

# Utvärdering av behandlingsprogrammet PRISM i Kriminalvården

VI BRYTER DEN ONDA CIRKELN



# Sammanfattning

## Bakgrund

PRISM är ett behandlingsprogram mot missbruksproblem bland kriminalvårdsklienter. Programmet bygger på principer från kognitiv beteendeterapi (KBT) och genomförs individuellt. PRISM utvecklades i Storbritannien och infördes i Sverige år 2003. Kriminalvården har tidigare utvärderat om programmet minskade återfall i brott bland programdeltagare 2003–2006. I den utvärderingen hade klienter som fullföljt PRISM signifikant lägre risk att återfalla i brott jämfört med klienter som inte deltagit i något behandlingsprogram. Däremot syntes ingen riskminskning bland samtliga som påbörjat programmet. Inga andra utvärderingar av PRISM:s effekt hittades. Däremot finns många utvärderingar av andra kriminalvårdsprogram mot missbruk. När resultatet av sådana utvärderingar vägdes samman i en översikt verkade missbruksprogram i genomsnitt minska återfall i brott med 20 % (från 35 % till 28 %). När utvärderingsresultat för KBT-baserade kriminalvårdsprogram, däribland missbruksprogram, vägdes samman i en annan översikt verkade de i genomsnitt minska återfall i brott med 25 % (från 40 % till 30 %). Den här utvärderingens syfte var att undersöka om PRISM minskar risken för missbruk och återfall i ny brottslighet bland kriminalvårdsklienter i Sverige.

## Metod

I utvärderingen ingick 654 klienter som påbörjat PRISM och som frigavs från anstalt eller skrevs in i frivård 2007–2009. Av dessa var det 68 % som fullföljde programmet. De jämfördes med fem gånger fler individuellt matchade klienter aktuella under samma period som inte påbörjat programmet. För att minska risken att bakgrundsskillnader gör resultaten missvisande användes så kallad propensity score-matchning. Det innebär att vi hanterade eventuella gruppsskillnader före programmet i till exempel ålder, antal tidigare domar och alkohol- eller narkotikaberoende.

## Resultat

Klienter som fullföljde PRISM hade signifikant lägre risk att återfalla i brott än en matchad jämförelsegrupp. Andelen som återföll i nya brott inom ett år var 58 % bland de som fullföljde programmet och 64 % i den matchade jämförelsegruppen. Ifall även de som påbörjat PRISM hade signifikant lägre risk att återfalla i brott berodde dock på vilket register som användes. Ingen statistiskt säkerställd minskad risk att vårdas för missbruk hittades, vare sig för de som fullföljt eller påbörjat programmet. Tvärt emot övriga resultat hade klienter som avbrutit högre risk både att återfalla i brott och vårdas för missbruk jämfört med klienter med liknande bakgrund.

## Diskussion

Sett tillsammans med tidigare studier på liknande program kan resultaten tolkas som att PRISM kan minska återfall i brott. Dock får man komma ihåg att resultat bör tolkas med viss försiktighet, det kan t.ex. fortfarande finnas bakgrundsskillnader mellan programdeltagare och jämförelsegrupp som snedvrider resultatet, trots att många viktiga bakgrundsfaktorer hanterades.

Det är svårt att dra några slutsatser om PRISM:s effekt på missbruk baserat på utvärderingen. Den indikator på missbruk som användes i utvärderingen är trubbig och fångar därför inte alla potentiell förändring i deltagarnas missbruk. Därför skulle ett mer känsligt mått behöva användas i en ny utvärdering, till exempel självrapporterad konsumtion av alkohol och andra droger.

# Innehållsförteckning

<b>Introduktion</b> .....	<b>1</b>
Tidigare forskning .....	2
Problemformulering och syfte.....	3
<b>Metod</b> .....	<b>4</b>
Undersökningsgrupp.....	4
Källor.....	4
Utfall.....	4
Bakgrundsfaktorer .....	5
Statistisk analys.....	6
<b>Resultat</b> .....	<b>7</b>
Huvudresultat.....	7
Känslighetsanalys.....	8
<b>Diskussion</b> .....	<b>9</b>
<b>Referenser</b> .....	<b>11</b>
<b>Bilagor</b> .....	<b>13</b>
Bilaga A. Metodfördjupning.....	13
Bilaga B. Bakgrundsfaktorer före och efter matchning.....	16

# Introduktion

PRISM (Programme for Reducing Individual Substance Misuse) är ett program mot missbruk bland kriminalvårdsklienter. Behandlingsprogrammet skapades av Philip Priestley, kriminolog, och Mary McMurren, psykologiprofessor. Målet med PRISM är att minska eller eliminera alkohol- och droganvändning och kriminalitet, samt att generellt verka för bättre hälsa. Programmet är avsett för missbruk av både olagliga och lagliga droger, t.ex. heroin och alkohol (Priestley & McMurren, 2011). Att missbruk ökar risken för kriminalitet är väletablerat (Bennett, Holloway, & Farrington, 2008). Genom att förändra deltagarnas konsumtion av alkohol och andra droger är tanken därför att man också kommer att minska risken för återfall i brott (McMurren & Priestley, 2003).

Programmet bygger på principer från kognitiv beteendeterapi (KBT) (McMurren & Priestley, 2003). PRISM innehåller bl.a. övningar i problemlösning, självkontroll, social färdighetsträning, kognitiv omstrukturering och återfallsprevention. Deltagarna uppmuntras även till livsstilsförändring och för att öka deras motivation används samtalsmetoden motiverande intervju (MI) (McMurren & Priestley, 2003; Priestley & McMurren, 2011).

PRISM består av tre delar: kartläggning, färdighetsträning och tillämpning/återfallsprevention. Först samlar man in information om klienterna och försöker öka deras motivation till förändring (McMurren & Priestley, 2003). Därefter får deltagarna lära sig färdigheter som ska hjälpa dem undvika att använda droger. I den sista delen får deltagarna tillämpa sina nya färdigheter på faktorer som bidrar till deras droganvändning och relaterad kriminalitet. PRISM är uppdelat i ett förmöte och 20 efterföljande möten där programledare och klient träffas enskilt under cirka 1,5 timme. Som minst bör PRISM utföras i tre månader och som mest nio, men sex månader är idealet (Priestley & McMurren, 2011).

Programledaren genomför PRISM enskilt med klienten. Det kan finnas flera skäl till att ge behandling enskilt, t.ex. personliga egenskaper som inte skulle fungera i grupp eller känsliga brottsdetaljer som inte bör tas upp i grupp (McMurren & Priestley, 2003). Det kan också vara svårt att få ihop tillräckligt många klienter för att starta en gruppbehandling, vilket ger upphov till väntetid (Priestley & Edström, 2010). PRISM ges av personer som utbildats i programmet och som måste certifieras inom ett år efter utbildningen för att få fortsätta genomföra detsamma. Certifiering innebär att programmets utförande granskas, bl.a. via videospelningar (Kriminalvården, 2009). Programledare som genomför PRISM får även löpande metodhandledning (Kriminalvården, 2008).

För varje ny kriminalvårdsklient upprättas en s.k. verkställighetsplan, där personal identifierar klientens risker och behov. Om det framkommer att klienten har behov av behandling föreslås ett lämpligt behandlingsprogram som man försöker motivera denne att delta i. Kriminalvårdens behandlingsprogram brukar alltså vara frivilliga, men för frivårdsklienter kan behandling ibland vara en del av påföljden.<sup>1</sup> PRISM utvecklades för frivården, men används i Sverige även på anstalt. Målgruppen för PRISM är kriminella med måttlig eller hög risk att återfalla i brott som dessutom har ett alkohol-/drogmissbruk som bidrar till deras kriminalitet (McMurren & Priestley, 2003). Det är dock inte uttalat att något specifikt bedömningsinstrument ska användas för att bedöma risken för återfall i brott. Klienter kan exkluderas från programmet p.g.a. allvarliga läs- och skrivsvårigheter, språksvårigheter eller pågående psykisk sjukdom (annat än missbruk/beroende) som kan vara till hinder för att tillgodogöra sig behandlingen (Kriminalvården, 2013).

PRISM började användas i svensk kriminalvård år 2003 och ackrediterades år 2006. Ackreditering innebär att programmet granskas av en panel med personer från Kriminalvårdens vetenskapliga råd. För att ackrediteras ska programmet bl.a. ha en vetenskapligt understödd förändringsmodell, tydlig målgrupp, utarbetade manualer samt vara riktat mot dynamiska riskfaktorer för kriminalitet. Tanken är att ackrediterade behandlingsprogram också ska utvärderas fortlöpande avseende effekter på återfall i brott (Kriminalvården, 2007).

---

<sup>1</sup> Inför rättegång kan frivårdspersonal genomföra en s.k. personutredning. Om det framkommer att det finns behov och förutsättningar för behandling kan det bli en del av påföljden. Detta kommer dock endast i fråga om personen också själv vill delta i behandling. Oftast rör det sig om personer med ett missbruk som tycks bidra till kriminalitet.

