h_{ii} är det *ite* diagonalelementet i hattmatrisen. h_{ii} kallas för **leverage** och har egenskapen att $0 \leq h_{ii} \leq 1$. Namnet leverage (hävstång) kommer av att en observation med högt leverage har sina värden långt ifrån $\bar{\mathbf{x}}$, där $\bar{\mathbf{x}}$ betecknar vektorn innehållande regressorernas respektive medelvärde. En observation med högt leverage kan alltså vara potentiellt inflytelserik på regressionslinjen, men det är inte säkert att vi ser det i en plot över residualer på grund av egenskapen (5) och att $0 \leq h_{ii} \leq 1$. [9, s. 169-170]

Studentiserade residualer också kallade *standardiserade residualer* fås ur residualerna genom att variansstandardisera dem med deras medelfel

$$r_i = \frac{\hat{\varepsilon}_i}{\sqrt{\hat{\sigma}^2(1 - h_{ii})}}.$$

Under normalfördelningsantagandet gäller att $r_i \sim N(0, 1)$. [8, s. 73]

En observation som har högt leverage och/eller stor residual kan vara inflytelserik, dvs ha stor påverkan på regressionslinjen och parameterskattningar. Ett mått som tar hänsyn till detta är **Cooks distance**,

$$C_i = \frac{1}{(p+1)} r_i \left(\frac{h_{ii}}{1 - h_{ii}} \right).$$

Weisberg [9, s. 198-200] föreslår att $C_i > 1$ bör undersökas närmare. Här tillämpar vi istället principen att i added variable plots och C+R plot märka observationer som får $C_i > \frac{4}{n}$, där n är antalet observationer [10].

2.5 Grafisk diagnostik för regressionsmodeller

2.5.1 Residual plot

Genom att plotta residualerna, $\hat{\varepsilon} = \mathbf{Y} - \hat{\mathbf{Y}}$, mot predikterade värden så kan vi upptäcka eventuell systematik som är ett tecken på att modellen inte är väl anpassad. Ett exempel på detta är om residualerna ökar med predikterade värden, icke-konstant varians, vilket skulle kunna tyda på att logtransformation av responsen vore lämplig. Residualerna är positivt korrelerade med responsvariabeln, men okorrelerade med predikterade värden varför det är enklare att upptäcka avvikelser från modellantagandena i en plot av residualer mot predikterade värden. [4, s. 3-5]

I residual plotarna nedan är observationer med studentiserade residualer större än 1.8 märkta.

2.5.2 Added variable plot

För att undersöka vilken effekt en regressor (låt säga x_j) har på responsen givet att det redan finns andra förklarande variabler i modellen kan man använda sig av en *added variable plot*. I enkel linjär regression så kan vi undersöka relationen mellan respons och regressor med hjälp av en scatterplot. I multipel linjär regression så har regressorerna relationer sinsemellan som simultant kan tänkas påverka responsen, varför man i en added variable plot har eliminerat den delen av responsen respektive x_j som kan predikteras linjärt av de andra regressorerna. En added variable plot kan fås på följande vis

- i) Gör regression av y på övriga x -variabler, exklusive x_j , och spara residualerna, $\hat{\varepsilon}_y$, som representerar den del av y som ej kan förklaras av övriga x -variabler.
- ii) Gör regression av x_j på övriga x -variabler och spara residualerna, $\hat{\varepsilon}_{x_j}$, som representerar den delen av x_j som ej kan förklaras av övriga x -variabler.
- iii) Plotta $\hat{\varepsilon}_y$ mot $\hat{\varepsilon}_{x_j}$ och lägg in regressionslinjen som fås vid enkel linjär regression av $\hat{\varepsilon}_y$ på $\hat{\varepsilon}_{x_j}$.

Linjens lutning motsvarar koefficienten för x_j i den fulla regressionsmodellen. Vi kan också se potentiellt inflytelserika observationer som påverkar skattningen av koefficienten. [4, s. 6-10]

Nedan så är även den *partiella korrelationen* inskriven i ploten, vilket är korrelationen mellan y och x_j justerad för övriga x -variabler och beräknas som korrelationen mellan $\hat{\varepsilon}_y$ och $\hat{\varepsilon}_{x_j}$.

2.5.3 C+R plot

För att undersöka om en term, x_j , i regressionsmodellen behöver transformeras så tittar vi på en C+R plot ("Component + Residual). Detta är en plot över $(\hat{\beta}_j \mathbf{x}_j + \hat{\varepsilon})$ mot \mathbf{x}_j . Icke linjära samband kan också visa sig i en added variable plot, men genom att lägga till komponenten $\hat{\beta}_j \mathbf{x}_j$ till residualerna $\hat{\varepsilon}$ så kan vi ställa eventuell icke-linearitet i relation till det linjära samband vi har. Om förhållandet inte ser linjärt ut kan det vara aktuellt att söka efter en transformation. [4, s. 10-11]

2.6 Indikatorvariabler

En indikatorvariabel är en variabel som kan anta värdet 0 eller 1. Låt säga att vi har en indikatorvariabel

$$k\ddot{o}n = \begin{cases} 0 & \text{f\"or kvinnor} \\ 1 & \text{f\"or m\"an} \end{cases}$$

så kan vi genom att inkludera $k\ddot{o}n$ i regressionsmodellen tillåta olika intercept för män och kvinnor. Om modellen är

$$y_i = \alpha_1 + \alpha_2 k\ddot{o}n + \beta x_i + \varepsilon_i$$

så blir skattningen

$$\begin{aligned} \hat{y}_i &= (\hat{\alpha}_1 + \hat{\alpha}_2) + \hat{\beta} x_i && \text{om deltagaren \u00e4r en man och} \\ \hat{y}_i &= \hat{\alpha}_1 + \hat{\beta} x_i && \text{om deltagaren \u00e4r en kvinna.} \end{aligned}$$

P\u00e5 samma s\u00e4tt kan vi till\u00e5ta regressionslinjen att ha olika lutning f\u00f6r de olika k\u00f6nen genom att introducera en samspelsvariabel $k\ddot{o}n \cdot x_i$ till modellen,

$$y_i = \alpha + \beta_1 x_i + \beta_2 k\ddot{o}n \cdot x_i + \varepsilon_i.$$

I multipel linj\u00e4r regression s\u00e5 skulle samspelsvariabeln $k\ddot{o}n \cdot x_i$ svara mot att vi till\u00e5ter olika hyperplan f\u00f6r de olika k\u00f6nen, eller med andra ord vi till\u00e5ter olika koefficienter, f\u00f6r den f\u00f6rklarande variabeln x , f\u00f6r de olika k\u00f6nen.

3 Beskrivning av datamaterialet

Datamaterialet består av fyra subgrupperingar som vi här väljer att kalla studie 1-4, där studie 1 och 2 innefattar data över försökspersoner som deltagit då Ekblom-Baks test utvecklades och studie 3 och 4 innefattar kompletterande data i form av nya försökspersoner som utfört både submaximalt och maximalt test. För att identifiera enskilda försökspersoner så används deras försökspersonnummer. För en närmare beskrivning, än vad som ges här, av de första datamaterialen/variablerna och utformningen av Ekblom-Baks test se Ekblom-Bak m.fl. [1] eller Elin Ekblom-Baks doktorsavhandling [3].

3.1 Responsvariabel - $VO_2\text{max}$

Responsvariabeln är $VO_2\text{max}$ uppmätt i L/min genom ett maximalt test på löpband (studie 1-3) respektive ett maximalt test på cykel (studie 4).

3.2 De olika studierna

Studie 1 (S1)

Studie 1 består av originalpopulationen som har använts för att anpassa EBs regressionsmodell och innefattar 143 personer (65 män och 74 kvinnor) där man har eftersträvat att få ett blandat stickprov med avseende på kön, ålder, angiven träningsnivå (inaktiv till mycket aktiv). Bland dessa personer har 131 utfört ett accepterat Åstrandtest. Resterande 12 personer hade för låg puls under det submaximala testet för att tillåta skattning av $VO_2\text{max}$ med Åstrands nomogram.

Studie 2 (S2)

16 män som exkluderades ur originalpopulationen eftersom de hade uppmätt $VO_2\text{max}$ högre än 4.5 L/min och det ansågs öka felskattningen för mycket för övriga individer att inkludera dem i modellenpassningen.

Studie 3 (S3)

Ytterligare 12 män som testats under våren 2014 för att försöka hitta några med lågt uppmätt $VO_2\text{max}$ (låg kondition).

Studie 4 (S4)

48 äldre personer (24 män och 24 kvinnor) som ingått i en studie som studenter på Gymnastik- och idrottshögskolan i Stockholm har utfört. Då det kunde innefatta risker att utföra det maximala testet (mätningen av $VO_2\text{max}$) på löpband så har dessa individer istället utfört det maximala testet på cykel. Mer om detta i [Avsnitt 4.4.1](#).

3.3 Förklarande variabler

EBs test bygger på den arbetsfysiologiska förutsättningen att pulsökningen mellan två givna belastningar är högre hos en person med lågt $VO_2\max$ jämfört med en person med högre $VO_2\max$. Deltagarna cyklar på två olika belastningar, en lägre standardiserad och en högre individuellt vald i samråd med testledaren. (Avsnitt 1.1.2)

Variabelnamn	Förklaring
<i>kön</i>	indikatorvariabel, 0 för kvinnor, 1 för män
ΔHF^*	(slag/min)
ΔPO^*	(watt)
$\frac{\Delta HF}{\Delta PO}$	
<i>ålder</i>	(år)
<i>längd</i>	(meter) till närmaste 0.5 cm
<i>vikt</i>	(kg) till närmaste 0.1 kg

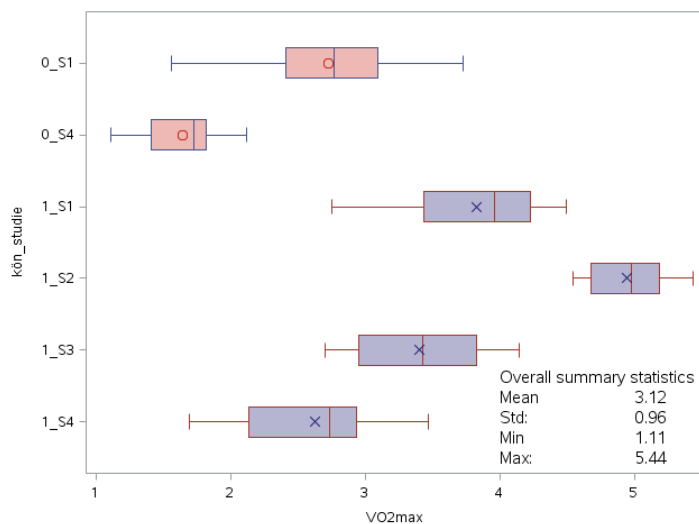
* ΔHF - Differensen mellan puls uppmätt på den högre individuellt valda belastningen och puls på den lägre standardiserade belastningen.

* ΔPO - (*Delta Power Output*) Differensen mellan effekt på den högre individuellt valda belastningen och den lägre standardiserade belastningen. Effekten är beräknad utifrån den inställda bromskraften (kp) och tramphastighet 60 varv/minut plus 8 % för att korrigera för friktion, huvudsakligen i kedjan, som ökar arbetets belastning [11] [1, s. 321].

