

Utvärdering av programmet VINN i Kriminalvården

VI BRYTER DEN ONDA CIRKELN



Sammanfattning

Bakgrund

VINN är ett program för kvinnor som utvecklades av norska kriminalvården och infördes i Sverige 2007. Programmets mål är att öka deltagarnas livskvalitet och att minska återfall i brott. Programmet bygger främst på förändringsmodellen Känsla Av SAMmanhang (KASAM) som har sitt ursprung i hälsoforskning om stressande händelser. Programmet använder även principer från social kognitiv teori, motiverande intervju och den transteoretiska modellen.

Inga tidigare utvärderingar av VINN:s effekt kunde hittas. Däremot finns effektutvärderingar av andra program för kvinnor, vilka har vägts samman i forskningsöversikter. Enligt en sådan översikt verkar kriminalvårdsprogram för kvinnor i genomsnitt minska återfall i brott med 29 %. Likt program för män tycks program för kvinnor bland annat ha större effekt om de använder metoder med stöd i kognitiv beteendeterapi (KBT) eller social kognitiv teori.

Syftet med den här utvärderingen var att undersöka om VINN minskar risken att återfalla i brott bland kvinnliga kriminalvårdsklienter i Sverige.

Metod

I utvärderingen ingick 194 kvinnor som påbörjat VINN och som frigavs från anstalt eller skrevs in i frivård 2007-2009. Av dessa var det 82 % som fullföljde programmet. Programdeltagarna jämfördes med lika många individuellt matchade kvinnor som var aktuella i Kriminalvården under samma period men som inte påbörjade VINN. För att minska risken att bakgrundsskillnader gör resultaten missvisande användes så kallad propensity score-matchning. Det innebär att vi hanterade eventuella skillnader mellan grupperna före programmet i till exempel ålder, antal tidigare domar och alkohol- eller narkotikaberoende.

Resultat

Bland de som fullföljt VINN var andelen återfall i brott inom ett år 49 % jämfört med 59 % i den matchade jämförelsegruppen. Både de som påbörjat och de som fullföljt VINN hade något minskad risk att återfalla i brott jämfört med matchade klienter med likvärdig bakgrund. Om riskminskningarna kunde säkerställas statistiskt eller inte berodde dock på vilket register som användes.

Diskussion

Eftersom alla analyser tyder på minskad risk att återfalla i brott är resultaten lovande. Men då få klienter ingick i undersökningen var det svårt att statistiskt säkerställa resultat skillnaderna. För att kunna dra säkrare slutsatser i en framtida utvärdering behövs därför fler programdeltagare. Att VINN verkar minska återfall i brott stöds av tidigare utvärderingar av kvinnoprogram som också baseras på social kognitiv teori. Utöver att minska återfall i brott har VINN som mål att öka deltagarnas livskvalitet. Den här utvärderingen säger dock inget om programmets effekt på detta.

Innehållsförteckning

Sammanfattning	1
Introduktion	3
Tidigare forskning	4
Problemformulering och syfte.....	5
Metod	6
Undersökningsgrupp.....	6
Källor.....	6
Utfall.....	7
Bakgrundsfaktorer.....	7
Statistisk analys.....	8
Resultat	10
Huvudresultat.....	10
Känslighetsanalys.....	10
Diskussion	11
Referenser	13
Bilagor	16
Bilaga A. Metodfördjupning.....	16
Bilaga B. Bakgrundsfaktorer före och efter matchning.....	19

Omslag: Kriminalvården
Tryckning: Kriminalvården Digitaltryck 2016
Beställningsnummer: 7184
ISBN: 978-91-86903-77-0
Kriminalvården
601 80 Norrköping
Tel: 077-22 80 800
E-post: hk@kriminalvarden.se
Hemsida: www.kriminalvarden.se

Introduktion

De flesta som begår brott är män och därför har behandlingsprogram i kriminalvård historiskt främst baserats på män. Kvinnor i kriminalvård har inte sällan något annorlunda problembild än kriminella män, t.ex. har de oftare narkotikamissbruk, posttraumatiskt stressyndrom (PTSD) och erfarenhet av sexuella övergrepp (Carlson & Shafer, 2010; Fazel, Bains, & Doll, 2006; Goff, Rose, Rose, & Purves, 2007; Wolff, Shi, & Siegel, 2009). Dessutom kan kvinnor i högre utsträckning ha en relationsbaserad inlärningsstil (Dowden & Andrews, 1999; Rettinger & Andrews, 2010). Utöver att missbruk är en viktig riskfaktor för kriminalitet kan dessa skillnader även påverka kvinnors mottaglighet för återfallsförebyggande program. I förhoppning att kunna förbättra resultaten har man därför utvecklat särskilda program för kvinnor. Utöver att minska återfall i brott har dessa många gånger som mål att också öka deltagarnas psykiska välbefinnande (Tripodi, Bledsoe, Kim, & Bender, 2011).

VINN är ett motivationsprogram för kvinnor inom kriminalvård eller missbruksbehandling som utvecklades av den norska kriminalvården i början av 2000-talet. Programmets mål är att motivera deltagarna att göra val som förbättrar deras livskvalitet och att förebygga återfall i brott. VINN är multimodalt, d.v.s. det bygger på flera teoretiska modeller och är riktat mot flera problemområden. Programmets huvudsakliga teori är Känsla Av SAMmanhang (KASAM) (Antonovsky, 2005), men kompletteras också med social kognitiv teori (Bandura, 1997), motiverande intervju (Miller & Rollnick, 2013) och den transteoretiska modellen (Prochaska & Velicer, 1997). KASAM har sitt ursprung i hälsorelaterad forskning om stressfulla händelser. Kortfattat går teorin ut på att ens förmåga att hantera svårigheter i livet ökar om man anser att:

- det som sker är förutsägbart, begripligt och strukturerat (*begriplighet*)
- man har de resurser som krävs (*hanterbarhet*)
- att livets krav och utmaningar är värda att engagera sig i (*meningsfullhet*)

Genom VINN ska deltagarnas motståndskraft mot svårigheter öka och därigenom deras livskvalitet förbättras (Højdahl, Magnus, Hagen, & Langeland, 2013; Højdahl & Størksen, 2012b).

