

Rapport Stöd för rätt sjukskrivning

Prediktion av risk för sjukskrivning respektive av långtidssjukskrivning bland sjukskrivna personer

Kristina Alexanderson
Matteo Bottai
Paolo Frumento
Elin Hinas
Kerstin Nilsson
Linnea Kjeldgård
Petter Tinghög

Sektionen för försäkringsmedicin samt Enheten för biostatistik
Karolinska Institutet

Innehåll

1	Inledning.....	3
2	Sammanfattning	4
3	Förkortningar och begrepp	7
4	Introduktion/Bakgrund	8
4.1	Syfte.....	9
5	Metod och material.....	9
5.1	Data.....	9
5.2	Personer med ett nytt sjukskrivningsfall ≥ 21 dagar i depressiv episod (F32) respektive i artros (M15-M19)	21
5.3	Personer med hanlovsfraktur (S62) eller bröstcancer (C50)	22
5.4	Kategorisering av variabler	25
5.5	Dataanalys	27
6	Resultat.....	30
6.1	Prediktionsmodeller för risken att ett sjukskrivningsfall blir längre än förväntat	30
6.2	Risk att blir sjukskriven vid specifika diagnoser	43
6.3	Statistical methods and suggestions for future work	63
7	Slutkommentarer	72
7.1	Metodologiska aspekter	72
7.2	Fortsatta studier	74
8	Tabell- och figurtexter.....	75
9	Referenser.....	78
10	Bilaga	81

1 Inledning

Försäkringskassans prognoser visar att sjukskrivningarna ökar både i omfattning och i längd. Sjukskrivning är en viktig del av vård och behandling, men forskning visar också att långvariga sjukskrivningar kan leda till sämre hälsa, ekonomi och relationer. Idén bakom projektet *Stöd för rätt sjukskrivning (SRS)* är att förbättra för alla inblandade i sjukskrivnings- och rehabiliteringsprocessen, inklusive individen själv. Ett gemensamt IT-baserat bedömningsstöd ska kunna bidra med ett samlat kunskapsunderlag för att läkare ska kunna utfärda läkarintyg med större träffsäkerhet samt hjälpa hälso- och sjukvården och Försäkringskassan att identifiera individer som har behov av specifika eller samordnade insatser. Bedömningsstödet kan också i en framtid användas som ett stöd för arbetsgivare och av individen själv för att kunna ta en aktiv del i sin egen rehabilitering.

Under förstudiearbetet har projektet utrett förutsättningarna för att skapa ett sådant samlat bedömningsstöd för olika aktörer i sjukskrivnings- och rehabiliteringsprocessen. Huvudsyftet har varit att undersöka möjligheten att bättre utnyttja kunskap om prognosfaktorer som påverkar sjukskrivningen. Utifrån denna kunskap kan mer träffsäkra prediktioner ges om sjukskrivningslängd och omfattning för en enskild individ.

Projektet har finansierats via Socialdepartementet genom överenskommelsen mellan regeringen och SKL om en kvalitetssäker och effektiv sjukskrivningsprocess, den så kallade sjukskrivningsmiljarden. Försäkringskassan och Sveriges Kommuner och Landsting har gemensamt ansvarat för förstudien. Socialstyrelsen och andra intressenter är representerade i projektets referensgrupper. Projektets utredningsarbete påbörjades i mars 2014.

Förstudiearbetet har varit indelat i följande delområden:

- Målarbetet har tagit fram förslag på övergripande gemensamma nationella mål för sjukskrivningsområdet och effektmål för projektet.
- Kunskapsanalysen har sammanställt vetenskaplig och annan relevant och aktuell kunskap om sjukskrivning, prognosfaktorer och insatser.
- Omvärldsanalysen har inventerat liknande bedömningsstöd i världen för att se om det finns relevant kunskap att dra lärdom av.
- Konzeptutredningen har utrett verksamhetsmässiga, tekniska och juridiska förutsättningar för att utveckla bedömningsstödet.

Utöver delområdena har projektövergripande arbete som styrning, ledning, uppföljning, kommunikation och kvalitetssäkring bedrivits.

Utredningsarbetet har utförts av en arbetsgrupp med bred kompetens och med stor samlad erfarenhet av nationella e-hälsoprojekt. Delar av utredningsarbetet har utförts av och med forskare och utredare från Karolinska Institutet, Lunds universitet och Linköpings universitet.

Denna rapport redovisar en del av förstudiens arbete. Förstudierapporten med en sammanfattning av hela resultatet, och alla delrapporter, finns på SKL:s webbplats:

<http://skl.se/halsasjukvard/sjukskrivningochrehabilitering/sjukskrivningsmiljarden/rattsjukskrivningstod.5229.html>

2 Sammanfattning

Syften med detta delprojekt inom SRS projektet var att

- (I) att fördjupa och validera tidigare genomförda preliminära analyser av prediktiva modeller för sjukskrivningslängd bland sjukskrivna personer
- (II) utveckla nya prediktiva modeller för risken att bli sjukskriven bland personer med en viss skada eller sjukdom samt
- (III) att ta fram en plan för fortsatt, fördjupad modellering av sannolikheten för sjukskrivning och sjuk- och aktivitetsersättning.

Flera sådana preliminära analyser har genomförts. I dessa har vi utgått från samtliga de personer i åldrarna 16-64 år som bodde i Sverige i december år 2009 respektive i december år 2010. Såväl retrospektiva som prospektiva data för dessa personer, hämtade från flera rikstäckande register administrerade av Försäkringskassan, Socialstyrelsen samt Statistiska centralbyrån, har ingått i olika typer av epidemiologiska och statistiska analyser. *Resultaten av analyserna visar på goda möjligheter att skapa bra prediktionsmodeller.* Nedan sammanfattas resultaten för vart och ett av de tre syftena.

I. Prediktionsmodeller för risken att bli långtidssjukskriven har utvecklats och validerats för två diagnosgrupper, nämligen för depressiv episod (ICD10-kod: F32) samt för artros (M15-M19). Dessa analyser baserades på information om samtliga personer som år 2010 hade ett nytt sjukskrivningsfall som blev minst 21 dagar långt i depressiv episod (n=23 079) respektive i artros (n=10 851). Utfallet i modellerna var att sjukskrivningsfallet blev längre än förväntat; 180 dagar för depressiv episod (detta gällde 33 % av alla) och 120 dagar för artros (38 % av alla). Dessa tidsgränser baserades på rekommendationerna för längsta sjukskrivningstid, givna i Socialstyrelsens diagnosspecifika försäkringsmedicinska beslutsstöd. Ett stort antal faktorer inkluderades i analyserna, avseende sociodemografi, sjukskrivningshistorik samt faktorer relaterade till hälso- och sjukvård och behandling. Efter val av optimal brytpunkt, som skiljer på dem som har hög respektive låg risk för utfallet, var den maximala andelen korrekt predicerade 72 procent för depressiv episod och 67 procent för artros.

Modellerna validerades sedan på liknande kohorter från år 2011 och befanns ha god validitet även för detta år.

Olika exklusions- och inklusionskriterier prövades och det framkom att använda kriterier hade stor inverkan på modellernas prediktiva förmåga. Fördjupade analyser behövs, till exempel för att få mer kunskap om vad som är av central betydelse för utfallet i högriskgrupper, det vill säga, om vilka i dessa grupper som inte får lång sjukfrånvaro.

Analys av olika subgrupper av prediktionsfaktorer visade att inklusion av faktorer relaterade till sjuklighet, hälso- och sjukvård och behandling förbättrar modellernas förmåga att diskriminera mellan dem som blev långtidssjukskrivna och dem som inte blev det. Speciellt gällde detta modellen för personer sjukskrivna i artros.

II. För personer som fått en handlovsfraktur (S62) alternativt en bröstcancerdiagnos (C50) har dels deskriptiva analyser av deras sjukskrivning och sjuk- och aktivitetsersättning genomförts

och dels har modeller utvecklats för att predicera dessa individers risk att bli sjukskriven vid diagnosen.

Av de 16 337 personer som år 2010 fick en *handlovsfraktur* hade en tredjedel (n=5 322) minst ett sjukskrivningsfall som översteg 14 dagar under det första året efter frakturen. Efter exklusion av dem som redan hade ett pågående sjukskrivningsfall eller hade sjuk- eller aktivitetsersättning om minst 75 procent vid frakturen, återstod 14 465 personer. En modell som predicerar risken att i samband med handlovsfrakturen bli sjukskriven i diagnos handlovsfraktur, utvecklades baserat på dessa personer, 20 procent fick utfallet. Modellens förmåga att predicera korrekt utfall blev 80 procent, givet en viss, optimal brytpunkt. AUC för prediktionsmodellen var 0,77.

I analyser av kvinnor med *bröstcancer* inkluderades de 3 536 kvinnor som år 2010 fick en sådan diagnos för första gången. Vid diagnosen var tre procent sjukskrivna sedan tidigare och 12 procent hade sjuk- eller aktivitetsersättning med en omfattning om minst 75 procent. Andelen kvinnor med sjukfrånvaro ökade kraftigt under året efter diagnosdatum, för att sedan åter avta mycket under de kommande två åren. I analyser av risk för att få ett nytt sjukskrivningsfall längre än 14 dagar i en cancerrelaterad eller psykisk diagnos inom en månad efter bröstcancerdiagnosen, fann vi att hälften av kvinnorna fick en sådan sjukskrivning. I de analyserna exkluderades de som inte var 'at risk' för detta (det vill säga, de som redan var sjukskrivna eller som hade sjuk- eller aktivitetsersättning på 75-100%). En procent av dem som var 'at risk' fick ett nytt sjukskrivningsfall i en annan diagnos inom en månad. I den preliminära prediktiva modellen av sannolikheten att få ett nytt sjukskrivningsfall uppnåddes en prediktionsförmåga på 67 procent rätt klassificerade personer; AUC för prediktionsmodellen var 0,73.

Dessa prediktionsmodeller vad gäller risken att bli sjukskriven i samband med de här två diagnoserna är ett första steg i att utveckla sådana modeller. Fler faktorer behöver inkluderas i analyserna och kategorisering av dem behöver undersökas, utfallet behöver justeras, både vad gäller tid till utfallet, utfallets duration och omfattning, det vill säga, hel- eller deltid.

III. En projektplan för att utveckla prediktiva modeller presenteras. Den baseras på mycket omfattande populationsbaserade registerdata om sjuklighet, vård, sociodemografi, socioekonomi, livssituation, sjukskrivningshistorik, etcetera. Registren är av god kvalitet och genereras av fyra olika myndigheter. De statistiska metoder som krävs för att utveckla prediktionsmodeller skiljer sig från de analyser som vanligen används i studier om associationer mellan möjliga riskfaktorer och utfallet. Här är syftet att, med så god träffsäkerhet som möjligt, predicera risken att bli sjukskriven alternativt få sjuk- eller aktivitetsersättning samt att predicera längden på ett pågående sjukskrivningsfall; det vill säga, minst tre olika utfall. Detta innebär att modellerna bör valideras, både genom så kallad intern och extern validering, samt att modellerna bör uppdateras och rekaliseras när senare data blir tillgängliga. Modeller för att predicera de tre ovan nämnda utfallen kommer att utvecklas separat, men kommer sedan att kunna kombineras när modellerna skall implementeras.

I projektplanen ingår att i analyserna även ta hänsyn till olika aspekter av tid. En sådan aspekt är att vissa variabler kan förändras över tid (samsjuklighet, familjesituation, sysselsättning, behandling, etcetera). En annan aspekt är att olika lång tid kan ha passerat sedan diagnosen/skadan alternativt sedan starten på sjukskrivningsfallet – vilket innebär att prognosen kan variera med detta. Det är, till exempel, skillnad på om läkaren träffar patienten dag 1, dag 21 eller dag 163 i ett sjukskrivningsfall. En tredje är att olika prediktorer kan ha olika stor betydelse över tid.

Projektgruppens tvärvetenskapliga expertis när det gäller denna typ av studier kommer att vara av central betydelse för att selektera och utvärdera möjliga variablers inkludering i de prediktiva modellerna liksom för lämplig kategorisering av variablerna, samt vid val av eventuella interaktioner.

Sammanfattningsvis finns det mycket goda möjligheter att, baserat på dessa data, med god precision predicera vilka personer med de studerade diagnoserna, som har hög risk för sjukskrivning eller för sjuk- eller aktivitetsersättning respektive har hög risk för att deras sjukskrivningsfall ska blir längre än förväntat – förutsatt att modellerna vidareutvecklas. Användbara prediktionsmodeller kan tas fram även för andra diagnoser. Fördelar med projektet är att länkade longitudinella individdata från flera olika rikstäckande register, med information om sjukfrånvaro, sjuk- och aktivitetsersättning, sjuklighet, vård, sociodemografi, socioekonomi, samt död är tillgängliga. Andra fördelar är att projektet är tvärvetenskapligt, att deltagarna har genuin kunskap om olika aspekter av området och att det innebär nära samarbete mellan forskare, kliniker och myndigheter.

En stor andel av personer som får en diagnos blir inte sjukskrivna i den – mer kunskap om vilka som blir sjukskrivna behövs, gällande olika sjukdomar och skador. Det finns en stor potential i att framgångsrikt fördjupa de här genomförda analyserna och framförallt i att utveckla dem enligt de föreslagna metoderna.

3 Förkortningar och begrepp

I rapporten används följande förkortningar och begrepp.

AUC	Area Under Curve, arean under ROC-kurvan
Faktor	Ett värde på en variabel, till exempel kvinna i variabeln kön
FK	Försäkringskassan
H-region	Indelning av kommuner, efter hur urbana de är, såsom i Storstad, Mellanstor stad eller Landsbygd
ICD-10	International Classification of Diagnoses, version 10 ⁽¹⁾
IMAS	Insurance Medicine All-Sweden
KI	Konfidensintervall
LCGM	Latent Class Growth Modelling
LISA	Longitudinell Integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier; rikstäckande register administrerat av Statistiska centralbyrån
Långtidssjukskrivning	I vissa analyser här definierat som ett sjukskrivningsfall som varar i minst 120 dagar vid sjukskrivning i artros eller minst 180 dagar vid sjukskrivning i depressiv episod
MiDAS	MikroData för Analys av Socialförsäkringen; rikstäckande register administrerat av Försäkringskassan
Musk	Muskuloskeletala diagnoser
Odds	Risken att en händelse sker dividerat med risken att händelsen inte sker
OR	Oddsquot (oddsrat), oddset för en händelse i en grupp dividerat med oddset för samma händelse i en annan (referens) grupp
PAR	Patientregistret; rikstäckande register administrerat av Socialstyrelsen
Psyk	Psykiska diagnoser
ROC-kurva	Plottar sensitiviteten mot 1-specificiteten
SA	Sjuk- och aktivitetsersättning
Sensitivitet	Andel sant positiva prediktioner
Sjukfrånvaro	Orden sjukfrånvaro och sjukskrivning används synonymt
Specificitet	Andel sant negativa prediktioner
SRS	Stöd för rätt sjukskrivning; namn på det projekt som denna rapport ingår i
Variabel	En grupp faktorer som utesluter varandra, till exempel kön eller ålder

4 Introduktion/Bakgrund

Detta är en andra delrapport i ett uppdrag som Sektionen för försäkringsmedicin, Karolinska Institutet fått från projektet Stöd för rätt sjukskrivning (SRS) och som ingår i den del av SRS-projektet som syftar till att ta fram prediktionsmodeller för risk för sjukskrivning respektive risk för att en sjukskrivning ska bli lång. Det tvärvetenskapliga projektet har genomförts i samarbete med forskare vid Sektionen för biostatistik, Karolinska Institutet.

Framförallt har följande sju personer vid Karolinska Institutet arbetat med projektet:

- Kristina Alexanderson, professor i socialförsäkring, PhD i medicinsk vetenskap, Sektionen 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)
- Paolo Frumento, forskarassistent, PhD in statistics, Enheten för biostatistik (IMM)
- Elin Hinas, statistiker, Sektionen för försäkringsmedicin (CNS)
- Kerstin Nilsson, statistiker, Sektionen för försäkringsmedicin (CNS)
- Linnea Kjeldgård, statistiker, Sektionen för försäkringsmedicin (CNS)
- Petter Tinghög, forskarassistent, fil dr, Sektionen för försäkringsmedicin (CNS)

Uppdraget har inneburit att, under mycket begränsad tid, arbeta med följande tre huvudområden:

- I. Validera och vidareutveckla redan de framtagna prediktiva modellerna för sjukskrivningslängd i två olika diagnoser⁽²⁾. Vi ombads även utvärdera gjorda exklusions- och inklusionskriterier samt göra motsvarande analyser med en mer begränsad uppsättning data – nämligen de data som Försäkringskassan har tillgång till.
- II. Utredda möjligheter att modellera sannolikheten att bli sjukskriven vid sjukdom eller skada. Dels analysera andel personer som redan var sjukskrivna eller hade sjuk- eller aktivitetsersättning vid diagnos, dels analysera andelen som blev detta.
- III. Föreslå en plan för att modellera sannolikheten för sjukskrivning respektive för att få sjuk- eller aktivitetsersättning vid olika tillstånd generellt samt för de sjukskrivningsdiagnoser som är vanligast inom primärvården.

Samtliga delar av analyserna har skett i dialog med beställaren. Sådana diskussioner har skett kontinuerligt, vanligen en eller flera gånger varje vecka, vad avser till exempel fokus, typ av analyser, utfallsmått, faktorer och kategorisering av dem.

Vissa delar av resultatdelen har skrivits på engelska.

4.1 Syfte

Projektets tre syften var (I) att fördjupa tidigare genomförda preliminära analyser av prediktiva modeller för sjukskrivningslängd bland sjukskrivna personer⁽²⁾ samt (II) att utveckla nya prediktiva modeller avseende risken att bli sjukskriven bland personer med en sjukdom eller en skada. Ett tredje (III) syfte var att ta fram en plan för fortsatt och fördjupad modellering av sannolikheten för sjukskrivning, som underlag för modeller för prototyper som kan användas i pilotdrift.

5 Metod och material

Flera olika typer av preliminära prospektiva kohortstudier har genomförts, enligt ovanstående syfte, för att få underlag till fortsatta analyser. De analyser, vars resultat presenteras i denna rapport, kan delas in i tre olika delar; en första del som är en uppföljning av analyser gjorda i en tidigare rapport⁽²⁾ och en andra del där personer i två olika diagnosgrupper studeras. I en tredje del ges förslag till fortsatta analyser.

I. I den första delen utvecklas och valideras tidigare framtagna prediktionsmodeller för långtidssjukskrivning för två grupper av sjukskrivna personer. En av grupperna utgörs av de personer som haft ett nytt sjukskrivningsfall som under år 2010 respektive under år 2011 varat i minst 21 dagar med en depressiv episod (ICD 10-kod⁽¹⁾: F32), den andra gruppen utgörs av dem som haft ett nytt sådant sjukskrivningsfall med en artrosdiagnos (M15-M19) som huvuddiagnos. Två olika kohorter av personer som uppfyller dessa kriterier har skapats för varje diagnostyp, en baserad på år 2010 och som användes för att utveckla modellerna och en baserad på år 2011 som användes för att validera modellerna.

II. I den andra delen studeras två andra grupper, nämligen de personer som under 2010 fick en handlovsfraktur och fick vård för denna inom specialiserad öppen- eller slutenvård i Sverige samt de kvinnor som under 2010 för första gången diagnosticerades med bröstcancer. För båda dessa grupper har dels deskriptiva analyser av deras sjukfrånvaro och sjuk- eller aktivitetsersättning utförts och dels prediktionsmodeller för risken att bli sjukskriven i samband med respektive diagnos.

III. Den tredje delen utgår från resultaten i del I och II samt avancerad epidemiologisk och statistik metod.

5.1 Data

Studierna baseras på information från den databas som kallas Insurance Medicine All-Sweden (IMAS) och som administreras av Sektionen för försäkringsmedicin, Karolinska Institutet. Databasen innehåller avidentifierade individdata länkade från elva olika rikstäckande register för personer boende i Sverige och har godkänts av den Regionala etikprövningsnämnden i Stockholm för denna typ av analyser. För detta projekt har data från sju av dessa register använts enligt följande:

Från Försäkringskassan

- MiDAS registret⁽³⁾: information om sjukskrivningsfall samt sjuk- eller aktivitetsersättningar som ersatts av Försäkringskassan under åren 2008-2013, för sjuk- eller aktivitetsersättningar inkluderades fall så långt tillbaka som registret täcker, det vill säga 1994, i vissa analyser. Ur registret hämtades information om varje falls startdatum (startdatum 1) och slutdatum, omfattning (hel- eller deltid), försäkradtyp (typ av sysselsättning) vid sjukskrivningsfallets start samt sjukskrivningsdiagnos (huvuddiagnosen i den första sjukskrivningsperioden i sjukskrivningsfallet). Detta innebär att information från både hel- och delfallsfiler i MiDAS har använts. (När data från MiDAS ”tvättades” slogs alla sjukskrivningsfall som startade inom fem dagar efter att ett tidigare fall avslutats samman med det tidigare fallet. Självklart räknades inte de mellanliggande dagarna med i beräkning av antal nettodagar.)
- Försäkringskassans register över personer som under åren 2009-2011 uppnått maximal tid i sjukförsäkringen och datum för detta.

Från Statistiska centralbyrån: information från LISA-registret⁽⁴⁾ om följande sociodemografiska variabler: ålder, kön, födelse-land, högsta utbildningsnivå, yrke, sektor, familjesituation, civilstånd, boendelän och H-region.

Från Socialstyrelsen information från:

- Dödsorsaksregistret om datum för död och dödsorsak 2010-2013.
- Patientregistret, dels Slutenvårdsregistret, vad avser tillfällen personer varit inlagda på sjukhus, avseende datum, antal dagar, samt huvud- och bidiagnoser under åren 2008-2011, dels Öppenvårdsregistret, vad avser datum för läkarbesök i specialiserad öppenvård för olika diagnoser under åren 2008-2011. (Detta register innehåller inte information om besök i primärvården).
- Cancerregistret, avseende diagnos, datum för diagnos, tumörtyp (malign/benign) och TNM-klassificeringsvariabler (vilka anger tumörutbredning vid diagnostillfället).
- Läkemedelsregistret om uthämtade receptbelagda läkemedel, vad avser datum och typ av läkemedel under 2008-2011.

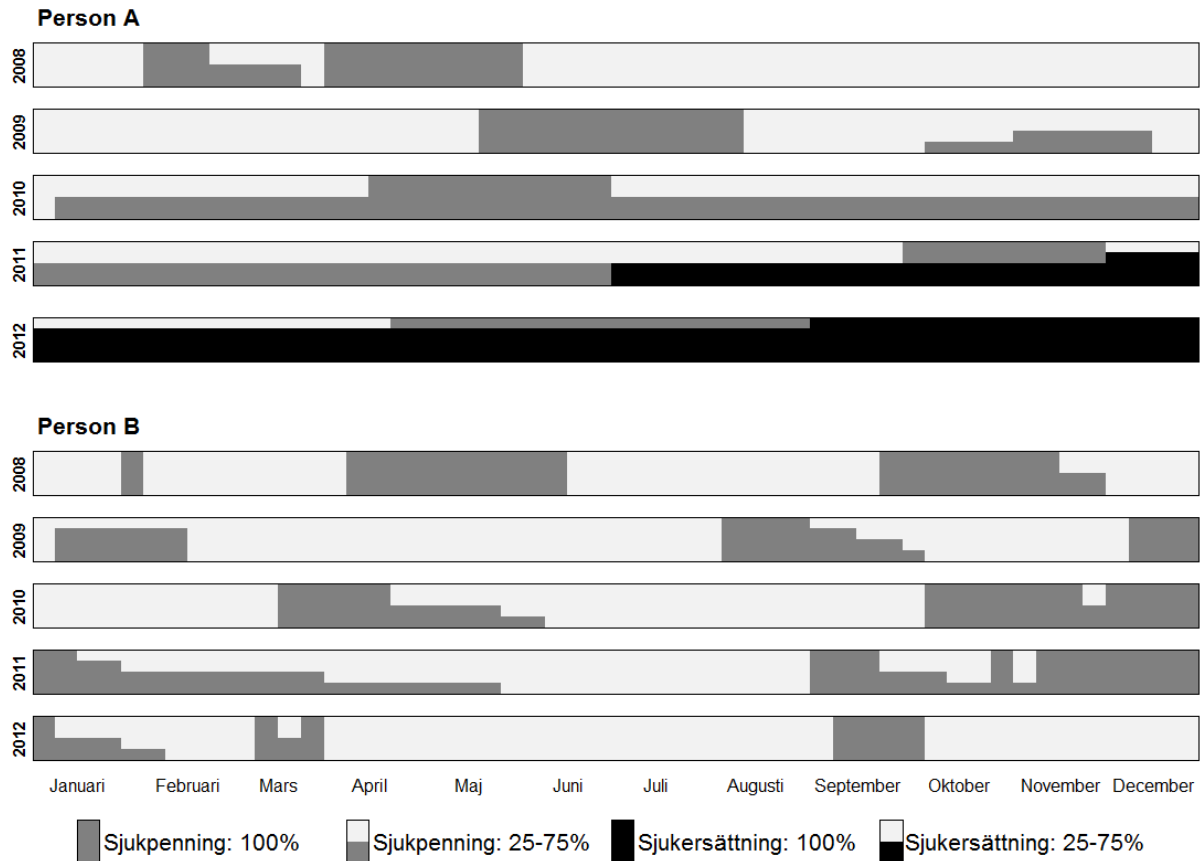
5.1.1 Om sjukskrivningsfall respektive personer med sjukskrivningsfall

Analyser av sjukfrånvarodata innebär många olika statistiska utmaningar^(5, 6) och några aspekter som behöver beaktas nämns nedan.

5.1.1.1 Komplexa processer

De komplexa processer som skall hanteras i dessa prediktionsmodeller kan illustreras på olika sätt. I Figur 1 ges exempel på hur två personers sjukfrånvaro och sjukersättning kan se ut över längre tid än vad som ges i typexempel av olika sjukskrivningsfall, som de beskrivs i Figur 3 och Figur 4. Person A har 2008 och 2009 fyra olika sjukskrivningsfall, ersatta av Försäkringskassan. Huvudelen av dessa perioder är på heltid men ett fall är först på 25 procent för att sedan blir på halvtid. Två år sedan har hen beviljats sjukersättning på halvtid, efter ett och ett halvt års sjukskrivning. Sjukersättningen övergår sedan till 75 procent för att

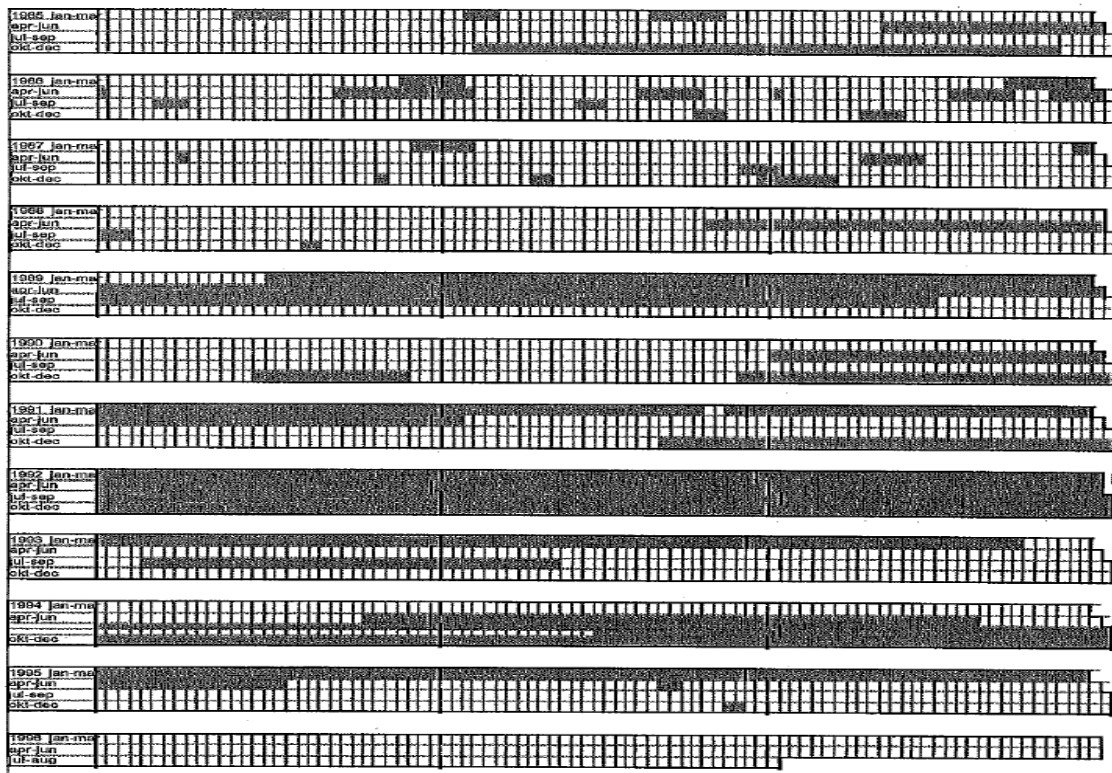
mot slutet av perioden vara på heltid. Person B har istället ett flertal sjukskrivningsfall, som alla börjar på heltid och där många avslutas på deltid.



Figur 1. Exempel på två fiktiva personers sjukskrivningsfall och sjukersättning, avseende duration och grad, under loppet av en femårsperiod.

Exemplet i Figur 2 har baserats på information från en persons faktiska sjukfrånvaro under drygt elva år, uppdelat per kvartal, från en period när även korta sjukskrivningsfall ersattes direkt av försäkringskassan. Just den personen hade ingen sjuk- eller aktivitetsersättning under dessa år.

Den finns även liknande komplexitet gällande andra faktorer, som kan variera över tid och påverka det studerade. Detta gäller till exempel sjuklighet, samsjuklighet, vårdkontakter, behandling, familjesituation, föräldradighet, arbetslöshet, typ av yrke och arbetsgivare. I de initiala analyser som presenterats i denna rapport har endast information om värden på sådana variabler vid baslinjen inkluderats, inte avseende förändring av dem över tid.

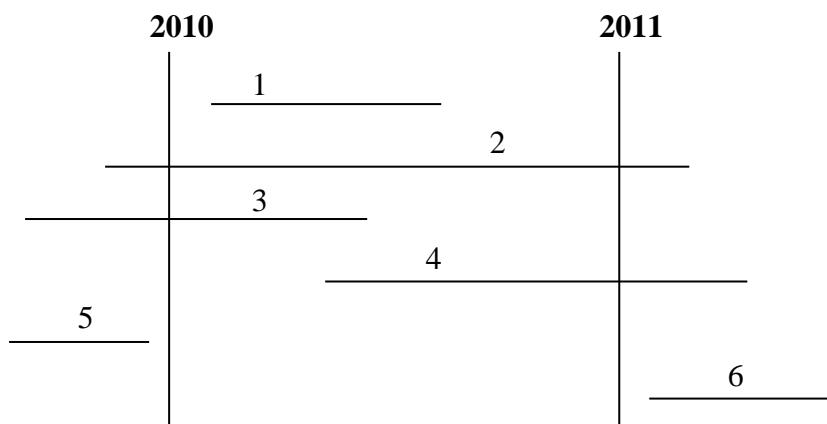


Figur 2. Diagram över en persons sjukfrånvaro år 1985 - augusti 1996

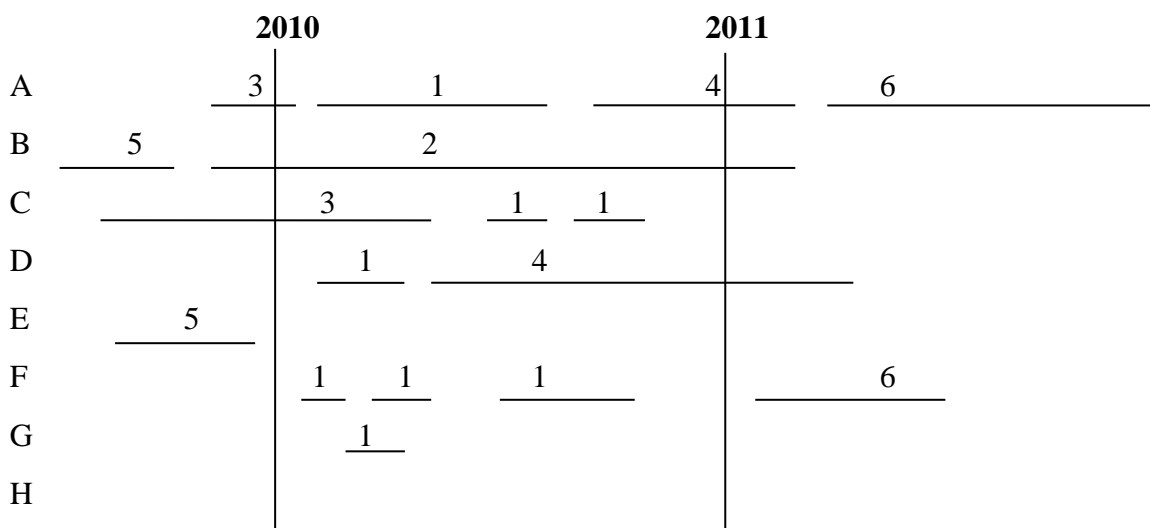
5.1.1.2 Sjukskrivningsfall respektive personer med sjukskrivningsfall under en viss period

Olika aspekter behöver beaktas när sjukfrånvaro under en viss period, till exempel ett visst år studeras⁽⁶⁾. Detta kan illustreras som i nedanstående Figur 3, där olika typer av sjukskrivningsfall representeras av olika siffror (1-6). En central sådan aspekt är att flera sjukskrivningsfall som pågår under ett år, här 2010, kan ha påbörjats under föregående år (sjukskrivningsfall 3 i Figur 3) och ibland till och med flera år innan. En annan aspekt är att sjukfrånvarofall som påbörjats under ett år, 2010, kan fortsätta även nästkommande år (sjukskrivningsfall 4 i Figur 3). Ytterligare en situation är att ett sjukskrivningsfall kan pågå under hela året, 2010 (sjukskrivningsfall 2 i Figur 3). Det vanligast förekommande är att ett sjukskrivningsfall både påbörjas och avslutas under samma år, 2010 (sjukskrivningsfall 1 i Figur 3). En person kan ha flera sjukskrivningsfall av denna typ eller kombinationer av flera olika typer (Figur 4).

