



**LUNDS**  
UNIVERSITET  
*Statistiska Institutionen*

# En jämförelse mellan resultaten vid analyser av kapillärt, venöst och arteriellt blod

**Håkan Tykesson**

16 juni 2010

Uppsats i statistik  
15 poäng  
Nivå 61-90 poäng  
VT 2010

Handledare: Jan Lanke

# Innehåll

<b>1</b>	<b>Inledning</b>	<b>4</b>
1.1	Syfte med rapporten . . . . .	4
<b>2</b>	<b>Datamaterialet</b>	<b>5</b>
<b>3</b>	<b>Teori</b>	<b>5</b>
3.1	Modellval 1 - Optimistisk metod . . . . .	7
3.2	Modellval 2 - Den försiktiga metoden . . . . .	8
3.3	Modellval 3 - ANOVA med slumpmässig effekt . . . . .	8
3.4	Prediktion . . . . .	9
<b>4</b>	<b>Resultat</b>	<b>10</b>
4.1	Kapillärt vs venöst . . . . .	10
4.1.1	Optimistisk metod . . . . .	11
4.1.2	Den försiktiga metoden . . . . .	11
4.1.3	ANOVA med slumpmässig effekt . . . . .	12
4.1.4	Prediktion . . . . .	12
4.2	Kapillärt vs arteriellt . . . . .	13
4.2.1	Optimistisk metod . . . . .	14
4.2.2	Den försiktiga metoden . . . . .	14
4.2.3	ANOVA med slumpmässig effekt . . . . .	15
4.2.4	Prediktion . . . . .	15
4.3	Kapillärt vs venöst/arteriellt . . . . .	16
4.3.1	Optimistisk metod . . . . .	17
4.3.2	Den försiktiga metoden . . . . .	17
4.3.3	ANOVA med slumpmässig effekt . . . . .	18
4.3.4	Prediktion . . . . .	19
4.4	Venöst vs arteriellt . . . . .	19
4.4.1	Oberoende . . . . .	20
4.4.2	Patientmedelvärde . . . . .	20
<b>5</b>	<b>Slutsats och diskussion</b>	<b>22</b>
<b>6</b>	<b>Referenser</b>	<b>23</b>
<b>A</b>	<b>Appendix -1</b>	<b>24</b>
<b>B</b>	<b>Appendix - 2</b>	<b>25</b>
<b>C</b>	<b>Appendix - Dataset</b>	<b>27</b>

### Sammanfattning

Hjärnskademarkören S100B är ett protein som huvudsakligen finns i stödjevävnaden i hjärnan och som läcker ut via blod-hjärnbarriären vid hjärnskada. För att kunna bestämma koncentrationen av S100B i blodet kan man använda tre olika blodprovsmetoder, kapillärt, venöst eller arteriellt. Det vanligaste sättet att ta ett blodprov är venöst men då det är både enklare och snabbare att ta kapillära prov är det önskvärt att använda sig av denna metod. Datamaterialet kommer från 71 olika mättillfällen, uppdelade på 52 olika patienter, där man tagit blodprov kapillärt och antingen venöst eller arteriellt. Målet med rapporten är att undersöka om det finns skillnader mellan de olika blodprovsmetoderna och ta fram prediktionsmodeller samt att beräkna prediktionsfel för att kunna göra en bedömning om det är lämpligt att ersätta venösa alt. arteriella prov med kapillära. För att analysera ifall det finns skillnader mellan kapillära och venösa/arteriella blodprov användes tre olika metoder: en optimistisk, en försiktig och ANOVA med slumpmässig effekt. Skillnaderna mellan venösa och arteriella prov studerades med t-test. Resultatet visade att det finns skillnader mellan kapillära och venösa alt. arteriella prov, i samtliga fall gav kapillära prov en högre koncentration. Då prediktionsfelen är stora kan det innebära problem om man vill ersätta venösa/arteriella prov med kapillära. Venösa och arteriella prov förväntas ge samma resultat enligt ANOVA.

### Abstract

The brain injury marker S100B is a protein found primarily in the supporting tissues of the brain and leaking through the blood brain barrier in case of a brain injury. In order to determine the concentration of S100B in the blood one can use three different methods of blood samples: capillary, venous or arterial. The most common way to take a blood sample is venous, but when it is both easier and faster to take capillary samples it is desirable to use this method. The dataset comes from 71 different measurement occasions, divided in 52 patients, where samples of blood were taken both capillary and either venous or arterial. The aim to this paper is to study if there are any differences between capillary and venous/arterial blood samples. Further on models for predictions will be done to calculate the errors of prediction in order to make an assessment if it is appropriate to replace venous/arterial samples with capillaries. To determine whether there are differences between the methods of blood samples three different approach will be done: an optimistic method, a cautious method and ANOVA with random effect. The difference between venous and arterial samples is made with ANOVA. The results show that there are differences between capillary and venous/arterial samples. In all three cases the capillary samples give a higher concentration. Since the error of prediction was large it can cause problem when replacing venous blood samples with capillary ones. According to ANOVA venous and arterial samples are expected to give similar results.

## 1 Inledning

Den här rapporten kommer att behandla ett problem där man mätt koncentrationen av ett visst protein i blodet med olika typer av blodprov. För att jämföra olika tekniker och analysmetoder är studier där man använder sig av två eller flera metoder vanliga. Syftet med dessa analyser är att undersöka ifall metoderna stämmer väl överens med varandra och i fall dessa kan användas omväxlande eller ersätta varandra. Går det till exempel att använda en billigare och snabbare metod som ger liknade resultat[1 sid 396-397].

Det analyserade proteinet som den här rapporten baseras på är en hjärn-skademarkör, S100B, där höga värden indikerar allvarlig skada. S100B finns huvudsakligen i stödjevävnaden i hjärnan och i händelse av hjärnskada läcker proteinet ut via blod-hjärnbarriären. S100B finns i låga nivåer hos friska personer,  $\leq 0.10\mu\text{g/L}$ , och kommer snabbt ut i blodet efter en skallskada[2].

Det vanligaste sättet att ta ett blodprov är venöst<sup>1</sup>. För vanliga blodprov går man inte in i en artär utan denna metod används vid akuta fall då man vill kontrollera syresättningen i blodet (måste vara arteriellt<sup>2</sup>)[3] och vid akuta fall på saltbalansen. Vid behandling av hjärnskadepatienter går man ibland, helt oberoende av S100B, in i en artär och i sådant fall är det naturligt att ta blodprovet från artären. För barn i åldern 0-10 år kan det vara krävande, tidsmässigt och jobbigt, att ta venösa prov och därför tar man kapillärprov<sup>3</sup> som är snabbare och enklare[3].

### 1.1 Syfte med rapporten

För att avgöra ifall det finns skillnader i analysresultaten beroende på vilken typ av blodprov som tagits kommer variansanalyser av datamaterialet att göras.