Programmet består av 14-16 gruppmöten om cirka 3 timmar (inklusive pauser) som sker 2-3 gånger i veckan. Dessutom ingår 4 individuella samtal där man bl.a. diskuterar den enskilde deltagarens behov. Mellan träffarna ska deltagarna använda ett arbetshäfte (Højdahl & Størksen, 2012a). VINN innehåller 13 teman som tas upp under ett möte vardera (se tabell 1). Inom dessa teman behandlas flera dynamiska kriminogena behov (föränderliga riskfaktorer för brott), t.ex. antisociala attityder, prokriminella relationer, missbruk och dåliga familjerelationer. Dessutom tar man även upp andra problemområden som ofta orsakar nedsatt livskvalitet hos kvinnor i kriminalvård, t.ex. låg självkänsla samt sex och samlevnad (Højdahl & Størksen, 2012b).

Tabell 1. Teman i VINN

1. Lära känna varandra	8. Sexualitet och kärlek
2. Identitet och självbild	9. Sorg och förluster
3. Kommunikation	10. Ilska
4. Brott	11. Våld
5. Förändring och val	12. Gränser
6. Ekonomi	13. Nätverk och relationer
7. Missbruk och beroende	

Not: Ämnet "missbruk och beroende" är inte obligatoriskt för deltagare utan missbruksproblem och kan också utökas till 2-3 möten om gruppdeltagarnas problem är stora.

VINN ges i grupper om 5-8 deltagare och leds av två programledare som är utbildade i programmet. Efter fem programomgångar måste programledare certifieras för att få fortsätta ge programmet. Certifiering innebär att programmets genomförande har granskats och bedömts som godkänt, bl.a. via videospelning och metodhandledning (Højdahl & Størksen, 2012b; Kriminalvården, 2009).

För varje kriminalvårdsklient upprättas en s.k. verkställighetsplan, där personal försöker identifiera klientens risker och behov. Om det framkommer att klienten har behov av behandling ska ett lämpligt program föreslås som man sedan försöker motivera denne att delta i (Kriminalvården, 2008). Klienter kan

även själva ta initiativ till att delta i VINN (Høj Dahl & Størksen, 2012b). Programledarna i VINN intervjuar alla potentiella programdeltagare för att avgöra deras lämplighet. För att delta i programmet ska personen vara öppen för att dela sina erfarenheter i grupp och motiverad för en livsstil utan kriminalitet och missbruk. Klienten får inte vara allvarligt psykisk sjuk och måste tala och förstå svenska för att kunna delta i programmet (Høj Dahl & Størksen, 2012b; Kriminalvården, 2013). Att delta i Kriminalvårdens program är frivilligt, men för frivårdsklienter kan behandling ibland vara en del av påföljden.¹

VINN är svenska Kriminalvårdens enda program riktat specifikt till kvinnor. Det infördes år 2007 och ackrediterades år 2010. Ackreditering innebär att programmet granskas av en panel erfarna kliniker och forskare från Kriminalvårdens vetenskapliga råd. För ackreditering krävs bl.a. en vetenskapligt understödd förändringsmodell, tydlig målgrupp, utarbetade manualer samt att programmet riktas mot dynamiska riskfaktorer för kriminalitet.

Tidigare forskning

Någon tidigare publicerad effektutvärdering av VINN finns inte till vår kännedom.² Däremot finns effektutvärderingar av andra kriminalvårdsprogram för kvinnor, vilka sammanfattats i systematiska forskningsöversikter. I en av dessa undersöktes betydelsen av principerna om risk, behov och mottaglighet (RBM) för behandling av kvinnliga klienter (Dowden & Andrews, 1999). RBM innebär kortfattat att rikta behandlingen till kriminella med högre risk att återfalla i brott, fokusera på kriminogena behov samt anpassa programmets utförande efter deltagarnas inlärningsstil. För att ingå i översikten skulle studien innehålla en jämförelsegrupp och de flesta deltagare vara kvinnor. Det innebar att 26 studier ingick i översikten, varav bara kvinnor ingick i 16 av studierna. De program som enbart gavs till kvinnor minskade överlag återfall i någon brottslighet med 29%, vilket är i samma storleksordning som i studier som främst baserats på män (Lipsey, Landenberger, & Wilson, 2007; Wilson, Bouffard, & Mackenzie, 2005). Störst effekt på återfall i brott bland kvinnor hade program som gavs till medel- och högrisk klienter, riktades mot kriminogena behov och använde metoder med stöd i kognitiv beteendeterapi eller social kognitiv teori (Dowden & Andrews, 1999). Översiktens slutsatser påverkades inte nämnvärt av studiernas varierande metodkvalitet (Andrews & Dowden, 1999).

I en senare forskningsöversikt ingick 6 behandlingsutvärderingar om återfall i brott hos kvinnor i anstalt (Tripodi et al., 2011). Alla studierna handlade om effekten av terapeutiska samhällen mot missbruk (eng: therapeutic communities). Enbart amerikanska utvärderingar ingick och de flesta var observationsstudier med jämförelsegrupp, d.v.s. det var inte slumpmässigt vem som ingick i program- respektive jämförelsegrupp. Baserat på översikten minskade programmen överlag kvinnornas risk att återfalla i brott med 30%.³

Sammanfattningsvis verkar effektiva återfallsförebyggande program för män i kriminalvård fungera lika väl för kvinnor. VINN är riktat mot flera kända kriminogena behov och använder principer från social kognitiv teori (Høj Dahl & Størksen, 2012b). Delar av programmet har därmed stöd i forskning på andra program för kvinnor (Dowden & Andrews, 1999). Även om VINN alltså har vissa likheter med tidigare utvärderade program så har den också unika egenskaper. Programmets främsta förändringsmodell är KASAM och någon utvärdering av program med samma utgångspunkt hittades inte. Det är därför osäkert om resultat från tidigare studier kan generaliseras till VINN.

¹ 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.

² Även om effektutvärderingar tycks saknas, så finns studier där man frågat VINN-deltagare om deras upplevelse av programmet (Granheim, 2007; Høj Dahl, Magnus, & Langeland, 2014). Enligt en utvärdering från norska kriminalvården var programdeltagare i Norge och Sverige överlag nöjda med programmets innehåll (Granheim, 2007). I en kvalitativ studie intervjuades VINN-deltagare i Norge, Sverige, Danmark, Estland och Ryssland (Høj Dahl et al., 2014). Även i denna studie uttryckte deltagarna uppskattning för programmet.

³ I översikten redovisades resultaten i form av oddskvot. För att få fram minskad risk att återfalla i brott konverterades oddskvoten till riskkvot genom en kalkylator (Kane, 2014).

