

Delrapport Stöd för rätt sjukskrivning 2017

# Prediktiva modeller för SRS piloten

Kristina Alexanderson

Gino Almondo

Matteo Bottai

Emilie Friberg

Paolo Frumento

Viktor Granvald

Pontus Josefsson

Kerstin Nilsson

Arvid Sondén

Avdelningen för försäkringsmedicin samt Enheten för biostatistik  
Karolinska Institutet

# Innehållsförteckning

## 1 Innehållsförteckning

Innehållsförteckning .....	2
2 Sammanfattning.....	3
Förord: Om SRS-projektet .....	4
3 och ett bedömningsstöd inom sjukskrivningsområdet.....	4
4 Förkortningar.....	5
5 Bakgrund.....	6
5.1 Syfte.....	7
6 Material och metod.....	7
6.1 Registerdata .....	7
6.1.2 Utfallsvariabel i A1 och A2.....	9
6.1.3 Inledande analyser och skapande av datasets .....	9
6.2 Del A1 .....	9
6.2.1 Variabler i Del A1.....	9
6.2.2 Analyser i A1 .....	11
6.3 Del A2 .....	12
6.3.1 Val av sjukskrivningsdiagnoser .....	13
6.3.2 Val av variabler i A2 analyserna.....	15
6.3.3 Från en initial till en specificerad modell med ännu färre variabler .....	16
6.3.4 Val av förklarande variabler för initiala samt specifika modeller .....	16
6.3.5 Utvärderingar av modellerna .....	17
6.3.6 Frågor för användning av allmänläkare.....	21
7 Slutkommentar .....	22
8 Referenser, urval.....	25
9 Bilaga 1 .....	26
10 Bilaga 2.....	28

## 2 Sammanfattning

I denna delrapport avrapporteras två delprojekt från Karolinska Institutet, gjorda inom ramen för det övergripande SRS-projektet. Det övergripande syftet har varit att ta fram prediktiva modeller för individens risk för lång sjukskrivningslängd, där modellerna kommer att användas i ett webverktyg i en pilotstudie inom primärvården.

Syftet med det första delprojektet (A1) var att utveckla och testa tidigare framtagna modeller för prediktion av duration av sjukskrivningsfall i två olika diagnoser (F43: Stressreaktion samt M75: skulder- och axelledsbesvär) bland de personer i Sverige som under 2011 haft ett sådant sjukskrivningsfall som översteg 14 dagar. Länkad information från rikstäckande register från Försäkringskassan, Statistiska centralbyrån samt Socialstyrelsen användes i analyserna. De modeller som tagits fram i tidigare delprojekt inom SRS prövades på fyra sätt: genom att ett stort antal ytterligare variabler togs med i analyserna, att olika kategoriseringar av dessa variabler prövades, samt att analyserna genomfördes separerade på olika strata. Det fjärde sättet var att ett flertal andra typer av analyser prövades, både vad avser bredare, mer övergripande/granskande regressionsmodeller och vad avser många olika maskininlärningsmodeller. Ingen av de olika testade modellerna gav bättre värden än den i föregående delprojekt föreslagna modellen; *Piecewise Constant Hazards Model* (PCH).

Baserat på dessa analyser identifierades 14 olika faktorer som hade störst betydelse för duration av sjukskrivningsfallet. Dessa användes sedan i det andra delprojektet, A2, där syftet var att ta fram prediktiva modeller baserat på de i första delprojektet identifierade faktorerna för 21 olika diagnoser. Urvalet av diagnoser gjordes i samverkan med SRS-projektets ledningsgrupp. Ett centralt kriterium var att de skulle vara vanligt förekommande sjukskrivningsdiagnoser i primärvården. För att underlätta implementeringen i primärvården beslöts det att nio prediktiva variabler skulle tas fram för var och en av diagnoserna, varav tre variabler skulle ingå i samtliga modeller. Dessa var kön, ålder och boenderegion.

Materialet i detta delprojekt (A2) var samtliga sjukskrivningsfall som påbörjades under perioden 1/1 2010-30/6 2012 (dvs 2,5 år) och som blev minst 15 dagar långa. Samtliga nya sådana fall i var och en av de 21 diagnoserna analyserades var för sig. Analyserna baserades på motsvarande information som i A1. I analyserna togs de sex övriga variabler som hade högst prediktivt värde fram för var och en av diagnoserna.

De olika modellernas prediktiva styrka befanns överlag vara god. Resultaten, i form av modeller att användas samt en sammanställning av information som ska läggas in i webverktyget har levererats till SRS-projektets ledning, för pilottestning i primärvård.

### 3 Förord: Om SRS-projektet och ett bedömningsstöd inom sjukskrivningsområdet

Statistik från Försäkringskassan visar att omfattningen och längden av sjukskrivningar har ökat under de senaste åren, vilket leder till ständigt stigande kostnader för samhället. Denna trend har observerats i alla yrken, men mest bland kvinnor och framförallt i psykiatriska diagnoser och allt tyder på att en fortsatt stigande trend. Samtidigt som sjukskrivning är en viktig del av vård och behandling visar forskning att långvariga sjukskrivningar kan leda till en försämrad hälsa, ekonomi och relationer för den enskilda individen.

I rapporten ”*Värdet av digital teknik i den svenska vården*” lyfter McKinsey & Company behovet av effektiva beslutstöd och avancerade datadrivna analyser för att uppnå förbättrade bedömningar och processer i hälso- och sjukvården. Detta gäller även sjukskrivningsförfarandet.

Sedan 2014 har därför ett bedömningsstöd inom sjukskrivningsområdet utretts, först i en förstudie och sedan i kompletterande utredningar. Arbetet har ingått som en del i överenskommelsen mellan staten och Sveriges Kommuner och Landsting (SKL) om en kvalitetssäker och effektiv sjukskrivningsprocess. Arbetet drivs i projektform genom projektet Stöd för rätt sjukskrivning (SRS) där SKL är beställare.

Tanken är att utveckla ett statistikbaserat bedömningsstöd för att underlätta för hälso- och sjukvården att tidigt i processen och mer träffsäkert kunna identifiera individer med ökad risk för längre sjukskrivning och det eventuella behovet av särskilda eller samordnade åtgärder, för en snabbare återgång i arbete/sysselsättning eller för att helt undvika sjukskrivning.

Under 2016 utvecklades en prototyp av bedömningsstödet som testades av användare i hälso- och sjukvården. Resultaten från förstudien, utredningarna och prototyptesterna visar att bedömningsstödet har möjlighet att skapa nytta för både individen och samhället, det fyller ett behov hos användare inom hälso- och sjukvården och är möjligt att realisera.

Arbetet fortsätter 2017-18 med att utveckla en första version av bedömningsstödet och därefter genomföra pilottester på vissa utvalda vårdenheter inom hälso- och sjukvården. Uppdraget ingår i och finansieras via överenskommelsen för 2017-18 (En kvalitetssäker och effektiv sjukskrivnings- och rehabiliteringsprocess)<sup>1</sup>.

SRS-projektet har uppdragit åt Karolinska Institutet att utveckla prediktiva modeller för pilotversionen av bedömningsstödet. Denna delrapport redovisar KI:s arbete.

Mer information om projektet Stöd för rätt sjukskrivning finns på SKL:s webbplats: <http://skl.se/halsasjukvard/sjukskrivningochrehabilitering/sjukskrivningsmiljarden/rattsjukskrivningsto d.5229.html>

Anne Snis, projektledare,  
SRS-projektet

---

<sup>1</sup> <https://skl.se/download/18.1e9e054d159029c422f4b7e/1481874824519/Overenskommelse-kvalitetssaker-effektiv-sjukskrivnings-rehabiliteringsprocess-2017-2018.pdf>

## 4 Förkortningar

I rapporten används följande förkortningar och begrepp:

AIC	Akaike Information Criterion. Statistiskt kvalitetsmått rörande modellens kvalitet i förmåga att beskriva data jämfört med andra möjliga modeller.
BIC	Bayesian Information Criterion. Statistiskt kvalitetsmått rörande modellens kvalitet
Bruttodagar	Antal dagar med ersättning från Försäkringskassan för sjukskrivning eller sjuk- och aktivitetsersättning
Faktor	Ett värde på en variabel, till exempel kvinna i variabeln kön
H-region	Indelning av kommuner, efter hur urbana de är, såsom i Storstad, Mellanstor stad respektive Glesbygd (1, 2)
HRM	Hazards risk modell
ICD10	International Classification of Diagnoses, version 10 (3)
IMAS	Insurance Medicine All Sweden; den databas projektet baseras på
LISA	Longitudinell Integrationsdatabas för Sjukförsäkrings- och Arbetsmarknads-studier; rikstäckande register administrerat av Statistiska centralbyrån
Log-Likelihood loss	Andelen information i utfallet givet modellen kan förklara som förloras ifall en enskild variabel raderas ifrån modellen.
MiDAS	MikroData för Analys av Socialförsäkringen; rikstäckande register administrerat av Försäkringskassan
Nettodagar	Antal <i>hela</i> dagar med ersättning från Försäkringskassan för sjukskrivning eller sjuk- och aktivitetsersättning. Dagar med partiell frånvaro är hopslagna till hela dagar.
PCH	Piecewise Constant Hazards Model
Pseudo-R-Squared	En generalisering av R-squared avseende proportionen av variationen i utfallet som kan förklaras av de tillgängliga prediktorerna.
SA	Sjuk- och aktivitetsersättning (tidigare kallat förtidspension respektive sjukbidrag)
SKL	Sveriges kommuner och landsting
SRS	Stöd för rätt sjukskrivning; namn på det projekt som denna rapport ingår i
Totala antalet dagar	Antal dagar mellan anmälningsdagen och sista dagen med ersättning i ett sjukskrivningsfall
Variabel	En grupp faktorer som utesluter varandra, till exempel kön som består av kategorierna kvinna och man.

## 5 Bakgrund

Detta delprojekt utgår från det arbete som gjorts av Karolinska Institutet i tre tidigare delprojekt inom ramen för SRS-projektet, avseende prediktion av fortsatt sjukskrivning bland personer som har påbörjat ett sjukskrivningsfall (4-6).