## Tidigare forskning

PRISM har tidigare utvärderats i Kriminalvården (Danielsson, Fors, & Freij, 2012). Deltagare i PRISM (n=380) jämfördes då med andra kriminalvårdsklienter som inte deltagit i något av Kriminalvårdens behandlingsprogram (n=1140). Resultatet justerades för bakgrundsskillnader som det fanns information om i Kriminalvårdens register och återfall i brott definierades som ny kriminalvårdspåföljd.<sup>2</sup> Klienter som påbörjat PRISM hade något högre risk att återfalla i brott än jämförelsegruppen men skillnaden var inte statistiskt säkerställd (hazardkvot [HR] 1,16, 95% konfidensintervall [KI] 0,96-1,39). De som fullföljt PRISM hade däremot signifikant lägre risk att återfalla i brott än jämförelsegruppen (HR 0,70, 95% KI 0,52-0,94). Att de som fullföljde PRISM hade lägre risk att återfalla i brott kan bero på att programmet minskar återfall i brott om det utförs i sin helhet. Resultatet skulle dock också kunna bero på att klienter som fullföljer behandlingsprogram har egenskaper som både ökar benägenheten att slutföra program och minskar risken att återfalla i brott. Exempelvis brukar klienter som fullföljer behandling vara mer motiverade redan innan programstart (Olver, Stockdale, & Wormith, 2011).

Vi hittade inte någon annan utvärdering av PRISM:s eventuella effekt på återfall i brott. PRISM är dock en individuell variant av grupprogrammet ASRO (Addressing Substance Related Offending) (McGuire, 2008).<sup>3</sup> I en engelsk utvärdering bestod programgruppen av 131 frivårdsklienter med obligatoriskt deltagande i ASRO som en del av påföljden (Palmer, Hatcher, McGuire, Bilby, & Ayres, 2011). Programdeltagarna jämfördes med 178 andra frivårdsklienter. Författarna hanterade vissa bakgrundsskillnader statistiskt.<sup>4</sup> De fann att de 41 klienter som fullföljt ASRO återföll i brott ungefär lika mycket som jämförelsegruppen. De 100 klienter som påbörjat men sedan avbrutit programmet återföll däremot signifikant oftare än jämförelsegruppen. Forskarna redovisade inte hur resultatet såg ut för samtliga som påbörjat programmet, men det borde innebära att programdeltagarna återföll i brott oftare än jämförelsegruppen. Författarna drog slutsatsen att bristen på positiv programeffekt, åtminstone delvis, kunde bero på att behandlingen var en obligatorisk del av straffet.

I nuläget finns få utvärderingar av om PRISM och systerprogrammet ASRO minskar återfall i brott.<sup>5</sup> Dessutom motsäger utvärderingarna delvis varandra. Det finns dock många studier på andra missbruksprogram för kriminella, vilka har sammanfattats i systematiska kunskapsöversikter med metaanalyser. Metaanalys innebär att man gör en samlad bedömning av flera enskilda studier genom att statistiskt sammanföra deras resultat. Missbruksprogram för kriminella verkar i genomsnitt minska återfall i brott med 20% (från 35% till 28%), men effekten varierar mellan olika typer av program (Mitchell, Wilson, & MacKenzie, 2012). I en systematisk översikt av KBT-baserade program för kriminella, däribland missbruksprogram, fann man att sådan behandling i genomsnitt minskar återfall i brott med 25% (från 40% till 30%) (Lipsey, Landenberger, & Wilson, 2007).

---

<sup>2</sup> Bakgrundsfaktorer som hanterades var: ålder, kön, tid i anstalt, tidigare verkställigheter, missbruk samt diverse brottstyper.

<sup>3</sup> PRISM liknar även det brittiska programmet Offender Substance Abuse Programme (OSAP), som dock inte verkar ha utvärderats ännu (McGuire, 2008).

<sup>4</sup> De hanterade bakgrundsfaktorerna: ålder, antal tidigare domar, brottstyp och resultat på riskinstrumentet OGRS2

<sup>5</sup> Två ytterligare utvärderingar av ASRO hittades men uteslöts ur genomgången av tidigare studier (Hollin et al., 2004; V. Hollis, 2007). I den första utvärderingen undersöktes effekten av flera behandlingsprogram, varav ASRO var ett (Hollin et al., 2004). Programmen redovisades dock inte separat och enbart 2,8% av programgruppen hade deltagit i ASRO. Resultaten säger därmed mer om de övriga behandlingsprogrammen än om ASRO och redovisas därför inte här. I den andra utvärderingen fanns ingen jämförelsegrupp, utan istället jämfördes deltagarnas faktiska brottsåterfall med deras predicerade brottsåterfall (V. Hollis, 2007). Utvärderingen uteslöts för att risken är stor att resultatet är missvisande när jämförelsegrupp saknas (Weisburd, Lum, & Petrosino, 2001).

## **Problemformulering och syfte**

I Kriminalvårdens förra utvärdering valde en hög andel av dem som påbörjade PRISM att avbryta programmet. Eventuella positiva programeffekter på de som påbörjat behandling kan försvagas om andelen som avbryter behandling är hög, eftersom de därmed exponerats mindre för programmet. Sedan förra utvärderingen har andelen som fullföljer PRISM ökat och det är bland annat därför angeläget med en ny utvärdering av programmet.

Det är viktigt att undersöka återfall i brott i effektutvärderingar av Kriminalvårdens program. För att ett program ska kunna minska återfall i brott måste det påverka faktorer av betydelse för risken för återfall i brott, t.ex. missbruk eller prokriminell attityd (Bonta & Andrews, 2007). Utfall som potentiellt är kopplade till brottsåterfall är därför också angelägna att utvärdera. Syftet med denna studie är därför att undersöka om PRISM minskar missbruk och återfall i brott bland kriminalvårdsklienter.

Studiens primära frågeställning är:

- Har klienter som påbörjat respektive fullföljt PRISM minskad risk för missbruk och återfall i brott jämfört med klienter som inte påbörjat programmet?

Dessutom undersöks den sekundära frågeställningen:

- Har de som påbörjat men sedan avbrutit PRISM ökad risk för återfall i brott och missbruk jämfört med klienter som inte påbörjat programmet?

## Metod

### Undersökningsgrupp

Programgruppen bestod av kriminalvårdsklienter som skrevs in i frivården eller ut från anstalt åren 2007 till 2009 och påbörjade PRISM under påföljden (n=654). De jämfördes med kriminalvårdsklienter aktuella under samma period men som inte deltagit i PRISM. För att minska risken för systematiska bakgrundsskillnader mellan grupperna användes propensity score-matchning (se sida 6).

Att jämföra de som *påbörjat* behandling med en jämförelsegrupp kallas för intention-to-treat. Både de som fullföljt och de som avbrutit behandlingen ingår alltså i undersökningen. När man erbjuder behandling kan man förstås inte på förhand säkert veta om en enskild person kommer att fullfölja programmet eller inte. Oönskade effekter av behandling kan dessutom leda till att deltagare väljer att avbryta denna. Att då exkludera dem som avbryter behandling från undersökningen skulle innebära att man sällar bort programdeltagare som drabbats av oönskade effekter. Det har därmed praktisk relevans vilken effekt behandling har på samtliga som påbörjat behandling. Nackdelen med att utgå ifrån de som påbörjat behandling är förstås att alla inte tagit del av *hela* behandlingen. Att ta del av hela behandlingssinnehållet kan vara en förutsättning för önskvärda effekter och i så fall skulle analysen enligt intention-to-treat underskatta eventuella positiva programeffekter, särskilt om andelen som avbrutit behandlingen är hög. Vi jämförde därför även de som *fullföljt* PRISM med en matchad jämförelsegrupp, vilket kallas för per-protokoll. Nackdelen med detta är att de som lyckas fullfölja behandling ofta har bättre förutsättningar redan innan behandlingen, t.ex. genom att vara mer motiverade att förändra sina liv. Med per-protokoll ökar därmed risken att behandlingen ser ut att ha positiva effekter när resultatet egentligen beror på bakgrundsskillnader. Sammanfattningsvis finns för- och nackdelar med både intention-to-treat och per-protokoll och det är därför som vi presenterar resultat från båda analyserna (Debidin & Lovbakke, 2005; S. Hollis & Campbell, 1999; J. A. Lewis & Machin, 1993).

### Källor

Undersökningen bygger på en avidentifierad databas med länkning mellan Kriminalvårdens register och andra nationella befolkningsregister.<sup>6</sup> Följande register användes: lagföringsregistret, misstankeregistret, Kriminalvårdens register, patientregistret, longitudinell integrationsdatabas för arbetsmarknadsstudier (LISA), folk- och bostadsräkningarna (FoB), flergenerationsregistret, registret över totalbefolkningen, migrationsregistret och dödsorsaksregistret.

### Utfall

Vård för missbruk och återfall i brott utgjorde utfall. Vård för missbruk användes som indikator på missbruk. Uppgifter om vård för missbruk hämtades från patientregistret och definierades som vård för akut intoxikation av alkohol eller andra droger alternativt inskrivning i slutna beroendevård.<sup>7</sup> Återfall i något nytt brott följdes i första hand upp i misstankeregistret. I misstankeregistret ingår enbart personer som av förundersökningsledare bedömts som minst skäligen misstänka för brott. Skälig misstanke kan ligga till grund för bl.a. anhållande och häktning. Internationellt är det vanligt att i behandlingsutvärderingar definiera återfall i brott som återarrestering (Lipsey et al., 2007).<sup>8</sup> Det verkar inte ha betydelse för resultaten om återfall i brott definieras som ny arrestering eller ny dom (Lipsey et al.,

---

<sup>6</sup> Registerinformation hämtades från databasen "Crime-2" vid Institutionen för medicinsk epidemiologi och biostatistik vid Karolinska Institutet. Detta gjordes inom ramen för en forskningsstudie som godkänns av Regional forskningsetisk kommitté vid Karolinska Institutet (2009/939-31/5). Databasen innehåller länkade, landsomfattande, longitudinell information från ett flertal befolkningsbaserade register, men är anonymiserad genom att personuppgifter är borttagna.