Idag finns relativt få välgjorda utvärderingar av kriminalvårdsprogram för kvinnor och därför förekommer kunskapsluckor på området. Utöver traditionella kriminogena behov, som missbruk och antisocial attityd, menar vissa att låg självkänsla och viktimisering också är kriminogena behov för kvinnor (Blanchette & Brown, 2006). Program där man höjer deltagarnas självkänsla och bearbetar deras viktimisering skulle i så fall minska återfall i brott för kvinnor, oavsett om de gör det för män eller inte.⁴ Idag finns dock inte tillräcklig forskning för att bedöma om låg självkänsla och viktimisering är kriminogena behov för kvinnor (Lart, Pantaziz, Pemberton, Turner, & Almeida, 2008). Det finns dock studier som tyder på att viktimisering har ingen eller mycket liten kausal betydelse för kvinnors benägenhet att begå brott (Forsman & Långstrom, 2012; Rettinger & Andrews, 2010). En annan kunskapslucka gäller kvinnors mottaglighet för program. Vissa menar att kvinnor oftare har en relationsbaserad inlärningsstil (Dowden & Andrews, 1999; Rettinger & Andrews, 2010). Kvinnor skulle i så fall vara mer mottagliga för program anpassade efter en relationsbaserad inlärningsstil, vilket skulle öka programeffekten. Inte heller om detta finns det tillräcklig forskning för man ska kunna dra några slutsatser (Lart et al., 2008). Att utvärdera VINN har därmed potential att fylla vissa kunskapsluckor om effekter av kriminalvårdsprogram för kvinnor.

Problemformulering och syfte

Det behövs fler väl genomförda effektutvärderingar av kriminalvårdsprogram för kvinnor (Tripodi et al., 2011). VINN ackrediterades i Kriminalvården år 2010 och ackrediterade program ska utvärderas med avseende på återfall i brott (Højdahl & Størksen, 2012a; Kriminalvården, 2007). Idag finns dock inga publicerade effektutvärderingar av programmet. Syftet med denna studie är därför att undersöka om VINN minskar återfall i brott bland kvinnliga kriminalvårdsklienter i Sverige.

⁴ Lågt självförtroende och viktimisering brukar traditionellt klassas som icke-kriminogena behov. I studier på mestadels män brukar behandlingsprogram riktade mot icke-kriminogena behov inte minska återfall i brott (Andrews & Bonta, 2010).

Metod

Undersökningsgrupp

Programgruppen bestod av kvinnliga kriminalvårdsklienter som skrevs in i frivården eller ut från anstalt år 2007 till 2009 och påbörjade VINN under påföljdstiden (n=194). Jämförelsegruppen bestod av kvinnliga klienter aktuella i Kriminalvården under samma period men som inte deltagit i programmet. För att minska risken för systematiska bakgrundsskillnader mellan grupperna användes propensity score-matchning (se sida 8).

I den här rapporten undersöks både de som påbörjat och fullföljt VINN. Att jämföra dem som *påbörjat* program, oavsett om de sedan fullföljde eller avbröt programmet, med en jämförelsegrupp kallas för *intention-to-treat*. Uteblivna eller oönskade effekter av programmet kan leda till att en deltagare vill avbryta densamma. Det går dock inte att på förhand säkert veta vilka som kommer att fullfölja programmet. Att då exkludera dem som avbryter programmet från undersökningen kan innebära att man sällar bort programdeltagare som drabbats av oönskade effekter. Det har därmed praktisk relevans att försöka fastställa vilken effekt programmet har på samtliga som påbörjat det. Nackdelen med att utgå ifrån alla som påbörjat programmet är å andra sidan att inte alla därmed fått *bela* programmet. Att man tar del av hela programinnehållet kan vara en förutsättning för att uppnå önskvärd effekt och man kan därför underskatta eventuella positiva effekter av ett program. Detta kan bli särskilt problematiskt om andelen som avbrutit programmet är högt. Därför jämförde vi även de som *fullföljt* VINN med en matchad jämförelsegrupp, vilket kallas för *per protokoll*. Nackdelen med detta är att de som klarar av att fullfölja program brukar ha bättre förutsättningar redan innan, t.ex. genom att vara mer motiverade och ha bättre impulskontroll. Med per-protokoll ökar därmed risken att programmet ser ut att ha positiva effekter när resultatet i själva verket beror på bakgrundsskillnader mellan deltagare som fullföljt programmet och jämförelsegrupp. Både intention-to-treat och per protokoll har alltså för- och nackdelar och det är därför som vi redovisar resultat från båda dessa analyser (Debidin & Lovbakke, 2005; Hollis & Campbell, 1999; J. A. Lewis & Machin, 1993).

På anstalt har en hög andel av aktuella kvinnor deltagit i VINN. Det ökar sannolikheten att kvinnliga anstaltsklienter som inte deltagit i VINN aktivt valts bort från programmet och därför inte är jämförbara med programdeltagare. Av den anledningen trodde vi inledningsvis att jämförelsegruppen skulle bli mer likvärdig om även klienter aktuella i Kriminalvården *före* införandet av VINN inkluderades. Vi provade därför att matcha programdeltagare med 1) samtida kvinnliga klienter och 2) både tidigare och samtida kvinnliga klienter. Därefter jämförde vi VINN-deltagarnas bakgrund med de matchade grupperna, för att se vilken av kontrollgrupperna som var mest lämplig som jämförelse. Det visade sig att de matchade grupperna var lika jämförbara med programgruppen. Att använda historiska jämförelsepersoner medför dessutom nackdelen att förändringar över tid i t.ex. klientsammansättning och verksamhetens innehåll och omfattning skulle kunna försämra jämförbarheten. Vi valde därför att bara inkludera samtida kvinnliga klienter i jämförelsegruppen.

Källor

Undersökningen bygger på en avidentifierad databas med länkning mellan Kriminalvårdens register och andra nationella befolkningsregister.⁵ Följande register användes: lagförings- och misstankeregistret, kriminalvårdsregistret (KVR), patientregistret, longitudinell integrationsdatabas för arbetsmarknadsstudier

⁵ Registerinformation hämtades från databasen "Crime-2" vid Institutionen för medicinsk epidemiologi och biostatistik vid Karolinska institutet. Detta gjordes inom ramen för en forskningsstudie som godkänts av Regional forskningsetisk kommitté vid Karolinska institutet (2009/939-31/5). Databasen innehåller länkad, landsomfattande, longitudinell information från flera befolkningsbaserade register, men anonymiserades innan studien genom att personuppgifter (namn, adress, personnummer) togs bort.

(LISA), folk- och bostadsräkningarna (FoB), flergenerationsregistret, registret över totalbefolkningen, migrationsregistret och dödsorsaksregistret.