I denna delrapport rapporteras de två första delarna av ett nu pågående delprojekt, dessa två delar benämns A1 och A2 i uppdraget och vi har valt att kalla dem så även i denna arbetsrapport. I avtalet formulerades dessa två delar så här:

### **“Development of predictive models for the SRS pilot operation**

A1. The aim of this task is to obtain more accurate predictive models than the ones developed in the previous project in 2016 (4), for diagnoses F43 (Reaction to severe stress, and adjustment disorders) and M75 (Shoulder lesions). This will be done by e.g. analysis of possible predictors for duration of a sick-leave spell; other categorizations of already included factors and some additional factors; stratification by different subgroups, for instance, age, sex, and comorbidity.

A2. Analysis and developing predictive models for other diagnoses, from 10 to 20 diagnoses. The choice of those diagnoses shall be done in discussion with the SRS project team. The choice of the first five diagnoses shall be done by the beginning of March 2017 and the rest in April 2017.”

Bakgrunden till ovanstående var att arbetet i det tidigare delprojektet skett med stor tidspress och att ytterligare analyser eventuellt skulle kunna ge bättre modeller (A1) Dessutom behövde motsvarande analyser göras även för andra sjukskrivningsdiagnoser (A2).

Projektet, som är tvärvetenskapligt, har genomförts i samarbete med forskare vid Avdelningen för försäkringsmedicin, Institutionen för klinisk neurovetenskap (CNS) samt vid Enheten för biostatistik vid Institutet för miljömedicin (IMM), dvs två institutioner inom Karolinska Institutet. Framför allt har följande nio personer vid Karolinska Institutet arbetat med projektet (i bokstavsordning):

- Kristina Alexanderson, professor i socialförsäkring, med dr, Avdelningen för försäkringsmedicin, Institutionen för klinisk neurovetenskap (CNS) (projektledare; kristina.alexanderson@ki.se)
- Matteo Bottai, professor i biostatistik, PhD in biostatistics, Enheten för biostatistik, Institutet för miljömedicin (IMM)
- Emilie Friberg, forskarassistent, Med dr, Avdelningen för försäkringsmedicin (CNS)
- Paolo Frumento, forskarassistent, PhD in statistics, Enheten för biostatistik (IMM)
- Viktor Granvald, statistiker, Avdelningen för försäkringsmedicin (CNS)
- Pontus Josefsson, statistiker, Avdelningen för försäkringsmedicin (CNS)
- Kerstin Nilsson, statistiker, Avdelningen för försäkringsmedicin (CNS)
- Syed Rahman, post doc, Med dr, läkare, Avdelningen för försäkringsmedicin (CNS)
- Arvid Sondén, statistiker, Avdelningen för försäkringsmedicin (CNS)

Dessutom har två personer från SRS-projektet; Gino Almondo och Marina Santini, medverkat i vissa delar av projektet. Projektledningen vid Karolinska Institutet har haft återkommande diskussioner med SRS-projektets ledning under hela projektiden.

## 5.1 Syfte

Syftet med detta delprojekt var att (A1) förfinas och utveckla de analyser som arbetats fram för två diagnosgrupper med många sjukskrivningsfall, dels F43 (Stressreaktion) dels M75 (skulder- och axelledsbesvär) samt att (A2) baserat på resultaten i A1, göra motsvarande analyser för ytterligare 18 diagnoser.

## 6 Material och metod

I detta projekt har fortsatta analyser genomförts avseende prediktion av vilka sjukskrivningsfall som kan riskera att bli långa/längre än förväntat. Analyserna utgår från det arbete som tidigare gjorts i tre delprojekt inom ramen för SRS-projektet (4-6) och har utvidgats betydligt från de tidigare analyserna. Projektet har bestått av följande två delar;

- dels (A1) avseende de personer som under år 2011 påbörjade ett nytt sjukskrivningsfall som blev 15 dagar eller längre, antingen i diagnosen F43 eller i diagnosen M75 (4-6),
- dels (A2) avseende de sjukskrivningsfall som påbörjades någon gång under perioden 1/1 2010 – 30/6 2012 och som blev 15 dagar eller längre, inom 21 olika diagnoser.

### 6.1 Registerdata

Analyserna i både A1 och A2 baseras på information inom forskningsprojektet Insurance Medicine All Sweden (IMAS), som leds och administreras inom Avdelningen för försäkringsmedicin, Karolinska Institutet. De data som arbetats med innehåller avidentifierade individdata, länkade från olika rikstäckande register för de personer i arbetsför ålder som registrerats som boende i Sverige vissa år. Projektet har godkänts av den Regionala etikprövningsnämnden i Stockholm. För detta projekt har data från följande sex register använts:

#### 6.1.1.1 Från Statistiska centralbyrån:

- LISA (7) (*Longitudinell Integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier*) för information om följande sociodemografiska variabler för året innan det studerade sjukskrivningsfallet började (2009-2011): ålder, kön, födelseland, högsta utbildningsnivå, yrke, sektor, familjesituation, civilstånd samt boendekommun.

#### 6.1.1.2 Från Socialstyrelsen

- *Dödsorsaksregistret* information om datum för död 2010 - oktober 2014.
- *Patientregistret*, dels från *Slutenvårdsregistret*, dvs sjukhusinläggningar: datum, antal dagar, samt huvud- och bidiagnoser under åren 2008-2012, dels från *Öppenvårdsregistret*, vad avser: datum samt diagnoser för läkarbesök i specialiserad öppenvård under åren 2008-2012. Detta register innehåller inte information om vårdbesök i primärvården.
- *Läkemedelsregistret* för information om utköp av receptbelagd medicin 2009-2011 (datum, ATC-kod).

#### 6.1.1.3 Från Försäkringskassan

- *MiDAS (MikroData för Analys av Socialförsäkringen)* (8): information om sjukskrivningsfall som ersatts av Försäkringskassan under åren 2008-2014. Ur registret hämtades information om sjukskrivningsfallens startdatum (startdatum 1/anmälningsdatum samt startdatum 2 vilket är den första dagen med ersättning från Försäkringskassan) och slutdatum, omfattning (grad

av deltid eller om heltid), antal dagar med ersättning från Försäkringskassan (både brutto- och nettodagar), totala antalet dagar i fallet (inklusive dagar utan ersättning från Försäkringskassan), försäkradtyp (typ av sysselsättning) vid sjukskrivningsfallets start, sjukskrivningsdiagnos (huvuddiagnosen i den första sjukskrivningsperioden i sjukskrivningsfallet) samt om fallet påbörjades med förebyggande sjukpenning eller rehabiliteringspenning. Motsvarande information om sjuk- och aktivitetsersättningar (tidigare kallat förtidspension respektive sjukbidrag) inhämtades också från MiDAS. Detta innebär att information från både hel- och delfallsfiler i MiDAS har använts.

- Försäkringskassans register över personer som under åren 2009-2011 uppnått maximal tid i sjukförsäkringen och datum för detta.

- Försäkringskassans register över personer som fått sin sjukpenning indragen samt datum för detta under åren 2009-2011.

Information om sjukskrivningsdiagnos från Försäkringskassans MiDAS databas finns endast på treställighetsnivå för ICD10 (International Classification of Diagnoses, version 10<sup>(3)</sup>). Det innebär att information om sjukskrivningsdiagnoser på mer detaljerad nivå inte funnits tillgängligt i projektet.

I MiDAS registreras sjukskrivningsfall där Försäkringskassan har betalat ut ersättning. Det innebär att för de flesta anställda, där arbetsgivaren betalar sjuklön under de första 14 dagarna i ett sjukskrivningsfall, saknas information i MiDAS om sjukskrivningsfallet om det inte överstiger 14 dagar. För arbetslösa och vissa andra grupper för vilka Försäkringskassan betalar ut ersättning tidigare i sjukskrivningsfallet finns dock kortare fall registrerade i MiDAS. För att inte introducera en bias relaterat till dessa grupper inkluderades i denna rapports analys endast de sjukskrivningsfall som blev 15 dagar eller längre.

#### 6.1.1.3.1 Tre mått på antal sjukskrivningsdagar

Tre variabler används för beskrivning av antal dagar i ett sjukskrivningsfall. Den första beskriver antalet dagar från anmälningsdagen (startdatum 1) till den sista dagen med ersättning, alltså då sjukskrivningsfallet avslutades, detta kallas här för *totala antalet dagar* i fallet. När data från MiDAS inom ramen för projektet behandlades (datacleaning på engelska) slogs alla sjukskrivningsfall, som startade inom fem dagar efter att ett tidigare fall avslutats, samman med det tidigare fallet. Detta på grund av den så kallade återinsjuknanderegeln, som innebär att personen får ersättning från Försäkringskassan redan från dag 1 i det nya fallet. För sjukskrivningsfall med ersättningstyp "förebyggande sjukskrivning" under fallet användes 14 dagar som gräns för sammanslagning, på grund av att sådana fall ofta upprepas inom det tidsintervallet. Vid beräkning av *totala antalet dagar* inkluderades inte de mellanliggande dagarna, det vill säga dagar utan ersättning från Försäkringskassan togs inte med.

Den andra variabeln är antal *bruttodagar*, som beskriver antalet dagar med ersättning från Försäkringskassan. Dessa dagar är i de flesta fall lika med antalet dagar mellan den första dagen med ersättning (startdatum 2) och den sista dagen med ersättning (slutdatum), men det finns sjukskrivningsfall där det förekommer dagar mitt i fallet där ingen ersättning har utbetalats, detta till exempel på grund av uttagen semester. Dessa dagar utan ersättning är inte medräknade i variabeln bruttodagar men de är medräknade i variabeln totala antalet dagar.



I den tredje variabeln har antal bruttodagar räknats om till antalet *nettodagar*, det vill säga, antalet hela dagar med ersättning från Försäkringskassan. Till exempel omräknas fyra bruttodagar med en omfattning om 25 % av heltid till en nettodag.