Uppsatsen kommer behandla olika fall där jämförelser mellan kapillära och venösa och/eller arteriella samt mellan venösa och arteriella prov. Med hjälp av beräknade p-värden och konfidensintervall avgörs det om blodproven har samma förväntade proteinkoncentrationer. I rapporten behandlas endast skillnader mellan metoderna men inte eventuella skillnader som finns mellan olika patienter.

Prediktion av koncentrationer av S100B i venöst blod kommer att skattas från kända analysresultat från de kapillära blodproven. För att få en uppfattning om hur bra prediktionen är beräknas prediktionsfelet och korrelationen.

---

<sup>1</sup>Blodet tas från en ven, syrefattigt blod till hjärta-lungsystem

<sup>2</sup>Provet tas från en artär, syrerikt blod från hjärta-lungsystem

<sup>3</sup>Blodprov tas genom att man sticker hål på huden (ofta ett finger eller öronsibb) och sedan suger upp några droppar blod.

## 2 Datamaterialet

Datamaterialet<sup>4</sup> kommer från 52 patienter som vid sammanlagt 71 tillfällen lämnat blodprov; 44 patienter har testas vid endast ett tillfälle medan de övriga 8 har testas vid mellan 2 och 10 olika tillfällen, se appendix dataset. Vid varje tillfälle har två blodprov tagits: ett på kapillärt blod och ett ytterligare på antingen venöst eller arteriellt blod. Patienterna är numrerade från 1 till 54 (från ursprungsmaterialet har två patienter strukits, en p.g.a borttappat analysresultat och en p.g.a uppenbar förväxling). Koncentrationen av S100B är mätt i  $\mu\text{g/L}$ . Detta har ingen större betydelse för den statistiska analysen då vi endast är intresserade av differensen mellan metoderna för att undersöka om man kan använda sig av kapillära prov när man analyserar halten av S100B.

För att analysera datamaterialet görs samtliga beräkningar med hjälp av MATLAB.

## 3 Teori

Varje observation  $X_{ij}$ , där  $i$  är patient och  $j$  är numret för provtillfället för patienten, består av ett medelvärde  $\mu$ , en patientspecifik effekt  $y_i$  och en mättillfällespecifik effekt  $z_{ij}$  [4, sid 415-416] och [5].

$$\text{Kapillärt: } x_{ij}^{kap} = \mu^{kap} + y_i^{kap} + z_{ij}^{kap}$$

$$\text{Venöst: } x_{ij}^{ven} = \mu^{ven} + y_i^{ven} + z_{ij}^{ven}$$

$$\text{Arteriellt: } x_{ij}^{art} = \mu^{art} + y_i^{art} + z_{ij}^{art}$$

Då de förväntade värdena av den patientspecifika och den mättillfällespecifika effekten är lika med noll

$$E[y_i] = E[z_{ij}] = 0$$

blir de förväntade värdena av blodproven lika med dess medelvärden

$$\text{Kapillärt: } E[x_{ij}^{kap}] = \mu^{kap}$$

$$\text{Venöst: } E[x_{ij}^{ven}] = \mu^{ven}$$

$$\text{Arteriellt: } E[x_{ij}^{art}] = \mu^{art}$$

Ekvationen nedan är om man ser proven tagna venöst och arteriellt som en enhet. I det här fallet gör man ingen skillnad på om proven tagits venöst eller arteriellt.

---

<sup>4</sup>För tillgång till datamaterialet tackar jag dr Ramona Åstrand, Neurokirurgisk klinik, Rigshospitalet, Köpenhamn.

$$\text{Venöst/arteriellt: } x_{ij}^{ven/art} = \mu^{ven/art} + y_i^{ven/art} + z_{ij}^{ven/art}$$

Precis som ovan är de förväntade värdena av  $y_i$  och  $z_{ij}$  lika med noll.

Differensen mellan kapillärt och venöst får följande uttryck

$$\begin{aligned} d_{ij}^{kap:ven} &= x_{ij}^{kap} - x_{ij}^{ven} \\ &= (\mu^{kap} - \mu^{ven}) + (y_i^{kap} - y_i^{ven}) + (z_{ij}^{kap} - z_{ij}^{ven}) \\ &= \mu^{kap:ven} + y_i^{kap:ven} + z_{ij}^{kap:ven} \end{aligned}$$

Eftersom de förväntade värdena av den patientspecifika effekten och mät-tillfällesspecifika effekten är lika med noll blir den förväntade differensen mellan kapillära och venösa prov

$$E \left[ d_{ij}^{kap:ven} \right] = \mu^{kap:ven}$$

Analogt med ovan får man

$$E \left[ d_{ij}^{kap:art} \right] = \mu^{kap:art}$$

och

$$E \left[ d_{ij}^{kap:ven/art} \right] = \mu^{kap:ven/art}$$

För att jämföra blodprov tagna venöst och arteriellt beräknas först differenserna mellan venöst och kapillärt resp. arteriellt och kapillärt.

$$\begin{aligned} d_{ij}^{ven:kap} &= x_{ij}^{ven} - x_{ij}^{kap} = -d_{ij}^{kap:ven} \\ d_{ij}^{art:kap} &= x_{ij}^{art} - x_{ij}^{kap} = -d_{ij}^{kap:art} \end{aligned}$$

Om blodprov tagna antingen venöst eller arteriellt kommer ge samma resultat måste de förväntade differenserna mellan venöst och kapillärt respektive arteriellt och kapillärt vara lika.

$$E \left[ d_{ij}^{ven:kap} \right] = E \left[ d_{ij}^{art:kap} \right]$$

Frågor som är av intresse

1. Blir de förväntade differenserna noll?

$$\mu^{kap:ven} = 0$$

$$\mu^{kap:art} = 0$$

$$\mu^{kap:ven/art} = 0$$

För att avgöra detta kommer tre olika metoder att användas.

2. Blir det förväntade resultatet av ett prov lika oavsett om det tagits venöst eller arteriellt?

$$E[x_{ij}^{ven}] = E[x_{ij}^{art}]$$

I det här fallet kommer två olika metoder att användas.

3. Finns det ett beroende mellan differenserna och medelvärdet av blodproven tagna vid samma tillfälle?

$$d_{ij}^{kap:NN} \text{ vs } \frac{x_{ij}^{kap} + x_{ij}^{NN}}{2}$$

där NN är antingen venöst, arteriellt eller venöst och arteriellt.

4. Hur bra kan man prediktera S100B-halten i venösa blodprov om man känner till analysresultaten från kapillära blodprov?

Problemet för de två första punkterna är att beräkna medelvärdet för punktskattningen av differensen mellan provtagningsmetoderna och därefter bestämma ett p-värde samt ett 95% konfidensintervall.

### 3.1 Modellval 1 - Optimistisk metod

Den här modellen betraktar samtliga differenser som oberoende och utifrån detta kommer medelvärdet och standardavvikelsen att skattas. Den här modellen tar inte hänsyn till att en del av patienterna har mer än ett mättillfälle och överskattar precisionen av skattningen. Vi betraktar

$$Z = \frac{\bar{d}_{..}^{kap:NN} - \mu_0}{s/\sqrt{n}}$$

där  $\bar{d}_{..}^{kap:NN} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} d_{ij}^{kap:NN}$ ,  $n$  är totala antalet mättillfällen,  $k$  är antal patienter,  $n_i$  är antal mättillfällen för patient  $i$ ,  $s$  är standardavvikelsen hos samtliga differenser,  $\mu_0$  är noll enligt nollhypotesen. Då blir  $Z$  en standardiserad variabel, där ett antagandet görs om att observationerna är ungefär normalfördelade.

