

Magisteruppsats i Statistik

**En retrospektiv studie av vilka  
patientgrupper som erhåller  
insulinpump**

Jonna Alnervik & Peter Nord Andersson



## Upphovsrätt

Detta dokument hålls tillgängligt på Internet – eller dess framtida ersättare – under 25 år från publiceringsdatum under förutsättning att inga extraordinära omständigheter uppstår. Tillgång till dokumentet innebär tillstånd för var och en att läsa, ladda ner, skriva ut enstaka kopior för enskilt bruk och att använda det oförändrat för ickekommersiell forskning och för undervisning. Överföring av upphovsrätten vid en senare tidpunkt kan inte upphäva detta tillstånd. All annan användning av dokumentet kräver upphovsmannens medgivande. För att garantera äktheten, säkerheten och tillgängligheten finns lösningar av teknisk och administrativ art. Upphovsmannens ideella rätt innefattar rätt att bli nämnd som upphovsman i den omfattning som god sed kräver vid användning av dokumentet på ovan beskrivna sätt samt skydd mot att dokumentet ändras eller presenteras i sådan form eller i sådant sammanhang som är kränkande för upphovsmannens litterära eller konstnärliga anseende eller egenart. För ytterligare information om Linköping University Electronic Press se förlagets hemsida <http://www.ep.liu.se/>.

## Copyright

The publishers will keep this document online on the Internet – or its possible replacement – for a period of 25 years starting from the date of publication barring exceptional circumstances. The online availability of the document implies permanent permission for anyone to read, to download, or to print out single copies for his/hers own use and to use it unchanged for non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional upon the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security and accessibility. According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement. For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its www homepage: <http://www.ep.liu.se/>.

© Jonna Alnervik och Peter Nord Andersson



## **Abstract**

### **Aim**

To investigate differences in the availability of continuous subcutaneous insulin infusion (CSII) between different patient groups as well as what causes a switch to CSII.

### **Method**

Data from 7224 individuals with Type 1 diabetes at ten different care units were analyzed to investigate the impact of gender, impairment, long-term blood glucose, dose of insulin, diabetes duration and age. A comparison between groups of patients was carried out using logistic regression as a cross-sectional study and Cox regression to investigate what precedes switching to CSII.

### **Results**

By using logistic regression the difference between patients currently using and not using CSII is shown. Cox regression furthermore handles time dependent variables and thus provides information on what precedes the switch to CSII. These analyses yielded similar results concerning variables fixed over time. Women use CSII more than men do and the proportion of CSII users differs in different care units. The cross-sectional study proves that old age reduces the likelihood of using CSII. This is confirmed by the time-dependent study that showed how the probability of switching to CSII is considerably lower at an older age. Long-term blood sugar also has a clear effect on the probability of a switch to CSII in which a high value gives a high probability of switching to CSII.

### **Conclusions**

In the current situation, there are differences in the proportion of CSII users between different groups of patients and differences also exist in the different groups' propensity to switch from other treatments to CSII. Depending on the patients' gender, impairment, long-term blood glucose, dose of insulin, diabetes duration and age, the probability of switching to CSII also varies.



## **Sammanfattning**

### **Målsättning**

Att utreda skillnader i tillgänglighet till insulinpump mellan olika patientgrupper samt vad som orsakar ett byte till insulinpump.

### **Metod**

Data från 7224 individer med typ 1 diabetes vid tio olika vårdenheter analyserades för att utreda effekterna av njurfunktion, kön, långtidsblodsocker, insulindos, diabetesduration samt ålder. Jämförelsen mellan patientgrupper utfördes med logistisk regression som en tvärsnittsstudie och Cox-regression för att utreda vad som föregått ett byte till pump.

### **Resultat**

Genom logistisk regression erhöles en bild av hur skillnader mellan patienter som använder insulinpump och patienter som inte gör det ser ut i dagsläget. Cox-regressionen tar med ett tidsperspektiv och ger på så sätt svar på vad som föregått ett byte till insulinpump. Dessa analyser gav liknande resultat gällande variabler konstanta över tiden. Kvinnor använder pump i större utsträckning än män och andelen pumpanvändare skiljer sig åt vid olika vårdenheter. I dagsläget visar sig hög ålder sänka sannolikheten att använda insulinpump, vilket bekräftas vid den tidsberoende studien som visade hur sannolikheten att byta till pump är avsevärt lägre vid hög ålder. Långtidsblodsockret har också tydlig effekt på sannolikheten att gå över till pump där ett högt långtidsblodsocker medför hög sannolikhet att byta till insulinpump.

### **Slutsatser**

I dagsläget finns det skillnader i andelen insulinpumpanvändare mellan olika patientgrupper och skillnader finns även i de olika gruppernas benägenhet att byta från andra insulinbehandlingar till insulinpump. Beroende av patienters njurfunktion, kön, långtidsblodsocker, insulindos, diabetesduration och ålder har dessa olika sannolikheter att byta till insulinpump.





## Förord

Denna uppsats är skriven inom ramen för Magisterprogrammet för statistik och dataanalys. Uppdragsgivare har varit Marcus Lind, diabetesläkare vid Uddevalla sjukhus.

Vi vill rikta ett stort tack till Marcus som varit mycket engagerad och hjälpt oss då vi haft frågor kring diabetes. Vi vill även tacka hans kollega Anders Odén, professor i biostatistik vid Chalmers universitet, som låtit oss använda en typ av Poissonmodell som han utvecklat för liknande data. Modellen är väl beprövad i liknande studier och används i denna studie för att validera övriga analyser. Samtidigt vill vi tacka för att vi fått använda programvaran Diab-Base och databasen den är kopplad till samt den hjälp vi fått kring data av Martin Fahlén och Jan Ramebäck.

Sist men inte minst vill vi tacka John Carstensen som varit vår handledare och fått agera bollplank kring statistiska metodval.

Vi hoppas att resultaten vi kommit fram till kommer att vara till användning i den fortsatta diabetesforskningen.

Linköping, maj 2010

Jonna Alnervik och Peter Nord Andersson



# Innehållsförteckning

<b>1 Inledning .....</b>	<b>1</b>
1.1 Bakgrund.....	1
1.1.1 Diabetes .....	1
1.1.2 HbA <sub>1c</sub> .....	2
1.1.3 Blodsockerkontroll .....	2
1.2 Syfte och frågeställning .....	4
<b>2 Material och Metod.....</b>	<b>5</b>
2.1 Material.....	5
2.2 Variabler.....	9
2.2.1 Konstanta variabler .....	10
2.2.2 Tidsberoende variabler.....	11
2.2.3 Kategorisering av kontinuerliga variabler .....	13
2.2.4 Val av tidsaxel till Cox-regression.....	13
2.3 Definition av population och rensning av data .....	14
2.4 Metod .....	15
2.4.1 Hypotesprövning .....	15
2.4.2 Logistisk regression .....	16
2.4.3 Cox-regression .....	17
2.4.4 Icke proportionella hasarder .....	20
<b>3 Resultat .....</b>	<b>21</b>
3.1 Logistisk regression .....	21
3.1.1 Val av modell till logistisk regression .....	21
3.1.2 Slutgiltig modell logistisk regression .....	22
3.2 Cox-regression .....	27
3.2.1 Val av modell till Cox-regression.....	27

3.2.2 Kontroll av log-linjärt beroende .....	28
3.2.3 Kategorisering av variabler .....	29
3.2.4 Kontroll av proportionalitetsantagandet .....	29
3.2.5 Stratifiering .....	30
3.2.6 Slutgiltig modell Cox-regression .....	30
3.2.7 Ytterligare analys av HbA <sub>1c</sub> -hasarden .....	33
<b>4 Slutsatser och metoddiskussion .....</b>	<b>35</b>
4.1 Metoddiskussion .....	35
4.1.1 Logistisk regression .....	35
4.1.2 Cox-regression .....	35
4.2 Slutsatser .....	37
4.3 Förslag till vidare studier.....	38
<b>Referenser .....</b>	<b>39</b>
<b>Bilagor</b>	
Bilaga 1 – Logistisk regression, test av interaktioner.....	I
Bilaga 2 – Cox-regression, kontinuerliga variabler .....	IV
Bilaga 3 – Tidsberoende Cox-Regression (Non-proportional), SAS-kod.....	V
Bilaga 4 – En speciell typ av Poissonregression .....	VIII

## Figurförteckning

Figur 1. Andel pumpanvändare bland diabetespatienterna grupperat på åldersintervall .....	5
Figur 2. Andel pumpanvändare bland diabetespatienterna grupperat på vårdenhet .....	6
Figur 3. Andel pumpanvändare bland diabetespatienterna grupperat på kön.....	6
Figur 4. Andel patienter med insulinpump vid de olika vårdenheterna .....	8
Figur 5. Andel av besöken gjorda av icke pumpanvändare som resulterar i utdelning av insulinpump .....	9
Figur 6. Andel blivande insulinpumpanvändare/ingångsnivå på E/kg.....	12
Figur 7. Oddsquoter för vårdenhet .....	25
Figur 8. Oddsquoter för ålder .....	25
Figur 9. Oddsquoter för diabetesduration .....	26
Figur 10. Andel patienter i varje grupp av HbA <sub>1c</sub> som har gått över till insulinpump .....	34

## Tabellförteckning

Tabell 1. Korrelationer mellan de kontinuerliga variabler som används i den logistiska regressionen .....	7
Tabell 2. Antal patienter i studien.....	11
Tabell 3. P-värden för variablerna i den grundläggande logistiska regressionen.....	21
Tabell 4. Walds test av kovariaterna i den logistiska regressionen .....	22
Tabell 5. Logistisk regression, oddsquoter och konfidensintervall .....	24
Tabell 6. P-värden för de grundläggande variablerna i Cox-regression.....	28
Tabell 7. Kontroll för icke log-linjärt beroende.....	28
Tabell 8. Kontroll för proportionalitet .....	29
Tabell 9. Stratifiering .....	30
Tabell 10. Walds test av kovariaterna i den slutgiltiga Cox-regressionen.....	31
Tabell 11. Cox-regression, hazardquoter och konfidensintervall .....	32
Tabell 12. Parameterskattningar för HbA <sub>1c</sub> och dess kvadrat .....	33



# 1 Inledning

## 1.1 Bakgrund

Ett av huvudmålen inom diabetesvården är att minska risken för följsjukdomar i form av ögon-, njur-, nervskador, stroke och hjärtinfarkter. Risken för följsjukdomar hos diabetespatienter ökar vid en högre blodsockernivå, som vanligtvis beskrivs av  $HbA_{1c}$ . För patienter med typ 1 diabetes sker insulinbehandling antingen med multipla dagliga insulininjektioner (MDI) eller insulinpump. När byte i klinisk praktik sker från MDI till insulinpump förbättras blodsockernivån beskrivet av lägre  $HbA_{1c}$ , en sänkning med i genomsnitt 0,4 procentenheter (Fahlén, Eliasson, & Odén, 2005). Om fler patienter med typ 1 diabetes behandlas med insulinpump skulle flera fall av ögon-, njur-, nervskador (DCCT Study Group, 1993), amputationer, strokes och hjärtinfarkter (Nathan, o.a., 2005), (Stratton, o.a., 2000) kunna undvikas varje år.

### 1.1.1 Diabetes

Diabetes mellitus, vardagligt kallad diabetes, finns i flera olika typer vilka alla innebär att blodsockernivån är för hög. Typ 1 diabetes medför att kroppens egen insulinproduktion helt eller nästan helt upphör. Typ 2 diabetes kan yttra sig på olika sätt för olika patienter. Patienterna har fortfarande viss insulinproduktion kvar men vissa har det inte i tillräckligt stor utsträckning för att kunna reglera kroppens blodsockernivåer medan andra fått en ökad motståndskraft mot insulin. Typ 2 diabetes behandlas ofta med tabletter som ökar kroppens känslighet för insulin eller ökar frisättningen av insulin från bukspottskörteln. En variant av typ 1 diabetes är LADA (Latent Autoimmune Diabetes in the Adult) där insjuknandet sker i vuxen ålder till skillnad från vanlig typ 1 diabetes där insjuknandet i regel sker före 35 års ålder (Hanås, 2008).

### **1.1.2 HbA<sub>1c</sub>**

HbA<sub>1c</sub> är ett mått på medelblodssockret under de två till tre senaste månaderna. Hemoglobin är ett ämne i den röda blodkroppen (erythrocyten) som transporterar syre från lungorna ut i vävnaderna. En röd blodkropp har en livstid på 120 dagar och beroende på hur högt blodssockret är binds olika mängd glykos till hemoglobinet under erythrocytens livstid (Hanås, 2008). HbA<sub>1c</sub> är måttet på hur många procent av hemoglobinet i blodets röda blodkroppar som är bundet till glukos.

Den svenska standardens mål för HbA<sub>1c</sub>-värdet är att barn och ungdomar med diabetes ska ligga mellan 6,0 och 6,5 procent och vuxnas värde under 6,0 procent (Hanås, 2008). Detta kan jämföras med normalvärdet för en frisk person som ligger mellan 3,6 och 5,0 procent. Undersökningar visar att patienter med ett HbA<sub>1c</sub> på 6 procent har en avsevärt minskad risk för långsiktiga komplikationer (DCCT Study Group, 1993), (Reichard, 1995). Vid ett HbA<sub>1c</sub>-värde över 8 procent är det hög risk för följsjukdomar som ögon-, njur- och nervskador, amputationer, hjärtinfarkter och strokes.

### **1.1.3 Blodsockerkontroll**

Insulin är det hormon som gör att cellerna får energi i form av glukos. Har kroppen nedsatt insulinproduktion stannar glukosen kvar i blodet vilket leder till högt blodsocker. Hos en frisk person insöndras det en liten mängd insulin i blodet dygnet runt, efter måltider insöndras extra insulin för att hantera glukosen som tillkommit.