Utfall

Utfallet var återfall i brott och följdes i första hand upp i misstankeregistret. I misstankeregistret ingår enbart personer som av förundersökningsledare bedömts som minst skäligen misstänkta 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).⁶ 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,⁷ d.v.s. ökad sannolikhet att statistiskt kunna säkerställa en faktisk effekt, och för att det gav ett års 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 slutdatum för 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 901 dagar (minst 0 dagar, som mest 1460 dagar). I misstankeregistret kunde 98,5% följas upp i minst ett år och 64,8% i minst två år. I lagföringsregistret kunde brottsåterfall däremot bara följas upp fram till 2009-12-31, vilket gav en genomsnittlig uppföljningstid på 546 dagar (minst 0, mest 1095).⁸ Därmed kunde enbart 65,6% följas upp i minst ett år och 32,7% i minst två år i lagföringsregistret.

Bakgrundsfaktorer

I studier där programgruppen inte är tillräckligt lik jämförelsegruppen före programmet 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. gruppskillnader före deltagande i programmet som påverkar risken för återfall i brott. Potentiella störfaktorer i den här undersökningen presenteras i tabell 2. De faktorer som i inledande analyser samvarierade såväl med programdeltagande som med återfall i brott hanterades statistiskt i undersökningen.

⁶ 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.

⁷ 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.

⁸ 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).

Tabell 2. Bakgrundsfaktorer i undersökningen.

Register	År	Bakgrundsfaktorer
Registret över totalbefolkningen	1968-2009	Ålder, bostadsområde, född i Sverige
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, arbetsinkomst
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	Anstalts- eller frivårdsklient, påbörjat allmänt kriminalitets-/våldsprogram, missbruksprogram, annat motivationsprogram än VINN
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.

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 är så lik(a) programdeltagarens som möjligt. Programgruppen jämförs därmed med personer som inte gått programmet, men som givet sin bakgrund (enligt kontrollvariablerna) hade 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 bakgrunder 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. Genom att hantera bortfall på det här sättet minskar 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 slutligen Coxregression (Friendship, Street, Cann, & Harper, 2005). I Coxregressionen jämfördes varje programdeltagare med den person i jämförelsegruppen som denne hade matchats med, s.k. stratifiering på

matchade par (Austin, 2008).⁹ Coxregression resulterar i effektstorleksmättet hazardkvot som indikerar programgruppens risk att 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 och tillhörande konfidensintervall redovisas även grafiskt genom s.k. forest plot (S. Lewis & Clarke, 2001). Dessutom anger vi resultaten i form av andelen återfall i brott inom ett år i programgruppen respektive den matchade jämförelsegruppen.¹⁰

⁹ 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”-korrigerings i Coxregressionen blir standardfelen robustare för specifikationsfel, och därmed blir inte precisionen i våra beräkningar falskt för hög eller med andra ord blir inte konfidensintervallen felaktigt för smala (Lin & Wei, 1989).

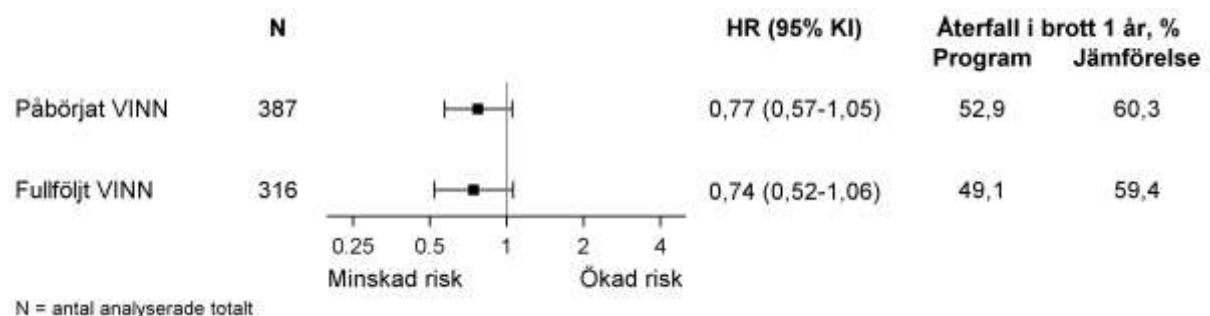
¹⁰ Då många programdeltagare saknade ett helt års uppföljningstid i lagföringsregistret erhöles andelen återfall i brott från Kaplan-Meier-överlevnadskurvor (Bland & Altman, 1998). Även om vi använde Kaplan-Meier innebär den korta uppföljningstiden dock fortfarande att resultatet blir mer osäkert för lagföringsregistret efter ett år (Rich et al., 2010).

Resultat

Huvudresultat

Klienter som påbörjat VINN (n=194) jämfördes med matchade kvinnliga klienter med likvärdig bakgrund (n=194).¹¹ De som påbörjat VINN hade minskad risk att återfalla i brott enligt misstankeregistret, men skillnaden kunde inte säkerställas statistiskt (se figur 1).

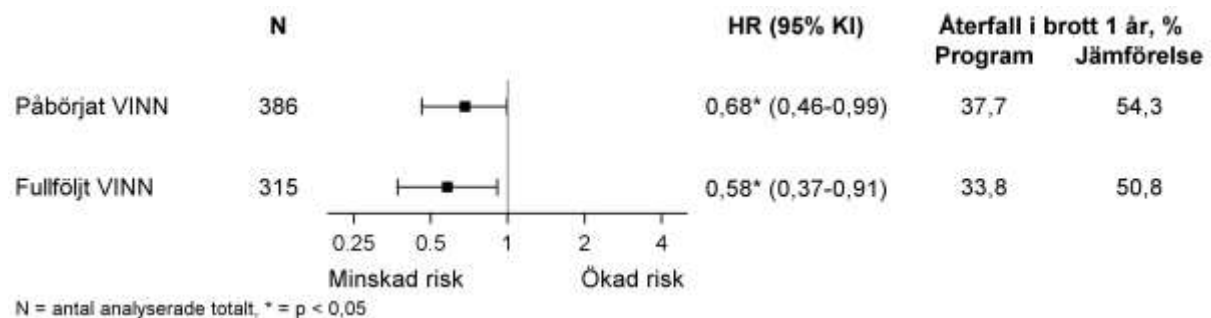
Av dem som påbörjade VINN år 2007-2009 fullföljde 82% programmet.¹² De som fullföljt VINN (n=158) jämfördes med en ny matchad jämförelsegrupp med liknande bakgrund (n=158). Även de som fullföljt VINN hade minskad risk att återfalla i brott jämfört med matchade klienter, men återigen gick inte skillnaden att säkerställa statistiskt.¹³



Figur 1. Hazardkvoter [HR] med 95% konfidensintervall [KI] för återfall i brott enligt misstankeregistret. Deltagare i VINN jämförda med matchade kvinnliga klienter som inte deltagit i programmet.