### 6.1.2 Utfallsvariabel i A1 och A2

Utfallet i analyserna i såväl A1 som A2 var durationen av ett sjukskrivningsfall. Detta mättes på olika sätt i de analyserna. I vissa analyser användes det totala antalet dagar från sjukskrivningsfallets dag ett tills det avslutades. För de fall som blev längre än 1000 dagar sattes då den längsta durationen till 1000 dagar.

### 6.1.3 Inledande analyser och skapande av datasets

Inledningsvis genomfördes övergripande regressionsanalyser med avseende att minska antalet variabler som skulle användas och studeras i nästa steg av analyser (såväl A1 som A2). Initialt granskades närmare tvåhundra variabler utifrån deras individuella relationer till utfallet (totala antalet sjukskrivningsdagar) via bl.a. många olika regressionsanalyser. Dessa övergripande analyser gjordes för att få en överblick av samtliga tillgängliga variabler och deras samband med och påverkan på utfallet, samt att skapa underlag för vidare avvägningar i kommande analyser. I de fall variablerna saknade relation (eller prediktiv kraft) till utfallsmåttet togs dessa bort. Ytterligare avgränsningar av variablerna i datasetet gjordes utifrån teoretiska argument rörande relevansen av att använda dessa som prediktorer av utfallsvariabeln.

## 6.2 Del A1

I tidigare publicerade rapporter från detta projekt har vi bl.a. studerat vilka faktorer som kan predicera sjukskrivningsfallets duration bland de personer som under året 2011 påbörjade ett nytt sjukskrivningsfall som blev minst 15 dagar långt inom två olika diagnoser (F43 respektive M75) (6). De variabler som då användes som möjliga prediktiva faktorer var patientens kön, ålder, sysselsättningsstatus, sjukskrivningens omfattning i början av fallet, månad då sjukskrivningen initierades samt om personen då hade en pågående partiell sjuk- eller aktivitetsersättning eller inte.

I analyserna i A1 var syftet att fördjupa dessa analyser, genom att:

- inkludera fler variabler än tidigare i analyserna
- testa olika kategoriseringar av variablerna
- genomföra stratifierade analyser samt
- pröva andra analysstrategier.

### 6.2.1 Variabler i Del A1

Utifrån de inledande analyserna beskrivna ovan inkluderades följande 28 variabler i de fortsatta analyserna inom del A1:

- Kön
- Ålder (år)
- Boenderegion (H-regioner) (1, 2)
- Boendekommun (geografisk indelning)
- Utbildningsnivå (tre nivåer)
- Familjesituation (gift/sambo med hemmavarande barn, gift/sambo utan hemmavarande barn/ensamstående med hemmavarande barn/ensamstående)
- Civilstånd (gift/i registrerat partnerskap, ej detta)
- Födelseland
- Typ av arbete (white/blue collar)
- Sektor (privat, offentlig arbetsgivare)
- Sjukskrivningsfallets duration (se nedan)
- Typ av sysselsättning vid sjukskrivningsfallets början (i arbete, arbetslös, föräldraledig, studerande)
- Omfattning av sjukskrivning vid fallets start (25, 50, 75 eller 100 %)
- Sjukskrivning under de 12 månaderna före det nuvarande fallets start (dvs år -1) (antal brutto- och nettodagar, sjukskrivningsdiagnos)
- Sjukskrivning under de 13-24 månaderna före det nuvarande fallets start (dvs år -2) (antal brutto- och nettodagar, sjukskrivningsdiagnos)
- Sjuk- och aktivitetsersättning under de 12 månaderna före det nuvarande fallets start (antal brutto- och nettodagar, diagnos) (dvs år -1)
- Sjuk- och aktivitetsersättning under de 13-24 månaderna före det nuvarande fallets start (antal brutto- och nettodagar, diagnos) (dvs år -2)
- Pågående partiell sjuk- eller aktivitetsersättning vid sjukskrivningsfallets start (grad och diagnos)
- Tidigare uppnått maximalt antal dagar i sjukförsäkringen (år 2010-2011; ja/nej)
- Tidigare fått sjukpenning indragen (mars 2009-2011)
- Emigration (årtal)
- Utköp av receptbelagd medicin under de 12 månaderna före det nuvarande sjukskrivningsfallets start (dvs år -1) (antal köp, typ av läkemedel)
- Besök i specialiserad öppenvård under de 12 månaderna före det nuvarande sjukskrivningsfallets start (dvs år -1) (datum, antal gånger, diagnoser)
- Besök i specialiserad öppenvård under de 13-24 månaderna före det nuvarande sjukskrivningsfallets start (dvs år -2) (datum, antal gånger, diagnoser)
- Sjukhusinläggning under de 12 månaderna före det nuvarande sjukskrivningsfallets start (dvs år -1) (datum, (omräknades till antal dagar inlagd) diagnoser)
- Sjukhusinläggning under de 13-24 månaderna före det nuvarande sjukskrivningsfallets start (dvs år -2) (datum (omräknades till antal dagar inlagd), diagnoser)
- Vårdbesök i specialiserad öppenvård eller inläggning på sjukhus när sjukskrivningsfallet påbörjas (ja/nej)
- Död (datum)

I analyserna av specialiserad öppenvård respektive av slutenvård (dvs, inläggning på sjukhus) exkluderades sådan vård som var pga okomplicerad förlossning (ICD10 kod O80) samt utredningar, screening etc (ICD10 koder Z00-Z99 – med undantag av Z73.0 'utbrändhet', som inkluderades bland psykiska diagnoser).

## 6.2.2 Analyser i A1

I analyserna i A1 studerades de ovan listade variablerna med utgångspunkt i prediktiv kraft.

### 6.2.2.1 Olika kategoriseringar

I ett första steg prövades flera olika kategoriseringar av variablerna. Ett stort antal olika typer av analyser genomfördes för att studera hur olika interna kategoriseringar av dessa variabler påverkade dess individuella prediktiva kraft samt modellens förklaringsförmåga överlag. Log-Lik-loss kriterier användes för att bedöma den individuella prediktiva kraften av de enskilda prediktorerna, och Akaike Information Criterion (AIC) samt Bayesian Information Criterion (BIC) värden jämfördes för att bedöma modellernas förmåga att förklara utfallet. Baserat på dessa analyser komprimerades dataseten till variabler kategoriserade i de nivåer som presenterade bäst prediktiv förmåga. En hörnsten i teorin bakom prediktiva modeller är att om det inte finns någon linjär relation mellan en potentiell prediktor och utfallet, kommer denna variabel inte heller ha någon prediktiv kraft på utfallet. Dock kan variabler tillsammans utgöra en roll i en förklarandemodell (exempelvis i interaktioner). Utifrån detta prövades även ett stort antal olika linjära prediktionsmodeller baserat på de olika variablernas nivåer av kategorier. Ett exempel på detta är variabeln Vårdbesök (läkarbesök i öppen specialistvård under de 12 månaderna före datum då det aktuella sjukskrivningsfallet påbörjades). Denna variabel kategoriserades först på flera olika sätt, såsom, när det gäller analyserna av de med ett första sjukskrivningsfall i diagnosen F43; (1) antal läkarbesök pga. annan psykisk diagnos (dvs alla psykiska diagnoser utom F43) samt besök vid psykiatrisk mottagning om diagnos saknas, (2) antal läkarbesök pga. muskuloskeletal diagnos, (3) antal läkarbesök pga. annan diagnos samt (4) antal läkarbesök där information om diagnosen saknades. Denna uppdelning av variabeln (med 4 kategorier) testades i jämförelse med färre kategorier (över/under/lika med median antalet besök) för att utreda om de skiljde sig åt med avseende på dess prediktiva kraft i modellen.

Denna typ av detaljerade analyser, där olika kategoriseringar av variabler genomfördes för majoriteten av variablerna (dvs för dem där flera olika kategoriseringar var möjliga), för att avgöra (a) ifall det innebar en statistisk skillnad att ha fler/färre/andra alternativ (t.ex. ifall modellen fick bättre eller sämre passform) samt (b) ifall det fanns signifikanta skillnader i den enskilda variabelns påverkan på utfallet. Log-likelihood loss (dvs hur mycket variation modellen kan beskriva, och som förloras om variabeln/faktorn/prediktorn tas bort från modellen) och pseudo-R-squared (uppskattning av andelen variation i utfallet som kan förklaras av variationen i prediktorn respektive modellen) användes som kriterier för att avgöra variablernas gemensamma påverkan på modellen i helhet samt individuell prediktiv kraft. Utifrån dessa avvägningar kollapsades de variabler där de olika nivåerna inte påverkade utfallet eller modellen signifikant, och nya variabler skapades vilka användes i de senare stegen av analyserna.

### 6.2.2.2 Stratifieringar

Dessutom genomfördes stratifierade analyser utifrån fyra olika aspekter. Dessa var kön (två strata - dvs personerna delades upp i två grupper, kvinnor för sig och män för sig), ålder (fem respektive två strata (16-44 samt 45-64 år)), sysselsättning (två strata), respektive om sjukskrivningsfallet antogs ha initierats från primärvården eller ej (två strata). Att analyserna

stratifierades för dessa variabler innebär t.ex. att analyserna gjordes för kvinnor separat och för män separat för att studera om detta innebar skillnader i prediktionsmodellerna – utifrån hypoteser om att mekanismerna bakom sjukskrivningsduration skulle kunna vara olika för kvinnor och män (9). Stratifieringen ledde inte till att modellernas prestanda förbättrades, dvs de variabler som hade betydelse för utfallet var desamma även i de olika specifika strata/grupperna. Personer som tidigare varit långtidssjukskrivna hade t.ex. högre risk för långtidssjukskrivning jämfört med personer, som när fallet påbörjades, var i arbete oavsett strata. De olika variabler som påverkade detta skilde sig alltså inte åt mellan de olika stratifierade grupperna.

#### 6.2.2.3 Val av modell

Ett flertal olika metoder och maskininlärnings algoritmer såsom Naive Bayes Classifier, Bayesian Network, Decision tree, Logistic Regression, Multilayer Perceptron, Random Forest, Support Vector Machines prövades även. Dessa modeller användes för att predicera om sjukskrivningsfallet t.ex. blev längre än 90 dagar.

