

Utvärdering av behandlingsprogrammet One-to-One i Kriminalvården

VI BRYTER DEN ONDA CIRKELN



Sammanfattning

Bakgrund

One-to-one (OTO) är ett behandlingsprogram med ursprung från England som infördes i Kriminalvården 1999. Programmet är baserat på principer från kognitiv beteendeterapi (KBT) och genomförs individuellt. Kriminalvården har tidigare utvärderat om OTO minskade återfall i brott bland programdeltagare 2003–2006. I den utvärderingen hade klienter som fullföljt programmet 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. Några andra utvärderingar av OTO:s effekt hittades inte. Däremot finns utvärderingar av kriminalvårdsprogram som likt OTO baseras på KBT. När sådana utvärderingar vägdes samman i en forskningsöversikt tydde de på att KBT-baserade behandlingsprogram i genomsnitt minskar återfall i brott med 25 % (från 40 % till 30 %). Program som likt OTO innehöll kognitiv omstrukturering eller hade individuella moment verkade dessutom minska återfall i brott mer än andra KBT-baserade program. Syftet med den här utvärderingen var att undersöka om OTO minskar risken att återfalla i brott bland kriminalvårdsklienter i Sverige.

Metod

I utvärderingen ingick 597 klienter som påbörjat OTO och som frigavs från anstalt eller skrevs in i frivård 2007–2009. Av dessa var det 67 % som fullföljde programmet. Programdeltagarna jämfördes med fem gånger fler individuellt matchade klienter aktuella under samma period men som inte påbörjat programmet. För att minska risken att bakgrundsskillnader gör resultaten missvisande användes så kallad propensity score-matchning. Alltså hanterades eventuella skillnader mellan grupperna före programmet i till exempel ålder, kön och antal tidigare domar.

Resultat

Klienter som fullföljde OTO hade signifikant lägre risk att återfalla i brott än en matchad jämförelsegrupp. Andelen som återföll i nya brott inom ett år var 54 % bland de som fullföljt programmet och 59 % i den matchade jämförelsegruppen. Däremot var risken att återfalla i brott inte signifikant lägre bland samtliga som påbörjade programmet. Det har att göra med att klienter som avbrutit OTO tvärtom hade ökad risk att återfalla i brott.

Diskussion

Tillsammans med tidigare utvärderingar på liknande program kan resultaten tolkas som att OTO kan minska återfall i brott, förutsatt att programmet genomförs i sin helhet. Det finns dock anledning att tolka resultatet med viss försiktighet, t.ex. kan resultatet påverkas av bakgrundsskillnader mellan de som fullföljt OTO och jämförelsegruppen, trots att många viktiga bakgrundsfaktorer hanterades. Utvärderingen tyder också på att det är viktigt att så många som möjligt av dem som påbörjar OTO fullföljer behandlingen.

Innehållsförteckning

Introduktion	1
Tidigare forskning	1
Problemformulering och syfte.....	2
Metod	3
Undersökningsgrupp.....	3
Källor.....	3
<i>Utfall</i>	3
<i>Bakgrundsfaktorer</i>	4
Statistisk analys.....	5
Resultat.....	6
Huvudresultat.....	6
Känslighetsanalys.....	6
Diskussion.....	8
Referenser.....	10
Bilagor	12
Bilaga A. Metodfördjupning.....	12
Bilaga B. Bakgrundsfaktorer före och efter matchning.....	15

Introduktion

One-to-One (OTO) är ett allmänt kriminalitetsprogram från England skapat av kriminologen Philip Priestley i början på 1990-talet. Målet med OTO är att minska återfall i brott. Allmänna kriminalitetsprogram är inte riktade mot någon speciell form av brott, utan passar de flesta brottskategorier. Programmet bygger på principer från kognitiv beteendeterapi (KBT) och social inlärningsteori (Priestley, 2003). KBT-baserade behandlingsprogram för kriminella utgår ifrån att brott uppstår p.g.a. inlärd kognitiva brister och förvrängningar. Därför ligger tonvikten vid att få klienter att förstå och förändra de tankeprocesser som föregår deras brott (Lipsey, Landenberger, & Wilson, 2007). OTO innehåller bl.a. övningar i problemlösning, självkontroll, moraliskt resonerande, social färdighetsträning och kognitiv omstrukturering. För att öka deltagarnas motivation till förändring används även samtalsmetoden motiverande intervju (MI) (Priestley, 2006).

OTO består av tre delar: kartläggning, färdighetsträning och tillämpning/återfallsprevention. Först kartläggs klienternas behov och styrkor. Därefter får deltagarna lära sig färdigheter som ska hjälpa dem undvika att begå nya brott. I den sista delen får deltagarna tillämpa sina nya färdigheter på faktorer som bidrar till deras kriminalitet. Programmet består av ett förmöte och 20 individuella möten om 60-90 minuter, en eller två gånger per vecka. OTO pågår i fyra till fem månader, beroende på antalet möten per vecka. Hemuppgifter och rollspelsövningar är en del av programmet (Priestley, 2006).

Programledaren utför OTO enskilt tillsammans med klienten, därav namnet One-to-One ("en-till-en"). Det kan finnas flera skäl till att ge behandling enskilt. Ibland kan det vara svårt att få tillräckligt många deltagare för grupprogram, vilket orsakar väntetid. Det kan också vara olämpligt att placera vissa klienter i gruppbehandling, exempelvis p.g.a. konflikter eller negativ gruppkultur (Priestley & Edström, 2010). OTO utförs av programledare som är utbildade i programmet. För att få fortsätta genomföra programmet ska man certifieras inom ett år efter utbildningens avslut. Certifieringen innebär att programmets genomförande granskas, bl.a. via videoinspelningar (Kriminalvården, 2009). Programledare som utför OTO ska även få löpande metodhandledning (Kriminalvården, 2008).