Varken sjukskrivningsfall nummer 5 eller 6 (i Figur 3), eller sjukskrivningsdagar under dessa sjukskrivningsfall, bör tas med när sjukfrånvaron för 2010 skall redovisas. Däremot kan tidigare sjukfrånvaro, till exempel sjukskrivningsfall nummer 5, påverka risken för sjukfrånvaro under 2010. Dessa aspekter kan hanteras på olika sätt och ger olika mått på sjukfrånvaro⁽⁵⁻¹⁰⁾.



Figur 3. Illustration av sjukfrånvarofall i relation till start- och slutdatum och årsgränser.



Figur 4. Illustration av sjukfrånvarofall i relation till start- och slutdatum och årsgränser för olika personer; A till H, där personen H inte har något sjukskrivningsfall alls under denna period och till exempel person E har ett sjukskrivningsfall som påbörjas och avslutas under föregående år, här 2009.

För att få kunskap om vilka personer som har hög eller låg risk för att sjukskrivningsfallet ska bli långt eller längre än förväntat görs i den här rapporten analyser där nya sjukskrivningsfall som blivit tre veckor långa under 2010 i en viss diagnos följs framåt. I en tidigare rapport inom ramen för SRS-projektet studerade vi även förloppsmönster av nettodagar i sjukfrånvaro samt sjuk- eller aktivitetsersättning i såväl det initiala fallet som i nya fall under det kommande året och särskilt de som hade högt antal dagar ⁽²⁾. I denna rapport fördjupas inte den typen av analyser, baserade på så kallade Group-Based Trajectories ^(11, 12).

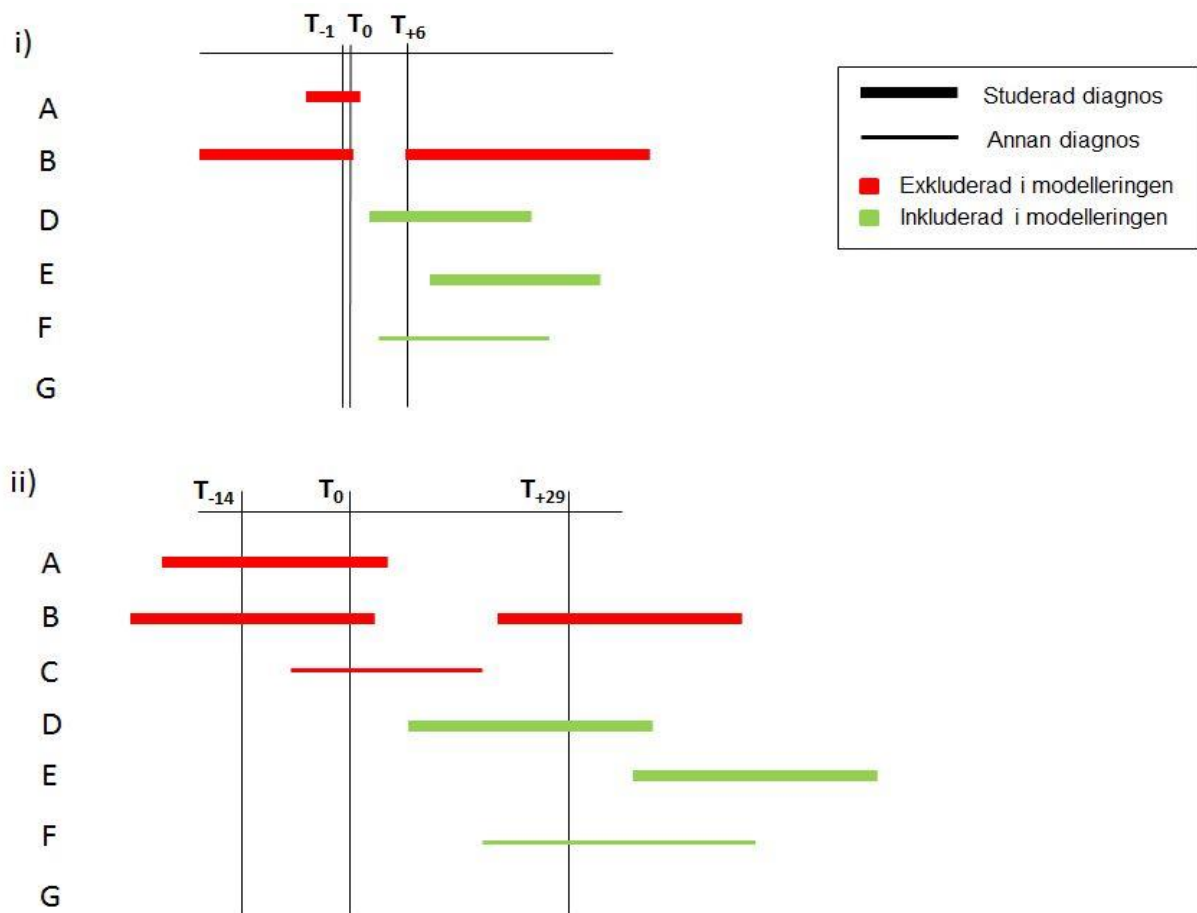
I denna rapport har vi även studerat risken för att personer med en viss diagnos överhuvudtaget blir sjukskriven med denna diagnos. I de analyserna kommer inte årsgränser att vara utgångspunkten, som i ovanstående figurer. Istället är utgångspunkten det datum när sjukskrivningsfallet började alternativt datum för när personen fick en diagnos – dessa datum kallas här T_0 .

I analyserna av risk för att ett sjukskrivningsfall skall bli längre än förväntat har T_0 satts till 'startdatum 1', enligt MiDAS-registrets terminologi, och endast fall som blivit minst 21 dagar långa har inkluderats. I analyserna inkluderades information om många faktorer, även eventuell tidigare sjukskrivning eller sjuk- och aktivitetsersättning. Endast sjukskrivningsfall som blivit längre än 14 dagar har då inkluderats i analyserna, för att inte introducera bias för de personer som har eller kan ha ersättning från Försäkringskassan i kortare sjukskrivningsfall. Detta är särskilt viktigt då personer i de grupperna kan ha ökad risk för längre sjukskrivning än andra personer. Det innebär att särskilda analyser för dessa personer kan behövas, för att få bättre kunskap om olika riskgrupper.

När det gäller risk för att bli sjukskriven i samband med en diagnos – i denna rapport handlovsfraktur respektive bröstcancer – kallas diagnosdatum för T_0 . Det finns dock alltid en viss osäkerhet gällande hur exakt detta datum är. En principiell bild av hur detta hanteras illustreras i Figur 5, dels för handlovsfrakturer (i), dels för bröstcancer (ii).

För personer med *handlovsfrakturer* studeras risken att få ett nytt sjukskrivningsfall i denna diagnos (S62) under perioden från dagen innan handlovsfrakturen till och med sjätte dagen efter. Personer med ett pågående sjukskrivningsfall (i detta fall definierat som att det påbörjades minst två dagar innan, och fortfarande pågick vid T_0) exkluderades från modelleringen (personer A och B i Figur 5). Personerna D-G inkluderas i modelleringen, där utfallet att få ett nytt sjukskrivningsfall i diagnosen S62 förekommer för person D, men inte för de övriga personerna.

För kvinnor med *bröstcancer* studeras risken att få ett nytt sjukskrivningsfall i cancer, cancerrelaterade eller psykiska diagnoser under perioden från dag 14 innan (här kallad T_{-14}) till och med dag 29 (här kallad T_{29}) efter datum för diagnos (det vill säga, T_0). Kvinnor som har pågående sjukskrivningsfall, här definierat som fall som startade mer än 14 dagar innan T_0 och som fortfarande pågick vid T_0 (personer A och B) samt fall som startade mindre än 14 dagar innan T_0 och pågick vid T_0 men i andra sjukskrivningsdiagnoser än de studerade, det vill säga i andra diagnoser än cancer, cancerrelaterade eller psykiska besvär (person C), exkluderades från modelleringen. Personerna D-G inkluderas alla i modelleringen, utfallet att få ett nytt sjukskrivningsfall i de studerade diagnoserna förekommer för person D men inte för de övriga.



Figur 5 Illustration av exempel av sjukskrivningsfall i relation till start- och slutdatum och den period relativt T_0 som studeras i modelleringen av sannolikheten att bli sjukskriven i samband med diagnos, för olika personer med i) handlovsfraktur respektive med ii) bröstcancer. T_0 = dagen för diagnosen.

5.1.2 Socialstyrelsens försäkringsmedicinska beslutsstöd

När det gäller om ett sjukskrivningsfall ska anses som långt eller som längre än förväntat har vi här valt att utgå från rekommendationerna i Socialstyrelsens försäkringsmedicinska beslutsstöd, den diagnosspecifika delen⁽¹³⁾. Om ett sjukskrivningsfall blivit längre än rekommendationerna där, har vi definierat fallet som långt.

Socialstyrelsens försäkringsmedicinska beslutsstöd består av två delar, dels en övergripande del om vad som generellt ska beaktas i samband med sjukskrivning⁽¹⁴⁾, dels en diagnosspecifik del⁽¹³⁾. I den senare ges övergripande rekommendationer om sjukskrivningstid och -grad för diagnoser som är vanliga sjukskrivningsdiagnoser. I de analyser vi tidigare genomfört inom ramen för projektet har bland annat diagnoserna depressiv episod, (ICD-10 kod¹: F32) och artros (M15-M19) studerats⁽²⁾. I denna rapport

¹ ICD 10; the International Classification of Diseases and Related Health Problems, version tio – det system för klassificering av diagnoser som Världshälsoorganisationen WHO antagit.

redovisas resultat av ytterligare, fördjupade analyser vad gäller dessa två diagnoser samt analyser för diagnoserna handlovsfraktur (S62) och bröstcancer (C50). I den diagnosspecifika delen av det försäkringsmedicinska beslutsstödet anges nedanstående för dessa diagnoser.

5.1.2.1 Det försäkringsmedicinska beslutsstödet för depressiv episod - F32

Depressiv episod; ICD-10 kod: F32

Försäkringsmedicinsk information

Depression bör redan initialt utredas och behandlas aktivt med tät uppföljning och med skattningar av sjukdomens svårighetsgrad. Depression uppträder ofta tillsammans med andra psykiska och somatiska åkommor som ångestsjukdom, missbruk etc. Samsjuklighet komplicerar behandlingen och kan försämra prognosen. Okomplicerade förstagångsdepressioner uppnår vid adekvat behandling ofta förbättrad funktion inom 3 månader. Full läkning kan dröja upp till 6 månader och i vissa fall betydligt längre. Stora individuella skillnader förekommer. När behandlingsalternativen i primärvården är uttömda, eller om patienten inte är helt återställd inom 3-6 månader, överväg om behandlingsansvaret bör övergå till psykiatrin.

Symptom, prognos, behandling

Depressiv episod (F32) betecknar ett första insjuknande i depression. Om diagnosen förekommit tidigare, ska diagnosen recidiverande depression (F33) användas. Diagnosen omfattar sannolikt en rad något olika tillstånd, med delvis skilda symtom, prognos och svar på behandling. Symtombilden omfattar både somatiska och psykiska symtom. Kognitiv störning (t.ex. koncentrationssvårigheter, minnesstörning och/eller ökad uttrötthet) är ofta framträdande. Identifierbar utlösande orsak utesluter inte att det kan handla om en egentlig depression som kräver behandling.

Symtomen kan behandlas farmakologiskt och/eller i kombination med evidensbaserade psykologiska behandlingsmetoder. Full effekt av evidensbaserad psykologisk behandling kan ske redan inom 4 veckor men mest sannolikt efter minst 8 veckor beroende på svårighetsgrad.

Vid livshotande eller svårbehandlad depression är elbehandling (ECT), kombinationsbehandling med olika preparat samt med psykologisk behandling alternativ. Målsättningen med behandlingen är fullständigt tillfrisknande. Depressionens svårighetsgrad bör kontinuerligt följas med skattning eller självskattning för bedömning och eventuell omprövning av behandlingen.

Funktionsnedsättning

Funktionen sätts ned av trötthet, energibrist, oförmåga att fatta beslut, oförmåga att planera, försämrat minne, bristande initiativförmåga, motivation och uthållighet samt ökad känslighet för stress. Social rädsla är vanlig. Vid svår depression kan den drabbade personen inte hantera vardagliga problem eller ta vård om sig själv.

Vägledning om sjukskrivning vid olika situationer

Lindrig förstagångsdepression kräver ofta inte sjukskrivning. Vid sjukskrivning kan återgång i arbete normalt ske innan avslutad behandling. Överväg partiell sjukskrivning.

- Lindrig depressiv episod upp till 3 månaders sjukskrivning.
- Medelsvår till svår depressiv episod upp till 6 månaders sjukskrivning eller mer.

5.1.2.2 Det försäkringsmedicinska beslutsstödet för artros (M15-M19)

Inom diagnosgruppen artros finns ett flertal olika diagnoser och rekommendationerna varierar något mellan dessa. För Höft- och knäartros (M16, M17) finns separata skrivningar och för artros i första karpometakarpalleden (M18), ofta kallat karpaltunnelsyndrom eller tumbasartros, gäller samma rekommendationer som för vissa andra typer av hand- och handlovsbesvär. För Polyartros (M15) och Andra artroser (M19) finns för närvarande inga rekommendationer i beslutsstödet.

Höftartros; ICD-10 kod: M16*Försäkringsmedicinsk information*

Mycket tungt arbete bör undvikas helt. Ett lagom rörligt arbete med möjlighet att ofta ändra arbetsställning är idealt.

Symptom, prognos, behandling

Artros i höften är en sjukdom som långsamt försämrar den drabbade leden och dess omgivande vävnader. Smärta är det vanligaste symtomet. Andra symtom är stelhet, ledsvullnader och felställning i leden. Symtomen utvecklas över tid med omväxlande förbättrings- och försämringsperioder. Det är vanligt med återkommande försämringsperioder. För artros som behandlas ortopedkirurgiskt är 2–4 månaders tillfrisknande normalt.

Funktionsnedsättning

Smärta. Ledstelhet.

Rehabiliteringsåtgärder

Det kan vara aktuellt med t.ex. ergonomiska förbättringsåtgärder, rehabskola, anpassade arbetsuppgifter eller omplacering. Diskutera åtgärder vid tungt arbete och efter artoplastik. Om det inte går att anpassa arbetsuppgifterna bör byte av arbete övervägas tidigt av patienten. Företagshälsovården och/eller arbetsgivaren och Försäkringskassan bör kopplas in tidigt.

Vägledning om sjukskrivning vid olika situationer

- Vid arbeten som inte belastar höftlederna i någon större omfattning är arbetsförmågan normalt inte nedsatt, oavsett artrosens svårighet.
- Vid försämringsperioder, lätta eller medelsvåra artrosbesvär och vid arbete som innebär hög belastning (som tunga eller upprepade lyft), kan arbetsförmågan vara helt nedsatt i upp till 3 veckor.
- Vid svåra artrosbesvär och belastande arbete kan deltidssjukskrivning bli aktuellt i avvaktan på åtgärder.
- Efter artoplastik (byte av led) kan arbetsförmågan vara nedsatt i upp till 8 veckor i lätta arbeten som inte belastar leden och i upp till 4 månader vid måttligt tunga arbeten.

Knäartros; ICD-10 kod: M17*Försäkringsmedicinsk information*

Mycket tungt arbete bör undvikas helt. Ett lagom rörligt arbete med möjlighet att ofta ändra arbetsställning är idealt.

Symptom, prognos, behandling

Artros i knä är en sjukdom som långsamt försämrar den drabbade leden och dess omgivande vävnader. Smärta är det vanligaste symtomet. Andra symtom är stelhet, ledsvullnader och felställning i leden. Symtomen utvecklas över tid med omväxlande förbättrings- och försämringsperioder. Det är vanligt med återkommande försämringsperioder. För artros som behandlas ortopedkirurgiskt är 2–4 månaders tillfrisknande normalt.

Funktionsnedsättning

Smärta. Ledstelhet.

Aktivitetsbegränsning

Artros sätter ned gång- och rörelseförmågan. Funktionen varierar ofta över tid.

Rehabiliteringsåtgärder

Det kan vara aktuellt med t.ex. ergonomiska förbättringsåtgärder, anpassade arbetsuppgifter eller omplacering. Diskutera åtgärder vid tungt arbete och efter artoplastik. Om det inte går att anpassa arbetsuppgifterna bör byte av arbete övervägas tidigt av patienten. Mycket tungt arbete bör undvikas helt. Företagshälsovården och/eller arbetsgivaren och Försäkringskassan bör kopplas in tidigt.

Vägledning om sjukskrivning vid olika situationer

- Vid arbeten som inte belastar knälederna i någon större omfattning är arbetsförmågan normalt inte nedsatt, oavsett artrosens svårighet.
- Vid försämringsperioder, lätta eller medelsvåra artrosbesvär och vid arbete som innebär hög belastning (som tunga eller upprepade lyft), kan arbetsförmågan vara helt nedsatt i upp till 3 veckor.
- Vid svåra artrosbesvär och belastande arbete kan deltidssjukskrivning bli aktuellt i avvaktan på åtgärder.
- Efter artoplastik (byte av led) kan arbetsförmågan vara nedsatt i upp till 8 veckor i lätta arbeten som inte belastar leden och i upp till 4 månader vid måttligt tunga arbeten.

Vissa hand- och handlovsbesvär; ICD-10 kod: G56.2, M18, M65.3, M65.4, M70.0, M72.0*Försäkringsmedicinsk information*

Vid långdragna besvär eller inflammationer på annat ställe bör reumatisk leddsjukdom övervägas. Vid triggerfinger, de Quervains sjukdom, peritendinitis crepitans och ulnariskompression återfår patienten funktionen efter upp till 3 veckor.

Se över arbetssituationen vid peritendinitis crepitans och ulnariskompression som är tillstånd som kan uppkomma på grund av belastning. Vid långdragna besvär eller inflammationer på annat ställe bör reumatisk leddsjukdom övervägas.

Symptom, prognos, behandling

Behandlingen kan vara farmakologisk, sjukgymnastisk och i vissa fall kirurgisk.

Triggerfinger är en degenerativ knuta på böjskenan som gör det svårt att röra fingret. de Quervains gör senskidan till tummen inflammerad och förträngd och ger smärta över tumbasen. Peritendinitis crepitans orsakas av inflammation av senskidorna och ger smärta utefter senskidan. Ulnariskompression orsakas av att armbågsnerven kläms och ger sensibilitetsstörning och senare kraftreduktion i ring- och lillfinger. Dupuytrens kontraktur är en förtjockning och skrumpling av handflatans senplatta. Tillståndet ger oförmåga att räta ut ett eller flera fingrar. Tumbasartros är en sjukdom som långsamt ger förändringar i den drabbade leden och dess omgivande vävnader. Behandlingen kan vara farmakologisk, sjukgymnastisk och i vissa fall kirurgisk. Triggerfinger, de Quervains, peritendinitis crepitans och ulnariskompression sätter tillfälligt ned funktionen i handen. Vid Dupuytrens kontraktur ökar problemen gradvis med bristande funktion som följd. Tumledsartros ger smärta och nedsatt rörelseförmåga.

Funktionsnedsättning

Samtliga sjukdomar påverkar gripförmågan och finmotorik i varierande omfattning.

Vägledning om sjukskrivning vid olika situationer

Vid bedömning av arbetsförmåga ska hänsyn tas till om tillståndet gäller dominant (höger för högerhänta, vänster för vänsterhänta) eller icke dominant sida, förutsatt att detta är relevant för det arbete som är aktuellt.

- Triggerfinger, tumbasartros och Dupuytrens kontraktur sätter sällan ned arbetsförmågan. Vid arbete som innebär hög belastning kan arbetsförmågan vara helt nedsatt i upp till 3 veckor.
- Peritendinitis crepitans och ulnariskompression är akuta tillstånd som sällan kräver mer än 1 veckas sjukskrivning. Vid arbete som innebär hög belastning kan arbetsförmågan vara helt nedsatt i upp till 3 veckor.
- de Quervains sjukdom kan sätta ned arbetsförmågan i upp till 1 månad.

5.1.2.3 Det försäkringsmedicinska beslutsstödet för handlovsfraktur (S62)

Skador på övre extremiteten (sårskada, luxation, distorsion och fraktur);

ICD-10 koder: S40, S42, S43, S50, S52, S53, S60, S62, S63

Försäkringsmedicinsk information

Funktionen blir mer begränsad om det är patientens dominerande sidan som skadats. Vid större skador som lednära frakturer uppstår ofta viss bestående funktionsnedsättning i form av svaghet och försämrad rörlighet i den skadade armen. Flertalet, om de inte har väldigt tunga manuella arbeten, kan dock återgå i arbete.

Symptom, prognos, behandling

Symtomen vid skador på övre extremiteten (armarna) är smärta som uppträder i anslutning till trauma (yttre våld). Direkt våld utan fraktur kallas kontusion och indirekt ledvåld utan fraktur kallas distorsion (stukning). Om det vid våldet uppstår en fraktur blir diagnosen normalt fraktur med angivelse av placering. Luxation är en svårare distorsion då leden helt glider isär. Svårare skador i exempelvis axelleden kan ta flera månader att läka och ge bestående funktionsnedsättning. Frakturer kan vara av olika svårighetsgrad från sprickor till operationskrävande splitterfrakturer. Läkningen tar oftast lång tid (månader).

Funktionsnedsättning

Under läkningen har patienten nedsatt förmåga att använda berörd axel eller arm i varierande grad, delvis beroende på skadans art och placering. Funktionen blir mer begränsande om det är patientens dominerande sidan som skadats.

Vägledning om sjukskrivning vid olika situationer

- Kontusion och distorsion medför i normalfallet nedsatt arbetsförmåga främst i manuella arbeten och i arbeten där armen måste användas – och då i upp till 3 veckor.
- Axeldistorsioner och frakturer medför i normalfallet i första hand nedsatt arbetsförmåga i manuella arbeten och i arbeten där armen måste användas. Arbetsförmågan kan i dessa fall vara nedsatt i upp till 4 månader.

5.1.2.4 Det försäkringsmedicinska beslutsstödet för bröstcancer (C50)

Bröstcancer; ICD-10 kod: C50

Försäkringsmedicinsk information

Psykiska reaktioner som sätter ned arbetsförmågan är vanligt såväl vid diagnosbesked som under och en tid efter behandlingen. Vanligt förekommande är akut och/eller försenad krisreaktion, depressiva symtom, ångest, oro för återfall och sömnstörning (se adekvat rekommendation för psykiska diagnoser). Etablera en samtalskontakt så tidigt som möjligt för de patienter som behöver hjälp och stöd.

Uppmuntra till fysisk aktivitet för att minska trötthet och ge ökad livskvalitet. Ökad fysisk aktivitet innebär ingen ökad risk för lymfödem. Om patienten behöver avstå från arbete för att behandlingen ska kunna genomföras kan sjukpenning i förebyggande syfte bli aktuellt.

Sjukdomsförloppet vid spridd bröstcancer varierar från patient till patient. Patienten kan ha perioder med normal funktion och arbetsförmåga, men förloppet är alltid progredierande (tilltagande).

Symptom, prognos, behandling

Vid icke-spridd bröstcancer är behandlingen nästan alltid kirurgisk. En del av, eller hela, bröstet tas bort och ett lymfkörtelgrepp görs i armhålan. Någon form av tilläggsbehandling ges i regel med cytostatika före eller efter operation (oftast 5 månader) och/eller strålbehandling (4-6 veckor) och/eller hormonell behandling (5 år). En del patienter får antikroppsbehandling under ett års tid.

Det är viktigt att försöka bejaka det friska, att informera om att sjukdomen redan genom operationen är behandlad och att övrig behandling är ett tillägg för säkerhets skull (och inte för någon känd sjukdomsmanifestation).

Spridd bröstcancer är ett tillstånd där spridningen finns från början eller patienten återfaller i tidigare känd bröstcancer. Tillståndet är kroniskt med en på sikt dålig prognos, men långa remissioner (förbättringar i patientens tillstånd) förekommer. Behandling krävs alltid och sker med, cytostatika, strålning eller hormoner. Kompletterande medicinsk behandling som smärtbehandling är ofta aktuell.

Funktionsnedsättning

En operation påverkar i regel inte det fysiska funktionstillståndet i någon större omfattning. Lättare smärtor, stramning och ömhet i operationsområdet förekommer. Efter ett större ingrepp i armhålan kan rörligheten i axelleden påverkas negativt. Postoperativa komplikationer som sårinfektion eller nervsmärta kan påverka funktionen under en längre tid. Cytostatikabehandling medför hos de flesta illamående, trötthet (fatigue), kognitiv nedsättning som minnesstörning och koncentrationssvårighet, infektionskänslighet och muskelsmärta, vilket ofta sätter ned funktionen under hela behandlingsperioden. Trötthet och kognitiv nedsättning kan hos vissa patienter kvarstå under flera månader efter avslutad behandling och är inte sällan av handikappande grad. Påverkan på nervfunktionen i händer och fötter, med nedsatt känsel och försämrad motorik, kan ge en nedsättning av funktionstillståndet, som kan vara bestående. Strålbehandling kan ge viss sveda och lättare smärta i huden under ett par veckor i slutet av behandlingen och upp till en månad efter avslutad behandling. Trötthet är vanligt förekommande. Under strålbehandlingstiden har de patienter som fått cytostatikabehandling ofta en kvarstående och ibland förvärrad trötthet. Under komplikationsfri strålbehandling är arbetsförmågan ofta god. Hormonell behandling kan ge övergångsbesvär (som blodvallningar och svettningar) som hos de flesta inte påverkar funktionen. För en del patienter medför hormonell behandling sömnstörningar, nedsatt kognitiv förmåga, koncentrationssvårighet och ökad uttrötthet. Hos en del av patienterna som får aromatashämmare uppstår förändringar i senor och leder som kan sätta ned funktionen betydligt under hela behandlingstiden. Dessa besvär går oftast över när behandlingen avslutas. Ange komplikationen som sjukskrivningsorsak om den orsakar funktionsnedsättningen (diagnos: artralgi/tenosynovit syndrom).

Antikroppsbehandling medför hos vissa patienter biverkningar i form av diffusa ledbesvär och lös avföring, som kan sätta ned funktionen delvis. Hos ett fåtal uppträder hjärtsvikt som nedsätter funktionstillståndet betydligt (se rekommendation för hjärtsvikt).

Psykiska reaktioner kan påverka funktionen under en längre tid. Lymfödem som sätter ned armfunktionen betydligt kan uppstå tidigt i förloppet och även flera år efter avslutad behandling. Besvären kan vara bestående. Vid spridd bröstcancer är funktionen ofta ned satt på grund av symtom som smärtor och trötthet samt biverkningar av behandlingen.

Rehabiliteringsåtgärder

Arbetsanpassning och successiv återgång i arbete med partiell sjukskrivning kan vara lämpligt. Upptrappning av arbetstiden får ofta ske med långtidsperspektiv.

En tidig dialog med patienten om sjukskrivningen rekommenderas. Kontakt med arbetsplats och arbetsgivare angående situationen kan underlätta patientens återgång i arbete. Avstämningsmöte mellan vårdgivare, patient, arbetsgivare och Försäkringskassan är att rekommendera, så att rehabiliteringen kan påbörjas tidigt. En rehabiliteringsplan bör upprättas tidigt i förloppet och revideras löpande.

Vägledning om sjukskrivning vid olika situationer

- Vid bröstbevarande operation med mindre lymfkörtelgrepp utan tilläggsbehandling kan 3 veckors sjukskrivning på 100 % vara lämpligt.
- Vid lymfkörtelgrepp och/eller mastektomi utan tilläggsbehandling kan 6 veckors sjukskrivning på 100 % vara lämpligt.
- Vid postoperativ cytostatikabehandling kan 5 månaders sjukskrivning behövas. Hos en mindre andel patienter finns viss arbetsförmåga mellan behandlingarna.
- Vid preoperativ cytostatikabehandling krävs oftast längre sjukskrivningsperiod, mer än 5 månader.
- Vid cytostatikabehandling före strålbehandling är arbetsförmågan oftast helt nedsatt under behandlingen. Efter 1–2 månader kan gradvis återgång i arbete ofta inledas.
- Hormonell behandling eller antikroppsbehandling medför för de flesta en möjlighet att återgå i arbete åtminstone på deltid. Efter avslutad behandling återställs funktionen i allmänhet inom några månader.
- Vid spridd bröstcancer krävs oftast sjukskrivning under lång tid – ofta 1 år eller längre.

Del I, de analyser som utgår från dem som genomfördes i vår tidigare rapport⁽²⁾ och som handlar om risk att ett sjukskrivningsfall blir längre än förväntat, avseende två diagnoser, dels depressiv episod (F32), dels artros (M15-M19).

5.2 Personer med ett nytt sjukskrivningsfall ≥ 21 dagar i depressiv episod (F32) respektive i artros (M15-M19)

I analyserna om risk för att ett sjukskrivningsfall i depressiv episod ska bli längre än förväntat har två olika kohorter använts; dels de personer som hade ett nytt sjukskrivningsfall som under år 2010 blev minst 21 dagar långt, räknat från första dagen med sjukskrivning (startdatum 1), dels motsvarande grupp som under 2011 hade ett sådant nytt sjukskrivningsfall. Åren 2010 och 2011 valdes dels för att praxis enligt rehabiliteringskedjan (som infördes i sjukförsäkringssystemet den 1/7 2008) skulle ha hunnit etableras inom Försäkringskassan och hälso- och sjukvården och dels för att större andel av de sjukskrivna skulle vara nya i systemet baserat på rehabiliteringskedjan. Kohorten för 2011 användes främst för att validera prediktionsmodellerna som utvecklades baserat på 2010-kohorten. Personer kan finnas med i båda dessa kohorter.

Totalt blev drygt 340 000 sjukskrivningsfall minst 21 dagar långa under år 2010 och drygt 350 000 sjukskrivningsfall uppfyllde samma kriterier år 2011 (Tabell 1). I sammanställningen i Tabell 1 ingår alla sjukskrivningsfall vars tjugoförsta dag inföll under perioden 1/1-31/12 2010 respektive under perioden 1/1-31/12 2011. Det vill säga personer med minst ett sjukskrivningsfall som startade mellan den 12/12 2009 och den 11/12 2010 och som blev minst 21 dagar långt inkluderades i kohorten som användes för att utveckla modellerna, och personer med minst ett sjukskrivningsfall som startade mellan den 12/12 2010 och den 11/12 2011 och som blev minst 21 dagar långt inkluderades i kohorten som användes för att validera modellerna.

Av dessa sjukskrivningsfall som blev minst 21 dagar långa hade sju procent depressiv episod (F32) som första huvuddiagnos och drygt tre procent hade en artrosdiagnos (M15-M19) som första huvuddiagnos. För de personer som hade mer än ett fall (depressiv episod: 951 individer 2010 och 964 individer 2011; artros: 498 individer 2010 och 555 individer 2011) som uppfyllde inklusionskriterierna valdes, efter diskussion med beställaren, det första av dessa att utgå från i analyserna.

I våra tidigare preliminära analyser⁽²⁾ exkluderades vissa grupper av personer ur kohorterna som användes för modelleringen av risken att bli långtidssjukskriven. De som då exkluderades var dels de som avled inom uppföljningstiden, dels de som hade tidigare eller pågående sjuk- eller aktivitetsersättning och dels de som tidigare uppnådd maximal tid i sjukförsäkringen. I analyserna i denna rapport exkluderades istället endast de som var äldre än 64 år den 31 december året innan inklusionsåret, alltså under år 2009 respektive år 2010, eller de som inte var folkbokförda i Sverige detta datum (depressiv episod: 55 personer 2010 och 49 personer 2011; artros: 270 personer 2010 och 304 personer 2011).