Initialt testades en *Piecewise Constant Hazard Model* (PCH) (10) baserat på resultaten från de tidigare delstudierna i projektet (4, 6). PCH-modellen är en flexibel modell för överlevnadsanalyser (på engelska "survival analysis") och beskriver hur risken för att en händelse inträffar är associerad med de studerade prediktorerna. Modellen hanterar även att tiden till händelsen inträffar kan varieras. Till exempel kan, i en typ av analyser, män ha en högre risk för förtida död än kvinnor, medan kvinnor lever längre. Det innebär att effekten av manligt kön i modellen kommer att öka eller minska risken, beroende på tidpunkten som väljs i den enskilda modellen. Då PCH modellen kan dela in tiden i flera subintervall, gör detta att en enskild Hazards risk modell (HRM) kan beräknas för varje enskilt tidsintervall. Detta ger en sannolikhet att en händelse skall inträffa i ett valt tidsintervall givet prediktorerna.

Ingen av de olika testade modellerna gav bättre värden än den i föregående delprojekt föreslagna PCH modellen.

Initialt genomfördes alltså ett mycket stort antal olika analyser för att studera om utfallet blev olika jämfört med tidigare analyser, om ytterligare variabler tillfördes, om de tidigare använda variablerna kategoriserades på andra sätt samt om stratifierade analyser genomfördes. Avsikten var, förutom att ta fram bästa modellen, även att begränsa den stora mängden variabler som fanns tillgänglig. Det finns ett stort teoretisk samt statistiskt värde i att studera hur olika faktorer kan bidra till att predicera längden på ett sjukskrivningsfall. Dock är det samtidigt problematiskt då stora mängder variabler kan uppge missvisande samband och påverkan i prediktioner som ej bidrar till att förbättra modellen i stort eller har god teoretisk grund. Ett uttryckligt krav från SRS-projektet på modellen var dessutom att få variabler skulle vara med. Ovanstående analyser ledde fram till att 14 variabler identifierades som havande högst prediktivt värde. Dessa listas nedan under rubrik 5.2.2.

### 6.3 Del A2

Parallellt med analysarbetet i A1 gjordes ett omfattande arbete för att etablera analysfiler för arbetet i A2, där samtliga nya sjukskrivningsfall som blev 15 dagar eller längre under en

period av 2,5 år (1/1 2010-30/6 2012) skulle analyseras, med användande av de variabler som identifierades med högt prediktiv värde i A1.

### 6.3.1 Val av sjukskrivningsdiagnoser

Ett flertal möten och diskussioner hölls med SRS-projektets ledning vad avser val av diagnoser för del A2. Diagnoserna skulle vara vanligt förekommande i primärvården, vara vanliga sjukskrivningsdiagnoser samt en relativt stor andel av sjukskrivningarna skulle bli längre än förväntat eller innebära olika typer av problem för allmänläkaren när det gäller hantering av sjukskrivning. Efter många diskussioner och genomgång av data utifrån dessa kriterier kom vi överens om nedanstående:

Av samtliga drygt en miljon sjukskrivningsfall som påbörjades någon gång under perioden 1/1 2010 till 30/6 2012 och som blev 15 dagar eller längre och där personen, när fallet påbörjades, var i åldersspannet 18-64 år, valdes samtliga fall som började med förmånstypen sjukpenning, där den sjukskrivne hade bott i Sverige året innan fallet påbörjades, inom 21 sjukskrivningsdiagnosgrupper. Vilka de är framgår av tabell 1 nedan. När det gäller depressioner så slog vi samman två ICD10 koder, F32 och F33, dvs depressiv episod och återkommande depressiv episod, då det framkommit att F32 inte sällan använts som diagnoskod även för återkommande depressioner. På motsvarande sätt slogs sjukskrivningsfall i de två diagnosgrupperna R52 och R53 samman, även om de inte är helt lika; Smärta och värk som ej klassificeras annorstädes respektive Sjukdomskänsla och trötthet. Båda dessa grupper är symtomdiagnoser och tillsammans utgjorde de en tillräckligt stor grupp för analyser.

I del A2 kodades de sjukskrivningsfall som avslutades med sjuk- eller aktivitetsersättning som att fallet blev 1000 dagar. Sjukskrivningsfall som avslutades i och med att personen avled eller emigrerade följdes till de slutade (slutdatum). Antal sjukskrivningsfall i respektive studerad diagnos, samt andel av dem som blev längre än 90 dagar respektive andel som blev längre än 1000 dagar, enligt ovanstående kriterier, framgår av tabell 1.

**Tabell 1.** Studerade sjukskrivningsdiagnoser; antal sjukskrivningsfall, andel av fallen som blev längre än 90 dagar respektive minst 1000 dagar, per inkluderad diagnosgrupp.

Diagnoskod, ICD10	Diagnos	Antal sjukskrivningsfall	Sjukskrivningsfall som blev >90 dagar n (%)	Sjukskrivningsfall som blev ≥1000 dagar <sup>2</sup> n (%)
F31	Bipolär sjukdom	6819	4436 (65)	75 (1,0)
F32+F33	Depressiv episod/Recidiverande depressioner	77721	40607 (52)	368 (0,4)
F41	Andra ångestsyndrom	26345	12650 (48)	152 (0,5)
F43	Anpassningsstörningar och reaktion på svår stress	83443	30868 (37)	200 (0,2)
G56	Mononeuropati i övre extremitet	19908	1733 (9)	10 (0,0)
I63	Cerebral infarkt orsakad av trombos i precerebrala artärer	3841	2768(72)	61(0,0)
M16	Höftledsartros	8694	5125 (59)	12 (0,1)
M17	Knäartros	12098	6391 (53)	10 (0,0)
M19	Andra artroser	6068	2813 (46)	15 (0,2)
M23	Andra sjukliga förändringar i knäled	11654	2395 (21)	<8 (0,0)
M51	Andra sjukdomar i mellankotskivorna	10269	4677 (46)	30 (0,2)
M53	Ryggsjukdomar som ej klassificeras annorstädes	8909	4075 (46)	24 (0,2)
M54	Ryggvärk	64048	16941 (26)	148 (0,2)
M75	Sjukdomstillstånd i skulderled	20049	6874 (34)	17 (0,0)
M77	Andra entesopatier	11745	2460 (21)	8 (0,0)
M79	Sjukdomstillstånd i mjukvävnad ej klassificerade annorstädes	20973	7729 (37)	48 (0,0)
R52+R53	Smärta och värk ej klassificerade annorstädes/ Sjukdomskänsla och trötthet	8656	3007 (35)	49 (0,5)
S52	Fraktur på underarm	14931	2496 (17)	8 (0,0)
S62	Fraktur på handled och hand	11557	1287 (11)	<8 (0,0)
S82	Fraktur på underben, inklusive fotled	15147	4760 (31)	20 (0,1)
S83	Luxation och distorsion av knäets leder och ligament	10916	2538 (23)	<8 (0,0)

<sup>2</sup> Här ingår även de sjukskrivningsfall som avslutades med sjuk- eller aktivitetsersättning

### 6.3.2 Val av variabler i A2 analyserna

I analyserna i A2 användes initialt följande 15 (14 prediktorer samt 1 utfall) variabler

- Kön (man/kvinna)
- Ålder det år sjukskrivningsfallet började
- Geografisk boenderegion (5 kategorier, motsvarande den regionindelning som Försäkringskassan använder)
- Tidigare sjukskrivning, under de 12 månaderna innan datum när detta sjukskrivningsfall påbörjades (antal bruttodagar, kategoriserat)
- Antal läkarbesök i specialiserad öppenvård under de 12 månaderna innan datum för när detta sjukskrivningsfall påbörjades (0, under eller på medianen respektive över median för antalet besök bland dem som hade besök)
- Antal dagar patienten varit inlagd på sjukhus under de 12 månaderna innan sjukskrivningsfallet påbörjades (0, under eller på medianen respektive över medianantalet dagar bland dem som varit inlagda)
- Sysselsättningsstatus när sjukskrivningsfallet påbörjades (föräldraledig/arbetslös/yrkesarbetande/studerande)
- Högsta utbildningsnivå (grundskola/gymnasium/universitet)
- Sjukskrivningsgrad dag 15 i sjukskrivningsfallet (25/50/75/100%)
- Om personen hade en pågående partiell sjuk- eller aktivitetsersättning när sjukskrivningsfallet påbörjades (ja/nej)
- Födelseland (Sverige/Övriga norden/Övriga EU25/Övriga världen)
- Familjesituation (Gift/sambo med hemmavarande barn/ Gift/sambo utan hemmavarande barn/Ensamstående med hemmavarande barn/Ensamstående utan hemmavarande barn)
- Antal uttag av receptbelagd medicin under de 12 månaderna före sjukskrivningsfallet påbörjades (minst två uttag i separat läkemedelskategori, definierat efter första positionen i ATC-koden)
- Om patienten varit inlagd respektive haft läkarbesök i specialiserad öppenvård i samband med att sjukskrivningsfallet startade (ja/nej)
- Som utfall användes variabeln totalantalet sjukskrivningsdagar i fallet, i vissa analyser för hela fallet, i andra analyser kategoriserat som att fallet blev över ett visst antal dagar, t.ex. 90 eller 180 dagar.

### 6.3.3 Från en initial till en specificerad modell med ännu färre variabler

De i delprojektet A1 identifierade 14 variablerna applicerades i analyser av sjukskrivningsfallen inom var och en av de valda 21 diagnosgrupperna. Samtliga modeller hade samma utfall (Totala antalet dagar med ett tak på 1000 dagar, respektive att fallet blev över 90 dagar eller över 180 dagar långt).

### 6.3.4 Val av förklarande variabler för initiala samt specifika modeller

Ett uttryckligt krav på modellen från SRS-projektet var att få variabler skulle vara med. Därför valdes ovanstående 14 variabler, de som hade bäst prediktivt värde. Dessa arbetade vi först med i analyserna. Inom SRS-projektet behövde variabler identifieras som allmänläkaren kan fylla i test-webversionsverktyget, få tillgång till inom SRS-projektet/via journalsystem, alternativt fråga patienten om.