För varje ny kriminalvårdsklient upprättas en s.k. verkställighetsplan, där klientens risker och behov identifieras. Om det framkommer att klienten har behov av behandling föreslås ett lämpligt behandlingsprogram som man försöker motivera denne att delta i. I Sverige används OTO på klienter i både frivård och anstalt (Priestley & Edström, 2010). Målgruppen är kriminalvårdsklienter med medelhög eller hög risk för återfall i brott. Dock är det inte uttalat att ett visst riskbedömningsinstrument ska användas för att bedöma detta. Exklusionskriterier är psykiska problem (annat än missbruk/beroende), kaotisk tillvaro, gravt missbruk och inlärningssvårigheter. Dömda för våld i nära relation eller sexualbrott inkluderas bara i OTO först efter att de deltagit i behandling specialiserad på dessa brottskategorier (Kriminalvården, 2013).

OTO började användas i svensk kriminalvård år 1999 och ackrediterades år 2003. Ackreditering innebär att programmet granskas av en panel bestående av medlemmar i 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 de ackrediterade behandlingsprogrammen fortlöpande ska utvärderas med avseende på återfall i brott (Kriminalvården, 2007).

Tidigare forskning

OTO har utvärderats tidigare inom Kriminalvården. I programgruppen (n=728) ingick då klienter utskrivna från anstalt eller inskrivna i frivård som deltog i OTO år 2003-2006 (Danielsson, Fors, & Freij, 2009). Jämförelsegruppen utgjordes av klienter som inte deltagit i något av Kriminalvårdens behandlingsprogram under samma period (n=7280). Resultatet justerades för bakgrundsskillnader som det fanns information om i Kriminalvårdens register och återfall i brott definierades som ny

kriminalvårdspåföljd.¹ Enligt utvärderingen hade klienter som påbörjat OTO ungefär lika hög risk som jämförelsegruppen att återfalla i något brott (hazardkvot [HR] 1,04, 95% konfidensintervall [KI] 0,92-1,17). Klienter som fullföljt programmet hade däremot en statistiskt säkerställd minskad risk att återfalla i brott (HR 0,75, 95% KI 0,62-0,90). Att de som fullföljt, men inte alla som påbörjat programmet, hade minskad risk att återfalla i brott kan bero på att OTO minskar återfall i brott först om det genomförs fullt ut. Men klienter som fullföljer behandling har å andra sidan ofta annorlunda bakgrund, t.ex. är de inte sällan välmotiverade redan innan behandlingen börjar (Olver, Stockdale, & Wormith, 2011). Man riskerar därför snedvridna resultat p.g.a. att fullföljares bakgrund både ökar sannolikheten att fullfölja program och minskar risken att återfalla i brott.

Vi har inte hittat någon annan utvärdering av OTO:s effekt på återfall i brott.² Det finns dock många studier på andra KBT-baserade program mot återfall i brott som sammanfattats i systematiska kunskapsöversikter med metaanalys. Metaanalys innebär att man gör en samlad bedömning av flera enskilda studier genom att statistiskt sammanföra deras resultat. I genomsnitt tycks KBT-baserade program minska återfall i brott (Lipsey et al., 2007; Wilson, Bouffard, & Mackenzie, 2005). Dessutom verkar innehåll av kognitiv omstrukturering eller individuella moment medföra större effekt på återfall i brott (Lipsey et al., 2007). Enligt utvärderingar av liknande behandlingsprogram borde programinnehållet i OTO därför kunna minska återfall i brott (Priestley, 2006).

Problemformulering och syfte

Idag finns enligt vår kännedom inga andra studier på OTO utöver Kriminalvårdens förra utvärdering, och där avbröts programmet av en hög andel av de som påbörjade OTO. När andelen som avbryter behandling är hög blir eventuella programeffekter för samtliga som påbörjat programmet svårare att påvisa, eftersom många då exponerats mindre för behandlingen. Andelen klienter som fullföljer OTO har sedan dess ökat och därför är det angeläget med en ny effektutvärdering av programmet. Syftet med denna studie är att undersöka om OTO minskar återfall i brott bland svenska kriminalvårdsklienter.

Studiens primära frågeställning är:

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

Dessutom undersöks den sekundära frågeställningen:

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

¹ Bakgrundsfaktorer som justerades för var: ålder, kön, tid i anstalt, tidigare verkställigheter samt diverse brottstyper.

² I två brittiska utvärderingar av ett flertal behandlingsprogram ingick även deltagare i OTO (Hollin et al., 2004; V. Hollis, 2007). Men i ingen av studierna redovisades behandlingsprogrammen separat och andelen OTO-deltagare var mycket låg (1,2% respektive 4,7%). Resultaten säger därmed mer om de övriga behandlingsprogrammen än om OTO och redovisas därför inte här. I den ena utvärderingen saknades dessutom jämförelsegrupp (V. Hollis, 2007), vilket ökar risken att resultaten är missvisande.

Metod

Undersökningsgrupp

Programgruppen bestod av kriminalvårdsklienter som skrevs in i frivården eller ut från anstalt åren 2007 till 2009 och påbörjade OTO under påföljden (n=597). De jämfördes med kriminalvårdsklienter aktuella under samma period men som inte deltagit i OTO. För att minska risken för systematiska bakgrundsskillnader mellan grupperna användes propensity score-matchning (se sida 4).