För patienter med typ 1 diabetes för vilka insulinproduktionen har upphört helt, sker behandling med insulininjektioner eller med insulinpump, även kallad CSII (Continuous Subcutaneous Insulin Infusion). Insulinpumpen utsöndrar insulin på ett sätt som ska efterlikna en frisk bukspottkörtels funktion för att hålla blodssockret på en bra nivå. Insulinpumpen avger en basdos i intervaller dygnet runt, och användaren ger extra insulin med hjälp av pumpen i samband med måltider.

När multipla dagliga insulininjektioner (MDI) används tar patienten en eller två dagliga injektioner med basinsulin för att täcka den grundnivå av insulin som kroppen alltid behöver. Därutöver tas insulininjektioner med kortverkande insulin till måltider för att ta hand om blodssockret som stiger i samband med måltider. I en insulinpump finns enbart kortverkande insulin där en regelbunden utsöndring från pumpen sker som utgör



basnivån. En potentiell fördel med insulinpump jämfört med MDI är att basnivån kan justeras för olika tider på dygnet och vid olika aktiviteter. Det behövs till exempel en lägre nivå av basinsulin i samband med fysisk aktivitet och vanligtvis behövs en högre nivå på efternatten än på förnatten (Hanås, 2008). Vid MDI kan däremot basnivån inte ändras då man ger en injektion som sedan varar 15 – 24 timmar beroende på typ av basinsulin.

Ett normalt blodsocker för en frisk person är ca 4,0 mmol/l före måltider och en lätt stegring sker vanligtvis en kortare period därefter (Hanås, 2008). Vid diabetes strävar man att nå så normala blodsockernivåer som möjligt och HbA1c sjunker ju lägre blodsocker som föreligger. Behandlingen är dock en svår balansgång mellan att inte ha alltför höga blodsockernivåer då risken för följsjukdomar ökar, men å andra sidan inte allt för låga då i värsta fall medvetlöshet kan inträda. Med en känning menar man att patienten känner av att blodsockret är lågt, vanliga symtom är till exempel hjärtklappning och svettningar vilket många patienter upplever som obehagligt. Vid en blodsockernivå på 2,0-3,0 mmol/l fungerar hjärnan sämre och man reagerar till exempel inte lika bra i trafiken. Vid ännu lägre nivåer kan medvetandet påverkas, det finns därför flera anledningar till att låga blodsockervärden (hypoglykemier) bör undvikas. Insulinpump minskar i regel risken för svåra känningar (Bode, Steed, & Davidson, 1994), (Dahl-Jørgensen, o.a., 1986), (Ludvigsson, Hermansson, Häger, Kernell, & Nordenskjöld, 1988).

## 1.2 Syfte och frågeställning

Syftet med denna studie är att undersöka hur tillgängligheten till insulinpumpbehandling skiljer sig åt utifrån kön, ålder, HbA<sub>1c</sub>-nivå, eventuell följsjukdoms svårighetsgrad, insulindos, diabetesduration och vårdenhet. Det har tidigare visats att användande av insulinpump har en större effekt vid höga HbA<sub>1c</sub>-värden men i övrigt samma effekt i olika patientgrupper. Kan det även visas att tillgängligheten till insulinpumpbehandling skiljer sig åt mellan olika patientgrupper tyder det på att tillgängligheten till insulinpump borde öka för gruppen med mindre andel pumpanvändare. Syftet leder till frågeställningarna:

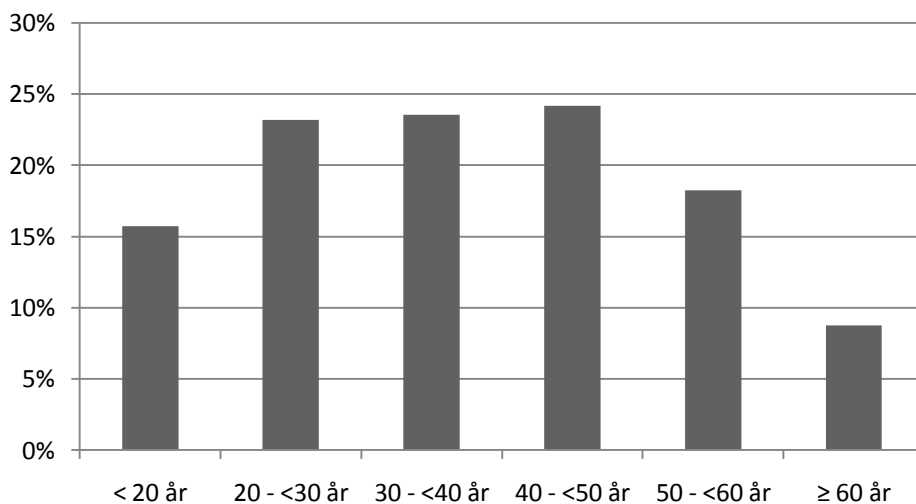
- Finns det några skillnader mellan de patienter som har insulinpump jämfört med de patienter som inte har pump i avseende på kön, ålder, diabetesduration och vårdenhet?
- Hur skiljer sig tillgängligheten till insulinpumpbehandling utifrån kön, ålder, HbA<sub>1c</sub>-nivå, grad av följsjukdom, insulindos, diabetesduration och vårdenhet?

## 2 Material och Metod

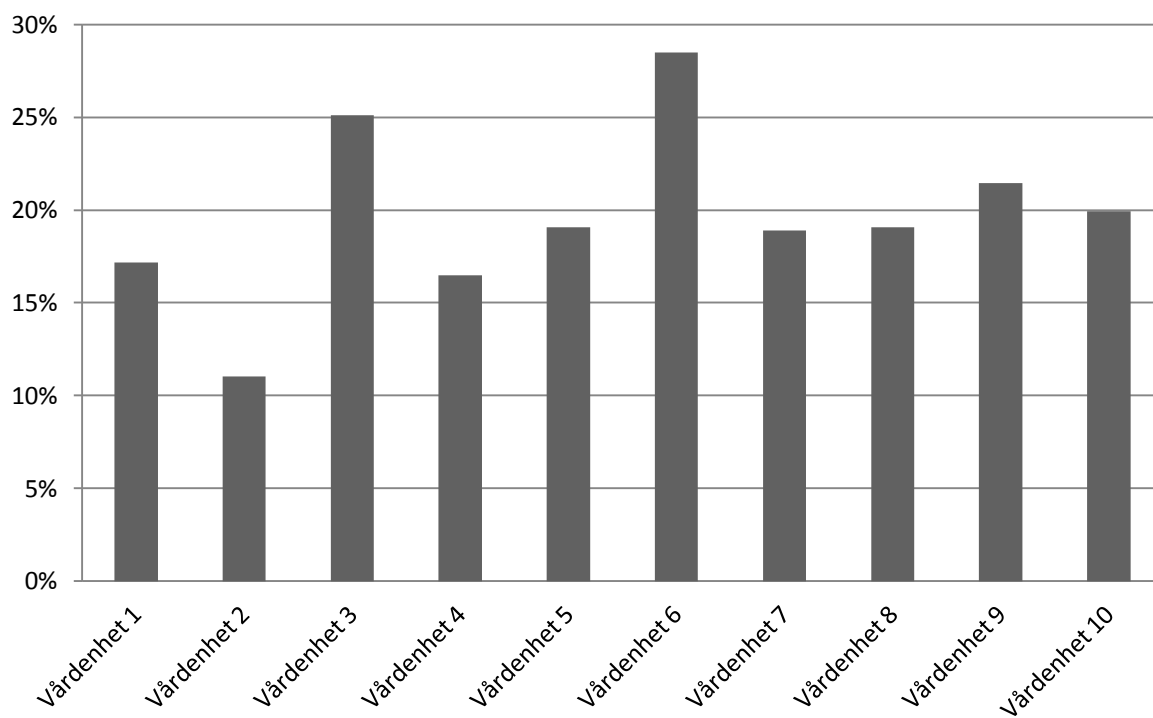
### 2.1 Material

Datamaterialet som ligger till grund för denna uppsats kommer från en databas i vilken sjuksköterskor och läkare för in diabetespatienters uppgifter med hjälp av Journalias programvara Diab-Base. Journalia är ett företag som sedan början av 90-talet utvecklat specialistsystem för vissa områden inom sjukvården, till exempel hjärt- och diabetesvården. Diab-Base används bland annat för att skapa journaler, leverera information till patienter samt skapa underlag för nationella register. Materialet som analyseras består av alla patienter med typ 1 diabetes från de tio vårdenheter som idag använder journalsystemet och det består dels av individuppgifter och dels av uppgifter från alla besök som gjorts under tiden patientens vårdenhet varit med i systemet. Besöken sträcker sig mellan fyra och femton år tillbaka i tiden.

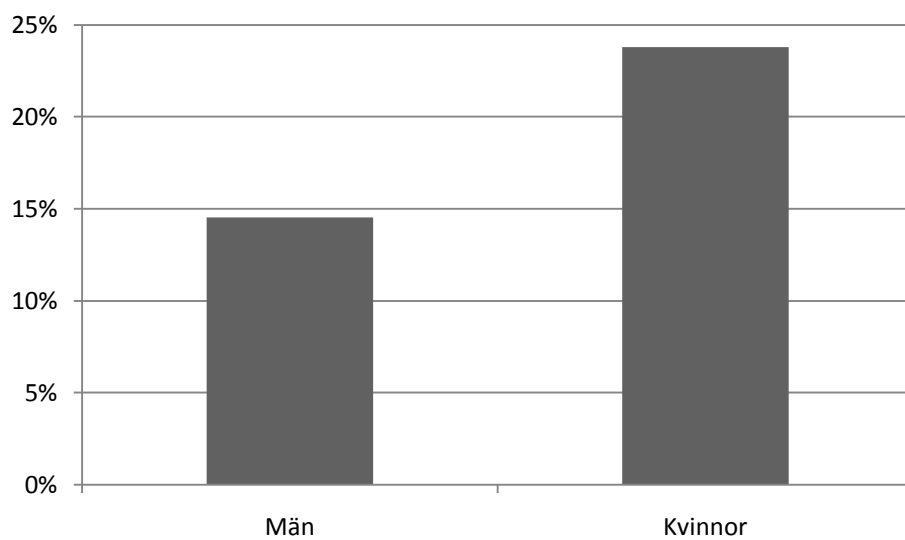
Figur 1, 2 och 3 visar andelen insulinpump användare inom varje grupp för variablerna Ålder, Vårdenhet och Kön. Figurerna indikerar att det finns skillnader i andelen insulinpump användare beroende på vilket värde variabeln antar.



Figur 1. Andel insulinpump användare bland diabetespatienterna grupperat på åldersintervall



Figur 2. Andel pumpanvändare bland diabetespatienterna grupperat på vårdenhet



Figur 3. Andel pumpanvändare bland diabetespatienterna grupperat på kön

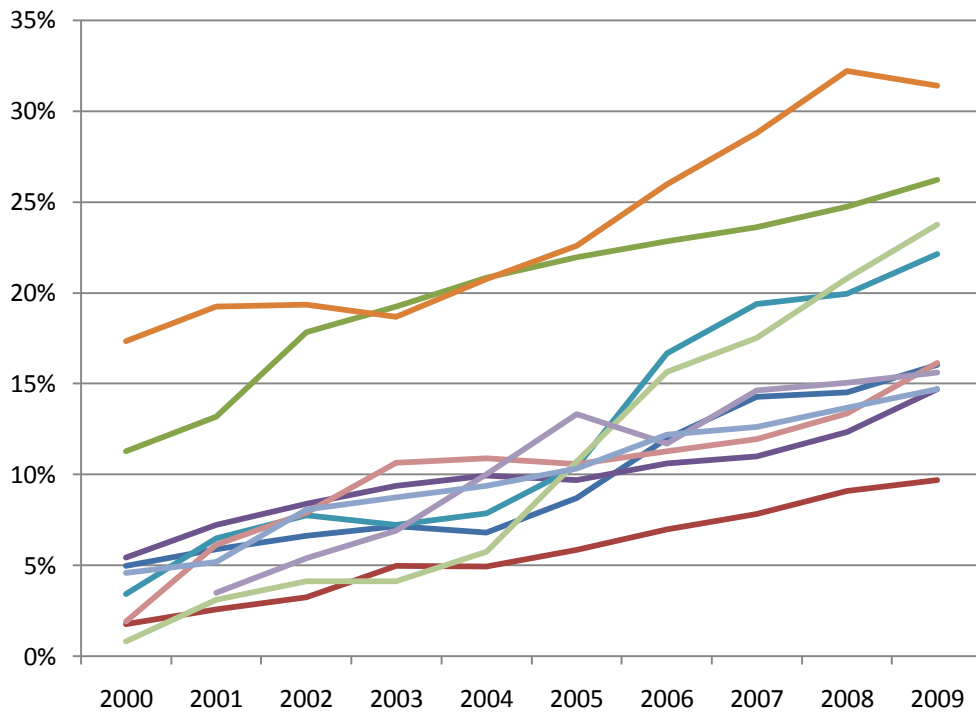
För att beskriva skillnader mellan patienter som använder insulinpump och patienter som inte använder insulinpump används logistisk regression. Variablerna som används i analysen är de som inte beror på typ av insulinbehandling. De variabler som är aktuella att börja analysen med är kön, ålder, diabetesduration och vårdenhet.

Korrelationer mellan variablerna som används i den logistiska regressionen visas i tabell 1. Korrelationsmatrisen bygger på data som består av patientens senaste besök. Mellan de förklarande variablerna är korrelationen låg i de flesta fall, det enda undantaget från detta är korrelationen mellan diabetesduration och ålder som är relativt hög ( $r = 0,51$ ). En del i den höga korrelationen kan vara det faktum att datamaterialet endast består av patienter med typ 1 diabetes och denna typ av sjukdomen bryter endast i undantagsfall ut på patienter som är äldre än 35 år. Det ger en naturlig förklaring till den höga korrelationen, äldre patienter har i regel haft diabetes längre eftersom sjukdomen sällan bryter ut i högre åldrar.