På begäran av SRS-projektets ledning sattes senare i projektet ett tak på totalt nio variabler per diagnos i respektive modell. Variablerna kön, region och ålder skulle då alltid ingå bland dessa nio. Målet med dessa analyser var då att utifrån grundmodellen (med 14 variabler) skapa specificerade modellerna bestående av 6 variabler utöver de tre automatiskt medtagna som representerade modellen med bäst prediktiv kraft för var och en av de 21 diagnosgrupperna. Vid tillämpning av modellen kommer information om kön och ålder att automatiskt hämtas i systemet och information om region att vara inlagd.

Utöver kön, ålder och region så valdes de sex variabler i de initiala modellerna som hade högst Log-likelihood loss, dvs. hur mycket information som modellen skulle förlora ifall just denna prediktor togs bort ifrån modellen. I alla modeller lades variabeln 'Sysselsättning vid sjukskrivningsfallets början' med (med dess fyra kategorier; anställd, student, föräldraledig och arbetslös) bland de sex då detta var en fråga som läkaren alltid måste fylla i sjukintyget i samband med en sjukskrivning. Sedan kördes modellerna om med endast de nio variablerna (dvs de "specifika" modellerna).

#### 6.3.4.1 Ytterligare bearbetning av modellerna

Ytterligare ett önskemål från SRS-projektets ledning var att undersöka om ett minsta möjliga antal kategorier inom varje variabel som tagits fram i ovanstående kunde uppnås utan att det negativt påverkade modellens kvalitet. Syftet var att minska antalet svarsalternativ för användning av modellerna för test i pilotstudien vid vårdcentraler.

Ett exempel på detta är variabeln familjesituation, som blev en av de sex variablerna för några av diagnoserna. I analyserna använde vi initialt fyra kategorier av denna variabel: [1] Gift/sambo utan hemmavarande barn, [2] Gift/sambo med hemmavarande barn, [3] Ensamstående med hemmavarande barn, samt [4] Ensamstående utan hemmavarande barn. Ifall man dikotomiserar dessa variabler kan man exempelvis endast ha de två variablerna [1] Gift/sambo, [2] Ensamstående, alternativt de två variablerna [1] leva med hemmavarande barn, [2] leva utan hemmavarande barn. I analyserna framkom att den första dikotomiseringen gav högst prediktivt värde (dvs [1] Gift/sambo, [2] Ensamstående). För dessa bearbetningar av modellvariablerna krävs det att det finns statistiska (dvs, att modellen passar/predicerar



bättre) samt teoretiska argument (tex ”Den mer specifika kategoriseringen ger en bättre/sämre bild av patientens bakgrundsinformation”).

För alla de variabler som här inkluderades som prediktorer som i de första modellerna hade flera kategoriseringsvärden (ex utbildningsnivå, familjesituation) så kördes dessa modeller om med kollapsade nivåer (dvs dikotomiserades enligt ovanstående exempel).

### 6.3.5 Utvärderingar av modellerna

Utvärderingar av modellerna har gjorts på flera sätt, nedan presenteras några av dessa, såsom framtagande av AIC och BIC värden för modellerna, modellernas Goodness-of-fit respektive modellernas informativitet.

#### 6.3.5.1 AIC och BIC värden för modellerna

AIC och BIC värden togs fram för samtliga modeller (för såväl de i A1 samt både för de initiala och för de mer specifika i A2, för var och en av diagnoserna, se tabell 2). Dessa kriteriemått avser generellt hur modellens förklaringsförmåga förbättras (mindre AIC i jämförelse mellan modellerna) eller försämras (högre AIC) då den valda kombinationen av variabler används för att beskriva data. BIC är ett liknande mått som AIC men anses vara mer restriktivt (11). I båda fallen är ett mindre värde ett tecken på en bättre modell. När antalet variabler reducerades var det inga stora förändringar i värdena för vare sig AIC eller BIC och i flertalet fall var de värdena t.o.m. lägre för de specificerade modellerna jämfört med för de initiala modellerna, vilket framgår av tabell 2, vilket alltså är positivt

**Tabell 2.** AIC och BIC värden för modellerna som inkluderade 14 respektive 9 variabler för de 21 diagnosgrupperna.

<b>Diagnoskod, ICD10</b>	<b>AIC 14 variabler</b>	<b>BIC 14 variabler</b>	<b>AIC 9 variabler</b>	<b>BIC 9 variabler</b>
F31	56594	61107	56459	59053
F32+F33	832682	838979	833306	837381
F41	269112	274772	269291	272988
F43	923202	929547	923792	927898
G56	171130	176485	171197	174996
I63	34584	38164	34515	37310
M16	96353	101040	96472	99180
M17	136801	141812	136956	139894
M19	70083	74606	69965	73012
M23	112582	117582	112452	115685
M51	120453	125360	120266	123146
M53	88095	92877	87872	90951
M54	647332	653652	647472	651616
M75	217912	223242	217858	220973
M77	115803	120794	115623	118844
M79	185336	190425	185181	188743
R52+R53	81181	85915	81074	84112
S52	144360	149513	144559	147665
S62	105741	110728	105655	108582
S82	159153	164308	159463	162482
S83	109408	114363	109301	112797

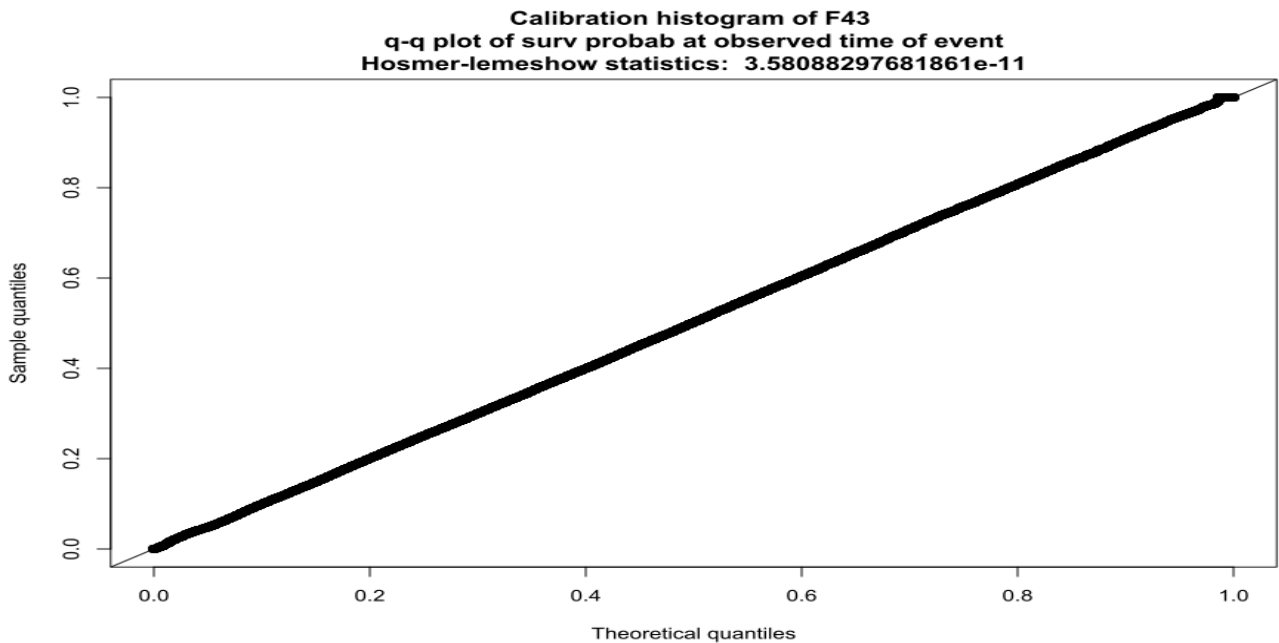
### 6.3.5.2 Goodness-of-fit

Ett goodness-of-fit test av 20 av modellerna genomfördes, för att utvärdera modellernas förmåga att uppskatta sannolikhet för långtidssjukskrivning. Resultat av dessa för diagnos F32 visas nedan i figur 1. Resultaten från motsvarande analys av 20 diagnoser framgår av Bilaga 1.

Modellernas goodness-of-fit uppskattades med det så kallade Hosmer-Lemeshow test genom att beräkna sjukskrivningsfallets sannolikhet att överstiga sin observerad längd, alltså  $P(T > t | x, \text{modell}, \theta)$ . Denna sannolikhet förväntas vara fördelad likformigt på intervallen  $[0,1]$  enligt teoremet Probability Integral Transform. Vi testade uniformiteten av modellens output grafiskt med en q-q graf. Q-q kurvan av en perfekt kalibrerad modell ligger på diagonalen.

Samtliga modeller visade sig passa data bra.