<sup>7</sup> Slutna beroendevård är avsedd för avgiftning/abstinensbehandling och indikerar därmed ett pågående missbruk.

<sup>8</sup> Arrestering finns inte som juridiskt begrepp i Sverige men motsvaras kanske bäst av gripande, eftersom båda innebär tillfälliga frihetsberövanden som polisen har rätt att besluta om vid misstanke om brott.

2007; Mitchell et al., 2012). Misstankeregistret användes för att det ger större statistisk power,<sup>9</sup> d.v.s. ökad sannolikhet att statistiskt säkerställa en faktisk effekt, och för att det gav ett år längre uppföljning än lagföringsregistret. Det tar också kortare tid för brott att registreras i misstankeregistret jämfört med lagföringsregistret. Deltagarna följdes även upp i lagföringsregistret då resultatens tillförlitlighet generellt anses stärkas om de kan bekräftas med en annan källa (Latessa & Holsinger, 1998; Lösel, 2001).

Programgruppen följdes upp med start från utskrivning från anstalt eller från slutdatum för programmet om de genomförde PRISM i frivård. Jämförelsegruppen följdes upp med start från utskrivning från anstalt eller från inskrivning i frivård. För att få reda på när klienten inte längre kunde följas upp p.g.a. utvandring eller dödsfall användes migrationsregistret och dödsorsaksregistret. Vård för missbruk följdes upp fram till år 2009-12-31, vilket gav en genomsnittlig uppföljningstid på 549 dagar (minst 0 dagar, maximalt 1095 dagar). Återfall i brott följdes upp i misstankeregistret fram till 2010-12-31, vilket gav en genomsnittlig uppföljningstid på 906 dagar (minst 0 dagar, maximalt 1460 dagar). I lagföringsregistret kunde återfall i brott följas upp fram till 2009-12-31, vilket gav samma uppföljningstid som för patientregistret.<sup>10</sup>

## Bakgrundsfaktorer

I studier där programgruppen inte är tillräckligt lik jämförelsegruppen före behandling riskerar resultatet att bli snedvridet (eng.: biased). I utvärderingar där det inte är slumpmässigt vem som deltar i programmet bör man därför försöka hantera störfaktorer (eng.: confounders), d.v.s. gruppkillnader före programmet som påverkar det man undersöker effekt på. Tänkbara störfaktorer i den här undersökningen presenteras i tabell 1. De faktorer i tabell 1 som i inledande analyser samvarierade såväl med programdeltagande som med utfallen (t.ex. återfall i brott) bedömdes vara störfaktorer och hanterades därför statistiskt i undersökningen.

**Tabell 1.** Bakgrundsfaktorer i undersökningen.

Register	År	Bakgrundsfaktorer
Registret över totalbefolkningen	1968-2009	Ålder, kön, bostadsområde
Lagföringsregistret	1973-2009	Antal domar, ålder vid första dom, någon gång dömd för våldsbrott, narkotikabrott, rattfylleri, mamma, pappa eller något syskon någon gång dömda för brott
LISA-databasen	1990-2009	Utbildningsnivå, civilstatus, antal år med arbetsinkomst senaste fem åren
Folk- och bostadsräkningarna (FoB)	1960, 1970, 1980, 1985, 1990	Socioekonomisk status i barndomen <sup>a</sup>
Patientregistret	Slutenvård 1973-2009. Öppenvård 2001-2009	Alkoholmissbruk/-beroende, narkotikamissbruk/-beroende, personlighetsstörning, ADHD, förälder med missbruk/-beroende
Kriminalvårdens register	2002-2009	Påbörjat allmänt kriminalitets-/våldsprogram, annat missbruksprogram än PRISM, motivationsprogram, sexualbrotts-/partnervåldsprogram, anstalt- eller frivårdsklient
Flergenerationsregistret	1961-2009 <sup>b</sup>	

<sup>a</sup> FoB 1965 ingick inte i databasen och socioekonomisk status ingick inte i FoB 1975 (Nilsson & Andersson, 1979).

Socioekonomisk status delades in i låg, medel och hög som i en tidigare studie (Frisell, Lichtenstein, & Långström, 2011).

<sup>b</sup> Personer som varit folkbokförda fr.o.m. 1961 och är födda tidigast 1932 ingår i flergenerationsregistret. Registret kopplades

<sup>9</sup> När man använder utfallsvariabler med två alternativ (ja/nej), påverkas den statistiska povern av antalet händelser, t.ex. återfall i brott. Eftersom misstankeregistret innehåller fler registrerade händelser än lagföringsregistret är den statistiska povern högre.

<sup>10</sup> Noll dagars uppföljningstid innebär att klienten dött eller utvandrat samma dag eller tidigare än uppföljningsstarten. Det kan också bero på att enstaka frivårdsklienter avslutade PRISM år 2010 och därför inte hade någon uppföljningstid i registren som innehåller data fram till 2009 (lagföringsregistret och patientregistret).



## Statistisk analys

För att hantera gruppskillnader före programmet mellan PRISM-gruppen och jämförelsegruppen, som skulle kunna göra resultaten missvisande, användes s.k. propensity score-matchning (Williamson, Morley, Lucan, & Carpenter, 2011). En ”propensity score” kan beskrivas som sannolikheten att gå ett program givet bakgrundsfaktorerna som hanteras i undersökningen (s.k. kontrollvariabler). Sannolikheten att gå programmet (propensity score) uträknas för varje individ utifrån dennes värden på kontrollvariablerna. För varje enskild person i programgruppen matchas därefter en eller flera jämförelsepersoner med propensity scores som ligger så nära programdeltagarens som möjligt. Programgruppen jämförs därmed med personer som inte gått programmet, men som givet sin bakgrund (enligt kontrollvariablerna) borde ha haft ungefär samma sannolikhet att göra det. Se bilaga A för fördjupad information om hur matchningen gick till och tabellerna i bilaga B för att jämföra gruppernas bakgrund före och efter matchning.

I de flesta statistiska analysmetoder utesluts individer som saknar någon data automatiskt från beräkningarna. När många bakgrundsfaktorer hanteras kan det innebära att en hög andel av individerna försvinner från undersökningen, då många har bortfall på någon av faktorerna, vilket sannolikt skulle göra resultatet missvisande. För att hantera sådant bortfall använde vi oss därför av en statistisk metod som kallas multipel imputation. Denna metod har gott forskningsstöd, men bör bara användas om vissa antaganden uppfylls (Schafer, 1999; Scheffer, 2002). När antagandena uppfylldes använde vi multipel imputation, annars tog vi istället hänsyn till bortfallet inom ramen för propensity score-matchningen. Att hantera bortfall på det här sättet innebär inte att försöka få fram vad det sanna värdet är, utan att minska den snedvridning av resultatet som bortfall ofta medför. För närmre beskrivning av hanteringen av bortfall se bilaga A.

För att kunna ta hänsyn till den varierande uppföljningstiden mellan individerna i studien användes Coxregression (Friendship, Street, Cann, & Harper, 2005). Då propensity score-matchning kombineras med Coxregression rekommenderas att man i Coxregressionen jämför varje programdeltagare med den/dem som de matchades med, genom s.k. stratifiering på matchade par (Austin, 2008).<sup>11</sup> Coxregression resulterar i effektstorleksmättet hazardkvot som indikerar programgruppens risk att t.ex. återfalla i brott i förhållande till jämförelsegruppen. En hazardkvot om 1 indikerar att grupperna har lika hög risk att återfalla i brott under uppföljningen. Är hazardkvoten under 1 har programdeltagare lägre risk att återfalla i brott än matchade klienter och är den över 1 har programdeltagare tvärtom förhöjd risk. För att redovisa om det finns en statistiskt säkerställd skillnad används konfidensintervall, där den egentliga gruppskillnaden med 95% sannolikhet finns inom intervallet. Om konfidensintervallen inte innefattar 1 så är skillnaden statistiskt säkerställd på 5%-nivån. Hazardkvoterna redovisas även grafiskt med s.k. ”forest plots” (S. Lewis & Clarke, 2001). Vi rapporterar också andelen med negativt utfall inom ett år i program- respektive jämförelsegrupp.<sup>12</sup> Ett annat sätt att försöka klargöra resultat är att redovisa ”numbers needed to treat” (NNT), d.v.s. antalet klienter som behöver behandlas med PRISM för att t.ex. undvika att en person återfaller i brott (Akobeng, 2005). När det fanns signifikanta skillnader mellan grupperna i huvudresultatet beräknades därför även NNT.<sup>13</sup>

---

<sup>11</sup> Det kan förekomma s.k. specifikationsfel i uträkningen av propensity scores, t.ex. bortfall, mätfel eller utelämnande av relevanta variabler. Men genom att använda ”robust sandwich”-korrigering i Coxregressionen blir standardfelen robustare för specifikationsfel, och därmed blir konfidensintervallen inte felaktigt för smala (Lin & Wei, 1989).

<sup>12</sup> Då många programdeltagare saknade ett helt års uppföljningstid i lagföringsregistret och patientregistret erhöles andelar från Kaplan-Meier-överlevnadskurvor (Bland & Altman, 1998). Även om vi använde Kaplan-Meier innebär den korta uppföljningstiden dock fortfarande att resultatet blir mer osäkert för lagföringsregistret och patientregistret efter ett år (Rich et al., 2010).