*Tabell 1. Korrelationer mellan de kontinuerliga variabler som används i den logistiska regressionen*

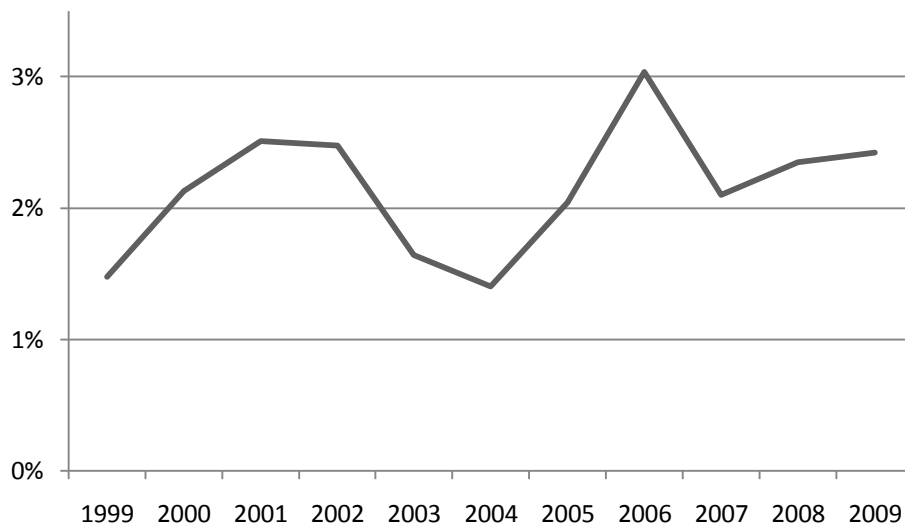
	<b>Kön</b>	<b>Duration</b>
<b>Duration</b>	0,040	-
<b>Ålder</b>	0,007	0,512

I figur 4 visas hur stor andel av de besökande patienterna som har insulinpump. Då besöksfrekvensen skiljer sig mycket åt mellan patienterna används endast det sista besöket varje patient har gjort de år som besök skett, detta för att patienter som gjort flera besök inte ska få för stor vikt. Diagrammet visar ett tydligt mönster på att användandet av insulinpump ökar och det är en klar skillnad mellan vårdenheterna i hur stor andel av patienterna som har insulinpump.



Figur 4. Andel patienter med insulinpump vid de olika vårdenheterna

Figur 5 visar andelen av patienterna som tilldelas insulinpump varje år. Data består av patientens sista besök det aktuella året och när en patient fått pump plockas individens fortsatta besök bort. Andelen patienter som inte redan har insulinpump och som börjat använda behandlingen varje år är under tidsperioden 1999 till 2009 stabilt mellan en och en halv och tre procentenheter.



Figur 5. Andel av besöken gjorda av icke pumpanvändare som resulterar i utdelning av insulinpump

## 2.2 Variabler

Då databasen innehåller all data som förs in i patienternas journaler är antalet variabler mycket stort. De variabler som kan vara av intresse för analyserna som ska utföras är de som kan tänkas påverka tilldelningen av insulinpump. Variablerna som används i studien redovisas nedan. När referensgrupp bestäms för de kontinuerliga variabler som i analysen används som dummyvariabler, görs valet mellan intervallen med minst och störst värden för att det ska bli lätt att illustrera. Då det gäller tidsvariabler, Ålder och Duration, väljs det sista intervallet och för variablerna HbA<sub>1c</sub> och enheter insulin per kg väljs det första intervallet.

## **Pump - Responsvariabel**

Responsvariabeln Pump antar värdet noll om patienten inte använt insulinpump under sin tid i studien och ett om patienten har gjort det.

### **2.2.1 Konstanta variabler**

Konstanta variabler är de som ej förändras med tiden, det vill säga är fixa över tid.

#### **Besöksdatum**

Datum för en individs besök på någon av vårdenheterna. Besöksdatum används för att beräkna Ålder, Duration samt tiden som patienten varit i studien.

#### **Kön**

Kön är en kategorisk variabel som antar värdet ett då det är en kvinna och noll då det är en man.

#### **Vårdenhet**

Vårdenhet är en kategorisk variabel som består av de tio vårdenheter som ingår i studien. Dessa vårdenheter är Alingsås sjukhus, Borås lasarett, Eksjö Högländssjukhus, Falu Lasarett, Gävle sjukhus, Kungälv sjukhus, Sahlgrenska universitetssjukhuset, Sandviken sjukhus, Skene lasarett och Uddevalla sjukhus. Samtliga är avidentifierade i rapporten. I analyserna representeras vårdenhet av nio dummyvariabler med Vårdenhet 10 som referensgrupp. Då Cox-regressionen utförs måste alla variabler ha uppmätta värden i början av tidsaxeln och på grund av saknade mätvärden i inledningen av tidsaxeln utförs analysen på ett datamaterial med något färre individer, se tabell 2.



Tabell 2. Antal patienter i studien

Vårdenhet	Population	Cox
Vårdenhet 1	441	365
Vårdenhet 2	1258	1085
Vårdenhet 3	625	451
Vårdenhet 4	1069	934
Vårdenhet 5	654	522
Vårdenhet 6	448	369
Vårdenhet 7	306	215
Vårdenhet 8	222	204
Vårdenhet 9	508	378
Vårdenhet 10	1693	1331
TOTAL	7224	5854

### 2.2.2 Tidsberoende variabler

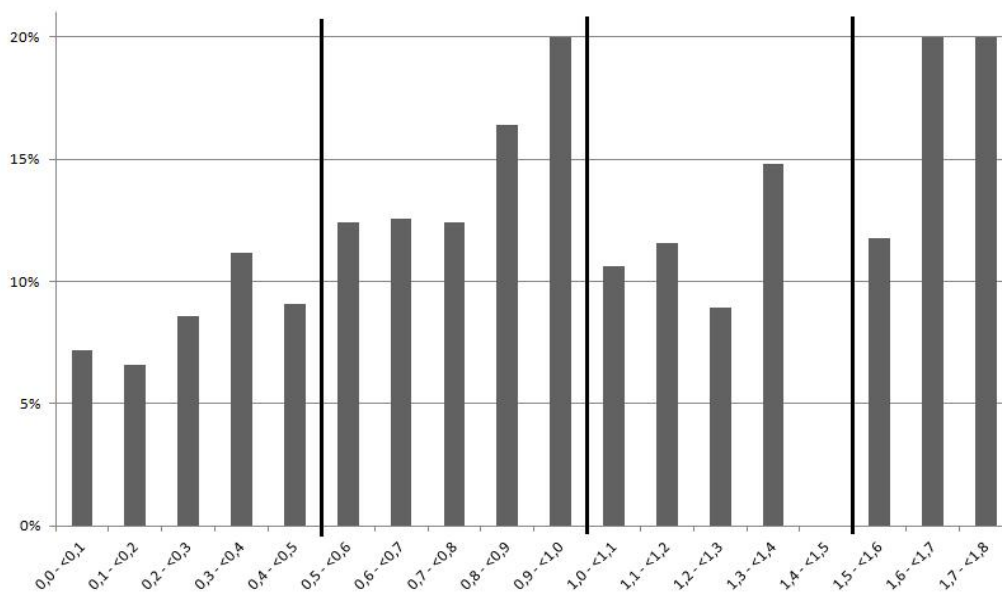
Tidsberoende är de variabler där värdet tillåts variera från en tidpunkt till en annan.

#### Duration

Duration är kontinuerlig och består av antal år som gått från att patienten diagnostiserades med diabetes till det aktuella besökets datum. För att tydligare kunna visa effekterna kategoriseras variabeln även till tioårs intervall: under 10; 10 - <20; 20 - <30; 30 - <40 samt 40 år eller längre (referensgrupp). Vid logistisk regression består durationsvariabeln av antalet år patienten haft diabetes vid sitt sista registrerade besök och vid Cox-regression används diabetesduration vid varje enskilt besök.

#### E/kg - Enheter insulin per kilo

Enheter insulin per kilo är en kvot mellan antal enheter insulin per dag och patientens vikt i kilo. Denna variabel är tidsberoende då både antal enheter insulin per dag och patientens vikt kan förändras över tiden. Det finns ej någon rekommenderad indelning varför den vid kategorisering delas upp utifrån fördelningen av patienternas E/kg vid första besökstillfället. Figur 6 visar andel patienter som fått pump i gruppen med liknande ingångsvärde på E/kg. Det finns en liten andel patienter, 0,35 procent, som har över 1,8 E/kg, dessa visas ej i figuren.



Figur 6. Andel blivande insulinpump användare/ingångsnivå på E/kg

Kategorisering har gjort så som visas i figur 6. Under 0,5 (referensgrupp); 0,5 - <1,0; 1,0 - <1,5 samt 1,5 enheter per kilo eller mer.

### HbA<sub>1c</sub>

HbA<sub>1c</sub> är en kontinuerlig variabel som förväntas ha stor inverkan på om en patient får insulinpump eller inte. Vid kategorisering delas denna variabel upp så att tolkning per procentenhet skall kunna utföras. Uppdelningen är: under 5 (referensgrupp); 5 - <6; 6 - <7; 7 - <8; 8 - <9 samt 9 eller högre. Denna uppdelning väljs då de flesta insulinpumpar delas ut till patienter med HbA<sub>1c</sub>-värden mellan fem och nio.

### Kreatinin

Kreatinin är en kontinuerlig, tidsberoende variabel som mäts genom blodprov och är ett mått på njurarnas funktion. Mätning görs i mikromol per liter,  $\mu\text{mol/l}$ . Referensvärdena ligger mellan 40 och 90  $\mu\text{mol/l}$  för kvinnor och 50 till 100  $\mu\text{mol/l}$  för män (Sjukvårdsupplysningen, 2010-05-20). Då mätvärdena är relativt stora tal skattas hasardkvoterna per tio enheter kreatinin.

## **Laserbehandling**

Laserbehandling är en tidsberoende variabel som har värdet noll tills patienten genomgått laserbehandling, för att sedan ha värdet ett. Laserbehandling är en indikator på grav retinopati, vilket är en typisk följsjukdom av diabetes.

## **Ålder**

Ålder används både i den logistiska regressionen och i Cox-regressionen som kontinuerlig i ett första skede för att sen kategoriseras då det blir lättare att tolka resultaten. Vid logistisk regression används ålder vid sista besök och vid Cox-regression används ålder vid aktuellt besök. Den kontinuerliga variabeln anges i hela år, den kategoriska i intervallen under 20; 20 – <30; 30 – <40; 40 – <50; 50 – <60 samt 60 år eller äldre (referensgrupp).

### **2.2.3 Kategorisering av kontinuerliga variabler**

För att använda linjär regressionsanalys på icke linjära variabler gäller generellt att dessa kan delas upp i grupper med liknande värden. Detta kallas kategorisering och genom att använda dummyvariabler skapade utifrån kategorierna blir effekterna lättare att tolka. Genom att till exempel dela in variabeln Ålder i åldersgrupper erhålls skattningar för åldersgrupperna istället för skattning av hur stor effekt varje år har. På detta sätt kan skillnader mellan olika åldersgrupper redovisas på ett tydligt sätt.

### **2.2.4 Val av tidsaxel till Cox-regression**

I en studie gjord över tid finns ofta flera möjliga tidsaxlar att tillgå. Val av tidsaxel påverkar samtliga skattningar och är av mycket stor vikt för analysen (Kirkwood & Sterne, 2003). Möjliga tidsaxlar att välja i en studie som denna är:

- Tid sedan rekrytering till studien, första besöksdatum till aktuellt datum.
- Ålder, tid sedan individen föddes.
- Tid sedan en viss händelse, tid sedan individen diagnostiserades med diabetes.
- Kalendertid, tid räknad från ett specifikt datum.

Tidsaxel ska generellt väljas så att den fångar upp de mest radikala förändringarna av hasardkvoter under ett begränsat tidsintervall. Vanligast vid Cox-regression är det första

alternativet, tid sedan rekrytering till studien, på grund av att Cox-regression inledningsvis utvecklades för studier av överlevnadsdata samt att statistikmjukvara först på senare tid blivit anpassad för att använda andra tidsaxlar.

I denna studie väljs det första alternativet, tid sedan rekrytering till studien, då det är den tidsaxel som förväntas fånga de mest radikala hasardförändringarna under ett begränsat tidsintervall. Ytterligare en anledning till valet av tidsaxel är att effekten av tid sedan diagnos (duration) och ålder är av intresse att skatta. Variabler vars effekt ska skattas måste användas som kovariater i regressionen och kan alltså inte användas som tidsaxel.

För att få en hanterbar tidsaxel räknas varje individs tid i månader sedan första besöksdatum. Saknade mätvärden har gjort att tidsaxeln har definierats som tid i månader sedan första besök då mätvärde registrerats minst en gång på samtliga tidsberoende variabler av intresse. Patienterna följs tills de har ett event (tilldelas insulinpump) eller censureras (försvinner ur studien på annat sätt).

### **2.3 Definition av population och rensning av data**

I denna kohortstudie definieras populationen som alla patienter med typ 1 diabetes vilka besökt någon av vårdenheterna under tiden vårdenheten ingått i studien.

För att analyserna inte ska påverkas av fel i data och för att säkerställa att analysresultaten blir korrekta genomförs viss rensning av data. Den första rensningen är att plocka bort uppenbart felaktiga observationer ur variabeln id-nummer vilken består av avidentifierade personnummer.

Steg två är att selektera patienter med typ 1 diabetes då ursprungsdata består av både patienter med typ 1 och typ 2 diabetes. Då inläsning till SAS sker skapas en unik nyckelvariabel för varje individ. Efter dessa initiala rensningar återstår 7266 individer som tillsammans gjort 264194 besök. Patienterna har sitt första besök vid olika tidpunkter och de har olika besöksfrekvens.

Rensningen av datamaterialet fortsätter med att rensa bort felaktigheter i variabeln med besöksdatum. Datum som inte existerar samt besök som utförts innan vårdenheten började använda Diab-Base tas bort. I nästa steg skapas en könsvariabel och personer med felaktiga personnummer rensas bort. Exempel på felaktiga personnummer är sådana

med datum eller månader som inte existerar. I det sista steget rensas de individer bort som enligt data fick diabetes innan de var födda.