Känslighetsanalys

Resultatens tillförlitlighet stärks om de kan bekräftas med en annan källa och därför följdes deltagarna i en känslighetsanalys även upp i lagföringsregistret. När återfall i brott följdes upp i lagföringsregistret istället för i misstankeregistret var resultaten i stort sett desamma (se figur 2). Men den minskade risken att återfalla i brott bland de som påbörjat respektive fullföljt VINN kunde för lagföringsregistret säkerställas statistiskt.



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

¹¹ Av de totalt 388 klienterna kunde 1 inte följas upp och därav är antalet analyserade 387 i figuren.

¹² 158 fullföljde programmet och 34 avbröt det. För 2 klienter saknades information om eventuellt fullföljande.

¹³ Kriminalvårdens programutvärderingar har ofta innehållit separat analys av dem som avbrutit behandlingen. I den här utvärderingen är dock antalet som avbrutit programmet bara 34, vilket skulle göra resultatet mycket osäkert. Dessutom visade det sig vara svårt att hitta klienter med likvärdig bakgrund att matcha med dem som avbrutit VINN. Vi valde därför att inte göra en separat analys på dem som avbrutit programmet i den här utvärderingen.

Diskussion

Inga statistiskt säkerställda skillnader hittades i huvudresultaten. Samtliga resultat pekar dock i riktning mot att svenska kriminalvårdsklienter som deltog i VINN 2007-2009 hade minskad risk att återfalla i brott. Att få klienter ingick i undersökningen medför låg statistisk power, d.v.s. låg sannolikhet att en faktisk skillnad kan säkerställas statistiskt. För att med säkerhet kunna dra slutsatsen att VINN minskar återfall i brott behövs därför en ny utvärdering med fler programdeltagare. Sammanfattningsvis är resultaten alltså osäkra men lovande.

VINN är riktat mot flera kända kriminogena behov och använder principer från social kognitiv teori (se introduktion, sida 3). Detta har stöd i tidigare forskning och talar därmed för att riskminskningen bland VINN-deltagare stämmer (Dowden & Andrews, 1999). Däremot har andra aspekter av VINN inte utvärderats tidigare, t.ex. programmets huvudsakliga förändringsmodell KASAM. Det är därför lämpligt att vara försiktig med att dra slutsatser om VINN från effektutvärderingar av andra program för kvinnliga klienter.

Även om vi hanterade många viktiga bakgrundsfaktorer genom propensity score-matchning kan det fortfarande kvarstå skillnader på andra faktorer som snedvrider resultatet. Vi hade t.ex. inte möjlighet att hantera skillnader i motivation. Om deltagare i VINN redan från början var mer motiverade än icke-deltagare kan det alltså ha snedvridit resultatet. Det enda sättet att vara säker på att resultat inte beror på systematiska skillnader mellan grupperna före programmet är att slumpmässigt dela in deltagare i program- eller jämförelsegrupp (s.k. randomiserad kontrollerad studie).

Då det brukar gå bättre för dem som fullföljer program är det viktigt att hålla ner andelen avbrott. I den här utvärderingen fullföljde 82% VINN, vilket kan jämföras med att i genomsnitt 73-76% av manliga och kvinnliga brottsdömda fullföljde behandling i två internationella översikter (McMurrin & Theodosi, 2007; Olver, Stockdale, & Wormith, 2011). Andelen som fullföljde VINN under undersökningsperioden kan alltså anses god.

En framgångsfaktor för återfallsförebyggande program för både kvinnor och män är att utgå ifrån riskprincipen, enligt vilken mer omfattande insatser bör ges till klienter med högre risk att återfalla i brott (Andrews & Dowden, 2006; Dowden & Andrews, 1999). VINN är inte uttryckligen utformat för klienter med medelhög eller hög risk för återfall i brott. Å ena sidan är det därmed möjligt att programmets effekt på återfall i brott kan förstärkas genom tydligare tillämpning av riskprincipen. Å andra sidan har programmet även ökat livskvalitet som mål och klienter med låg risk för återfall i brott kan ändå ha stora problem på det området. Trots att det inte är ett uttalat urvalskriterium verkar dock den genomsnittliga VINN-deltagaren under undersökningsperioden ha haft åtminstone medelhög risk för återfall i brott (se bilaga B, tabell B1).

De tidigaste programdeltagarna i vår utvärdering var de första som deltog i VINN i Kriminalvården. Det kan ha försvagat programmets effekt eftersom programledarna då inte hade lika stor erfarenhet av VINN som de har idag. Dessutom var VINN ännu inte ackrediterat i Kriminalvården under undersökningsperioden och manualerna har sedan dess förändrats något (Højdahl & Størksen, 2012a). Det är med andra ord möjligt att VINN:s inverkan på återfall i brott har förändrats efter undersökningsperiodens slut.

I utvärderingar med författare som haft inflytande över programmet hittar man ofta starkare effekt på återfall i brott (Petrosino & Soydan, 2005). I sådana studier brukar programtroheten vara bättre, d.v.s. programmet utförs som avsett, vilket i sin tur påverkar hur stark effekt man hittar på återfall i brott (Andrews & Dowden, 2005). Eftersom det här är en retrospektiv utvärdering kunde vi inte kontrollera hur VINN utfördes under undersökningsperioden. Det är därmed möjligt att vi hade hittat en tydligare programeffekt om det här vore en prospektiv studie med kontroll över bl.a. programtrohet.

Man brukar skilja mellan två typer av effektutvärderingar. I den första typen undersöks om det alls är möjligt för ett program att få önskvärd effekt, genom att se till så omständigheterna i övrigt är så idealiska

som möjligt (eng: efficacy trial eller explanatory trial). I den andra typen undersöker man istället om programmet har effekt så som det utförs i praktiken (eng: effectiveness trial eller pragmatic trial). Fördelen med den första typen är att det blir lättare att urskilja en eventuell programeffekt. Nackdelen är att sådana utvärderingar har lägre praktisk relevans, eftersom program sällan utförs så i praktiken (Trewick & Zwarenstein, 2009). Den här rapporten är ett exempel på en *effectiveness trial*.