**Figur 1.** Resultat av Hosmer-Lemeshow testet för diagnosen F43.



### 6.3.5.3 Modellernas informativitet

Inom SRS-projektet har man även utvärderat modellernas förväntade beteende för avsett praktiskt användande. Modellerna kommer i praktiskt bruk initialt att användas för att estimeras den sjukskrivne patientens relativa risk (här oddskvot) för att vara sjukskriven längre en 90 dagar. Syftet är att hjälpa läkaren identifiera patienter med förhöjd risk relativt till "snittpatienten" med samma sjukskrivningsdiagnos genom att datorprogrammet tar fram en 'riskflagga' eller uppmärksamhetssignal. Vidare har SRS-projektgruppen bestämt att riskflaggan ska ha tre nivåer (enligt de "gyllene regler" för beslutsstödsystem som nämns i en rapport av Löfström (12, kapitel 4.2)). De tre nivåerna kommer att bli följande:

1. Lätt förhöjd risk
2. Måttligt förhöjd risk
3. Starkt förhöjd risk

Modellens output är en sannolikhet (en siffra i intervallet [0,1]) och översätts till de ovannämnda nivåerna genom att:

1. omvandla modellens output till odds
2. omvandla prevalens av långtidssjukskrivning (andel sjukskrivningsfall med samma huvuddiagnos som överstiger 90 dagar) till odds
3. Definiera oddskvoten mellan patientens beräknade risk (enligt modellen) och prevalensen bland de faktiska fallen i modellen
4. Tilldela patienten en riskflagga enligt följande
  - a. Om oddskvoten är  $<2$ : Lätt förhöjd risk
  - b. Om oddskvoten ligger i intervallet 2 till  $<4$ : Måttligt förhöjd risk
  - c. Om oddskvoten är 4 eller däröver: Starkt förhöjd risk

Utifrån detta har SRS-projektgruppen identifierat två mått för informativiteten, nämligen:

1. Larmfrekvens av risknivån ”Måttligt förhöjd risk”
2. Larmfrekvens av risknivån ”Starkt förhöjd risk”

Genom att ta fram siffror för dessa två mått kan man därför svara på frågan ”Hur stor andel av patienter sjukskrivna 15 dagar i respektive diagnos tilldelas respektive riskflaggning?”

Tabell 3 innehåller de ovannämnda larmfrekvenserna per diagnos samt den totala larmfrekvensen per diagnos, det vill säga summan av de två larmfrekvenserna. Observera att larmfrekvenserna är uttryckta i decimalform (dvs tex 0,1 motsvarar 10%). I tabell 3 har detta separerats för de två sammanslagna diagnosgrupperna, dvs för dem som modellerna är framtagna för: F32+F33 samt R52+R53.

**Tabell 3.** Beräknade 'larmfrekvenser' för respektive diagnos, baserade på av SRS-projektet framtagna nivåer

Diagnoskod; ICD10	Larmfrekvens; måttligt förhöjd risk	Larmfrekvens; starkt förhöjd	Larmfrekvens, totalt
<b>F31</b>	0,23	0,07	0,3
<b>F32</b>	0,13	0,02	0,15
<b>F33</b>	0,13	0,02	0,15
<b>F41</b>	0,15	0,06	0,2
<b>F43</b>	0,06	0,03	0,08
<b>G56</b>	0,14	0,03	0,16
<b>I63</b>	0,03	0	0,03
<b>M16</b>	0,11	0,01	0,12
<b>M17</b>	0,06	0	0,06
<b>M19</b>	0,09	0,03	0,12
<b>M23</b>	0,06	0,01	0,07
<b>M51</b>	0,09	0,04	0,13
<b>M53</b>	0,05	0,12	0,18
<b>M54</b>	0,06	0,04	0,1
<b>M75</b>	0,05	0,01	0,06
<b>M77</b>	0,07	0,01	0,08
<b>M79</b>	0,03	0,15	0,18
<b>R52</b>	0,02	0,12	0,14
<b>R53</b>	0,02	0,12	0,14
<b>S52</b>	0,03	0	0,03
<b>S62</b>	0,02	0	0,02
<b>S82</b>	0,01	0	0,01
<b>S83</b>	0,01	0	0,01

### 6.3.6 Frågor för användning av allmänläkare

Ett dokument togs fram med de sex variablerna med högst prediktivt värde för var och en av de 21 diagnosgrupperna (se bilaga 1). Variablerna är där listade i ordning utifrån prediktivt värde – de med högst värde listades först osv. Enda undantaget från detta är att variabeln sysselsättning listades först för samtliga diagnoser, eftersom läkaren vid en sjukskrivning, i sjukintyget ändå ska ge information vad avser sysselsättningsituation. Denna variabel finns därför redan ifyllt i föregående sjukintyg, dvs det som gäller för sjukskrivningsfallets första delperiod och skulle på motsvarande sätt som information om kön och ålder kunna hämtas direkt från systemet.

De 21 modellerna liksom dokumentet med frågorna har levererats till SRS-projektet.

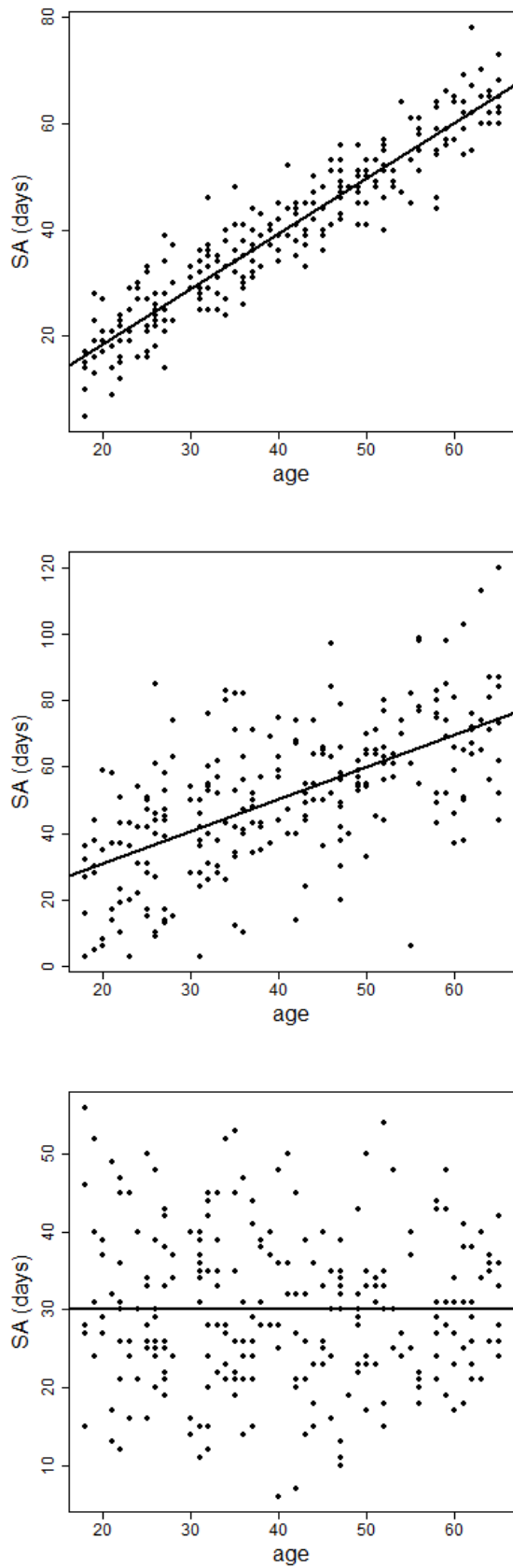
## 7 Slutkommentar

I dessa explorativa och omfattande analyser har syftet varit att ta fram användbara modeller för prediktion av om huruvida ett sjukskrivningsfall som pågått 15 dagar riskerar att bli långt/längre än förväntat. Avsikten är alltså att modellerna ska användas på individnivå.

En prediktiv modell syftar till en god estimering av distributionen av en utfallsvariabel ( $y$ ), baserat på en uppsättning möjliga prediktorer ( $x$ ). Vi kan t.ex. predicera durationen av ett sjukskrivningsfall, baserat på individbaserad information om t.ex. ålder, kön, sjukvårdshistorik, yrke, etc. En prediktiv modells prestanda beror på styrkan i associationen mellan tillgängliga prediktorer och utfallsvariabeln. Beakta t.ex. punkterna i de tre nedanstående spridningsdiagrammen (figur 2). I den första delfiguren finns det en mycket stark association mellan tillgängliga prediktorer och utfallsvariabeln. Om en person, George, är 30 år gammal kan vi vara säkra på att hans sjukskrivningsfall kommer att vara mellan 20 - 40 dagar. Om istället en annan person, Lukas, är 60 år gammal, kommer hans sjukskrivningsfall troligen att vara längre, ungefär 50 - 70 dagar. Även om sjukskrivningsfallens längd varierar starkt - mellan 10 - 80 dagar, kan vi predicera deras längd relativt säkert baserad på uppgift om ålder.

I den andra figuren är associationen svagare och graden av osäkerhet är mycket större. Till exempel är det möjligt att såväl George som Lukas sjukskrivningsfall blir mellan någonstans mellan 40 och 70 dagar.

I mönstret i den tredje delfiguren predicerar ålder inte sjukskrivningsfallens duration och det finns inget sätt att, baserat på deras ålder, predicera skillnad i duration mellan George och Lukas, baserat på deras ålder. Vi kan utifrån de data, för alla åldrar, predicera en sjukskrivningsduration om mellan 10 och 50 dagar. Här är det viktigt att komma ihåg att denna prediktion är korrekt. Däremot är det inte, om data ser ut som i den sista figuren, möjligt att utifrån en persons ålder få ett smalare intervall.

**Figur 2.** Tre exempel på hypotetiska spridningsdiagram.

Alla datasets har sina styrkor och tillkortakommanden. Om vi hade haft ytterligare information, såsom om sjukdomens/skadans allvarlighetsgrad, mer ingående information om samsjuklighet, primärvårdskontakter, patientens behandling och rehabilitering (både vad som erbjudits och vad patienten deltar i), patientens livsstil (avseende t.ex. tobaksvanor, alkohol, droger, matvanor, misshandel, etc.), krav i betalt och obetalt arbetet, möjligheter till anpassning av arbetet eller till annan typ av arbete, om sjukskrivande läkares och annan sjukvårdspersonals inställning till patienten och arbete/sjukfrånvaro, *kanske* vi hade kunnat få ännu bättre prediktiva modeller. Men bara kanske. Mekanismerna bakom dessa synnerligen komplexa fenomen är än så länge endast undersökta i begränsad utsträckning (13).

Ett problem med att, som i detta projekt använda MiDAS-data är att endast första sjukskrivningsdiagnosen i sjukskrivningsfallet finns registrerad, dvs det saknas information om huruvida sjukskrivningsdiagnosen ändras under sjukskrivningsfallets gång, respektive om huruvida det finns information i läkarintyget om ytterligare diagnoser som kan ha bidragit till arbetsförmågan. Ett annat problem är att diagnosen endast är angiven på treställighetsnivå. Inom t.ex. F43 finns det fyra olika diagnosgrupper på fyrställighetsnivå, och det är stora skillnader i förväntad sjukskrivningslängd mellan dessa fyra diagnoser (14). Detta är två begränsningar med här använda data. En styrka är att den faktiska sjukskrivningslängden ingår, dvs den sjukskrivning (längd och grad) som patienten använt och som har godkänts av Försäkringskassan.