Då ovanstående rensningar utförts återstår 7224 patienter med sammanlagt 250789 besök.

## **2.4 Metod**

Denna studie ska svara på frågor om hur en binär responsvariabel påverkas av flertalet förklarande variabler, såväl kvalitativa som kvantitativa. Detta leder till att följande metoder kommer att användas: Logistisk regression samt Cox-regression.

För att utreda om det i dagsläget finns några skillnader mellan patienter som har insulinpump jämfört med patienter som inte har det utförs en tvärsnittsstudie med logistisk regression. Endast data från senaste besök analyseras.

Cox-regression är den vanligaste regressionsmetoden för analys av överlevnadsdata (Kirkwood & Sterne, 2003). Metoden kan till skillnad från den logistiska regressionen ta med ett tidsperspektiv i beräkningen och variabler som förändras över tid kan inkluderas vilket gör att det med metoden kan utredas vad som föregår ett byte till insulinpump.

### **2.4.1 Hypotesprövning**

För att bedöma variabelers lämplighet i modeller utförs test av signifikant inverkan på responsvariabeln. Detta test kan utföras på en eller flera variabler simultant genom flera olika metoder, exempel på sådana är Likelihood-kvottest, scoretest eller Walds test. Likelihood-kvottestet bygger direkt på log-likelihood-kvoten, Walds test bygger på dess kvadratiska approximation och score-testet bygger på en alternativ kvadratisk approximation av log-likelihood-kvoten. Likelihood-kvoten är det mest stabila av testen medan övriga är goda approximationer så länge parameterskattningarna görs på log-skala (Kirkwood & Sterne, 2003). Samtliga regressioner i denna studie utförs på log-skala vilket gör att de tre testen kommer ge lika resultat. Likelihood-kvottestet kan liknas med F-test vid vanlig regression där jämförelse görs mellan en fullständig och en reducerad modell. Detta test kan ej utföras direkt i programvaran som används (SAS) utan vissa beräkningar måste utföras för hand. Walds test utförs som standard för varje skattad parameter och kan även göras för flera parametrar simultant direkt i programvaran, varför denna metod

kommer att användas genomgående i denna studie. Walds test jämförs med en  $\chi^2$ -fördelning varför dess teststatistika blir  $\chi^2_{\text{Wald}}$ .

$$\chi^2_{\text{Wald}} = -2 * \log (\text{LR})_{\text{kvadratisk}} = \left( \frac{\text{MLE} - \theta_{\text{null}}}{S} \right)^2 = \left( \frac{\text{MLE}}{S} \right)^2, \text{ om } \theta_{\text{null}} = 0$$

LR = Likelihood-kvot.

MLE = Skattning av maximum likelihood, utförs av programvara.

$\theta_{\text{null}}$  = Logaritmerad hasard- eller oddskvot beroende på typ av regression.

## 2.4.2 Logistisk regression

För att belysa eventuella skillnader mellan patienter som använder insulinpump och patienter som inte gör det utförs logistisk regression. Denna metod väljs då responsvariabeln är binär vilket innebär att den endast kan anta två värden, patienterna har fått insulinpump eller så har de det inte. I analysen undersöks data från det senaste besöket patienten har gjort och endast de variabler som inte påverkas av eventuellt användande av insulinpump inkluderas i modellen. De variabler som används är kön, ålder, vårdenhet och diabetesduration.

Modellen är multiplikativ och kan skrivas:

$$\text{Odds}(y = 1) = \exp (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$$

$b_p$  skattar  $\beta_p$  med "Maximum Likelihood". Skattningen utifrån datamaterialet görs genom att iterera till det värde på parametrarna hittats som maximerar sannolikheten för aktuellt datamaterial.

Oddskvoten (OR) för var och en av variablerna fås genom att använda betaskattningen för respektive variabel i formeln nedan:

$$\text{OR}_{x_p} = \frac{\text{Odds}(y = 1)}{\text{Odds}(y = 0)} = \frac{\exp (\beta_0 + \beta_p)}{\exp (\beta_0)} = \exp (\beta_p)$$

Resultaten kommer att presenteras med oddskvoter och dess konfidensintervall. För att beräkna de 95 procentiga konfidensintervallen används formeln:

$$\exp(\beta \pm 1,96 * S. E. (\beta))$$

Om oddskvotens konfidensintervall exkluderar talet ett kan slutsatsen dras att med 95 procents säkerhet påverkar den aktuella variabeln sannolikheten att patienten med givna egenskaper har pump. Är värdet på oddskvoten större än ett betyder det att för varje ökning i den aktuella kovariaten, givet att övriga kovariater hålls konstanta, så ökar oddset för en viss händelse med oddskvoten minus ett. Är oddskvoten istället mindre än ett minskar oddset för händelsen med ett minus oddskvoten för varje ökning i aktuell kovariat givet att de andra kovariaterna hålls konstanta. (Kutner, Nachtsheim, Neter, & Li, 2005)

### 2.4.3 Cox-regression

Regressionsmodellen för hasardkvoter (Cox, 1972) introducerades först av professor D.R. Cox och den refereras ofta till som "Cox regression model", "Cox proportional hazards model" eller helt enkelt "Cox model". Cox-regression är en semi-parametrisk modell eftersom baselinehasarden kan anta vilken form som helst men kovariaterna kommer in i modellen linjärt och modellen kräver inte något antagande om specifik fördelning på överlevnadstiden. Dock görs ett starkt antagande om att effekterna av de olika variablerna är konstanta över tiden och är additiva i en viss skala (Altman, 1991).

Cox-regression är idag den vanligaste metoden för att jämföra överlevnadsdata mellan en eller flera grupper (Kirkwood & Sterne, 2003). Metoden jämför andelen observationer vid varje tidpunkt där ett event inträffat mot de individer som fortfarande är kvar i studien.

Den momentana risken är ofta väldigt liten och av svagt intresse. Vad som är intressant att utreda är om det över tiden finns systematiska skillnader mellan hasarderna beroende på individers olika egenskaper (Petrie & Sabin, 2009), har en kvinnlig patient större sannolikhet att gå över till insulinpump än en manlig? Dessa systematiska skillnader skattas som avvikelser från baselinehasarden ( $h_0(t)$ ), den underliggande hasardfunktionen för en individ där alla kovariater har värdet noll. Baselinehasarden i sig kan variera över tid

och skattas ej vid regressionen, endast det relativa förhållandet mellan denna och de påverkande faktorerna skattas.

Skattningar  $b_1, \dots, b_p$  av  $\beta_1, \dots, \beta_p$  erhålls genom att använda en variant av maximum likelihood som kallas "partial likelihood". Den partiella likelihood-skattningen tillåter variabler som förändras över tiden vilket gör det möjligt att använda tidsberoende kovariater i regressionen. Exponenten av betaskattningarna ( $e^b$ ) ger skattningar av de relativa hasarderna eller hasardkvoterna (HR) enligt formeln nedan:

$$HR(t) = \frac{h(t)}{h_0(t)} = \frac{h_0(t) * e^{(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}}{h_0(t)} = e^{(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}$$

Hasardkvoter tolkas på samma sätt som oddskvoter tolkas vid logistisk regression. Saknas effekt blir hasardkvoten ett, vilket kommer från  $e^{\beta_i} = 1$  då  $\beta_i = 0$ . Är hasardkvoten 1,12 innebär det att om värdet på variabeln ökar med en enhet så ökar sannolikheten att gå över till insulinpump med 12 procent, förutsatt att alla andra variabler hålls konstanta.

Till skillnad från alternativa metoder som till exempel exponentiell, Weibull och Gompertz modeller så behöver inget antagande om specifik fördelning göras (Petrie & Sabin, 2009). Cox-regressionen gör antagande om att hasardkvoterna är proportionella mot baselinehasarden,  $h_0(t)$ , över tiden. För att kontrollera detta finns flertalet metoder så som visuell analys, Kaplan-Meier överlevnadsplot (log-log plot) eller så testas varje variabel genom test för signifikanta tidsinteraktioner vilket är metoden som används i denna studie. Även antagande om linjäritet på en viss skala görs, linjäritet på log-skala är en förutsättning vid användning av Cox-regression.

Som i alla överlevnadsanalyser görs antagandet om att den informativa censureringen är helt oberoende av sannolikheten att individen skall ha event. Informativ censurering är då patienter flyttar eller av annan anledning försvinner från studien före uppföljningstidens slut. Administrativ censurering sker vid sista registrerade besök då detta anses vara slutet för patientens tid i studien men den är generellt icke informativ (Petrie & Sabin, 2009).

### **Linjäritet på log-skala**

Linjäritetsantagandet i Cox-regressionen kontrolleras genom att för varje tidsberoende variabel inkludera kvadraten av variabeln i modellen. Genom att testa den kvadratiska



termen för signifikant effekt erhålls svar på om variabeln är log-linjär eller ej. Om linjäritetsantagandet faller kan modellen justeras genom att kategorisera den aktuella variabeln vilket också gör det lättare att tolka resultaten.

### **Proportionalitet**

Proportionalitetsantagandet testas genom att inkludera interaktioner mellan tidsvariabeln och var och en av de tidsberoende variablerna. En interaktionsterm för varje sådan kombination skapas för att sedan testas med Walds test. Blir detta test signifikant är variabeln av intresse ej proportionell mot baselinehasarden och proportionalitetsantagandet faller (Kirkwood & Sterne, 2003). Vid användning av kategoriserade variabler är tillvägagångssättet snarlikt. Tidsinteraktioner skapas och testas sedan med ett Wald-test, skillnaden är att samtliga dummyvariabler skapade från den aktuella variabeln testas simultant med nollhypotesen att samtliga dummyvariabler är noll.

## 2.4.4 Icke proportionella hasarder

När proportionalitetsantagandet faller måste detta åtgärdas, det finns flera metoder för att hantera problemet och de vanligaste är följande två varianter.

### 1. Dela upp uppföljningstiden.

Tidsaxeln delas upp så att proportionalitet råder inom varje partition av tidsaxeln vilket gör att Cox-regression kan utföras där de olika partitionerna tillåts ha olika baselinehasarder. Tidsaxeln kan delas i två eller fler delar så att proportionalitet råder inom varje intervall (Petrie & Sabin, 2009). Bästa uppdelning erhålls på enklast sätt genom visuell granskning. Sammanfattningen av dessa proportionella Cox-regressioner ger sedan skattningen på hasardkvoter avseende hela datamaterialet.

### 2. Stratifiera

För att justera för icke-proportionalitet så kan data stratifieras utifrån variabeln som är orsaken till att det inte är proportionellt. Teorin bakom detta tillvägagångssätt är detsamma som i fallet då uppföljningstiden delas upp i partitioner, data delas i mindre grupper som tillåts ha olika baselinehasarder. Skattningarna inom varje intervall vägs samman för att kunna presenteras som skattningar av totala effekter. Metoden förutsätter att variabeln som används för stratifieringen ej är av stort intresse att analysera, då analys på effekten av denna ej kan utföras.

$$\sum_{i=1}^n \delta_i \left\{ \sum_{j=1}^p \beta_j x_{ji}(t_i) - \log \sum_{h=1}^H \sum_{l \in R_h(t_i)} \exp \left( \sum_{j=1}^p \beta_j x_{jl}(t_i) \right) \right\}$$

I datamaterialet som analyseras finns variabeln Vårdenhet. Denna analys täcker endast populationen av de tio vårdenheter som ingår i studien vilket gör att slutsatser inte kan dras om andra liknande vårdenheter. Utvecklingen mellan olika vårdenheter kan inte förväntas vara densamma och därför är skillnader mellan vårdenheter ej av största intresse i denna analys. Att vårdenheterna ej utvecklas på samma sätt över tidsaxeln visar figur 4 tecken på. Det indikerar att det eventuellt kan råda icke-proportionalitet i variabeln. Vårdenheter har i kalendertid kommit in i studien vid olika tidpunkter men på grund av den valda tidsaxeln hanteras dessa som om de kommit in vid samma tidpunkt.

## 3 Resultat

### 3.1 Logistisk regression

För att finna den modell som på bästa sätt beskriver skillnader mellan de patienter som har insulinpump jämfört med de som inte har insulinpump utförs logistisk regression. I den logistiska regressionen används endast variabler som inte påverkas av om patienten använder insulinpump eller inte, så som till exempel insulindos eller HbA<sub>1c</sub>-värde gör.

#### 3.1.1 Val av modell till logistisk regression

De första variablerna som testas i modellen är kön, ålder, duration och vårdenhet. Tabell 3 visar variablernas p-värden från denna modell. Walds test utförs simultant på samtliga kovariater tillhörande den kategoriserade variabeln Vårdenhet för att utreda om den påverkar responsvariabeln. Då testet ger ett signifikant resultat med ett p-värde som är mindre än 0,0001 har minst en av de ingående dummyvariablerna för vårdenhet betydande effekt och alltså ska dessa behållas i modellen. De resterande variablerna duration, kön samt ålder blir alla signifikanta och påverkar alltså sannolikheten att en patient har insulinpump. Ingen av variablerna exkluderas.

Tabell 3. P-värden för variablerna i den grundläggande logistiska regressionen

Variabel	$\chi^2_{Wald}$	$p > \chi^2_{Wald}$
Vårdenhet 1	0,1	0,7652
Vårdenhet 2	23,1	<0,0001
Vårdenhet 3	12,4	0,0004
Vårdenhet 4	4,1	0,0436
Vårdenhet 5	0,0	0,8488
Vårdenhet 6	25,1	<0,0001
Vårdenhet 7	0,0	0,8913
Vårdenhet 8	3,1	0,0796
Vårdenhet 9	3,0	0,0850
Kön	86,9	<0,0001
Duration	124,3	<0,0001
Ålder	250,8	<0,0001

För att undersöka om sambandet mellan responsvariabeln och kovariaten Duration är linjärt på logskala eller inte utökas modellen med dess kvadratiska term. Resultatet blir signifikant med ett p-värde som är mindre än 0,0001 vilket tyder på att oddsen för

insulinpump och duration inte är linjärt på log-skala. På samma sätt testas variabeln Ålder, även denna blir signifikant och behålls i modellen.