<sup>13</sup> NNT räknades ut genom att jämföra gruppernas ”failure rate” enligt Kaplan-Meier, vilket rekommenderats för överlevnadsanalys (Altman & Andersen, 1999). När en programeffekt är negativ kallas det för ”numbers needed to harm” (NNH) istället för ”numbers needed to treat” (NNT).

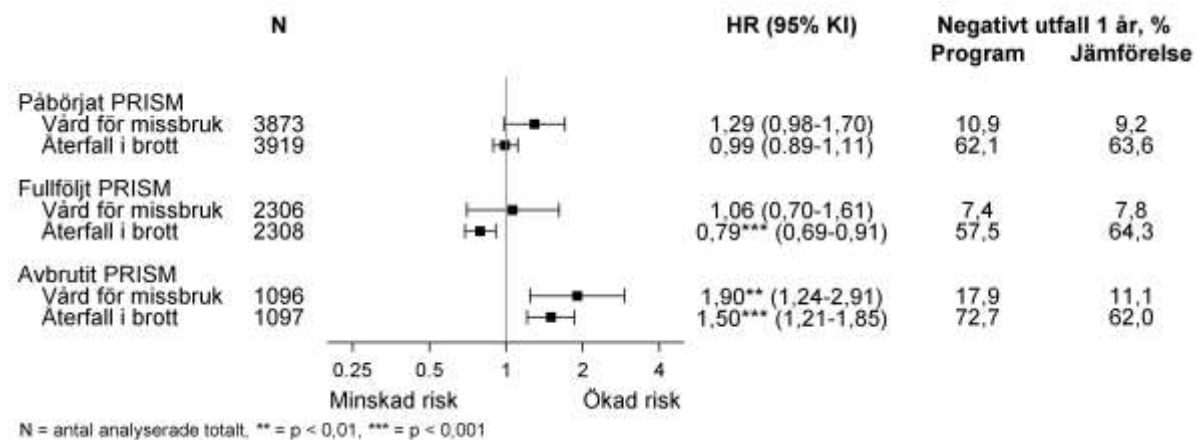
# Resultat

## Huvudresultat

Kriminalvårdsklienter som **påbörjat** PRISM (n=654) jämfördes med individuellt propensity score-matchade klienter som inte påbörjat PRISM (n=3270). Matchningen innebär att grupperna har likvärdig bakgrund på de faktorer som hanterades i undersökningen. Klienter som påbörjat PRISM hade marginellt högre risk att vårdas för missbruk jämfört med den matchade jämförelsegruppen, men skillnaden var inte statistiskt säkerställd (se figur 1). Programgruppens risk att återfalla i brott enligt misstankeregistret var däremot lika hög som för jämförelsegruppen.

Av dem som påbörjat PRISM fullföljde 68 % programmet.<sup>14</sup> Klienter som **fullföljt** PRISM (n=385) jämfördes med en ny matchad jämförelsegrupp (n=1925), eftersom de inte har samma bakgrund som alla som påbörjat programmet. Fullföljare av PRISM hade i stort sett lika hög risk att vårdas för missbruk som matchade klienter. Däremot hade de som fullföljt PRISM statistiskt säkerställd minskad risk att återfalla i brott enligt misstankeregistret. Programfullföljare hade 21 % minskad risk att skäligen misstänkas för nya brott (absolut riskminskning 6,8 procentenheter). Inom ett år innebär det att om 15 klienter fullföljer PRISM kommer 1 person förhindras från att misstänkas för nya brott.<sup>15</sup>

Klienter som påbörjat men sedan **avbrutit** PRISM (n=183) jämfördes med ännu en matchad jämförelsegrupp (n=915). De som avbrutit PRISM hade signifikant högre risk att vårdas för missbruk eller misstänkas för nya brott än jämförelsegruppen. Risken att vårdas för missbruk var 90 % högre och risken att återfalla i brott 50 % högre (absolut riskökning 6,8 respektive 10,7 procentenheter). Det kan uttryckas som att om 15 klienter avbryter PRISM kommer 1 behöva vårdas för missbruk som inte skulle det annars. Motsvarande för återfall i brott innebär att om 9 klienter avbryter PRISM kommer 1 misstänkas för brott inom ett år som inte skulle det utan programmet.



**Figur 1.** Hazardkvoter [HR] med 95% konfidensintervall [KI] för utfallen vård för missbruk och återfall i brott, enligt patientregistret respektive misstankeregistret. Deltagare i PRISM jämfördes med matchade klienter som inte deltagit i programmet.<sup>16</sup>

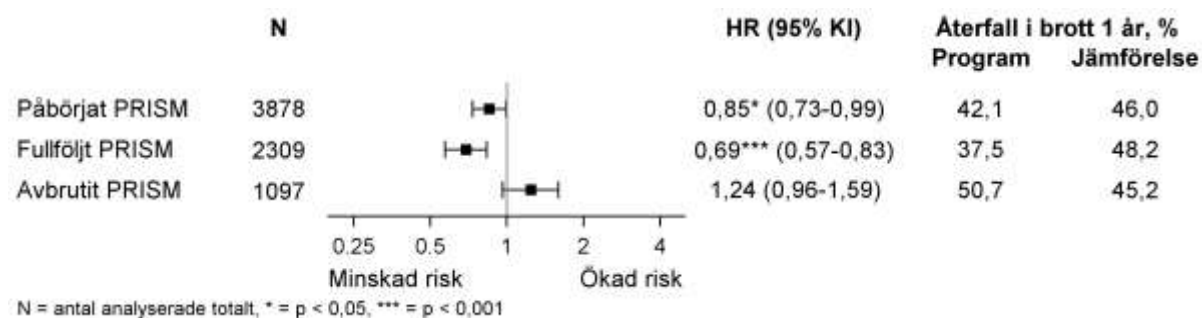
<sup>14</sup> 385 klienter fullföljde PRISM och 183 avbröt programmet. Uppgift om fullföljande saknades om 86 deltagare i PRISM.

<sup>15</sup> Denna uträkning är känd som Numbers Needed to Treat (NNT) vid positiv effekt respektive Numbers Needed to Harm (NNH) vid negativ effekt (Altman & Andersen, 1999).

<sup>16</sup> Att antalet individer i figuren är något färre än summan av programgruppen och jämförelsegruppen beror på att enstaka klienter saknade uppföljningstid och därför föll bort från analyserna.

## Känslighetsanalys

Resultatets tillförlitlighet stärks om de kan bekräftas med en annan källa och därför följdes deltagarna i en känslighetsanalys också upp i lagföringsregistret. Även vid uppföljning i lagföringsregistret hade de som fullföljt PRISM signifikant lägre risk att återfalla i brott (se figur 2). Till skillnad från uppföljningen i misstankeregistret kunde dock en statistiskt säkerställd riskminskning även hittas för de som påbörjat PRISM. De som avbrutit PRISM hade även enligt lagföringsregistret högre risk att dömas för nya brott jämfört med matchade klienter, men skillnaden var inte statistiskt säkerställd.



**Figur 2.** Hazardkvoter [HR] med 95% konfidensintervall [KI] för återfall i brott enligt lagföringsregistret. Deltagare i PRISM jämfördes med matchade klienter som inte deltagit i programmet.

## Diskussion

De som påbörjat eller fullföljt PRISM hade i stort sett lika hög risk att vårdas för missbruk som matchade klienter med likvärdig bakgrund. Det är dock svårt att dra säkra slutsatser om vilken effekt PRISM har på missbruk utifrån resultatet. För de som fullföljt PRISM har utfallet vård för missbruk ett brett konfidensintervall, som varken utesluter betydelsefull riskminskning eller riskökning. Med andra ord är resultatet osäkert (Haynes, Sackett, Guyatt, & Tugwell, 2006).<sup>17</sup> Vård för missbruk är dessutom en indikator på förhållandevis grava missbruksproblem. Om programdeltagarnas totala konsumtion av alkohol och andra droger förändrades skulle det inte nödvändigtvis synas på indikatorn. Resultatet kan inte heller relateras till någon tidigare utvärdering av PRISM:s effekt på missbruk eller konsumtion av alkohol och andra droger, eftersom någon sådan inte verkar finnas. För att kunna dra säkra slutsatser om PRISM:s effekt på missbruk behövs därför en ny utvärdering med mer känsligt mått, t.ex. självrapporterad konsumtion av alkohol och andra droger.

De som påbörjat PRISM hade oförändrad risk att återfalla i brott enligt misstankeregistret, men något minskad risk enligt lagföringsregistret.<sup>18</sup> Storleken på riskminskningen enligt lagföringsregistret kan dock betraktas som försumbar.<sup>19</sup> Däremot hade de som fullföljt PRISM minskad risk att återfalla i brott oavsett register. Det kan tolkas som att PRISM minskar återfall i brott om programmet genomförs i sin helhet. Denna slutsats är dock inte helt säker. Att klienter som fullföljt PRISM hade minskad risk att återfalla i brott kan nämligen bero på annat än behandlingsprogrammet. Klienter som fullföljer behandlingsprogram har ofta en bakgrund som minskar risken att återfalla i brott (Olver et al., 2011). Trots att många viktiga bakgrundsfaktorer hanterades i utvärderingen kan det därför kvarstå bakgrundsskillnader mellan de som fullföljt PRISM och jämförelsegruppen, som gör resultatet missvisande (s.k. bias). Vi kunde t.ex. inte hantera eventuella skillnader i motivation mellan programdeltagare och jämförelsegrupp. Klienter som fullföljt PRISM kan därför ha varit mer motiverade redan från början, vilket gjorde dem mer benägna både att slutföra programmet och avhålla sig från återfall i brott. Att tidigare utvärderingar av liknande behandlingsprogram funnit minskade brottsåterfall talar dock emot att resultatet bara skulle bero på bakgrundsskillnader mellan grupperna (Lipsey et al., 2007; Mitchell et al., 2012).