Det är möjligt att variablernas betydelse i modellerna kan variera med förändrade socialförsäkringsregler, förändrad tillämpning av regler, samt förändring av sjuklighet och krav på arbetsmarknaden. Det vill säga, uppföljning av modellerna behövs.

Inom detta delprojekt, liksom inom det tidigare har det ett flertal gånger diskuterats om vi ska utgå från sjukskrivningsfall i analyserna, eller från personer med nya sjukskrivningsfall. Nackdelen med att utgå från sjukskrivningsfall är att bias kan introduceras i analyserna om det genomgående är personer med vissa egenskaper som har fler än ett fall. Fördelen med att utgå från sjukskrivningsfall, som vi gjort i A2, är framför allt att samtliga fall, dvs inte bara de som inträffar först under en inkluderingsperiod, bidrar med information i analyserna. Eftersom det endast var en liten andel personer som hade fler än ett sjukskrivningsfall i A1 (för F43: 3%) bedömdes det viktigare att ha inkludera fler fall i analyserna (6).

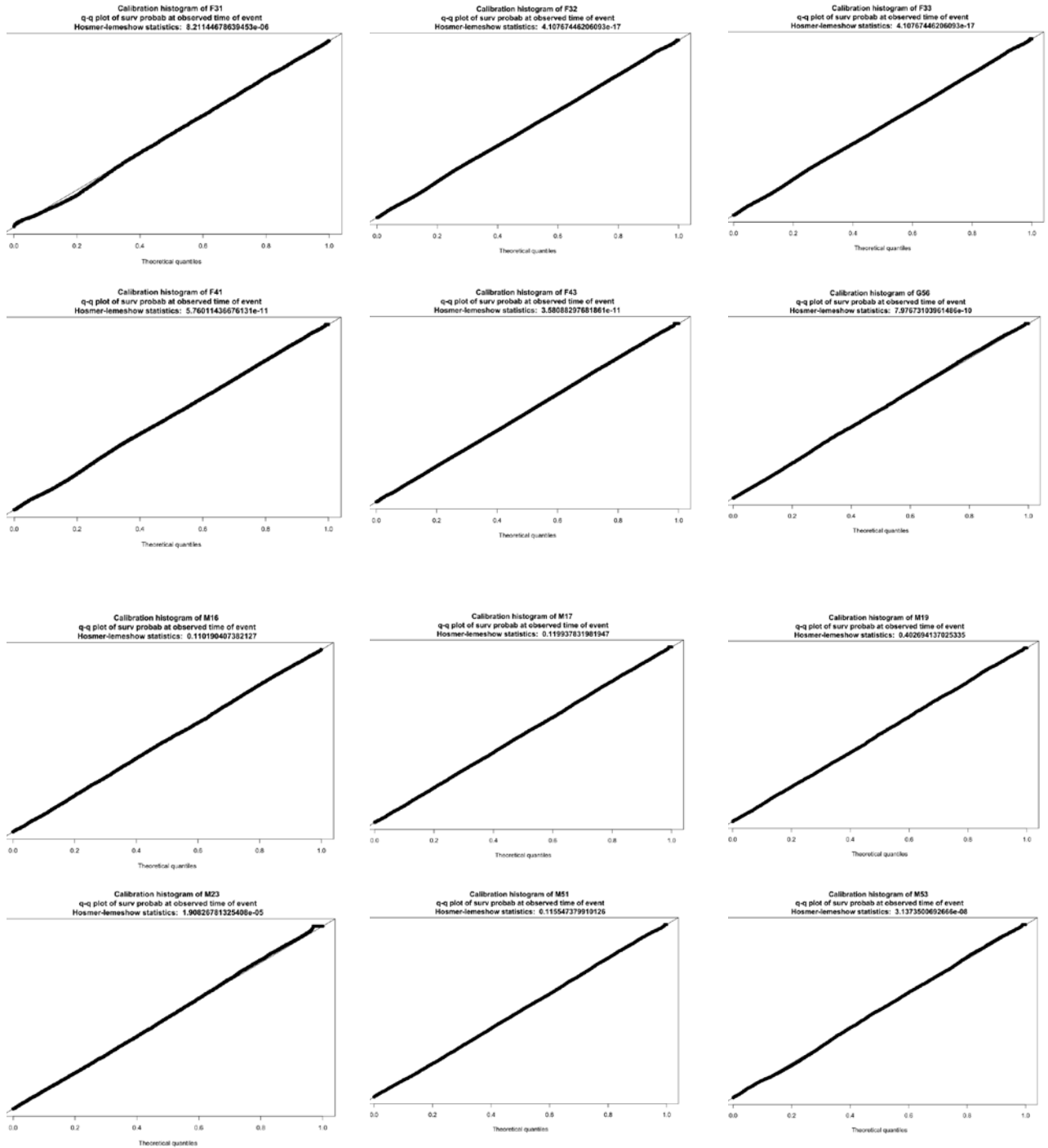


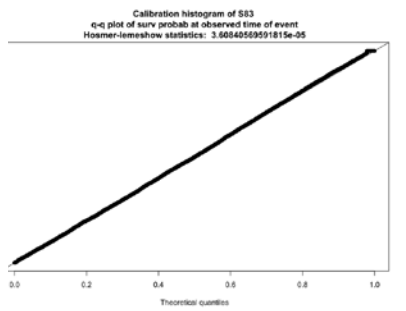
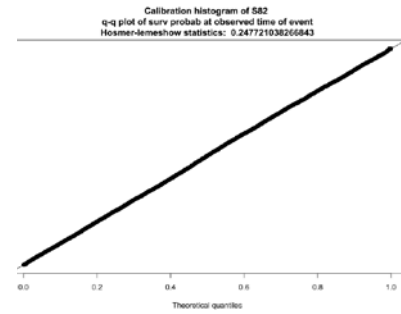
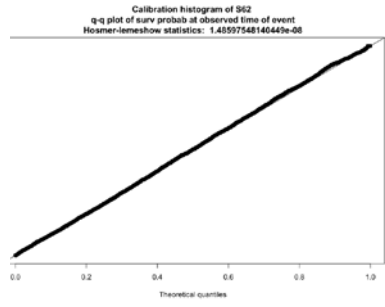
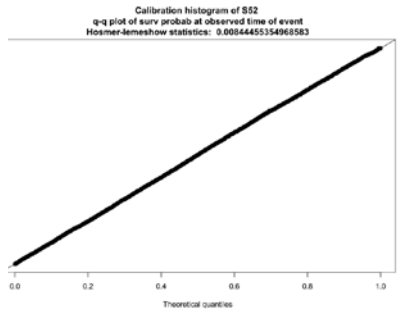
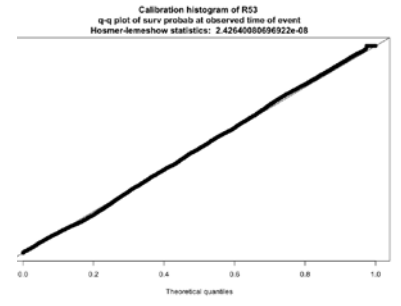
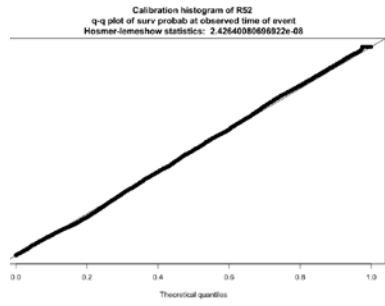
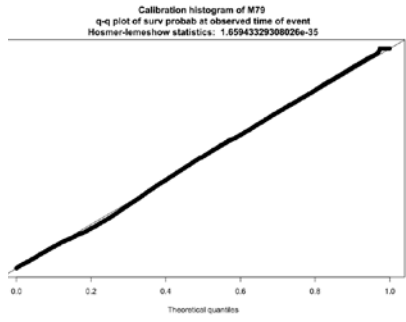
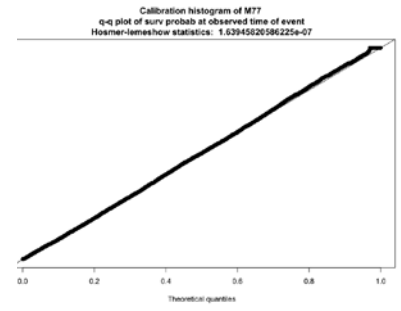
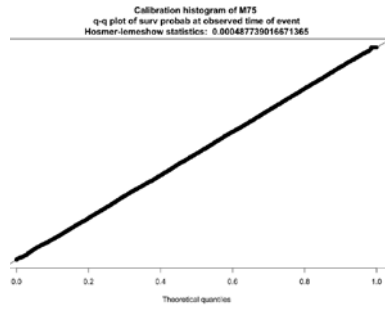
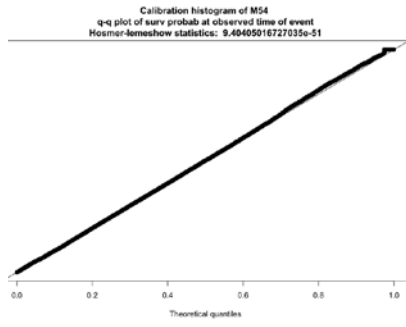
## 8 Referenser, urval

