
Modellversuch: Risikoorientierter Sanktionsvollzug

Bemerkungen und Analysen zum Projekt des Amtes für Justizvollzug des Kantons Zürich

Martin Killias & Raquel Rosés Brüngger
Killias Research & Consulting AG, Lenzburg

Zielsetzung der vorliegenden Bemerkungen

Anfangs Juli wurde der Öffentlichkeit ein neues Instrument zur Risiko-Abklärung unter Personen im Justizvollzug vorgestellt. Es basiert auf einem 95-seitigen Dokument mit dem Datum vom 23. Mai 2014¹. Federführend war das Amt für Justizvollzug des Kantons Zürich. Weiter beteiligt waren die Strafvollzugsämter der Kantone Luzern, Thurgau und St. Gallen. Der Bund hat das Projekt als Modellversuch finanziert. Ziel ist bzw. war, im Anschluss an verschiedene „kritische“ Vorfälle im Straf- und Massnahmenvollzug über ein verbessertes Instrument zur Risikobeurteilung zu verfügen. Beabsichtigt ist wohl die Übernahme dieses Instruments durch alle Kantone – dies im Zeichen der Vereinheitlichung des Vollzugswesens oder zumindest der forensischen Beurteilung der Gefährlichkeit von Personen im Freiheitsentzug.

Die vorliegenden Bemerkungen sollen – im Sinne eines *Peer reviewing* – die Stärken und Schwächen des vorgeschlagenen Instruments ROS („risiko-orientierter Strafvollzug“) einer kritischen Prüfung unterziehen. Es geht darum, einerseits offene oder noch nicht völlig geklärte Fragen aufzuzeigen und andererseits die Treffsicherheit des vorgestellten Prognose-Instruments aufzuzeigen. Dies dient letztlich der sachlichen Diskussion. Wie wir zeigen werden, sind einige wichtige Fragen noch nicht geklärt oder anhand der zugänglichen Dokumente noch nicht zu beurteilen. Weiter ist die Treffsicherheit des verwendeten (und anderen Kantonen nun vorgeschlagenen) Instruments anhand der vorgelegten Daten zu diskutieren. Dies auch im Vergleich zu anderen Instrumenten, zu denen wir in der Literatur geeignete Daten zur Treffsicherheit haben finden können. Zum Schlussbericht liegen zwei Beilagen vor, wovon wir vor allem die Ergebnisevaluation von Christian Schwarzenegger, Patrick Manzoni und Mirjam Baur beiziehen werden².

Im Folgenden halten wir uns weitgehend an den Aufbau des Schlussberichts vom 23. Mai 2014. Wir beginnen mit einer kurzen Zusammenfassung des wesentlichen Inhalts und diskutieren anschliessend die Ergebnisse der Evaluation. Dabei sollen die Stärken und Schwächen wie auch die offenen

¹ Der Schlussbericht vom 23.05.2014 (nachstehend zitiert SB) ist erhältlich unter http://www.ejpd.admin.ch/content/dam/data/sicherheit/straf_und_massnahmen/ber-mv/ros-schlussber-d.pdf

² Dieser wurde uns freundlicherweise von Herrn Prof. Schwarzenegger zur Verfügung gestellt. Er ist abrufbar unter https://www.bj.admin.ch/content/bj/de/home/themen/sicherheit/straf_und_massnahmenvollzug/modellversuche/evaluationsberichte.html#ROS%20-%20Risikoorientierter%20Sanktionenvollzug

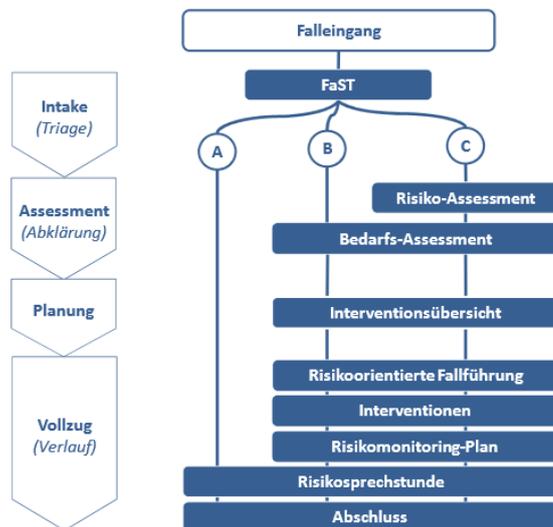
Fragen angesprochen werden. *Unsere kritischen Fragen und Bemerkungen sind jeweils kursiv an den entsprechenden Stellen eingefügt.*