Resultaten i vår utvärdering liknar Kriminalvårdens tidigare utvärdering av PRISM. Där var risken att återfalla i brott något högre, men inte signifikant, bland samtliga som påbörjat PRISM medan de som fullföljde programmet hade signifikant lägre risk att återfalla i brott än jämförelsegruppen (Danielsson et al., 2012). I en utvärdering av PRISM:s systerprogram ASRO återföll däremot även de som fullföljt programmet i samma utsträckning som jämförelsegruppen (Palmer et al., 2011).

I en forskningsöversikt verkade missbruksprogram i anstalt i genomsnitt minska återfall i brott med 20% (från 35% till 28%). Enligt en annan forskningsöversikt verkade KBT-baserade behandlingsprogram för kriminella, däribland missbruksprogram, i genomsnitt minska återfall i brott med 25% (från 40% till 30%) (Lipsey et al., 2007). I en stor majoritet av studierna i översikten utgick man från de som *påbörjat* programmen. I utvärderingar där man istället fokuserat på dem som fullföljt behandling hittar man oftast ännu större skillnader i återfall i brott. Enligt både våra huvudresultat och Kriminalvårdens tidigare PRISM-utvärdering verkar de som påbörjat PRISM inte ha minskad risk att återfalla i brott. Däremot hade

---

<sup>17</sup> Konfidensintervallen är bred även för de som påbörjat PRISM, men eftersom konfidensintervall sträcker sig från i stort sett ingen skillnad till måttlig riskökning är resultatet inte lika osäkert som för fullföljare.

<sup>18</sup> Baserat på tidigare utvärderingar verkar det sakna betydelse för resultaten vilket register som används för att följa upp återfall i brott (Lipsey et al., 2007; Mitchell et al., 2012). En förklaring till den lilla resultatskillnaden i den här utvärderingen kan vara att lagföringsregistret hade ett år kortare uppföljningstid än misstankeregistret (i genomsnitt 549 dagar istället för 906 dagar). Skillnad i uppföljningstid kan ha betydelse eftersom programeffekter kan variera över tid. Men eftersom viss skillnad syns även sett till andelen återfall inom ett år förklarar det förmodligen inte hela skillnaden. En annan möjlig förklaring är att tiden mellan brott och registrering är längre i lagföringsregistret. Vissa brott som registrerats i misstankeregistret hann alltså inte registreras i lagföringsregistret. För tillförlitliga resultat borde uppföljningstiden därför egentligen vara längre i analyser baserade på lagföringsregistret.

<sup>19</sup> Riskminskningen enligt lagföringsregistret motsvarar en effektstorlek om HR 1,18 (om programgruppen istället utgör referensgrupp), vilket är under gränsen för en liten effektstorlek och räknas alltså som försumbar (Bedard, Krzyzanowska, Pintilie, & Tannock, 2007; Olivier & Bell, 2013).

de som fullföljt PRISM 21-31% minskad risk att återfalla i brott i vår utvärdering, beroende på i vilket register som uppföljningen skedde. Att vi inte hittade lika stor och otvetydig riskminskning som i många andra studier av KBT-baserade behandlingsprogram kan bero på annat än själva programinnehållet. Behandling som blivit en del av verksamhetens normala programutbud har ofta lägre effekt på återfall i brott än behandlingsprogram som införs på prov (Andrews, 2011; Lipsey et al., 2007). När behandlingsprogram införs i stor skala blir det svårare att kontrollera att de utförs som avsett, s.k. programtrohet (Andrews, 2011), vilket påverkar effekten på återfall i brott (Andrews & Dowden, 2005).<sup>20</sup> Även om vi inte hade tillgång till systematiska data om programtrohet i genomförandet av PRISM, arbetar Kriminalvården generellt med certifiering och handledning för att upprätthålla programtroheten. Program som är del av ordinarie verksamhet brukar dessutom mer sällan leva upp till principerna om risk, behov och mottaglighet (RBM) (Andrews, 2011), vilket också minskar programmets effekt på återfall i brott (Dowden & Andrews, 2000). RBM innebär kortfattat att prioritera behandling till kriminella med hög risk att återfalla i brott, fokusera på brottsdrivande faktorer samt anpassa programmet efter deltagarnas inlärningsstil. Det finns tecken på att PRISM utfördes i enlighet med principerna om RBM under undersökningsperioden.<sup>21</sup> Även behandlingsprogram som utförs i enlighet med RBM-principerna brukar dock få sämre effekt på återfall i brott *i ordinarie verksamhet*, vilket kan bero på bristande programtrohet (Andrews, 2011). Utifrån tidigare behandlingsforskning kan det alltså vara så att behandlingsprogrammet skulle få bättre effekt i Kriminalvården om programtroheten ökade.

Effektutvärderingar kan delas in i två typer. I den första typen utvärderas om en intervention har avsedd effekt under *idealiska omständigheter*, med strikt kontroll över bl.a. urvalskriterier, programutförande och programledarnas kunskap och erfarenhet (eng.: efficacy eller explanatory trial). I den andra utvärderingstypen analyseras istället om interventionen har önskad effekt under *normala omständigheter*, d.v.s. så som det brukar utföras i praktiken (eng.: effectiveness trial eller pragmatic trial). Den här rapporten är en utvärdering av PRISM under normala omständigheter. Det är möjligt att resultaten skulle vara annorlunda om PRISM utvärderades under idealiska omständigheter men det skulle inte ha lika stor praktisk relevans, eftersom behandlingsprogram sällan utförs så i praktiken (Trewick & Zwarenstein, 2009).

---

<sup>20</sup> I de utvärderingar av PRISM och systerprogrammet ASRO som hittills utförts har programmet varit del av ordinarie verksamhet (Danielsson et al., 2012; Palmer et al., 2011).

<sup>21</sup> Målgruppen är klienter med medel eller hög risk för återfall i brott och enligt bakgrundsstatistik om programdeltagarna verkar PRISM oftast ha getts till den målgruppen (se tabellerna i bilaga B). Däremot vet vi inte hur ofta riskbedömningsinstrument använts för att välja ut deltagare till programmet. PRISM är riktat mot missbruk, vilket anses vara ett viktigt kriminogent behov. Dessutom anses användning av tekniker från KBT, som i PRISM, stärka deltagarnas mottaglighet (Andrews et al., 1990). Att PRISM genomförs individuellt borde även göra det lättare för programledare att anpassa programmets utförande efter den enskilde deltagarens inlärningsstil.

## Referenser

- Akobeng, A. K. (2005). Understanding measures of treatment effect in clinical trials. *Archives of Disease in Childhood*, 90(1), 54-56. doi: 10.1136/adc.2004.052233
- Altman, D. G., & Andersen, P. K. (1999). Calculating the number needed to treat for trials where the outcome is time to an event. *BMJ*, 319(7223), 1492-1495. doi: 10.1136/bmj.319.7223.1492
- Andrews, D. A. (2011). The impact of nonprogrammatic factors on criminal-justice interventions. *Legal and Criminological Psychology*, 16(1), 1-23. doi: 10.1348/135532510X521485
- Andrews, D. A., & Dowden, C. (2005). Managing correctional treatment for reduced recidivism: A meta-analytic review of programme integrity. *Legal and Criminological Psychology*, 10(2), 173-187. doi: 10.1348/135532505X36723
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28(3), 369-404. doi: 10.1111/j.1745-9125.1990.tb01330.x
- Austin, P. C. (2008). A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Statistics in Medicine*, 27(12), 2037-2049. doi: 10.1002/sim.3150
- Bedard, P. L., Krzyzanowska, M. K., Pintlilie, M., & Tannock, I. F. (2007). Statistical power of negative randomized controlled trials presented at American Society for Clinical Oncology annual meetings. *Journal of Clinical Oncology*, 25(23), 3482-3487.
- Bennett, T., Holloway, K., & Farrington, D. (2008). The statistical association between drug misuse and crime: A meta-analysis. *Aggression and Violent Behavior*, 13(2), 107-118. doi: 10.1016/j.avb.2008.02.001
- Bland, J. M., & Altman, D. G. (1998). Survival probabilities (the Kaplan-Meier method). *BMJ*, 317(7172), 1572-1580. doi: 10.1136/bmj.317.7172.1572
- Bonta, J., & Andrews, D. A. (2007). Risk-need-responsivity model for offender assessment and rehabilitation *Corrections Research User Report No. 2007-06*. Ottawa, Ontario: Public Safety Canada.
- Danielsson, M., Fors, A., & Freij, I. (2012). *Behandlingsprogrammet PRISM i Kriminalvården. Utvärdering av återfall i brott för programdeltagare 2003-2006*. Norrköping: Kriminalvården.
- Debidin, M., & Lovbakke, J. (2005). Offending behaviour programmes in prison and probation. I G. Harper & C. Chitty (red.), *The impact of corrections on re-offending: a review of 'what works'*. London: Home Office.
- Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology* 42(4), 449-467.
- Friendship, C., Street, R., Cann, J., & Harper, G. (2005). Introduction: The policy context and assessing the evidence. I G. Harper & C. Chitty (red.), *The impact of corrections on re-offending: A review of 'what works'* (s. 1-16). London: Home Office.
- Frisell, T., Lichtenstein, P., & Långström, N. (2011). Violent crime runs in families: A total population study of 12.5 million individuals. *Psychological Medicine*, 41(1), 97-105. doi: 10.1017/S0033291710000462
- Haynes, R. B., Sackett, D. L., Guyatt, G. H., & Tugwell, P. (2006). *Clinical epidemiology: How to do clinical practice research*. Philadelphia: Lippincott Williams & Wilkins.
- Hollin, C., Palmer, E., McGuire, J., Hounscome, J., Hatcher, R., Bilby, C., & Clark, C. (2004). *Pathfinder programmes in the Probation Service: A retrospective analysis*. London: Home Office.
- Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ*, 319(7211), 670-674. doi: 10.1136/bmj.319.7211.670
- Hollis, V. (2007). *Reconviction analysis of programme data using interim accredited programmes software (IAPS)*. London: RDS NOMS Research and Evaluation.
- Kriminalvården. (2007). *Ackreditering av brotts- och missbruksrelaterade program i svensk kriminalvård*. Norrköping: Kriminalvården.
- Kriminalvården. (2008). *Q-BoM: Kvalitetsledningssystem för brotts- och missbruksrelaterade program*. Norrköping: Kriminalvården.
- Kriminalvården. (2009). *Uttagning och certifiering av programledare för brotts- och missbruksrelaterade program*. Norrköping: Kriminalvården.
- Kriminalvården. (2013). *Kvalitetskriterier för behandlingsprogram. PRISM, programme for reducing individual substance misuse*. Norrköping: Kriminalvården.