Från den beräknade  $Z$ -variabeln kan ett p-värde bestämmas och därigenom kan man avgöra ifall nollhypotesen kan förkastas eller ej. För att få en uppfattning om precisionen i punktskattningen av differensernas medelvärde beräknas ett konfidensintervall

$$\bar{d}_{..}^{kap:NN} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

där vi använder  $z_{\alpha/2} = 1.96$  för få ett 95% konfidensintervall.

### 3.2 Modellval 2 - Den försiktiga metoden

Första steget i den här modellen är att beräkna ett medelvärde av differenserna inom varje patient. Dessa medelvärden betraktas sedan som ett oberoende stickprov och utifrån detta kommer ett nytt medelvärde och standardavvikelsen att beräknas. Den försiktiga metoden ger en väntevärdesriktig skattning av  $\mu^{kap:NN}$ , se appendix 1, och en korrekt uppfattning om precisionen. Då beräkningar görs som om samtliga varianser inom patienterna är lika innebär det att man gör en precisionsförlust.

Vi betraktar

$$Z = \frac{\bar{d}^{kap:NN} - \mu_0}{s/\sqrt{k}}$$

där  $\bar{d}^{kap:NN} = \frac{1}{k} \sum_{i=1}^k \bar{d}_i$ ,  $\bar{d}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} d_{ij}$ ,  $k$  är antal patienter,  $s = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{d}_i - \bar{\bar{d}})^2}$ . Där vi gör ett antagande om att observationerna är ungefär normalfördelade.

Precis som för den optimistiska metoden kan ett konfidensintervall beräknas

$$\bar{d}^{kap:NN} \pm z_{\alpha/2} * \frac{s}{\sqrt{k}}$$

där vi använder  $z_{\alpha/2}=1.96$  för ett få ett 95% konfidensintervall.

### 3.3 Modellval 3 - ANOVA med slumpmässig effekt

För att få en korrekt och effektiv jämförelse mellan blodproven måste man använda information om varianser både inom patienter (varianserna i mätresultat för de patienter som har mer än ett provtagningsstillfälle) och mellan patienter.

Den totala variansen för differenserna mellan provtagnings-metoderna blir

$$\hat{\tau}_x^2 = \hat{\tau}_y^2 + \hat{\tau}_z^2$$

Variansen inom patienter,  $\tau_z^2$ , skattas med hjälp av ANOVA. Detta förutsätter ett antagande om att denna varians är densamma för alla patienter [5]. Vi betraktar

$$\hat{\tau}_z^2 = \sum \sum (d_{ij} - \bar{d}_i)^2 / (n_0 - k)$$

där  $n_0 = \sum_{i=1}^k n_i$  och  $k$  är antal patienter.

För att skatta den andra variansen  $\tau_y^2$ , som representerar heterogeniteten [5], skattas först variansen mellan patienterna med hjälp av ANOVA. Denna varians subtraheras med variansen inom patienter och sedan divideras med en faktor som beror på antalet mättillfällen inom varje patient



$$\hat{\tau}_y^2 = \frac{\sum n_i (\bar{d}_i - \bar{d})^2 / (k-1) - \sum \sum (d_{ij} - \bar{d}_i)^2 / (n-k)}{\frac{(\sum n_i)^2 - \sum n_i^2}{(k-1) \sum n_i}}$$

där  $n_i$  är antalet mättillfällen inom patient  $i$ ,  $n$  är totala antalet mättillfällen och  $k$  är antalet patienter. Se appendix 2 för mer utförlig teori.

För att beräkna medelfelet för differensernas medelvärde måste man vikta  $\hat{\tau}_y^2$  beroende på antalet mättillfällen patient  $i$  har och dividera  $\hat{\tau}_z^2$  med totala antalet mättillfällen. Medelfelet blir då

$$\hat{V}[\bar{d}_{..}] = \frac{\sum_i n_i^2}{n^2} \hat{\tau}_y^2 + \frac{1}{n} \hat{\tau}_z^2$$

vilket leder till följande uttryck för att bestämma ett p-värde

$$Z = \frac{\bar{d}_{..}^{kap:NN} - \mu_0}{\sqrt{\hat{V}[\bar{d}_{..}]}}$$

där  $Z$  är en standardiserad variabel (precis som tidigare ett antagande om normalfördelade variabler),  $n$  antal mättillfällen och  $\mu_0$  är noll enligt nollhypotesen.

Precis som för de tidigare metoderna kan ett konfidensintervall beräknas.

$$\bar{d}_{..}^{kap:NN} \pm z_{\alpha/2} \cdot \sqrt{\hat{V}[\bar{d}_{..}]}$$

där vi använder  $z_{\alpha/2} = 1.96$  för att få ett 95% konfidensintervall

### 3.4 Prediktion

För att prediktera koncentrationen av S100B i venösa och/eller arteriella blodprov från kapillära prov kommer en linjär regression att göras där givna värden på venösa och/eller arteriella prov plottas mot kapillära. Genom att använda sig av den räta linjens ekvation kan sedan värden på venösa prov predikteras från kända kapillära provtagningsresultat. Ett prediktionsfel kommer att beräknas enligt

$$\text{prediktionsfel} = \sqrt{\frac{\sum_i (y_i - \hat{y})^2}{n}}$$

där  $y$  och  $\hat{y}$  i detta fall är de observerade respektive skattade värdet på venösa och/eller arteriella blodprov och  $n$  är antal mättillfällen.

## 4 Resultat

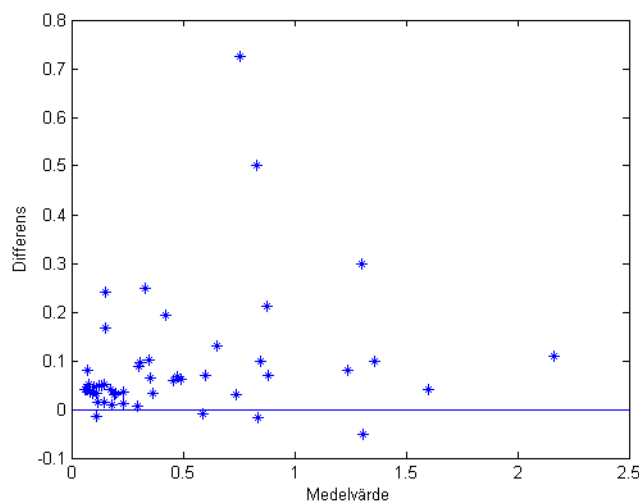
Följande fyra fall kommer att studeras.