Ett exempel på en parvis interaktion som testas är Kön\*Duration då durationsvariabeln är kontinuerlig och interaktionstermen blir signifikant med ett p-värde på 0,0375. Variablerna Ålder och Duration kategoriseras och när ett simultant test sedan utförs på de ingående termerna i interaktionen mellan variablerna Kön och Duration blir inte testet signifikant. Modellen med interaktionstermen baserad på den kontinuerliga durationsvariabeln samt det simultana testet visas i bilaga 1. Resultatet tyder på att interaktionstermen har viss påverkan men då målet är att ha en enkel modell och att grundvariablerna har tydligare effekt beslutas att inte inkludera interaktionstermer i modellerna i denna studie.

### 3.1.2 Slutgiltig modell logistisk regression

Modellen innehållande variablerna Kön, Vårdenhet, Ålder samt Duration kommer att användas som slutgiltig modell. Testet med den globala nollhypotesen  $\beta=0$  ger signifikant resultat, modellens variabler kan användas för att påvisa skillnader mellan patienter som använder insulinpump och patienter som inte gör det.

För att få en mer lättolkad modell kategoriseras variablerna Duration och Ålder. Test på de kategoriserade variablerna som ingår visar att samtliga är signifikanta. Resultaten redovisas i tabell 4 där värdena på  $\chi^2_{Wald}$  visar att den variabel som inverkar mest på om patienten använder insulinpump eller inte är Ålder.

Tabell 4. Walds test av kovariaterna i den logistiska regressionen

Variabel	DF	$\chi^2_{Wald}$	$P > \chi^2_{Wald}$
Vårdenhet (kategoriserad)	9	98,1	<0,0001
Ålder (kategoriserad)	5	196,7	<0,0001
Duration (kategoriserad)	4	155,3	<0,0001
Kön	1	87,3	<0,0001

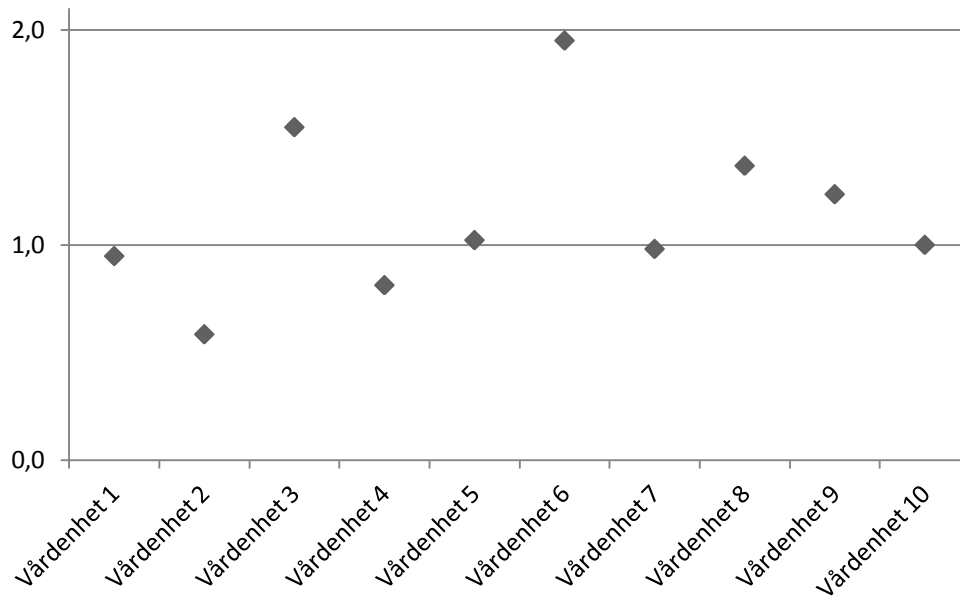
Tabell 5 visar oddskvoterna samt konfidensintervallen för alla variablerna som inkluderas i den slutgiltiga logistiska regressionen. Det visar att kvinnor i större utsträckning än män

har insulinpump och det finns skillnader mellan vårdenheterna i hur stor andel av patienterna som har insulinpump. Minst andel patienter med pump har Vårdenhet 2 där det är ungefär hälften så stor andel som har insulinpump jämfört med Vårdenhet 10. Den vårdenhet med störst andel är Vårdenhet 6 som har nästan dubbelt så stor andel om alla andra kovariater hålls konstanta. Resultatet visar också att det finns signifikant skillnad i andelen insulinpump användare mellan de olika ålderskategorierna. Minst är andelen i den översta ålderskategorin, då patienten vid sitt senaste besök var 60 år eller äldre. Andelen är störst i åldersintervallet mellan 20 och 30 år och ju äldre patienten blir ju mindre blir andelen patienter som har insulinpump. Det finns även skillnader i andelen insulinpump användare beroende på hur lång tid patienten haft diabetes, ju längre tid sedan diabetesdiagnos ju större andel har insulinpump. Dock finns det en grupp som inte följer trenden. Patienter med väldigt lång diabetesduration, över 40 år, använder insulinpump i mindre utsträckning än gruppen med duration 30 till 40 år.

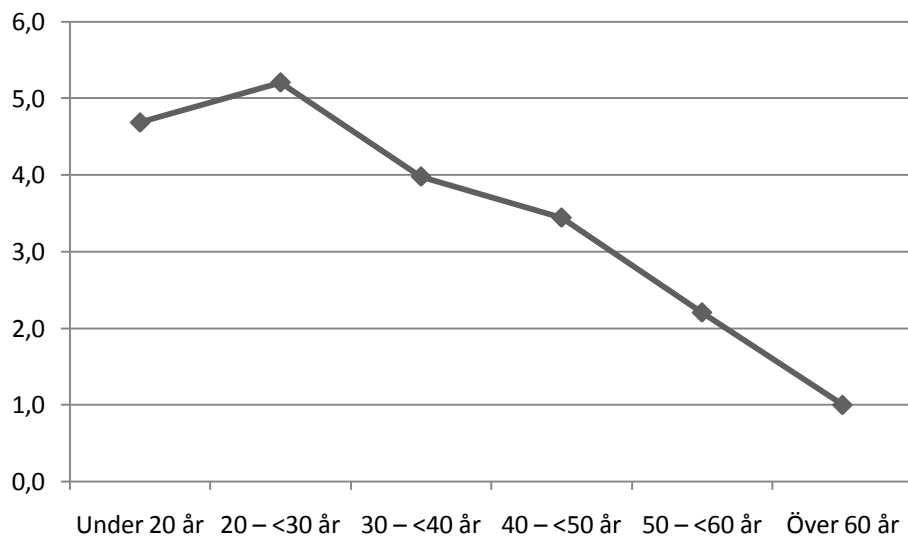
Tabell 5. Logistisk regression, oddskvoter och konfidensintervall

Variabel	OR	95% KI <sub>OR</sub>	
Kön	1,83	1,62	2,08
Vårdenhet 1	0,95	0,71	1,27
Vårdenhet 2	0,58	0,47	0,73
Vårdenhet 3	1,55	1,22	1,95
Vårdenhet 4	0,81	0,66	1,00
Vårdenhet 5	1,02	0,80	1,30
Vårdenhet 6	1,94	1,51	2,50
Vårdenhet 7	0,98	0,71	1,37
Vårdenhet 8	1,36	0,94	1,97
Vårdenhet 9	1,23	0,95	1,60
Vårdenhet 10	1		
Ålder: <20	4,68	2,54	8,60
Ålder: 20 - <30	5,22	4,06	6,71
Ålder: 30 - <40	3,96	3,13	5,01
Ålder: 40 - <50	3,42	2,75	4,26
Ålder: 50 - <60	2,20	1,76	2,75
Ålder: ≥60	1		
Duration: <10	0,26	0,20	0,35
Duration: 10 - <20	0,76	0,60	0,97
Duration: 20 - <30	0,94	0,74	1,19
Duration: 30 - <40	1,27	1,01	1,60
Duration: ≥40	1		

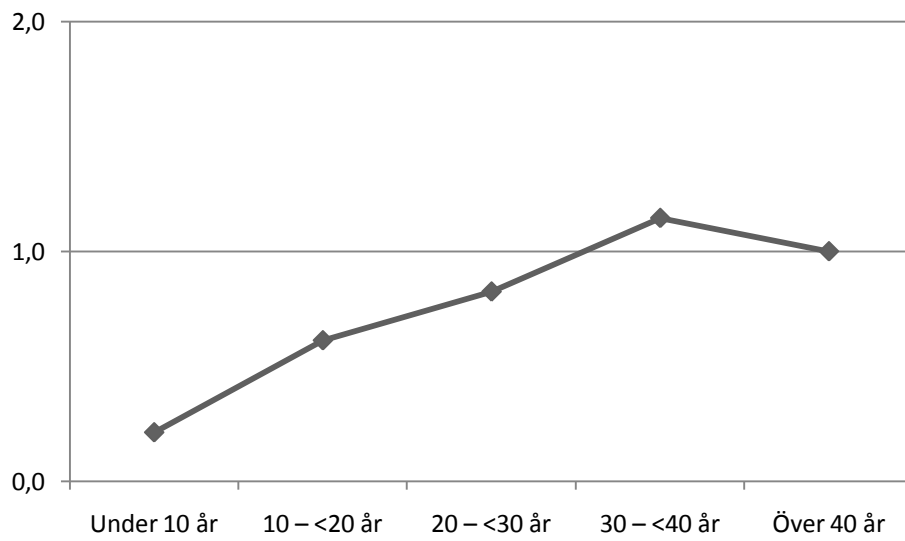
För att lättare se hur oddskvoterna förhåller sig mellan kategorierna inom variablerna illustreras dessa i figur 7, 8 och 9.



Figur 7. Odds kvoter för vårdenhet



Figur 8. Odds kvoter för ålder



*Figur 9. Oddskvoter för diabetesduration*



## 3.2 Cox-regression

I den logistiska regressionen visas skillnader mellan patientgrupper i dagsläget. För att utreda hur olika variabler påverkar sannolikheten att en patient ska gå över till insulinpump används Cox-regression. Denna regressionsmetod kan hantera kovariater som förändras med tiden vilket gör att även tidsberoende variabler kan undersökas. Data innehåller 698 intressanta händelser (event) varav 17 censureras på grund av saknade mätvärden i duration. Censureringen beror på att datumet då patienten diagnostiserades med diabetes är okänt. Totalt så har 681 patienter av de 5854 som är med i Cox-regressionen event under den studerade tidsperioden.

### 3.2.1 Val av modell till Cox-regression

I den första modellen används de variabler som med logistisk regression visats ha signifikant effekt på vilka patientgrupper som använder pump. Modellen från den logistiska regressionen utökas med de tidsberoende variablerna HbA<sub>1c</sub>, Kreatinin och E/kg. Variablerna Ålder och Duration tillåts förändras över tiden medan Vårdenhet och Kön hålls fixa som i den logistiska regressionen. Variabeln Laser testas också som en tidsberoende kovariat men då den saknar signifikant påverkan på responsvariabeln med ett  $\chi^2_{Wald}$ -värde på 1,9 vilket ger ett p-värde på 0,1675, har variabeln uteslutits ur modellen.

Resultatet av den första Cox-regressionen visas i tabell 6, samtliga kovariater i regressionen har signifikant effekt. Ett simultant Wald-test på den kategoriserade variabeln Vårdenhet med nio frihetsgrader ger signifikant resultat med ett p-värde mindre än 0,0001 vilket påvisar att även denna ska vara med i modellen. Fullständig utskrift med hasardkvoter och parameterskattningar finns i bilaga 2.

Tabell 6. P-värden för de grundläggande variablerna i Cox-regression

Variabel	$\chi^2_{\text{Wald}}$	$p > \chi^2_{\text{Wald}}$
Vårdenhet 1	1,0	0,3059
Vårdenhet 2	35,3	<0,0001
Vårdenhet 3	30,0	<0,0001
Vårdenhet 4	6,6	0,0101
Vårdenhet 5	0,4	0,5502
Vårdenhet 6	10,0	0,0016
Vårdenhet 7	4,0	0,0447
Vårdenhet 8	0,2	0,0965
Vårdenhet 9	3,4	0,0635
HbA <sub>1c</sub>	97,3	<0,0001
Kreatinin	6,1	0,0136
E/kg	93,4	<0,0001
Duration	7,4	0,0064
Ålder	181,7	<0,0001
Kön	29,4	<0,0001

### 3.2.2 Kontroll av log-linjärt beroende

Ett, på log-skala, icke linjärt beroende mellan responsvariabel och tidsberoende kovariater skulle påvisa att den log-linjära modellen som testats i föregående avsnitt ej är korrekt. Genom att utöka modellen med kvadraten till en tidsberoende kovariat testas om sambandet mellan den aktuella kovariaten och responsvariabeln har ett linjärt beroende eller ej. Detta utförs individuellt för varje tidsberoende kovariat i modellen.

De resultat som presenteras i tabell 7 visar att det finns signifikant krökt samband mellan responsvariabel och kovariaterna HbA<sub>1c</sub>, E/kg, Duration och Ålder. Dessa samband måste modelleras för att erhålla korrekta hasardkvoter.

Tabell 7. Kontroll för icke log-linjärt beroende

Variabel	$p > \chi^2_{\text{Wald}}$
HbA <sub>1c</sub> <sup>2</sup>	0,0003
E/kg <sup>2</sup>	<0,0001
Kreatinin <sup>2</sup>	0,9996
Duration <sup>2</sup>	<0,0001
Ålder <sup>2</sup>	0,0015

### 3.2.3 Kategorisering av variabler

De icke log-linjära samband som upptäcktes genom analys av kvadratiska termer i föregående avsnitt måste modelleras för att erhålla korrekta skattningar. Kategorisering av dessa variabler utförs för att hantera problemet och det resulterar även i att hasardkvoterna i den slutgiltiga modellen blir mer lättolkade. Kategorisering gjordes enligt uppdelning beskriven i variabelavsnittet.