- Latessa, E. J., & Holsinger, A. (1998). The importance of evaluating correctional programs: Assessing outcome and quality. *Corrections Management Quarterly*, 2, 22-29.
- Lewis, J. A., & Machin, D. (1993). Intention to treat - who should use ITT? *British Journal of Cancer*, 68(4), 647-650.
- Lewis, S., & Clarke, M. (2001). Forest plots: Trying to see the wood and the trees. *BMJ*, 322(7300), 1479-1480. doi: 10.1136/bmj.322.7300.1479
- Lin, D. Y., & Wei, L. J. (1989). The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association*, 84(408), 1074-1078. doi: 10.1080/01621459.1989.10478874
- Lipsey, M. W., Landenberger, N. A., & Wilson, S. J. (2007). Effects of cognitive-behavioral programs for criminal offenders. *Campbell Systematic Reviews*, 3(6). doi: 10.4073/csr.2007.6
- Lösel, F. (2001). Evaluating the effectiveness of correctional programs: Bridging the gap between research and practice. I G. A. Bernfeld, D. P. Farrington & A. W. Leschied (red.), *Offender rehabilitation in practice: implementing and evaluating effective programs* (s. 67-96). Chichester: Wiley.
- McGuire, J. (2008). Programmes for probationers. I G. McIvor & P. Raynor (red.), *Developments in Social Work with Offenders* (s. 153-183). London: Jessica Kingsley.
- McMurrin, M., & Priestley, P. (2003). PRISM: Ett program för att minska sitt drogmissbruk. Teori. Norrköping: Kriminalvården.
- Mitchell, O., Wilson, D., & MacKenzie, D. L. (2012). The effectiveness of incarceration-based drug treatment on criminal behavior: A systematic review. *Campbell Systematic Reviews*, 8(18). doi: 10.4073/csr.2012.18
- Nilsson, I.-B., & Andersson, R. (1979). Folk- och bostadsräkningen 1975. Del 6:2. Förvärvsarbetande nattbefolkning i hela riket och länen m.m. Stockholm: SCB.
- Olivier, J., & Bell, M. L. (2013). Effect sizes for 2x2 contingency tables. *PloS one*, 8(3). doi: 10.1371/journal.pone.0058777
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2011). A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *Journal of Consulting and Clinical Psychology*, 79(1), 6-21. doi: 10.1037/a0022200
- Palmer, E., Hatcher, R., McGuire, J., Bilby, C., & Ayres, T. (2011). Evaluation of the Addressing Substance-Related Offending (ASRO) program for substance-using offenders in the community: A reconviction analysis. *Substance Use & Misuse*, 46(8), 1072-1080. doi: 10.3109/10826084.2011.559682
- Priestley, P., & Edström, E. (2010). One-to-one programmet i England, Wales och Sverige. I A. H. Berman & C. Å. Farbring (red.), *Kriminalvård i praktiken: Strategier för att minska återfall i brott och missbruk* (s. 439-466). Lund: Studentlitteratur.
- Priestley, P., & McMurrin, M. (2011). PRISM: Ett program för att minska sitt drogmissbruk. Manual. Norrköping: Kriminalvården.
- Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C. J., Nussenbaum, B., & Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves. *Otolaryngology - Head and Neck Surgery*, 143(3), 331-336. doi: 10.1016/j.otohns.2010.05.007
- Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, 8(1), 3-15. doi: 10.1177/096228029900800102
- Scheffer, J. (2002). Dealing with missing data. *Research Letters in the Information and Mathematical Sciences*, 3(1), 153-160.
- Treweek, S., & Zwarenstein, M. (2009). Making trials matter: Pragmatic and explanatory trials and the problem of applicability. *Trials*, 10(1), 37. doi: 10.1186/1745-6215-10-37
- Weisburd, D., Lum, C. M., & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *The ANNALS of the American Academy of Political and Social Science*, 578(1), 50-70. doi: 10.1177/000271620157800104
- Williamson, E., Morley, R., Lucan, A., & Carpenter, J. (2011). Propensity scores: From naïve enthusiasm to intuitive understanding. *Statistical Methods in Medical Research*, 23(3), 273-293. doi: 10.1177/0962280210394483

# Bilagor

## ***Bilaga A. Metodfördjupning***

### **Matchning**

Matchningen utfördes i statistikprogrammet SAS via ett publicerat s.k. makro (Coca-Perrillon, 2007). De bakgrundsfaktorer som skulle användas i matchningen valdes utifrån om variabeln uppvisade ett signifikant samband ( $p < 0,05$ ) med programdeltagande och *dessutom* var signifikant associerad med någon av utfallsvariablerna. Det beror på att det bara är variabler som både påverkar utfallen och har olika fördelning i program- och jämförelsegrupp som riskerar förvränga resultat (eng.: bias) (Bonita, Beaglehole, & Kjellström, 2006).

För att veta om en matchning resulterat i att program- och jämförelsegrupperna blivit tillräckligt lika jämfördes grupperna före och efter matchning. Enligt rekommendation redovisas procent, medelvärden och effektstorleksmättet Cohens  $d$  (Austin, 2009). Om matchningen är bra så ska skillnaderna mellan grupperna minska. Det finns inte samsyn om var gränsen går för när grupperna är tillräckligt lika. Det rekommenderas att Cohens  $d$  bör vara lägre än 0,20, vilket i andra sammanhang brukar vara minimigränsen för en liten effektstorlek. Den mest konservativa rekommendationen är dock att Cohens  $d$  bör vara lägre än 0,10 (Apel & Sweeten, 2010). Vi eftersträvade att Cohens  $d$  skulle vara lägre än 0,10, men åtminstone under 0,20. För jämförelse av grupperna före och efter matchning se bilaga B.

Vi använde metoden nearest neighbor-matchning. Det innebär att till varje programdeltagare matchades den/dem i jämförelsegruppen med mest lik propensity score till programdeltagaren. Klienterna i jämförelsegruppen kunde enbart matchas till en programdeltagare var (eng.: matching without replacement). Eftersom fler individer innebär större chans att med rimlig säkerhet kunna statistiskt säkerställa en faktisk effekt (statistisk power) är det önskvärt att så många individer som möjligt ingår. Därför matchades fem personer till varje programdeltagare.

### **Hantering av bortfall**

I de flesta statistiska analyser exkluderas individer med bortfall på någon variabel. För att detta inte ska förvränga resultaten så måste bortfallet vara slumpmässigt, vilket det sällan är (Yuan, 2000). En rekommenderad metod för att hantera bortfall är s.k. multipel imputation, som innebär att bortfall av information ersätts med "sannolika värden" (Schafer, 1999). Vid multipel imputation skapas flera nya dataset, där värden i det ursprungliga datasetet kopieras medan bortfall ersätts med sannolika värden, som skiljer sig från datasetet till dataset. Målet med multipel imputation är inte att ersätta bortfall med det sanna värdet, utan att möjliggöra giltig statistisk slutledning (Rubin, 1996). Det kan t.ex. handla om att få fram en så korrekt skattning av relationen mellan två variabler som möjligt, trots bortfallet. Under förutsättning att bortfallet beror på de faktorer som man tar hänsyn till vid multipel imputation, kommer bortfallet inte längre leda till förvrängda resultat. I flera s.k. simuleringsstudier har multipel imputation lett till bättre estimering av det sanna resultatet än att exkludera individer med bortfall (Marshall, Altman, Royston, & Holder, 2010; Osborne, 2012; Raghunathan, 2004).