För dem som avled under uppföljningstiden (depressiv episod: 32 personer 2010 och 33 personer 2011; artros: 6 personer 2010 och 4 personer 2011), det vill säga inom 180 respektive 120 dagar efter sjukskrivningsfallets start ändrades utfallet till 1. Det vill säga i analyserna sågs det som om deras sjukskrivningsfall blev mer än 180 respektive 120 dagar. För två tredjedelar av dem som var sjukskrivna i depressiv episod och som avled inom 180 dagar var dödsorsaken suicid. Bland de personer som var sjukskrivna i artros och som avled inom 120 dagar, var suicid dödsorsak för en tredjedel.

Tabell 1. Antal nya sjukskrivningsfall som blev minst 21 dagar långa under 2010 och 2011, samtliga sådana fall samt sjukskrivningsfall i depressiv episod respektive i artros. Antal personer med minst ett sådant sjukskrivningsfall samt antal personer 16-64 år med minst ett sådant sjukskrivningsfall i depressiv episod respektive i artros.

	2010	2011
Antal nya sjukskrivningsfall som översteg 21 dagar	343 167 fall	355 122 fall
Antal sjukskrivningsfall med depressiv episod (F32) som första diagnos	24 125 fall	24 339 fall
Antal personer med minst ett sjukskrivningsfall i depressiv episod	23 134 ind	23 287 ind
Antal personer (16-64 år) med minst ett sjukskrivningsfall i depressiv episod	23 079 ind	23 238 ind
Antal sjukskrivningsfall med artros (M15-M19) som första diagnos	11 640 fall	11 979 fall
Antal personer med minst ett sjukskrivningsfall i artros	11 121 ind	11 393 ind
Antal personer (16-64 år) med minst ett sjukskrivningsfall i artros	10 851 ind	11 089 ind

Del II, analyser om risk att bli sjukskriven vid två olika diagnoser; dels handlovsfraktur (S52), dels bröstcancer (C50)

5.3 Personer med handlovsfraktur (S62) eller bröstcancer (C50)

5.3.1 Personer med handlovsfraktur (S62)

I den övergripande diagnosgruppen S62 ingår flera undergrupper, relaterade till frakturer i olika ben i handlov och fingrar. Under 2010 registrerades drygt 56 000 besök inom specialiserad öppenvård där handlovsfraktur (ICD10-kod S62) var huvuddiagnos (87,5 %) eller en bidiagnos. Under samma period startade 2 300 inläggningar inom slutenvården där handlovsfraktur var huvuddiagnos (56,6 %) eller bidiagnos. Samtliga dessa besök utgjordes av ca 30 000 personer. Efter exklusion av dem som inte var folkbokförda i Sverige den 31 december 2009 eller var yngre än 16 år eller äldre än 64 detta datum var det 18 678 personer i kohorten. Därefter exkluderades även de som hade besök inom specialiserad öppen- eller slutenvården på grund av en handlovsfraktur under de två åren (730 dagarna) innan första noteringen av en handlovsfraktur under 2010. Detta för att tidpunkten för det första besöket

eller inläggningen under 2010 med större sannolikhet skulle vara tidpunkten för när handlovsfrakturen skedde, det vill säga, inte vara på grund av ett återbesök för en handlovsfraktur som inträffat tidigare. Dessutom exkluderades de personer som samtidigt hade en eller flera andra allvarliga skador, till exempel en skallskada eller fraktur på annat ställe på kroppen då handlovsfrakturen för dessa troligen inte var den primära orsaken till en kommande sjukskrivning. Detta innebar att kohorten till slut bestod av 16 337 personer. Datum för handlovsfrakturen, här definierat som det första vårdtillfället för handlovsfraktur under år 2010, kallas här för T_0 .

Inför skapandet av en prediktionsmodell för risken att bli sjukskriven på grund av handlovsfrakturen exkluderades de som redan var sjukskrivna samt de som hade sjuk- eller aktivitetsersättning om minst 75 procent (1 872 personer; 11,5 %). Detta innebar att kohorten som låg till grund för dessa modelleringar bestod av 14 465 personer. Utfallet i prediktionsmodellen var att få ett nytt sjukskrivningsfall med handlovsfraktur (S62) som huvuddiagnos, mellan dagen före vård dagen (här kallad T_{-1}) och dag sex efter vård dagen (T_6). Orsaken till att även dagen före den första registrerade vård dagen inkluderades var att några personers sjukskrivningsfall i handlovsfraktur påbörjats redan då, möjligen på grund av att personen av olika anledningar inte sökt vård direkt efter skadan alternativt först besökt primärvården som lett till remiss till specialiserad vård. Orsak till valet av att även inkludera sjukskrivningsfall som påbörjats först några dagar senare är att få med dem som på grund av till exempel helgdagar väntat några dagar med att sjukanmäla sig.

5.3.2 Personer med bröstcancer (C50)

I analyser av personer med bröstcancer valde vi att endast inkludera de kvinnor som fått denna diagnos. Den absoluta majoriteten av personer som får en sådan diagnos är kvinnor.

Ur cancerregistret selekterades de kvinnor som under året 2010 registrerats med en första malign bröstcancer under 2010 (ICD10-kod: C50) och som i december 2009 var högst 63 år gamla. Personer som inte fanns i LISA någon gång under åren 2007-2013 exkluderades (1 person) samt personer där grunden för diagnos befanns vara obduktion, det vill säga, efter dödsfall (2 personer). Den slutliga kohorten bestod av 3536 kvinnor, den yngsta var 18 år.

Indelning i cancerstadium gjordes baserat på T-, N- och M- klassificeringsvariabler från cancerregistret, enligt WHO⁽¹⁵⁻¹⁷⁾. T-, N- och M-variablerna anger tumörutbredning vid diagnostillfället, där T: tumörens storlek, N: lymfkörtlar, M: fjärrmetastaser. Om värden för alla dessa variabler saknades för en person kunde inte indelning i cancerstadium göras (25 kvinnor). Om information för en eller två av dessa variabler saknades eller var kodade med ett 'X' (=kan ej bedömas) sattes värdet till 0. Sjutton av kvinnorna hade värdet X på alla tre variabler, i dessa fall ersattes X med 0. I de fall flera poster (tumörer) fanns med i cancerregistret för samma person år 2010 men med olika datum, valdes den tumör och datum för diagnos med mest avancerade cancerstadium. I ett fall valdes dock den första tumören även om stadium för denna hade lägre grad, då det var 252 dagar mellan datumen för de två diagnoserna under 2010. I övrigt var det mycket små skillnader mellan diagnosdatum då en annan tumör än den första har valts. Cancerstadium delades in i följande kategorier: alla T, N,

M saknas; TONOMO+Stadium 0 + I; Stadium II; Stadium III+IV. Ju högre stadium desto allvarligare är sjukdomen.

Avseende vilka diagnoser som skulle ingå i utfallet i modelleringen för sannolikheten att bli sjukskriven i samband med bröstcancerdiagnos inkluderades, efter diskussion med SRS-projektet, förutom denna specifika diagnos (C50) även annan cancer (C00-D48), diagnoser relaterade till bröstcancer (Z80, Z85, N61-N63) samt psykiska diagnoser (F00-F99 och Z73). Orsaken till att även andra cancerdiagnoser togs med är att kvinnan även kan ha haft andra cancerdiagnoser, som angivits först på intyget samt att ibland anges en mer övergripande än en specifik diagnos på ett sjukintyg⁽¹⁸⁾. Då några sjukskrivningsdiagnoser istället var andra, men cancerrelaterade diagnoser, såsom Z80- malign tumör i familjens sjukhistoria, Z85- malign tumör i den egna sjukhistorien, N61- Inflammatoriska sjukdomar i bröstkörtel, N62- Hypertrofi av bröstkörtel samt N63 - Icke specificerad knuta i bröstkörtel, inkluderades även dessa, då de kan anses relaterade till cancerdiagnosen och till exempel ibland angivits innan det var klart från provsvar om det var en malign cancer. På motsvarande sätt leder en cancerdiagnos till stor oro, ångest eller stress för vissa personer^(19, 20), varför även sjukskrivning i psykiska diagnoser i samband med bröstcancerdiagnosen togs med. Däremot togs inte andra somatiska diagnoser, såsom frakturer, med.

Tidsperioden för modelleringen valdes, baserat på data om sjukfrånvaro i dessa diagnoser före diagnosdatum och efter diskussioner med SRS ledningen, till 14 dagar innan diagnosdatum till och med 29 dagar efter. I kommande analyser finns det anledning att även pröva andra tidsperioder. Sammantaget definierades utfallet i prediktionsmodellen som att få ett nytt sjukskrivningsfall som varade längre än 14 dagar i någon av de ovan beskrivna diagnoserna (C00-D48, Z80, Z85, N61-N63, F00-F99 och Z73) under tidsperioden från 14 dagar innan bröstcancerdiagnosen (T_0) till och med 29 dagar efter bröstcancerdiagnosen. *I tabeller, figurer och i resultatbeskrivningarna nedan använder vi termen "cancer" på så sätt att den innefattar även de cancerrelaterade diagnoserna, det vill säga med "cancer" avses alltså följande ICD-koder: C00-D48, Z80, Z85, N61-N63.*

Inför framtagandet av modellen exkluderades de kvinnor som hade ett pågående sjukskrivningsfall samt de som hade pågående sjuk- eller aktivitetsersättning om minst 75 procent (totalt 521 kvinnor; 14,7 %). Som pågående sjukskrivningsfall definierades fall som startade mer än 14 dagar innan bröstcancerdiagnosen och som fortfarande pågick vid diagnosdatum, samt fall som startade mindre än 14 dagar innan diagnosdatum och pågick vid diagnosdatum men i andra sjukskrivningsdiagnoser än de studerade, det vill säga i andra diagnoser än cancer eller psykiska besvär. Ytterligare 61 kvinnor exkluderades efter detta, dels på grund av saknade värden för variabler ingående i modelleringen och dels på grund av extrema värden på vissa av de kontinuerliga variablerna. Femton kvinnor saknade värden på alla T, N, M-variabler och därmed kunde ingen tilldelning i cancerstadium göras; två kvinnor bodde ej i Sverige året innan och därmed kunde inte variabler för tidigare sjukskrivning, sjuk- eller aktivitetsersättning samt tidigare vård beräknas. Extrema värden, alltså sådana värden på observationer som ligger långt från de övriga och därmed skulle kunna riskera att göra modelleringen ostabil förekom i följande variabler; Antal besök i öppenvården under två år

innan diagnos >20: 9 personer, Antal dagar i slutenvård under två år innan diagnos >20: 9 personer, Antal nettodagar i sjukskrivning två år innan diagnos >300: 18 personer, Antal nettodagar i sjuk- eller aktivitetsersättning två år innan diagnos >500: 12 personer. Antal kvinnor som ingick i kohorten som prediktionsmodellen baserades på uppgick till 2954 efter dessa exkluderingar.

5.4 Kategorisering av variabler

Det finns ett stort antal faktorer, på olika strukturella nivåer, som kan ha samband med om en person blir sjukskriven respektive ha samband med om ett sjukskrivningsfall blir långt eller inte⁽⁵⁾. I analyserna vars resultat presenteras i denna rapport inkluderades några sådana faktorer på individnivå, nämligen sociodemografiska faktorer som kön, ålder, födelseland, utbildningsnivå, boenderegion (både geografisk region och så kallad H-region, enligt Statistiska centralbyråns indelning av kommuner efter hur urbana de kan anses vara), familjesituation, civilstånd och typ av yrke/sysselsättning. Dessutom användes information om personens sjuklighet både under tiden före sjukskrivningen/skadan/diagnosen samt under de första 21 dagarna av sjukskrivningsfallet. För information om sjuklighet har vi använt information om slutenvård och specialiserad öppenvård, tidigare sjukskrivningar (varaktighet och diagnoser), omfattning (hel- eller deltid) i början av det studerade sjukskrivningsfallet samt uttag av receptbelagda läkemedel.

Nedan följer beskrivningar av hur dessa faktorer har kategoriserats. Några av faktorerna har kategoriserats på något olika sätt i de olika analyserna, detta för att de analyserade grupperna har olika sammansättningar av personer.

För slutenvård samt för besök inom specialiserad öppenvård har två typer av vårdtillfällen exkluderats då dessa inte kan sägas vara relaterade till sjuklighet enligt denna typ av analys. Det gäller dels normala förlossningar (ICD-kod O80) dels vårdkontakter på grund av andra orsaker än sjukdom eller skada, till exempel screeningundersökningar (ICD-koder Z00-Z99). Av dessa Z-koder inkluderades dock en som en psykisk diagnos i analys av personer med handlovsfrakturer eller bröstcancer, nämligen Z73.0; ”utbrändhet”.

Vissa personer som inkluderades i kohorten av personer med handlovsfrakturer hade fler än ett öppenvårdsbesök den aktuella dagen eller både besök inom öppenvården och slutenvården under samma dag. Information från samtliga dessa besök har legat till grund vid skapandet av de faktorer som baseras på öppenvårds- och slutenvårdsregistret, till exempel antal operationskoder och om handlovsfraktur (ICD-10 kod S62) var en huvuddiagnos eller inte.

Vid uppdelningar av ålderskategorier har olika uppdelningar använts för de olika delstudierna. Såväl personer sjukskrivna med artros och kvinnor med bröstcancer är vanligen något äldre; ålderskategorierna har anpassats efter detta. En följd av den skeva åldersfördelning i dessa kohorter är att även andra faktorer påverkas, till exempel består gruppen ’hemmaboende <20 år’ av få personer, dessa har därför kategoriserats som ’ensamstående utan barn’.

Även typ av sysselsättning vid sjukskrivningsfallets start kategoriserades på något olika sätt för de två diagnosgrupperna depressiv episod och artros. För dem sjukskrivna i depressiv

episod kategoriserades studenterna till samma grupp som egenföretagarna, då studenterna var få och lika stora andelar i dessa två grupper hade sjukskrivningsfall som varade i minst 180 dagar. Bland dem sjukskrivna i artros fanns inga studenter, istället utgjorde föräldralediga personer en mycket liten grupp och dessa kategoriserades tillsammans med gruppen anställda.

Variabeln sysselsättning vid sjukskrivningsfallets start användes inte i analyser av personer som drabbats av en handlovsfraktur eller diagnostiserats med bröstcancer, då den variabeln är baserad på data från MiDAS och gäller vid tidpunkten för start av ett nytt sjukskrivningsfall – och inte alla personer i de analyserna blev sjukskrivna. I dessa analyser användes istället information från LISA om yrke på två sätt, dels typ av yrke kategoriserat till tre grupper; utan yrke, arbetare respektive tjänstemän, dels typ av sektor även den kategoriserad i tre grupper: ej anställd, anställd inom privata sektorn och anställd inom offentliga sektorn.

För några personer saknades information om utbildningsnivå respektive födelseland. För de personer där information om utbildningsnivå saknades sattes den till den lägsta nivån (grundskola) och de personer där information om födelseland saknades sattes i gruppen övriga världen. Personer som inte fanns i LISA under två på varandra följande år och, enligt dödsorsaksregistret, ännu ej hade avlidit kategoriserades som emigrerade. Emigrationsdatum sattes till 31 december det första av dessa två år. Inflyttning till Sverige (före 2010): Om en person inte fanns registrerad i LISA under två efterföljande år innan det aktuella året, definierades personen som att den inte bodde i Sverige året innan det aktuella året.

Antalet tidigare nettodagar med sjukpenning eller med sjuk- och aktivitetsersättning har beräknats inför analyserna. Då har antalet ersatta dagar (bruttodagar) relaterats till grad. Till exempel blir fyra dagar med sjukskrivning på 25 procent av ordinarie arbetstid en nettodag. Det sammanlagda antalet nettodagar med sjukpenning under de 365 respektive de 730 dagarna innan det aktuella sjukskrivningsfallet startade alternativt första noterade händelse av handlovsfraktur eller bröstcancer beräknades inför analyserna. I beräkningarna inkluderades då endast dagar från sjukskrivningsfall som varat längre än 14 dagar och endast dagar från och med dag 15 i fallet. Detta för att undvika bias för dem som var arbetslösa eller av annan anledning blev ersatta från Försäkringskassan innan dag 15 i sjukskrivningsfallet. I modellerna för prediktion av långtidssjukskrivning inkluderades en faktor som indikerade om man hade haft sjuk- eller aktivitetsersättning (tidigare kallat förtidspension eller sjukbidrag) från och med år 1994.

Geografisk regionindelning har gjorts i enlighet med den indelning som används inom Försäkringskassan år 2015:

- Nord: Gävleborg, Västernorrland, Jämtland, Västerbotten, Norrbotten
- Mitt: Södermanland, Uppsala, Västmanland, Östergötland
- Stockholm: Stockholm, Gotland
- Väst: Västra Götaland, Örebro, Halland, Dalarna, Värmland
- Syd: Skåne, Blekinge, Kronoberg, Jönköping, Kalmar

Kategorisering av övriga variabler framgår av tabeller och nedanstående text.

5.5 Dataanalys

Denna rapport består, som framgår ovan, av tre delar, varav analyser görs i de två första och förslag till analyser ges i den tredje delen. I den första delen har de modeller som presenterades i en tidigare rapport⁽²⁾ vidareutvecklats något och sedan validerats på en kohort baserad på kommande år. I den andra delen presenteras dels deskriptiva resultat för personer i två olika diagnosgrupper samt prediktionsmodeller för personernas risk att bli sjukskriven i respektive diagnos. I båda dessa delar används logistisk regression⁽²¹⁾ för att utveckla prediktionsmodellerna. I den andra delen inkluderades även splines⁽²²⁾ för kontinuerliga variabler i prediktionsmodellerna.

Om sjukskrivningsfall: i samtliga analyser har endast sjukskrivningsfall som blivit minst 15 dagar långa inkluderats. I analyser av risk för att få ett nytt sjukskrivningsfall har personer med pågående sjuk- eller aktivitetsersättning exkluderats, om ersättningen hade en omfattning om 100 eller 75 procent av heltid. Den senare gruppen exkluderades för att inte introducera en könsbias i analyserna, en större andel kvinnor har fått sjuk- eller aktivitetsersättning om 75 procent även om de inte kunnat yrkesarbeta på resterande 25 procent. Effekter av detta tillvägagångsätt behöver naturligtvis utvärderas i kommande analyser.

5.5.1 Utveckling av prediktionsmodeller

När man, som i denna delstudie väljer att ha en dikotom utfallsvariabel som kan anta två värden (0=personen fick inte utfallet respektive 1=personen fick utfallet) och flera olika möjliga förklarande variabler, använder man som standardmetod multivariabel logistisk regression för att skatta en prediktionsmodell⁽²¹⁾. Utfallsvariabeln är här dels långtidssjukskrivning; mer än 180 dagar för depressiv episod och mer än 120 dagar för artros och dels att bli sjukskriven i specifika diagnoser i samband med en handlovsfraktur respektive en bröstcancerdiagnos. De testade förklarande variablerna beskrivs mer ingående i resultatdelen.

Den predicerade sannolikheten att en person blir långtidssjukskriven ($y=1$) givet vissa förklarande variabler (X) kan skrivas $P(y=1|\mathbf{x})$, detta är ett värde mellan 0 och 1 och beräknas på följande sätt:

$$P(y = 1|\mathbf{x}) = \frac{e^{g(\mathbf{x})}}{1 + e^{g(\mathbf{x})}}$$

$$g(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

där $\beta_0, \beta_1, \dots, \beta_k$ är konstanter (Beta-koefficienter) som skattas från data och presenteras i resultatdelen (Tabell 2 och 5) och x_1, x_2, \dots, x_k är förklarande variabler/faktorer. Samtliga variabler i den första delen i denna rapport är kategoriska, det vill säga varje värde på variablerna får en egen beta-koefficient, utom referenskategorierna. En annan aspekt som behöver beaktas är att x_i ofta innebär förändringar av tillgängliga prediktorer, i form av interaktioner, polynomer och splines. Här har "natural cubic splines" (se till exempel^(22, 23)) använts i avsnitt II för att modellera potentiella ickelinjära effekter av kontinuerliga variabler.

Det är möjligt att komma fram till den "bästa" möjliga prediktiva modellen på flera olika sätt. För prediktionsmodellerna om risk för långtidssjukskrivning i den första delen har vi valt att använda den så kallade "backward stepwise selection metoden". Det innebär att vi har börjat med en multivariabel modell innehållande de variabler som inkluderats och sedan stegvis exkluderat de variabler som inte var signifikanta ($p > 0,05$) i den multivariabla modellen⁽²⁾. Detta leder fram till en modell som endast innehåller statistiskt signifikanta variabler.

5.5.2 Utvärdering och validering av prediktionsmodeller

I den första delen i resultatet har så kallad temporal validering⁽²⁴⁾ av prediktionsmodellerna genomförts, vilket innebär att kohorten för valideringen är framtagen på samma sätt som kohorten som användes för utvecklingen av modellen men för en annan tidsperiod. Valideringen gick till så att modellerna (för risken att sjukskrivningsfallet blev mer än 180 respektive 120 dagar långt för dem sjukskrivna minst 21 dagar i ett nytt fall i depressiv episod respektive i artros) som utvecklades baserat på 2010-kohorterna applicerades på 2011-kohorterna genom att prediktionsvärden (risker) beräknades för samtliga personer i dessa valideringskohorter. Därefter beräknades flera olika mått utifrån dessa prediktionsvärden som jämförs med motsvarande mått baserade på prediktionsvärdena från utvecklingskohorterna, det vill säga kohorterna av personer från 2010.

För att ha användning av en prediktionsmodell där utfallet, som i dessa analyser, är binärt behöver man välja en brytpunkt som skiljer på dem som har hög risk respektive låg risk för utfallet. Detta kan göras på olika sätt, beroende på hur viktigt det är att man inte klassificerar personer till 'fel' grupp när det gäller hög eller låg risk för utfallet. En optimal brytpunkt kan väljas där totala antalet korrekt predicerade blir så stort som möjligt, men man kan till exempel också man välja en brytpunkt där andelen falskt negativa blir lika stor som andelen falskt positiva.

För att utvärdera de framtagna modellernas interna validitet genomfördes en serie statistiska tester⁽²⁵⁾. Sammantaget ger dessa en nyanserad uppfattning av prediktionsmodellernas prestanda och egenskaper. De använda måtten kan klassificeras som antingen; a) diskrimineringsmått, vilka syftar till att undersöka hur väl modellen kan särskilja mellan dem som fick det undersökta utfallet gentemot dem som inte fick utfallet, det vill säga AUC (Area under curve)^(26, 27) samt diskrimineringslutningen⁽²⁸⁾; b) kalibreringsmått⁽²⁹⁾, vilka syftar till att visa hur väl modellens predicerade utfallsvärden överensstämmer med de faktiskt observerade utfallen, det vill säga Hosmer-Lemeshow test, "calibration-in-the-large" och "calibration slope"⁽³⁰⁾; samt c) övergripande mått, vilka sammanfattar modellens diskriminerings- och kalibreringsegenskaper i ett enskilt mått; R^2 ⁽³¹⁾ och Brier värdet⁽³²⁾.

AUC är arean under ROC-kurvan som plottar 1-specificiteten (andelen bland dem som *inte* fick utfallet men som predicerades att få utfallet) mot sensitiviteten (andelen bland dem som fick utfallet och som predicerades att få utfallet) med varje persons prediktion som brytpunkt. Arean under kurvan kan anta värden mellan 0,5 och 1, där 0,5 innebär att modellen inte är bättre på att predicera utfallet än att singla slant skulle varit, och 1 innebär att modellen är en perfekt beskrivning av verkligheten. Diskrimineringslutningen är skillnaden mellan

medelvärden av prediktionsvärdena för den grupp som fick utfallet och medelvärdet för den grupp som inte fick utfallet. Lutningen kan alltså variera mellan 1 för en perfekt modell och 0 för en modell som inte alls kan urskilja dessa två grupper.

Hosmer-Lemeshow's test är ett "goodness-of-fit test" avser att testa överensstämmelse mellan prediktionsvärdet och utfallet. Personerna i kohorten delas in i deciler efter deras prediktionsvärden, därefter jämförs medelvärdet av prediktionsvärdena med andelen som fick utfallet i varje decil med hjälp av ett χ^2 test. Om modellen har en bra prediktiv förmåga så är skillnaderna mellan dessa två värden inte signifikant.

Ett annat sätt att ta reda på om prediktionsvärdet (sannolikheten för utfallet) stämmer överens med det faktiska utfallet kan man göra en så kallad "calibration plot". Den plottar prediktionsvärdena från modellen mot andelen som fick utfallet bland de personer som fick samma prediktionsvärde. Bland dem som fick ett prediktionsvärde på 0,67 bör alltså 67 procent ha fått utfallet, detta medför att plotten bör bli en rak linje med lutning ("calibration slope") 1 och intercept ("calibration-in-the-large") 0. Denna metod kan användas för att se om modellen som utvecklats på en kohort fungerar att använda på en annan kohort. Om lutningen då blir mindre än 1 tyder det på att modellen ger för extrema prediktionsvärden. Om interceptet är större än 0 är prediktionsvärdena genomgående för låga och om interceptet är mindre än 0 är prediktionsvärdena genomgående för höga.

Analysen är utförd med SAS, SPSS samt R.

6 Resultat

Resultaten i denna rapport är uppdelade i tre delar. Först presenteras modeller som predicerar risken att ett påbörjat sjukskrivningsfall blir långt bland personer sjukskrivna i depressiv episod respektive i artros. Detta är en uppföljning och fördjupning av analyser gjorda i en tidigare rapport ⁽²⁾. Här presenteras även resultat från valideringar av dessa modeller med en liknande kohort som de är utvecklade med men från ett senare år. Dessutom ges resultat från en utvärdering av användande av de exklusionskriterier som användes i den tidigare rapporten ⁽²⁾, samt från analyser av modeller baserade på olika grupper av prediktionsfaktorer.

I den andra delen av resultatet presenteras först deskriptiva analyser om sjukfrånvaro dels bland personer som vårdats för en handlovsfraktur under 2010 och dels bland kvinnor som diagnostiserades med bröstcancer under 2010. Därefter presenteras prediktionsmodeller för risken att bli sjukskriven bland personer som fått dessa diagnoser.

I den tredje delen presenteras förslag till framtagande av prediktionsmodeller, dels av risken att bli sjukskriven respektive sjukskriven i olika längd vid olika diagnoser, dels för att personer som är sjukskrivna i psykiska eller muskuloskeletala diagnoser vid en vårdcentral ska blir långvarigt sjukskrivna.

Resultat, del I

6.1 Prediktionsmodeller för risken att ett sjukskrivningsfall blir längre än förväntat

Ett av syftena med denna rapport var att fördjupa och förbättra tidigare genomförda analyser ⁽²⁾. En av skillnaderna mellan de analyser som presenteras i denna rapport jämfört med de i den tidigare är att i utvecklingen av modellerna inkluderades sjukskrivna personer under hela år 2010 istället för endast det första halvåret. En annan stor skillnad är vilka exklusionskriterier som användes inför skapandet av kohorterna. Efter dialog med SRS:s projektledning beslutades det att inte exkludera de personer som tidigare uppnått maximal tid i sjukförsäkringen samt de som hade pågående partiell sjuk- eller aktivitetsersättning eller som haft sådan partiell eller på heltid någon gång under det föregående 12 månaderna. Efter dessa ändringar bearbetades det fram två nya prediktionsmodeller, en för risken att bli långtidssjukskriven (mer än 180 dagar) bland dem som hade ett nytt sjukskrivningsfall i depressiv episod (F32) som blev minst 21 dagar långt och en för risken att bli långtidssjukskriven (mer än 120 dagar) bland dem som hade ett nytt sjukskrivningsfall i artros (M15-M19) som blev minst 21 dagar långt. Dessa två modeller utvecklades baserat på en kohort från 2010 och validerades på en kohort från 2011.

I avsnitt 6.1.3 nedan undersöks om de exkluderingar som tidigare gjordes var rimliga. Anledningen till att de exkluderades var att det redan fanns viss kunskap om att dessa personer hade en hög risk att bli långtidssjukskrivna.

I det sista avsnittet i denna första del av resultatet jämförs modeller med olika grupper av prediktionsfaktorer med varandra. I några modeller exkluderas de faktorer som Försäkringskassan inte skulle kunna applicera i modellerna utifrån de data de har tillgång till. I andra modeller läggs ytterligare faktorer till. Anledningen till att dessa nya faktorer (om antal barn samt om yrke) inte är med i modellen som validerades är för att tidsåtgången för de olika analyserna var osäker och efter diskussion med beställarna prioriterade att få analyserna klara i tid än att skapa nya faktorer till modellerna.

6.1.1 Depressiv episod – sjukskrivningsfall som varat i minst 21 dagar

Totalt 23 079 personer som var 16-64 år gamla hade minst ett nytt sjukskrivningsfall i depressiv episod som under år 2010 blev minst 21 dagar långt. Motsvarande antal personer som inkluderades på liknande sätt år 2011 var 23 238. Några av dessa personer hade två eller fler sjukskrivningsfall som uppfyllde inklusionskriterierna; det första av dessa fall valdes inför analyserna. Av de sjukskrivningsfall från år 2010 som analyserades blev 33,4 procent längre än 180 dagar. Bland sjukskrivningsfallen från år 2011 blev 33,8 procent längre än 180 dagar. I båda kohorterna var nästan 70 procent kvinnor och 37 procent var ensamstående utan barn.

Utfallet som prediceras i modellen är att fortfarande vara sjukskriven (i någon omfattning) efter 180 dagar från när fallet påbörjades. De som avled inom 180 dagar efter sjukskrivningsfallets start (32 personer år 2010 och 33 personer år 2011) kategoriserades som att de varit sjukskrivna i minst 180 dagar. Två tredjedelar av dem som avled inom uppföljningstiden avled på grund av suicid.

I modellen inkluderades 17 olika variabler (Tabell 2), vilka kan delas in i tre kategorier:

- Sociodemografiska: Ålder, Födelseland, Utbildningsnivå, H-region, Sysselsättning
- Relaterade till sjukförsäkringen: Uppnådd maxtid, Tidigare sjuk- eller aktivitetsersättning, Nettodagar med sjukpenning året innan, Omfattning vid start, Tidigare sjukskrivning med samma diagnos
- Relaterade till sjukvård och medicinering: Slutenvård under de första 21 dagarna av sjukskrivningsfallet, Öppenvård under de första 21 dagarna av fallet, Diagnosspecifik öppenvård (tre variabler) dels under de 365 dagarna innan sjukskrivningsfallet började och dels under de första 21 dagarna, Uttag av antidepressiva läkemedel dels under de 365 dagarna innan sjukskrivningsfallet började och dels under de 21 första dagarna.

Hur variablerna operationaliserats framgår av Tabell 2. En tredjedel (n=7 713) fick utfallet att sjukskrivningen blev mer än 180 dagar år 2010, denna andel var samma även år 2011 (n=7 845). Personer yngre än 35 år hade en lägre risk att bli långtidssjukskrivna (>180 dagar) jämfört med äldre (Tabell 2). Faktorerna tidigare uppnådd maxtid, sjuk- eller aktivitetsersättning samt tidigare nettodagar med sjukskrivning indikerade högre risk, likaså att ha påbörjat det aktuella sjukskrivningsfallet på heltid. Faktorn att tidigare ha varit sjukskriven i depressiv episod gav dock en lägre risk att bli långtidssjukskriven.

Tabell 2. Antal personer i varje grupp bland de personer som hade ett nytt sjukskrivningsfall i depressiv episod (F32) som år 2010 blev minst 21 dagar långt, och andel av dessa vars sjukskrivningsfall blev längre än 180 dagar samt betavärdet och oddsraten (OR) med 95 % konfidensintervall (KI) för de variabler som inkluderades i modellen som predicerar risken att bli sjukskriven längre än 180 dagar för dessa personer.