Zusammenfassung des Versuchskonzepts

Der **Modellversuch** wurde zwischen 2010 und 2013 im risikoorientierten Sanktionsvollzug in den Kantonen Luzern, St. Gallen, Thurgau und Zürich durchgeführt (SB³, S. 6) und zielt auf die Reduktion der Rückfälligkeit von Straftätern nach dem Vollzug einer Sanktion (SB, S. 32). Die **zentrale Fragestellung** betrifft die Durchführbarkeit einer flächendeckenden und organisationübergreifenden Risikobeurteilung im Sanktionsvollzug. Die Evaluation umfasst eine Ergebnis- und eine Prozessevaluation (SB, S. 9).

Die **ROS-Konzeption** ist durch ein mehrstufiges Prozessmodell abgebildet (Abbildung 7⁴), wobei die

Abbildung 1: Ursprüngliche Fassung des ROS-Arbeitsprozesses bis Mai 2011 (übernommen aus SB, Abb. 6, S. 41)



jeweiligen Arbeitsschritte instrumentengestützt und standardisiert erfolgen (SB, S. 23). Zur deutlicherung der nachfolgenden Ausführungen übernehmen wir hier dieses Schema.

Der erste Schritt erfolgt durch die Anwendung des **Fall-Screening-Tools** (FaST) zur Gruppeneinteilung der Fälle (SB, S. 24). FaST umfasst verschiedene Merkmalsbereiche aus dem Vollzugsauftrag (d.h. wohl dem Anlassdelikt), dem Strafregisterauszug und dem Rechtsinformationssystem (RIS) des Kantons Zürich, um die Einteilung der in den Strafvollzug Eingewiesenen in drei Gruppen durchzuführen. Berücksichtigt werden somit der aktuelle deliktbezogene Entscheid, die Vorstrafen, zusätzliche problematische Aspekte (wie zum Beispiel, ob mindestens ein Gewaltdelikt oder Sexualdelikt vor dem vollendeten 18. Lebensjahr begangen wurde), und

risikomindernde Faktoren (wie aktuelles Alter über 50). *Allerdings sind für uns die Kriterien für die Einteilung in die drei Gruppen (und insbesondere das Gewicht, das dabei einer Vorverurteilung wegen eines Gewalt- und/oder Sexualdelikts zukommt) nicht vollständig transparent. Wir werden unten darauf zurückkommen.*

Die Einteilung erfolgt je nach den genannten Merkmalsbereichen in folgende drei Gruppen:

Gruppe A: Keine oder wenige Hinweise für einen erhöhten Abklärungsbedarf

Gruppe B: Hinweise für einen erhöhten Abklärungsbedarf in Bezug auf allgemeine Delikte

Gruppe C: Hinweise für einen erhöhten Abklärungsbedarf in Bezug auf Sexual und/oder Gewaltdelikte.

Der Zweite Schritt, die **ROS-Abklärung** (R-A) dient der Feststellung des Rückfallrisikos und des Interventionsbedarfs, sowie der Identifizierung von Problembereichen durch einen Erfassungsbogen

³ SB: Amt für Justizvollzug Kanton Zürich (2014). Schlussbericht. Modellversuch: Risikoorientierter Sanktionsvollzug. 23.05.2014 http://www.ejpd.admin.ch/content/dam/data/sicherheit/straf_und_massnahmen/ber-mv/ros-schlussber-d.pdf

⁴ SB, S. 43

(SB, 24). Die Abteilung für Forensisch-Psychologische Abklärung (AFA) fasst die fallspezifischen Informationen zusammen, um einen Vollzug zu ermöglichen, der an wissenschaftlichen Wirksamkeitsprinzipien orientiert ist. Diese Phase beinhaltet die Informationssammlung, ein standardisiertes Risk-Assesment (fallspezifische Auswahl des Instrumentes, SB 27), ein Fallkonzept (in Zusammenhang mit dem Delikt), eine Risikoqualifizierung und eine Schlussfolgerung mit den Erkenntnissen (SB, S. 25-28). *Nicht erläutert wird im SB, nach welchen Kriterien welches Prognose-Instrument (PCL-R, VRAG, SORAG, Static-99, ODARA/DVRAG, FOTRES) eingesetzt werden soll, noch wann ein solches überhaupt in Frage kommt. Dies wäre wünschbar, weil auch bei den Gruppen B und (wie wir verstehen) auch C nicht immer eine vertiefte Abklärung erfolgen soll. Anzumerken ist bereits an dieser Stelle, dass die zusätzlichen Abklärungen nur dann die Treffsicherheit von ROS erhöhen können, wenn die Einteilung in die drei Gruppen A, B und C tatsächlich der unterschiedlichen Risikodisposition der beurteilten Personen Rechnung trägt.*