Vi hanterade bortfall på bakgrundsfaktorer på två sätt; multipel imputation och bortfallsindikatorer. Att vi inte enbart använde multipel imputation beror på att det främst är anpassat för kvantitativa (numeriska) variabler. För att multipel imputation ska kunna utföras på kvalitativa (kategoriska) variabler brukar krävas att bortfallet är monotont (Ake, 2005). Ett exempel på monotont bortfall är då en individ i en longitudinell enkätstudie upphör att besvara enkäten efter ett visst antal mättillfällen. Monotona bortfall är därmed ovanliga. Den metod för multipel imputation som var tillgänglig för oss och kan hantera bortfall som inte är monotont kallas MCMC (Markov Chain Monte Carlo). Metoden är utvecklad för kvantitativa variabler men det har gjorts försök att använda MCMC på kvalitativa variabler genom att behandla variablerna som om de vore kvantitativa. Tyvärr har simuleringsstudier funnit att detta tillvägagångssätt leder till bias (Ake,



2005; Horton, Lipsitz, & Parzen, 2003). I vår studie hade vi inte ett monotont bortfall och multipel imputation var därmed bara lämpligt att använda på kvantitativa variabler.

Bortfall på kvalitativa variabler hanterades genom att använda bortfallsindikatorer inom ramen för propensity score-matchning. Metoden har beskrivits utförligt tidigare (Rosenbaum, 2010). Bortfall ersätts med ett godtyckligt värde och samtidigt skapas en ny variabel som indikerar om variabeln har bortfall. Det innebär att det godtyckliga värdet inte påverkar vad individen får för propensity score, men att bortfallet däremot gör det. På så sätt blir programgruppen och jämförelsegruppen likvärdiga på både tillgänglig information och bortfall. Under förutsättning att bortfallet beror på kontrollvariablerna så ska bortfallet därmed inte längre leda till bias.

I en simuleringsstudie på bortfallshantering vid propensity score-matchning jämfördes multipel imputation och användning av bortfallsindikatorer (Mattei, 2009). Författaren drog slutsatsen att resultaten inte gav starkt stöd för att den ena metoden skulle vara bättre än den andra, men att multipel imputation verkar vara något bättre. Att multipel imputation skulle vara att föredra förklarades med att det går att ta hänsyn till fler faktorer bakom bortfallet. Exempelvis rekommenderas att man tar hänsyn till utfallsvariabeln vid multipel imputation, vilket man inte kan vid uträkning av propensity scores. Att vi inte enbart hanterade bortfall genom bortfallsindikatorer beror alltså på att multipel imputation är att föredra.

För att minska risken för kvarstående bias efter multipel imputation rekommenderas att man inkluderar:

1. variabler som ska ingå i senare analys (inklusive utfallsmått)
2. variabler som korrelerar med bortfall
3. variabler som korrelerar med imputerade variabler (van Buuren, Boshuizen, & Knook, 1999).

Ju fler variabler som inkluderas vid multipel imputation desto lägre sannolikhet att resultatet blir skevt. Samtidigt är det inte nödvändigt att inkludera mer än de 15-25 viktigaste variablerna (van Buuren et al., 1999). Vi inkluderade därför alla potentiella kontrollvariabler och utfallsvariabler i multipel imputation. Multipel imputation (MI) utfördes via proceduren Proc MI i SAS. Variablerna som vi utförde multipel imputation på var: ålder vid första dom, antal domar och antal år med arbetsinkomst senaste fem åren. Man kan inte utföra multipel imputation på variabler där bortfallet beror på ett sant värde. I en studie av normalbefolkningen vore det därför olämpligt att utföra multipel imputation på antal domar, eftersom personer som aldrig dömts inte kan ha några lagföringsdata. I det här fallet var dock alla i studien kriminalvårdsklienter, och därmed kan bortfall på antal domar inte bero på att de aldrig dömts.

Samtliga variabler som skulle ingå i någon av analyserna kontrollerades för vid multipel imputation. Dessutom ingick bortfallsindikatorer för samtliga kvalitativa variabler med bortfall. Därmed kunde vi kontrollera för värden på kvalitativa variabler med bortfall, utan att deras bortfall oavsiktligt imputerades. Fem imputationer utfördes enligt Proc MI:s standard. När bortfallet är lågt bör fem imputationer räcka (Schafer, 1999). För att kunna utföra propensity score-matchning efter multipel imputation beräknades propensity scores för alla fem imputerade dataset. Därefter räknades genomsnittlig propensity score ut över alla fem dataset och användes vid matchning. I en jämförelse mellan två sätt att utföra propensity score-matchning efter multipel imputation var detta tillvägagångssätt att föredra (Mitra & Reiter, 2012).

## Referenser till bilaga A

- Ake, C. (2005). *Rounding after multiple imputation with non-binary categorical covariates*. Paper presented at the SAS Users Group International 30 Proceedings. <http://www2.sas.com/proceedings/sugi30/112-30.pdf>
- Apel, R. J., & Sweeten, G. (2010). Propensity score matching in criminology and criminal justice. I A. R. Piquero & D. Weisburd (red.), *Handbook of Quantitative Criminology* (s. 543-562). New York: Springer.
- Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*, 28(25), 3083-3107. doi: 10.1002/sim.3697
- Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic epidemiology*. Geneva: World Health Organization.
- Coca-Perrillon, M. (2007). *Local and Global Optimal Propensity Score Matching*. Paper presented at the SAS Global Forum 2007. <http://www2.sas.com/proceedings/forum2007/185-2007.pdf>
- Horton, N. J., Lipsitz, S. R., & Parzen, M. (2003). A potential for bias when rounding in multiple imputation. *The American Statistician*, 57(4), 229-232. doi: 10.1198/0003130032314
- Marshall, A., Altman, D. G., Royston, P., & Holder, R. (2010). Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. *BMC Medical Research Methodology*, 10(1), 7. doi: 10.1186/1471-2288-10-7
- Mattei, A. (2009). Estimating and using propensity score in presence of missing background data: an application to assess the impact of childbearing on wellbeing. *Statistical Methods and Applications*, 18(2), 257-273. doi: 10.1007/s10260-007-0086-0
- Mitra, R., & Reiter, J. P. (2012). A comparison of two methods of estimating propensity scores after multiple imputation. *Stat Methods Med Res*. doi: 10.1177/0962280212445945
- Osborne, J. W. (2012). *Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data*. Thousand Oaks: Sage Publications.
- Raghunathan, T. E. (2004). What do we do with missing data? Some options for analysis of incomplete data. *Annual Review of Public Health*, 25(1), 99-117. doi: 10.1146/annurev.publhealth.25.102802.124410
- Rosenbaum, P. R. (2010). *Design of observational studies*. New York: Springer.
- Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, 91(434), 473-489. doi: 10.1080/01621459.1996.10476908
- Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, 8(1), 3-15. doi: 10.1177/096228029900800102
- van Buuren, S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, 18(6), 681-694. doi: 10.1002/(SICI)1097-0258(19990330)18:6<681::AID-SIM71>3.0.CO;2-R
- Yuan, Y. (2000). *Multiple imputation for missing data: Concepts and new development*. Paper presented at the SAS Users Group International 25 Proceedings. <http://www2.sas.com/proceedings/sugi25/25/st/25p267.pdf>

## Bilaga B. Bakgrundsfaktorer före och efter matchning

**Tabell B1.** Jämförelse av eventuella bakgrundsskillnader mellan dem som *påbörjat* PRISM och jämförelsegruppen före och efter propensity score-matchning (1:5 nearest neighbor).

	Påbörjat PRISM (n=654)	Omatchad jämförelsegrupp (n=56375)	Matchad jämförelsegrupp (n=3270)	Cohens <i>d</i>	
				Före	Efter
Ålder år 2013, m	38,7 år	41,8 år	38,5 år	0,23	-0,02
Kön, %				0,01	0,01
Kvinna	13,0	11,6	13,4		
Man	87,0	88,5	86,6		
Anstalt, %	34,3	30,8	35,8	0,02	0,02
Civilstatus, %				0,05	0,01
Gift/sambo med gemensamma barn	11,0	19,7	11,3		
Ogift/sambo utan gemensamma barn	89,0	79,8	88,8		
Saknas	0,0	0,6	0,0		
Bostadsområde, %				0,09	0,04
Stad/förort	80,9	60,7	82,8		
Annan ort	18,8	38,1	16,9		
Information saknas	0,3	1,2	0,3		
Högsta utbildningsnivå, %				0,07	0,06
Folk- eller grundskola < 9 år	6,1	6,6	6,4		
Folk- eller grundskola 9 år	50,9	36,2	50,9		
Gymnasial utbildning 1-3 år	35,3	45,6	36,6		
Universitet och högskola, inkl. gymnasium 4 år	4,7	8,0	3,3		
Information saknas	2,9	3,6	2,8		
SES i barndom, %				0,03	0,04
Låg	36,5	35,6	37,7		
Medel	28,3	20,5	26,5		
Hög	15,6	13,2	17,0		
Information saknas	20,0	30,7	18,8		
<i>Tidigare brottslighet</i>					
Antal tidigare domar, m	15,4	8,8	15,2	-0,49	-0,01
Ålder vid första dom, m	19,4 år	24,2 år	19,2 år	0,44	-0,04
Ngn gång dömd för våldsbrott, %	69,3	61,4	69,2	0,03	0,00
Ngn gång dömd för narkotikabrott, %	84,6	38,4	85,6	0,20	0,02
Ngn gång dömd för rattfylleri, %	59,3	50,6	59,8	0,04	0,01
Information om tidigare brottslighet saknas, %	0,2	0,4	0,2		
<i>Familjär brottslighet/missbruk</i>					
Mamma ngn gång dömd, %	31,0	21,1	31,3	0,05	0,00
Pappa ngn gång dömd, %	56,1	43,9	56,7	0,05	0,01
Ngt syskon som ngn gång dömts, %	39,5	33,6	39,5	0,03	0,00
Ngn förälder med missbruk/beroende, %	21,4	17,0	20,7	0,03	0,01
<i>Psykiatri</i>					
Alkoholmissbruk/-beroende, %	26,3	22,2	25,6	0,02	0,01
Narkotikamissbruk/-beroende, %	51,2	21,5	51,4	0,15	0,00
Personlighetsstörning, %	7,3	5,8	9,0	0,01	0,04
ADHD, %	6,3	3,4	6,8	0,03	0,02