	Antal 2010	Andel (%) >180 dagar	Beta	OR (95 % KI)
Samtliga personer / Intercept (β_0)	23079	33,4%	-1,367	
Åldersgrupper ¹				
16-24 år	1432	22,1%	-0,517	0,60 (0,52-0,69)
25-34 år	5061	29,1%	-0,196	0,82 (0,76-0,89)
35-44 år	6840	35,0%	0	1
45-54 år	5956	36,4%	-0,041	0,96 (0,89-1,04)
55-64 år	3790	35,8%	-0,012	0,99 (0,90-1,08)
Födelseland				
Sverige	19251	32,4%	0	1
Övriga Norden	695	31,4%	-0,191	0,83 (0,69-0,98)
Övriga EU25	573	35,8%	-0,030	0,97 (0,80-1,17)
Övriga världen	2560	41,4%	0,101	1,11 (1,01-1,22)
Utbildningsnivå ¹				
Grundskola	3212	36,6%	0	1
Gymnasium	11707	32,6%	-0,078	0,92 (0,85-1,01)
Eftergymnasial	8160	33,3%	0,080	1,08 (0,99-1,19)
H-region ¹				
Storstad	9036	34,3%	0	1
Mellanstor stad	8023	32,4%	-0,125	0,88 (0,82-0,95)
Landsbygd	6020	33,5%	-0,072	0,93 (0,86-1,00)
Sysselsättning vid sjukskrivningsfallets början				
Anställd	18802	29,1%	0	1
Arbetslös	3134	58,7%	0,696	2,01 (1,83-2,19)
Egenföretagare/Student	693	38,7%	0,291	1,34 (1,13-1,58)
Föräldraledig	450	28,2%	-0,121	0,89 (0,71-1,10)
Uppnått maxtid i sjukförsäkringen ³	1648	83,1%	1,794	6,01 (5,12-7,07)
Pågående eller tidigare SA ^{3,4}	2401	47,8%	0,163	1,18 (1,06-1,30)
Omfattning vid sjukskrivningsfallets början				
Deltid 25-75%	3450	27,6%	0	1
Heltid 100 %	19629	34,4%	0,372	1,45 (1,33-1,59)
Tidigare sjukskrivning, nettodagar ²				
Ingen tidigare sjukskrivning	16159	29,7%	0	1
0,25 - 49,75 Nettodagar	3711	30,2%	0,096	1,10 (1,01-1,20)
50 - 89,75 Nettodagar	954	37,3%	0,315	1,37 (1,18-1,59)
90 - 179,75 Nettodagar	996	51,1%	0,654	1,92 (1,66-2,23)
180 - 358 Nettodagar	1259	74,1%	0,668	1,95 (1,63-2,33)
Diagnoser i tidigare sjukskrivningsfall ²				
Ingen sjukskrivning i F32	20936	32,4%	0	1
F32 - depressiv episod	2143	43,8%	-0,304	0,74 (0,65-0,83)

Forts tabell 2

Slutenvård de första 21 dagarna				
Ingen slutenvård	21445	32,9%	0	1
Minst en dag (samtliga diagnoser)	1634	40,3%	0,315	1,37 (1,23-1,53)
Besök öppenvården, de första 21 dagarna				
Inga besök	18924	32,1%	0	1
Ett besök	3043	39,5%	0,258	1,29 (1,16-1,44)
2 eller fler besök	1112	38,8%	0,291	1,34 (1,14-1,57)
Besök öppenvården, de första 21 dagarna				
Övriga diagnoser (somatiska) ³	2083	37,1%	-0,173	0,84 (0,73-0,96)
Tidigare besök öppenvården ²				
F32 - depressiv episod ³	1584	50,9%	0,203	1,23 (1,08-1,38)
Övriga psykiska diagnoser ³	2433	46,7%	0,115	1,12 (1,01-1,24)
Tidigare läkemedelsuttag ^{2, 5}				
Inga uttag	12758	29,5%	0	1
1-2 uttag	4264	33,3%	0,021	1,02 (0,94-1,11)
3 eller fler uttag	6057	41,8%	0,139	1,15 (1,07-1,24)
Läkemedelsuttag ⁵ de första 21 dagarna				
Inga uttag	13138	33,7%	0	1
1-2 uttag	9647	32,6%	0,135	1,14 (1,08-1,22)
3 eller fler uttag	294	45,6%	0,392	1,48 (1,15-1,91)

¹ Gällande december 2009.

² Med tidigare avses under ett år (365 dagar) innan det studerade sjukskrivningsfallet påbörjades

³ Referensgrupp är att **inte** ha uppnått maxtid, **inte** haft tidigare eller pågående SA respektive att **inte** haft något besök inom specialiserad öppenvård med respektive diagnos som huvuddiagnos.

⁴ Med tidigare avses från och med 1994 till och med dag 21 i sjukskrivningen.

⁵ Uttag av antidepressiva läkemedel.

Jämfört med den modell som presenterades i vår tidigare rapport⁽²⁾ tillkom här tre variabler, nämligen utbildningsnivå samt två variabler om diagnosspecifik tidigare specialiserad öppenvård.

De variabler som testades men som inte inkluderades i modellen var bland annat kön, familjesituation samt vissa diagnosspecifika variabler om tidigare öppen eller slutenvård^(2, tabell 12-15).

Fördelningen av antal personer inom respektive variabel i tabellen ovan var relativt lika under de två åren 2010 och 2011. Den största skillnaden mellan åren var att bland dem som hade ett sjukskrivningsfall i depressiv episod som uppnådde 21 dagar under år 2010 var det en mindre andel som inte hade några sjukskrivningsfall under året innan som översteg 14 dagar (70 %), jämfört med bland dem som hade ett liknande sjukskrivningsfall under år 2011 (73 %). Det fanns även en liten skillnad i andel som var anställda respektive arbetslösa när sjukskrivningsfallet startade de båda åren (anställda 2010: 81 %; 2011: 84 %).

För vissa grupper fanns det även mindre skillnader mellan åren i andel vars sjukskrivningsfall blev längre än 180 dagar. Bland de i 2010-kohorten som uppnått maxtid i sjukförsäkringen innan det studerade sjukskrivningsfallet startade var det 83 procent som blev långtidssjukskrivna, motsvarande andel i 2011-kohorten var 79 procent. Bland dem som hade

pågående eller tidigare (sedan 1994) sjuk- eller aktivitetsersättning var skillnaden mellan kohorterna den omvända, i 2010-kohorten blev 48 procent långtidssjukskrivna och i 2011-kohorten 53 procent.

6.1.1.1 Validering av modellen

Den optimala brytpunkten där andelen korrekt predicerade blir så stort som möjligt är för denna modell 0,47 (Tabell 3). Detta innebär att alla som av modellen får ett prediktionsvärde som är större än 0,47 anses ha en hög risk att bli långtidssjukskrivna (här definierat som >180 dagar). Andelen som enligt modellen anses ha en hög risk men som sedan inte blev långtidssjukskrivna är 3,5 procent, och andelen som anses ha en låg risk men som ändå blev långtidssjukskrivna är 24,7 procent. Om man vill minska andelen falskt negativa kan man välja en brytpunkt där denna andel blir lika stor som andelen falskt positiva, denna brytpunkt blir här 0,32. Appliceras dessa brytpunkter på prediktionsvärdena i valideringskohorten minskar totala andelen rättpredicerade och andel falskt negativa ökar.

Tabell 3. Andel korrekt predicerade vid val av olika brytpunkter, dels i träningskohorten 2010 och dels i valideringskohorten 2011.

Brytpunkt:	2010		2011	
	0,47	0,32	0,47	0,32
Andel rättpredicerade (%)	71,8	66,8	70,3	65,7
Andel falskt positiva (%)	3,5	16,5	3,3	15,6
Andel falskt negativa (%)	24,7	16,6	26,3	18,5

De övergripande måtten indikerar att det finns ytterligare faktorer som inte har inkluderats i dessa analyser som har samband med utfallet - att sjukskrivningsfallet blir längre än 180 dagar (Tabell 4). Brier värdet kan variera mellan 0 (en perfekt modell) och 0,223 för 2010-kohorten, respektive 0,224 för valideringskohorten baserad på år 2011. Här blir det viktade Brier värdet 12 procent, det viktade Brier värdet kan variera mellan 0 och 100 procent, där 100 procent indikerar perfekt modell.

Diskrimineringsmåtten visar hur bra modellen är på att skilja de som blev långtidssjukskrivna från de som inte blev det. För både AUC och diskrimineringslutningen gäller att 1,000 indikerar en perfekt modell, medan värdena 0,5 respektive 0 indikerar att modellen är oanvändbar.

Samtliga mått var något sämre för valideringsdatasetet, vilket är förväntat.

Kalibreringsmåtten, som visar hur bra prediktionsvärdet stämmer överens med det observerade utfallet för varje person, visar att prediktionsmodellen fungerar bra även för 2011 ("Calibration-in-the-large" nära 0 och "Calibration slope" nära 1).

Tabell 4. Olika mått för prestanda på modellen som predicerar risken att bli sjukskriven längre än 180 dagar bland de personer som 2010 alternativt 2011 hade ett nytt sjukskrivningsfall i depressiv episod (F32) som blev minst 21 dagar långt. Modellen är utvecklad baserad på 2010-kohorten (n=23 079) och validerad på 2011-kohorten (n=23 238).

	2010	2011
<i>Övergripande mått</i>		
R ² (Nagelkerke)	0,155	0,123
Brier värde	0,196	0,202
Viktat Brier värde	12,0 %	9,5 %
<i>Diskrimineringsmått</i>		
AUC (95% KI)	0,676 (0,668-0,683)	0,656 (0,649-0,664)
Diskrimineringslutning	0,120	0,099
<i>Kalibreringsmått</i>		
Hosmer-Lemeshow test χ^2 (p-värde)	12,981 (0,113)	26,391 (0,001)
"Calibration-in-the-large"	0	0,034
"Calibration slope"	1	0,958

6.1.2 Artros – sjukskrivningsfall som varat i minst 21 dagar

Totalt 10 851 personer i åldrarna 16-64 år hade minst ett nytt sjukskrivningsfall i artros som blev minst 21 dagar under år 2010. Några av dessa personer hade två eller fler sjukskrivningsfall som uppfyllde dessa kriterier, det första av dessa fall valdes inför analyserna. Av de sjukskrivningsfall från 2010 som analyserades blev 37,9 procent längre än 120 dagar. Totalt var 54,5 procent av personerna i kohorten kvinnor. År 2011 var det 11 089 personer som uppfyllde inklusionskriterierna. Bland dessa var 54,1 procent kvinnor, och för 37,1 procent blev sjukskrivningsfallet längre än 120 dagar.

Utfallet som prediceras i modellen var att fortfarande vara sjukskriven (i någon omfattning) efter 120 dagar. De som avled inom 120 dagar efter sjukskrivningsfallets start (6 personer år 2010 och 4 personer år 2011) kategoriserades som att ha varit sjukskrivna i mer än 120 dagar.

De 25 variabler som inkluderades i modellen (Tabell 5) kan delas in i tre kategorier:

- Sociodemografiska: Ålder, Utbildningsnivå, Region, Familjesituation, Sysselsättning
- Relaterade till sjukförsäkringen: Uppnådd maxtid, Tidigare eller pågående sjuk- eller aktivitetsersättning, Typ av artros, Nettodagar med sjukpenning året innan, Omfattning vid start, Tidigare sjukskrivning med samma diagnos eller annan muskuloskeletal diagnos.
- Relaterade till sjukvård eller medicinering: Diagnosspecifik slutenvård (fem variabler) och öppenvård (sex variabler) dels under de 365 dagarna innan sjukskrivningsfallet började och dels under de första 21 dagarna, Uttag av två olika typer av läkemedel under de 365 dagarna innan sjukskrivningsfallet började.

Hur dessa variabler operationaliserats framgår av Tabell 5. Trettioåtta procent (n=4 111) fick utfallet att sjukskrivningen blev mer än 120 dagar år 2010, denna andel var 37 procent år 2011 (n=4 109). Ju äldre personerna var, desto högre blev risken för långtidssjukskrivning. Att tidigare ha uppnått maxtid, tidigare haft sjuk- eller aktivitetsersättning samt tidigare varit sjukskriven gav högre risk, likaså att ha påbörjat sjukskrivningen på deltid. Att tidigare ha varit sjukskriven i artros eller andra muskuloskeletala diagnoser gav dock en lägre risk att det pågående fallet skulle bli längre än 120 dagar.

Jämfört med den modell som presenterades i vår tidigare rapport⁽²⁾ inkluderades här fler diagnosspecifika variabler gällande tidigare sjukskrivning, öppenvård samt slutenvård.

De variabler som testades men som inte inkluderades i modellen var bland annat kön, födelseland, samt variabler som indikerade tidigare sjukskrivningar eller vård på grund av psykiska sjukdomar^(2, tabell 23-27).

Tabell 5. Antal personer i varje grupp, bland de personer som hade ett nytt sjukskrivningsfall i artros (M15-M19), som år 2010 blev minst 21 dagar långt, och andel av dessa vars sjukskrivningsfall blev längre än 120 dagar samt betavärdet och oddsraten (OR) med 95 % konfidensintervall (KI) för de variabler som inkluderades i modellen som predicerar risken att bli sjukskriven längre än 120 dagar för dessa personer.

	Antal 2010	Andel (%) >120 dagar	Beta	OR (95 % KI)
Samtliga personer / Intercept (β_0)	10851	37,9%	-0,768	
Åldersgrupper ¹				
20-39 år	522	25,9%	-0,510	0,60 (0,48-0,76)
40-49 år	1835	35,6%	-0,072	0,93 (0,81-1,07)
50-54 år	1986	38,0%	0	1
55-59 år	3100	39,7%	0,093	1,10 (0,97-1,24)
60-64 år	3408	39,2%	0,109	1,12 (0,98-1,27)
Utbildningsnivå ¹				
Grundskola	2277	44,5%	0	1
Gymnasium	5815	40,0%	0,581	1,79 (1,58-2,03)
Eftergymnasial	2759	28,1%	0,501	1,65 (1,49-1,83)
Region ¹				
Nord	1577	36,7%	0	1
Mitt	1610	39,4%	0,158	1,17 (1,01-1,36)
Stockholm	1882	34,8%	0,039	1,04 (0,90-1,21)
Väst	3384	39,3%	0,168	1,18 (1,04-1,35)
Syd	2398	38,0%	0,116	1,12 (0,98-1,29)
Familjesituation ¹				
Gift/sambo, utan barn	3822	37,6%	0	1
Gift/sambo, med barn	3098	35,4%	0,063	1,07 (0,95-1,20)
Ensamstående, utan barn	3125	40,6%	0,140	1,15 (1,04-1,28)
Ensamstående, med barn	806	38,2%	0,129	1,14 (0,96-1,35)

Forts tabell 5

Syssetsättning vid sjukskrivningsfallets början

Anställd/Föräldraledig	9410	36,2%	0,	1
Arbetslös	800	51,0%	0,092	1,10 (0,93-1,30)
Egenföretagare	641	45,7%	0,274	1,32 (1,11-1,56)
Uppnått maxtid i sjukförsäkringen ³	443	86,0%	1,783	5,94 (4,38-8,07)
Pågående eller tidigare SA ^{3, 4}	1606	46,8%	0,199	1,22 (1,08-1,37)
Typ av artros				
M15 Polyartros	159	36,5%	-0,353	0,70 (0,49-1,01)
M16 Höftledsartros	3417	37,9%	0,158	1,17 (1,03-1,33)
M17 Knäartros	4247	38,1%	0,077	1,08 (0,96-1,22)
M18 Artros i första karpometakarpalleden	884	34,3%	-0,245	0,78 (0,66-0,94)
M19 Andra artros	2144	39,0%	0	1
Omfattning vid sjukskrivningsfallets start				
Deltid 25-75 %	1252	53,6%	0	1
Heltid 100 %	9599	35,8%	-0,489	0,61 (0,54-0,70)
Tidigare sjukskrivning, nettodagar ²				
Ingen tidigare sjukskrivning	7596	34,1%	0	1
0,25 - 49,75 Nettodagar	1682	39,2%	0,213	1,24 (1,08-1,42)
50 - 89,75 Nettodagar	547	38,9%	0,208	1,23 (1,00-1,52)
90 - 179,75 Nettodagar	576	57,3%	0,839	2,31 (1,87-2,86)
180 - 365 Nettodagar	450	71,3%	0,858	2,36 (1,80-3,09)
Diagnoser för tidigare sjukskrivning ²				
Artros (M15-M19) ³	1183	47,3%	-0,166	0,85 (0,71-1,00)
Andra musk. sjukdomar (M30-99) ³	691	45,7%	-0,198	0,82 (0,67-1,00)
Slutenvård året innan ²				
Artros (M15-M19) ³	1566	33,0%	-0,378	0,69 (0,60-0,78)
Andra musk. sjukdomar (M30-99) ³	102	33,3%	-0,481	0,62 (0,39-0,98)
Slutenvård 21 första dagarna				
Alla (exkl Z och O80) ³	4757	31,4%	-0,393	0,68 (0,60-0,76)
Andra ledsjukdomar (M00-14, M20-25) ³	61	39,3%	0,709	2,03 (1,15-3,60)
Övriga diagnoser (exkl musk, psyk, Z och O80) ³	245	43,7%	0,554	1,74 (1,32-2,29)
Öppenvård året innan ²				
Alla (exkl Z och O80) ³	8774	36,4%	-0,158	0,85 (0,74-0,98)
Artros (M15-M19) ³	6542	36,2%	0,190	1,21 (1,07-1,37)
Andra ledsjukdomar (M00-14, M20-25) ³	917	30,0%	-0,395	0,67 (0,57-0,80)
Andra musk. sjukdomar (M30-99) ³	1361	35,4%	-0,224	0,80 (0,70-0,92)
Öppenvård 21 första dagarna				
Artros (M15-M19) ³	2141	33,8%	-0,145	0,87 (0,77-0,97)
Andra ledsjukdomar (M00-14, M20-25) ³	204	21,1%	-0,682	0,51 (0,35-0,73)
Läkemedel året innan ²				
Smärtstillande ³	5781	41,6%	0,236	1,27 (1,16-1,38)
Preparat mot magsår ³	2096	42,8%	0,124	1,13 (1,02-1,26)

¹ Gällande december 2009.² Med tidigare avses under ett år (365 dagar) innan det studerade sjukskrivningsfallet påbörjades³ Referensgrupp är att **inte** uppnått maxtid, **inte** haft tidigare eller pågående SA, att **inte** haft någon tidigare sjukskrivning, inga dagar inom slutenvården, inga besök inom specialiserad öppenvård med respektive diagnos som huvuddiagnos samt att **inte** gjort uttag av detta läkemedel.⁴ Med tidigare avses från och med 1994 till och med dag 21 i sjukskrivningen.

Fördelningen av antal personer inom respektive variabel i ovanstående Tabell 5 var relativt lika under de två åren 2010 och 2011. En något större andel av dem som hade ett nytt sjukskrivningsfall som varat minst 21 dagar år 2011 hade haft tidigare sjukvård än de 2010. Bland dem som hade en sjukskrivning om minst 21 dagar år 2010 var 44 procent inlagda på sjukhus åtminstone någon gång under de första 21 dagarna i fallet, för 2011 var motsvarande andel något högre; 46 procent. År 2010 hade 81 procent haft minst ett öppenvårdsbesök under året innan sjukskrivningen och 60 procent hade haft minst ett besök med artros som diagnos, år 2011 var motsvarande andelar 84 respektive 64 procent.

För några, relativt små, grupper fanns det även vissa skillnader mellan åren i andel personer vars sjukskrivningsfall blev längre än 120 dagar. Bland annat var det år 2011 en större andel (48 %) som blev långtidssjukskrivna bland dem som hade polyartros (M15) som huvuddiagnos i sjukskrivningen jämfört med 2010 (36 %). Bland dem som var egenföretagare samt bland dem som tidigare uppnått maxtid i sjukförsäkringen var det en lägre andel som blev långtidssjukskrivna 2011 (40 % resp. 78 %) jämfört med 2010 (46 % resp. 86 %).

6.1.2.1 Validering av modellen

För denna modell är den optimala brytpunkten, där totala antalet korrekt predicerade personer blir så stort som möjligt 0,46 och andelen korrekt predicerade blir 67 procent i båda kohorterna (Tabell 6). Detta innebär att alla som av modellen får ett prediktionsvärde som är större än 0,46 anses ha en hög risk att bli långtidssjukskrivna (här definierat som >120 dagar). Andelen som enligt modellen anses ha en hög risk men som sedan inte blev långtidssjukskrivna (falskt positiva) är 8,7 procent, och andelen som anses ha en låg risk men som ändå blev långtidssjukskrivna (falskt negativa) är 24,3 procent. Om man vill minska andelen falskt negativa kan man välja en brytpunkt där denna andel blir lika stor som andelen falskt positiva, denna brytpunkt blir här 0,39.

Tabell 6. Andel korrekt predicerade vid val av olika brytpunkter dels i träningskohorten 2010 och dels i valideringskohorten 2011.

	2010		2011	
Brytpunkt:	0,46	0,39	0,46	0,39
Andel rättpredicerade (%)	67,0	64,7	66,6	63,9
Andel falskt positiva (%)	8,7	18,5	8,6	18,3
Andel falskt negativa (%)	24,3	16,8	24,7	17,8

De övergripande måtten indikerar att det finns ytterligare faktorer som inte har inkluderats i dessa analyser som har samband med utfallet att bli sjukskriven i mer än 120 dagar (Tabell 7). Brier värdet kan variera mellan 0 (en perfekt modell) och 0,235 för 2010-kohorten, respektive 0,233 för valideringskohorten baserad på år 2011. Här blir det viktade Brier värdet tio procent, det viktade Brier värdet kan variera mellan 0 och 100 procent, där 100 procent indikerar perfekt modell. Samtliga mått var något sämre för valideringsdatasetet, vilket är förväntat.

Kalibreringsmått, som visar hur bra prediktionsvärdet stämmer överens med det observerade utfallet för varje person, visar att prediktionsmodellen fungerar bra även för 2011 ("Calibration-in-the-large" nära 0 och "Calibration slope" nära 1).

Tabell 7. Olika mått för prestanda på modellen som predicerar risken att bli sjukskriven längre än 120 dagar bland de personer som hade ett nytt sjukskrivningsfall i artros (M15-M19) som blev minst 21 dagar långt. Modellen är utvecklad baserad på 2010-kohorten (n=10 851) och validerad på 2011-kohorten (n=11 089).

	2010	2011
<i>Övergripande mått</i>		
R ² (Nagelkerke)	0,135	0,116
Brier värde	0,211	0,213
Viktat Brier värde	10,1 %	8,7 %
<i>Diskrimineringsmått</i>		
AUC (95 % KI)	0,678 (0,667-0,688)	0,666 (0,655-0,676)
Diskrimineringslutning	0,101	0,091
<i>Kalibreringsmått</i>		
Hosmer-Lemeshow chi ² (p-värde)	10,875 (0,209)	7,978 (0,436)
"Calibration-in-the-large"	0	-0,027
"Calibration slope"	1	0,952

6.1.3 Utvärdering av exklusionskriterier

I våra tidigare, preliminära analyser⁽²⁾ exkluderades vissa grupper av personer ur kohorterna som användes för modelleringen av risken att bli långtidssjukskriven, då det redan fanns viss kunskap om att de hade hög risk för att det påbörjade sjukskrivningsfallet skulle bli långt. Att dessa grupper hade en hög risk framkom även i de initiala regressionsanalyserna som genomfördes innan exkluderingarna. Inklusionskriteriet var att ha haft ett nytt sjukskrivningsfall som under 2010 blev 21 dagar långt, den 21:a dagen efter den första dagen med sjukskrivning (startdatum 1 i MiDAS) inföll alltså under 2010, vilket innebär att några fall kunde ha påbörjats i december 2009. Kriterierna för att exkluderas i de tidigare analyserna var:

- De som vara äldre än 64 den 31 december 2009 (året innan inklusion) De som fyllde 65 år under uppföljningstiden var alltså med i kohorten.
- De som inte var folkbokförda i Sverige den 31 december 2009.
- De som avled under uppföljningstiden, det vill säga, inom 120 respektive 180 dagar efter sjukskrivningsfallets start.
- De som innan den studerade sjukskrivningen uppnått maximal tid inom sjukförsäkringen.
- De som någon gång under de föregående 365 dagarna hade haft sjuk- eller aktivitetsersättning (oavsett omfattning) samt de som när sjukskrivningsfallet startade hade partiell sjuk- eller aktivitetsersättning.

I de analyser vars resultat presenteras i denna rapport exkluderades endast de som uppfyllde de två första punkterna ovan. De få som avled behölls i analyserna, men deras utfall ändrades till att deras sjukskrivning blev längre än 120 respektive 180 dagar. De som tidigare exkluderats på grund av tidigare eller pågående sjuk- eller aktivitetsersättning eller uppnådd maximal tid i sjukförsäkringen fick även vara kvar i dessa analyser efter diskussion med SRS:s projektledning. Detta innebär att två nya prognosfaktorer inkluderades, nämligen att ha haft tidigare eller pågående sjuk- eller aktivitetsersättning samt att tidigare ha uppnått maximal tid inom sjukförsäkringen.

För att utvärdera om dessa två exkluderingskriterier var befogade eller inte studerades prediktionsvärdena specifikt för de personer som uppfyllde dessa två kriterier. Det visade sig att samtliga som uppnått maximal tid inom sjukförsäkringen fick ett prediktionsvärde över 0,50 och att medianen var över 0,80 i alla fyra kohorter för denna grupp (Tabell 8). Detta innebär att sannolikheten att sjukskrivningsfallet blir längre än 180 respektive än 120 dagar är över 50 procent för samtliga personer som tidigare uppnått maximal tid inom sjukförsäkringen, för hälften av dessa personer är sannolikheten över 80 procent.

Bland dem som hade tidigare eller pågående sjuk- eller aktivitetsersättning var det inte lika tydligt, men medianen av dessa personers prediktionsvärden var betydligt högre än medianen för andra personer.

Tabell 8. Deskriptiva analyser av de prediktionsvärden som genereras av modellerna ovan (Tabell 2 och Tabell 5), uppdelat på dem som uppfyller respektive inte uppfyller tidigare använda exklusionskriterier⁽²⁾, för de personer som haft ett sjukskrivningsfall i depressiv episod respektive i artros som respektive år blev minst 21 dagar långt. Använt utfall är att det fallet blev >180 dagar respektive >120 dagar långt.

	2010				2011			
	N	Långtids-sjukskrivna	Median	Min-Max	N	Långtids-sjukskrivna	Median	Min-Max
Depressiv episod								
Samtliga	23079	33,4 %	0,28	0,10 - 0,95	23238	33,8 %	0,28	0,09 - 0,96
Maxtid ¹	1648	83,1 %	0,85	0,56 - 0,95	1475	79,1 %	0,82	0,54 - 0,96
SA ²	1828	50,7 %	0,42	0,16 - 0,95	1757	56,5 %	0,54	0,14 - 0,95
Övriga ³	19655	29,0 %	0,27	0,10 - 0,78	20141	30,0 %	0,27	0,09 - 0,78
Artros								
Samtliga	10851	37,9 %	0,36	0,04 - 0,96	11089	37,1 %	0,35	0,06 - 0,96
Maxtid ¹	443	86,0 %	0,88	0,50 - 0,96	411	78,1 %	0,84	0,38 ⁴ - 0,96
SA ²	1386	47,3 %	0,44	0,09 - 0,95	1196	49,7 %	0,45	0,07 - 0,96
Övriga ³	8994	34,9 %	0,34	0,04 - 0,83	9471	34,5 %	0,34	0,06 - 0,82

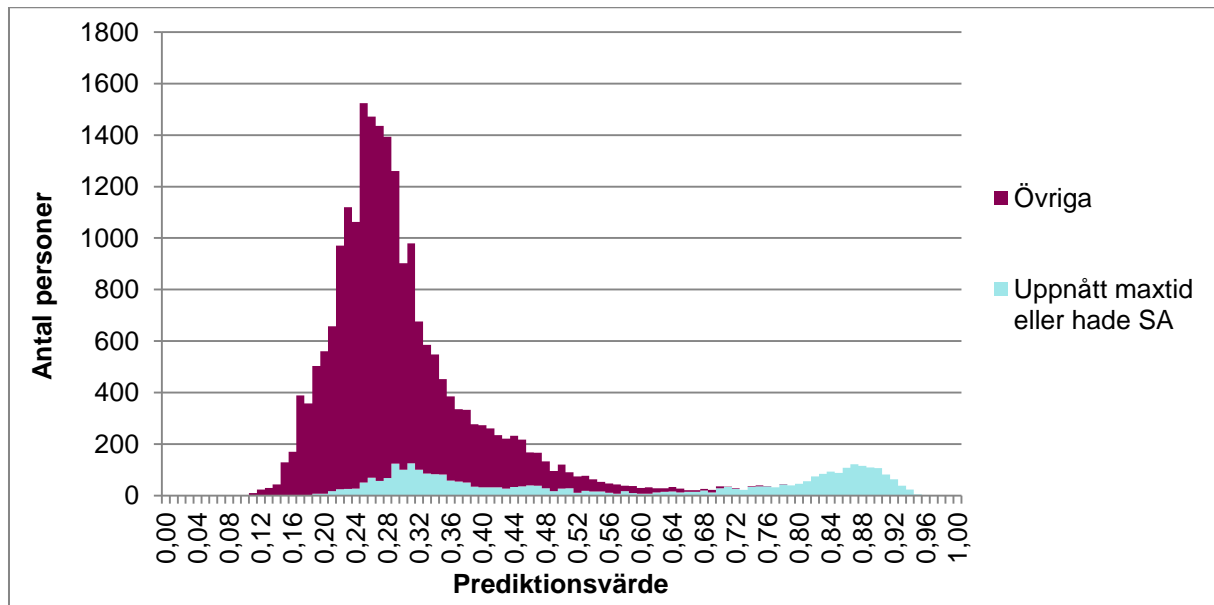
¹Har tidigare, sedan 2009, uppnått maximal tid i sjukförsäkringen

²Haft sjuk- eller aktivitetsersättning (SA) någon gång under de 365 dagarna innan sjukskrivningsfallet började alternativt pågående partiell SA när sjukskrivningsfallet började.

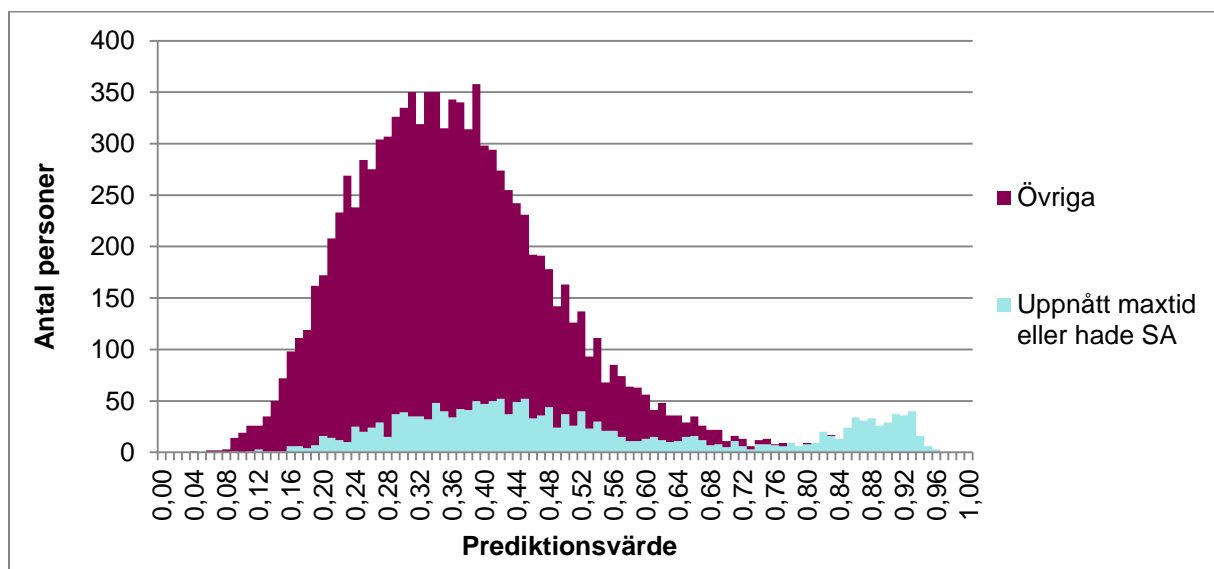
³De av de samtliga som inte haft tidigare eller pågående SA eller uppnått maxtid i sjukförsäkringen.

⁴En person hade detta (0,38) prediktionsvärde, övriga personer i denna grupp hade ett prediktionsvärde >0,50.

Detta kan också visas tydligt med hjälp av histogram (Figur 6 och Figur 7) där det framgår att prediktionsvärdet är betydligt högre för personer som antingen uppnått maximal tid i sjukförsäkringen eller hade tidigare eller pågående sjuk- eller aktivitetsersättning (SA) (de turkosa staplarna).



Figur 6. Fördelning av prediktionsvärden som genereras av modellen i avsnitt 6.1.1, för de personer som under 2010 hade ett nytt sjukskrivningsfall som blev minst 21 dagar långt i diagnosen depressiv episod, uppdelat på dem som uppfyllde respektive inte uppfyllde två av de exklusionskriterier som använts i tidigare analyser⁽²⁾ (tidigare ha uppnått maxtid i sjukförsäkringen alternativt ha eller ha haft sjuk- eller aktivitetsersättning (SA) senaste året).



Figur 7. Fördelning av prediktionsvärden som genereras av modellen i avsnitt 6.1.2, för de personer som under 2010 hade ett nytt sjukskrivningsfall som blev minst 21 dagar långt i en artrosdiagnos, uppdelat på de som uppfyllde respektive inte uppfyllde två av de exklusionskriterier som använts i tidigare analyser⁽²⁾ (tidigare ha uppnått maxtid i sjukförsäkringen alternativt ha eller ha haft sjuk- eller aktivitetsersättning (SA) senaste året).