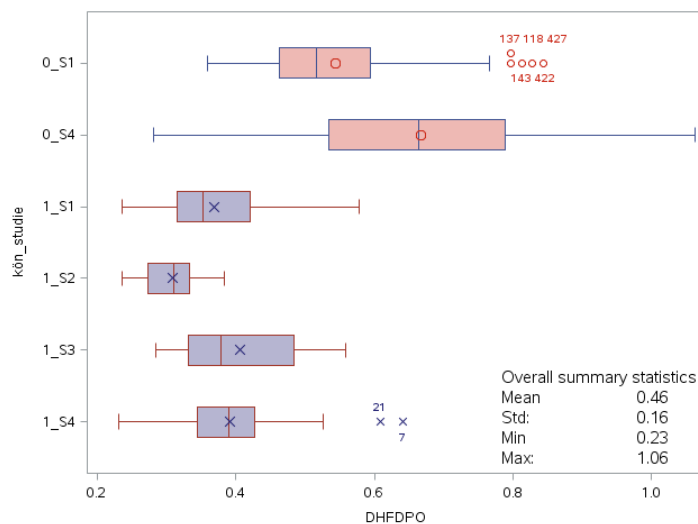
Då den högre belastningen är individuellt vald så behöver pulsökningen, ΔHF , standardiseras för att kunna jämföras mellan olika individer. Ett sätt att göra detta är att relatera den till det ökade syrebehovet ΔVO_2 . Bland de män och kvinnor, i studie 1 och 2, som även mätte VO_2 under submaximalt arbete på cykel så finner Ekblom-Bak m.fl. [3] [1] att ΔPO kan användas som företrädare (*proxy*) för ΔVO_2 och förhållandet antas vara linjärt samt liknande för män och kvinnor (jämför Åstrands Figur 1a). En multipel linjär regressionsmodell utformades med $\frac{\Delta HF}{\Delta PO}$, *kön* och *ålder* som förklarande variabler.

3.4 Scatterplot och boxplot av datamaterialet

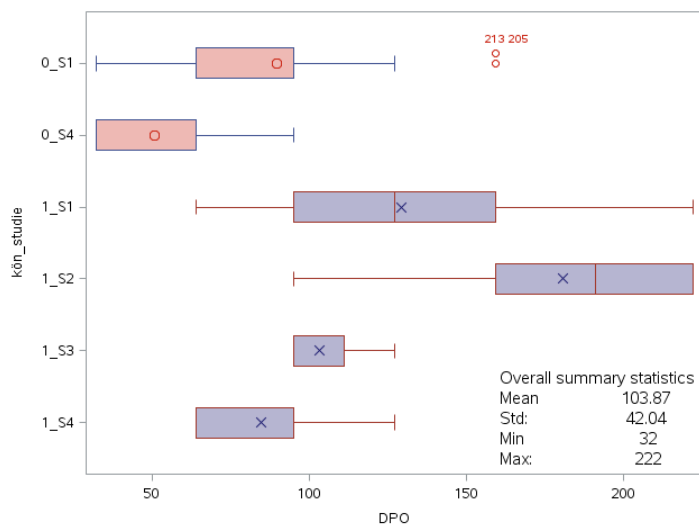
(a) $VO_2\max$



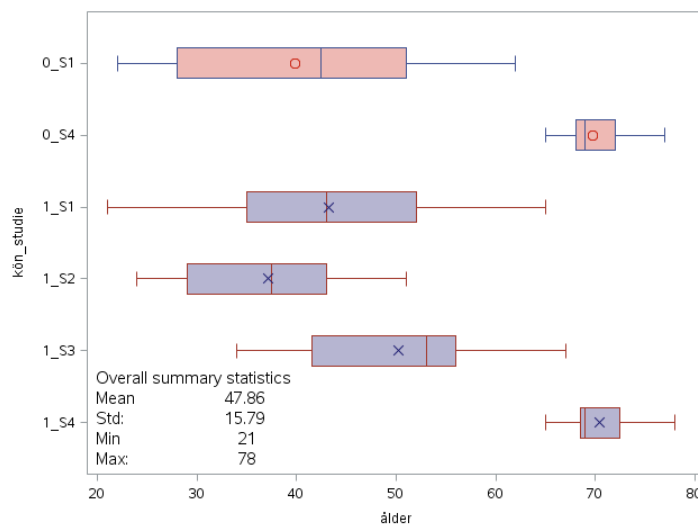
(b) $\frac{\Delta HF}{\Delta PO}$



(c) ΔPO

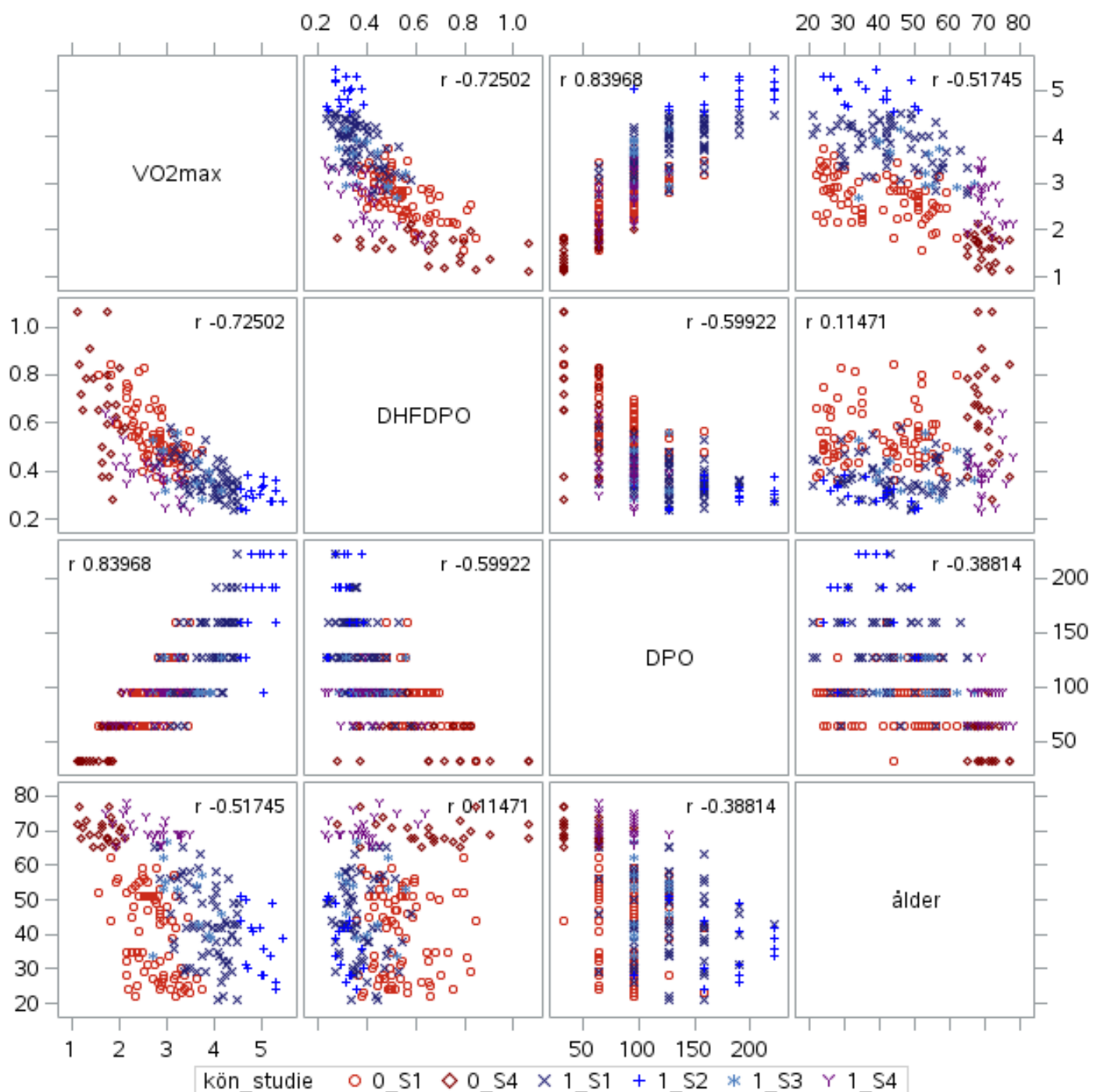


(d) $\text{\AA}lder$



Figur 2: Boxplot över vissa av variablerna

Kanterna på boxen ges av första (Q_1) och tredje (Q_3) kvartilen, markeringen inuti boxen visar medelvärdet, linjen medianen. Linjerna som går ut från boxen visar värden som ligger inom $1.5 \cdot (Q_3 - Q_1)$ från första respektive tredje kvartilen och de observationer som ligger utanför dessa kan betraktas som avvikande. För lägesmått och spridningsmått se [Appendix](#).



Figur 3: Scatterplot över vissa av variablerna samt Pearsons korrelationskoefficient (r). Notera att deltagarna i studie 4 (0_S4 och 1_S4) ser ut att ligga längre ned än övriga deltagare i ploten av $VO_2\text{max}$ mot $\frac{\Delta HF}{\Delta PO}$.

4 Statistisk analys

4.1 Studie 1 - Ekblom-Baks regressionsmodell

Ekblom-Baks modell

Respons $VO_2\max$
143 observationer, studie 1

	R^2	$RMSE$	$RMSEP$		
	0.8232	0.3045	0.3086		
Variabel	Skattning	Medelfel	t-värde	Pr	$> t $
Intercept	4.9820	0.1686	29.54	<	.0001
<i>kön</i>	0.6502	0.0683	9.53	<	.0001
$\frac{\Delta HF}{\Delta PO}$	-2.8862	0.2554	-11.30	<	.0001
<i>ålder</i>	-0.0171	0.0022	-7.87	<	.0001

Genom att inkludera variabeln *kön* i regressionsmodellen så tillåter vi olika intercept för de olika könen. I **Tabell 1** ser vi att korrelationen mellan *ålder* och $VO_2\max$ inte är så hög, -0.161, jämfört med de fulla datamaterialet, -0.517 (**Figur 3**), men trots det så är *ålder* en signifikant variabel. Den partiella korrelationen, korrelationen mellan *ålder* och $VO_2\max$ justerat för de övriga variablerna *kön* och $\frac{\Delta HF}{\Delta PO}$, är högre 0.555, så det är troligen så att den förklarande förmåga *ålder* har är specifik för just *ålder*.

	$VO_2\max$	$\frac{\Delta HF}{\Delta PO}$	ΔPO	<i>ålder</i>	<i>kön</i>
$VO_2\max$	1.000	-0.804	0.749	-0.161	0.765
$\frac{\Delta HF}{\Delta PO}$	-0.804	1.000	-0.618	-0.107	-0.658
ΔPO	0.749	-0.618	1.000	-0.009	0.568
<i>ålder</i>	-0.161	-0.107	-0.009	1.000	0.144
<i>kön</i>	0.765	-0.658	0.568	0.144	1.000

Tabell 1: Korrelationsmatris över variablerna i studie 1.