- differensen mellan kapillärt och venöst
- differensen mellan kapillärt och arteriellt
- differensen mellan kapillärt och venöst/arteriellt
- differensen mellan venöst och arteriellt

För de tre första fallen kommer tre olika metoder användas för att analysera datamaterialet: en optimistisk metod, en försiktig metod och en metod med hjälp av ANOVA. För differensen mellan venöst och arteriellt kommer två olika metoder användas för att analysera datamaterialet. Dessa båda metoder kommer att diskuteras senare.

### 4.1 Kapillärt vs venöst

Datamaterialet som ligger till grund för jämförelsen mellan kapillära och venösa prov kommer från 49 olika mättillfällen uppdelade på 31 patienter. Endast 7 av patienterna har mer än ett provtagningstillfälle.



Figur 1: Differensen vs medelvärdet av kapillära och venösa provpar

I figur 1 kan man se att kapillära blodprov tenderar att ge en högre proteinkoncentration än venösa prov. Differensen i proteinkoncentrationen mellan provtagningsmetoderna är oberoende av hur hög koncentrationen är.

#### 4.1.1 Optimistisk metod

Medelvärden av differenserna mellan provtagningsmetoderna beräknades till  $0.0920 \mu\text{g/L}$  och standardavvikelsen till  $0.1316 \mu\text{g/L}$ . Då blir

$$Z = \frac{0.0920}{0.1316/\sqrt{49}} = 4.8945$$

vilket ger

$$\text{p-värde} = 4.9 \cdot 10^{-7}$$

och för ett 95% konfidensintervall blir skillnaden mellan kapillära och venösa prov

$$0.0920 \pm 1.96 \cdot \frac{0.1316}{\sqrt{49}}$$

dvs

$$I = [0.055 \ 0.129]$$

Med den optimistiska metoden kan nollhypotesen förkastas. Med denna metod får man resultatet att det finns skillnader mellan provtagningsmetoderna. Detta är precis vad man kan förvänta sig genom att studera figur 1 där man kunde konstatera att kapillära prov ger generellt en högre koncentration än venösa.

#### 4.1.2 Den försiktiga metoden

Medelvärden beräknades till  $0.1043 \mu\text{g/L}$  och standardavvikelsen till  $0.1420 \mu\text{g/L}$ . Då blir

$$Z = \frac{0.1043}{0.1420/\sqrt{31}} = 4.0886$$

vilket ger

$$\text{p-värde} = 2.2 \cdot 10^{-5}$$

och för ett 95% konfidensintervall blir skillnaden mellan kapillära och venösa prov

$$0.1043 \pm 1.96 \cdot \frac{0.1420}{\sqrt{31}}$$

dvs

$$I = [0.054 \ 0.154]$$

Med den försiktiga metoden kan nollhypotesen förkastas. Även med denna metod blir resultatet att det finns skillnader mellan provtagningsmetoderna.

Tabell 1: ANOVA kapillärt vs venöst

Variations- orsak	Kvadrat- summa	Frihets- grader	Medelkvadrat- summa
Mellan patient	0.6939	30	0.0231
Inom patient	0.1371	18	0.0076
Totalt	0.8310	48	

### 4.1.3 ANOVA med slumpmässig effekt

Följande storheter har beräknats:

$$\frac{(\sum n_i)^2 - \sum n_i^2}{(k-1) \sum n_i} = 1.5211$$

$$\hat{\tau}_y^2 = 0.0102$$

$$\hat{\tau}_z^2 = 0.0076$$

$$\bar{d}_{..} = 0.0920$$

$$\hat{V}[\bar{d}_{..}] = 8.5638 \cdot 10^{-4}$$

Då blir

$$Z = \frac{0.0920}{\sqrt{8.5638 \cdot 10^{-4}}} = 3.1438$$

vilket ger

$$\text{p-värde} = 8.3 \cdot 10^{-4}$$

och för ett 95% konfidensintervall blir skillnaden mellan kapillära och venösa prov

$$0.0920 \pm 1.96 \cdot \sqrt{8.5638 \cdot 10^{-4}}$$

dvs

$$I = [0.035 \ 0.149]$$

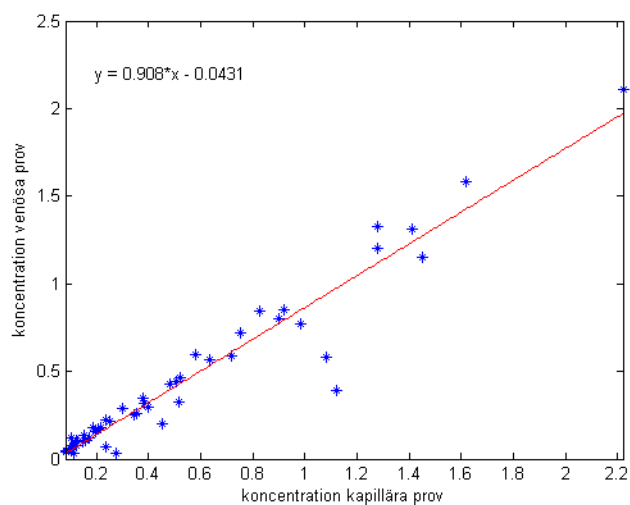
Denna metod visar precis som de övriga att det finns skillnader mellan kapillära och venösa prov.

### 4.1.4 Prediktion

För att prediktera koncentrationen i venösa prov från kapillära prov görs linjär regression där resultaten från venösa prov plottades mot kapillära prov se figur 2.

Den linjära regressionen beräknades till:

$$\text{venöst} = -0.0431 + 0.908 \cdot \text{kapillärt}$$

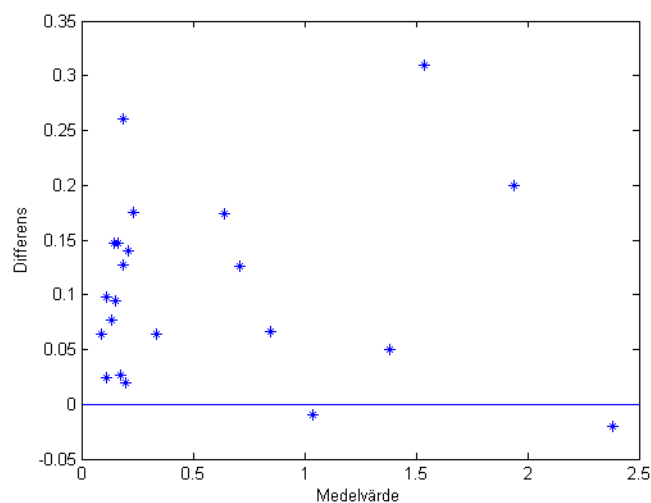


Figur 2: Linjär regression venösa vs kapillära prov

Ekvationen innebär att prediktionen av koncentrationen av S100B i venösa blodprov kommer bli lägre än i kapillära blodprov. Prediktionsfelet beräknades till 0.122. Korrelationen mellan venösa och kapillära prov beräknades till 0.963.

## 4.2 Kapillärt vs arteriellt

Datamaterialet kommer från 22 olika mättillfällen uppdelade på 21 patienter.