6.1.4 Modeller baserade på olika grupper av prediktionsfaktorer

Några av de möjliga prediktionsfaktorer som är inkluderade i modellerna som presenteras ovan har handläggare alternativt analytiker på Försäkringskassan inte direkt tillgång till information om. De skulle behöva fråga den försäkrade eller läkaren om dessa faktorer. Därför har en analys gjorts för att särskilja hur mycket dessa faktorer bidrar till modellen. Dessa faktorer var utbildningsnivå, födelse-land samt alla faktorer skapade ur socialstyrelsens patientregister och läkemedelsregister. Det finns en skillnad i hur bra en modell med dessa faktorer (Modell A) predicerar jämfört med en modell utan dessa faktorer (Modell B) mätt med AUC, speciellt för artros (Tabell 9).

Vi fick även i uppdrag att testa några ytterligare faktorer, dessa var antal hemmavarande barn i olika åldrar samt två faktorer relaterade till yrke. Modell A inklusive dessa faktorer predicerar något bättre än modell A utan dessa faktorer (Tabell 9). Det är faktorerna om yrke; utan yrke, arbetare och tjänstemän, samt om sektor; ej anställd, anställd inom privata sektorn och anställd inom offentliga sektorn, som står för denna förbättring. Faktorerna om antal barn i olika åldrar bidrog inte till någon förbättring av modellen.

Utöver detta testades även en alternativ kategorisering av antal nettodagar med tidigare sjukskrivning där två år innan starten på det studerade sjukskrivningsfallet inkluderades. Även detta förbättrade modellerna något (Tabell 9). Modellerna vars AUC presenteras i Tabell 9 har ännu inte validerats, med till exempel hjälp av 2011-kohorterna.

Tabell 9. Jämförelser mellan modeller som inkluderar olika faktorer, för att predicera risken att bli långtidssjukskriven i depressiv episod respektive i artros bland personer sjukskrivna i minst 21 dagar i respektive diagnos under 2010; arean under ROC-kurvan (AUC) presenteras för samtliga modeller.

	Depressiv episod (F32)	Artros (M15-M19)
	AUC (95 % KI)	AUC (95 % KI)
Modell A ¹	0,676 (0,668-0,683)	0,678 (0,667-0,688)
Modell B ²	0,661 (0,653-0,668)	0,639 (0,628-0,650)
Modell A + barn ³ och yrke ⁴	0,683 (0,676-0,691)	0,691 (0,680-0,701)
Modell B + barn ³	0,664 (0,656-0,671)	0,640 (0,630-0,651)
Modell A 2 år SS ⁵	0,682 (0,672-0,692)	0,680 (0,673-0,688)
Modell A 2 år SS ⁵ + yrke ⁴	0,686 (0,678-0,693)	0,693 (0,683-0,703)

¹ Modeller som presenteras i avsnitt 6.1.1 (Ålder, Födelse-land, Utbildningsnivå, H-region, Sysselsättning, Uppnådd maxtid, Tidigare sjuk- eller aktivitetsersättning, Nettodagar med sjukpenning året innan, Omfattning vid start, Tidigare sjukskrivning med samma diagnos, Slutenvård under de första 21 dagarna av sjukskrivningsfallet, Öppenvård under de första 21 dagarna av fallet, Diagnosspecifik öppenvård (tre variabler) dels under de 365 dagarna innan sjukskrivningsfallet började och dels under de första 21 dagarna, Uttag av antidepressiva läkemedel dels under de 365 dagarna innan sjukskrivningsfallet började och dels under de 21 första dagarna) samt 6.1.2 (Ålder, Utbildningsnivå, Region, Familjesituation, Sysselsättning, Uppnådd maxtid, Tidigare eller pågående sjuk- eller aktivitetsersättning, Typ av artros, Nettodagar med sjukpenning året innan, Omfattning vid start, Tidigare sjukskrivning med samma diagnos eller annan muskuloskeletal diagnos, Diagnosspecifik slutenvård (fem variabler) och öppenvård (sex variabler) dels under de 365 dagarna innan sjukskrivningsfallet började och dels under de första 21 dagarna, Uttag av två olika typer av läkemedel under de 365 dagarna innan sjukskrivningsfallet började).

² Modell A exklusive variabler om utbildningsnivå, födelse-land samt alla variabler skapade ur Socialstyrelsens patientregister och läkemedelsregister.

⁴ Tre variabler om antal barn i åldrarna; 0-6 år, 7-15 år samt 16-20 år.

⁵ Två variabler om yrke; utan yrke, arbetare och tjänstemän, samt om sektor; ej anställd, anställd inom privata sektorn och anställd inom offentliga sektorn.

³ De variabler som handlar om tidigare sjukskrivning inkluderar två år (730 dagar) innan den studerade sjukskrivningen startade.

Resultat, del II.

6.2 Risk att blir sjukskriven vid specifika diagnoser

Denna del av resultaten handlar om personer som får en sjukdom eller skada och deras risk för sjukskrivning. Detta innebär en helt annan utgångspunkt än i analyserna i del I och i de tidigare rapporterna om prediktiva modeller i SRS projektet^(2, 33).

En central fråga i SRS projektet har varit att om de framtagna prediktionsmodellerna endast baseras på de personer som redan är sjukskrivna, kommer prediktionsmodellerna inte att innehålla information om hur vanligt det är att bli sjukskriven vid olika diagnoser eller vid samsjuklighet – och att resultaten därmed blir mindre intressanta för hälso- och sjukvården. I USA används till exempel system för rekommendationer om sjukskrivning, som helt baserats på sjukskrivningsdata där endast sjukskrivna personer inkluderats^(34, 35). Detta har ofta kritiserats och Sverige valde efter mycket överväganden att inte anamma detta vid utformandet av det försäkringsmedicinska beslutstödet^(14, 36). Många personer i arbetsföra åldrar som får en sjukdom eller skada blir inte sjukskrivna på grund av den, medan andra personer redan kan vara sjukskrivna eller ha sjuk- eller aktivitetsersättning i en annan diagnos när sjukdomen eller skadan uppstår. I vilken omfattning detta är fallet saknas det dock i stor utsträckning kunskap om.

Av denna anledning fick vi även i uppdrag att göra två typer av analyser för personer med en viss diagnos,

- dels en deskription av personernas sjukfrånvaro och sjuk- eller aktivitetsersättning under åren före och efter diagnosen,
- dels ta fram en prediktionsmodell för personernas risk att bli sjukskrivna med just den diagnosen.

I diskussion med SRS-projektledning föll valet då på två av alla möjliga olika typer av diagnoser, nämligen handlovsfraktur, och bröstcancerdiagnos – det vill säga, en skade- och en sjukdomsdiagnos. Orsaken till valet av handlovsfrakturer var att i ett annat delprojekt inom SRS-projektet hade analyser av sjukskrivningsfall i handlovsfraktur som överstigit tre veckor med fokus på faktorer som predicerade att fallet skulle bli längre än 180 dagar⁽³³⁾. Frågan uppstod då om hur stor del av dem med handlovsfraktur som faktiskt blev sjukskrivna i denna diagnos. Orsakerna till val av bröstcancer som den andra diagnosen var dels att personer med denna diagnos kan identifieras via Socialstyrelsens cancerregister, dels att projektgruppen tidigare gjort vissa studier av kvinnor med bröstcancer, vilket innebär att det, inom den mycket korta tidsramen för detta delprojekt, var möjligt att använda kunskaper från de analyserna.

Nedan redovisas först resultat från de deskriptiva analyserna av förekomst av sjukfrånvaro respektive sjuk- och aktivitetsersättning generellt och i olika diagnosgrupper bland personer

som drabbats av en handlovsfraktur respektive av bröstcancer. Sedan följer ett avsnitt där modeller som predicerar risken att bli sjukskriven för personer i dessa två grupper.

6.2.1 Handlovsfrakturer

Under år 2010 fick 18 678 personer som var 16-64 år gamla och folkbokförda i Sverige i december 2009 vård på sjukhus (specialiserad öppenvård eller slutenvård) på grund av en handlovsfraktur. Efter exkludering av dem som även hade allvarligare skador vid det vårdtillfället samt av dem som hade fått vård för handlovsfraktur under de två åren (730 dagar) före datum för det första sådana vårdtillfället under år 2010 bestod kohorten av 16 337 personer. Datum för handlovsfrakturen, här definierat som det första vårdtillfället för handlovsfraktur under år 2010, kallas här för T_0 . Fördelning av bakgrundsfaktorer för dessa personer, såsom ålder, kön med mera, framgår i Tabell 11 nedan.

6.2.1.1 Sjukskrivning och sjuk- eller aktivitetsersättning två år innan och tre år efter handlovsfrakturen, beskrivande analyser

Varje persons antal nettodagar med sjukskrivning samt antal nettodagar med sjuk- och aktivitetsersättning har beräknats för fem olika tidsperioder, dels för året efter handlovsfrakturen, det vill säga de kommande 365 dagarna från T_0 till T_{+1} som är tidpunkten för den 365:e dagen efter vårdtillfället T_0 , dels för år två (T_{+1} till T_{+2}) respektive år tre (T_{+2} till T_{+3}) efter T_0 , samt för de två åren innan vårdtillfället T_0 (T_{-2} till T_{-1} respektive T_{-1} till T_0). De första 14 dagarna i varje sjukskrivningsfall exkluderades i dessa beräkningar.

Under året efter handlovsfrakturen är endast en tredjedel av alla personer i kohorten sjukskrivna i någon diagnos, och 23 procent är någon gång under året sjukskrivna i diagnos S62, handlovsfraktur (Tabell 10). Andelen med sjukskrivning i psykiska diagnoser förändras inte över tid (2-3 %). Under de två åren före handlovsfrakturen samt under år 2 och 3 efter handlovsfrakturen är andelen med minst någon dag med sjukskrivning (första 14 dagarna i varje sjukskrivningsfall exkluderade) 9 till 12 procent. Även andelen med sjuk- eller aktivitetsersättning är i princip oförändrad över dessa fem år (7-8 %). Andelen som varken hade sjukskrivning eller sjuk- eller aktivitetsersättning under året efter handlovsfrakturen var 61 procent. Övriga år varierade denna andel mellan 83 till 85 procent.

Tabell 10. Antal och andel personer med sjukskrivning (SS) respektive sjuk- eller aktivitetsersättning (SA), för samtliga diagnoser respektive för olika diagnosgrupper, uppdelat på år relativt datum för handlovsfrakturen (T₀) bland de 16 337 personer som år 2010 var 16-65 år gamla och fick en handlovsfraktur (ICD 10 kod: S62).

SS/SA/ Diagnos	Nettodagar per person	T ₋₂ -T ₋₁ n (%)	T ₋₁ -T ₀ n (%)	T ₀ -T ₊₁ n (%)	T ₊₁ -T ₊₂ n (%)	T ₊₂ -T ₊₃ n (%)
Totala antalet personer		16187 (100)	16337 (100)	16337 (100)	16056 (100)	15746 (100)
SS ¹						
Samtliga	0	14783 (91,3)	14872 (91,0)	11015 (67,4)	14306 (89,1)	13919 (88,4)
diagnoser	>0-30	629 (3,9)	750 (4,6)	2669 (16,3)	763 (4,8)	769 (4,9)
	>30-90	370 (2,3)	358 (2,2)	1887 (11,6)	478 (3,0)	503 (3,2)
	>90-180	197 (1,2)	172 (1,1)	479 (2,9)	246 (1,5)	276 (1,8)
	>180	208 (1,3)	185 (1,1)	287 (1,8)	263 (1,6)	279 (1,8)
S62	0	16178 (99,9)	16201 (99,2)	12652 (77,4)	15955 (99,4)	15671 (99,5)
	>0-30	≤8	121 (0,7)	2153 (13,2)	50 (0,3)	34 (0,2)
	>30-90	≤8	13 (0,1)	1263 (7,7)	31 (0,2)	28 (0,2)
	>90-180	0	≤8	211 (1,3)	≤8	≤8
	>180	0	0	58 (0,4)	15 (0,1)	≤8
Psykiska	0	15851 (97,9)	15998 (97,9)	15997 (97,9)	15674 (97,6)	15280 (97,0)
	>0-30	119 (0,7)	121 (0,7)	95 (0,6)	127 (0,8)	136 (0,9)
	>30-90	80 (0,5)	85 (0,5)	104 (0,6)	81 (0,5)	133 (0,8)
	>90-180	47 (0,3)	58 (0,4)	62 (0,4)	75 (0,5)	82 (0,5)
	>180	90 (0,6)	75 (0,5)	79 (0,5)	99 (0,6)	115 (0,7)
Övriga ¹	0	15104 (93,3)	15299 (93,6)	14607 (89,4)	14738 (91,8)	14397 (91,4)
	>0-30	526 (3,2)	549 (3,4)	831 (5,1)	638 (4,0)	648 (4,1)
	>30-90	293 (1,8)	271 (1,7)	585 (3,6)	368 (2,3)	367 (2,3)
	>90-180	148 (0,9)	112 (0,7)	186 (1,1)	166 (1,0)	179 (1,1)
	>180	116 (0,7)	106 (0,6)	128 (0,8)	146 (0,9)	155 (1,0)
SA						
Samtliga	>0	1217 (7,5)	1235 (7,6)	1181 (7,2)	1105 (6,9)	1063 (6,8)
S62	>0	≤8	≤8	≤8	≤8	≤8
Psykiska	>0	440 (2,7)	444 (2,7)	430 (2,6)	412 (2,6)	417 (2,6)
Övriga ²	>0	787 (4,9)	799 (4,9)	762 (4,7)	699 (4,4)	655 (4,2)
Ej inkluderade ³		150	0	0	281	591
Blev >65 år					215 ⁴	434 ⁵
Avled ≤65 år					40	86
Bodde ej i Sverige		150			26	71

¹ De 14 första dagarna i samtliga sjukskrivningsfall är ej medräknade.

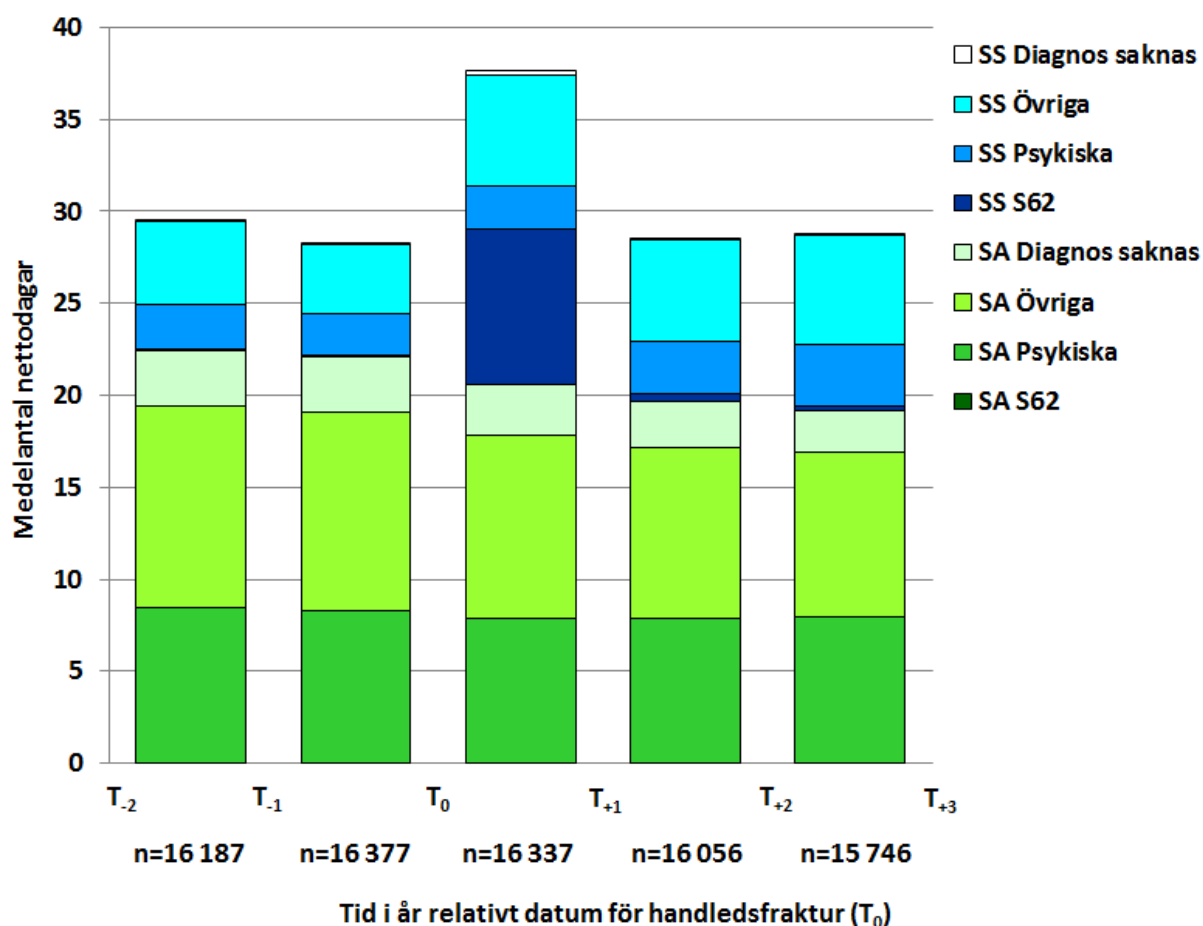
² I diagnosgruppen Övriga ingår även de SS och SA fall där information om diagnos saknas i MiDAS.

³ Personer som fyller 65, avlider eller emigrerar inkluderas till och med det år för när detta inträffar. Personer som inte bodde i Sverige innan datum för handlovsfrakturen är ej inkluderade motsvarande år.

⁴ Ålder över 65 år definieras efter vilket kalenderår som personerna blir över 65 år, då den information vi använt för ålder baseras på LISA som gäller för 31 december 2009. Dessa 215 personer blev alltså 66 år under år 2011.

⁵ Dessa 434 personer blev 66 år under år 2012 (se föregående fotnot).

Därefter har medelantalet nettodagar för samtliga personer i kohorten beräknats för varje tidsperiod uppdelat på sjukskrivning (SS) respektive sjuk- och aktivitetsersättning (SA) i olika diagnosgrupper (Figur 8). Medelantalet nettodagar uppgick till 37,6 dagar under året efter vårdtillfället (17,1 dagar med SS och 20,6 dagar med SA). För de två åren före T_0 samt för de två åren efter T_{+1} var medelantalet något färre än 30 nettodagar. Majoriteten av dessa dagar var med sjuk- eller aktivitetsersättning (SA).



Figur 8. Medelantal nettodagar uppdelat på sjukskrivning (SS) respektive sjuk- eller aktivitetsersättning (SA), olika diagnosgrupper och år relativt datum för handledsfraktur (T_0) bland de 16 337 personer som fick detta år 2010 och då var 16-65 år gamla. Antal inkluderade personer skiljer sig något mellan åren till följd av att personer blir >65 år, avlider samt på grund av in- och utflyttning till Sverige.

6.2.1.2 Sjukskrivning och sjuk- eller aktivitetsersättning i samband med handledsfrakturen

Inför modellering av risken att bli sjukskriven i samband med en handledsfraktur exkluderades följande två grupper av personer som inte var under risk för att påbörja ett nytt sjukskrivningsfall:

- dels de personer som redan hade ett pågående sjukskrivningsfall (1 023 personer) vid T_0 , här definierat som att det skulle ha påbörjats minst två dagar innan T_0 .
- dels de personer som vid T_0 hade pågående sjuk- eller aktivitetsersättning om minst

75 procent omfattning (851 personer, varav 23 personer hade 75 %). Två av dessa personer som hade sjuk- eller aktivitetsersättning om 75 procent var även sjukskrivna vid T₀.

Således bestod den kohort som användes till utvecklingen av en modell för risken att bli sjukskriven, av 14 465 personer.

Utfallet i modellen som predicerar risken att bli sjukskriven i samband med en handlovsfraktur var att få en ny sjukskrivning med handlovsfraktur (S62) som diagnos, som varade minst 14 dagar och som startade någon gång från och med dagen innan vårdtillfället för handlovsfrakturen till och med den sjätte dagen efter vårdtillfället. Totalt fick 20,5 procent (2 959 personer) av samtliga i kohorten ett nytt sådant sjukskrivningsfall. Denna till synes låga andel skulle delvis kunna vara relaterad till att en tredjedel av kohorten var yngre än 25 år. Många unga vuxna har ännu inte tillräckligt hög sjukpenninggrundande inkomst (SGI) för att omfattas av sjukpenningförsäkringen och kan därför inte få ersättning från Försäkringskassan även om deras arbetsförmåga är nedsatt mer än 25 procent under mer än 14 dagar. Andelen som får utfallet bland dem som fyllt 25 år var 24 procent.

Om utfallet skulle utökas till ett nytt sjukskrivningsfall under samma tidsperiod i förhållande till T₀ men oavsett sjukskrivningsfallets längd (det vill säga, om även sjukskrivningsfall kortare än 14 dagar och ersatta av Försäkringskassan inkluderades) eller sjukskrivningsdiagnos (det vill säga, om även sjukskrivning i andra diagnoser än handlovsfraktur inkluderades), så ökar andelen personer som fick utfallet till 25 procent (3 607 personer).

Enligt MiDAS registret fick cirka 4 500 personer som var 16-64 år en ny sjukskrivning med handlovsfraktur (S62) som diagnos under 2010. Denna diskrepans på ca 1 500 personer kan ha flera förklaringar. I denna studie exkluderades de personer som samtidigt vårdades för andra allvarigare skador samt de som hade vårdats för handlovsfraktur under de två föregående åren, även dem med pågående sjukskrivning exkluderades, några av dessa skulle kunna ha haft nya sjukskrivningsfall som startade under 2010 med handlovsfraktur som diagnos. Vi har inte heller tagit med de sjukskrivningsfall i handlovsfraktur som påbörjades mer än en vecka efter frakturen. Det kan även finnas personer som blir sjukskrivna som inte vårdats på sjukhus eller i specialiserad öppenvård och därmed inte registreras i patientregistret.

6.2.1.3 Bakgrundsfaktorer för personer med en handlovsfraktur under 2010

I hela kohorten med personer som fick vård för en handlovsfraktur under 2010 (16 337 personer) var det 70 procent män och en knapp majoritet var under 35 år (Tabell 11). Ungefär samma köns- och åldersfördelning fanns i kohorten som användes till modelleringen, men där var det en något större andel män och yngre personer. Att en stor del av personerna i kohorten var unga personer, har samband med relativt stora andelar som var ensamstående utan barn (41 %), hemmaboende < 20 år (19 %) och/eller ej förvärvsarbetande (36 %).

Tabell 11. Socioekonomiska och sociodemografiska bakgrundsfaktorer för alla personer i kohorten med en handlovsfraktur (S62) under år 2010, samt för den kohorten från 2011 som användes för modelleringen.

	<i>Hela kohorten</i>		<i>Kohorten som användes för modellering</i>	
	<i>Antal</i>	<i>Andel (%)</i>	<i>Antal</i>	<i>Andel (%)</i>
Alla	16337	100,0	14465	100,0
Kön				
Män	11433	70,0	10353	71,6
Kvinnor	4904	30,0	4112	28,4
Åldersgrupper ¹				
16-24 år	5289	32,4	5055	34,9
25-34 år	3148	19,3	2872	19,9
35-44 år	2932	17,9	2567	17,7
45-54 år	2605	15,9	2156	14,9
55-64 år	2363	14,5	1815	12,5
Födelseland				
Sverige	14016	85,8	12446	86,0
Övriga	2321	14,2	2019	14,0
Utbildningsnivå ¹				
Grundskola	4932	30,2	4382	30,3
Gymnasium	7869	48,2	6824	47,2
Eftergymnasial	3536	21,6	3259	22,5
Boenderegion ¹				
Nord	2037	12,5	1798	12,4
Mitt	2271	13,9	2018	14,0
Stockholm	3485	21,3	3129	21,6
Väst	4622	28,3	4010	27,7
Syd	3922	24,0	3510	24,3
H-region ¹				
Storstad	5918	36,2	5272	36,4
Mellanstor stad	5804	35,5	5161	35,7
Landsbygd	4615	28,2	4032	27,9
Familjesituation ¹				
Gift/sambo, utan barn	1551	9,5	1288	8,9
Gift/sambo, med barn	4238	25,9	3786	26,2
Ensamstående, utan barn	6739	41,2	5777	39,9
Ensamstående, med barn	766	4,7	633	4,4
Hemmaboende <20 år	3043	18,6	2981	20,6
Civilstånd ¹				
Ogift, skild, änka/änkling	11646	71,3	10356	71,6
Gift, registrerat partnerskap	4691	28,7	4109	28,4
Typ av yrke ¹				
Inget yrke	4752	29,1	4132	28,6
Tjänsteman	3592	22,0	3346	23,1
Arbetare	7993	48,9	6987	48,3
Sektor ¹				
Ej förvärvsarbetande	5814	35,6	4881	33,7
Offentligt anställd	2280	14,0	2033	14,1
Privat anställd	8243	50,5	7551	52,2

¹ Gällande december 2009

Sex procent av de i kohorten hade handlovsfraktur som bidiagnos och en annan diagnos som huvuddiagnos vid vårdbesöket (Tabell 12), vanligast var andra typer av skador, speciellt sårskada på handlov eller hand (S61). Knappt fyra procent hade slutenvård, det vill säga, blev inlagda på sjukhus i samband med handlovsfrakturen, nästan hälften av dessa hade endast slutenvård, det vill säga, hade inga öppenvårdsbesök under dagen för T₀. Ett fåtal personer var redan inlagda på grund av en annan sjukdom eller skada när de fick sin handlovsfraktur.

Lite drygt sju procent hade vårdats för en tidigare handlovsfraktur, det vill säga, för mer än två år sedan. De som fått vård för handlovsfraktur inom två år innan det första vårdtillfället med en handlovsfraktur under år 2010 exkluderades, detta på grund av att öka säkerheten att handlovsfrakturen skedde i samband med detta vårdtillfälle och att vården inte gällde uppföljande vård från en tidigare handlovsfraktur. Det kan dock vara så att en del av dessa personer inte hade en ny handlovsfraktur utan att detta vårdtillfälle är en uppföljning av en fraktur som skedde för mer än två år sedan och personen har inte haft denna typ av vård för handlovsfraktur under de två föregående åren.

Tabell 12. Faktorer relaterade till vårdtillfället för alla personer i kohorten med en handlovsfraktur (S62) under år 2010 (n=16 337), samt för den kohorten som användes för modelleringen (n=14 465).

	Hela kohorten		Kohorten som användes för modellering	
	Antal	Andel (%)	Antal	Andel (%)
S62 som huvuddiagnos	15329	93,8	13623	94,2
Slutenvårdad vid vårdtillfället	635	3,9	528	3,7
Tidigare handlovsfraktur >2 år sedan	1229	7,5	1071	7,4
Antal operationskoder ¹				
0	10464	64,1	9309	64,4
1	3230	19,8	2842	19,6
2	1626	10,0	1412	9,8
3	653	4,0	570	3,9
>3	364	2,2	332	2,3

¹ Maximala antalet operationskoder i samband med handlovsfrakturen var 30 stycken (en person)

Fjorton procent av hela kohorten hade minst ett sjukskrivningsfall som översteg 14 dagar under de två åren innan handlovsfrakturen, denna andel minskade ner till elva procent efter exklusion av dem som redan var sjukskrivna alternativt hade sjuk- eller aktivitetsersättning, det vill säga, i kohorten som användes för modelleringen (Tabell 13). På samma sätt minskade andelen med mer än 365 nettodagar med sjuk- eller aktivitetsersättning under de två föregående åren, men andelen med 365 eller färre nettodagar med sjuk- eller aktivitetsersättning består, detta på grund av att dessa personer troligtvis har sjuk- eller aktivitetsersättning på deltid (50 % eller 25 %) och därmed var under risk att bli sjukskrivna.

En majoritet av hela kohorten hade minst ett besök inom specialiserad öppenvård under de två föregående åren och knappt 15 procent hade minst en dag inom slutenvård inom samma period. Dessa andelar var mindre i kohorten som användes för modellering av risk att bli sjukskriven.

Tabell 13. Tidigare sjukskrivning och sjuk- eller aktivitetsersättning samt tidigare öppen- och slutenvård för alla personer i kohorten med en handlovsfraktur (S62) under år 2010 (n=16 337), samt för den kohorten som användes för modelleringen (n=14 465). Med tidigare avses under två år (730 dagar) innan handlovsfrakturen 2010; T₀.

	Hela kohorten		Kohorten som användes för modellering	
	Antal	Andel (%)	Antal	Andel (%)
Tidigare sjukskrivning ¹ , nettodagar ²				
Ingen tidigare sjukskrivning	14082	86,2	12826	88,7
0,25-90 Nettodagar	1626	10,0	1257	8,7
Mer än 90 Nettodagar	629	3,9	382	2,6
Tidigare sjukskrivning ¹ , diagnoser				
Psyksiska diagnoser	524	3,2	359	2,5
Övriga diagnoser	1879	11,5	1342	9,3
Tidigare sjuk- eller aktivitetsersättning ¹ , nettodagar				
Ingen tidigare sjuk- eller aktivitetsersättning	15066	92,2	14088	97,4
0,25-365 Nettodagar	364	2,2	292	2,0
Mer än 365 Nettodagar	907	5,6	85	0,6
Tidigare sjuk- eller aktivitetsersättning ¹ , diagnoser				
Psyksiska diagnoser	470	2,9	109	0,8
Övriga diagnoser	821	5,0	274	1,9
Tidigare besök i öppenvården ¹				
Inga besök	6805	41,7	6456	44,6
1-2 besök	5159	31,6	4595	31,8
3 eller fler besök	4373	26,8	3414	23,6
Tidigare besök i öppenvården ¹ , diagnoser				
Psyksiska diagnoser	1235	7,6	840	5,8
Övriga diagnoser	9272	56,8	7818	54,0
Tidigare slutenvård ¹				
Ingen slutenvård	13994	85,7	12725	88,0
En dag	1520	9,3	1239	8,6
Två eller fler dagar	823	5,0	501	3,5
Tidigare slutenvård ¹ , diagnoser				
Psyksiska diagnoser	504	3,1	317	2,2
Övriga diagnoser	2067	12,7	1542	10,7

¹ Med tidigare avses två år innan datum för handlovsfrakturen

² De 14 första dagarna i samtliga sjukskrivningsfall är ej medräknade

6.2.2 Bröstcancer

I analyser av personer nydiagnostiserade med bröstcancer valde vi att endast inkludera kvinnor med denna diagnos, det vill säga, inga män. Den absoluta majoriteten av personer som får en bröstcancerdiagnos är kvinnor, (av dem i arbetsför ålder som fick en första bröstcancerdiagnos år 2010 var 99,6 % kvinnor). Under år 2010 var det 3 538 kvinnor upp till 64 år som fick en första bröstcancerdiagnos i Sverige (efter att en person exkluderats på grund av att den inte fanns registrerad i LISA någon gång under 2007-2013). I våra analyser exkluderades sedan ytterligare 2 kvinnor där grunden för diagnos befanns vara obduktion, det vill säga, efter dödsfall (vilket innebär att en sjukskrivning för dessa kvinnor troligen inte kunde anses vara relaterad till bröstcancer eller vård för bröstcancer, då en diagnos inte fanns tidigare). Den slutliga kohorten för deskriptiva analyser bestod därmed av 3 536 kvinnor. Fördelning av sociodemografiska faktorer för dessa kvinnor framgår i Tabell 16 nedan.

6.2.2.1 Sjukskrivning och sjuk- eller aktivitetsersättning två år innan och tre år efter bröstcancerdiagnosen, beskrivande analyser

Antal nettodagar med sjukskrivning och sjuk- eller aktivitetsersättning beräknades per person och år relativt datum för bröstcancerdiagnos, här kallad T_0 , dels för samtliga diagnoser men även uppdelat på cancer (ICD-koder: C00-D48, Z80, Z85, N61-N63), psykiska (F00-F99, Z73) respektive övriga diagnoser. I dessa analyser exkluderades de 14 första dagarna i samtliga sjukskrivningsfall. Notera att det är olika antal kvinnor varje år i beräkningarna, detta är en följd av att personer blir >65 år, avlider, samt på grund av in- och utflyttning till Sverige.

Under året efter bröstcancerdiagnosen har 28 procent av kvinnorna inget sjukskrivningsfall som överstiger 14 dagar, därmed har de inga nettodagar med sjukskrivning (Tabell 14).

Andelen kvinnor med en sjukskrivning i cancer var 67 procent under året efter T_0 , knappt hälften av dessa (32 %) var sjukskrivna i cancer mer än halva året. Andra året (T_{+1} - T_{+2}) efter T_0 hade 35 procent av de då inkluderade kvinnorna minst ett sjukskrivningsfall, oavsett diagnos, som översteg 14 dagar, för sjukskrivningsfall i cancer var denna andel 25 procent. Tredje året efter diagnos (T_{+2} - T_{+3}) var motsvarande andelar 25 procent respektive 12 procent av kvinnorna; andelen kvinnor sjukskrivna i cancer minskade alltså avsevärt under åren efter diagnos.