Utöver att minska återfall i brott har VINN som mål öka deltagarnas livskvalitet. I en retrospektiv registerstudie som denna är det inte möjligt att undersöka livskvalitet. Våra resultat säger därför inget om vilken effekt VINN kan ha på livskvalitet. Eftersom det inte tycks finnas andra effektutvärderingar av VINN är det alltså för närvarande okänt om programmet ökar livskvalitet. För att undersöka det i framtiden behöver man prospektivt samla in information om livskvalitet hos både programdeltagare i VINN och en lämplig jämförelsegrupp.

Som nämdes i introduktionen finns kunskapsluckor inom forskningen om kriminalvårdsprogram för kvinnor. Låg självkänsla och brottsutsatthet har framförts som möjliga kriminogena behov för kvinnor (Blanchette & Brown, 2006). Kvinnor anses också ha en relationsbaserad inlärningsstil i högre utsträckning (Dowden & Andrews, 1999; Rettinger & Andrews, 2010). VINN kan anses väl anpassat efter vad som kan vara specifika behov för kvinnliga klienter. Samtidigt finns idag inte tillräckligt med studier för att veta om dessa hypoteser om könsskillnader i behandlingsbehov stämmer. Eftersom det fanns en tendens mot minskat återfall i brott i den här utvärderingen kan vi inte utesluta att hypoteserna stämmer.¹⁴ Samtidigt är VINN riktat mot många olika problemområden och det är därför svårt att urskilja exakt vad som driver en eventuell programeffekt. Ett sätt att ta reda på det vore att i en forskningsöversikt jämföra kriminalvårdsprogram för kvinnor med t.ex. relationsfokus med program som inte har det. Även om den här utvärderingen alltså inte ensam kan fylla dessa kunskapsluckor kommer man i framtiden förhoppningsvis kunna göra det genom att sammanföra och jämföra våra resultat med andra utvärderingar.

¹⁴ Det finns dock studier som talar emot att utsatthet för våld skulle vara betydelsefull kausal riskfaktor för kriminalitet bland kvinnor (Forsman & Långstrom, 2012).

Referenser

- Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct*. New Providence, NJ: LexisNexis/Matthew Bender.
- Andrews, D. A., & Dowden, C. (1999). A meta-analytic investigation into effective correctional intervention for female offenders *FORUM on Corrections Research*, 11(3), 18-21.
- 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., & Dowden, C. (2006). Risk principle of case classification in correctional treatment: A meta-analytic investigation. *International Journal of Offender Therapy and Comparative Criminology*, 50(1), 88-100. doi: 10.1177/0306624X05282556
- Antonovsky, A. (2005). *Hälsans mysterium*. Stockholm: Natur & Kultur.
- 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
- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: W. H. Freeman.
- Blanchette, K., & Brown, S. L. (2006). *The assessment and treatment of women offenders: An integrative perspective*. Chichester, West Sussex, England J. Wiley & Sons.
- 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
- Carlson, B. E., & Shafer, M. S. (2010). Traumatic histories and stressful life events of incarcerated parents: Childhood and adult trauma histories. *The Prison Journal*, 90(4), 475-493. doi: 10.1177/0032885510382224
- 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. (1999). What works for female offenders: A meta-analytic review. *Crime & Delinquency*, 45(4), 438-452. doi: 10.1177/001128799045004002
- Fazel, S., Bains, P., & Doll, H. (2006). Substance abuse and dependence in prisoners: A systematic review. *Addiction*, 101(2), 181-191. doi: 10.1111/j.1360-0443.2006.01316.x
- Forsman, M., & Långström, N. (2012). Child maltreatment and adult violent offending: Population-based twin study addressing the 'cycle of violence' hypothesis. *Psychological Medicine*, 42(9), 1977-1983. doi: 10.1017/s0033291711003060
- 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
- Goff, A., Rose, E., Rose, S., & Purves, D. (2007). Does PTSD occur in sentenced prison populations? A systematic literature review. *Criminal Behaviour and Mental Health*, 17(3), 152-162. doi: 10.1002/cbm.653

- Granheim, P. K. (2007). Evalueringsrapport VINN samtalegruppe for kvinner - basert på datamateriale fra norske og svenske enheter 2006-2007. Oslo: Kriminalomsorgens utdanningscenter KRUS.
- 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
- Højdahl, T., Magnus, J. H., Hagen, R., & Langeland, E. (2013). "VINN" - An accredited motivational program promoting convicted women's sense of coherence and coping. *EuroVista*, *2*(3), 177-190.
- Højdahl, T., Magnus, J. H., & Langeland, E. (2014). "A bridge to change": Experiences of participation in "VINN"— a motivational program for convicted women. A qualitative study. *EuroVista*, *3*(2), 81-96.
- Højdahl, T., & Størksen, M. (2012a). VINN - Motivationsprogram för kvinnor. Programmanual. Norrköping: Kriminalvården.
- Højdahl, T., & Størksen, M. (2012b). VINN - Motivationsprogram för kvinnor. Teori-, genomförande- och utvärderingsmanual för programledare. Norrköping: Kriminalvården.
- Kane, S. P. (2014). Odds ratio to risk ratio. Hämtat 3 juni, 2014, från <http://clincalc.com/Stats/ConvertOR.aspx>
- 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: VINN. Norrköping: Kriminalvården.
- Lart, R., Pantaziz, C., Pemberton, S., Turner, W., & Almeida, C. (2008). Interventions aimed at reducing re-offending in female offenders: a rapid evidence assessment (REA). London: Ministry of Justice.
- 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

- Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change*. New York, NY: Guilford Press.
- 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
- Petrosino, A., & Soydan, H. (2005). The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1, 435-450.
- Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behavior change. *American journal of health promotion*, 12(1), 38-48. doi: 10.4278/0890-1171-12.1.38
- Rettinger, L. J., & Andrews, D. A. (2010). General risk and need, gender specificity, and the recidivism of female offenders. *Criminal Justice and Behavior*, 37(1), 29-46. doi: 10.1177/0093854809349438
- 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
- 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
- Tripodi, S. J., Bledsoe, S. E., Kim, J. S., & Bender, K. (2011). Effects of correctional-based programs for female inmates: A systematic review. *Research on Social Work Practice*, 21(1), 15-31. doi: 10.1177/1049731509352337
- 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
- Wolff, N., Shi, J., & Siegel, J. A. (2009). Patterns of victimization among male and female inmates: Evidence of an enduring legacy. *Violence & Victims*, 24(4), 469-484. doi: 10.1891/0886-6708.24.4.469

Bilagor

Bilaga A. Metodfördjupning

Matchning

Matchningen utfördes i statistikprogrammet SAS via ett publicerat s.k. makro (Coca-Perraillon, 2007). De bakgrundsfaktorer som skulle användas i matchningen valdes utifrån om variabeln uppvisade ett signifikant samband med både att delta i VINN och med återfall i brott. 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). Gränsen för signifikans sattes i det här sammanhanget vid $p < 0,10$ istället för det konventionella $p < 0,05$. Anledningen var att vi inte ville missa att hantera betydelsefulla skillnader p.g.a. låg statistisk power.