Figur 3: Differensen vs medelvärdet av kapillära och arteriella provpar

I figur 3 kan man se att proven tagna kapillärt tenderar att ge en högre proteinkoncentration. Differensen i koncentrationen mellan provtagningsmetoderna är oberoende av hur hög koncentrationen är.

#### 4.2.1 Optimistisk metod

Medelvärden för differenserna beräknades till  $0.1073 \mu\text{g/L}$  och standardavvikelsen till  $0.0839 \mu\text{g/L}$ . Då blir

$$Z = \frac{0.1073}{0.0839/\sqrt{22}} = 5.9989$$

vilket ger

$$\text{p-värde} = 9.9 \cdot 10^{-10}$$

och för ett 95% konfidensintervall blir skillnaden mellan kapillära och arteriella prov

$$0.1073 \pm 1.96 \cdot \frac{0.0839}{\sqrt{22}}$$

dvs

$$I = [0.072 \ 0.142]$$

Med den optimistiska metoden kan nollhypotesen förkastas. Resultatet av denna metod visar att det finns skillnader mellan provtagningsmetoderna. Detta är precis vad man kan förvänta sig genom att studera figur 3 där man kunde konstatera att kapillära prov ger generellt en högre koncentration än arteriella.

#### 4.2.2 Den försiktiga metoden

Medelvärdet beräknades till  $0.1067 \mu\text{g/L}$  och standardavvikelsen till  $0.0855 \mu\text{g/L}$ . Då blir

$$Z = \frac{0.1067}{0.0855/\sqrt{21}} = 5.7150$$

vilket ger

$$\text{p-värde} = 5.5 \cdot 10^{-9}$$

och för ett 95% konfidensintervall blir skillnaden mellan kapillära och arteriella prov

$$0.1067 \pm 1.96 \cdot \frac{0.0855}{\sqrt{21}}$$

dvs

$$I = [0.070 \ 0.143]$$

Med den försiktiga metoden kan nollhypotesen förkastas. Även med denna metod blir resultatet att det finns skillnader mellan provtagningsmetoderna. Konfidensintervallet för den här metoden och den optimistiska är nästan identiska vilket var förväntat då endast en patient hade mer än ett mät-tillfälle.

### 4.2.3 ANOVA med slumpmässig effekt

Tabell 2: ANOVA kapillärt vs arteriellt

Variations- orsak	Kvadrat- summa	Frihets- grader	Medelkvadrat- summa
Mellan patient	0.1465	20	0.0073
Inom patient	0.0014	1	0.0014
Totalt	0.1479	21	

Följande storheter har beräknats:

$$\frac{(\sum n_i)^2 - \sum n_i^2}{(k-1) \sum n_i} = 1.0455$$

$$\hat{\tau}_y^2 = 0.0057$$

$$\hat{\tau}_z^2 = 0.0014$$

$$\bar{d}_{..} = 0.1073$$

$$\hat{V}[\bar{d}_{..}] = 3.4477 \cdot 10^{-4}$$

då blir

$$Z = \frac{0.1073}{\sqrt{3.4477 \cdot 10^{-4}}} = 5.7797$$

vilket ger

$$\text{p-värde} = 3.7 \cdot 10^{-9}$$

och för ett 95% konfidensintervall blir skillnaden mellan kapillära och arteriella prov

$$0.1073 \pm 1.96 \cdot \sqrt{3.4477 \cdot 10^{-4}}$$

dvs

$$I = [0.071 \ 0.144]$$

Även med denna metod kan nollhypotesen förkastas. Med denna metod blir resultatet att det finns skillnader mellan provtagningsmetoderna. Konfidensintervallet för denna metod blev i stort sett identiskt med de föregående.

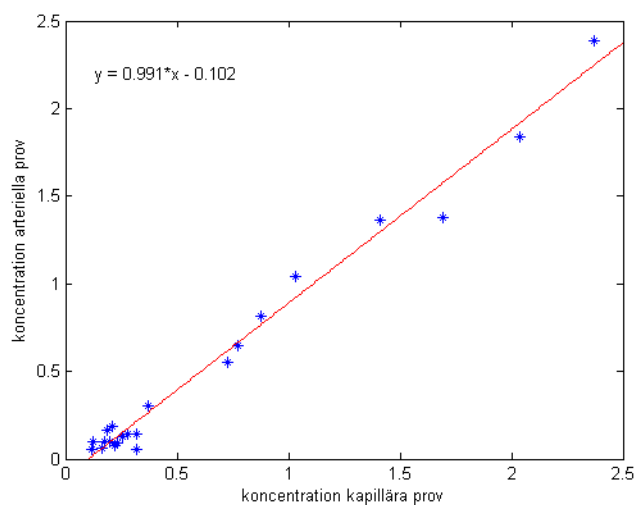
### 4.2.4 Prediktion

För att prediktera koncentrationen i arteriella prov från kapillära prov görs linjär regression precis som för prediktionen av venösa prov, se figur 4.

Den linjära regressionen beräknades till:

$$\text{arteriellt} = -0.102 + 0.991 \cdot \text{kapillärt}$$

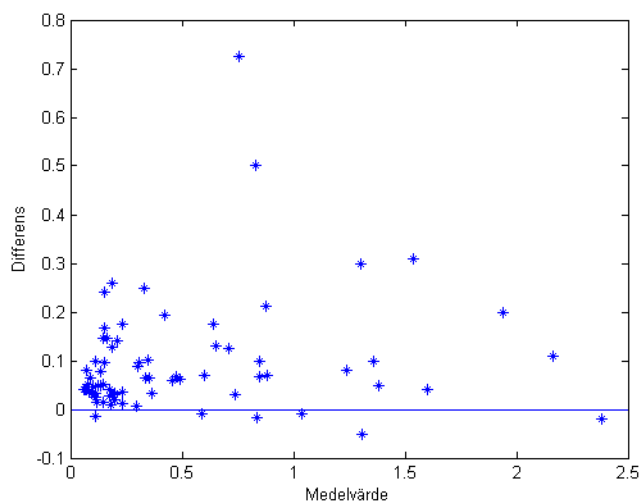
Ekvationen innebär att prediktionen av arteriella prov kommer ge en lägre proteinkoncentration än kapillära. Prediktionsfelet beräknades till 0.082. Korrelationen mellan kapillära och arteriella prov beräknades till 0.992.



Figur 4: Linjär regression arteriella vs kapillära prov

### 4.3 Kapillärt vs venöst/arteriellt

Datamaterialet kommer från 71 olika mättillfällen uppdelat på 52 patienter. Endast 8 av patienterna har mer än ett mättillfälle. I det här fallet görs ingen skillnad på om proven är tagna venöst eller arteriellt.



Figur 5: Differensen vs medelvärdet av kapillära och venösa/arteriella provpar

I figur 5 kan man se att proven tagna kapillärt tenderar att ha en högre



koncentration än prov tagna antingen venöst eller arteriellt. Differensen i proteinkoncentrationen mellan de olika provtagningsmetoderna är oberoende av den hur hög proteienkoncentrationen är.