### 3.2.4 Kontroll av proportionalitetsantagandet

Cox-regression kallas ofta ”Cox proportional hazards regression” av den enkla anledning att hasardkvoterna måste vara proportionella mot baselinehasarden över tid. Detta antagande kontrolleras enklast genom att testa om kovariaterna interagerar med den aktuella tidsvariabeln. Testet utförs individuellt för varje tidsberoende kovariat i modellen.

*Tabell 8. Kontroll för proportionalitet*

Variabel	DF	$p > \chi^2_{\text{Wald}}$
Tid*Kreatinin	1	0,1626
Tid*Kön	1	0,1413
Tid* HbA <sub>1c</sub> (kategoriserad)	5	0,0953
Tid*Vårdenhet (kategoriserad)	9	<0,0001
Tid*E/kg (kategoriserad)	3	0,6691
Tid*Duration (kategoriserad)	4	0,0263
Tid*Ålder (kategoriserad)	5	0,9581

Walds test på Tid\*Vårdenhet och Tid\*Duration ger signifikant resultat vilket innebär att kovariaterna Vårdenhet och Duration ej uppfyller proportionalitetsantagandet.

### 3.2.5 Stratifiering

För att hantera icke-proportionalitet påvisad av signifikanta tidsinteraktioner så stratifieras data på variabeln Vårdenhet. Denna stratifiering tillåter att varje stratum i regressionen använder olika baselinelinjesrisker vilket löser problemet med icke-proportionalitet. Hur denna stratifiering delar upp datamaterialet visas i tabell 9.

Tabell 9. Stratifiering

Vårdenhet	Total	Event	Censurerade	Procent Censurerade
Vårdenhet 1	365	42	323	88,49
Vårdenhet 2	1085	79	1006	92,72
Vårdenhet 3	451	82	369	81,82
Vårdenhet 4	934	98	836	89,51
Vårdenhet 5	522	80	442	84,67
Vårdenhet 6	369	90	279	75,61
Vårdenhet 7	215	19	196	91,16
Vårdenhet 8	204	37	167	81,86
Vårdenhet 9	378	34	344	91,01
Vårdenhet 10	1331	120	1211	90,98
TOTAL	5854	681	5173	88,37

Interaktionstermen Tid\*Duration som i föregående analys gav signifikant resultat, har efter stratifiering ett p-värde på 0,17. Efter stratifiering håller alltså proportionalitetsantagandet i Cox-regressionen.

### 3.2.6 Slutgiltig modell Cox-regression

Efter gruppering av tidsberoende variabler samt stratifiering utifrån variabeln Vårdenhet utförs Cox-regression på den slutgiltiga modellen. Test av den globala nollhypotesen ger signifikant resultat, variablerna i modellen har signifikant inverkan på sannolikheten att gå över till insulinpump.

För att kontrollera att samtliga variabler påverkar den totala effekten så testas varje variabel med Walds test. Samtliga kovariater har i regressionen tydlig effekt på sannolikheten att gå över till insulinpump vilket visas i tabell 10. Ju högre värde på  $\chi^2_{\text{Wald}}$  ju mer påverkar variabeln sannolikheten att en patient går över till insulinpump. Störst

inverkan har variabeln E/kg, enheter insulin per kg kroppsvikt och därefter kommer variabeln Ålder.

Tabell 10. Walds test av kovariaterna i den slutgiltiga Cox-regressionen

Variabel	$\chi^2_{\text{Wald}}$	DF	$p > \chi^2_{\text{Wald}}$
Kreatinin	9,6	1	0,0020
Kön	25,8	1	<0,0001
HbA <sub>1c</sub> (kategoriserad)	78,7	5	<0,0001
E/kg (kategoriserad)	200,9	3	<0,0001
Duration (kategoriserad)	9,5	4	0,0507
Ålder (kategoriserad)	154,1	5	<0,0001

Resultatet visar att samtliga kovariater i modellen ska vara kvar och tolkning av hasardkvoter kan nu göras.

Hasardkvoter som redovisas i tabell 11 är grundade på data som stratifierats utifrån Vårdenhet. Varje vårdenhet tillåts ha olika baselinehasarder i regressionen, hasardkvoterna visar alltså skattning av olika parametrars effekt oberoende av vårdenhet.

Tabell 11. Cox-regression, basardkvoter och konfidensintervall

Variabel	HR	95% KI <sub>HR</sub>	
Kreatinin	0,94	0,91	0,98
Kön	1,52	1,29	1,79
HbA <sub>1c</sub> : <5	1		
HbA <sub>1c</sub> : 5 - <6	1,98	1,05	3,75
HbA <sub>1c</sub> : 6 - <7	2,70	1,45	5,02
HbA <sub>1c</sub> : 7 - <8	3,75	2,02	6,96
HbA <sub>1c</sub> : 8 - <9	4,82	2,58	9,03
HbA <sub>1c</sub> : ≥9	5,65	2,98	10,72
E/kg: <0,5	1		
E/kg: 0,5 - <1,0	0,29	0,25	0,35
E/kg: 1,0 - <1,5	0,17	0,12	0,25
E/kg: ≥1,5	0,40	0,18	0,86
Duration: <10	0,82	0,56	1,20
Duration: 10 - <20	1,05	0,74	1,50
Duration: 20 - <30	1,12	0,79	1,59
Duration: 30 - <40	1,24	0,88	1,77
Duration: ≥40	1		
Ålder: <20	9,37	3,58	24,50
Ålder: 20 - <30	8,63	5,91	12,59
Ålder: 30 - <40	6,18	4,35	8,77
Ålder: 40 - <50	3,98	2,80	5,64
Ålder: 50 - <60	2,63	1,86	3,73
Ålder: ≥60	1		

Resultatet av Cox-regressionen visar på många sätt liknande resultat som den logistiska regressionen. Kvinnor har betydligt större chans än män att gå över till insulinpump samtidigt som unga patienter i större utsträckning byter behandlingsmetod än vad äldre gör. Lång diabetesduration ökar chansen att gå över till insulinpump upp till en duration på 40 år då den minskar något.

De variabler som ej undersöktes med logistisk regression är HbA<sub>1c</sub>, E/kg och Kreatinin. Kreatinin har avtagande effekt, ett högt mätvärde minskar sannolikheten att gå över till insulinpump något. Att gå över till insulinpump är troligare för en patient som antingen har stor eller liten dagsdos insulin per kilo, grupperna under 0,5 enheter per dag och 1,5 eller fler enheter har störst sannolikhet att gå över till insulinpump. HbA<sub>1c</sub> som är

blodsockernivån över en längre tid har stor effekt och analysen visar att en hög nivå på HbA<sub>1c</sub> ökar sannolikheten att gå över till insulinpump.

### 3.2.7 Ytterligare analys av HbA<sub>1c</sub>-hasarden

Att ett högt långtidsblodsocker ofta leder till att patienten får insulinpump är inget konstigt då insulinpumpen förväntas sänka värdet till en acceptabel nivå. Något som har diskuterats är dock om denna ökning fortsätter även vid högre nivåer eller om det finns en brytpunkt i ökningen där ett växande HbA<sub>1c</sub> inte längre leder till ökad sannolikhet att gå över till insulinpump.

För att utreda detta utfördes en mindre studie, en Cox-regression på endast den tidsberoende kontinuerliga variabeln HbA<sub>1c</sub> samt dess kvadrat. Genom att använda dessa parameterskattningar utreds med hjälp av derivering om det finns en maxpunkt för den ökande trenden och vad maxpunkten i så fall är.

Tabell 12. Parameterskattningar för HbA<sub>1c</sub> och dess kvadrat

Variabel	$\beta$	$p > \chi^2_{\text{Wald}}$
HbA <sub>1c</sub>	0,68015	<0,0001
HbA <sub>1c</sub> <sup>2</sup>	-0,03152	0,0004

$$f(x) = \beta_1 * x + \beta_2 * x^2$$

$$\frac{d(f(x))}{dx} = \beta_1 + 2\beta_2 * x = 0 \rightarrow x = \frac{-\beta_1}{2\beta_2} \rightarrow x = \frac{-0,68015}{2(-0,03152)} \approx 10,7892$$

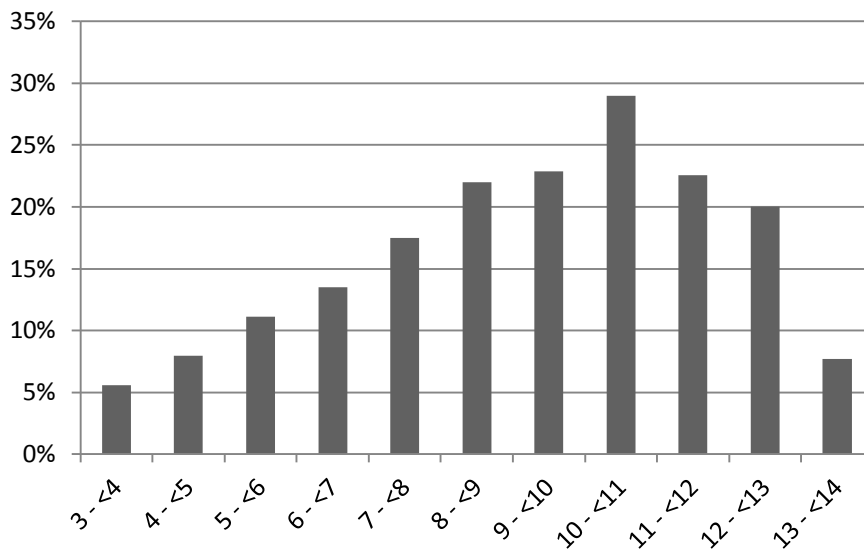
$$\frac{d^2(f(x))}{dx^2} = 2\beta_2 = -0,06304 \rightarrow \text{Maxpunkt funnen där HbA}_{1c} \text{ har värdet } 10,7892.$$

Eftersom andraderivatet är negativt är det en maxpunkt.

Det finns tecken på att brytpunkten är vid ett HbA<sub>1c</sub>-värde på 10,79. Har patienten ett värde över 10,79 ökar inte längre sannolikheten att gå över till insulinpump med ett växande HbA<sub>1c</sub>.

Figur 10 visar andel av patienter i varje grupp av HbA<sub>1c</sub>-nivå som gått över till insulinpump under studiens uppföljningstid. HbA<sub>1c</sub>-nivån grundar sig på sista registrerade besök för dem som ej gått över till pump och HbA<sub>1c</sub>-nivå vid event för dem som gått över till pump. Diagrammet är alltså ej grundat på hasarder och skall alltså ej tolkas som sådana. Men vad som visas även i denna figur är att risken för att gå över till insulinpump

är som högst vid ett värde på HbA<sub>1c</sub> mellan tio och elva, vilket ger stöd till resultatet av ett maximum vid 10,79.



Figur 10. Andel patienter i varje grupp av HbA<sub>1c</sub> som har gått över till insulinpump



## **4 Slutsatser och metoddiskussion**

### **4.1 Metoddiskussion**

För att utreda olika förklarande variabelers effekt på en responsvariabel finns ett antal statistiska metoder att välja mellan. Då det handlar om flertalet förklarande variabelers påverkan på en responsvariabel står valet främst mellan variansanalys och regressionsanalys. Om analys endast ska utföras på kategoriska variabler är dessa metoder väldigt lika och valet mellan dem är indifferent. Ska hänsyn tas till variabelers kvantitativa natur utan att använda kategorisering kan endast regressionsanalys användas (Kutner, Nachtsheim, Neter, & Li, 2005).

I denna studie antar responsvariabeln Pump värdet ett då patienten har gått över till insulinpump och har annars värdet noll. Då responsvariabeln är binär går det ej att använda vanlig multipel regressionsanalys. För analys av överlevnadsdata finns ett antal regressionsmodeller att välja mellan varav logistisk regression och Cox-regression används i denna studie.

#### **4.1.1 Logistisk regression**

Den logistiska regressionen används som en tvärsnittsstudie för att beskriva vilka skillnader som finns mellan patienter som använder insulinpump och patienter som inte gör det.

Då den huvudsakliga anledningen till att använda logistisk regression är att på ett enkelt sätt beskriva skillnader mellan insulinpumpanvändare och patienter som inte använder insulinpump önskas modellen hållas så enkel som möjligt. Detta för att kunna tolka effekterna som de olika variablerna har. Interaktionstermer gör modellen mer svårtolkad varför det inte inkluderas några sådana. Om mer avancerade analyser över skillnader mellan patienter som använder insulinpump och patienter som inte använder pump ska utföras bör interaktionstermer inkluderas.

#### **4.1.2 Cox-regression**

Att använda "Cox proportional hazards regression" på ett datamaterial där proportionalitet ej råder kan självklart diskuteras. Kravet på att variablerna ska komma in linjärt på log-skala kan också skapa diskussion då det finns signifikanta kvadratiske termer.

Men till dessa problem finns lösningar i form av kategorisering och stratifiering. Efter dessa åtgärder håller samtliga antaganden och det finns ej längre några problem med att använda denna regression.

Interaktionsvariabler är något som medvetet valts bort i denna metod trots att dessa med stor säkerhet skulle kunna göra modellen bättre på att förklara vad som leder till att patienter går över till pump. Att dessa väljs bort beror på att frågan som utreds i denna studie är olika faktorerers effekt på sannolikheten att gå över till pump. Interaktionstermer bör ge en bättre modell då den förklarar en större del av variationen i responsvariabeln, men den blir också svårare att tolka. För att få en lättolkad modell med tydliga effekter väljs alltså interaktionstermer medvetet bort.

En ytterligare analys av  $HbA_{1c}$  utfördes för att finna den punkt där hasarden inte längre ökar genom ökat  $HbA_{1c}$ . Denna analys verifieras ej genom att kontrollera för proportionalitet vilket försvaras genom att den antas vara proportionell då  $HbA_{1c}$  ej haft signifikant interaktion med tid i föregående analys. Denna analys är dock ej tillräcklig och proportionalitetsantagandet bör verifieras även i denna modell för att få en tillförlitlig analys.