Der nächste Schritt ist die **Fallübersicht** (FÜ). Diese ist das Hilfsmittel zur fall- bzw. personenbezogenen Planung und Steuerung des Straf- und Massnahmenvollzugs aus rückfallpräventiver Sicht. Die FÜ stellt die Verknüpfung zwischen der Abklärung und den konkreten Interventionsmöglichkeiten dar. Es beinhaltet die folgenden Bereiche: Zieldelikt, deliktrelevante personenbezogene Problembe- reiche, Deliktrelevante umweltbezogene und psychische Faktoren, deliktverhalten, Risikoquantifizierung, Veränderungsbedarf, Unterstützungsbedarf, Kontrollbedarf (SB, S. 28-29). *Der Versuch, die gesammelten Informationen und Daten zu bündeln und für die Umsetzung auf der Ebene der konkreten Vollzugsplanung nutzbar zu machen ist zweifellos sinnvoll. Allerdings wird die Umsetzung der konkreten Gefährlichkeit einer Person nur insoweit Rechnung tragen können, als die eingesetzten Instrumente zukünftige Risiken korrekt einzuschätzen erlauben.*

Zum Schluss dient die **ROS – Verlaufsliste** (VL) der regelmässigen Standortbestimmung während des Vollzuges einer Strafe oder Massnahme (SB, 29-30). Je nach Auftragsart (d.h. wohl des Anlassdelikts) sind sechs verschiedene Listen in Form von Berichtsformularen erstellt worden: die Alarmliste, die VL Therapie, die VL Motivation, die VL Interventionsprogramme, die VL Vollzug und die VL Bewährung. Wir werden unten auf die Evaluation dieser Verlaufsliste zurückkommen. *Grundsätzlich ist sehr zu begrüessen, dass die Verläufe des Vollzugs in den verschiedenen Bereichen systematisch und möglichst vollständig dokumentiert werden.*

Evaluationskonzept und Ergebnisse (SB, 45ff.)

Die **zentrale Fragestellung** des Modellversuchs lautet: „Wenn es gelingt, Rückfallrisiko und Interventionsbedarf valide zu erfassen, diese in eine Vollzugsplanung zu überführen und den Vollzug entsprechend dieser Planung durchzuführen, dann führt dies zu einem Rückgang der Rückfallraten bei den hiervon betroffenen Straffälligen“ (SB, 32).

Die **Prozessevaluation** (SB, S. 45-46) umfasst die Prüfung der Implementierung, sowie der Übertragung auf andere Kantone und die Bewertung der Kosten-Nutzen-Relation. Die nötigen Informationen sollten durch Gruppeninterviews, Einzelinterviews und Analyse der Akten erhoben werden. Sie wurde einer externen Firma anvertraut. Die Umsetzbarkeit von ROS in den beteiligten Kantonen wurde bejaht. Des Weiteren wurden begünstigende und hinderliche Faktoren identifiziert. Die Kosten-Nutzen-Analyse ergibt, dass die Prozessqualität im Sanktionsvollzug durch ROS optimiert werden konnte. Die Optimierung der Kosten erfolgt durch die Minderung der „unnötigen“ Abklärungen und der „Vermeidung der kriminalitätsbedingten Folgekosten für die Gesellschaft“ (SB, 49-54). *Die im SB zusammengefassten Ergebnisse und namentlich die Kosten-Nutzen-Analyse sind qualitativ und daher vorliegend nicht eigentlich überprüfbar.*

Die **Ergebnisevaluation** (SB, 46-48) umfasst die Prüfung der Validität von FaST im Hinblick auf die künftigen Verläufe. Die Ergebnisevaluation wird in vier Teilaufgaben durchgeführt.