1. Regionala indelningar i Sverige den 1 januari 2003. Del 1. Statistiska centralbyrån; 2003.
2. Karta över H-regionernas omfattning: Statistiska Centralbyrån; 2015  
[[https://www.scb.se/Grupp/Hitta\\_statistik/Regional%20statistik/Kartor/\\_Dokument/H-region\\_farg\\_karta.pdf](https://www.scb.se/Grupp/Hitta_statistik/Regional%20statistik/Kartor/_Dokument/H-region_farg_karta.pdf)].
3. International Statistical Classification of Diseases and Related Health Problems 10th Revision, Version for 2007: WHO; 2010 [<http://www.who.int/classifications/icd/en/>].
4. Alexanderson K, Bottai M, Frumento P, Hinas E, Nilsson K, Kjeldgård L, Tinghög P. Prediktion av fortsatt sjukskrivning respektive av långtidssjukskrivning bland sjukskrivna personer. Försäkringskassan och Sverige Kommuner och Landsting (SKL); 2016.
5. Hinas E, Nilsson K, Tinghög P, Mittendorfer-Rutz E, Alexanderson K. Prediktion av fortsatt sjukfrånvaro bland personer sjukskrivna i depressiv episod respektive i artros (Bilaga 5, Förstudierapport Stöd för rätt sjukskrivning; SRS). Försäkringskassan och Sverige Kommuner och Landsting (SKL); 2015.
6. Alexanderson K, Bottai M, Frumento P, Hinas E, Nilsson K. Prediktionsmodeller för sjukskrivningsduration. Bilaga 5, Stöd för rätt sjukskrivning. Stockholm: Karolinska Institutet/SKL; 2016.
7. Longitudinell Integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier (LISA) 1990 till 2009. SCB; 2011.
8. MiDAS Sjukpenning och rehabiliteringspenning. Försäkringskassan; 2011.
9. Alexanderson K, Marklund S, Mittendorfer-Rutz E, Svedberg P. Studier om kvinnors och mäns sjukfrånvaro. Sektionen för försäkringsmedicin, Karolinska Institutet; 2011.
10. Frumento P. pch: Piecewise Constant Hazards Models for Censored and Truncated Data.: R package version 1.2, url: <http://CRAN.Rproject.org/package=pch>; 2016.
11. Burnham K, Anderson D. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, Second Edition. 2nd ed, 2002, XXVI, 488 p. ed: Springer Verlag; 2002.
12. Löfström R. Varningsinformation Etapp 2. Slutrapport. Sjukvårdsrådgivningen SVR AB. 2008.
13. Sjukskrivning - orsaker, konsekvenser och praxis. En systematisk litteraturoversikt. Stockholm: Statens beredning för medicinsk utvärdering (SBU); 2003.
14. Försäkringsmedicinskt beslutsstöd; diagnosspecifika rekommendationer <http://www.socialstyrelsen.se/riktlinjer/forsakringsmedicinsktbeslutsstod>: Socialstyrelsen; 2017

## 9 Bilaga 1

Figurer över resultaten från Hosmer-Lemeshow test för var och en av 20 diagnoserna.





## 10 Bilaga 2

### Faktorer att ta med för 21 diagnoser

I samtliga modeller läggs följande tre faktorer in, hämtade från systemet:

**Kön** (kvinna, man)

**Ålder**: följande kategorier: 17-29, 30-39, 40-49, 50-56, 57 - år

**Region** – utgår från den region vårdcentralen/vårdenheten finns inom; en av dessa fem: Nord, Väst, Mitt, Stockholm, Syd

(enligt denna indelning, som Försäkringskassan använder:

Nord: Gävleborg, Västernorrland, Jämtland, Västerbotten, Norrbottnen.

Mitt: Södermanland, Uppsala, Västmanland, Östergötland

Stockholm: Stockholm, Gotland

Väst: Västra Götaland, Örebro, Halland, Dalarna, Värmland

Syd: Skåne, Blekinge, Kronoberg, Jönköping, Kalmar)

**För samtliga diagnoser har sysselsättningsstatus lagts högst upp, eftersom läkaren ändå måste fråga om detta. Övriga fem faktorer är listade i ordning av deras betydelse för utfallet.**

D = Default – dvs det värde som automatiskt läggs in om inte läkaren ändrar det.

Diagnos	Faktor	Variabler	Underlag för fråga att ställa till patienten
<b>M75</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Tidigare sjukskrivning	D: 0 - eller korta fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alternativ formulering: Om du tänker tillbaka till juni förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Född i Sv	D: Ja Nej	Är du född i Sverige?

<b>M54</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller korta fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Samsjuklighet	D: 0 eller 1 2 eller fler	Hur många <b>andra</b> långvariga sjukdomar har du?
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
<b>F41</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Partiell sjuk- eller aktivitetsersättning	D: Nej Ja	Har du deltids förtidspension (sjukersättning alt aktivitetsersättning) nu?
<b>F32/33</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös	Vilken är din nuvarande sysselsättning?

		Föräldraledig Studerar	
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Samsjuklighet	D: 0 eller 1 2 eller fler	Hur många <b>andra</b> långvariga sjukdomar har du?
<b>F43</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Född i Sverige	D: Ja Nej	Är du född i Sverige?
<b>M79</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös	Vilken är din nuvarande sysselsättning?

		Föräldraledig Studerar	
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Samsjuklighet	D: 0 eller 1 2 eller fler	Hur många <b>andra</b> långvariga sjukdomar har du?
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Partiell sjuk- eller aktivitetsersättning	D: Nej Ja	Har du deltids förtidspension (sjukersättning alt aktivitetsersättning) nu?
<b>M53</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Partiell sjuk- eller aktivitetsersättning	D: Nej Ja	Har du deltids förtidspension (sjukersättning alt aktivitetsersättning) nu?
	Samsjuklighet	D: 0 eller 1 2 eller fler	Hur många <b>andra</b> långvariga sjukdomar har du?
<b>M17</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig	Vilken är din nuvarande sysselsättning?

		Studerar	
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Tidigare sjukskrivning	D: 0 - eller kortare fall  15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Inlagd sjukhus	D: Nej Ja	Har du varit inlagd på sjukhus mer än en dag de senaste 12 månaderna? (räkna ej pga okomplicerad förlossning)
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
<b>M16</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Tidigare sjukskrivning	D: 0 - eller kortare fall  15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Inlagd sjukhus	D: Nej Ja	Har du varit inlagd på sjukhus mer än en dag de senaste 12 månaderna? (räkna ej pga okomplicerad förlossning)
<b>M19</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?



	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Inlagd sjukhus	D: Nej Ja	Har du varit inlagd på sjukhus mer än en dag de senaste 12 månaderna? (räkna ej pga okomplicerad förlossning)
<b>M77</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Född i Sverige	D: Ja Nej	Är du född i Sverige?
	Familjesituation	D: Ja, sambo/gift Nej	Är du sambo/gift?
<b>M23</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?

	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Läkarbesök, sjukhus senaste 12 månaderna – ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
<b>G56</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Född i Sverige	D: Ja Nej	Är du född i Sverige?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
<b>S82</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Högsta utbildningsnivå	D: Nej	Har du gått på universitet eller högskola?

		Ja	(inklusive ha gått kurser)
	Inlagd sjukhus	D: Nej Ja	Har du varit inlagd på sjukhus mer än en dag de senaste 12 månaderna? (räkna ej pga okomplicerad förlossning)
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Läkarbesök, sjukhus senaste 12 månaderna – (ej primärvård)	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
<b>FS+83</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Familjesituation	D: Ja, sambo/gift Nej	Är du sambo/gift?
	Född i Sverige	D: Ja Nej	Är du född i Sverige?
<b>S62</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)

	Född i Sverige	D: Ja Nej	Är du född i Sverige?
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Inlagd sjukhus	D: Nej Ja	Har du varit inlagd på sjukhus mer än en dag de senaste 12 månaderna? (räkna ej pga okomplicerad förlossning)
<b>S52</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Inlagd sjukhus	D: Nej Ja	Har du varit inlagd på sjukhus mer än en dag de senaste 12 månaderna? (räkna ej pga okomplicerad förlossning)
	Född i Sverige	D: Ja Nej	Är du född i Sverige?
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
<b>M51</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller kortare fall	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till

		15-90 dagar 91-180 181-365	maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Född i Sverige	D: Ja Nej	Är du född i Sverige?
	Samsjuklighet	D: 0 eller 1 2 eller fler	Hur många <b>andra</b> långvariga sjukdomar har du?
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
<b>F31</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker
	Familjesituation	D: Ja, sambo/gift Nej	Är du sambo/gift?
	Partiell sjuk- eller aktivitetsersättning	D: Nej Ja	Har du deltidspension (sjukersättning alt aktivitetsersättning) nu?
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
<b>R52+R53</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs

			specificera samma månad som besöket sker
	Grad av sjukskrivning tidigare i fallet?	D: 100% 75% 50% 25%	Var du först sjukskriven på hel- eller deltid i det här fallet?
	Samsjuklighet	D: 0 eller 1 2 eller fler	Hur många <b>andra</b> långvariga sjukdomar har du?
	Läkarbesök, sjukhus senaste 12 månaderna - ej primärvård	D: Nej Ja	Har du haft mer än ett läkarbesök utanför primärvården de senaste 12 månaderna? (t.ex. på sjukhus)
	Högsta utbildningsnivå	D: Nej Ja	Har du gått på universitet eller högskola? (inklusive ha gått kurser)
<b>I63</b>	Sysselsättningsstatus	D: Yrkesarbetar Arbetslös Föräldraledig Studerar	Vilken är din nuvarande sysselsättning?
	Tidigare sjukskrivning	D: 0 - eller kortare fall 15-90 dagar 91-180 181-365	Ungefär hur många dagar har du varit sjukskriven under de senaste 12 månaderna. (Alt: om du tänker tillbaka till maj förra året, ungefär hur många dagar har du varit sjukskriven sedan dess?) (dvs specificera samma månad som besöket sker)
	Familjesituation	D: Ja, sambo/gift Nej	Är du sambo/gift?
	Född i Sverige	D: Ja Nej	Är du född i Sverige?
	Påbörjades det pågående sjukskrivningsfallet inom primärvården?	D: Ja Nej	Sjukskrevs du från primärvården i början av fallet?
	Inlagd sjukhus	D: Nej Ja	Har du varit inlagd på sjukhus mer än en dag de senaste 12 månaderna? (räkna ej pga okomplicerad förlossning)