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 deltagare kommer att fullfölja programmet eller inte. Uteblivna eller oönskade effekter av behandling kan dessutom leda till att en deltagare väljer att avbryta densamma. 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 exponerats för *bela* behandlingen. Att ta del av hela behandlingssinnehållet kan vara en förutsättning för önskvärda effekter och i så fall skulle analyser enligt intention-to-treat kunna 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 ändra sina liv. Med per-protokoll ökar därmed risken att behandlingen ser ut att ha positiva effekter när resultatet i själva verket skulle kunna bero på bakgrundsskillnader. Sammanfattningsvis har alltså både intention-to-treat och per-protokoll för- och nackdelar och det är därför som vi utförde och här presenterar 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.³ 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

Utfallet bestod i återfall i något nytt brott och följdes i första hand upp i misstankeregistret. Enbart personer som av förundersökningsledare bedömts som minst skäligen misstänkta för brott ingår i misstankeregistret. Skäligen 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).⁴ Det verkar inte ha betydelse för resultaten om återfall i brott definieras som ny arrestering eller ny dom (Lipsey et al., 2007; Mitchell, Wilson, & MacKenzie, 2012). Misstankeregistret användes för att det ger större statistisk power,⁵ d.v.s. ökad sannolikhet att statistiskt säkerställa en faktisk effekt, och för att det

³ Registerinformation hämtades från databasen "Crime-2" vid Institutionen för medicinsk epidemiologi och biostatistik vid Karolinska Institutet. Detta gjordes inom ramen för en forskningsstudie som godkännts av Regional forskningsetisk kommitté vid Karolinska Institutet (2009/939-31/5). Databasen innehåller länkad, landsomfattande, longitudinell information från ett flertal befolkningsbaserade register, men är anonymiserad genom att personuppgifter är borttagna.

⁴ 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.

⁵ När man använder utfallsvariabler med två alternativ (ja/nej), påverkas den statistiska poweren av antalet händelser, t.ex. återfall i brott. Eftersom misstankeregistret innehåller fler registrerade händelser än lagföringsregistret är den statistiska poweren högre.

gav ett år längre uppföljning än lagföringsregistret. Det tar också kortare tid för brott att registreras i misstankeregistret jämfört med lagföringsregistret. Resultatets tillförlitlighet stärks dock om de kan bekräftas med en annan källa och därför följdes deltagarna även upp i lagföringsregistret (Latessa & Holsinger, 1998; Lösel, 2001).

Programgruppen följdes upp med start från frigivning från anstalt eller, om de var frivårdsklienter, från att de avslutade programmet. Jämförelsegruppen följdes upp med start från frigivning 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. Återfall i brott kunde följas upp i misstankeregistret fram till 2010-12-31, vilket gav en genomsnittlig uppföljningstid på 905 dagar (minst 0 dagar, mest 1460 dagar). I lagföringsregistret kunde brottsåterfall däremot bara följas upp fram till 2009-12-31, vilket gav en genomsnittlig uppföljningstid på 548 dagar (minst 0, mest 1095).⁶

Bakgrundsfaktorer

I studier där det inte är slumpmässigt vilka deltagare som ingår i programgrupp respektive jämförelsegrupp riskerar bakgrundskillnader mellan grupperna att göra resultaten missvisande (eng.: biased). I den här typen av undersökningar bör man därför hantera störfaktorer (eng.: confounders), d.v.s. gruppskillnader 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 visade sig samvariera både med deltagande i programmet och med återfall i brott bedömdes som 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 arbete i november månad senaste fem åren
Folk- och bostadsräkningarna (FoB)	1960, 1970, 1980, 1985, 1990	Socioekonomisk status i barndomen ^a
Patientregistret	Slutenvård 1973-2009. Öppenvård 2001-2009	Alkoholmissbruk/-beroende, narkotikamissbruk/-beroende, personlighetsstörning, ADHD, förälder med missbruk/-beroende
Kriminalvårdens register	2002-2009	Påbörjat annat allmänt kriminalitets-/våldsprogram än OTO, missbruksprogram, motivationsprogram, sexualbrotts-/partnervåldsprogram, anstalt- eller frivårdsklient
Flergenerationsregistret	1961-2009 ^b	

^a 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).

^b Personer som varit folkbokförda fr.o.m. 1961 och föddes tidigast 1932 ingår i flergenerationsregistret. Registret kopplades med lagförings- och patientregistret för information om brottslighet och missbruk i familjen.

⁶ Noll dagars uppföljningstid innebär att klienten dött eller utvandrat samma dag eller tidigare än uppföljningsstarten. Det kan också bero på att enstaka frivårdsklienter avslutade OTO år 2010 och därför inte hade någon uppföljningstid i lagföringsregistret (som innehöll data fram till år 2009).

Statistisk analys

För att minska risken att gruppskillnader innan programmet gör resultatet 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).⁷ Coxregression resulterar i effektstorleksmättet hazardkvot, som indikerar hur mycket högre eller lägre risk för utfallet som programdeltagare har i förhållande till jämförelsegruppen under undersökningsperioden. 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 i denna utvärdering även grafiskt med s.k. ”forest plots” (S. Lewis & Clarke, 2001). Vi rapporterar också andelen brottsåterfall inom ett år i program- respektive jämförelsegrupp.⁸ Ett annat sätt att försöka klargöra resultat är att redovisa ”numbers needed to treat” (NNT), d.v.s. antalet klienter som behöver behandlas med OTO för att t.ex. undvika att en person återfaller i brott (Akobeng, 2005). När det fanns signifikanta skillnader mellan grupperna i huvudresultaten beräknades därför även NNT.⁹

⁷ 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”-korrigerig i Coxregressionen blir standardfelen robustare för specifikationsfel, och därmed blir konfidensintervallen inte felaktigt för smala (Lin & Wei, 1989).

⁸ Då många programdeltagare saknade ett helt års uppföljningstid i lagföringsregistret erhöles andelen återfall i brott 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 efter ett år (Rich et al., 2010).

⁹ 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).