Respons	Regressorer	R^2	R^2_{adj}	$RMSE$	$RMSEP$
$VO_2\max$	<i>kön</i> $\frac{\Delta HF}{\Delta PO}$	0.7443	0.7407	0.3648	0.3685
$VO_2\max$	<i>kön</i> $\frac{\Delta HF}{\Delta PO}$ <i>ålder</i>	0.8232	0.8194	0.3045	0.3086
$VO_2\max$	<i>kön</i> $\frac{\Delta HF}{\Delta PO}$ <i>ålder</i> ΔPO	0.8673	0.8635	0.2647	0.2692
$\ln(VO_2\max)$	<i>kön</i> $\frac{\Delta HF}{\Delta PO}$ <i>ålder</i> ΔPO	0.8614	0.8573	0.0873	0.0890

Tabell 2: Anpassningsmått och prediktionsmått

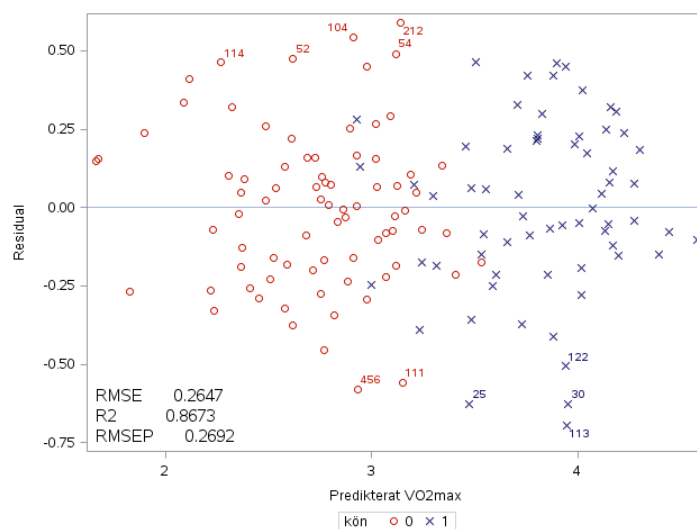
Om vi tar bort *ålder* så minskar förklaringsgraden (R^2), R^2_{adj} och modellens prediktiva förmåga ($RMSEP$ ökar), vilket talar för att *ålder* ska

vara med i modellen. Vi kan också se att anpassningsmått och modellens prediktiva förmåga ökar om vi lägger till variabeln ΔPO . ΔPO står som täljare i $\frac{\Delta HF}{\Delta PO}$ och en anledning att ΔPO ökar anpassningen av modellen skulle kunna vara att $VO_2\max$ inte är helt linjär i beroendet av $\frac{\Delta HF}{\Delta PO}$. Den nya modellen ser ut som följer

Respons $VO_2\max$
143 observationer, studie 1

	R^2	RMSE	RMSEP
	0.8673	0.2647	0.2696

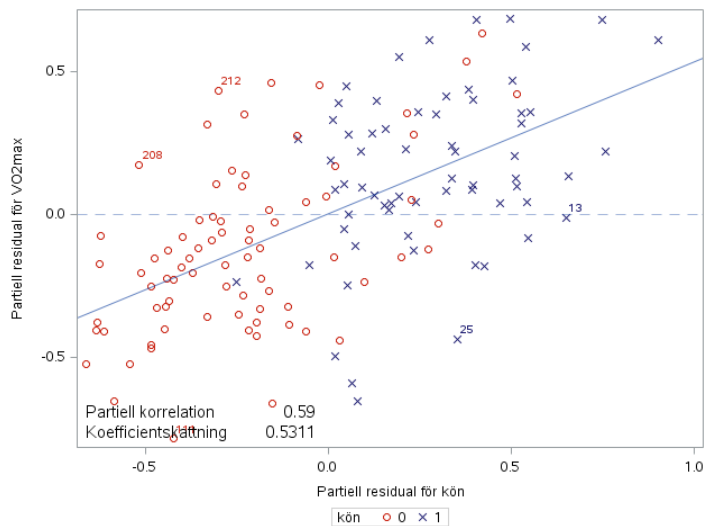
Variabel	Skattning	Medelfel	t-värde	Pr > t
Intercept	4.0412	0.2019	20.02	< .0001
<i>kön</i>	0.5311	0.0619	8.58	< .0001
$\frac{\Delta HF}{\Delta PO}$	-2.2311	0.2422	-9.22	< .0001
ΔPO	0.00578	0.00085	6.78	< .0001
<i>ålder</i>	-0.0155	0.0019	-8.11	< .0001



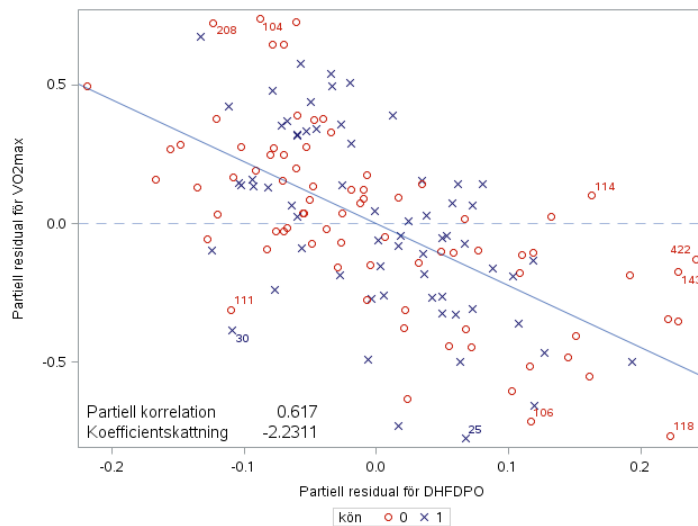
Figur 4: Residual mot predikerat värde

I **Figur 4** skulle man kunna ana ökande varians med ökande predikerade värden men logaritmering av responsen ger ej ökad förklaringsgrad, varken då regressorerna är otransformerade ($R^2 = 0.8614$) eller då de icke-kategoriska regressorerna är logtransformerade ($R^2 = 0.8532$). En orsak till att det inte är så stor skillnad mellan en modell med logaritmerad respons och ursprunglig respons är att logaritmfunktionen är nästan linjär i det intervall där vi har uppmätt $VO_2\max$, 1.5 till 4.5 L/min.

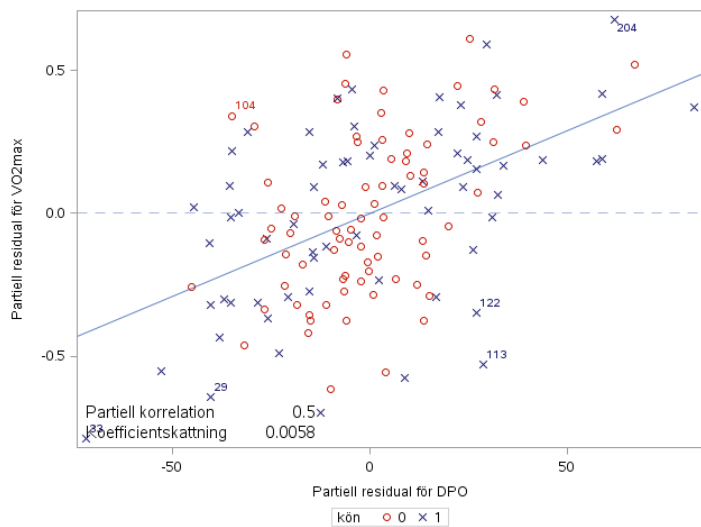
(a) Added variable plot för *kön*



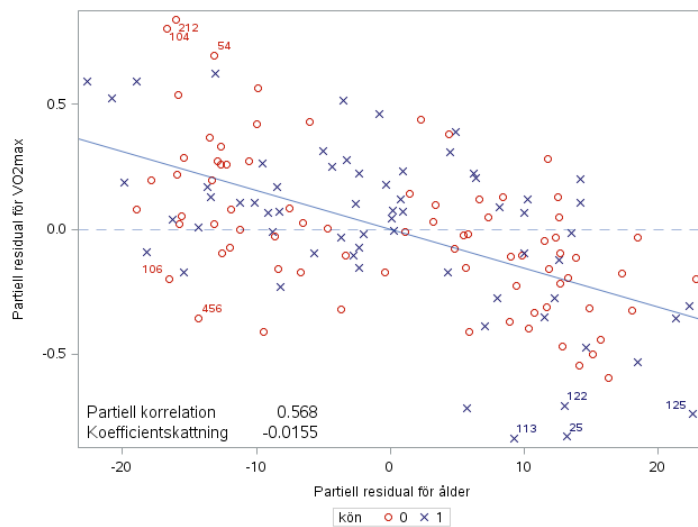
(b) Added variable plot för $\frac{\Delta HF}{\Delta PO}$



(c) Added variable plot för ΔPO

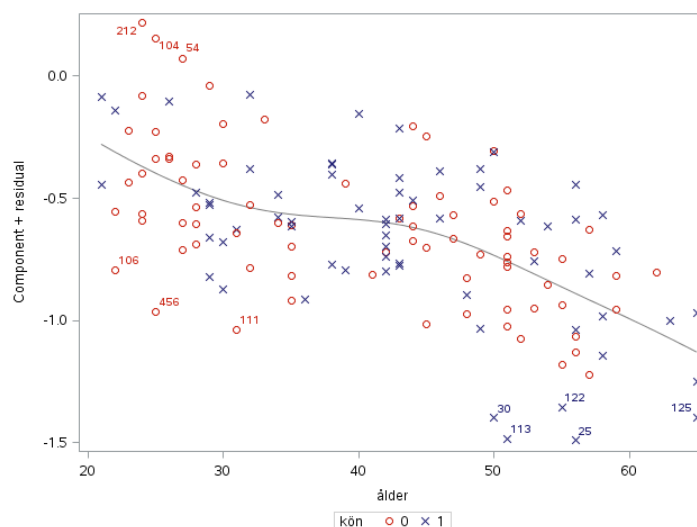


(d) Added variable plot för *ålder*



Figur 5: Added variable plots, studie 1
Respons: $VO_2\max$, Regressorer: *kön*, $\frac{\Delta HF}{\Delta PO}$, ΔPO , *ålder*

I added variable plots så kan vi se att den partiella korrelationen är ≥ 0.5 för alla regressorer. Det ser inte ut som att det är skillnad mellan könen, vad gäller påverkan på regressionslinjen vid tillägg av en regressor, om vi bortser från regressorn *kön*. Det ser inte heller ut att finnas några extrema outliers. Däremot ser relationen mellan $VO_2\max$ och *ålder*, justerat för de övriga variablerna, inte helt linjär ut. Vi tittar närmare på den.



Figur 6: C+R plot för *ålder*

Respons: $VO_2\max$, Regressorer: *kön*, $\frac{\Delta HF}{\Delta PO}$, ΔPO , *ålder*

I ploten är också en kubisk penalized b-spline inlagd som stöd för ögat [6]

I C+R plot för *ålder* så ser markeringarna ut att kunna tänkas följa ett tredjegradspolynom. Vi testar att lägga till $\textit{ålder}^2$ och $\textit{ålder}^3$ till modellen.