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 (även kallat standardiserad bias) (Austin, 2009).¹⁵ Om matchningen är fallit ut väl 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 eftersträlvade att Cohens d skulle vara lägre än 0,10, men åtminstone under 0,20.

Matchningarna resulterade inte oväntat i kraftigt minskade bakgrundsskillnader mellan grupperna. Mätt i Cohens d minskade skillnaden i genomsnitt med 79%.¹⁶ Men trots kraftigt minskade bakgrundsskillnader, och likvärdig fördelning av propensity scores, kunde Cohens d inte minskas mer än till under 0,20 för flera variabler. Detta innebär att matchningarna var acceptabla, men inte perfekta. Att Cohens d i flera fall inte kunde minskas till under 0,10 kan bero på att det lättare uppstår små slumpmässiga skillnader på enstaka variabler när antalet programdeltagare är litet, oavsett om man använder propensity score-matchning eller randomisering (Austin, 2009). Den matchningsmetod som vi använde kallas för nearest neighbor matchning (även känt som greedy matchning). Det innebär att varje programdeltagare matchas med den i jämförelsegruppen med mest lik propensity score. Klienterna i jämförelsegruppen kunde enbart matchas till en programdeltagare var och ”återanvändes” alltså inte (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. Vi försökte därför matcha fem personer till varje programdeltagare. Det visade sig dock bara möjligt att matcha en klient till varje programdeltagare. När vi försökte matcha mer än en klient till varje programdeltagare blev gruppernas bakgrund inte tillräckligt likvärdig (Cohens $d > 0,20$).

För klienter som fullföljt VINN blev grupperna efter matchning inte tillräckligt lika på variabeln *påbörjat missbruksprogram* (Cohens $d=0,24$). Vi försökte förbättra matchningen genom att tillämpa flera i litteraturen föreslagna metoder:

- Ange ett maxavstånd för hur stor skillnaden i propensity scores som mest fick vara. Enligt rekommendation användes 0,2 av standardavvikelsen som maxavstånd (Austin, 2011b).
- Tillåta återanvändning av deltagare (eng: matching with replacement) (Austin, 2011a)
- Matcha deltagarna på både propensity score och variabeln påbörjat missbruksprogram via Mahalanobis metric matching (Coca-Perraillon, 2007).
- Exkludera programdeltagare med högre propensity scores än det högsta värdet i jämförelsegruppen.

¹⁵ Cohens d benämns ibland även som ”standardized bias” eller ”standardized difference” och beräknades utifrån publicerade formler (Austin, 2011a).

¹⁶ Gruppskillnaden gick från att i genomsnitt ha storleksordningen Cohens $d \pm 0,34$ till $\pm 0,07$ (se bilaga B).

Men samtliga försök resulterade fortfarande i att grupperna blev obalanserade på samma eller någon annan variabel efter matchning (Cohens $d > 0,20$). Till sist hanterade vi problemet genom att kontrollera för variabeln *påbörjat missbruksprogram* i efterföljande Coxregression (Austin, 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 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 förutsäga vad det sanna värdet är på bortfallsvariabeln*, 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 kan ta hänsyn till vid multipel imputation, kommer bortfallet inte längre leda till förvrängda resultat. I flera s.k. simuleringstudier 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 använde en form av multipel imputation som kallas för Fully Conditional Specification (FCS). Imputeringen utfördes via proceduren Proc MI i statistikprogrammet SAS version 9.4. Det är inte nödvändigt att inkludera mer än de 15-25 viktigaste variablerna i multipel imputation (van Buuren, Boshuizen, & Knook, 1999). Vi inkluderade därför alla potentiella kontrollvariabler och utfallsvariabler. För metoden FCS rekommenderas att imputera minst 20 dataset vid 10-15% bortfall (Van Buuren, Brand, Groothuis-Oudshoorn, & Rubin, 2006). Bortfallet i vår studie varierade mellan variabler från 0% till 16% (medel 1%, median 0%) och vi imputerade 20 dataset. För att kunna utföra propensity score-matchning efter multipel imputation beräknades propensity scores för alla 20 imputerade dataset. Därefter räknades genomsnittlig propensity score ut över alla dataseten 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

- 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. (2007). Propensity-score matching in the cardiovascular surgery literature from 2004 to 2006: A systematic review and suggestions for improvement. *The Journal of Thoracic and Cardiovascular Surgery*, *134*(5), 1128-1135. doi: 10.1016/j.jtcvs.2007.07.021
- 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
- Austin, P. C. (2011a). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399-424. doi: 10.1080/00273171.2011.568786
- Austin, P. C. (2011b). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics*, *10*(2), 150-161. doi: 10.1002/pst.433
- Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic epidemiology*. Geneva: World Health Organization.

- 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>
- 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
- 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
- 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
- Van Buuren, S., Brand, J. P. L., Groothuis-Oudshoorn, C. G. M., & Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. *Journal of Statistical Computation and Simulation*, 76(12), 1049-1064. doi: 10.1080/10629360600810434
- 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* VINN och jämförelsegruppen före och efter propensity score-matchning (1:1 nearest neighbor).