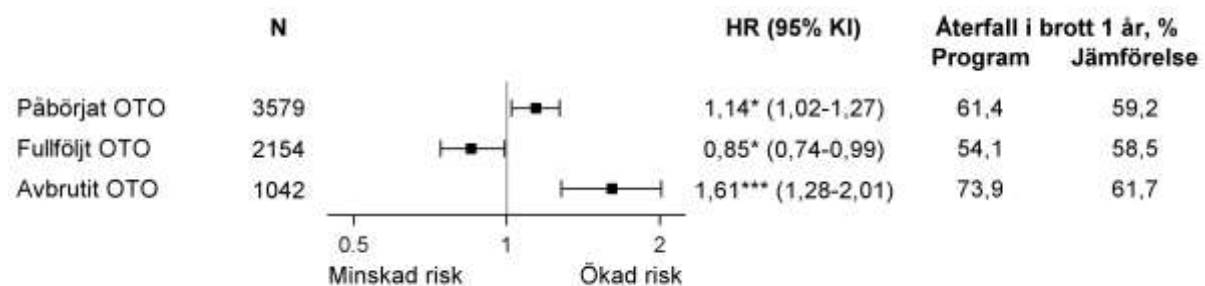
Resultat

Huvudresultat

Kriminalvårdsklienter som *påbörjat* OTO (n=597) jämfördes med matchade klienter som inte påbörjat OTO (n=2985). Matchningen innebär att grupperna har likvärdig bakgrund på de faktorer som hanterades i undersökningen. Klienter som påbörjat OTO hade jämfört med matchade klienter en marginellt om än signifikant ökad risk att återfalla i brott enligt misstankeregistret (se figur 1). Risken att återfalla i brott var 14% högre för dem som påbörjat OTO (absolut riskökning 2,2 procentenheter). Det motsvarar inom ett år att om 45 klienter påbörjar OTO kommer en att misstänkas för nya brott som inte skulle det annars.

Bland de som startade OTO fullföljde 67 procent programmet.¹⁰ Klienter som *fullföljt* OTO (n=359) jämfördes med en ny matchad jämförelsegrupp (n=1795), eftersom deras bakgrund skiljer sig från samtliga som påbörjat programmet. Jämfört med matchade klienter hade de som fullföljt OTO signifikant lägre risk att återfalla i brott enligt misstankeregistret. Fullföljares risk att återfalla i brott var 15% lägre (absolut riskminskning 4,4 procentenheter). Inom ett år innebär det att om 23 klienter fullföljer OTO förhindras att en av dem återfaller i brott.

Klienter som påbörjat men sedan *avbrutit* OTO (n=176) jämfördes med ytterligare en matchad jämförelsegrupp med liknande bakgrund (n=867). Deltagare i OTO som avbrutit programmet hade signifikant ökad risk att misstänkas för nya brott jämfört med den matchade jämförelsegruppen. Risken att återfalla i brott var 61% högre för dem som avbröt OTO (absolut riskökning 12,2 procentenheter). Inom ett år innebär det att om 8 klienter avbryter OTO skulle en misstänkas för nya brott som inte skulle det annars.



N = antal analyserade totalt, * = p < 0,05, *** = p < 0,001

Figur 1. Hazardkvoter [HR] med 95% konfidensintervall [KI] för återfall i brott enligt misstankeregistret. Deltagare i One-to-One jämfördes med matchade klienter som inte deltagit i programmet.¹¹

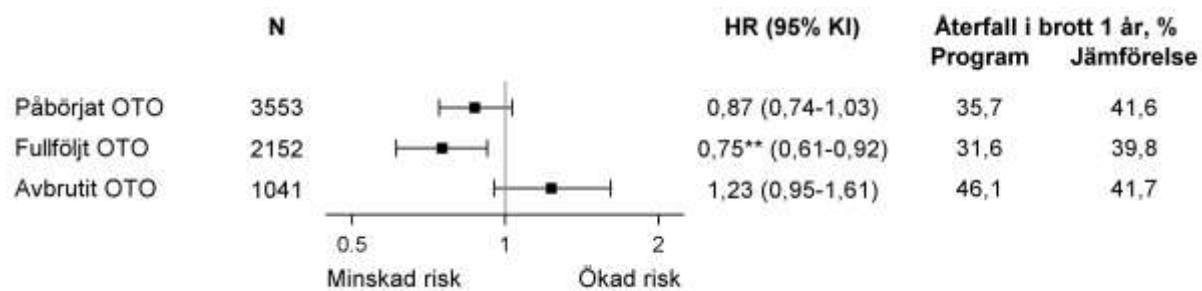
Känslighetsanalys

Resultatens tillförlitlighet stärks om de kan bekräftas med en annan källa och därför följdes deltagarna i en känslighetsanalys även upp i lagföringsregistret. När deltagarna följdes upp i lagföringsregistret istället för i misstankeregistret var resultaten likartade (se figur 2). Enligt lagföringsregistret fanns till skillnad från misstankeregistret ingen signifikant skillnad mellan de som påbörjat OTO och matchade klienter.

Däremot hade de som fullföljt OTO åter signifikant lägre risk att återfalla i brott. Klienter som avbrutit OTO hade som tidigare ökad risk att dömas för brott, men skillnaden var inte statistiskt säkerställd.

¹⁰ 359 klienter fullföljde OTO och 177 avbröt programmet. För 61 programdeltagare saknades information om de fullföljt eller avbrutit OTO.

¹¹ 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.



N = antal analyserade totalt, ** = $p < 0,01$

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

Diskussion

För de som påbörjat OTO kunde ingen statistiskt säkerställd riskminskning hittas.¹² De som fullföljde programmet hade däremot en statistiskt säkerställd minskad risk att återfalla i brott oavsett vilket register som användes för uppföljning. Våra resultat stämmer därmed överens med Kriminalvårdens tidigare utvärdering av OTO, där de som påbörjat programmet hade i stort sett lika hög risk att återfalla i brott som jämförelsegruppen, medan klienter som fullföljt OTO hade signifikant minskad risk att återfalla i brott (Danielsson et al., 2009). Enligt systematiska översikter minskar KBT-baserad behandling för kriminella i genomsnitt återfall i brott (Lipsey et al., 2007; Wilson et al., 2005). KBT-baserade program som likt OTO innehåller kognitiv omstrukturering har dessutom visat större effekt på återfall i brott (Lipsey et al., 2007).