Andelen kvinnor som hade sjukskrivningsfall >14 dagar i psykisk diagnos varierade inte mycket mellan de studerade åren (mellan 3-4 %). För sjukskrivning i övriga diagnoser var motsvarande andel mellan åtta till elva procent för alla år utom för året efter bröstcancerdiagnosen (T_0 - T_1), då den var sex procent.

Andelen kvinnor med sjuk- eller aktivitetsersättning varierade mellan 15 till 18 procent, andelen minskade något efter T_{+1} . Under året efter bröstcancerdiagnosen var det 15 procent som varken hade sjukskrivning eller sjuk- eller aktivitetsersättning. De två åren föregående diagnosen var denna andel 73 procent för vart och ett av åren. Under andra och tredje året efter diagnosen var motsvarande andel 52 respektive 62 procent.

Tabell 14. Antal och andel personer med sjukskrivning (SS) och sjuk- eller aktivitetsersättning (SA), för samtliga diagnoser respektive uppdelat på diagnos, per år utgående från datum för bröstcancerdiagnos (T₀) för samtliga kvinnor med denna diagnos under 2010 och som då var upp till 64 år gamla (N=3536), under två år innan och upp till tre år efter.

SS/SA/ Diagnos	Nettodagar per person	T ₋₂ -T ₋₁ n(%)	T ₋₁ -T ₀ n(%)	T ₀ -T ₊₁ n(%)	T ₊₁ -T ₊₂ n(%)	T ₊₂ -T ₊₃ n(%)
		3522 (100)	3534 (100)	3536 (100)	3492 (100)	3191 (100)
SS ¹						
Samtliga	0	3115 (88,4)	3134 (88,7)	978 (27,7)	2282 (65,3)	2401 (75,2)
	>0-30	198 (5,6)	206 (5,8)	485 (13,7)	422 (12,1)	324 (10,2)
	>30-90	98 (2,8)	105 (3,0)	440 (12,4)	297 (8,5)	165 (5,2)
	>90-180	55 (1,6)	50 (1,4)	366 (10,4)	213 (6,1)	133 (4,2)
	>180	56 (1,6)	39 (1,1)	1267 (35,8)	278 (8,0)	168 (5,3)
Cancer ²	0	3509 (99,6)	3506 (99,2)	1165 (32,9)	2610 (74,7)	2801 (87,8)
	>0-30	≤8	19 (0,5)	486 (13,7)	306 (8,8)	124 (3,9)
	>30-90	≤8	≤8	408 (11,5)	224 (6,4)	89 (2,8)
	>90-180	≤8	≤8	337 (9,5)	151 (4,3)	75 (2,4)
	>180	≤8	≤8	1140 (32,2)	201 (5,8)	102 (3,2)
Psykiska ³	0	3423 (97,2)	3435 (97,2)	3430 (97,0)	3354 (96,0)	3074 (96,3)
	>0-30	42 (1,2)	48 (1,4)	25 (0,7)	44 (1,3)	50 (1,6)
	>30-90	22 (0,6)	30 (0,8)	22 (0,6)	43 (1,2)	21 (0,7)
	>90-180	16 (0,5)	11 (0,3)	15 (0,4)	19 (0,5)	21 (0,7)
	>180	19 (0,5)	10 (0,3)	44 (1,2)	32 (0,9)	25 (0,8)
Övriga ⁴	0	3219 (91,4)	3246 (91,9)	3331 (94,2)	3208 (91,9)	2848 (89,3)
	>0-30	158 (4,5)	153 (4,3)	85 (2,4)	151 (4,3)	205 (6,4)
	>30-90	74 (2,1)	73 (2,1)	29 (0,8)	62 (1,8)	63 (2,0)
	>90-180	37 (1,1)	36 (1,0)	16 (0,5)	30 (0,9)	35 (1,1)
	>180	34 (1,0)	26 (0,7)	75 (2,1)	41 (1,2)	40 (1,3)
SA						
Samtliga	>0	613 (17,4)	619 (17,5)	606 (17,1)	553 (15,8)	492 (15,4)
Cancer	>0	≤8	≤8	12 (0,3)	13 (0,4)	21 (0,7)
Psykiska	>0	171 (4,9)	170 (4,8)	160 (4,5)	148 (4,2)	135 (4,2)
Övriga ²	>0	445 (12,6)	449 (12,7)	436 (12,3)	393 (11,3)	338 (10,6)
Ej inkluderade ⁵		14	2	0	44	345
Anledning						
Över 65 år						248 ⁶
Avlidna ≤ 65 år					43	91
Bor ej i Sverige ≤ 65 år		14	2		1	6

¹ De 14 första dagarna är borttagna i samtliga sjukskrivningsfall.

² ICD-koder: C00-D48, Z80, Z85, N61-N63

³ ICD-koder: F00-F99, Z73

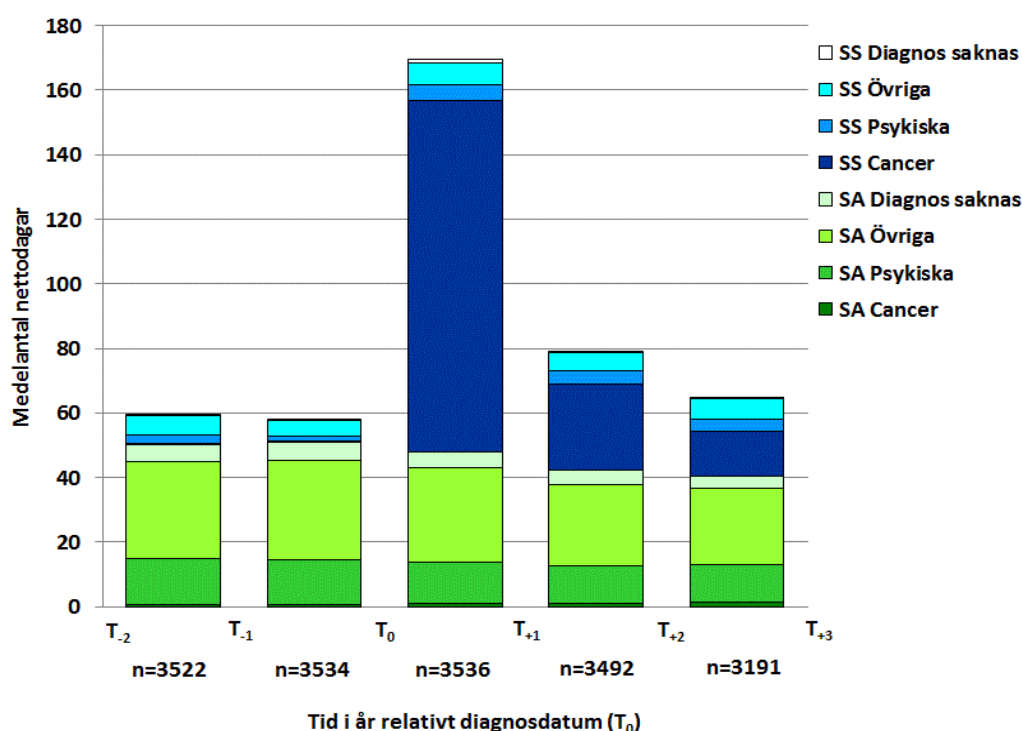
⁴ I diagnosgruppen Övriga ingår även de SS och SA fall där information om diagnos saknas i MiDAS.

⁵ Personer som fyllde 65, avled eller emigrerade inkluderas till och med det år för när detta inträffar. Personer som inte bodde i Sverige innan diagnosdatum är ej inkluderade motsvarande år.

⁶ Ålder över 65 år definieras efter vilket kalenderår som personerna blir över 65 år, då den information vi använt för ålder baseras på LISA som gäller för 31 december 2009. Dessa 248 kvinnor blev alltså 66 år under år 2012.

Under året efter bröstcancerdiagnos (T_0 - T_{+1}) var medelantalet nettodagar med sjukskrivning (SS) 121,3 dagar. Detta var betydligt högre än åren innan diagnos då medelantalet dagar var 6,7 (T_{-1} - T_0) respektive 9,0 (T_{-2} - T_{-1}) dagar (Figur 9). Av sjukskrivningsdagarna under första året efter diagnos var 108,8 nettodagar i cancer. Under andra året efter diagnos (T_{+1} - T_{+2}) var medelantalet dagar med sjukskrivning lägre, 36,5 dagar (varav cancer: 26,6 dagar) och under tredje året (T_{+2} - T_{+3}) var medelantalet dagar 24,1 (varav cancer: 14,0 dagar). Medelantal nettodagar med sjuk- och aktivitetsersättning uppgick till runt 50 under vart och ett av de två åren innan diagnos och till runt 40 dagar under andra och tredje året efter diagnos.

Minskningen i medelantal SA dagar de senare åren beror framförallt på att äldre kvinnor med sjukersättning hade fått ålderspension.



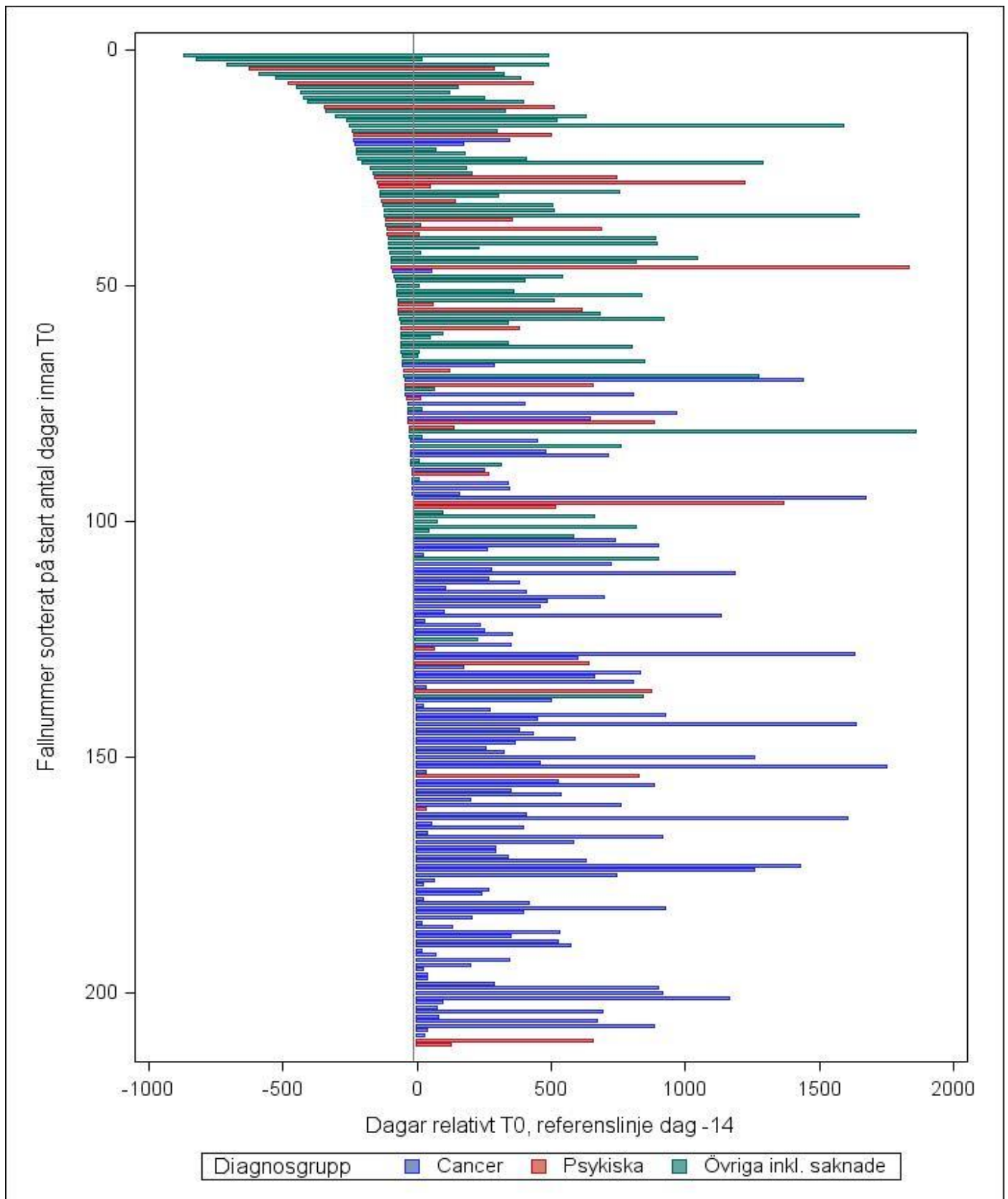
Figur 9. Medelantal nettodagar uppdelat på sjukskrivning (SS), sjuk- eller aktivitetsersättning (SA), diagnosgrupp och år relativt diagnosdatum (T_0) bland samtliga kvinnor med en första bröstcancerdiagnos under 2010 (N=3536) som då var högst 64 år. Antal inkluderade personer skiljer sig mellan åren till följd av att personer blivit >65 år, avlidit samt på grund av in- och utflyttning till Sverige.

6.2.2.2 Sjukskrivning och sjuk- eller aktivitetsersättning i samband med bröstcancerdiagnosen, beskrivande analyser

För modellering av risken att bli sjukskriven i samband med bröstcancerdiagnos valdes som utfall att få ett nytt sjukskrivningsfall som varade längre än 14 dagar i cancer (C00-D48), cancerrelaterade (Z80, Z85, N61-N63) eller psykiska diagnoser (F00-F99 samt Z73) under

tidsperioden från 14 dagar innan bröstcancerdiagnosen (T_0) till och med 29 dagar efter bröstcancerdiagnosen. Valet av 14 dagar innan diagnosdatum motiverades av att ett sjukskrivningsfall i samband med bröstcancerdiagnos skulle kunna påbörjas inom en (kortare) tidsperiod innan diagnosdatum, bland annat för att det kan ta tid innan provsvar om typ av tumör kan dröja från det att provet tas. Det är vanligen i samband med provsvaret som tumören registreras i cancerregistret. För att undersöka vilken dag som skulle kunna vara rimlig att sätta som gräns för detta, specialstuderades samtliga de sjukskrivningsfall som startade innan T_0 och som fortfarande pågick vid T_0 , efter att ha exkluderat personer som vid T_0 hade en pågående sjuk- eller aktivitetsersättning med omfattning om minst 75 procent. Antalet kvinnor som hade ett sådant sjukskrivningsfall var 211. I Figur 10 visualiseras dessa 211 sjukskrivningsfall med avseende på deras start- och slutdatum relativt T_0 samt sjukskrivningsdiagnos. Drygt hälften av fallen (108 fall) startade inom 14 dagar innan T_0 , och majoriteten av dessa sjukskrivningsfall var i en cancerdiagnos. I analyserna, liksom i fallen som visas i Figur 8, finns information från MiDAS för åren 2008-2013, det vill säga, fallen kan ha börjat innan 2008 eller slutat efter 2013. Bland fallen i figuren började ett av sjukskrivningsfallen redan under år 2007. Åtta av fallen pågick fortfarande i början av år 2015.

Inför framtagandet av prediktionsmodellen av risk för ny sjukskrivning exkluderades de kvinnor som vid T_0 hade en pågående sjuk- eller aktivitetsersättning om minst 75 procent (415 kvinnor, 11,7 %) samt de som hade ett pågående sjukskrivningsfall (106 kvinnor, 3,0 %). Som pågående sjukskrivningsfall definierades då fall som startade mer än 14 dagar innan T_0 och som fortfarande pågick vid T_0 (103 kvinnor), samt fall som startade mindre än 14 dagar innan T_0 och pågick vid T_0 men i andra sjukskrivningsdiagnoser än de studerade, det vill säga i andra diagnoser än cancer eller psykiska besvär (3 kvinnor). Av de 103 sjukskrivningsfall som startade mer än 14 dagar innan bröstcancerdiagnosen var tio med diagnosen C50, sju med annan cancerdiagnos (C00-D48, exklusive C50) och 23 med psykiska diagnoser. Två av kvinnorna som hade sjuk- eller aktivitetsersättning om 75 procent var även sjukskrivna vid T_0 .



Figur 10. De 211 sjukskrivningsfall >14 dagar som startade innan T₀ (datum för bröstcancerdiagnos) och som fortfarande pågick vid T₀, i förhållande till antal dagar innan T₀ fallen

startade respektive avslutades. Kvinnor med sjuk- eller aktivitetsersättning omfattande minst 75 % vid T_0 är exkluderade. Det vertikala strecket markerar 14 dagar innan T_0 .

Av de 3 015 kvinnorna som inte hade pågående sjukskrivning eller sjuk- eller aktivitetsersättning om minst 75 procent fick 1 496 (49,6 %) ett nytt sjukskrivningsfall som varade mer än 14 dagar i cancer eller psykiska diagnoser under perioden 14 dagar innan till 29 dagar efter bröstcancerdiagnosdatum (T_0). Gränsen om 29 dagar efter diagnosdatum sattes efter diskussion med beställarna. En liten andel av kvinnorna (1,3 %, 39 till antalet) fick under samma period ett nytt sjukskrivningsfall i andra diagnoser.

Inför modelleringen exkluderades sedan ytterligare 61 kvinnor, dels på grund av att värden saknades i cancerregistret för de variabler som skulle ingå i modelleringen och dels på grund av extrema värden på vissa av de kontinuerliga variablerna². Ytterligare en orsak till exkludering var att inte ha bott i Sverige under åren före diagnos (T_0). Efter dessa exkluderingar uppgick antalet kvinnor som ingick i kohorten som prediktionsmodellen baserades på till 2 954. Av dessa 2 954 kvinnor fick 49,5 procent ett sjukskrivningsfall i cancer eller psykiska diagnoser under perioden 14 dagar innan till 29 dagar efter datum för bröstcancerdiagnos, det vill säga i princip samma andel som innan detta sista exkluderingssteg.

Av de kvinnor som ingick i kohorten för modellering avled ingen inom den tidsperiod som där studerades (upp till och med 29 dagar efter T_0). Inte heller beviljades någon av kvinnorna sjuk- eller aktivitetsersättning med omfattning om 75 procent eller mer inom denna period.

I modelleringen var utfallet att få ett nytt sjukskrivningsfall som påbörjats under tidsperioden 14 dagar innan T_0 till och med 29 dagar efter T_0 och som blev minst 14 dagar långt. I Tabell 15 visas antal och andel personer som fick sitt första nya sjukskrivningsfall som blev längre än 14 dagar från dag 14 innan T_0 till och med 31 december 2013, uppdelat på olika tidsintervall från T_0 . Hur stor andel av fallen som var i cancer, psykisk respektive i övrig diagnos redovisas också. Personer med pågående sjukskrivning eller sjuk- eller aktivitetsersättning om minst 75 procent vid T_0 är här exkluderade. Av dessa kvarvarande 3 015 kvinnor fick 51 procent ett första sjukskrivningsfall från 14 dagar innan till och med 29 dagar efter bröstcancerdiagnosen, det vill säga den tidsperiod vi studerar i modelleringen, varav 98 procent i cancer eller psykiska diagnoser. Ser man till nästkommande 30 dagar, det vill säga från 30 dagar till och med 59 dagar efter datum för bröstcancerdiagnosen, fick ytterligare 20 procent av kvinnorna ett första sjukskrivningsfall, varav 97 procent i cancer eller psykiska diagnoser. Lite mindre än en femtedel av kvinnorna, 18 procent, fick inget sjukskrivningsfall under den period som här studeras, från och med dag 14 innan diagnosdatum till och med 31 december 2013. Nio kvinnor fick sjuk- eller aktivitetsersättning under den studerade perioden.

² När det gäller exkludering av personer på grund av att deras värden var extrema (så kallade outliers), så gjordes även analyserna med dem inkluderade. Inkludering/exkludering av misstänkta outliers förefaller inte påverka analysresultaten. Dock rekommenderar vi att dessa outliers hanteras med särskild varsamhet i kommande analyser. Dels är det av vikt att då vara uppmärksam på eventuella mätfel i data dels att studera om dessa outliers har starkare samband med riken att blir sjukskriven eller att ha långa sjukskrivningsfall.

Tabell 15. Antal och andel kvinnor per tidsintervall efter T₀ för vilket deras första sjukskrivningsfall inträffade, räknat från och med dag 14 innan T₀ upp till minst 3 år efter, för kvinnor med en första bröstcancerdiagnos under 2010 (N=3015). Kvinnor med pågående sjukskrivning eller pågående sjuk- eller aktivitetsersättning om minst 75 procent är här exkluderade. Sjukskrivningsfall kortare än 15 dagar är inte medräknade.

Dagar relativt T ₀	Antal (% av alla kvinnor; kolumn%)	Cancer ¹ (rad%)	Psykiska (rad%)	Övriga diagnoser (rad%)
Inget SS-fall fr.o.m. 14 dagar innan T ₀ till uppföljningstidens slut ^{2,3}	536 (17,8)			
14 dagar innan-29 dagar efter T ₀	1535 (50,9)	95	3	2
30 dagar efter-59 dagar efter T ₀	599 (19,9)	96	1	3
60 dagar efter-89 dagar efter T ₀	152 (5,0)	93	3	4
90 dagar efter-119 dagar efter T ₀	78 (2,6)	92	1	6
120 dagar efter-179 dagar efter T ₀	48 (1,6)	90	0	10
180 dagar efter-364 dagar efter T ₀	19 (0,6)	68	0	32
365 dagar efter-729 dagar efter T ₀	21 (0,7)	57	14	29
730 eller fler dagar efter T ₀ ³	27 (0,9)	18	15	67

¹ ICD-koder: C00-D48, Z80, Z85, N61-N63

² Nio av dessa kvinnor fick sjuk- eller aktivitetsersättning under perioden från 14 dagar innan T₀ och framåt

³ Då sjukskrivningsdata från MiDAS hämtades för perioden 2008-2013 varierar uppföljningstiden mellan personer beroende på när under 2010 de fick sin diagnos, från 3 års uppföljning (1095 dagar) till 4 år (1460 dagar).

6.2.2.3 Bakgrundsfaktorer för kvinnor med bröstcancer

I Tabell 16 beskrivs dels hela kohorten (3 536 kvinnor) och dels den grupp som inkluderades i modelleringen (2 954 kvinnor) utifrån socioekonomiska och sociodemografiska bakgrundsfaktorer. Omkring 40 procent av kvinnorna var 56 år eller äldre. För kvinnornas arbetsrelaterade sektortillhörighet var det stor skillnad mellan de två grupperna. Detta beror till stor del på att de kvinnor som exkluderades inför utvecklandet av modellen hade pågående sjuk- eller aktivitetsersättning om minst 75 procent vid T₀; majoriteten av dessa hade omfattning 100 procent, det vill säga, de var inte förvärvsarbetande.

Tabell 16. Socioekonomiska och sociodemografiska bakgrundsfaktorer för alla kvinnor <65 år i kohorten med en första bröstcancerdiagnos år 2010, samt för den grupp som sedan inkluderades i modelleringen.

	<i>Hela kohorten</i>		<i>Kohorten som användes för modellering</i>	
	<i>Antal</i>	<i>Andel (%)</i>	<i>Antal</i>	<i>Andel (%)</i>
Alla	3536	100,0	2954	100,0
Åldersgrupper ¹				
18-35	122	3,5	111	3,8
36-45	665	18,8	610	20,7
46-50	583	16,5	515	17,4
51-55	630	17,8	525	17,8
56-60	844	23,9	664	22,5
61-63	692	19,6	529	17,9
Födelseland				
Sverige	2950	83,4	2509	84,9
Övriga länder	586	16,6	445	15,1
Utbildningsnivå ¹				
Grundskola	551	15,6	378	12,8
Gymnasium	1549	43,8	1265	42,8
Eftergymnasial	1436	40,6	1311	44,4
Boenderegion ¹				
Nord	425	12,0	344	11,7
Mitt	486	13,7	403	13,6
Stockholm	823	23,3	710	24,0
Väst	934	26,4	767	26,0
Syd	868	24,6	730	24,7
H-region ¹				
Storstad	1376	38,9	1172	39,7
Mellanstor stad	1252	35,4	1051	35,6
Landsbygd	908	25,7	731	24,8
Familjesituation ¹				
Gift/sambo, utan barn	1042	29,5	852	28,8
Gift/sambo, med barn	1235	34,9	1129	38,2
Ensamstående, utan barn	885	25,0	656	22,2
Ensamstående, med barn	374	10,6	317	10,7
Civilstånd ¹				
Ogift, skild, änka/änkling	1523	43,1	1228	41,6
Gift, registrerat partnerskap	2013	56,9	1726	58,4
Sektor ¹				
Ej förvärvsarbete	810	22,9	379	12,8
Offentligt anställd	1409	39,9	1330	45,0
Privat anställd	1317	37,3	1245	42,2

¹ Gällande december 2009.

Majoriteten av kvinnorna tillhörde kategorin för de tidigaste stadierna av bröstcancer, TONOMO + Stadium 0+I, (60 % respektive 61 %) (Tabell 17). En majoritet hade ingen

tidigare sjukskrivning (81 % respektive 83 % för modelleringsgruppen) eller tidigare sjuk- eller aktivitetsersättning (82 % respektive 94 %) under de två åren innan bröstcancerdiagnosen.

En majoritet (56 %) av hela kohorten hade minst ett besök inom specialiserad öppenvård under de två föregående åren och knappt 12 procent hade minst en dag inom slutenvård inom samma period. Dessa andelar minskade efter exkluderingen av personer inför modellering av sannolikhet att bli sjukskriven i samband med diagnos.

Tabell 17. Cancerstadium och tidigare sjukskrivning och sjuk- eller aktivitetsersättning samt tidigare öppen- och slutenvård för alla kvinnor <65 år i kohorten med en första bröstcancerdiagnos år 2010 respektive för den grupp som inkluderades i modelleringen.

	Hela kohorten		Kohorten som användes för modellering	
	Antal	Andel (%)	Antal	Andel (%)
Cancerstadium				
Alla T, N, M saknas	25	0,7		
T0N0M0 + Stadium 0+I	2120	60,0	1799	60,9
Stadium II	1162	32,9	990	33,5
Stadium III+IV	229	6,5	165	5,6
Tidigare sjukskrivning, nettodagar ^{1,2}				
Ingen tidigare sjukskrivning	2870	81,2	2444	82,7
0,25-90 nettodagar	485	13,7	417	14,1
Mer än 90 nettodagar	181	5,1	93	3,2
Tidigare sjukskrivning, diagnoser ¹				
Psykiska diagnoser	161	4,6	124	4,2
Övriga diagnoser	540	15,3	412	14,0
Tidigare sjuk- eller aktivitetsersättning, nettodagar ¹				
Ingen tidigare sjuk- eller aktivitetsersättning	2908	82,2	2771	93,8
0,25-365 nettodagar	180	5,1	153	5,2
Mer än 365 nettodagar	448	12,7	30	1,0
Tidigare sjuk- eller aktivitetsersättning, diagnoser ¹				
Psykiska diagnoser	181	5,1	47	1,6
Övriga diagnoser	466	13,2	137	4,6
Tidigare besök i öppenvården ¹				
Inga besök	1549	43,8	1390	47,1
1-2 besök	1039	29,4	898	30,4
3 eller fler besök	948	26,8	666	22,6
Tidigare besök i öppenvården, diagnoser ¹				
Psykiska diagnoser	166	4,7	72	2,4
Övriga diagnoser	1825	51,6	1441	48,8
Tidigare slutenvård ¹				
Ingen slutenvård	3123	88,3	2709	91,7
1-14 dagar	354	10,0	235	8,0
Mer än 14 dagar	59	1,7	10	0,3
Tidigare slutenvård, diagnoser ¹				
Psykiska diagnoser	32	0,9	11	0,4
Övriga diagnoser	391	11,1	238	8,1

¹Med tidigare avses två år innan diagnosdatum fram till dag 15 innan diagnosdatum (se Figur 5).

²De 14 första dagarna i samtliga sjukskrivningsfall är ej medräknade.

6.2.3 Prediction models; preliminary analysis of cohorts of people with breast cancer or wrist fracture

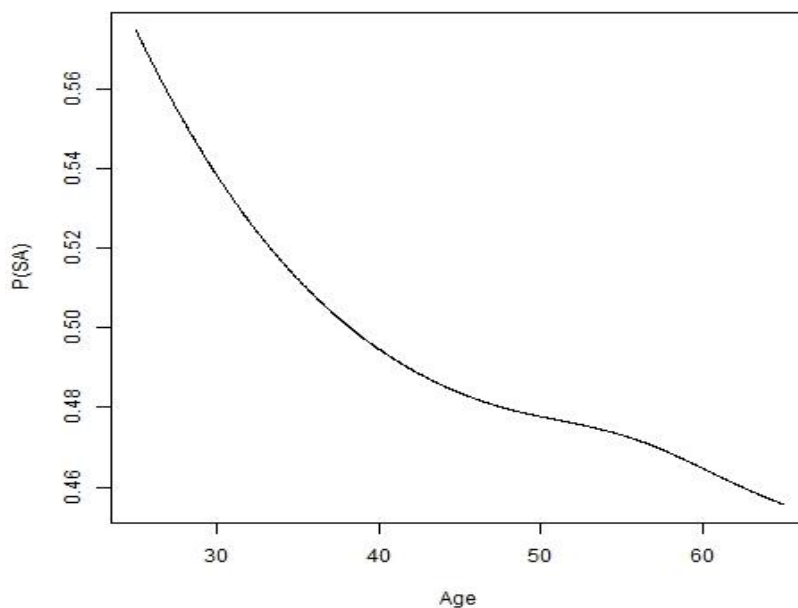
One of the aims of this project was to predict the binary outcome; a new sick-leave spell related to a diagnose, using a rich set of covariates that include demographic and socio-economic indicators, familiar status, health indicators, healthcare, and previous history of sickness absence and disability pension. The covariates are presented in Tabell 11 to Tabell 13 for wrist fractures and in Tabell 16 and Tabell 17 for breast cancer. In these preliminary analyses, the outcome is binary and the covariates are measured at baseline. A more informative approach would be to model the entire pattern of sickness absence and disability pension, as suggested in Section 6.3.

The risk of being sickness certified with a new sick-leave spell due to the same or related diagnoses was modeled using logistic regression (e.g., ^(21, 37)), that permits describing the effect of covariates on a binary outcome variable. A logistic model is formulated as follows

$$\log[p(y_i = 1)/p(y_i = 0)] = x_i'\beta \quad (1)$$

where y_i denotes the sick-leave status of individual i , and x_i is a vector of observed covariates.

Typically, x_i includes transformations of the available predictors, such as interactions, polynomials, and splines. Natural cubic splines (e.g., ^(22, 23, 38)) were here used to model potentially nonlinear effects of continuous covariates. For instance, the effect of age on the risk of having a new sick-leave spell in the breast cancer cohort was described using a 3-dimensional cubic spline that is represented graphically in Figur 11. The five variables that here were modeled using splines were, age, number of net days with sick leave and disability pension, respectively, number of visits to outpatient care and number of days in inpatient care during the two previous years.



Figur 11. Association between age and risk of a new sick-leave spell (SA) in the cohort of women diagnosed with breast cancer, modeled using a natural cubic spline.

Of the variables included in the multivariable model (Tabell 11 to Tabell 13) for the risk of having a new sick-leave spell after a wrist fracture (S62), the following were statistically significant ($p < 0.05$): age, H-region (= type of living area), educational level, sector of work, type of occupation, whether S62 was the main healthcare diagnosis at T_0 or not, if the patient had surgery at T_0 , as well as whether the patient had in-patient care (hospitalization) at T_0 . The variables concerning previous disability pension and country of birth were at the border of being statistically significant. The outcome in the model was to have a new sick-leave spell, due to wrist fracture (S62) as the main sick-leave diagnoses, starting between day T_{-1} and day T_{+6} .

Out of the variables included in the multivariable model (Tabell 16 and Tabell 17) for the risk of having a new sick-leave spell in connection to a first breast cancer diagnosis (C50), the following variables were statistically significant ($p < 0.05$): age, educational level, sector of work, geographical living area, cancer stage, and number of previous net sick-leave days. The outcome in this model was to have a new sick-leave spell starting between day T_{-14} and day T_{+29} , due to a cancer diagnosis, cancer-related diagnosis, or a mental sick-leave diagnosis (C00-D48, Z80, Z85, N61-N63, F00-F99, and Z73).

Possibly other variables might have been statistically significant if the variables had been categorized in other ways, if another number of knots had been chosen for the spines of the continuous variables, alternatively, if any of the variables had been completely excluded.