#### 4.3.1 Optimistisk metod

Medelvärde för differenserna mellan provtagningsmetoderna beräknades till  $0.0967 \mu\text{g/L}$  och standardavvikelsen till  $0.1185 \mu\text{g/L}$ . Då blir

$$Z = \frac{0.0967}{0.1185/\sqrt{71}} = 6.8813$$

vilket ger

$$\text{p-värde} = 3.0 \cdot 10^{-12}$$

och för ett 95% konfidensintervall blir skillnaden mellan kapillära och venösa/arteriella prov

$$0.0967 \pm 1.96 \cdot \frac{0.1185}{\sqrt{71}}$$

dvs

$$I = [0.069 \ 0.124]$$

Med den optimistiska metoden kan nollhypotesen förkastas. Med denna metod blir resultatet att det finns skillnader mellan kapillära och venösa/arteriella prov. Detta är väntat då figur 5 visade att kapillära prov tenderar att ge en högre koncentration och då tidigare jämförelser i den här rapporten visade på skillnader mellan kapillära och venösa samt mellan kapillära och arteriella prov.

#### 4.3.2 Den försiktiga metoden

Medelvärde beräknades till  $0.1052 \mu\text{g/L}$  och standardavvikelsen till  $0.1214 \mu\text{g/L}$ . Då blir

$$Z = \frac{0.1052}{0.1214/\sqrt{52}} = 6.8813$$

vilket ger

$$\text{p-värde} = 2.0 \cdot 10^{-10}$$

och för ett 95% konfidensintervall blir skillnaden mellan kapillära och venösa/arteriella prov

$$0.1052 \pm 1.96 \cdot \frac{0.1214}{\sqrt{52}}$$

dvs

$$I = [0.072 \ 0.138]$$

Även med den försiktiga metoden kan nollhypotesen förkastas. Med denna metod blir resultatet att det finns skillnader mellan kapillära och venösa/arteriella prov.

### 4.3.3 ANOVA med slumpmässig effekt

Tabell 3: ANOVA kapillärt vs venöst/arteriellt

Variations- orsak	Kvadrat- summa	Frihets- grader	Medelkvadrat- summa
Mellan patient	0.8440	51	0.0165
Inom patient	0.1384	19	0.0073
Totalt	0.9824	70	

Följande storheter har beräknats:

$$\frac{(\sum n_i)^2 - \sum n_i^2}{(k-1) \sum n_i} = 1.0455$$

$$\hat{\tau}_y^2 = (0.0165 - 0.0073) / 1.34 = 0.0069$$

$$\hat{\tau}_z^2 = 0.0073$$

$$\bar{d}_{..} = 0.0967$$

$$\hat{V}[\bar{d}_{..}] = 3.6180 \cdot 10^{-4}$$

Då blir

$$Z = \frac{0.0967}{\sqrt{3.6180 \cdot 10^{-4}}} = 5.0863$$

vilket ger

$$\text{p-värde} = 1.8 \cdot 10^{-7}$$

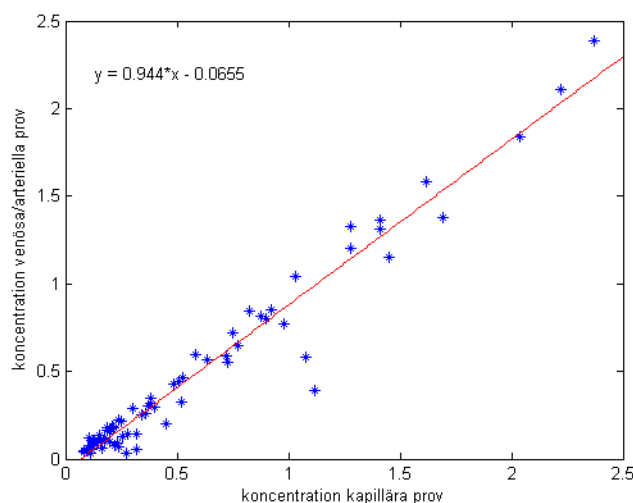
och för ett 95% konfidensintervall blir skillnaden mellan kapillära och venösa/arteriella prov

$$0.0967 \pm 1.96 \cdot \sqrt{3.6180 \cdot 10^{-4}}$$

dvs

$$I = [0.060 \ 0.134]$$

Precis som för de föregående metoderna kan man med den här metoden förkasta nollhypotesen. Med denna metod blir resultatet att det finns skillnader mellan kapillära och venösa/arteriella prov.



Figur 6: Linjär regression venösa/arteriella vs kapillära prov

#### 4.3.4 Prediktion

För att prediktera koncentrationen i venösa/arteriella prov från kapillära prov görs linjär regression precis som tidigare, se figur 6

Den linjära regressionen beräknades till:

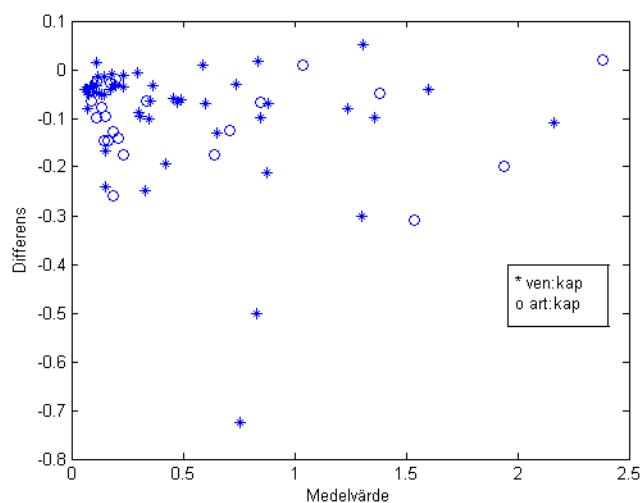
$$\text{venöst/arteriellt} = -0.0655 + 0.944 \cdot \text{kapillärt}$$

Ekvationen innebär att prediktionen av venösa/arteriella prov kommer ge en lägre proteinkoncentration än kapillära. Prediktionsfelet beräknades till 0.114. Korrelationen mellan kapillära och arteriella prov beräknades till 0.976.

#### 4.4 Venöst vs arteriellt

Då man inte tagit venöst och arteriellt prov samtidigt vid de olika mättillfällena är metoderna ovan inte användbara i detta fall. För att lösa problemet delar man in mättillfällena i två grupper, en grupp som innehåller differenserna i proteinkoncentration mellan venösa och kapillära blodprov och en grupp för differensen mellan arteriella och kapillära blodprov. Med hjälp av t-test kan sedan datamaterialet analyseras. För att avgöra om proteinkoncentrationen i prov tagna antingen venöst eller arteriellt kommer ge samma resultat kommer två olika metoder att användas.

I figur 7 kan man se att differenserna mellan venösa och kapillära prov respektive arteriella och kapillära prov inte skiljer sig nämnvärt åt. Vilket skulle indikera att venösa och arteriella blodprov kommer ge samma resultat.