Respons	Regressorer	R^2	R_{adj}^2	RMSE	RMSEP
$VO_2\max$	<i>kön</i> $\frac{\Delta HF}{\Delta PO}$ <i>ålder</i> ΔPO	0.8673	0.8635	0.2647	0.2692
$VO_2\max$	<i>kön</i> $\frac{\Delta HF}{\Delta PO}$ ΔPO <i>ålder</i> $\textit{ålder}^2$ $\textit{ålder}^3$	0.8705	0.8648	0.2635	0.2694

R_{adj}^2 , ökar något, men $RMSEP$ är nästan oförändrad. Dessutom visar det sig att vid test av om koefficienterna, β_4 respektive β_5 , för $\textit{ålder}^2$ respektive $\textit{ålder}^3$, är noll, dvs vid test av

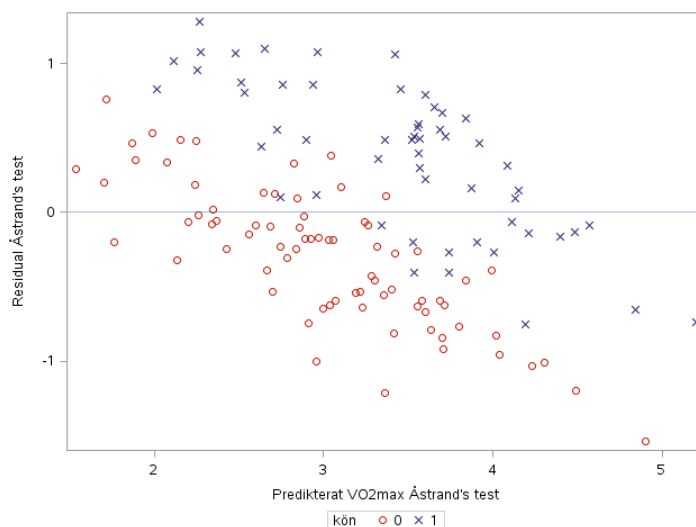
$$H_0 : \beta_4 = \beta_5 = 0,$$

så är p-värdet 0.1981. Hur är det då med splinen i **Figur 6**? Om vi ersätter komponenten $\hat{\beta}$ *ålder* med motsvarande värden för splinen så ökar förklaringsgraden till 0.8725, dvs resultatet är liknande som vid att lägga till andra och tredjegrads term för *ålder*. Även fast relationen inte ser helt linjär ut så är det svårt att hitta någon passande transformation.

4.1.1 Jämförelse med Åstrands test

Då vi har godkänt Åstrands test för enbart 131 av deltagarna, 76 kvinnor och 55 män, så utesluts övriga deltagare i jämförelserna nedan om inte annat

anges. Ekblom Bak m.fl [1, s 319] skriver att tidigare studier på Åstrands test har visat att Åstrands test tenderar både att underskatta $VO_2\max$ (främst för män) och överskatta $VO_2\max$ (främst för kvinnor).



Figur 7: Residual mot predikerat värde för Åstrands test

Om vi tittar på residualer mot predikerade värden för Åstrands test så ser det mycket riktigt ut som att männen underskattas och kvinnorna överskattas. Inför en ny variabel $Pred_Åstrand$ för predikerat $VO_2\max$ med Åstrands test och granska **Figur 7** i termer av linjär regressionsanalys. Det ser ut som att koefficienten för $Pred_Åstrand$ inte ska vara ett, samt att vi behöver ha olika intercept för män och kvinnor. Alltså borde vi kunna korrigera Åstrands test genom att utföra linjär regression med $VO_2\max$ som responsvariabel och $Pred_Åstrand$ och $kön$ som regressorer.

Skattningsmetod	R^2	$RMSE$	$RMSEP$
Åstrands test	0.4550*	-	-
Åstrands korrigerade test	0.8138*	0.3101	0.3131
EBs test	0.8258*	0.3045 \diamond	0.3086 \diamond
EBs test med tillägg av ΔPO	0.8751*	0.2647 \diamond	0.2692 \diamond

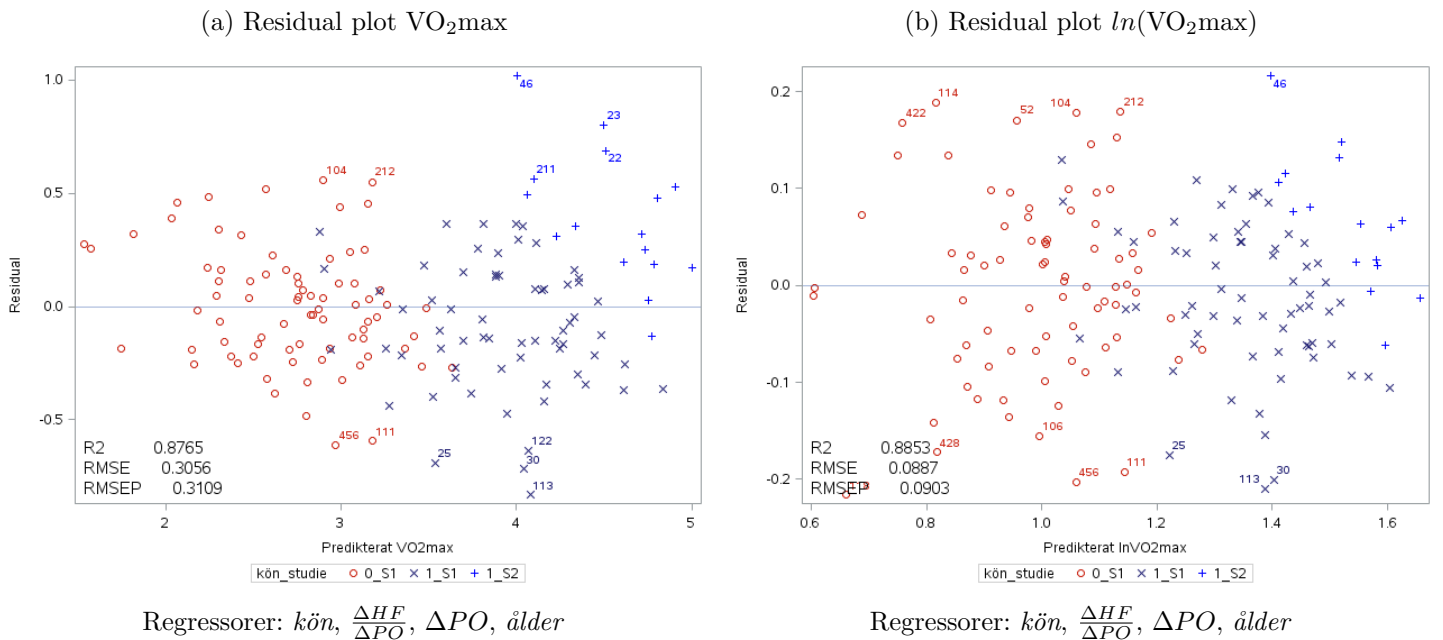
* R^2 är beräknad som korrelationskoefficienten i kvadrat mellan $VO_2\max$ och predikerat värde av $VO_2\max$ för de 131 försökspersoner som har godkänt Åstrands test. \diamond $RMSE$ och $RMSEP$ är skattningarna då alla 143 försökspersonerna är inkluderade i modellenanpassningen.

Vi kan se att då vi korrigerar Åstrands test med hjälp av delar av datamaterialet som används för att anpassa EBs test så är skillnaden i anpassning och prediktiv förmåga mellan Åstrands test och EBs test liten. För parame-

terskattningar, residual plot och QQ-Plot för Åstrands korrigerade test se [Appendix](#).

4.2 Studie 2 - Logaritmering av responsen

Studie 2 innehåller 16 män som exkluderades ur originalpopulationen eftersom de hade uppmätt $VO_2\max$ högre än 4.5 L/min och det ansågs öka felskattningen för mycket för övriga individer att inkludera dem i modellenpassningen. I [Figur 2](#) (boxplot över datamaterialet) kan vi också se att flertalet av männen i studie 2 har högre ΔPO och lägre pulsökning per effekttökning $\frac{\Delta HF}{\Delta PO}$, jämfört med övriga deltagare. När vi lägger till männen i studie 2 fyller vi alltså på med fler värden i ändarna av intervallen för $VO_2\max$, ΔPO och $\frac{\Delta HF}{\Delta PO}$.



Figur 8: Residual plots med deltagarna i studie 1 och 2

Om vi lägger till männen i studie 2 till modellen med $kön$, $\frac{\Delta HF}{\Delta PO}$, ΔPO och ålder som regressorer så ser det ut som att vi får ökande residualer med ökande predikterade värden, men genom att logaritmera responsen så verkar vi får mer konstant varians. Detta ökar också förklaringsgraden.

Respons	Regressorer	R^2	$RMSE$	$RMSEP$
$VO_2\max$	$kön$, $\frac{\Delta HF}{\Delta PO}$, ålder	0.8082	0.3796	0.3840
$VO_2\max$	$kön$, $\frac{\Delta HF}{\Delta PO}$, ΔPO , ålder	0.8765	0.3056	0.3110
$\ln(VO_2\max)$	$kön$, $\frac{\Delta HF}{\Delta PO}$, ΔPO , ålder	0.8853	0.0887	0.0903

Genom att inkludera ΔPO så blir effekten av att lägga till männen i studie 2 mindre. Nu verkar också en modell med logaritmerad respons bättre. Observera att $RMSEP$ inte är jämförbar mellan transformerad och ursprunglig respons. Istället så svarar det nu mot en uppskattning av prediktionsfelet i procent av predikterat värde, vilket här blir ungefär 10% ($e^{0.09} = 1.09$ och $e^{-0.09} = 0.91$). Vad händer om vi inkluderar fler försökspersoner i modellen?

4.3 Studie 3 - Transformation av regressor?

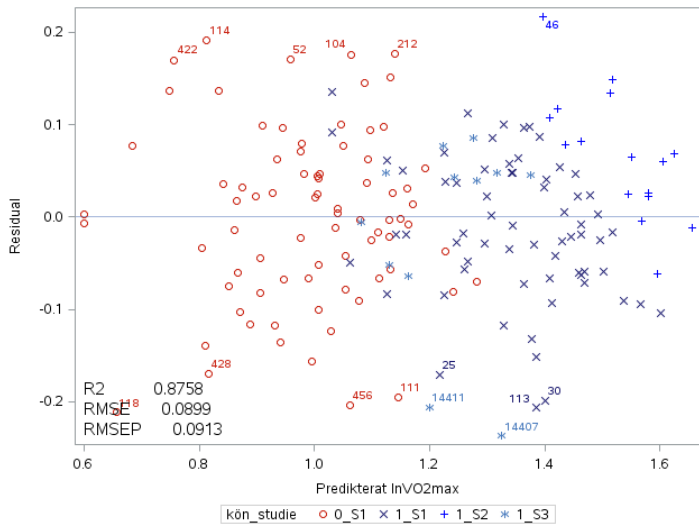
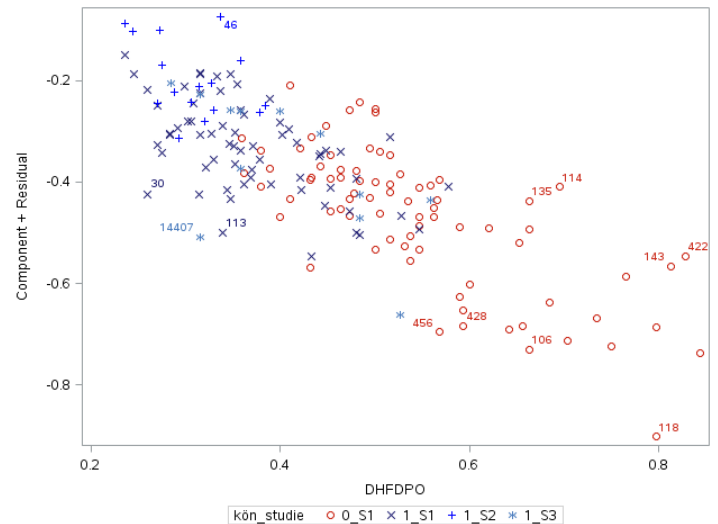
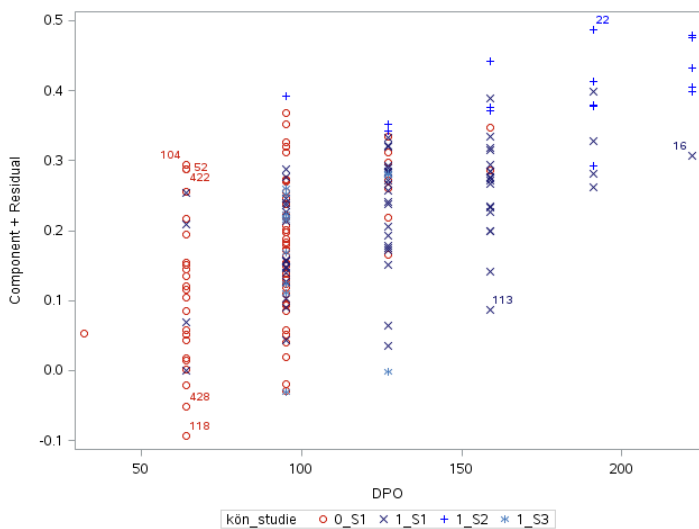
Under vårterminen 2014 annonserades det efter män som var lågaktiva/inaktiva för att inkludera i anpassningen av testet. De 12 män som anmälde intresse och deltog hade i medel inte så mycket lägre uppmätt $VO_2\max$ (3.40 L/min) jämfört med männen i originalpopulationen (studie 1, 3.83 L/min) men genom att inkludera dem så lägger vi till fler försökspersoner med $VO_2\max$ i det lägre intervallet för männen.