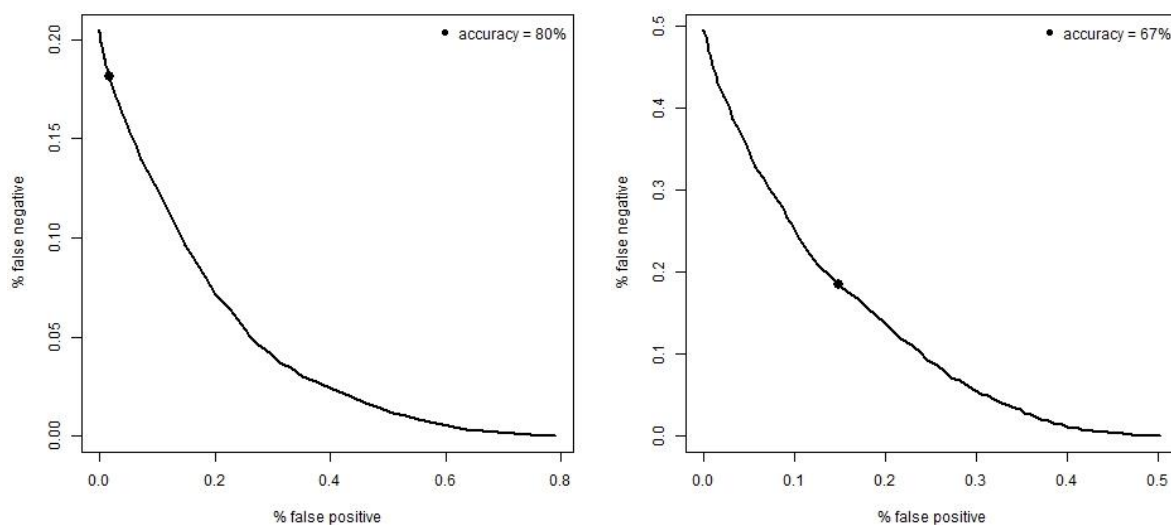
Details on model coefficients are provided in the appendix (bilaga, tabell 1 and 2). Statistical significance can be used to assess which variables are important in determining the risk of sickness absence for the two selected diagnoses. The R code that has been implemented to analyze the data is provided as supplementary material.

The resulting model can be used to predict a new observation, given covariates. We selected an optimal threshold c such that predicting *sick leave* = 1 whenever the fitted probability was above c , minimized the sum of false positive (FP) and false negative (FN) and maximizing the proportion of correctly classified observations. Results are summarized in Tabell 18, where we also report the area under the curve (AUC)^(26, 27). The predictive model for individuals with wrist fractures could correctly predict 80 per cent of the 14 465 individuals included, and the model for women with breast cancer could correctly predict 67 per cent out of 2 954 women, given the optimal thresholds (0.49 for wrist fractures and 0.54 for breast cancer).

Tabell 18. Optimal threshold (c), proportion of false positive (FP) and false negative (FN), proportion of correctly classified observations, and area under the curve (AUC) in the studies of people with a wrist fracture ($n=14\ 465$) and in the cohort of women with breast cancer ($n=2\ 954$), respectively.

	Wrist fracture	Breast cancer
Threshold (c)	0.49	0.54
% False positive	2	15
% False negative	18	18
% Correct	80	67
Area under the curve (AUC)	0.77	0.73

In Figur 12 below, we illustrate the receiver operating characteristic (ROC) curves for the two studies. The dots indicate the coordinates (FP , FN) corresponding to the selected value of threshold c .



Figur 12. Proportion of false positive/negative at different values of c (left: the cohort of people with wrist fracture; right: the cohort of women with breast cancer). The dots correspond to the optimal choice of threshold c .

Resultat, del III

6.3 Statistical methods and suggestions for future work

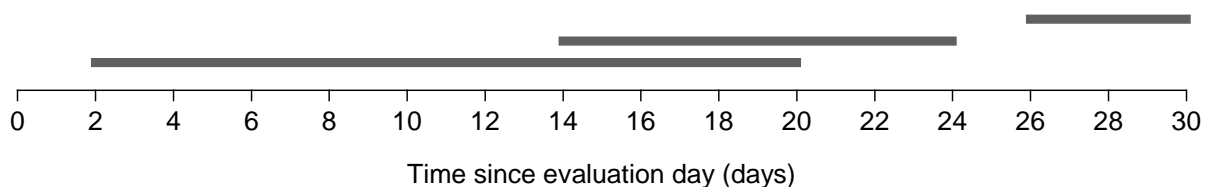
This chapter contains a description of a possible methodological approach to the derivation and implementation of statistical models for predicting the probability of being granted disability pension (sjuk- eller aktivitetsersättning), probability of starting a new sick-leave spell (sjukskrivningsfall), or ending an ongoing sick-leave spell. The models are intended for use in clinical healthcare practice or the Försäkringskassan.

The following sections introduce several important methodological issues that need to be taken into account in the process of developing and implementing the predictive models, such as estimation of prediction accuracy, internal and external validation, time-varying predictors, and periodic updating and recalibration as new data become available. These issues and their possible relevance on the prospective predictive models are summarized, and literature is referred to for more details.

The timeframe for the development of the project is assumed to be one year with support by Försäkringskassan. The work will require close collaboration between researchers at the Karolinska Institutet, representatives from Försäkringskassan, and possibly individuals with other expertise.

6.3.1 Methodological approach

The suggested project will develop and implement predictive models for the probability of starting a new sick-leave spell and for estimating its duration as well as for the probability of becoming disability pensioned. To better understand the problem, Figur 13 gives a visual representation of three fictitious individuals followed up over a time period of one month.



Figur 13. Three fictitious individuals followed up for a period of one month. The solid lines indicate the duration of a sick-leave spell.

Day zero indicates the day an individual is evaluated, for instance at a healthcare center. Day 1 indicates the day in the future for which the probability of starting a new sick-leave spell can be calculated. The solid lines indicate the duration of a sick-leave spell. For the individual

at the bottom of the plot the sick-leave spell starts on day 2 and ends on day 20. The middle individuals' sick-leave spell starts on day 14 and ends on day 24. The top individuals' sick-leave spell starts on day 26 and is still ongoing at the end of the one-month follow-up period.

All that was said above with respect to starting a sick-leave spell could be extended to be granted disability pension, with the exception that disability pension will last for a longer period of time. From age 30, disability pensions can since some years only be granted permanently. For simplicity, we discuss only sick-leave spells in the remainder of this chapter.

6.3.1.1 Time-invariant and time-varying characteristics

On any given day, data on sick leave and many other individual and environmental characteristics are available. The individual's characteristics include, for example, sex, age, type of work, previous diagnosis of cancer, multi morbidity, medication, family situation, economic situation, etcetera. Environmental characteristics include, for example, type of living area, type, size and economic situation of employer, type of caregiver, changes in insurance practices or rules, changes in healthcare treatment policies.

Some of these characteristics are constant over time, while others vary over time. Examples of time-invariant characteristics are sex, date of birth, place of birth, and genotypes. Examples of time-varying characteristics are age, education, morbidities, type of job and employer.

The number of days that already have passed in a sick-leave spell will be an important time-varying characteristic when the future duration of the spell is evaluated. This is clearly a strong predictor of the duration of a sick-leave spell. For example, the probability that its duration is over one year is smaller on the first day of a spell than on the 364th day.

6.3.1.2 Time scale

For any given individual observed on any given day the envisioned predictive models will predict the probability of being on sick leave on any of the subsequent days or weeks. The time scale, therefore, will be the time from the day an individual is evaluated and onward. This choice is dictated by the intended use of the predictive models. Having predictive models that predict the probability of starting a new sick-leave spell or the duration of an ongoing spell should facilitate its use in clinical practice and at Försäkringskassan.

The predicting ability of virtually any individual or environmental characteristic can be expected to change over time. For example, the predictors of long-term sick-leave spells that are important on the first day of a spell may be not the same as those that are important after one year into a spell.

6.3.1.3 Data källor

Data från den så kallade IMAS-databasen (Insurance Medicine All Sweden) kommer att användas i analyserna. Databasen är synnerligen omfattande vad gäller antal variabler, personer och tid den täcker. IMAS innehåller avidentifierade data, länkade på individnivå, från 11 rikstäckande register, och för detta projekt planeras data från ytterligare ett, regionalt, register att inkluderas. Vi avser att i analyserna använda data från nedan listade register.

Registren listas utifrån organisation som byggt upp och administrerar dem. Efter varje register listas, inom parantes, övergripande information om variabler som avses ingå i analyserna och för vilka år de för närvarande finns i IMAS databasen:

Statistiska centralbyrån

- *Statistiska centralbyråns individdatabas* för åren 1985–2013 (kön, födelseår, födelse-land, utbildningsnivå, civilstånd, familjesituation, antal hemmavarande barn i olika åldrar, boendeort, H-region, sysselsättning, yrke, anställningssektor – genom *flergenerationsregistret* erhålls motsvarande information om deras föräldrar)
- *FAD-databasen* (1986-2013) (information om individernas arbetsställen, till exempel vad avser storlek, andel anställda kvinnor, ekonomisk situation)
- *Undersökning om levnadsförhållanden* (ULF) (1985-2013) (cirka 125.000 av personerna i IMAS databasen har deltagit i ULF-studierna åtminstone en gång. Det innebär att för dessa personer finns det självrapporterade data om sjuklighet, livsstil etcetera, som kan användas i vissa analyser^(39, 40)).

Försäkringskassan

- *MikroData för Analys av Socialförsäkringen (MiDAS)* data om sjukfrånvaro och sjuk- och aktivitetsersättning för perioden (1994-2014) (datum, grad (hel eller deltid), sysselsättning vid sjukskrivningsfallets början, diagnoser från 1994 för SA fall och från 2005 för sjukskrivningsfall)
- *Register över personer som uppnått maximal tid i sjukförsäkringen* (2010-2014) (datum)
- *Register över personer som nekats förlängning av sjukpenning* (mars 2009-juli 2010) (datum)

Socialstyrelsen

- *Patientregistret (PAR), avseende slutenvårdsregistret (1964 för vissa, från 1987-2013 för samtliga), polikliniska operationer (1997-2000) samt öppenvårdsregistret (2001-2013)* (datum för öppen och slutenvård, diagnoser, operationskoder, typ av klinik, etcetera.)
- *Cancerregistret* (1958 – 2012) (datum och typ av cancer samt sjukdomens allvarlighetsgrad)
- *Medicinska Födelseregistret (MFR)* (1973-2013) (datum för förlossning och typ av förlossning)
- *Socialstyrelsens dödsorsaksregister* (1961-2013 vad avser dödsorsak, till och med 2014 vad avser datum för död) (datum för död och dödsorsak)
- *Läkemedelsregistret* (juli 2005-september 2014) (datum för utköp av receptbelagt läkemedel, typ av läkemedel och dos)

Stockholms läns landsting

- *Svenska MS registret* (2001-2013), rikstäckande kvalitetsregister över patienter med MS, (behandling, funktionsnedsättning, etcetera)
- *VAL-databaser* (1995-2014)⁽⁴¹⁾ – data som vi avser inkludera, men ännu inte har i IMAS databasen (avseende besök i primärvården inom Stockholms läns landsting; datum, diagnos, vård).

För att kunna utveckla modeller som predicerar risken att bli sjukskriven behöver man först identifiera personer med en sjukdom eller skada som är under risk att bli sjukskriven i denna diagnos. I IMAS kan det göras på flera olika sätt. Några exempel är genom diagnoser eller åtgärds-koder vid sjukhusinläggningar eller besök i specialiserad öppenvård, registrerade i patientregistret eller genom typ av medicinsk behandling läkemedelsregistret, cancerregistret, MS registret, diagnoser vid tidigare sjukskrivning eller sjuk- eller aktivitetsersättning. En länkning av VAL-databasen innebär ytterligare möjligheter.

Initialt behöver såväl stratifierade analyser som analyser för samtliga personer under risk för utfallet att genomföras. Sådana stratifieringar bör till exempel utgå från kön, ålderskategorier, födelse-region, sysselsättningsstatus (såsom arbetslös/i arbete), pågående partiell sjuk- eller aktivitetsersättning, etcetera.

6.3.2 Statistical models

This section provides some technical details about the statistical modeling framework. The described framework represents a starting point that will be adjusted to the data and continuous findings as these will develop during the course of the one-year project. One single statistical modeling approach could be used to model both the probability of starting a new sick-leave spell and for estimating its duration. However, we expect the modeling flexibility of this possible approach to be insufficient. We, therefore, plan to derive separate models for the probability of starting a new sick-leave spell and for estimating its duration and grade, that is, being for full- or part-time. These are detailed in the following subsections.

6.3.2.1 Modeling the probability of starting a new sick-leave spell

Suppose subject i is observed on day t_0 and it is possible that this individual starts a new sick-leave spell in the subsequent days. The probability that this subject will start a new sick-leave spell on any subsequent day $t > t_0$ can be modeled as

$$\Pr(\text{start SA}_{i,t}) = \text{expit}(\beta_0 + \beta_1 \text{age}_{i,t} + \beta_2 \text{sex}_{i,t} + \dots)$$

The function “expit” is defined as

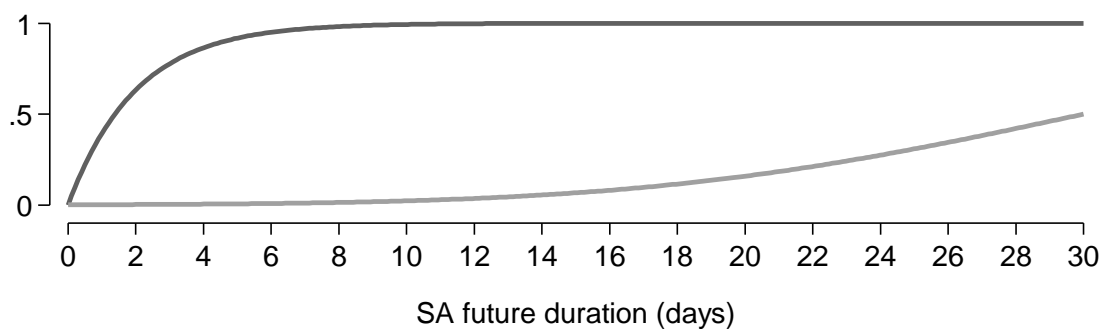
$$\text{expit}(\eta) = 1/(1 + \exp(-\eta))$$

The probability can be calculated for those who are at-risk of starting a new sick-leave spell. For example, those who are already on sick leave or on full-time disability pension are not at-risk. The data for estimating the above model will include all at-risk days and the first day of all sick-leave spells from all individuals. Death and emigration are considered censoring events as are old-age pension or turning 65 years old, C_i . In analyses with sick leave as the outcome also date when granted full-time disability pension is considered a censoring event. Censoring is incorporated in the at-risk sets, $A_t = \{i : C_i > t\}$. Censoring is assumed conditionally independent of the outcome, $C_i \perp T_i$.

The above is a logistic regression model. This is a popular statistical regression method that ensures that all estimated probabilities and confidence intervals will be in the feasible range (0, 1). Our modeling approach, however, is a non-standard application of logistic regression. By considering all at-risk days, our approach is strictly related to other regression methods to estimate the hazard of an event in survival analysis (e.g., Cox regression, Poisson regression).

6.3.2.2 Modeling the duration of a sick-leave spell

The predictive model described in this section will permit to estimate the probability of ending a sick-leave spell and returning back to work on any day after an individual on sick-leave – or e.g., being granted disability pension - is evaluated. Suppose subject i and that this individual is on a sick-leave spell on day t_0 . In Figur 14 the cumulative probability of ending a sick-leave spell on any future day $t > t_0$ for two fictitious individuals are shown. While the individual corresponding to the top curve has a probability of about 50% of ending the sick-leave spell by day 2, the other individual has a probability of 50% of being on sick-leave for over a month.



Figur 14. Cumulative probability of ending a sick-leave spell in any of future days for two fictitious individuals

Let D_{i,t_0} denote the duration of a sick-leave spell counting the time from day t_0 . A model for the duration of a sick-leave spell is

$$D_{i,t_0} \sim F(d \mid \gamma; \text{days on SA}_{i,t_0}, \text{age}_{i,t_0}, \text{sex}_i, \dots)$$

In the above expression, the function $F(d \mid \gamma)$ indicates the cumulative distribution function from a flexible family of functions, indexed by a parameter vector γ . For estimating the duration of a sick-leave spell all days on sick leave from all individuals will be included in the analyses. Conversely, the days not on sick-leave will be excluded. All the individual's and environmental characteristics as measured at the starting time point t_0 will be evaluated as possibly relevant predictors of the $F(d \mid \gamma)$.

Although nonparametric estimation of conditional distribution function, such as $F(d \mid \gamma)$, are available, we plan to start from parametric families of functions^(42, 43). These models, however,

can be made as flexible as necessary by using, for example, regression splines. We plan to consider regression spline basis to help model nonlinear relationships in the development of the models.

6.3.2.3 *A combined likelihood*

As mentioned earlier, we will derive separate predictive models for the probability of starting a new sick-leave spell and for future duration on any given day of an ongoing spell. This is to increase flexibility of the modeling approach. Once the two sets of predictive models have been derived, these can be combined in a single likelihood. This will permit extracting a great amount of information that may be extremely useful to clinicians and Försäkringskassan. Virtually any possible quantity can be estimated by combining the two sets of models. The following bulleted list shows only some examples of possible inferential quantity of interest:

- Prevalence of sick leave on any day
- Incidence of sick leave in any time period
- Summaries of the distribution of duration of sick-leave spells
- Calendar trends, time cycles, and seasonality
- All the important predictors of any of the above

The above inferences can be obtained for the entire population or for specific subpopulations like, for example, individuals on the day of a cancer diagnosis or of a bone fracture.

This will be conducted for the whole population as well as for primary health care patients regarding the most prevalent reasons for sickness absence generated from general practitioners, that is, common mental disorders and long-term pain related to musculoskeletal disorders⁽⁴⁴⁻⁴⁷⁾.

6.3.3 **Development and validation of predictive models**

The methodological approach described above will represent a working framework that will be subjected to continuous evaluation and fine-tuning during the course of the project. Within this framework, we expect that much work will go into selecting and evaluating important predictors. This crucial step will be carried out by a close collaboration between all researchers, analysts and experts involved in the project. What variables might be important predictor, what are their features, how practicable they might be in clinical settings, their epidemiologic relevance, and more, can be assessed by the joint effort of different experts.

In general, the identifying of an optimal set of individual and environmental characteristics that jointly may represent a set of optimal predictors are essential steps in the development phase of the predictive models. The selection of sets of optimal predictor will be based equally on statistical performance and clinical practicability. The former will ensure that the predictive models will satisfy statistical standards such as high precision and accuracy, while the latter will ensure that the predictive models can be effectively used in practice.

A vast literature is available in biostatistics and other public health disciplines on methods for assessing both statistical performance and practicability of predictive models⁽⁴⁸⁻⁵¹⁾. The following short subsections summarize few salient points that are particularly relevant to the suggested project. Some bibliographical references are reported for convenience. The list is however incomplete. The interested reader can peruse easily accessible literature on the topic for more information.

6.3.3.1 Identifying potential predictors

It will be essential to identify what individual and environmental characteristics are potentially relevant predictors. In this phase, the list of potential predictors should be as inclusive as possible. It would be fine to include predictors that will later turn out to be irrelevant, but it would be an issue if important predictors were left out. The potential predictors will be selected among those that are available in the data. The predictors that are easier to measure and more widely available will be preferred over the others.

6.3.3.2 Selecting sets of optimal predictors

Different sets of predictors will be selected and contrasted. Many techniques have been proposed for selecting an optimal set of predictors, such as machine-learning techniques like random forests^(52, 53) and penalized approaches like the lasso method⁽⁵⁴⁾. We will make use of all these available techniques to guide the selection of sets of optimal predictors. The decision on which predictors will be included in the final predictive models, however, great importance will be given to their practicability in the settings where these models will be utilized.

6.3.3.3 Coding process of predictors

The selected predictors will be measured on different scales. For example, some will be binary (e.g., sex, previous diagnosis of some disease), some categorical (e.g., type of job, sector, geographical location, or type of living area), some discrete (e.g., number of previous sick-leave spells, number of comorbidities), and some continuous (e.g., age, previous number of net days with sick-leave, days in hospital, years in Sweden, years unemployed, time since a specific diagnoses). The different types of variables will require different coding strategies for including them in the predictive models. The binary variables will be coded with values zero and one. The categorical variables that have sufficient number of valid observation in each category will be introduced by means of indicator variables or in some cases be categories, e.g., categories of number of sick-leave days during a year. As usual, a referent category will be left out of the predictive models for statistical identification in the estimation process. Categories with insufficient number of observations will be regrouped with other or omitted from the analyses. Discrete and continuous variables will be entered as numeric predictors. To model possible nonlinear associations we will use regression natural cubic splines⁽⁵⁵⁻⁵⁹⁾. The number of splines knots and location will be determined by optimizing the goodness of fit of the model as measured by the likelihood function.

6.3.3.4 Modeling interaction and effect modification

We will explore the joint effects of multiple predictors. For example, one predictor alone may be important only if in the presence of another interacting predictor. In epidemiologic literature this is often referred to as “effect modification” or “statistical interaction”⁽⁶⁰⁾. Major effect modification will be analyzed by stratification, while minor effect modification will be modeled by introducing two-way interaction terms in the predictive models. Higher order interaction will also be explored if warranted.

6.3.3.5 *Exploring certainty of predictions*

The certainty by which the probability of sickness absence or disability pension can be assessed will greatly vary across subpopulations. For example, if a person is run over by a bus and is laying in a bed at a hospital in a coma he or she will surely not be at work the next day. An otherwise healthy subject who is seen by a doctor for a minor skin infection is not expected to have a higher probability than anyone else of having to be sickness absent the next day. We expect that the predictors that may be important in one subpopulation may be irrelevant in other subpopulations. For example, the relevant determinants of the probability in young individuals may not be the same as those in the older age spectra. The set of important predictors and the magnitude of their effect are expected to vary across age, sex, type of job, social strata, and many other individual and environmental characteristics. Identifying subpopulations in which assessing probability and duration of sickness absence will be an extremely important task in the suggested project, which will entail a close collaboration between statisticians, epidemiologists, behavioral scientists, medical scientists and insurance medicine researchers with relevant background and skills that are available in the team for the suggested project.

6.3.3.6 *Statistical performance of the predictive models*

The predictive models will be thoroughly assessed in terms of their statistical performance. The selection of predictors and their coding will be studied carefully. For the predictive models for the residual duration of a sick-leave spell, the choice of the parametric family $F(d | \gamma)$ will also be intensively studied.

The tools to assess statistical performance of predictive models are widely known. Among many other performance indicators, we will focus on prediction error, sensitivity, specificity, and predictive values. To assess the above we will use methods like cross-validation⁽⁶¹⁾ and the “0.632+” bootstrap⁽⁶²⁾ among others.

In addition to assessing the ability of the models to predict start and duration of a sick-leave spell on new patients, we will consider the issues of forecasting in the future. More specifically, we will explore how models that are developed based on data from some calendar period perform when applied to subsequent calendar periods. A number of statistical tools permit determining forecast error (among others: ⁽⁶³⁾).

6.3.4 Deliverables

The suggested study will yield a number of important findings and tools that will be equally relevant in clinical settings and in research. The immediate deliverables will be predictive

models for estimating the probability of for instance, being granted disability pension, starting a new sick-leave spell, or ending an ongoing sick-leave spell. These models will permit identifying not only important predictors of any of the above probabilities but also to determine the level of certainty in the estimates. For example, it will help identify groups of individuals that are very likely to return to work within a week, those that are very likely to be on sick leave for, for instance, at least a month, and those who's returning to work is hard to predict. The predictive models will be optimized for use in practical settings. After the tools will be delivered, they will not require further assistance to be used. A web-calculator can also be developed if considered useful. (For an example of a predictive model developed at the Unit of Biostatistics at Karolinska Institutet on an entirely different topic, see www.imm.ki.se/biostatistics/calculators/prisk/.)

7 Slutkommentarer

Resultaten har sammanfattats i den inledande sammanfattningen. Här diskuteras de kort.

Det är viktigt att beakta att de här genomförda analyserna är explorativa och mycket preliminära. De utgör framförallt exempel på vad som kan vidareutvecklas i kommande analyser än någon typ av slutresultat. Trots det ganska rudimentära angreppssätt vi hittills haft för att utveckla de presenterade modellerna, lyckades vi korrekt predicera runt 70 procent av individerna. Detta indikerar att det finns goda möjligheter att i framtiden få fram mycket träffsäkra prediktionsmodeller, speciellt för vissa grupper. Framtagandet av sådana modeller kräver emellertid andra och betydligt mer tidskrävande arbetssätt samt mer avancerade statistiska metoder än vad som har funnits utrymme för att använda i denna rapport. Även om 70 procent korrekt predicering är ett förhållandevis bra resultat, vill vi understryka att det innebär att 30 procent blir felklassificerade. Det vill säga, det finns potential att utveckla modellerna vidare, vilket vi ger förslag till hur det kan ske.

Det finns ett stort behov av kunskap om riskfaktorer för att en person blir sjukskriven, för att det sjukskrivningsfallet blir långt alternativt att personen får många upprepade korta fall eller blir mer permanent marginalisering från arbetsmarknaden i form av sjuk- eller aktivitetsersättning. Detta behov finns både inom hälso- och sjukvård, inom Försäkringskassan, Arbetsförmedlingen och bland arbetsgivare. Sådan kunskap kan även användas som bas för att studera effekter av åtgärder för att påverka utfallet. Kunskapen kan bidra till mer effektivt användning av knappa resurser och det torde vara kostnadseffektivt att utveckla denna typ av förfinade prediktionsmodeller.

Resultaten visar på vikten av att i studier av sjukskrivning inte bara ha med dem som blivit sjukskrivna, utan samtliga personer i arbetsföra åldrar. Detta stämmer med klinisk erfarenhet och är även i linje med överenskommelser mellan olika aktörer och myndigheter (till exempel inom Nationellt försäkringsmedicinskt forum), avseende bland annat hur Socialstyrelsens diagnosspecifika försäkringsmedicinska beslutsstöd skulle utformas⁽⁶⁴⁾. En stor andel av personerna som har olika sjukdomar och skador är inte alls sjukskrivna.

I olika sammanställningar framkommer att den absoluta merparten av alla sjukskrivningsfall inte blir längre än några dagar och ungefär 80 procent av sjukskrivningsfallen har avslutats inom en vecka^(6, 65). Försäkringskassan har oftast inte information om sjukskrivningsfall som är kortare än två veckor.

Resultaten visar även att information om sjuklighet, vård och behandling ökar modellernas prediktionsförmåga.

7.1 Metodologiska aspekter

Styrkor i denna studie är att den är populationsbaserad, innehåller reella data för samtliga personer i arbetsför ålder i Sverige och baseras på data från rikstäckande register av god kvalitet⁽⁶⁶⁻⁶⁸⁾. Andra styrkor är att materialet är mycket omfattande; innehåller information om många personer, många variabler och för många år och därmed tillåter analyser på subgruppsnivå, inklusive för specifika diagnoser. Det finns heller inget bortfall, vi kan alltså

följa samtliga personer under hela uppföljningstiden, alternativt tills de avlidit eller emigrerat. Materialet innebär även mycket goda möjligheter att validera de framtagna prediktionsmodellerna på annan kohort. Detta är mycket viktigt för att inte landa i en överoptimistisk prediktion som möjligen inte är optimal i klinisk verksamhet.

I databasen finns information om ett stort antal faktorer, både socioekonomiska och sociodemografiska faktorer samt information om sjuklighet, vård och behandlingar. Att vi kunnat inkludera information om sjuklighet i dessa analyser är en central styrka i projektet, då Försäkringskassan inte har den informationen. En annan styrka är att vi kunnat inkludera personerna som har fått en viss diagnos, inte bara de som blivit sjukskrivna i den diagnosen. Det innebär att resultaten och modellerna även kan vara mer användbara i praxis.

En annan styrka är att vi i prediktionsanalyserna av långa sjukskrivningsfall har utgått från sjukskrivningsfallets startdatum, inte från datum när ersättning utbetalats från Försäkringskassan. Vi har alltså utgått från 'startdatum 1' i MiDAS när vi identifierat fall som pågått i minst 21 dagar. Om vi inte hade gjort det, hade vi introducerat en bias, till exempel i förhållande till arbetslösa personer samt till personer med vissa kroniska eller långvariga diagnoser som innebär att de får rätt till ersättning från Försäkringskassan från dag ett i ett sjukskrivningsfall.

Om flera personer hade ett nytt sjukskrivningsfall som uppfyllde inklusionskriterierna, valde vi att utgå från det första sådant fall, efter diskussion med SRS projektgruppen. Vi kunde givetvis valt det sista, eller gjort ett slumpmässigt urval av fallen, för de få personer som hade mer än ett. Betydelsen av att välja det första fallet kan behöva analyseras i kommande studier.

Gränsen om 29 dagar för ett nytt sjukskrivningsfall efter datum för bröstcancerdiagnos kan diskuteras och andra tidsgränser kan behöva testas. Det är ofta effekter av behandlingen för bröstcancer, snarare än själva bröstcancersjukdomen, som leder till nedsatt arbetsförmåga för kortare eller längre tid – och det kan dröja längre än fyra veckor efter diagnos innan behandlingar sätts in ⁽⁶⁹⁻⁷³⁾.

Validitet i sjukskrivningsdiagnoser diskuteras ibland men det finns ytterst få studier om detta. I en studie jämfördes sjukskrivningsdiagnos på 300 sjukintyg, där sjukskrivande läkare arbetade i primärvården, med information från de medicinska journalerna ⁽¹⁸⁾. Validiteten befanns vara god, särskilt på högre kapitelnivåer. I en annan studie från Inspektionen för socialförsäkringen, granskades validiteten i de data om diagnos som kodats in i MiDAS ⁽⁷⁴⁾ och den befanns vara god.

Ibland ifrågasätts det om diagnosen depressiv episod används när istället diagnosen recidiverande depressioner (ICD-kod: F33) borde användas, särskilt eftersom depressiv episod ibland till och med förekommer som huvuddiagnos för sjuk- och aktivitetsersättningar. I det försäkringsmedicinska beslutsstödet står följande ”Depressiv episod (F32) betecknar ett första insjuknande i depression. Om diagnosen förekommit tidigare, ska diagnosen recidiverande depression (F33) användas.” Trots detta var det cirka en av tio (9,3 %) av dem med depressiv episod som, under de 365 dagarna innan det analyserade sjukskrivningsfallet startade, hade ett tidigare sjukskrivningsfall med samma diagnos (F32). En annan aspekt här är att det i MiDAS endast

finns information om den första diagnosen på sjukintyget, det vill säga, den diagnos som bidrog mest till den nedsatta aktiviteten och arbetsförmågan den dag sjukintyget skrevs. Det innebär att depressiv episod kan ha angivits som bidragande orsak i ett antal sjukintyg. När ett sjukskrivningsfall förlängs med ett nytt intyg kan ibland en ny diagnos ges. Detta registreras dock ytterst sällan i MiDAS. För detta projekt betyder det att några av sjukskrivningsfallen som studerats kan ha fått en ny diagnos. Det betyder även att några fall kan ha missats, om till exempel ett sjukskrivningsfalls första huvuddiagnos var en fraktur och fallet senare förlängdes i depressiv episod och den delen av fallet översteg 21 dagar.

En svaghet är att vi hittills endast analyserat enskilda riskfaktorer och inte undersökt det faktum att olika kombinationer av riskfaktorer kan samverka. I kommande steg behöver sådana interaktioner identifieras vilket i sin tur avsevärt kan förbättra modellernas prestanda. Detta kräver dock ett systematiskt arbete för att kunna säkerställa att sådana kombinationer som identifieras och inkluderas är sanna och inte slumpmässiga.

En annan svaghet är att vi för samtliga fyra studerade diagnoser utgått från ICD-10 på treställighetsnivå, eftersom Försäkringskassan endast efterfrågat diagnos på treställighetsnivå i sina förtryckta sjukintygsblanketter. Även om en läkare anger diagnos på fyr- eller femställighetsnivå inkluderas sällan sådan information i MiDAS.

Detta innebär att det kan finnas undergrupper av diagnoser, som kan ha olika typer av påverkan på en individs funktion, aktivitet och arbetsförmåga. Rekommendationerna i försäkringsmedicinska beslutstödet kan även variera med underdiagnoser.

Definition av vad som är ett 'långt sjukskrivningsfall' kan diskuteras och andra längder kommer att användas i kommande analyser. I de föreslagna modelleringarna avses en modell tas fram som kan användas för vilken sjukskrivningslängd som helst. Att som här utgå från rekommendationerna i det försäkringsmedicinska beslutstödet var ett pragmatiskt val. Självklart kan andra längder användas – och det kan finnas anledning att ta fram modeller som inte bara predicerar olika duration, utan även detta i olika sjukskrivningsgrader. På motsvarande sätt behöver andra val i dessa preliminära analyser hanteras i kommande studier, till exempel vad gäller att exkludera inte bara de med hel sjuk- och aktivitetsersättning utan även de med en omfattning om 75 procent.