Figur 7: Differensen vs medelvärdet av venösa:kapillära och arteriella:kapillära prov

#### 4.4.1 Oberoende

I detta fall kommer samtliga 71 differenser att betraktas som oberoende. I gruppen venöst-kapillärt finns det 49 mättilfällen och i gruppen arteriellt-kapillärt 22 mättilfällen. Dessa båda grupper jämförs sedan med hjälp av t-test.

Tabell 4: ANOVA venöst vs arteriellt, oberoende

Variations- orsak	Kvadrat- summa	Frihets- grader	Medelkvadrat- summa	F-värde	p-värde
Mellan klasser	0.0036	1	0.0036	0.2511	0.6179
Inom klasser	0.9788	69	0.0142		
Totalt	0.9824	70			

I tabell 4 kan man se att p-värdet beräknades till 0.62 vilket gör att nollhypotesen kan inte förkastas. Resultatet är att man inte kan påvisa någon statistisk säkerställd skillnad mellan venösa och arteriella prov med denna metod.

#### 4.4.2 Patientmedelvärde

Det första steget i den här metoden är att först beräkna medelvärdet för differenserna i proteinkoncentrationen för varje patient. Medelvärdena betraktas sedan som oberoende och beräkningar görs med ANOVA precis som

i förra metoden. I gruppen venöst-kapillärt finns 31 patienter och i gruppen arteriellt-kapillärt finns 21 patienter.

Tabell 5: ANOVA venöst vs arteriellt, patientmedelvärde

Variations- orsak	Kvadrat- summa	Frihets- grader	Medelkvadrat- summa	F-värde	p-värde
Mellan klasser	0.0001	1	0.0001	0.0049	0.9447
Inom klasser	0.7510	50	0.0150		
Totalt	0.7510	51			

I tabell 5 kan man se att p-värdet med denna metod beräknades till 0.95. Detta gör att nollhypotesen kan inte förkastas. Precis som för den föregående metoden blir resultatet att man inte kan påvisa någon statistiskt säkerställd skillnad mellan venösa och arteriella prov.

## 5 Slutsats och diskussion

Vid jämförelsen mellan kapillära och venösa prov kan man konstatera att det finns skillnader mellan dessa båda provtagningsmetoder, samtliga tre modeller visade att kapillära prov ger en högre koncentration av S100B. ANOVA med slumpmässig effekt och den försiktiga metoden ger ett bredare konfidensintervall än för den optimistiska metoden. Samtliga modeller visade dock liknande resultat. Prediktionen med linjära regression visade att prediktionsfelet blev relativt stort vilket kan innebära problem då man vill använda sig av kapillära prov istället för venösa. Prediktionen har också en begränsning då den inte är användbar vid låga koncentrationer för kapillära prov. Korrelationen mellan venösa och kapillära prov var däremot nära värdet ett.

Jämförelsen mellan kapillära och arteriella prov visade på skillnader, samtliga modeller visade även här att kapillär provtagning ger en högre koncentration. Samtliga modeller gav nästan samma konfidensintervall orsaken till detta är att materialet innehåller endast en patient som har mer än ett mättillfälle. Prediktionen visar även här att prediktionsfelet är stort och problem vid prediktion vid låga koncentrationer. Även här blev korrelationen nära ett.

Även mellan kapillära och antingen venösa eller arteriella prov blir resultatet att det finns skillnader mellan provtagningsmetoderna. Som konstaterats tidigare ger kapillära prov högre koncentrationer. Prediktionsfelet blev även i detta fall relativt stort.

Vid jämförelsen mellan venösa och arteriella prov kan man inte påvisa någon statistiskt säkerställd skillnad mellan provtagningsmetoderna. Detta innebär att om man väljer att ta proven venöst eller arteriellt spelar ingen roll för analysresultatet, utan valet av provtagningsmetod görs med som är mest praktiskt vid mättillfället.

## 6 Referenser

### Referenser

- [1] Altman D.G (1991), *Practical Statistics for Medical Research (1 ed)* London: Chapman & Hall
- [2] Undén J, Ingebrigtsen T, Romner B (2008), Handläggning av lättare skullskador med hjälp av ett blodprov. *Läkartidningen Nr 24-25*, 1846-1848
- [3] Åstrand Ramona, Rigshospitalet Köpenhamn Personlig kontakt
- [4] Lindgren B.W (1968), *Statistical Theory, (2 ed)* New York: Macmillan
- [5] Bland J.M and Altman D (2007), Agreement between methods of measurement with multiple observations per individual. *Journal of Biopharmaceutical Statistics 17*, 571-582.

## A Appendix -1

Bakgrund till den försiktiga metoden.

Betrakta följande observationer:  $x_1, \dots, x_n$  Då blir väntevärde och varians för patient  $i$ .

$$E[x_i] = \mu$$

$$V[x_i] = \sigma_i^2$$

Betrakta följande

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

vilket leder fram till följande

$$\begin{aligned} E[s^2/n] &= \frac{1}{n} E[s^2] \\ &= \frac{1}{n} \frac{1}{n-1} \left[ E \sum (x_i - \bar{x})^2 \right] \\ &= \frac{1}{n} \frac{1}{n-1} \left[ \sum E(x_i^2) - \sum E(\bar{x}^2) \right] \\ &= \frac{1}{n} \frac{1}{n-1} \left[ \sum (\sigma_i^2 + \mu_i^2) - \sum \left( \frac{\sigma_i^2}{n} + \mu_i^2 \right) \right] \\ &= \frac{1}{n} \frac{1}{n-1} (n-1) \frac{\sum \sigma_i^2}{n} \\ &= \frac{\sum \sigma_i^2}{n^2} = \text{Var}(\bar{x}) \end{aligned}$$

är sant även när  $x_1, \dots, x_n$  har varierande varianser.



## B Appendix - 2

Skattningarna av  $\tau_y^2$  och  $\tau_z^2$  för ANOVA med slumpmässig effekt.

Differenserna mellan de olika provtagningsmetoderna ges av

$$d_{ij} = \mu + y_i + z_{ij}$$

där

$$i = 1, \dots, k$$

$$j = 1, \dots, n_i$$

För väntevärdet och varians gäller

$$E[y_i] = E[z_{ij}] = 0$$

$$V[y_i] = \tau_y^2$$

$$V[z_i] = \tau_z^2$$

Medelvärdet för differenserna inom patient och samtliga mätilfällen beräknas så:

$$\bar{d}_i = \frac{1}{n_i} \sum_j d_{ij}$$

$$\bar{d}_{..} = \frac{1}{n} \sum_i \sum_j d_{ij}$$

Betrakta följande kvadratsummor

$$Q = \sum_i \sum_j (d_{ij} - \bar{d}_{..})^2$$

$$Q_0 = \sum_i \sum_j (d_{ij} - \bar{d}_i)^2$$

$$Q_1 = \sum_i \sum_j (\bar{d}_i - \bar{d}_{..})^2 = \sum_i n_i (\bar{d}_i - \bar{d}_{..})^2$$

Q kan också bestämmas genom

$$Q = Q_0 + Q_1$$

Att

$$E[Q_0] = (n - k) \tau_z^2$$

då  $y_i$  försvinner ur  $d_{ij} - \bar{d}_i$ .