För att validera resultatet av Cox-regressionen utnyttjas en metod som använts i flertalet liknande studier (Lind, Odén, Fahlén, & Eliasson, 2010), (Simonsson, Sjölund, Bümbling, Ahlman, Nilsson, & Odén, 2007). Det är en speciell typ av Poissonregression som utvecklats just för denna typ av studier. Resultaten från denna typ av Poissonregression, som finns i bilaga 2, visar på liknande effekter som Cox-regressionen vilket medför att denna metod kan anses validerad.

## 4.2 Slutsatser

Den logistiska regressionen visar att det finns tydliga skillnader mellan patientgrupper. Beroende på kön, ålder, duration och vårdenhet använder patienter insulinpump i olika utsträckning. Kvinnor, unga och patienter med lång diabetesduration är grupper som använder insulinpump i större utsträckning än andra.

Vid analys av vad som föregår byte till insulinpump så bekräftas de resultat som den logistiska regressionen antytt. Genom användandet av tidsberoende variabler i Cox-regressionen visas också att högt långtidsblodssocker,  $HbA_{1c}$ , kraftigt ökar sannolikheten att gå över till insulinpump. Resultaten bekräftas av en speciell typ av Poissonregression som även den skattar hazardkvoter liknande dem från Cox-regressionen.

Insulinpump sänker  $HbA_{1c}$  och ett lägre värde på  $HbA_{1c}$  minskar risken för följsjukdomar. Då användandet av insulinpump har samma effekt i olika patientgrupper kan det tänkas att en satsning för att öka användandet i grupper som har låg andel pumpanvändare skulle kunna öka den totala nyttan av insulinpumpsanvändande.

Sammantaget så kan slutsatsen dras att det finns skillnader mellan patientgrupper i dagsläget och även i de olika gruppernas benägenhet att byta från andra insulinbehandlingar till insulinpump.

### 4.3 Förslag till vidare studier

Arbetet med denna studie har under vägens gång lett till flera nya frågeställningar. Flertalet av dessa ryms inte inom ramen för uppsatsen men tros vara av intresse för den fortsatta diabetesforskningen.

Vårdpersonal och patienter är båda delaktiga i beslutet att byta till insulinpump eller ej. Det är därför av intresse att utreda mer ingående om de skillnader mellan patientgrupper som visats i denna studie beror på patienternas preferenser eller om det snarare är vårdpersonalens attityd som påverkar. Enligt resultaten från tvärsnittsstudien genomförd med logistisk regression så skiljer sig andelen insulinpumpsanvändare åt mellan olika vårdenheter, vilket skulle kunna vara ett tecken på att det är attityden hos vårdpersonalen som skiljer sig åt mellan olika vårdenheter. En kvalitativ studie av attityden till insulinpump hos vårdpersonal och patienter skulle därför vara av intresse.

Denna studie har utrett vad som skiljer insulinpumpanvändare och patienter som inte använder insulinpump åt samt vad som föregår ett byte till insulinpump. På samma sätt skulle en studie kunna utreda vad som skiljer patienter som slutar med insulinpump åt från patienter som fortsätter med behandlingen samt vad orsaken till att patienter slutar med insulinpump är.

## Referenser

### Litteraturförteckning

Altman, D. (1991). *Practical statistics for medical research*. London: Chapman and Hall/CRC.

Bode, B., Steed, D., & Davidson, P. (1994). Long-term pump use and SMBG in 205 patients. 43.

Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 34, 187-220.

Dahl-Jørgensen, K., Brinchmann-Hansen, O., Hanssen, H., Ganes, T., Kierulf, P., Smeland, E., et al. (1986). Effect of near normoglycaemia for two years on progression of early diabetic retionopathy, nepropathy and neuropathy: the Oslo study. *BMJ* , 1195-1199.

DCCT Study Group. (1993). The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitu. *N Engl J Med*, 977-986.

Fahlén, M., Eliasson, B., & Odén, A. (2005). Optimization of basal insulin delivery in Type 1 diabetes: a retrospective study on the use of continuous subcutaneous insulin infusion and insulin glargine. *Diabet Med*, 22, 382-386.

Hanås, R. (2008). *Typ 1 diabetes hos barn, ungdomar och unga vuxna* (Fjärde uppl.). Västerås: BetaMed.

Kirkwood, B., & Sterne, J. (2003). *Essential Medical Statistics* (Andra uppl.). Malden, MA: Blackwell Publishing.

Kutner, M., Nachtsheim, C., Neter, J., & Li, W. (2005). *Applied Linear Statistical Models*. London: McGraw-Hill Higher Education.

Lind, M., Odén, A., Fahlén, M., & Eliasson, B. (2010). The shape of the metabolic memory of HbA1c: re-analyzing the Diabetes Control and Complications Trial with respect to time-dependent effects. *Diabetologia* .

Ludvigsson, J., Hermansson, G., Häger, A., Kernell, A., & Nordenskjöld, K. (1988). Adekvat substitution av insulinbristen grundläggande i behandlingen av diabetes hos unga. *Läkartidningen*, 2004-2008.

Nathan, D., Cleary, P., Backlund, J., Genuth, S., Orchard, T., Raskin, P., et al. (2005). Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N Engl J Med*, 2643-2653.

Petrie, A., & Sabin, C. (2009). *Medical Statistics at a Glance* (Tredje uppl.). Chichester: John Wiley & Sons Ltd.

Reichard, P. (1995). Are there any glycemic thresholds for the serious microvascular diabetic complications? *J Diabetes Complications*, 25-30.

Simonsson, T., Sjölund, K., Bümning, P., Ahlman, H., Nilsson, B., & Odén, A. (2007). Reducing uncertainty in health-care resource allocation. *British Journal of Cancer*, 1834-1838.

Stratton, I., Adler, A., Neil, H., Matthews, D., Manley, S., Cull, C., et al. (2000). Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational stud. *BMJ*, 405-412.

## **Övriga källor**

Sjukvårdsupplysningens hemsida:

<http://www.1177.se/artikel.asp?CategoryID=35955>, 2010-05-20

## Bilagor

### Bilaga 1 – Logistisk regression, test av interaktioner

För att komma fram till modellen som används i den slutgiltiga logistiska regressionen inkluderas kvadratiske termer på variablerna Ålder och Duration och interaktioner mellan de olika variablerna testas. Interaktionen Kön\*Duration blir signifikant då Duration är kontinuerlig men inte då variabeln är kategoriserad. Interaktionerna mellan Vårdenhet\*Ålder och Vårdenhet\*Duration får en respektive två signifikanta termer bland de nio interaktionerna skapade utifrån Vårdenhets dummyvariabler. När det utförs simultana test på dessa variabler blir de signifikanta och alltså finns viss påverkan av de två interaktionerna. Dock beslutas att inte inkludera dess i modellen då målet är att få en enkel modell.

Modellerna där de kvadratiske termerna på variablerna Ålder samt Duration testas.

Parameter	Wald Chi-Square	Pr > ChiSq
Intercept	72,2	<0,0001
Vårdenhet 1	0,0	0,8234
Vårdenhet 2	19,8	<0,0001
Vårdenhet 3	13,5	0,0002
Vårdenhet 4	3,4	0,0654
Vårdenhet 5	0,1	0,7902
Vårdenhet 6	24,5	<0,0001
Vårdenhet 7	0,0	0,9583
Vårdenhet 8	4,0	0,0467
Vårdenhet 9	2,8	0,0970
Kön	90,8	<0,0001
Duration	111,3	<0,0001
Ålder	14,2	0,0002
Ålder <sup>2</sup>	51,4	<0,0001

Parameter	Wald Chi-Square	Pr > ChiSq
Intercept	103,3	<0,0001
Vårdenhet 1	0,1	0,7148
Vårdenhet 2	21,6	<0,0001
Vårdenhet 3	12,8	0,0003
Vårdenhet 4	4,0	0,0466
Vårdenhet 5	0,0	0,8794
Vårdenhet 6	27,1	<0,0001
Vårdenhet 7	0,0	0,8472
Vårdenhet 8	3,9	0,0493
Vårdenhet 9	3,1	0,0770
Kön	86,2	<0,0001
Duration	176,2	<0,0001
Duration <sup>2</sup>	103,0	<0,0001
Ålder	218,9	<0,0001

Modellerna innehållande interaktionerna Vårdenhet\*Ålder och Vårdenhet\*Duration samt de test som simultant utreder om interaktionstermen ska inkluderas i modellen.

Parameter	Wald Chi-Square	Pr > ChiSq
Intercept	51,1	<0,0001
Vårdenhet 1	0,2	0,6836
Vårdenhet 2	4,0	0,0463
Vårdenhet 3	0,1	0,7285
Vårdenhet 4	2,6	0,1038
Vårdenhet 5	0,1	0,7009
Vårdenhet 6	0,1	0,7031
Vårdenhet 7	0,9	0,3311
Vårdenhet 8	0,0	0,9246
Vårdenhet 9	2,1	0,1475
kvinna	88,3	<0,0001
Ålder	0,5	0,4726
Ålder2	13,3	0,0003
Duration	126,2	<0,0001
Duration2	71,7	<0,0001
Vårdenhet 1 * Ålder	0,1	0,7429
Vårdenhet 2 * Ålder	12,6	0,0004
Vårdenhet 3 * Ålder	0,8	0,3796
Vårdenhet 4 * Ålder	1,2	0,2712
Vårdenhet 5 * Ålder	0,2	0,6343
Vårdenhet 6 * Ålder	1,7	0,1987
Vårdenhet 7 * Ålder	1,0	0,3242
Vårdenhet 8 * Ålder	0,3	0,5779
Vårdenhet 9 * Ålder	0,9	0,3338

Parameter	Wald Chi-Square	Pr > ChiSq
Intercept	64,0	<0,0001
Vårdenhet 1	3,6	0,0562
Vårdenhet 2	0,4	0,5529
Vårdenhet 3	0,6	0,4497
Vårdenhet 4	0,1	0,7271
Vårdenhet 5	0,0	0,9906
Vårdenhet 6	2,2	0,1392
Vårdenhet 7	0,0	0,9909
Vårdenhet 8	2,5	0,1139
Vårdenhet 9	12,0	0,0005
kvinna	88,6	<0,0001
Age	0,1	0,7087
Ålder2	11,2	0,0008
Duration	112,2	<0,0001
Duration2	73,8	<0,0001
Vårdenhet 1 * Duration	5,1	0,0240
Vårdenhet 2 * Duration	2,8	0,0920
Vårdenhet 3 * Duration	1,3	0,2540
Vårdenhet 4 * Duration	0,3	0,5742
Vårdenhet 5 * Duration	0,0	0,9058
Vårdenhet 6 * Duration	1,5	0,2238
Vårdenhet 7 * Duration	0,0	0,9536
Vårdenhet 8 * Duration	0,4	0,5177
Vårdenhet 9 * Duration	8,4	0,0037

Linear Hypotheses Testing Results			
Label	Wald Chi-Square	DF	Pr > ChiSq
Vårdenhet * Ålder	26,5	9	0,0017

Linear Hypotheses Testing Results			
Label	Wald Chi-Square	DF	Pr > ChiSq
Vårdenhet * Duration	23,3	9	0,0056



Modellen där interaktionstermen Kön\*Duration är inkluderad samt de simultana test som utförs då variablerna Duration och Ålder kategoriserats.

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept	1	-2,5481	0,2978	73,2292	<0,0001	0,078
Vårdenhet 1	1	-0,0456	0,1478	0,0951	0,7577	0,955
Vårdenhet 2	1	-0,5115	0,1144	19,9791	<0,0001	0,600
Vårdenhet 3	1	0,4341	0,1189	13,3366	0,0003	1,544
Vårdenhet 4	1	-0,2029	0,1084	3,5039	0,0612	0,816
Vårdenhet 5	1	0,0295	0,1256	0,0551	0,8143	1,030
Vårdenhet 6	1	0,6712	0,1302	26,5839	<0,0001	1,957
Vårdenhet 7	1	-0,0259	0,1686	0,0236	0,8780	0,974
Vårdenhet 8	1	0,4028	0,1922	4,3928	0,0361	1,496
Vårdenhet 9	1	0,2288	0,1324	2,9869	0,0839	1,257
Kön	1	0,868	0,1396	38,6859	<0,0001	2,382
Duration	1	0,1129	0,00993	129,32	<0,0001	1,119
Duration <sup>2</sup>	1	-0,0014	0,00017	71,4822	<0,0001	0,999
Ålder	1	0,0081	0,0139	0,3411	0,5592	1,008
Ålder <sup>2</sup>	1	-0,0005	0,00015	12,5608	0,0004	0,999
Kön*Duration	1	-0,0104	0,00502	4,3271	0,0375	0,990

Linear Hypotheses Testing Results			
Label	Wald Chi-Square	DF	Pr > ChiSq
Vårdenhet	98,2	9	<0,0001
Ålder	197,4	5	<0,0001
Duration	92	4	<0,0001
Kön*Duration	4,2	4	0,3822

Övriga interaktionstermer som testas blir inte signifikanta.

## Bilaga 2 – Cox-regression, kontinuerliga variabler

Som ett steg på vägen till att finna en korrekt och lättolkad modell utfördes Cox-regression på intressanta variabler i dess ursprungliga kontinuerliga form. Denna modell är ej justerad till proportionalitet och test av log-linjäritet mellan kovariater och responsvariabel har ej utförts. Utskriften visar dock att de valda variablerna verkar ha signifikant effekt på övergång till pump, varför dessa variabler används i den fortsatta analysen.