7.2 Fortsatta studier

Det ingick i uppdraget att ge förslag till fortsatta studier, vilket har gjorts i avsnitt III; 6.3. Vårt förslag innebär unika möjligheter att i en tvärvetenskaplig miljö och med hjälp av omfattande data om hela befolkningen i arbetsför ålder, genomföra analyser och ta fram modeller som kan användas som bas för fortsatt arbete inom området. Vi har här valt att utgå från sjukskrivningsdiagnoser, då såväl Försäkringskassans data och andra studier som försäkringsmedicinska beslutstödet visar att sjukskrivningslängd varierar mycket med typ av diagnos. Även andra utgångspunkter kan naturligtvis tas. Mycket grundläggande utvecklingsarbete behövs, baserat i gedigen kunskap om statistik, denna typ av data, den praktiska situationen med en patient/försäkrad samt om sjukförsäkringen och försäkringsmedicin.

8 Tabell- och figurtexter

8.1.1 Tabellförteckning

Tabell 1. Antal nya sjukskrivningsfall som blev minst 21 dagar långa under 2010 och 2011, samtliga sådana fall samt sjukskrivningsfall i depressiv episod respektive i artros. Antal personer med minst ett sådant sjukskrivningsfall samt antal personer 16-64 år med minst ett sådant sjukskrivningsfall i depressiv episod respektive i artros.....	22
Tabell 2. Antal personer i varje grupp bland de personer som hade ett nytt sjukskrivningsfall i depressiv episod (F32) som år 2010 blev minst 21 dagar långt, och andel av dessa vars sjukskrivningsfall blev längre än 180 dagar samt betavärdet och oddsraten (OR) med 95 % konfidensintervall (KI) för de variabler som inkluderades i modellen som predicerar risken att bli sjukskriven längre än 180 dagar för dessa personer.	32
Tabell 3. Andel korrekt predicerade vid val av olika brytpunkter, dels i träningskohorten 2010 och dels i valideringskohorten 2011.	34
Tabell 4. Olika mått för prestanda på modellen som predicerar risken att bli sjukskriven längre än 180 dagar bland de personer som 2010 alternativt 2011 hade ett nytt sjukskrivningsfall i depressiv episod (F32) som blev minst 21 dagar långt. Modellen är utvecklad baserad på 2010-kohorten (n=23 079) och validerad på 2011-kohorten (n=23 238).....	35
Tabell 5. Antal personer i varje grupp, bland de personer som hade ett nytt sjukskrivningsfall i artros (M15-M19), som år 2010 blev minst 21 dagar långt, och andel av dessa vars sjukskrivningsfall blev längre än 120 dagar samt betavärdet och oddsraten (OR) med 95 % konfidensintervall (KI) för de variabler som inkluderades i modellen som predicerar risken att bli sjukskriven längre än 120 dagar för dessa personer.	36
Tabell 6. Andel korrekt predicerade vid val av olika brytpunkter dels i träningskohorten 2010 och dels i valideringskohorten 2011.	38
Tabell 7. Olika mått för prestanda på modellen som predicerar risken att bli sjukskriven längre än 120 dagar bland de personer som hade ett nytt sjukskrivningsfall i artros (M15-M19) som blev minst 21 dagar långt. Modellen är utvecklad baserad på 2010-kohorten (n=10 851) och validerad på 2011-kohorten (n=11 089).....	39
Tabell 8. Deskriptiva analyser av de prediktionsvärden som genereras av modellerna ovan (Tabell 2 och Tabell 5), uppdelat på dem som uppfyller respektive inte uppfyller tidigare använda exklusionskriterier ⁽²⁾ , för de personer som haft ett sjukskrivningsfall i depressiv episod respektive i artros som respektive år blev minst 21 dagar långt. Använt utfall är att det fallet blev >180 dagar respektive >120 dagar långt.	40
Tabell 9. Jämförelser mellan modeller som inkluderar olika faktorer, för att predicera risken att bli långtidssjukskriven i depressiv episod respektive i artros bland personer sjukskrivna i minst 21 dagar i respektive diagnos under 2010; arean under ROC-kurvan (AUC) presenteras för samtliga modeller.	42
Tabell 10. Antal och andel personer med sjukskrivning (SS) respektive sjuk- eller aktivitetsersättning (SA), för samtliga diagnoser respektive för olika diagnosgrupper, uppdelat på år relativt datum för handlovsfrakturen (T ₀) bland de 16 337 personer som år 2010 var 16-65 år gamla och fick en handlovsfraktur (ICD 10 kod: S62).	45
Tabell 11. Socioekonomiska och sociodemografiska bakgrundsfaktorer för alla personer i kohorten med en handlovsfraktur (S62) under år 2010, samt för den kohorten från 2011 som användes för modelleringen.	48

Tabell 12. Faktorer relaterade till vårdtillfället för alla personer i kohorten med en handlovsfraktur (S62) under år 2010 (n=16 337), samt för den kohorten som användes för modelleringen (n=14 465).	49
Tabell 13. Tidigare sjukskrivning och sjuk- eller aktivitetsersättning samt tidigare öppen- och slutenvård för alla personer i kohorten med en handlovsfraktur (S62) under år 2010 (n=16 337), samt för den kohorten som användes för modelleringen (n=14 465). Med tidigare avses under två år (730 dagar) innan handlovsfrakturen 2010; T_0 .	50
Tabell 14. Antal och andel personer med sjukskrivning (SS) och sjuk- eller aktivitetsersättning (SA), för samtliga diagnoser respektive uppdelat på diagnos, per år utgående från datum för bröstcancerdiagnos (T_0) för samtliga kvinnor med denna diagnos under 2010 och som då var upp till 64 år gamla (N=3536), under två år innan och upp till tre år efter.	52
Tabell 15. Antal och andel kvinnor per tidsintervall efter T_0 för vilket deras första sjukskrivningsfall inträffade, räknat från och med dag 14 innan T_0 upp till minst 3 år efter, för kvinnor med en första bröstcancerdiagnos under 2010 (N=3015). Kvinnor med pågående sjukskrivning eller pågående sjuk- eller aktivitetsersättning om minst 75 procent är här exkluderade. Sjukskrivningsfall kortare än 15 dagar är inte medräknade.	57
Tabell 16. Socioekonomiska och sociodemografiska bakgrundsfaktorer för alla kvinnor <65 år i kohorten med en första bröstcancerdiagnos år 2010, samt för den grupp som sedan inkluderades i modelleringen.	58
Tabell 17. Cancerstadium och tidigare sjukskrivning och sjuk- eller aktivitetsersättning samt tidigare öppen- och slutenvård för alla kvinnor <65 år i kohorten med en första bröstcancerdiagnos år 2010 respektive för den grupp som inkluderades i modelleringen.	59
Tabell 18. Optimal threshold (c), proportion of false positive (FP) and false negative (FN), proportion of correctly classified observations, and area under the curve (AUC) in the studies of people with a wrist fracture (n=14 465) and in the cohort of women with breast cancer (n=2 954), respectively.	62

8.1.2 Figurförteckning

Figur 1. Exempel på två fiktiva personers sjukskrivningsfall och sjukersättning, avseende duration och grad, under loppet av en femårsperiod.	11
Figur 2. Diagram över en persons sjukfrånvaro år 1985 - augusti 1996	12
Figur 3. Illustration av sjukfrånvarofall i relation till start- och slutdatum och årsgränser.	13
Figur 4. Illustration av sjukfrånvarofall i relation till start- och slutdatum och årsgränser för olika personer; A till H, där personen H inte har något sjukskrivningsfall alls under denna period och till exempel person E har ett sjukskrivningsfall som påbörjas och avslutas under föregående år, här 2009.	13
Figur 5 Illustration av exempel av sjukskrivningsfall i relation till start- och slutdatum och den period relativt T_0 som studeras i modelleringen av sannolikheten att bli sjukskriven i samband med diagnos, för olika personer med i) handlovsfraktur respektive med ii) bröstcancer. T_0 = dagen för diagnosen.	15
Figur 6. Fördelning av prediktionsvärden som genereras av modellen i avsnitt 6.1.1, för de personer som under 2010 hade ett nytt sjukskrivningsfall som blev minst 21 dagar långt i diagnosen depressiv episod, uppdelat på dem som uppfyllde respektive inte uppfyllde två av de exklusionskriterier som använts i tidigare analyser ⁽²⁾ (tidigare ha uppnått maxtid i sjukförsäkringen alternativt ha eller ha haft sjuk- eller aktivitetsersättning (SA) senaste året)....	41

- Figur 7. Fördelning av prediktionsvärden som genereras av modellen i avsnitt 6.1.2, för de personer som under 2010 hade ett nytt sjukskrivningsfall som blev minst 21 dagar långt i en artrosdiagnos, uppdelat på de som uppfyllde respektive inte uppfyllde två av de exklusionskriterier som använts i tidigare analyser⁽²⁾ (tidigare ha uppnått maxtid i sjukförsäkringen alternativt ha eller ha haft sjuk- eller aktivitetsersättning (SA) senaste året). 41
- Figur 8. Medelantal nettodagar uppdelat på sjukskrivning (SS) respektive sjuk- eller aktivitetsersättning (SA), olika diagnosgrupper och år relativt datum för handlovsfraktur (T_0) bland de 16 337 personer som fick detta år 2010 och då var 16-65 år gamla. Antal inkluderade personer skiljer sig något mellan åren till följd av att personer blir >65 år, avlider samt på grund av in- och utflyttning till Sverige. 46
- Figur 9. Medelantal nettodagar uppdelat på sjukskrivning (SS), sjuk- eller aktivitetsersättning (SA), diagnosgrupp och år relativt diagnosdatum (T_0) bland samtliga kvinnor med en första bröstcancerdiagnos under 2010 (N=3536) som då var högst 64 år. Antal inkluderade personer skiljer sig mellan åren till följd av att personer blivit >65 år, avlidit samt på grund av in- och utflyttning till Sverige. 53
- Figur 10. De 211 sjukskrivningsfall >14 dagar som startade innan T_0 (datum för bröstcancer-diagnos) och som fortfarande pågick vid T_0 , i förhållande till antal dagar innan T_0 fallen startade respektive avslutades. Kvinnor med sjuk- eller aktivitetsersättning omfattande minst 75 % vid T_0 är exkluderade. Det vertikala strecket markerar 14 dagar innan T_0 55
- Figur 11. Association between age and risk of a new sick-leave spell (SA) in the cohort of women diagnosed with breast cancer, modeled using a natural cubic spline. 61
- Figur 12. Proportion of false positive/negative at different values of c (left: the cohort of people with wrist fracture; right: the cohort of women with breast cancer). The dots correspond to the optimal choice of threshold c 62
- Figur 13. Three fictitious individuals followed up for a period of one month. The solid lines indicate the duration of a sick-leave spell. 63
- Figur 14. Cumulative probability of ending a sick-leave spell in any of future days for two fictitious individuals 67

9 Referenser

1. International Statistical Classification of Diseases and Related Health Problems 10th Revision, Version for 2007: WHO; 2010 [cited 2010 March 26]. <http://www.who.int/classifications/icd/en/>.
2. 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.
3. MiDAS Sjukpenning och rehabiliteringspenning. Försäkringskassan, 2011.
4. Longitudinell Integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier (LISA) 1990 till 2009. SCB, 2011.
5. Sjukskrivning - orsaker, konsekvenser och praxis. En systematisk litteraturöversikt. Stockholm: Statens beredning för medicinsk utvärdering (SBU), 2003.
6. Kjeldgård L, Ekmer A, Vaez M, Alexanderson K. Sjukfrånvaro bland kvinnor och män inom Polismyndigheten i Stockholms län. Rapport 2010. Stockholm: Institutionen för klinisk neurovetenskap, försäkringsmedicin Sf; 2010.
7. Hensing G, Alexanderson K, Allebeck P, Bjurulf P. How to measure sickness absence? Literature review and suggestion of five basic measures. *Scandinavian Journal of Social Medicine*. 1998;26(2):133-45.
8. Borg K, Goine H, Söderberg E, Marnetoft S-U. Measures of Sickness Absence: a Comparison between Seven Measures Based on Data from three Swedish Counties. *European Journal of Public Health*. 2004;14(4 Supplement):109.
9. Voss M, Stark S, Alexanderson K, Vingård E. The effect of using different measures of sickness absence. *European Journal of Public Health*. 2004;14(4 Supplement):46.
10. Borg K, Hensing G, Alexanderson K. Risk factors for disability pension over 11 years in a cohort of young persons initially sick-listed with low back, neck, or shoulder diagnoses. *Scand J Public Health*. 2004;32(4):272-8.
11. Jones B, Roeder K. A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. *Sociological Methods & Research*. 2001;29(3):374-93.
12. Jones B, Nagin D. Advances in Group-Based Trajectory Modeling and an SAS Procedure for Estimating Them. *Sociological Methods & Research*. 2007;35(4):542-71.
13. Försäkringsmedicinskt beslutsstöd, diagnosspecifikt. www.socialstyrelsen.se: Socialstyrelsen.
14. Försäkringsmedicinskt beslutsstöd - vägledning för sjukskrivning (reviderad 2012). Stockholm: Socialstyrelsen; 2012.
15. Pathology and genetics of tumours of the breast and female genital organs. Lyon: WHO; IARC Press; 2003.
16. Sobin L, Wittekind C. TNM Classification of Malignant Tumours. Hoboken, New Jersey: John Shiley & Sons; 2002.
17. Spiessl B, Scheibe O, Wagner G. TNM-Atlas. Illustrated Guide to the classification of Malignant Tumours. Heidelberg New York: Springer-Verlag Berlin 1982.
18. Ljungdahl L, Bjurulf P. The accordance of diagnoses in a computerized sick-leave register with doctor's certificates and medical records. *Scandinavian Journal of Social Medicine*. 1991;19(3):148-53.
19. Saboonchi F, Petersson LM, Wennman-Larsen A, Alexanderson K, Brannstrom R, Vaez M. Changes in caseness of anxiety and depression in breast cancer patients during the first year following surgery: Patterns of transiency and severity of the distress response. *European*

- journal of oncology nursing : the official journal of European Oncology Nursing Society. 2014.
20. Saboonchi F, Petersson LM, Wennman-Larsen A, Alexanderson K, Vaez M. Trajectories of Anxiety Among Women with Breast Cancer: A Proxy for Adjustment from Acute to Transitional Survivorship. *J Psychosoc Oncol*. 2015;33(6):603-19.
 21. Hosmer D, Lemeshow S, Sturdivant R. *Applied logistic regression*. 3rd ed: Wiley; 2013.
 22. Hastie T. Generalized additive models. In: Chambers J, Hastie T, editors. *Statistical Models in S: Wadsworth & Brooks/Cole*; 1992.
 23. Olsson D, Alexanderson K, Bottai M. Sickness absence and the time-varying excess risk of premature death: a Swedish population-based prospective cohort study. *Journal of epidemiology and community health*. 2015.
 24. Karel ea. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Ann Intern Med*. 2015;162(1):W1-W73. doi:10.7326/M14-0698.
 25. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128-38.
 26. Hastie T, Tibshirani R, Friedman JH. *The Elements of Statistical Learning: data mining, inference, and prediction* 2009.
 27. Fawcett T. An Introduction to ROC Analysis. *Pattern Recognition Letters* 2006;27(8):861 - 74.
 28. Yates J. External correspondence: decomposition of the mean probability score. *Organ Behav Hum Perform*. 1982;30:132-56.
 29. Hilden J, Habbema J, Bjerregaard B. The measurement of performance in probabilistic diagnosis. Part II: Trustworthiness of the exact values of the diagnostic probabilities. *Methods Inf Med*. 1978;17:227-37.
 30. Miller ME, Langefeld CD, Tierney WM, Hui SL, McDonald CJ. Validation of probabilistic predictions. *Med Decis Making*. 1993;13(1):49-58.
 31. Nagelkerke N. A note on a general definition of the coefficient of determination. *Biometrika*. 1991;78:691-2.
 32. Brier G. Verification of forecasts expressed in terms of probability. *Mon Wea Rev*. 1950;78:1-3.
 33. Timpka T, Dahlström Ö, Spreco A, Karlgren J. Initiala epidemiologiska analyser av sjukskrivningsprognos. Förstudierapport Stöd för rätt sjukskrivning. Stockholm: Försäkringskassan, Sveriges Kommuner och Landsting, 2015 Contract No.: Dnr FK: 028522-2014, Dnr SKL: 14/3621.
 34. Reed P, editor. *The Medical Disability Advisor. Workplace Guidelines for Disability Duration*. Singapore: Reed Group Holdings, Ltd.; 2004.
 35. Delclos G. Comparative overview of disability guidelines in the United States; part of a workshop titled: Diagnosis-specific sickness certification guidelines/recommendations: their background and implications. 5th European Public Health Conference; 7-10 November 2012; Malta: *European Journal of Public Health*; 2012. p. 96-7.
 36. Försäkringsmedicinskt beslutsstöd - vägledning för sjukskrivning. Socialstyrelsen, 2007.
 37. Agresti A. *Categorical Data Analysis*. New York: Wiley-Interscience; 2002.
 38. De Boer C. *A Practical Guide to Splines*: Springer; 2001.
 39. Pierre PV, Fridberger A, Wikman A, Alexanderson K. Self-reported hearing difficulties, main income sources, and socio-economic status; a cross-sectional population-based study in Sweden. *BMC Public Health*. 2012;12:874.

40. Gustafsson K, Marklund S, Aronsson G, Wikman A, Floderus B. Which role does social isolation and low participation in society play for risk of disability pension among young adults? *PLoS ONE*. 2015;10(6:e0130361).
41. VAL databaserna. Datalager för uppföljning av vårdhändelser i SLL. Översiktlig beskrivning. 2014.
42. Friedman M. Piecewise Exponential Models for Survival Data with Covariates. *The Annals of Statistics*. 1982;10(1):101-13.
43. Breslow NE. Analysis of Survival Data under the Proportional Hazards Model. *International Statistical Review / Revue Internationale de Statistique*. 1975;43(1):45-57.
44. Sjukskrivningar 60 dagar eller längre. En beskrivning av sjukskrivna åren 1999-2014 efter kön, ålder, arbetsmarknadsstatus, yrke, sjukskrivningslängd och diagnospanorama. Stockholm: Försäkringskassan, 2015.
45. Socialförsäkringen i siffror. Stockholm: Försäkringskassan, 2015.
46. Sjukfrånvarons utveckling. Delrapport 2. Stockholm: Försäkringskassan, 2015.
47. Söderberg E, Lindholm C, Kärrholm J, Alexanderson K. Läkares sjukskrivningspraxis; en systematisk litteraturöversikt: Socialdepartementet; SOU 2010:107; 2010.
48. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27(2):157-72; Comments 207-12.
49. Geisser S. Predictive Inference: An Introduction. New York: Chapman & Hall; 1993.
50. Gerds T, Cai T, Schumacher M. The performance of risk prediction models. *Biometrical Journal*. 2008;50(4):457-79.
51. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer; 2008.
52. Breiman L. Random Forests. *Machine Learning*. 2001;45(1):5-32.
53. Liaw A, Wiener M. Classification and Regression by random Forest. *R News*. 2002;2(3):18-22.
54. Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*. 1996;58(1):267-88.
55. Harrell F. Regression Modeling Strategies With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis: Springer International Publishing Switzerland; 2015.
56. Bottai M, Geraci M, Lawson A. Testing for unusual aggregation of health risk in semiparametric models. *Stat Med*. 2008;27(15):2902-21.
57. Bottai M, Pistelli F, Di Pede F, Baldacci S, Simoni M, Maio S, Carrozzi L, Viegi G. Percentiles of inspiratory capacity in healthy nonsmokers: a pilot study. *Respiration*. 2011;82(3):254-62.
58. Pistelli F, Bottai M, Viegi G, Di Pede F, Carrozzi L, Baldacci S, Pedreschi M, Giuntini C. Smooth reference equations for slow vital capacity and flow-volume curve indexes. *Am J Respir Crit Care Med*. 2000;161(3 Pt 1):899-905 (Erratum in: *Am J Respir Crit Care Med* 2001 Nov 1;164(9):1740. PubMed PMID: 10712340).
59. Pistelli F, Bottai M, Carrozzi L, Baldacci S, Simoni M, Di Pede F, Viegi G. Reference equations for spirometry from a general population sample in central Italy. *Respir Med*. 2007;101(4):814-25.
60. Rothman KJ, Lash TL, Greenland S. Modern Epidemiology. Third Edition (Revised) ed: Lippincott, Williams, & Wilkins 2012.
61. Stone M. Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society*. 1974;36(11):1-147.

62. Efron B, Tibshirani R. Improvements on Cross-Validation: The .632+ Bootstrap Method. *Journal of the American Statistical Association*. 1997;92(438):548-60.
63. Gilchrist W. *Statistical Forecasting*. London: John Wiley & Sons; 1976.
64. Larsson J. Ett litet steg för läkarkåren - ett stort steg för sjukskrivningsprocessen (temanummer: Försäkringsmedicin - i gränssnittet mellan medicin och försäkring). *Socialmedicinsk tidsskrift*. 2009;86(6):529-38.
65. Riksförsäkringsverket. *Försäkringsmedicin*. Stockholm: Försäkringskassaförbundet; 1988.
66. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekblom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol*. 2009;24(11):659-67.
67. Ludvigsson JF, Andersson E, Ekblom A, Feychting M, Kim JL, Reuterwall C, Heurgren M, Olausson PO. External review and validation of the Swedish national inpatient register. *BMC Public Health*. 2011;11:450.
68. Svenska register en unik resurs för hälsa och välfärd. Stockholm: 2013.
69. de Boer AGEM. The European Cancer and Work Network: CANWON. *Journal of occupational rehabilitation*. 2013;23(3):393-8.
70. de Boer A, Taskila T, Tamminga S, Feuerstein M, Frings-Dresen M, Verbeek J. Interventions to enhance return-to-work for cancer patients (Review). *The Cochrane Library* 2015, Issue 9: The Cochrane Collaboration, 2015.
71. Bradley CJ, Neumark D, Oberst K, Luo Z, Brennan S, Schenk M. Combining registry, primary, and secondary data sources to identify the impact of cancer on labor market outcomes. *Med Decis Making*. 2005;25(5):534-47.
72. Wennman-Larsen A, Olsson M, Alexanderson K, Nilsson M, Petersson L. Arm morbidity and sick leave among working women shortly after breast cancer surgery. *European Journal of Oncology Nursing*. 2012.
73. Petersson LM, Wennman-Larsen A, Nilsson M, Olsson M, Alexanderson K. Work situation and sickness absence in the initial period after breast cancer surgery. *Acta Oncol*. 2011;50(2):282-8.
74. Läkarintygens betydelse för sjukfrånvaron. En registerstudie av till Försäkringskassan inkomna läkarintyg. Inspektionen för socialförsäkringen, 2014 17.

10 Bilaga

Bilaga tabell 1. Skattade koefficienter inklusive standardfel, z-värden och p-värden för den framtagna logistiska regressionsmodellen att predicera risk för ny sjukskrivning för kvinnor <65 år med en första bröstcancerdiagnos år 2010, bland dem som inte redan var sjukskrivna eller hade sjuk- eller aktivitetsersättning vid diagnos, N=2954¹. Ns står för natural cubic spline.

	Skattad koefficient	Standardfel	z-värde	p-värde
Intercept	-29,921	922,125	-0,032	0,974
Ålder				
ns(Ålder) ^a	-1,176	0,284	-4,142	<0,001
ns(Ålder) ^b	-3,083	0,933	-3,304	<0,001
ns(Ålder) ^c	-1,131	0,183	-6,191	<0,001
Utbildningsnivå ²				
Grundskola	0			
Gymnasium	-0,127	0,136	-0,930	0,352
Eftergymnasial	-0,397	0,140	-2,824	0,005
Familjesituation ²				
Gift/sambo utan hemmaboende barn	0			
Gift/sambo med hemmaboende barn	0,129	0,123	1,046	0,296
Ensamstående utan hemmaboende barn	-0,083	0,169	-0,490	0,624
Ensamstående med hemmaboende barn	0,091	0,197	0,463	0,644
H-region ²				
Storstad	0			
Mellanstor stad	1,308	0,128	10,225	<0,001
Landsbygd	1,096	0,133	8,270	<0,001
Födelseland				
Sverige	0			
Övriga länder	-0,095	0,121	-0,791	0,429
Sektor ²				
Ej förvärvsarbetande	0			
Offentligt anställd	2,369	0,175	13,516	<0,001
Privat anställd	1,909	0,171	11,134	<0,001
Civilstånd ²				
Ogift, skild, änka/änkling	0			
Gift, registrerat partnerskap	-0,142	0,134	-1,058	0,290
Boenderegion ²				
Nord	0			
Mitt	-0,332	0,164	-2,030	0,042
Stockholm	0,964	0,181	5,329	<0,001
Väst	0,258	0,149	1,725	0,085
Syd	0,211	0,148	1,424	0,154
Cancerstadium				
TONOM0 + Stadium 0+I	0			
Stadium II-IV	-0,480	0,085	-5,647	<0,001

*Forts Bilaga tabell 1*Tidigare sjukskrivning, nettodagar^{3,4}

Ingen tidigare sjukskrivning	0,285	0,579	0,493	0,622
ns(mer än noll nettodagar) ^a	-0,473	0,607	-0,780	0,435
ns(mer än noll nettodagar) ^b	1,719	0,702	2,450	0,014
ns(mer än noll nettodagar) ^c	0,499	0,884	0,565	0,572

Tidigare sjukskrivning, diagnoser³

Psykiska diagnoser	-0,012	0,504	-0,023	0,981
Övriga diagnoser	0,008	0,536	0,015	0,988

Tidigare sjuk- eller aktivitetsersättning, nettodagar³

Ingen tidigare sjuk- eller aktivitetsersättning	16,202	882,744	0,018	0,985
ns(mer än noll nettodagar) ^a	0,790	0,797	0,991	0,322
ns(mer än noll nettodagar) ^b	4,292	2,506	1,713	0,087
ns(mer än noll nettodagar) ^c	1,827	1,197	1,527	0,127

Tidigare sjuk- eller aktivitetsersättning, diagnoser³

Psykiska diagnoser	14,480	882,744	0,016	0,987
Övriga diagnoser	14,622	882,744	0,017	0,987

Tidigare besök i öppenvården³

Inga besök	-0,010	0,447	-0,023	0,982
ns(ett eller flera besök) ^a	0,456	0,343	1,329	0,184
ns(ett eller flera besök) ^b	-0,912	0,935	-0,975	0,330
ns(ett eller flera besök) ^c	-1,103	0,731	-1,509	0,131

Tidigare besök i öppenvården, diagnoser³

Psykiska diagnoser	0,403	0,308	1,307	0,191
Övriga diagnoser	0,300	0,223	1,350	0,177

Tidigare slutenvård³

Ingen slutenvård	12,477	266,602	0,047	0,963
ns(en eller flera dagar) ^a	-1,162	0,779	-1,490	0,136
ns(en eller flera dagar) ^b	-3,100	2,011	-1,541	0,123
ns(en eller flera dagar) ^c	-1,106	1,201	-0,921	0,357

Tidigare slutenvård, diagnoser³

Psykiska diagnoser	-0,666	1,301	-0,512	0,608
Övriga diagnoser	13,633	266,600	0,051	0,959

¹Ytterligare exkludering av 61 kvinnor gjordes, dels på grund av att värden saknades i cancerregistret för de variabler som skulle ingå i modelleringen och dels på grund av extrema värden på vissa av de kontinuerliga variablerna. Ytterligare en orsak till exkludering var att inte ha bott i Sverige under åren före diagnos (T₀).

²Gällande december 2009.

³Med tidigare avses två år innan diagnosdatum fram till dag 15 innan diagnosdatum.

⁴De 14 första dagarna i samtliga sjukskrivningsfall är ej medräknade.

^{a,b,c}Koefficienter för splines funktioner.

Bilaga tabell2. Skattade koefficienter inklusive standardfel, z-värden och p-värden för den framtagna logistiska regressionsmodellen att predicera risk för ny sjukskrivning för personer mellan 16-65 år gamla med en handlovsfraktur år 2010, bland dem som inte redan var sjukskrivna eller hade sjuk- eller aktivitetsersättning vid diagnos, N=14 465. Ns står för natural cubic spline.

	Skattad koefficient	Standardfel	z-värde	p-värde
Intercept	-5,415	7,389	-0,733	0,464
Kön				
Män	0			
Kvinnor	-0,052	0,056	-0,931	0,352
Ålder				
ns(Ålder) ^a	0,298	0,132	2,252	0,024
ns(Ålder) ^b	1,128	0,391	2,884	0,004
ns(Ålder) ^c	0,359	0,118	3,050	0,002
Födelseland				
Sverige	0			
Övriga länder	0,122	0,068	1,776	0,076
Civilstånd ¹				
Ogift, skild, änka/änkling	0			
Gift, registrerat partnerskap	-0,041	0,075	-0,554	0,580
Familjesituation ¹				
Gift/sambo utan hemmaboende barn	0			
Gift/sambo med hemmaboende barn	-0,084	0,093	-0,908	0,364
Ensamstående utan hemmaboende barn	-0,103	0,108	-0,952	0,341
Ensamstående med hemmaboende barn	-0,103	0,143	-0,722	0,470
Hemmaboende <20 år	-0,224	0,172	-1,299	0,194
H-region ¹				
Storstad	0			
Mellanstor stad	0,115	0,072	1,608	0,108
Landsbygd	0,224	0,072	3,121	0,002
Boenderegion ¹				
Nord	0			
Mitt	0,087	0,087	1,002	0,316
Stockholm	-0,087	0,101	-0,862	0,389
Väst	0,049	0,078	0,624	0,532
Syd	0,015	0,079	0,189	0,850
Utbildningsnivå ¹				
Grundskola	0			
Gymnasium	0,150	0,062	2,437	0,015
Eftergymnasial	-0,202	0,086	-2,356	0,018
Sektor ¹				
Ej förvävsarbetande	0			
Offentligt anställd	1,701	0,100	16,972	<0,001
Privat anställd	1,610	0,081	19,808	<0,001

Forts Bilaga tabell 2

Typ av yrke ¹				
Inget yrke	0			
Tjänsteman	-0,693	0,110	-6,314	<0,001
Arbetare	0,497	0,088	5,643	<0,001
S62 som huvuddiagnos	1,186	0,118	10,026	<0,001
Antal operationskoder				
0	0			
1-4	0,271	0,049	5,572	<0,001
5-30	0,183	0,211	0,865	0,387
Slutenvårdad vid vårdtillfället	0,823	0,120	6,877	<0,001
Tidigare handledsfraktur >2 år sedan	-0,088	0,091	-0,974	0,330
Tidigare sjukskrivning ² , nettodagar ³				
Ingen tidigare sjukskrivning	-0,468	0,345	-1,356	0,175
ns(mer än noll nettodagar) ^a	-0,319	0,424	-0,752	0,452
ns(mer än noll nettodagar) ^b	-0,559	0,442	-1,266	0,206
ns(mer än noll nettodagar) ^c	-0,640	0,725	-0,883	0,377
Tidigare sjukskrivning ² , diagnoser				
Psykiska diagnoser	0,213	0,299	0,711	0,477
Övriga diagnoser	0,171	0,323	0,530	0,596
Tidigare sjuk- eller aktivitetsersättning ² , nettodagar				
Ingen tidigare sjuk- eller aktivitetsersättning	1,009	1,254	0,805	0,421
ns(mer än noll nettodagar) ^a	-0,746	0,578	-1,290	0,197
ns(mer än noll nettodagar) ^b	-2,878	1,698	-1,695	0,090
ns(mer än noll nettodagar) ^c	-2,493	1,700	-1,467	0,142
Tidigare sjuk- eller aktivitetsersättning ² , diagnoser				
Psykiska diagnoser	1,798	1,105	1,628	0,104
Övriga diagnoser	1,720	1,133	1,517	0,129
Tidigare besök i öppenvården ²				
Inga besök	-0,090	0,322	-0,279	0,781
ns(ett eller flera besök) ^a	-0,356	0,371	-0,960	0,337
ns(ett eller flera besök) ^b	-0,815	0,879	-0,928	0,353
ns(ett eller flera besök) ^c	-0,917	1,563	-0,587	0,557
Tidigare besök i öppenvården ² , diagnoser				
Psykiska diagnoser	0,025	0,149	0,165	0,869
Övriga diagnoser	-0,113	0,254	-0,446	0,656
Tidigare slutenvård ²				
Ingen slutenvård	0,468	7,255	0,064	0,949
ns(en eller flera dagar) ^a	-1,635	4,469	-0,366	0,715
ns(en eller flera dagar) ^b	-0,905	12,042	-0,075	0,940
ns(en eller flera dagar) ^c	N/A			
Tidigare slutenvård ² , diagnoser				
Psykiska diagnoser	-0,190	0,431	-0,441	0,659
Övriga diagnoser	0,356	0,495	0,719	0,472

¹Gällande december 2009. ²Med tidigare avses två år innan datum för handledsfrakturen. ³De 14 första dagarna i samtliga sjukskrivningsfall är ej medräknade. ^{a,b,c}Koefficienter för splines funktioner.