Beräkning av  $E[Q_1]$  kräver en del förberedelse

$$V[\bar{d}_i] = V[y_i + \bar{z}_i] = \tau_y^2 + \frac{1}{n_i} \tau_z^2$$

$$V[\bar{d}_{..}] = V\left[\frac{1}{n} \sum_i n_i \bar{d}_i\right]$$

$$\begin{aligned}
&= \frac{1}{n^2} \sum_i n_i^2 V[\bar{d}_i] \\
&= \frac{1}{n^2} \sum_i n_i^2 (\tau_y^2 + \frac{1}{n_i} \tau_z^2) \\
&= \frac{1}{n^2} \sum_i n_i^2 \tau_y^2 + \frac{1}{n} \tau_z^2
\end{aligned}$$

Kovariansen för  $[d_{i.}, \bar{d}_{..}]$  ges av

$$\begin{aligned}
C[\bar{d}_i., \bar{d}_{..}] &= C[\bar{d}_i., \frac{1}{n} \sum n_k \bar{d}_k.] \\
&= C[\bar{d}_i., \frac{1}{n} n_i \bar{d}_i.] \\
&= \frac{n_i}{n} V[\bar{d}_i.] \\
&= \frac{n_i}{n} (\tau_y^2 + \tau_z^2) \\
&= \frac{n_i}{n} \tau_y^2 + \frac{1}{n} \tau_z^2
\end{aligned}$$

Beräkning av  $E[Q_1]$  ges av följande

$$\begin{aligned}
E[Q_1] &= E[\sum_i n_i (\bar{d}_i. - \bar{d}_{..})^2] \\
&= \sum_i n_i V[\bar{d}_i. - \bar{d}_{..}] \\
&= \sum_i n_i (V[\bar{d}_i.] + V[\bar{d}_{..}] - 2C[\bar{d}_i., \bar{d}_{..}])
\end{aligned}$$

Med insatta värden på varianser och kovariansen får man följande (samtliga summor går från  $i$  till  $k$ )

$$\begin{aligned}
E[Q_1] &= \left( \sum n_i + \frac{\sum n_i \sum n_i^2}{n^2} - 2 \frac{\sum n_i n_i}{n} \right) \tau_y^2 + \left( \sum \frac{n_i}{n_i} + \frac{\sum n_i}{n} - 2 \sum n_i \frac{1}{n} \right) \tau_z^2 \\
&= \left( \sum n_i + \frac{\sum n_i^2}{\sum n_i} - 2 \frac{\sum n_i^2}{\sum n_i} \right) \tau_y^2 + \left( \sum \frac{n_i}{n_i} + \frac{\sum n_i}{n} - 2 \frac{\sum n_i}{n} \right) \tau_z^2 \\
&= \left( \frac{(\sum n_i)^2 - \sum n_i^2}{\sum n_i} \right) \tau_y^2 + (k-1) \tau_z^2
\end{aligned}$$

Vi får följande

$$\hat{\tau}_y^2 = \frac{Q_1 / (k-1) - \hat{\tau}_z^2}{\frac{(\sum n_i)^2 - \sum n_i^2}{(k-1) \sum n_i}}$$

som leder fram till

$$\hat{\tau}_y^2 = \frac{\sum_i n_i (\bar{d}_i. - \bar{d}_{..})^2 / (k-1) - \sum_i \sum_j (d_{ij} - \bar{d}_i.)^2 / (n-k)}{\frac{(\sum n_i)^2 - \sum n_i^2}{(k-1) \sum n_i}}$$

## C Appendix - Dataset

Patient	Obs.nr/Patient	Antal obs/Patient	Konc. kapillärt	Konc. venöst	Konc. arteriellt
1	1	1	0.299	0.292	.
2	1	1	0.250	0.215	.
3	1	1	0.082	0.041	.
4	1	1	0.720	0.590	.
5	1	2	1.410	1.310	.
5	2	2	0.128	0.096	.
6	1	3	0.485	0.427	.
6	2	3	0.343	0.255	.
6	3	3	0.086	0.044	.
7	1	10	1.620	1.580	.
7	2	10	1.280	1.200	.
7	3	10	0.920	0.850	.
7	4	10	0.524	0.461	.
7	5	10	0.188	0.178	.
7	6	10	0.216	0.182	.
7	7	10	0.197	0.157	.
7	8	10	0.148	0.100	.
7	9	10	0.159	0.109	.
7	10	10	0.151	0.137	.
8	1	4	1.080	0.578	.
8	2	4	0.636	0.566	.
8	3	4	0.509	0.441	.
8	4	4	0.519	0.326	.
9	1	1	0.125	0.079	.
10	1	2	0.827	0.845	.
10	2	2	0.583	0.593	.
11	1	1	0.772	.	0.646
12	1	1	0.235	.	0.088
14	1	1	1.450	1.150	.
15	1	1	2.220	2.110	.
16	1	1	0.398	0.297	.
17	1	2	1.280	1.330	.
17	2	2	0.752	0.721	.
18	1	2	0.384	0.320	.
18	2	2	0.357	0.262	.
19	1	1	1.120	0.394	.
20	1	1	0.111	0.078	.
21	1	1	0.380	0.346	.
22	1	1	0.106	0.120	.
23	1	1	0.175	0.123	.
24	1	1	0.237	0.070	.
25	1	1	0.098	0.063	.
26	1	1	0.453	0.203	.
27	1	1	0.275	0.034	.
28	1	1	0.237	0.225	.
29	1	1	2.370	.	2.390
30	1	1	0.900	0.801	.
31	1	1	0.205	0.175	.

Patient	Obs.nr/Patient	Antal obs/Patient	Konc. kapillärt	Konc. venöst	Konc. arteriellt
32	1	1	0.879	.	0.813
33	1	1	1.690	.	1.380
34	1	1	1.410	.	1.360
35	1	1	0.189	.	0.162
36	1	1	0.163	.	0.065
37	1	1	0.113	0.033	.
39	1	1	0.982	0.770	.
40	1	1	0.119	.	0.055
41	1	1	0.318	.	0.058
42	1	1	0.253	.	0.126
43	1	2	0.197	.	0.102
43	2	2	0.222	.	0.075
44	1	1	0.370	.	0.306
45	1	1	0.318	.	0.143
46	1	1	0.122	.	0.098
47	1	1	1.030	.	1.040
48	1	1	0.728	.	0.554
49	1	1	2.040	.	1.840
50	1	1	0.174	.	0.097
51	1	1	0.123	0.108	.
52	1	1	0.106	0.054	.
53	1	1	0.280	.	0.140
54	1	1	0.210	.	0.190

Kommentar till dataset. Koncentrationen är mätt i  $\mu\text{g/L}$ . En punkt i datasetet betyder att man inte har något analysresultat av blodprovsmetoden. Patient 13 och 38 har strukits p.g.a. förväxling och p.g.a ett uppenbart fel.