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
Vårdenhet 1	1	-0,18739	0,18302	1,0	0,3059	0,83
Vårdenhet 2	1	-0,90463	0,15226	35,3	<0,0001	0,41
Vårdenhet 3	1	0,79558	0,14522	30,0	<0,0001	2,22
Vårdenhet 4	1	-0,36694	0,14267	6,6	0,0101	0,69
Vårdenhet 5	1	-0,08972	0,15017	0,4	0,5502	0,91
Vårdenhet 6	1	0,45530	0,14429	10,0	0,0016	1,58
Vårdenhet 7	1	0,49956	0,24888	4,0	0,0447	1,65
Vårdenhet 8	1	0,09651	0,19350	0,2	0,6180	1,10
Vårdenhet 9	1	0,36151	0,19479	3,4	0,0635	1,44
HbA <sub>1c</sub>	1	0,08116	0,00823	97,3	<0,0001	1,09
Kreatinin	1	-0,04210	0,01700	6,1	0,0136	0,96
E/kg	1	-1,86576	0,19308	93,4	<0,0001	0,16
Duration	1	0,01046	0,00384	7,4	0,0064	1,01
Ålder	1	-0,05171	0,00384	181,7	<0,0001	0,95
Kön	1	0,44374	0,08187	29,4	<0,0001	1,56

### Bilaga 3 – Tidsberoende Cox-Regression (Non-proportional), SAS-kod

Följande kod är ett exempel på hur Cox-regression har utförts i SAS. Just denna kod utför stratifierad tidsberoende Cox-regression och ger resultatet som presenteras som den slutgiltiga modellen, i avsnitt 3.2.6.

```
/*- Macro som skapar tidsaxel, samt tilldelar macrovariabeln n det totala antalet tidpunkter då event inträffar --*/  
/*- Selektar samtliga tidpunkter där event inträffar för någon individ för att sedan endast utföra analys vid dessa tidpunkter.  
    Detta utförs då Cox-regression endast analyserar tidpunkter då event inträffat. --*/
```

```
%MACRO COX_n /STORE;  
%GLOBAL n dura;  
PROC SORT DATA=diab.event_cox OUT=tid(keep=dur_month) NODUPKEY;  
  BY dur_month;  
RUN;  
DATA tid;  
  SET tid;  
  n=_n_;  
RUN;  
  
PROC SQL NOPRINT;  
  SELECT DISTINCT  
    't' || left(put(tid.n, 4.)) || '=' || left(put(tid.dur_month, 4.)) || ','  
    left(put(grp.n, 4.))  
  INTO :dura SEPARATED BY ',' :n  
  FROM tid, (select distinct max(n) as n from tid) as grp  
  ;  
QUIT;  
%MEND COX_n;
```

```
/*- Följande macron skapar dummyvariabler till Cox-regressionen utifrån olika kategoriska variabler. --*/
```

```
%MACRO HbA1c_dummy(j) / STORE; /* Skapar dummyvariabler för HbA1c */  
  IF hb[&j.] >= 5 AND hb[&j.] < 6 THEN N_HbA1c_5_6 =1; ELSE N_HbA1c_5_6 =0;  
  IF hb[&j.] >= 6 AND hb[&j.] < 7 THEN N_HbA1c_6_7 =1; ELSE N_HbA1c_6_7 =0;  
  IF hb[&j.] >= 7 AND hb[&j.] < 8 THEN N_HbA1c_7_8 =1; ELSE N_HbA1c_7_8 =0;  
  IF hb[&j.] >= 8 AND hb[&j.] < 9 THEN N_HbA1c_8_9 =1; ELSE N_HbA1c_8_9 =0;  
  IF hb[&j.] >= 9 THEN N_HbA1c_9__ =1; ELSE N_HbA1c_9__ =0;  
%MEND HbA1c_dummy;
```

```
%MACRO Upkg_dummy(j) / STORE; /* Skapar dummyvariabler för E/kg */  
  IF Upkg[&j.] >= 0.5 AND Upkg[&j.] < 1.0 THEN N_Upkg_1 =1; ELSE N_Upkg_1 =0;  
  IF Upkg[&j.] >= 1.0 AND Upkg[&j.] < 1.5 THEN N_Upkg_2 =1; ELSE N_Upkg_2 =0;  
  IF Upkg[&j.] >= 1.5 THEN N_Upkg_3 =1; ELSE N_Upkg_3 =0;  
%MEND Upkg_dummy;
```

```
%MACRO Dur_dummy(j) / STORE; /* Skapar dummyvariabler för Duration */  
/*- ELSE IF på grund av saknade värden i denna variabel --*/  
  IF dur[&j.] < 10 THEN N_Dur_0_10 =1; ELSE IF dur[&j.] = . THEN N_Dur_0_10 = .; ELSE N_Dur_0_10=0;  
  IF dur[&j.] >= 10 AND dur[&j.] < 20 THEN N_Dur_10_20 =1; ELSE IF dur[&j.] = . THEN N_Dur_10_20 = .;  
    ELSE N_Dur_10_20=0;  
  IF dur[&j.] >= 20 AND dur[&j.] < 30 THEN N_Dur_20_30 =1; ELSE IF dur[&j.] = . THEN N_Dur_20_30 = .;  
    ELSE N_Dur_20_30=0;  
  IF dur[&j.] >= 30 AND dur[&j.] < 40 THEN N_Dur_30_40 =1; ELSE IF dur[&j.] = . THEN N_Dur_30_40 = .;  
    ELSE N_Dur_30_40=0;  
%MEND Dur_dummy;
```

```

%MACRO Age_dummy(j) / STORE; /* Skapar dummyvariabler för Age */
  IF Age[&j.] < 20 THEN N_Age_0_20 =1; ELSE N_Age_0_20=0;
  IF Age[&j.] >= 20 AND Age[&j.] < 30 THEN N_Age_20_30 =1; ELSE N_Age_20_30 =0;
  IF Age[&j.] >= 30 AND Age[&j.] < 40 THEN N_Age_30_40 =1; ELSE N_Age_30_40 =0;
  IF Age[&j.] >= 40 AND Age[&j.] < 50 THEN N_Age_40_50 =1; ELSE N_Age_40_50 =0;
  IF Age[&j.] >= 50 AND Age[&j.] < 60 THEN N_Age_50_60 =1; ELSE N_Age_50_60 =0;
%MEND Age_dummy / STORE;

/*- Time Dependent Cox regression (Non-proportional) --*/
/*- Data justeras till proportionalitet genom stratifiering --*/
/*- Endast kod efter denna kommentar behöver köras efter att ovanstående macron körts lagrats --*/

%Cox_n;
ODS LISTING close;
ODS RESULTS off;
ODS PDF FILE="&home\pdf\Cox_dummy_tid_strata.pdf" CONTENTS=no;
TITLE "Time Dependent COX Regression on Diabetes Data";

PROC PHREG DATA=diab.cox ;
  MODEL Time*Pump(0)=N_Creat Kvinna /* N_Laser */
    N_HbA1c_5_6 N_HbA1c_6_7 N_HbA1c_7_8 N_HbA1c_8_9 N_HbA1c_9__ /* HbA1c */
    N_Upkg_1 N_Upkg_2 N_Upkg_3 /* Enheter per kilo */
    N_Dur_0_10 N_Dur_10_20 N_Dur_20_30 N_Dur_30_40 /*Duration - Tid sedan diagnosdatum i år */
    N_Age_0_20 N_Age_20_30 N_Age_30_40 N_Age_40_50 N_Age_50_60 /* Alder i år */
    /RISKLIMITS;

/*- Stratifierar på Vårdenhet --*/
  STRATA Sjukhus_namn;

/*- Walds simultana test på kategoriserade variabler --*/
  HbA1c: TEST N_HbA1c_5_6=N_HbA1c_6_7=N_HbA1c_7_8=N_HbA1c_8_9=N_HbA1c_9__=0;
  Upkg: TEST N_Upkg_1=N_Upkg_2=N_Upkg_3=0;
  Duration: TEST N_Dur_0_10=N_Dur_10_20=N_Dur_20_30=N_Dur_30_40=0;
  Age: TEST N_Age_0_20=N_Age_20_30=N_Age_30_40=N_Age_40_50=N_Age_50_60=0;

/*- Arayer vilka refererar till tidsberoende variabler --*/
  ARRAY tt{*} t1-t&n.;
  ARRAY hb{*} HbA1c1-HbA1c&n.;
  ARRAY Creat{*} Creat1-Creat&n.;
  ARRAY Upkg{*} Upkg1-Upkg&n.;
  ARRAY Dur{*} Dur1-Dur&n.;
  ARRAY Age{*} Age1-Age&n.;

/*- Fyller array tt{*} med tidsaxeln genom macro "Cox_n" --*/
  &dura.;

/*- Tilldelar alder och durations arrayer värden, justering för att göra beräkning i hela år utförs --*/
  DO i=1 TO dim(tt);
    dur[i]=first_dura_y + (tt[i]/12);
    Age[i]=first_age_y + (tt[i]/12);
  END;

/*- Om datum för event är tidigare än första tidpunkt i studien --*/
  IF Time < tt[1] THEN DO;
    N_Creat=0;
    * HbA1c - Under 5 är referens;
    N_HbA1c_5_6=0; N_HbA1c_6_7=0; N_HbA1c_7_8=0; N_HbA1c_8_9=0; N_HbA1c_9__=0;
    * U/kg - Under 5 är referens;
    N_Upkg_1=0; N_Upkg_2=0; N_Upkg_3=0;
    * Duration - 40+ är referens;
    N_Dur_0_10=0; N_Dur_10_20=0; N_Dur_20_30=0; N_Dur_30_40=0;
  END;

```

```

* AGE - 60+ är referens;
N_Age_0_20=0; N_Age_20_30=0; N_Age_30_40=0; N_Age_40_50=0; N_Age_50_60=0;
END;

/*- Om datum för event är efter sista tidpunkt - tilldela sista mätvärde --*/
ELSE IF time >= tt[&n.] THEN DO;
  N_Creat=Creat&n.;

  * HbA1c - Under 5 är referens;
%HbA1c_dummy(&n.);
  * Upkg - Under 5 är referens;
%Upkg_Dummy(&n.);
  * Duration - 40+ är referens;
%Dur_dummy(&n.);
  * Age - 60+ är referens;
%Age_dummy(&n.);
END;

/*- Om datum för event är mellan aktuell tidpunkt och nästkommande - tilldela aktuellt mätvärde --*/
ELSE DO i=1 TO dim(hb)-1;
  IF tt[i] <= Time < tt[i+1] then do;
    N_Creat= Creat[i];
    * HbA1c - Under 5 är referens;
%HbA1c_dummy(i);
    * Upkg - Under 5 är referens;
%Upkg_Dummy(i);
    * Duration - 40+ är referens;
%Dur_dummy(i);
    * Age - 60+ är referensgrupp;
%Age_dummy(i);
  END;
END;
RUN;
ODS PDF CLOSE;

```

## Bilaga 4 – En speciell typ av Poissonregression

I denna speciella typ av Poissonregression används samma variabler som i Cox-regressionen. Tidsberoende variabler uppdateras med senaste uppmätta värde. Den enda skillnaden i denna modell är att variabeln Kreatinin skall tolkas per enhet kreatinin istället för per tio enheter i Cox-regressionen.

### Vårdenhet 2 och Vårdenhet 10

Variabel	$\beta$	SE	HR	95% KI <sub>HR</sub>		p-värde
Konstant	-5,6477	1,7649				
Ålder	-0,0257	0,0072	0,97	0,96	0,99	0,0004
Kön	0,6098	0,1538	1,84	1,36	2,49	0,0001
min(tid sedan 1979, 25)	0,0506	0,0705	1,05	0,92	1,21	0,4724
max(tid sedan 1979-25,0)	0,5864	0,0599	1,80	1,60	2,02	0,0000
Duration	0,0130	0,0071	1,01	1,00	1,03	0,0658
E/kg	-0,6247	0,3809	0,54	0,25	1,13	0,1010
HbA <sub>1c</sub>	0,2038	0,0492	1,23	1,11	1,35	0,0000
Kreatinin	-0,0031	0,0023	1,00	0,99	1,00	0,1886

### Vårdenhet 1, Vårdenhet 5 och Vårdenhet 7

Variabel	$\beta$	SE	HR	95% KI <sub>HR</sub>		p-värde
Konstant	-6,7059	2,4202				
Ålder	-0,0263	0,0082	0,97	0,96	0,99	0,0014
Kön	0,4671	0,1836	1,60	1,11	2,29	0,0110
min(tid sedan 1979, 25)	0,1078	0,0958	1,11	0,92	1,34	0,2605
max(tid sedan 1979-25,0)	0,5444	0,0621	1,72	1,53	1,95	0,0000
Duration	0,0030	0,0077	1,00	0,99	1,02	0,6931
E/kg	-1,6288	0,472	0,20	0,08	0,49	0,0006
HbA <sub>1c</sub>	0,2380	0,069	1,27	1,11	1,45	0,0006
Kreatinin	-0,0006	0,0026	1,00	0,99	1,00	0,8146

### Vårdenhet 3, Vårdenhet 4, Vårdenhet 6, Vårdenhet 8 och Vårdenhet 9

Variabel	$\beta$	SE	HR	95% KI <sub>HR</sub>	p-värde	
Konstant	-1,8961	1,2729				
Ålder	-0,0137	0,0054	0,99	0,98	1,00	0,0113
Kön	0,2589	0,1259	1,30	1,01	1,66	0,0397
min(tid sedan 1979, 25)	-0,0738	0,0506	0,93	0,84	1,03	0,1448
max(tid sedan 1979-25,0)	0,5681	0,0504	1,76	1,60	1,95	0,0000
Duration	0,0104	0,0049	1,01	1,00	1,02	0,0347
E/kg	-0,7711	0,2697	0,46	0,27	0,78	0,0043
HbA <sub>1c</sub>	0,1872	0,0416	1,21	1,11	1,31	0,0000
Kreatinin	-0,0095	0,0035	0,99	0,98	1,00	0,0071

### Antal patienter i de olika grupperna

Grupp av vårdenheter	Antal patienter	Antal som fått pump
Vårdenhet 2 och 10	2615	194
Vårdenhet 1, 5 och 7	1146	136
Vårdenhet 3, 4, 6, 8 och 9	2432	296