Att de som fullföljt OTO mindre ofta återfall i brott kan tolkas som att programmet har brottsreducerande effekt, under förutsättning att det genomförs fullt ut. Det finns dock aspekter som gör den här slutsatsen osäker. De som fullföljer behandlingsprogram brukar nämligen ofta ha 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 OTO och jämförelsegruppen, som gör resultatet missvisande (s.k. bias). Eftersom vi t.ex. inte kunde hantera eventuella gruppskillnader i motivation kan de som fullföljde OTO ha varit mer motiverade redan från början, och därmed mer benägna både att slutföra programmet och återfalla mindre i brott. Att tidigare forskning tyder på att behandlingsprogram med liknande innehåll som OTO minskar brottsåterfall talar dock emot att den positiva effekten enbart skulle bero på bias.

Den ökade risken att återfalla i brott bland dem som avbröt OTO stämmer överens med studier av andra behandlingsprogram (McMurrin & Theodosi, 2007; Olver et al., 2011). Att misslyckas med att slutföra sin behandling kan tänkas öka klientens uppgivenhet och förstärka antiauktoritära och antisociala attityder. Det är dock oklart om riskökningen verkligen är en effekt just av behandlingsavbrottet; den skulle även kunna bero på att de som avbryter behandling har svårare problematik och mindre motivation redan innan programstart, trots att vi med propensity score-matchning försökte hantera sådana bakgrundsskillnader.

Enligt en forskningsöversikt minskar KBT-baserade behandlingsprogram i genomsnitt återfall i brott med 25% (från 40% till 30%) (Lipsey et al., 2007). I de flesta utvärderingarna i översikten jämfördes klienter som *påbörjat* behandling med en jämförelsegrupp. I utvärderingar där man istället utgått från dem som fullföljt behandling hittar man oftast ännu större skillnader i återfall i brott. I den här studien fann vi inte någon tydlig minskning bland dem som påbörjat OTO, men *fullföljande* av programmet minskade risken för återfall i brott med 15-25% beroende på vilket register som användes. Att vi inte hittade lika stor och tydlig programeffekt som tidigare påvisats för andra återfallsreducerande program kan bero på sammanhanget. Behandling som blivit del av verksamhetens ordinarie programutbud har ofta mindre effekt på återfall i brott än behandlingsprogram som införs på prov (Andrews, 2011; Lipsey et al., 2007). Ibland hittar man ingen effekt alls när behandlingsprogram utförs i ordinarie verksamhet, även om man hittat programeffekter på återfall i brott i tidigare studier (Goggin & Gendreau, 2006). När behandlingsprogram införs i stor skala i rutinverksamhet blir det bl.a. svårare att kontrollera att de utförs som avsett, s.k. programtrohet (Andrews, 2011) vilket påverkar effekten på återfall i brott (Andrews & Dowden, 2005). Vi undersökte inte om OTO utfördes med programtrohet, men i Kriminalvården försöker man upprätthålla programtrohet genom certifiering och handledning. Utan att undersöka saken kan vi dock inte veta om dessa åtgärder är tillräckliga för att OTO ska utföras med programtrohet. Utifrån tidigare behandlingsforskning kan det alltså vara så att programmet skulle få bättre effekt i Kriminalvården om programtroheten ökade.

¹² Klienter som påbörjat OTO hade marginellt högre risk att återfalla i brott enligt misstankeregistret, men inte enligt lagföringsregistret. Skillnader i den storleksordningen som hittades i misstankeregistret brukar betraktas som försumbara (Bedard, Krzyzanowska, Pintilie, & Tannock, 2007). 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. Skillnad i uppföljningstid kan ha betydelse eftersom programeffekter kan variera över tid

Behandlingsprogram i ordinarie verksamhet brukar utöver att ofta ha sämre programtrohet mer sällan leva upp till principerna om risk, behov och mottaglighet (RBM) (Andrews, 2011) vilket också reducerar den återfallsförebyggande effekten (Dowden & Andrews, 2000). Kortfattat innebär RBM att prioritera behandling till kriminella med medelhög till hög risk att återfalla i brott, fokusera på brottsdrivande faktorer samt att anpassa programmets utförande efter deltagarnas inlärningsstil. Det finns dock tecken på att OTO utfördes i enlighet med principerna om RBM under undersökningsperioden.¹³

Effektutvärderingar kan delas in i två typer. I den första utvärderas om en behandling har avsedd effekt under *idealiska omständigheter* (eng. efficacy trial eller explanatory trial). I den andra utvärderingstypen undersöker man istället om behandlingen har önskad effekt under *normala omständigheter* (eng. effectiveness trial eller pragmatic trial). För att kunna uttala sig om ifall OTO överhuvudtaget kan minska återfall i brott behöver programmet alltså utvärderas under idealiska omständigheter. Den här rapporten är dock en utvärdering av OTO så som programmet normalt utförs i Kriminalvården. Å ena sidan är det därför möjligt att resultaten skulle vara annorlunda i en utvärdering med strikt kontroll över bl.a. urvalskriterier, programutförande och programledarnas kompetens. Å andra sidan har utvärderingar som utförs under idealiska omständigheter inte lika stor praktisk relevans, eftersom behandling sällan utförs så i praktiken (Trewick & Zwarenstein, 2009).