Respons $\ln(VO_2\max)$				
studie 1 och 2			studie 1, 2 och 3	
159 observationer			171 observationer	
R^2			R^2	
$RMSEP$			$RMSEP$	
0.8853			0.8758	
0.0903			0.09134	
Variabel	Skattning	Medelfel	Skattning	Medelfel
Intercept	1.5054	0.0636	1.5161	0.0629
$kön$	0.1563	0.0205	0.1521	0.0196
$\frac{\Delta HF}{\Delta PO}$	-0.8539	0.0789	-0.8660	0.0772
ΔPO	0.00185	0.00023	0.00185	0.00023
$ålder$	-0.00544	0.00062	-0.00554	0.00060

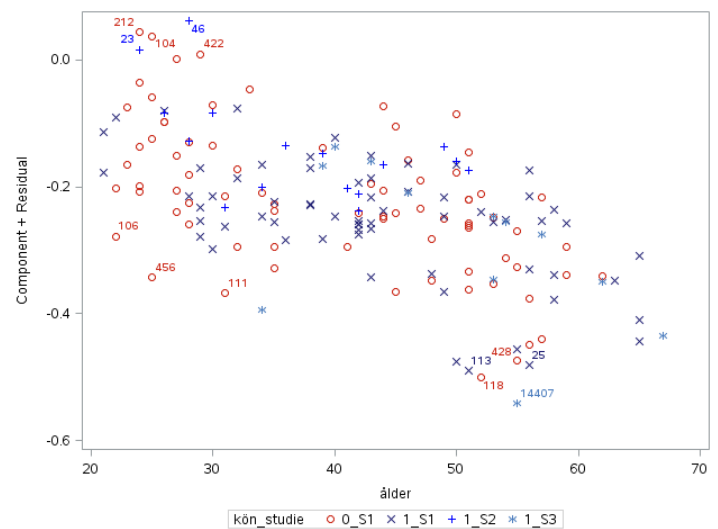
Tabell 3: Parameterskattningar och anpassningsmått
Hypotesen att en parameterskattning är noll, givet att de andra parametrarna är med modellen, kan förkastas på 1% signifikansnivå för alla variabler

Detta minskar koefficientskattningen för $kön$ något samt även för $\frac{\Delta HF}{\Delta PO}$. Den observanta läsaren kanske noterar att R^2 minskar trots att vi lägger till observationer. Det kommer av att vi lägger till försökspersoner som, för de förklarande variablerna, har sina värden nära medelvärdena för männen och kvinnorna tillsammans, samtidigt som den totala variationen kring $\overline{VO_2\max}$ ökar. Två av männen i studie 3, försökspersonnummer 14411 och 14407, får höga residualer vilket kan ses i [Figur 9a](#). Man kan även se på added variable plots ([Appendix](#)) att de verkar ha inflytande på parameterskattningarna, men då vi har många observationer för att anpassa regressionslinjen så ser de inte extremt inflytelserika ut.

(a) Residual mot predikterade värden

(b) C+R plot för $\frac{\Delta HF}{\Delta PO}$ (c) C+R plot för ΔPO 

(d) C+R plot för ålder



Figur 9: Residual plot och C+R plots, studie 1-3

Respons: $\ln(\text{VO}_2\text{max})$, Regressorer: kön , $\frac{\Delta HF}{\Delta PO}$, ΔPO , ålder

För att undersöka om någon av regressorerna behöver transformeras så tittar vi på C+R plots för de kontinuerliga variablerna. Förhållandet ser nästan linjärt ut för både $\frac{\Delta HF}{\Delta PO}$ och ΔPO och det visar sig att transformering, t.ex. logaritmering av en av dem eller bägge (antagande om multiplikativa effekter) inte förbättra modellens prediktiva förmåga. $RMSEP$ ökar om vi logaritmerar någon eller flera av regressorerna. Det är också bara i

änden som C+R plot för $\frac{\Delta HF}{\Delta PO}$ ser något olinjär ut och logaritmering ändrar inte detta förhållande. C+R plot för *ålder* ser som tidigare inte helt linjär ut. Men resultatet, om vi lägger till tredje- och andrageradsterm för *ålder*, är liknande de vi fick då bara originalpopulationen var inkluderad i modelanpassningen. Vid tillägg av andra och tredje gradsterm för ålder ökar R_{adj}^2 något, men så gör även *RMSEP*.

Hur är det då med de olika könen? Borde vi ha en könsspecifik modell utöver olika intercept för de olika könen? I added variable plots ([Appendix](#)) så ser spridningen mellan könen jämn ut vad gäller påverkan på regressionslinjen men vi kan ändå testa om vi borde ha olika koefficienter, för de förklarande variablerna, för de olika könen. Inför samspelsvariabler $kön \cdot \frac{\Delta HF}{\Delta PO}$, $kön \cdot \Delta PO$ samt $kön \cdot ålder$ med koefficienter β_4 , β_5 respektive β_6 och testa hypotesen

$$H_0 : \beta_4 = \beta_5 = \beta_6 = 0$$

för modellen som också innehåller $kön$, $\frac{\Delta HF}{\Delta PO}$, ΔPO och *ålder* som förklarande variabler. P-värdet för detta test blir 0.7, förklaringsgraden för modellen, 0.8768, och QQ-plot över residualer ger inte tecken på att strida mot normalfördelningsantagandet. Det tyder på att vi inte behöver könsspecifik modell utöver olika intercept för de olika könen.

4.4 Studie 4 - Inkludering av alla försökspersoner

4.4.1 Korrigering för annan mätmetod av $VO_2\max$

Männen och kvinnorna i studie 4 har utfört mätningen av $VO_2\max$ på cykel istället för på löpband. I möte med Elin Ekblom-Bak så berättade hon att maximalt test och mätning av $VO_2\max$ på cykel kunde ge 5-10 % lägre uppmätt $VO_2\max$ än om samma person utförde ett maximalt test på löpband. För deltagarna i studie 4, som i medel är äldre än resterande deltagare i detta material, var inte ett maximalt test på löpband möjligt. Det fanns också en tveksamhet om kvinnorna bröt testet tidigare, på grund av bristande benstyrka, än männen. 94 personer i originalpopulationen utförde två mätningar av $VO_2\max$ och 59 av dem hade två godkända mätningar. Bland de 59 som hade två godkända mätningar så var inte variationen stor mellan tillfällena; *The CV was 2.70 %, which is regarded as small and in concordance with internal laboratory measurements and other previous reports...* [1, s. 321]. Däremot var spridningen större för de deltagarna i studie 4 som utförde mer än ett maximalt test. Högst uppmätt $VO_2\max$ är inkluderat i detta datamaterial. Samtidigt vore det bra att kunna inkludera dessa deltagare eftersom det handlar om personer som antas (oavsett mätmetod) ha lågt $VO_2\max$ vilket behövs för att göra modellen mer generell och mer användbar i en blandad population.

För att kunna inkludera deltagarna i studie 4 så behöver vi kontrollera om det finns systematisk skillnad i uppmätt VO₂max mot övriga deltagare och korrigera för detta. I och med att vi har responsen på logskala så kan vi undersöka detta genom att införa en indikatorvariabel

$$studie4 = \begin{cases} 1 & \text{för Studie 4} \\ 0 & \text{för Studie 1-3} \end{cases}$$

och utföra linjär regression med *studie4* inkluderad som regressor. Be-teckna parameterskattningen för *studie4* med $\hat{\beta}_{s4}$. Logaritmlagarna ger att $\ln(VO_{2max} \cdot pomax) = \ln(VO_{2max}) + \ln(pomax)$ där vi låter *pomax* be-teckna procent av VO₂max. Om en modell med $\ln(VO_{2max})$ som respons och *kön*, $\frac{\Delta HF}{\Delta PO}$, ΔPO och *ålder* som regressorer kan antas ha god prediktiv förmåga för äldre försökspersoner med lågt VO₂max, så ger $\hat{\beta}_{s4}$ en skatt-ning av $\ln(pomax)$. $\exp(\hat{\beta}_{s4})$ ger då i medel, procent av VO₂max som cyklat VO₂max har gett.

Med liknande argument så kan vi införa en till term *kön* · *studie4* för att undersöka om det är skillnad mellan könen, med avseende på hur många procent cyklat VO₂max har gett av vad motsvarande test på löpband skulle ha gett. Det visar sig att koefficientskattningen för *kön* · *studie4* inte är signi-fikant skild från noll på 20 %s-nivån. Däremot om vi utför linjär regression med vår tidigare anpassade modell, alla försökspersoner inkluderade och regressorn *studie4* inkluderad så är koefficientskattningen för *studie4* signi-fikant skild från noll på 1%^s nivå och fås till -0.15037. $R^2 = 0.9027$ och QQ-plot över residualer ser inte ut att strida mot normalfördelningsantagandet. Detta motsvarar 86.0 % av VO₂max ($e^{-0.15037} = 0.86$).

Variabel	Respons $\ln(VO_{2max}^*)$					
	studie 1-2		studie 1-3		studie 1-4	
	159 observationer		171 observationer		219 observationer	
	R^2	$RMSEP$	R^2	$RMSEP$	R^2	$RMSEP$
	0.8853	0.0903	0.8758	0.09134	0.8824	0.10610
Skattning	Medelfel	Skattning	Medelfel	Skattning	Medelfel	
Intercept	1.5054	0.0636	1.5161	0.0629	1.4137	0.0581
<i>kön</i>	0.1563	0.0205	0.1521	0.0196	0.1690	0.01200
$\frac{\Delta HF}{\Delta PO}$	-0.8539	0.0789	-0.8660	0.0772	-0.7227	0.0648
ΔPO	0.00185	0.00023	0.00185	0.00023	0.00226	0.00024
<i>ålder</i>	-0.00544	0.00062	-0.00554	0.00060	-0.00599	0.00051

Tabell 4: Parameterskattningar och anpassningsmått

* VO₂max korrigerat med 14 % för studie 4.