**Fortsättning tabell B1.**

	Påbörjat PRISM (n=654)	Omatchad jämförelsegrupp (n=56375)	Matchad jämförelsegrupp (n=3270)	Cohens <i>d</i>	
				Före	Efter
<i>Behandlingsprogram</i>					
Påbörjat <i>annat</i> missbruksprogram än PRISM, %	18,0	3,4	16,7	0,17	0,03
Påbörjat sexbrotts-/partnervåldsprogram, %	1,5	1,1	0,6	0,01	0,09
Påbörjat motivationsprogram, %	34,3	6,7	32,9	0,23	0,02
Påbörjat allmänt kriminalitets-/våldsprogram, %	22,0	5,1	20,7	0,16	0,02
Antal år med arbetsinkomst senaste fem åren, m	1,7 år	2,6 år	1,6 år	0,46	-0,02
Information om arbete saknas, %	0,0	0,6	0,0		

**Tabell B2.** Jämförelse av eventuella bakgrundsskillnader mellan dem som *fullföljt* PRISM och jämförelsegruppen före och efter matchning (1:5 nearest neighbor).

	Fullföljt PRISM (n=385)	Omatchad jämförelsegrupp (n=56375)	Matchad jämförelsegrupp (n=1925)	Cohens <i>d</i>	
				Före	Efter
Ålder år 2013, m	39,2 år	41,8 år	38,9 år	0,20	-0,02
Kön, %				0,00	0,01
Kvinna	10,9	11,6	11,2		
Man	89,1	88,5	88,8		
Anstalt, %	47,0	30,8	46,8	0,06	0,00
Civilstatus, %				0,04	0,02
Gift/sambo med gemensamma barn	10,4	19,7	11,3		
Ogift/sambo utan gemensamma barn	89,6	79,8	88,7		
Information saknas	0,0	0,6	0,0		
Bostadsområde, %				0,06	0,00
Stad/förort	78,7	60,7	78,7		
Annan ort	21,0	38,1	21,1		
Information saknas	0,3	1,2	0,2		
Högsta utbildningsnivå, %				0,06	0,09
Folk- eller grundskola < 9 år	4,7	6,6	6,2		
Folk- eller grundskola 9 år	52,2	36,2	47,7		
Gymnasial utbildning 1-3 år	34,8	45,6	38,8		
Universitet och högskola, inkl. gymnasium 4 år	4,9	8,0	3,8		
Information saknas	3,4	3,6	3,5		
SES i barndom, %				0,02	0,11
Låg	37,4	35,6	36,8		
Medel	28,8	20,5	22,4		
Hög	16,1	13,2	11,4		
Information saknas	17,7	30,7	29,5		
<i>Tidigare brottslighet</i>					
Antal tidigare domar, m	15,8	8,8	15,7	-0,52	0,00
Ålder vid första dom, m	19,2 år	24,2 år	19,1 år	0,46	-0,02
Ngn gång dömd för våldsbrott, %	72,2	61,4	71,5	0,04	0,01
Ngn gång dömd för narkotikabrott, %	85,2	38,4	87,4	0,16	0,04
Ngn gång dömd för rattfylleri, %	61,8	50,6	61,7	0,04	0,00
Information om tidigare brottslighet saknas, %	0,3	0,4	0,1		
<i>Familjär brottslighet/missbruk</i>					
Mamma ngn gång dömd, %	32,0	21,1	32,4	0,04	0,01
Pappa ngn gång dömd, %	57,4	43,9	57,2	0,04	0,00
Ngt syskon som ngn gång dömts, %	40,5	33,6	40,9	0,02	0,01
Ngn förälder med missbruk/beroende, %	20,8	17,0	21,4	0,02	0,01
<i>Psykiatri</i>					
Alkoholmissbruk/-beroende, %	23,9	22,2	25,8	0,01	0,03
Narkotikamissbruk/-beroende, %	45,7	21,5	47,2	0,10	0,02
Personlighetsstörning, %	7,3	5,8	9,1	0,01	0,05
ADHD, %	6,0	3,4	6,5	0,02	0,02

**Fortsättning tabell B2.**

	Fullföljt PRISM (n=385)	Omatchad jämförelsegrupp (n=56375)	Matchad jämförelsegrupp (n=1925)	Cohens <i>d</i>	
				Före	Efter
<i>Behandlingsprogram</i>					
Påbörjat <i>annat</i> missbruksprogram än PRISM, %	21,6	3,4	19,2	0,16	0,04
Påbörjat sexbrotts-/partnervåldsprogram, %	1,3	1,1	0,5	0,00	0,07
Påbörjat motivationsprogram, %	36,4	6,7	36,5	0,19	0,00
Påbörjat allmänt kriminalitets-/våldsprogram, %	27,8	5,1	25,8	0,17	0,03
Antal år med arbetsinkomst senaste fem åren, m	1,6 år	2,6 år	1,6 år	0,50	-0,01
Information om arbete saknas, %	0,0	0,6	0,0		

**Tabell B3.** Jämförelse av eventuella bakgrundsskillnader mellan dem som *avbrutit* PRISM och jämförelsegruppen före och efter matchning (1:5 nearest neighbor).

	Avbrutit PRISM (n=183)	Omatchad jämförelsegrupp (n=56375)	Matchad jämförelsegrupp (n=915)	Cohens <i>d</i>	
				Före	Efter
Ålder år 2013, m	37,8 år	41,8 år	37,8 år	0,30	0,00
Kön, %				0,02	0,01
Kvinna	16,4	11,6	15,7		
Man	83,6	88,5	84,3		
Anstaltsklient, %	23,0	30,8	25,3	0,02	0,04
Civilstatus, %				0,02	0,04
Gift/sambo med gemensamma barn	12,6	19,7	11,0		
Ogift/sambo utan gemensamma barn	87,4	79,8	89,0		
Information saknas	0,0	0,6	0,0		
Bostadsområde, %				0,05	0,04
Stad/förort	84,7	60,7	86,6		
Annan ort	15,3	38,1	13,4		
Information saknas	0,0	1,2	0,0		
Högsta utbildningsnivå, %				0,04	0,11
Folk- eller grundskola < 9 år	10,9	6,6	7,9		
Folk- eller grundskola 9 år	49,7	36,2	54,5		
Gymnasial utbildning 1-3 år	32,2	45,6	31,7		
Universitet och högskola, inkl. gymnasium 4 år	3,8	8,0	2,7		
Information saknas	3,3	3,6	3,2		
SES i barndom, %				0,01	0,04
Låg	37,7	35,6	37,5		
Medel	25,1	20,5	22,6		
Hög	14,2	13,2	13,0		
Information saknas	23,0	30,7	26,9		
<i>Tidigare brottslighet</i>					
Antal tidigare domar, m	15,3	8,8	15,3	-0,48	0,00
Ålder vid första dom, m	19,6 år	24,2 år	19,9 år	0,42	0,04
Ngn gång dömd för våldsbrott, %	68,3	61,4	70,0	0,02	0,03
Ngn gång dömd för narkotikabrott, %	88,0	38,4	88,4	0,12	0,01
Ngn gång dömd för rattfylleri, %	55,2	50,6	56,2	0,01	0,03
Information om tidigare brottslighet saknas, %		0,4			
<i>Familjär brottslighet/missbruk</i>					
Mamma ngn gång dömd, %	29,5	21,1	29,6	0,02	0,00
Pappa ngn gång dömd, %	53,6	43,9	52,5	0,02	0,02
Ngt syskon som ngn gång dömts, %	37,7	33,6	36,2	0,01	0,02
Ngn förälder med missbruk/beroende, %	18,0	17,0	24,4	0,00	0,11
<i>Psykatri</i>					
Alkoholmissbruk/-beroende, %	26,2	22,2	26,6	0,01	0,01
Narkotikamissbruk/-beroende, %	61,8	21,5	62,2	0,11	0,01
Personlighetsstörning, %	8,7	5,8	9,7	0,01	0,02
ADHD, %	6,0	3,4	6,7	0,02	0,02

Fortsättning tabell B3.

	Avbrutit PRISM (n=183)	Omatchad jämförelsegrupp (n=56375)	Matchad jämförelsegrupp (n=915)	Cohens <i>d</i>	
				Före	Efter
<i>Behandlingsprogram</i>					
Påbörjat <i>annat</i> missbruksprogram än PRISM, %	15,9	3,4	16,8	0,08	0,02
Påbörjat sexbrotts- /partnervåldsprogram, %	1,1	1,1	0,8	0,00	0,03
Påbörjat motivationsprogram, %	34,4	6,7	34,1	0,12	0,01
Påbörjat allmänt kriminalitets-/våldsprogram, %	15,9	5,1	16,4	0,06	0,01
Antal år med arbetsinkomst senaste fem åren, m	1,7 år	2,6 år	1,7 år	0,44	0,00
Saknas, %	0,0	0,6	0,0		