¹³ Målgruppen är klienter med medel eller hög risk för återfall i brott. Inspektion av bakgrundsdata för programdeltagarna i denna utvärdering antyder att den genomsnittliga deltagaren hade medelhög eller hög risk (se tabellerna i bilaga B). Däremot vet vi inte i vilken utsträckning som riskbedömningsinstrument användes för att välja ut deltagare till programmet. OTO ska bl.a. förbättra deltagarnas självkontroll; ett viktigt kriminogent behov. Dessutom anses användning av tekniker från KBT, som i OTO, stärka deltagarnas mottaglighet (Andrews et al., 1990). Att OTO genomförs individuellt borde även göra det lättare för programledare att anpassa programmets utförande efter den enskilda 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
- 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., Pintilie, 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.
- 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
- Danielsson, M., Fors, A., & Freij, I. (2009). *Behandlingsprogrammet One to One. Utvärdering av återfall i ny brottslighet för programdeltagare för åren 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
- Goggin, C., & Gendreau, P. (2006). The implementation and maintenance of quality services in offender rehabilitation programmes. I C. Hollin & E. Palmer (red.), *Offending behaviour programmes: Development, application and controversies*. Chichester: John Wiley.
- 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. One to one*. 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.
- McMurrin, M., & Theodosi, E. (2007). Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime & Law*, 13(4), 333-343. doi: 10.1080/10683160601060374
- 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.
- 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
- Priestley, P. (2003). *One-to-one: Ett kognitivt och beteendebaserat program för att reducera återfall i brott. Teori och Empiri*. Norrköping: Kriminalvården.
- Priestley, P. (2006). *OTO: Ett program för att minska återfall i brott. Manual*. Norrköping: Kriminalvården.
- 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.
- 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
- 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
- Wilson, D. B., Bouffard, L. A., & Mackenzie, D. L. (2005). A quantitative review of structured, group-oriented, cognitive-behavioral programs for offenders. *Criminal Justice and Behavior*, 32(2), 172-204. doi: 10.1177/0093854804272889

Bilagor

Bilaga A. Metodfördjupning

Matchning

Matchningen utfördes i statistikprogrammet SAS via ett publicerat s.k. makro (Coca-Perrailon, 2007). De bakgrundsfaktorer som skulle användas i matchningen valdes utifrån om variabeln uppvisade ett signifikant samband ($p < 0,05$) med programdeltagande och *desutom* var signifikant associerad med återfall i brott. Det beror på att det bara är variabler som både påverkar utfallet och har olika fördelning i program- och jämförelsegrupp som riskerar förvränga (eng.: bias) resultat (Bonita, Beaglehole, & Kjellström, 2006). Variabeln anstalts-/frivårdsklient hade för klienter som avbrutit OTO inte signifikant gruppsskillnad *före* matchning, och ingick därmed inte i beräkning av propensity scores, men blev skevt fördelad mellan grupperna *efter* matchning. Eftersom variabeln hade signifikant samband med återfall i brott användes den som kontrollvariabel i propensity score-beräkningen. Detta för att förhindra att matchningen resulterade i en skillnad mellan grupperna som i sin tur riskerat att göra resultatet skevt.

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 för varje programdeltagare matchades den individ 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. För klienter som avbrutit OTO angavs ett maxavstånd (eng.: "caliper") för hur stor skillnaden mellan programdeltagaren och jämförelsepersonen som mest fick vara för att de skulle matchas. Om grupperna skiljer sig mycket åt före matchning kan vissa i jämförelsegruppen som är mest lik någon av programdeltagarna ändå vara relativt olik denne. För att undvika dålig matchning kan man därför ange ett maxavstånd för hur stor skillnaden i propensity score får vara (Coca-Perrailon, 2007).

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 dataset till dataset. Målet med multipel imputation är inte att ersätta bortfall med det sanna värdet, utan 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 då 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:

- variabler som ska ingå i senare analys (inklusive utfallsmått)
- variabler som korrelerar med bortfall
- 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 arbete registrerat i november 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 (Mittra & 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
- Coca-Perraillon, 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

Bilaga B. Bakgrundsfaktorer före och efter matchning

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

	Före matchning			Efter matchning		
	Påbörjat OTO (n=597)	Omatchad jämförelsegrupp (n=56232)	Cohens d	Påbörjat OTO (n=597)	Matchad jämförelsegrupp (n=2985)	Cohens d
Ålder år 2013, m	35,2 år	41,8 år	0,50	35,2 år	35,1 år	-0,01
Kön, %			0,01			0,01
Kvinna	10,7	11,6		10,7	10,4	
Man	89,3	88,4		89,3	89,6	
Anstaltsklient, %	34,8	30,7	0,02	34,8	35,8	0,01
Civilstatus, %			0,04			0,00
Gift/sambo med gemensamma barn	11,9	19,7		11,9	11,9	
Ogift/sambo utan gemensamma barn	87,8	79,8		87,8	88,0	
Information saknas	0,3	0,6		0,3	0,2	
Bostadsområde, %			0,06			0,01
Stad/förort	75,2	60,8		75,2	75,8	
Annan ort	24,5	38,0		24,5	24,1	
Information saknas	0,3	1,2		0,3	0,1	
Högsta utbildningsnivå, %			0,09			0,07
Folk- eller grundskola < 9 år	5,0	6,6		5,0	6,3	
Folk- eller grundskola 9 år	56,1	36,1		56,1	52,7	
Gymnasial utbildning 1-3 år	32,0	45,7		32,0	33,7	
Universitet och högskola, inkl. gymnasium 4 år	3,2	8,0		3,2	3,2	
Information saknas	3,7	3,6		3,7	4,1	
SES i barndom, %			0,02			0,03
Låg	42,0	35,5		42,0	41,3	
Medel	20,4	20,6		20,4	21,5	
Hög	12,6	13,2		12,6	11,7	
Information saknas	25,0	30,6		25,0	25,6	
<i>Tidigare brottslighet</i>						
Antal tidigare domar, m	12,2	8,9	-0,25	12,2	11,8	-0,03
Ålder vid första dom, m	18,5 år	24,3 år	0,52	18,5 år	18,5 år	0,00
Ngn gång dömd för våldsbrott, %	76,9	61,2	0,06	76,9	77,6	0,01
Ngn gång dömd för narkotikabrott, %	63,2	38,6	0,10	63,2	63,3	0,00
Ngn gång dömd för rattfylleri, %	43,2	50,8	0,03	43,2	43,3	0,00
Information om tidigare brottslighet saknas, %	0,0	0,4		0,0	0,0	
<i>Familjär brottslighet/missbruk</i>						
Mamma ngn gång dömd, %	34,0	21,1	0,06	34,0	33,8	0,00
Pappa ngn gång dömd, %	62,8	43,8	0,08	62,8	62,8	0,00
Ngt syskon som ngn gång dömts, %	40,5	33,6	0,03	40,5	39,5	0,02
Ngn förälder med missbruk/beroende	21,1	17,0	0,02	21,1	20,2	0,02
<i>Psykatri</i>						
Alkoholmissbruk/-beroende, %	16,6	22,3	0,03	16,6	16,0	0,01
Narkotikamissbruk/-beroende, %	30,3	21,7	0,04	30,3	30,2	0,00
Personlighetsstörning, %	7,4	5,6	0,01	7,4	6,5	0,03
ADHD, %	9,1	3,4	0,06	9,1	9,8	0,02