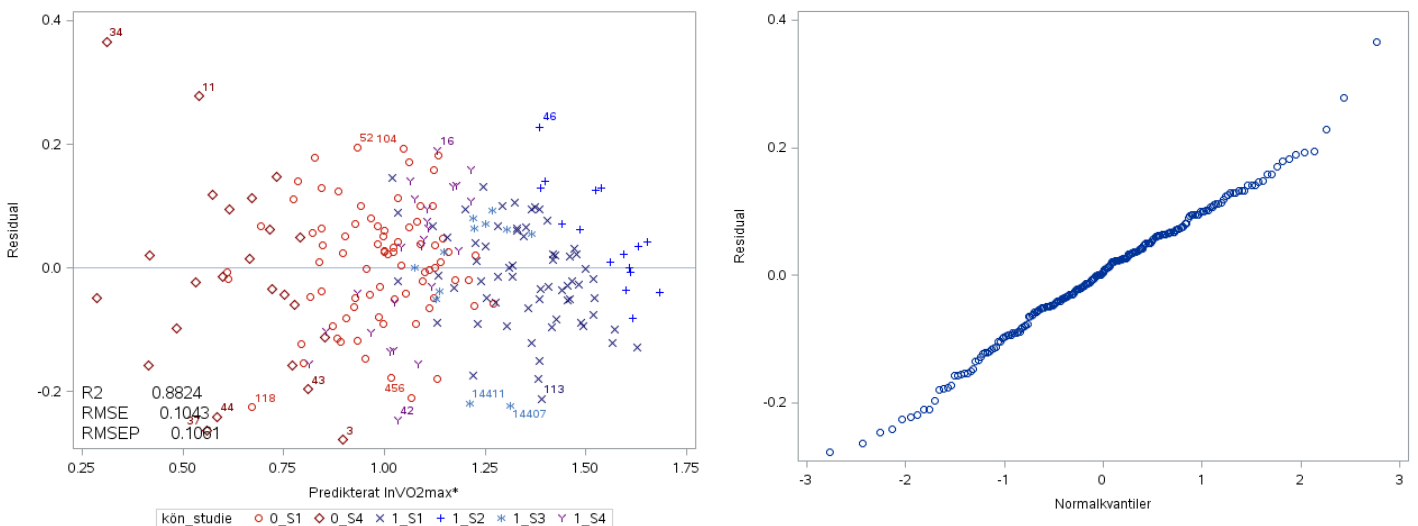
Hypotesen att en parameterskattning är noll, givet att de andra parametrarna är med i modellen, kan förkastas på 1%^s signifi-kansnivå för alla variabler.

Om vi hade behållt variabeln $kön \cdot studie4$ så hade motsvarande skattningar blivit 87.5 % för männen och 84.3 % för kvinnorna. Vi korrigerar försökspersonernas VO_2max i studie 4 genom att höja det med 14 % för att kunna inkludera dem i modellenpassningen. **Tabell 3** utvidgas med parameterskattningar när hela datamaterialet är inkluderat.

När vi inkluderar deltagarna i studie 4 så ökar koefficientskattningen för $kön$ vilket kan förklaras av att kvinnorna i studie 4 har lågt uppmätt VO_2max i relation till övriga deltagarna. Detta kan ses i boxploten **Figur 2** eller i lägesmått och spridningsmått i **Appendix**. Deltagarna i studie 4 är äldre personer, $ålder \geq 65$, med lågt uppmätt VO_2max vilket också medför att koefficientskattningen för $ålder$ blir mindre. $RMSEP$ blir nu 0.10610 vilket motsvarar en uppskattning av prediktionsfelet till 11 % av predikerat värde på ursprunglig skala.

(a) Residual mot predikerat värde

(b) QQ-Plot över residualerna



Figur 10: Residual plot och QQ-plot, studie 1-4
 Respons: $\ln(VO_2max^*)$, Regressorer: $kön$, $\frac{\Delta HF}{\Delta PO}$, ΔPO , $ålder$

En deltagare, försöksperson nummer 34 i studie 4 får hög residual. Det är värt att notera att detta är en av personerna som vi korrigerat VO_2max med 14% för. Höjningen med 14% behöver inte vara korrekt på individnivå utan stämmer bättre på gruppnivå för att få rätt parameterskattningar. Försöksperson nummer 34 är en av de kvinnor som har högst noterat $\frac{\Delta HF}{\Delta PO}$ och lägsta belastningsökning ΔPO . Den enda andra kvinnan som har lika hög $\frac{\Delta HF}{\Delta PO}$ är försöksperson nummer 11. Dessa två individer är de som får högst Cooks D och som även ligger långt ifrån övriga i added variable plots (**Appendix**). Om vi utesluter försökspersonnummer 34 så får vi en liten minskning av $RMSEP$ till 0.1026. Om vi dessutom också utesluter

försöksperson nummer 11 så är motsvarande värde 0.1005. Dessa individer är alltså inflytelserika, men vi vet inte om det finns anledning att utesluta dem, såsom mätfel eller liknande.

`proc reg` i SAS 9.4 tillhandahåller ett alternativ för test av hypotesen att de ej observerbara felen är oberoende och har konstant varians. För närmare detaljer se [5]. P-värdet för detta test, med alla försökspersoner inkluderade, blir 0.0405. Vi kan på 5%^s signifikansnivå förkasta hypotesen om konstant varians och oberoende fel. Om vi däremot utesluter försöksperson nummer 34 så blir motsvarande p-värde 0.0586. I **Figur 10** kan vi se att den grupp av individer som ser ut att få störst residualer är kvinnor med lågt predikerat $\ln(\text{VO}_2\text{max})$, vilket motsvaras av främst kvinnorna studie 4. Att de tillhör de individer som vi har korrigerat VO_2max med 14 %, vilket inte behöver passa bra på individnivå, gör tolkningen av residualerna i **Figur 10** svår.

4.4.2 Vikt, längd och andra utförda test

Deltagarna i studie 4 är alla individer med $\text{ålder} \geq 65$. I de tidigare studierna finns bara 4 deltagare med $\text{ålder} \geq 65$. Om vi utför samma test som för deltagarna i studie 4, dvs om det verkar som att modellen systematisk över eller underskattar de 4 deltagarna med $\text{ålder} \geq 65$, som ej ingår i studie 4, så blir inte parameterskattningen för indikatorvariabeln ej signifikant skild från noll. Detsamma gäller om vi upprepar testet för individer med ålder i intervallet 20 - 40 år samt 40 - 60 år.

Indikatorvariabler med indelning per vart 20 år har också använts för att undersöka om koefficientskattningarna behöver variera med åldern för de kontinuerliga variablerna. Ingen samspelsterm mellan åldersindikatorn och kontinuerlig variabel har fått parameterskattning skild från noll på 5%^s nivån och detsamma gäller för samspelsvariabler mellan ålder och $\frac{\Delta HF}{\Delta PO}$ respektive ålder och ΔPO .

Då även vikt och längd fanns inkluderade i datamaterialet så kändes det naturligt att testa om de, eller kanske ännu hellre BMI , hade något att tillföra som förklarande variabler.

Korrelation med VO_2max^*								
Kvinnor			Män			Totalt		
$n = 102$			$n = 117$			$n = 219$		
<i>vikt</i>	<i>längd</i>	<i>BMI</i>	<i>vikt</i>	<i>längd</i>	<i>BMI</i>	<i>vikt</i>	<i>längd</i>	<i>BMI</i>
0.1365	0.3076	-0.0452	0.1781	0.4590	-0.1552	0.5315	0.6774	0.1367

Trots att korrelation med VO_2max är rätt hög för *vikt* och *längd* för det fulla datamaterialet så kan vi se att den inte är så hög för de separata könen. Att den är hög för det fulla datamaterialet beror troligen på att männen i

datamaterialet har högre uppmätt $VO_2\text{max}$ samtidigt som de också väger mer och är längre än kvinnorna. Den möjligt prediktiva förmåga som *vikt*, *längd* och *BMI* har tas om hand av de förklarande variabler vi redan har inkluderade i vår modell varför de inte har något att tillföra.

5 Diskussion

Syftet med detta kandidatarbete var att undersöka Ekblom-Baks redan anpassade regressionsmodell samt att jämföra mot Åstrand's test i de fall där försökspersoner har utfört bägge testen. Då vi korrigerar Åstrand's test så är skillnaden i prediktiv förmåga inte så stor mellan de olika testen. Något som skulle kunna tala för Ekblom-Baks test är om det inte behövs upprepade submaximala test för att skatta $VO_2\text{max}$.

När vi inkluderar fler försökspersoner i modellenpassningen så kan vi se att parameterskattningarna, inte helt oväntat, ändras. Eftersom modellen är tänkt att kunna användas i en blandad population så vore det bra att kunna inkludera, så nära det går, ett tvärsnitt av populationen med avseende på kön, ålder och träningsnivå. Men hur vet man vad ett tvärsnitt av populationen är? Sen tillkommer svårigheten att de som kanske är mest intresserade att delta i utformandet, och göra både submaximalt test och maximalt test, mycket möjligt är träningsintresserade individer. Även om syftet främst är att kunna prediktera $VO_2\text{max}$ för icke-elitidrottare så kan det ändå vara bra att inkludera en del individer med bra kondition (tex deltagarna i studie 2) för att inte anpassa en regressionsmodell på ett alltför homogent material. På detta vis kan vi, när vi fyller på värden i ändarna av intervallen se om modellen varit korrekt eller om vi behöver transformera variabler (tex ha responsen på logskala). Om vi anpassar en regressionsmodell på ett alltför homogent datamaterial så blir modellen bra på att prediktera för individer som liknar de som används för att anpassa modellen, men inte lika säkert bra på att prediktera $VO_2\text{max}$ för individer som skiljer sig från de som användes för att anpassa modellen.

Det finns också mycket osäkerhet angående uppmätt $VO_2\text{max}$ för deltagarna i studie 4. Hur stora var skillnaderna i $VO_2\text{max}$ bland de individer som utförde två maxtest i studie 4? Är variationen lika stor för yngre mer tränade individer vid mätning av $VO_2\text{max}$ på cykel? Med yngre avses här individer som är yngre än 65 år. När vi korrigerar $VO_2\text{max}$ för deltagarna i studie 4 så förutsätter vi att modellen med den respons och de regressorer vi använder också kan fungera för äldre individer med lågt uppmätt $VO_2\text{max}$. Det kan vara möjligt, med säkrare mätning av $VO_2\text{max}$ för denna målgrupp, att det visar sig att det inte är en helt lämplig modell för dem. Under hösten 2014 kommer fler försökspersoner med lågt $VO_2\text{max}$ att inkluderas i studien och de kommer att utföra maxtestet, mätning av $VO_2\text{max}$, på löpband.