	Påbörjat Vinn (n=194)	Omatchad jämförelsegrupp (n=6416)	Matchad jämförelsegrupp (n=194)	Cohens d	
				Före	Efter
Ålder vid uppföljningsstart, m	40,2 år	38,0 år	41,1	0,19	-0,09
Anstaltsklient, %	82,5	16,5	88,1	1,76	-0,16
Civilstatus, %				0,00	-0,02
Gift/sambo med gemensamma barn	21,1	21,1	22,2		
Ogift/inte sambo med gemensamma barn	78,9	78,6	77,8		
Information saknas	0,0	0,3	0,0		
Bostadsort, %				0,23	0,17
Stad/förort	70,6	60,3	63,4		
Övrig ort	28,4	39,0	36,6		
Information saknas	1,0	0,7	0,0		
Högsta utbildningsnivå, %					
Folk- eller grundskola < 9 år	8,8	6,2	8,8	0,10	0,00
Folk- eller grundskola 9 år	43,3	36,6	44,9	0,13	-0,03
Gymnasial utbildning 1-3 år	42,3	41,9	40,7	0,02	0,03
Universitet och högskola, inkl. gymnasium 4 år	3,6	11,9	3,6	-0,32	-0,02
Information saknas	2,1	3,5	2,1		
Högsta SES i barndom, %					
Låg	18,6	19,5	23,7	-0,07	-0,15
Medel	37,63	33,8	37,1	0,03	-0,01
Hög	34,02	30,2	26,8	0,03	0,15
Information saknas	9,79	16,6	12,4		
Född i Sverige, %	84,5	79,3	80,9	0,14	0,10
Arbetsinkomst året innan, %	20,6	47,6	17,5	-0,60	0,08
Information saknas, %	0,0	0,3	0,0		
<i>Tidigare brottslighet</i>					
Antal tidigare domar, m	15,8	6,8	14,7	0,65	0,07
Ålder vid första dom, m	23,3 år	28,3 år	23,5	-0,46	-0,02
Ngn gång lagförd för våldsbrott, %	49,0	40,2	49,0	0,18	0,01
Ngn gång lagförd för narkotikabrott, %	77,8	36,4	80,9	0,93	-0,07
Ngn gång lagförd för rattfylleri, %	52,6	45,2	54,6	0,15	-0,04
Information om tidigare brottslighet saknas, %	0,5	0,4	0,0		
<i>Familjär brottslighet/missbruk</i>					
Mamma ngn gång dömd, %	28,9	25,7	30,9	0,06	-0,06
Pappa ngn gång dömd, %	49,0	46,2	51,0	0,05	-0,05
Ngt syskon som ngn gång dömts, %	38,1	34,6	32,0	0,07	0,13
Ngn förälder med missbruk/beroende, %	25,8	22,5	27,8	0,08	-0,05

Fortsättning tabell B1.

	Påbörjat Vinn (n=194)	Omatchad jämförelsegrupp (n=6416)	Matchad jämförelsegrupp (n=194)	Cohens d	
				Före	Efter
<i>Psykiatri</i>					
Alkoholmissbruk/-beroende, %	30,9	28,3	33,0	0,06	-0,04
Narkotikamissbruk/-beroende, %	55,7	31,1	56,7	0,51	-0,02
Personlighetsstörning, %	10,8	11,0	16,5	0,00	-0,17
ADHD, %	2,6	3,3	3,6	-0,04	-0,06
<i>Behandlingsprogram</i>					
Påbörjat missbruksprogram, %	36,1	2,8	34,5	0,93	0,03
Påbörjat <i>annat</i> motivationsprogram än Vinn, %	53,6	4,2	47,4	1,30	0,12
Påbörjat allmänt kriminalitets- /våldsprogram, %	24,7	1,9	19,6	0,71	0,12
Propensity score, m	0,24	0,02	0,23		

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

	Fullföljt Vinn (n=158)	Omatchad jämförelsegrupp (n=6416)	Matchad jämförelsegrupp (n=158)	Cohens d	
				Före	Efter
Ålder vid uppföljningsstart, m	40,2 år	38,0 år	41,2 år	0,20	-0,08
Anstaltsklient, %	85,4	16,5	91,1	1,91	-0,18
Civilstatus, %				-0,01	0,10
Gift/sambo med gemensamma barn	20,9	21,1	17,1		
Ogift/inte sambo med gemensamma barn	79,1	78,6	82,9		
Information saknas	0,0	0,3	0,0		
Bostadsort, %				0,18	0,06
Stad/förort	68,4	60,3	65,8		
Övrig ort	30,4	39,0	33,5		
Information saknas	1,3	0,7	0,6		
Högsta utbildningsnivå, %					
Folk- eller grundskola < 9 år	7,6	6,2	8,9	0,05	-0,05
Folk- eller grundskola 9 år	44,3	36,6	46,8	0,15	-0,06
Gymnasial utbildning 1-3 år	42,4	41,9	38,6	0,00	0,07
Universitet och högskola, inkl. gymnasium 4 år	3,8	11,9	2,5	-0,31	0,07
Information saknas	1,9	3,5	3,2		
Högsta SES i barndom, %					
Låg	17,7	19,5	21,5	-0,10	-0,12
Medel	38,0	33,8	37,3	0,02	-0,02
Hög	35,4	30,2	29,1	0,06	0,12
Information saknas	8,9	16,6	12,0		
Född i Sverige, %	85,4	79,3	82,3	0,16	0,09
Arbetsinkomst året innan, %	22,2	47,6	22,8	-0,56	-0,02
Information saknas, %	0,0	0,3	0,0		
<i>Tidigare brottslighet</i>					
Antal tidigare domar, m	15,4	6,8	16,6	0,60	-0,07
Ålder vid första dom, m	23,6 år	28,3 år	24,1 år	-0,43	-0,05
Ngn gång lagförd för våldsbrott, %	44,3	40,2	43,7	0,08	0,01
Ngn gång lagförd för narkotikabrott, %	76,6	36,4	75,3	0,90	0,03
Ngn gång lagförd för rattfylleri, %	50,6	45,2	55,1	0,15	-0,09
Information om tidigare brottslighet saknas, %	0,4	0,4	0,6		
<i>Familjär brottslighet/missbruk</i>					
Mamma ngn gång dömd, %	27,9	25,7	34,2	0,05	-0,14
Pappa ngn gång dömd, %	49,4	46,2	51,3	0,06	-0,04
Ngt syskon som ngn gång dömts, %	38,6	34,6	32,9	0,08	0,12
Ngn förälder med missbruk/beroende, %	26,0	22,5	28,5	0,08	-0,06
<i>Psykiatri</i>					
Alkoholmissbruk/-beroende, %	29,1	28,3	30,4	0,02	-0,03
Narkotikamissbruk/-beroende, %	53,2	31,1	51,9	0,46	0,03
Personlighetsstörning, %	8,9	11,0	13,9	-0,07	-0,16
ADHD, %	1,3	3,3	2,5	-0,14	-0,09

Fortsättning tabell B2.

	Fullföljt Vinn (n=158)	Omatchad jämförelsegrupp (n=6416)	Matchad jämförelsegrupp (n=158)	Cohens d	
				Före	Efter
<i>Behandlingsprogram</i>					
Påbörjat missbruksprogram, %	38,6	2,8	27,2	0,99	0,24*
Påbörjat <i>annat</i> motivationsprogram än Vinn, %	53,8	4,2	55,1	1,30	-0,03
Påbörjat allmänt kriminalitets- /våldsprogram, %	25,3	1,9	24,7	0,73	0,01
Propensity score, m	0,22	0,02	0,21		

* Kontrollerat för i avslutande Coxregression p.g.a. obalans i matchning som inte kunde hanteras på annat sätt.