Fortsättning tabell B1.

	Före matchning			Efter matchning		
	Påbörjat OTO (n=597)	Omatchad jämförelsegrupp (n=56232)	Cohens d	Påbörjat OTO (n=597)	Matchad jämförelsegrupp (n=2985)	Cohens d
<i>Behandlingsprogram</i>						
Påbörjat missbruksprogram, %	17,6	6,1	0,10	17,6	18,1	0,01
Påbörjat sexualbrotts-/partnervåldsprogram, %	0,8	1,2	0,01	0,8	0,5	0,03
Påbörjat motivationsprogram, %	36,9	6,6	0,24	36,9	36,7	0,00
Påbörjat <i>annat</i> allmänt kriminalitets-/våldsprogram än OTO, %	20,6	4,3	0,16	20,6	20,5	0,00
Antal år med arbete i november senaste 5 åren, m	0,9 år	1,8 år	0,46	0,9 år	0,8 år	-0,02
Information om arbete saknas, %	0,3	0,6		0,3	0,2	

Tabell B2. Jämförelse av eventuella bakgrundsskillnader mellan dem som *fullföljt* One-to-One (OTO) och jämförelsegruppen före och efter propensity score-matchning (1:5 nearest neighbor).

	Före matchning			Efter matchning		
	Fullföljt OTO (n=359)	Omatchad jämförelsegrupp (n=56232)	Cohens d	Fullföljt OTO (n=359)	Matchad jämförelsegrupp (n=1795)	Cohens d
Ålder år 2013, m	35,4 år	41,8 år	0,48	35,4 år	35,2 år	-0,02
Kön, %			0,02			0,03
Kvinna	7,8	11,6		7,8	8,9	
Man	92,2	88,4		92,2	91,3	
Anstaltsklient, %	43,5	30,7	0,04	43,5	42,8	0,01
Civilstatus, %			0,02			0,02
Gift/sambo med gemensamma barn	13,7	19,7		13,7	14,6	
Ogift/sambo utan gemensamma barn	85,8	79,8		85,8	84,9	
Information saknas	0,6	0,6		0,6	0,6	
Bostadsområde, %			0,06			0,01
Stad/förort	78,6	60,8		78,6	79,4	
Annan ort	20,9	38,0		20,9	20,1	
Information saknas	0,6	1,2		0,6	0,5	
Högsta utbildningsnivå, %			0,06			0,06
Folk- eller grundskola < 9 år	3,6	6,6		3,6	5,1	
Folk- eller grundskola 9 år	51,8	36,1		51,8	50,1	
Gymnasial utbildning 1-3 år	36,5	45,7		36,5	36,1	
Universitet och högskola, inkl. gymnasium 4 år	4,2	8,0		4,2	4,0	
Information saknas	3,9	3,6		3,9	4,7	
SES i barndom, %			0,01			0,03
Låg	41,8	35,5		41,8	36,0	
Medel	20,9	20,6		20,9	19,6	
Hög	13,7	13,2		13,7	11,9	
Information saknas	23,7	30,6		23,7	32,5	
<i>Tidigare brottslighet</i>						
Antal tidigare domar, m	12,1	8,9	-0,25	12,1	12,3	0,01
Ålder vid första dom, m	18,6 år	24,3 år	0,51	18,6 år	18,6 år	-0,01
Ngn gång dömd för våldsbrott, %	74,1	61,2	0,04	74,1	74,8	0,01
Ngn gång dömd för narkotikabrott, %	63,8	38,6	0,08	63,8	62,9	0,01
Ngn gång dömd för rattfylleri, %	40,7	50,8	0,03	40,7	41,0	0,00
Information om tidigare brottslighet saknas, %		0,4				
<i>Familjär brottslighet/missbruk</i>						
Mamma ngn gång dömd, %	33,2	21,1	0,05	33,2	33,1	0,00
Pappa ngn gång dömd, %	62,4	43,8	0,06	62,4	63,5	0,02
Ngt syskon som ngn gång dömts, %	40,4	33,6	0,02	40,4	40,4	0,00
Ngn förälder med missbruk/beroende	20,9	17,0	0,02	20,9	22,7	0,01
<i>Psykatri</i>						
Alkoholmissbruk/-beroende, %	15,0	22,3	0,03	15,0	16,3	0,03
Narkotikamissbruk/-beroende, %	26,2	21,7	0,02	26,2	27,5	0,02
Personlighetsstörning, %	5,6	5,6	0,00	5,6	6,2	0,02
ADHD, %	7,5	3,4	0,04	7,5	7,0	0,01

Fortsättning tabell B2.