Det vore intressant om försöksperson nummer 34 kunde upprepa både sub-

maximalt och maximalt test för att se om både uppmätt $VO_2\text{max}$ och $\frac{\Delta HF}{\Delta PO}$ blir liknande eller kan det vara så att någonting påverkade henne så att differensen mellan puls på lägre och högre belastning var påverkad? Är hennes värden extrema eller är modellen inte väl anpassad i ändarna där kvinna 34 har sina värden? Intressant vore också att inkludera fler kvinnor med högt uppmätt $VO_2\text{max}$ för att se vad som händer med parameterskattningar. Det skulle kunna visa sig att man tex behöver könsspecifika koefficienter vid inkludering av fler kvinnor med högt $VO_2\text{max}$, vilket nuvarande modell och datamaterial inte ger stöd för.

Att ΔPO inte tidigare inkluderats i modellen hade delvis med att göra att vilken belastningsnivå som, i samråd med testledaren, valdes för den högre belastningen kan variera med ett steg från dag till dag, medan $\frac{\Delta HF}{\Delta PO}$ anses vara konstant vid ändring av ett belastningssteg. Ett steg motsvarar en förändring i $\pm 32 \Delta PO$. Vi kan se att inkludering av ΔPO minskar prediktionsfelet ($RMSEP$) för de individer som är med och anpassar modellen och en anledning till detta skulle kunna vara att $VO_2\text{max}$ inte är helt linjär i beroendet av $\frac{\Delta HF}{\Delta PO}$. ΔPO används i kvoten $\frac{\Delta HF}{\Delta PO}$ som proxy för ΔVO_2 och förhållandet mellan $VO_2\text{max}$ och $\frac{\Delta HF}{\Delta VO_2}$ antas vara linjärt likvärd som förhållandet mellan ΔVO_2 och ΔPO . Något man skulle kunna undersöka vidare är relationen mellan ΔVO_2 och ΔPO för att undersöka om transformation av ΔPO i kvoten $\frac{\Delta HF}{\Delta PO}$ kan ge en säkrare skattning av $VO_2\text{max}$.

På grund av osäkerheten i responsen för deltagarna i studie 4 så kan slutgiltig modell i denna uppsats och vägen till den ses som en möjlig metod att vidareutveckla Ekblom-Baks test.

Referenser

- [1] Ekblom-Bak, E. m.fl. “A new submaximal cycle ergometer test for prediction of VO₂max”. I: *Scandinavian Journal of Medicine & Science in Sports* 24.2 (2012), s. 319–326.
- [2] Ekblom-Bak, Elin. “Nytt cykeltest för att mäta konditionen”. I: *Svensk Idrottsforskning* 4.1 (2013), s. 33–35.
- [3] Ekblom-Bak, Elin. “Physical activity, cardiorespiratory fitness, and abdominal obesity in relation to cardiovascular disease risk - epidemiological studies”. Diss. Karolinska Institutet, 2013.
- [4] Ohlsson, Esbjörn. *Grafisk diagnostik för regressionsmodeller*. Matematiska institutionen, Stockholms universitet, febr. 1998.
- [5] SAS Institute Inc. *Reg procedure: Testing for Heteroscedasticity*. Hämtat 27 juli 2014. URL: http://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_reg_sect042.htm.
- [6] SAS Institute Inc. *Transreg procedure: Penalized B-Splines*. Hämtat 27 juli 2014. URL: http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_transreg_sect020.htm.
- [7] Sen, Ashis och Srivastava, Muni. *Regression Analysis - Theory, Methods, and Applications*. Springer, 1990.
- [8] Sundberg, Rolf. *Lineära Statistiska Modeller*. Matematiska institutionen, Stockholms universitet, okt. 2013.
- [9] Weisberg, Sanford. *Applied Linear Regression*. Third edition. John Wiley & Sons, Inc, 2005.
- [10] Wikipedia. *Cook's distance*. Hämtat 16 augusti 2014. URL: http://en.wikipedia.org/wiki/Cook's_distance.
- [11] Åstrand, Per-Olof. *Work tests with the bicycle ergometer*. Varberg, 1965.
- [12] Åstrand, Per-Olof och Ryhming, Irma. “A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work”. I: *Journal of Applied Physiology* 7.2 (1954), s. 218–221.

6 Appendix

Lägesmått och spridningsmått

Studie 1												
Variabel	Kvinnor				Män				Totalt			
	<i>n</i> = 78				<i>n</i> = 65				<i>n</i> = 143			
	mean	std.	min	max	mean	std.	min	max	mean	std.	min	max
VO ₂ max	2.73	0.45	1.56	3.73	3.83	0.47	2.75	4.49	3.23	0.72	1.56	4.49
$\frac{\Delta HF}{\Delta PO}$	0.54	0.12	0.36	0.84	0.37	0.08	0.24	0.58	0.46	0.13	0.24	0.84
ΔPO	89.56	22.54	32	159	129	34.65	64	222	107.50	34.72	32	222
<i>ålder</i>	39.85	12.12	22	62	43.26	11.40	21	65	41.40	11.88	21	65

Studie 2					Studie 3			
Variabel	Män				Män			
	<i>n</i> = 16				<i>n</i> = 12			
	mean	std.	min	max	mean	std.	min	max
VO ₂ max	4.94	0.28	4.54	5.44	3.40	0.49	2.70	4.14
$\frac{\Delta HF}{\Delta PO}$	0.31	0.04	0.24	0.38	0.41	0.09	0.28	0.56
ΔPO	180.69	39.65	95	222	103.00	14.47	95	127
<i>ålder</i>	37.19	8.85	24	51	50.25	9.92	34	67

Studie 4												
Variabel	Kvinnor				Män				Totalt			
	<i>n</i> = 24				<i>n</i> = 24				<i>n</i> = 48			
	mean	std.	min	max	mean	std.	min	max	mean	std.	min	max
VO ₂ max*	1.87	0.33	1.27	2.41	2.99	0.58	1.93	3.95	2.43	0.74	1.27	3.95
$\frac{\Delta HF}{\Delta PO}$	0.67	0.20	0.28	1.06	0.39	0.10	0.23	0.64	0.53	0.21	0.23	1.06
ΔPO	50.63	18.57	32	95	84.71	17.61	64	127	67.67	24.84	32	127
<i>ålder</i>	69.79	3.34	65	77	70.38	3.35	65	78	70.08	3.32	65	78

* VO₂max höjt med 14% för studie 4

Lägesmått och spridningsmått, succesiv inkludering av de olika studierna

Studie 1

Variabel	Kvinnor				Män				Totalt			
	<i>n</i> = 78				<i>n</i> = 65				<i>n</i> = 143			
	mean	std.	min	max	mean	std.	min	max	mean	std.	min	max
VO ₂ max	2.73	0.45	1.56	3.73	3.83	0.47	2.75	4.49	3.23	0.72	1.56	4.49
$\frac{\Delta HF}{\Delta PO}$	0.54	0.12	0.36	0.84	0.37	0.08	0.24	0.58	0.46	0.13	0.24	0.84
ΔPO	89.56	22.54	32	159	129	34.65	64	222	107.50	34.72	32	222
<i>ålder</i>	39.85	12.12	22	62	43.26	11.40	21	65	41.40	11.88	21	65

Studie 1 och 2

Variabel	Kvinnor				Män				Totalt			
	<i>n</i> = 78				<i>n</i> = 81				<i>n</i> = 159			
	mean	std.	min	max	mean	std.	min	max	mean	std.	min	max
VO ₂ max	2.73	0.45	1.56	3.73	4.05	0.63	2.75	5.44	3.40	0.86	1.56	5.44
$\frac{\Delta HF}{\Delta PO}$	0.54	0.12	0.36	0.84	0.36	0.08	0.24	0.58	0.45	0.14	0.24	0.84
ΔPO	89.56	22.54	32	159	139.22	41.04	64	222	114.86	41.48	32	222
<i>ålder</i>	39.85	12.12	22	62	42.06	11.16	21	65	40.97	11.66	21	65

Studie 1, 2 och 3

Variabel	Kvinnor				Män				Totalt			
	<i>n</i> = 78				<i>n</i> = 93				<i>n</i> = 171			
	mean	std.	min	max	mean	std.	min	max	mean	std.	min	max
VO ₂ max	2.73	0.45	1.56	3.73	3.96	0.65	2.70	5.44	3.40	0.84	1.56	5.44
$\frac{\Delta HF}{\Delta PO}$	0.54	0.12	0.36	0.84	0.36	0.08	0.24	0.58	0.45	0.13	0.24	0.84
ΔPO	89.56	22.54	32	159	134.55	40.48	64	222	114.03	40.27	32	222
<i>ålder</i>	39.85	12.12	22	62	43.12	11.30	21	67	41.63	11.76	21	67

Studie 1, 2, 3 och 4

Variabel	Kvinnor				Män				Totalt			
	<i>n</i> = 102				<i>n</i> = 117				<i>n</i> = 219			
	mean	std.	min	max	mean	std.	min	max	mean	std.	min	max
VO ₂ max*	2.53	0.56	1.27	3.73	3.76	0.75	1.93	5.44	3.19	0.91	1.27	5.44
$\frac{\Delta HF}{\Delta PO}$	0.57	0.15	0.28	1.06	0.37	0.08	0.23	0.64	0.46	0.16	0.23	1.06
ΔPO	80.40	27.23	32	159	124.32	42.07	64	222	103.87	42.04	32	222
<i>ålder</i>	46.89	16.65	22	77	48.71	15.02	21	78	47.86	15.79	21	78

* VO₂max höjt med 14% för studie 4

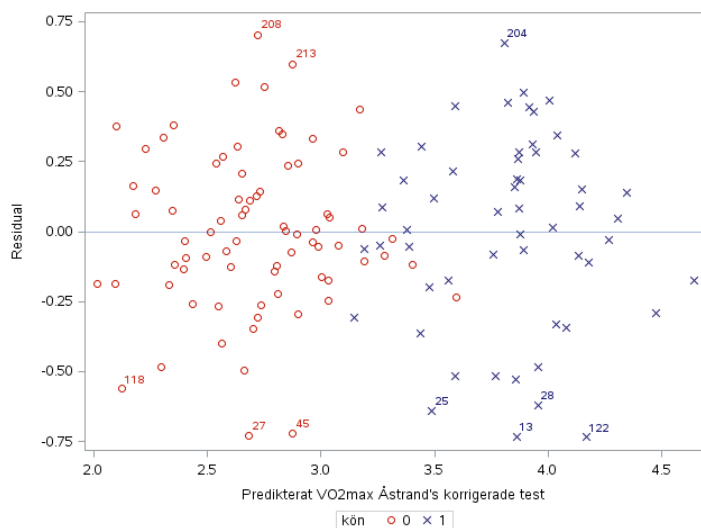
Åstrands korrigerade test

Respons VO₂max
131 observationer från studie 1

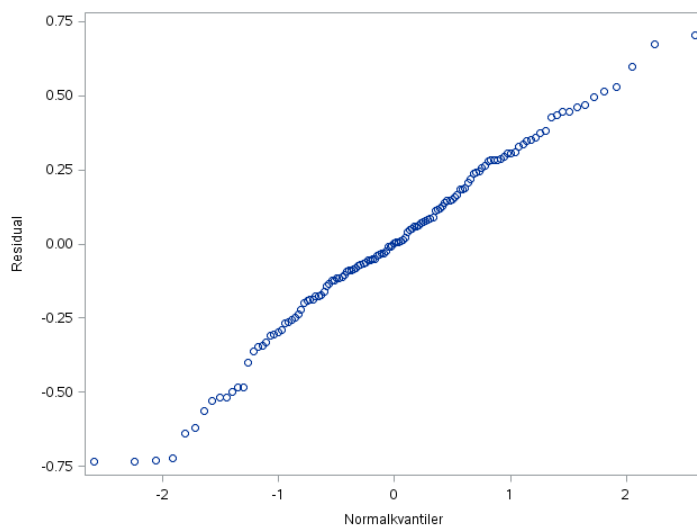
R^2	$RMSE$	$RMSEP$
0.81384	0.31014	0.31311

Variabel	Skattning	Medelfel	t-värde	Pr > t
Intercept	1.29443	0.12187	10.62	< .0001
<i>kön</i>	0.90481	0.05760	15.71	< .0001
<i>Pred_Åstrand</i>	0.46985	0.03880	12.11	< .0001

Residual plot

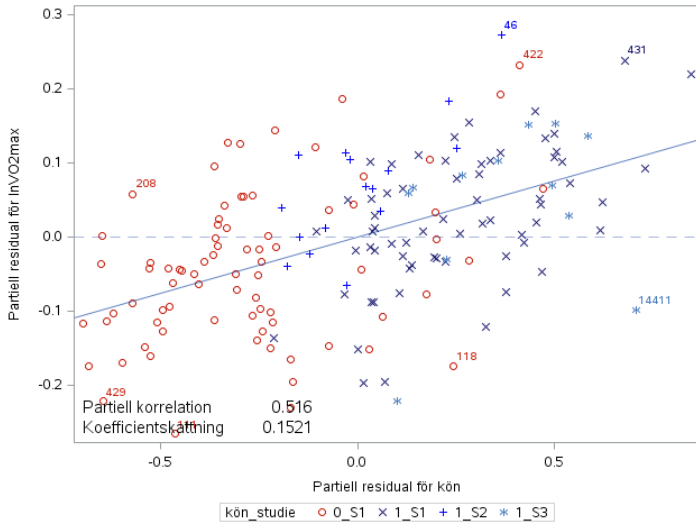


QQ-plot över residualerna

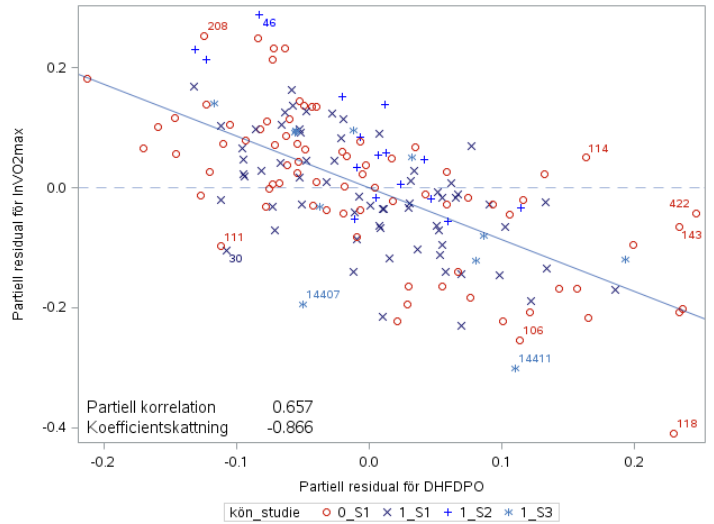


Added variable plots, studie 1-3

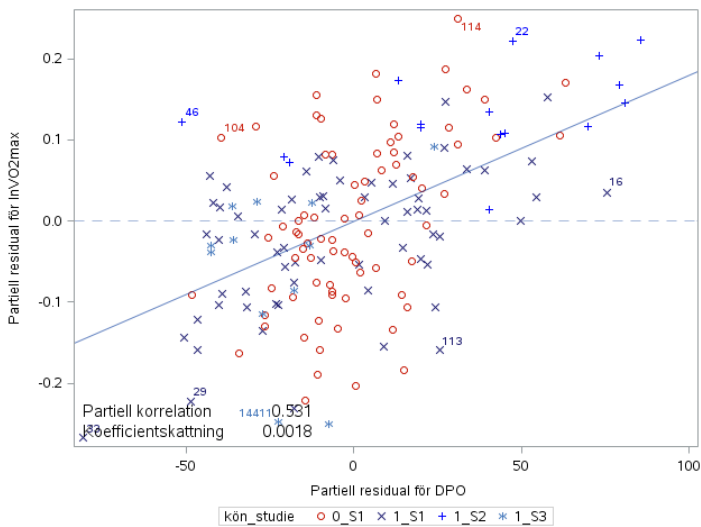
Added variable plot för *kön*



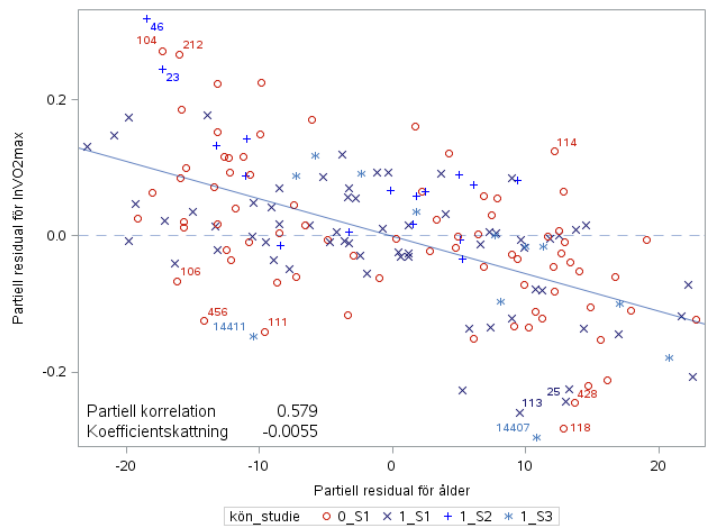
Added variable plot för $\frac{\Delta HF}{\Delta PO}$



Added variable plot för ΔPO



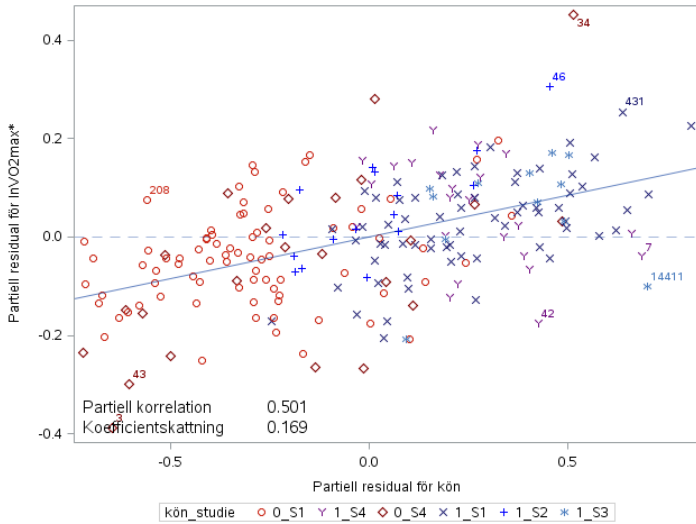
Added variable plot för *ålder*



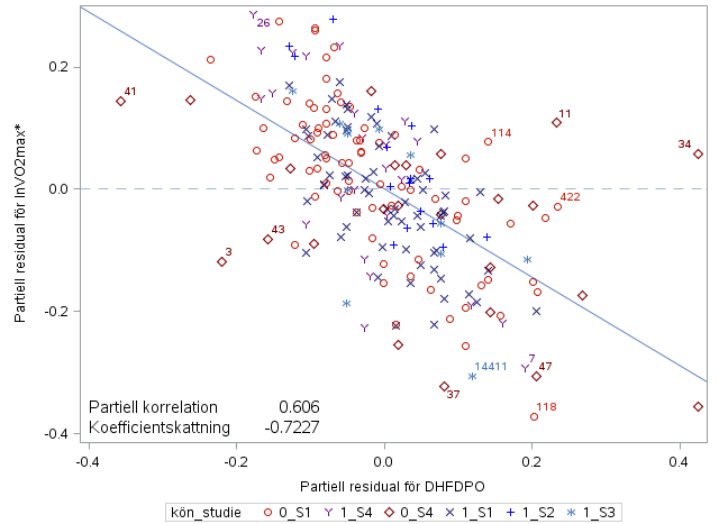
Respons: $\ln(VO_2 \max)$, Regressorer: *kön*, $\frac{\Delta HF}{\Delta PO}$, ΔPO , *ålder*

Added variable plots, studie 1-4

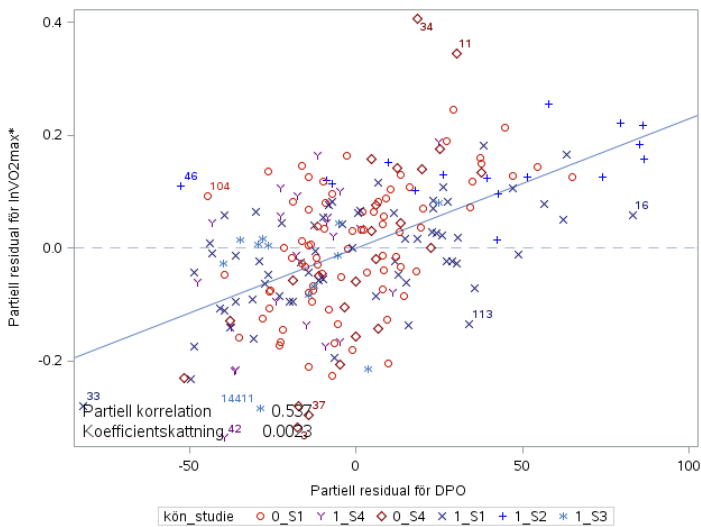
Added variable plot för *kön*



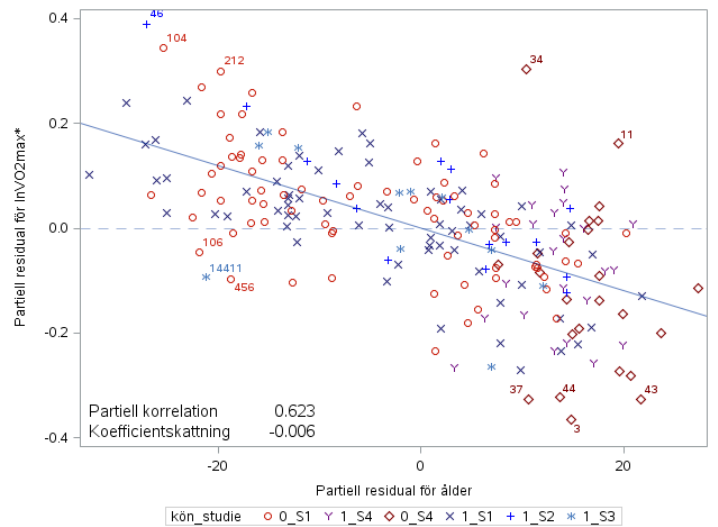
Added variable plot för $\frac{\Delta HF}{\Delta PO}$



Added variable plot för ΔPO



Added variable plot för *ålder*



Respos: $\ln(VO_2 \max^*)$, Regressorer: *kön*, $\frac{\Delta HF}{\Delta PO}$, ΔPO , *ålder*