	Före matchning			Efter matchning		
	Fullföljt OTO (n=359)	Omatchad jämförelsegrupp (n=56232)	Cohens d	Fullföljt OTO (n=359)	Matchad jämförelsegrupp (n=1795)	Cohens d
<i>Behandlingsprogram</i>						
Påbörjat missbruksprogram, %	19,8	6,1	0,09	19,8	18,9	0,02
Påbörjat sexbrotts- /partnervåldsprogram, %	0,6	1,2	0,01	0,6	1,1	0,04
Påbörjat motivationsprogram, %	36,8	6,6	0,19	36,8	37,5	0,01
Påbörjat <i>annat</i> allmänt kriminalitets- /våldsprogram än OTO, %	22,8	4,3	0,14	22,8	21,8	0,02
Antal år med arbete i november senaste 5 åren, m	0,9 år	1,8 år	0,43	0,9 år	0,9 år	-0,02
Information om arbete saknas, %	0,6	0,6		0,6	0,6	

Tabell B3. Jämförelse av eventuella bakgrundsskillnader mellan dem som *avbrutit* One-to-One (OTO) och jämförelsegruppen före och efter propensity score-matchning (1:5 nearest neighbor med caliper).

	Före matchning			Efter matchning		
	Avbrutit OTO (n=177)	Omatchad jämförelsegrupp (n=56232)	Cohens d	Avbrutit OTO (n=176)	Matchad jämförelsegrupp (n=867)	Cohens d
Ålder år 2013, m	35,0 år	41,8 år	0,51	35,0 år	34,5 år	-0,05
Kön, %			0,01			0,02
Kvinna	14,1	11,6		13,6	12,8	
Man	85,9	88,4		86,4	87,2	
Anstaltsklient, %	28,3	30,7	0,01	28,4	29,0	0,01
Civilstatus, %			0,03			0,02
Gift/sambo med gemensamma barn	9,6	19,7		9,7	10,6	
Ogift/sambo utan gemensamma barn	90,4	79,8		90,3	89,4	
Information saknas	0,0	0,6		0,0	0,0	
Bostadsområde, %			0,02			0,03
Stad/förort	72,3	60,8		72,2	73,7	
Annan ort	27,7	38,0		27,8	26,3	
Information saknas		1,2		0,0	0,0	
Högsta utbildningsnivå, %			0,07			0,10
Folk- eller grundskola < 9 år	8,5	6,6		8,5	9,8	
Folk- eller grundskola 9 år	65,5	36,1		65,3	65,3	
Gymnasial utbildning 1-3 år	20,3	45,7		20,5	21,0	
Universitet och högskola, inkl. gymnasium 4 år	2,3	8,0		2,3	0,9	
Information saknas	3,4	3,6		3,4	3,0	
SES i barndom, %			0,02			0,03
Låg	42,4	35,5		42,6	38,4	
Medel	18,6	20,6		18,8	18,5	
Hög	11,3	13,2		11,4	10,6	
Information saknas	27,7	30,6		27,3	32,5	
<i>Tidigare brottslighet</i>						
Antal tidigare domar, m	12,4	8,9	-0,26	12,2	11,8	-0,03
Ålder vid första dom, m	18,3 år	24,3 år	0,54	18,3 år	18,3 år	-0,02
Ngn gång dömd för våldsbrott, %	80,2	61,2	0,04	80,1	81,4	0,03
Ngn gång dömd för narkotikabrott, %	63,8	38,6	0,06	63,6	61,6	0,03
Ngn gång dömd för rattfylleri, %	48,0	50,8	0,01	47,7	43,5	0,06
Information om tidigare brottslighet saknas, %	0,0	0,4		0,0	0,0	
<i>Familjär brottslighet/missbruk</i>						
Mamma ngn gång dömd, %	35,6	21,1	0,04	35,2	35,0	0,00
Pappa ngn gång dömd, %	63,3	43,8	0,04	63,1	63,1	0,00
Ngt syskon som ngn gång dömts, %	41,2	33,6	0,02	41,5	39,9	0,02
Ngn förälder med missbruk/beroende	20,3	17,0	0,01	19,9	24,9	0,09
<i>Psykatri</i>						
Alkoholmissbruk/-beroende, %	19,8	22,3	0,01	19,3	20,1	0,02
Narkotikamissbruk/-beroende, %	37,3	21,7	0,04	36,9	35,5	0,02
Personlighetsstörning, %	10,2	5,6	0,02	9,7	9,0	0,02
ADHD, %	9,0	3,4	0,03	8,5	10,3	0,04

Fortsättning tabell B3.

	Före matchning			Efter matchning		
	Avbrutit OTO (n=177)	Omatchad jämförelsegrupp (n=56232)	Cohens d	Avbrutit OTO (n=176)	Matchad jämförelsegrupp (n=867)	Cohens d
<i>Behandlingsprogram</i>						
Påbörjat missbruksprogram, %	14,1	6,1	0,04	13,6	13,4	0,01
Påbörjat sexbrotts-/partnervåldsprogram, %	1,7	1,2	0,01	1,7	0,7	0,08
Påbörjat motivationsprogram, %	36,2	6,6	0,13	35,8	33,2	0,04
Påbörjat <i>annat</i> allmänt kriminalitets-/våldsprogram än OTO, %	16,4	4,3	0,07	16,5	13,0	0,08
Antal år med arbete i november senaste 5 åren, m	0,7	1,8	0,55	0,7	0,6	-0,03
Information om arbete saknas, %						