Die **Interrater-Reliabilität des FaST** prüft die Objektivität, die Reliabilität und die Validität. Zwei Mitarbeiter des Kriminologischen Instituts erhielten das FaST-Manual und die gleiche Schulung wie Mitarbeitende des Zürcher Bewährungs- und Vollzugsdienstes (BVD), um sich auf den FaST Einsatz vorzubereiten. Der Einfluss der FaST Ausprägungen (d.h. der Einteilung in die drei Gruppen A, B und C) einer Stichprobe von 40 Fällen, kodiert durch die Mitarbeiter aus dem Kriminologischen Institut, wurde auf ihre Übereinstimmung mit den Kodierungen der BVD-Mitarbeitenden überprüft. Die Kombination zwischen den Ratern ergibt einen Krippendorff's Alphawert von 0.84, der laut dem Schlussbericht als sehr gut übereinstimmend interpretiert wird (SB, 54). *Für uns sind diese Ergebnisse nachvollziehbar. Offensichtlich ist die Einteilung in die drei Gruppen nur zu einem geringen Teil durch die Person bedingt, die die Beurteilung vornimmt.*

Die **Retrospektive Studie** zur Erfassung der Rückfälligkeit befasst sich mit zwei Leitfragen: Werden Personen mit einem Erhöhten Abklärungsbedarf auch zuverlässig als C-Fall eingestuft? Wie gross ist die Übereinstimmung zwischen einem FaST-Ergebnis und einem VRAG-Ergebnis?

Die Informationen zur Stichprobenziehung wurden aus dem RIS (Rechtsinformationssystem des Kantons Zürich) gewonnen, zusätzlich wurde zur bundesweiten Information auf das „vollautomatisierte Strafregister“ (VOSTRA) zugegriffen. Die Stichprobe bildeten alle Fälle, die zwischen 2004-2005 beim BVD abgeschlossen werden konnten (Start der Risiko-Exposition, also der „time-at-risk“ als Folge einer Entlassung). Zum Zeitpunkt der Auswertung waren die meisten Personen in der Stichprobe seit mindestens 6 Jahren in Freiheit. Die wegen Gewalt- und/oder Sexualdelikten verurteilten Personen wurden ausnahmslos in die Stichprobe aufgenommen, um dem besonderen Interesse dieser relativ kleinen Gruppe Rechnung zu tragen. Von den wegen anderen Straftaten Verurteilten wurde demgegenüber eine Zufallsstichprobe von 193 zufällig ausgewählt (EB, S. 14⁵). Nach der Einteilung mussten einige Fälle aus verschiedenen Ausschlussgründen aus der Stichprobe entfernt werden. Die gültige Stichprobe beinhaltet je nach Auswertungs- und Ausschlusskriterien 621 bis 733 Fälle.

Zur Beantwortung der **ersten Leitfrage**, ob nämlich die Gruppeneinteilung der tatsächlichen Gefährlichkeit Rechnung trägt, dienen die Rückfallraten. Das steuerbare Einkommen nach der Sanktion diente weiter als Indikator der sozialen Integration. Um zu prüfen, ob es sich bei den C-Fällen tatsächlich um solche handelt, die einen erhöhten Abklärungsbedarf bezüglich Gewalt- und Sexualdelikte aufweisen, wurde untersucht, ob C-Fälle häufiger mit einem Gewalt- oder Sexualdelikt rückfällig werden als A- und B- Fälle. Wir übernehmen hier die im EB figurierende Tabelle 12.

⁵ EB: Schwarzenegger, C., Manzoni, P. und Baur, M. (2013) Modellversuch. Risikoorientierter Sanktionsvollzug (ROS). Ergebnisevaluation Schlussbericht.

Tabelle 1: Rückfallart, differenziert nach FaST-Kategorien (übernommen aus EB, Tab. 12, S. 26)

		A	B	C	Gesamt
Kein Rückfall oder Rückfall allgemeine Delinquenz	Absolute Häufigkeiten:	430	86	39	555
	Spaltenprozente:	90.0%	86.0%	75.0%	88.1%
	Residualwerte:	0.4	-0.2	-1.0	
Rückfall G/S-Delinquenz	Absolute Häufigkeiten:	48	14	13	75
	Spaltenprozente:	10.0%	14.0%	25.0%	11.9%
	Residualwerte:	-1.2	0.6	2.7	
Gesamt	Absolute Häufigkeiten:	478	100	52	630
	Spaltenprozente:	100.0%	100.0%	100.0%	100.0%

C-Fälle wurden mit 25% signifikant häufiger mit einem G/S-Delikt rückfällig als die anderen Gruppen. Allerdings wurden von den A-Fällen ebenfalls 10% und von den B-Fällen gar 14% wegen eines Gewalt- oder Sexualdelikts (G/S) erneut verurteilt. Die Mehrheit der mit einem G/S-Delikt rückfälligen Personen wies zum Zeitpunkt der FaST-Zuteilung keinen Eintrag wegen eines G/S-Delikts auf – nach FaST wären sie andernfalls, zumindest bei Vorliegen eines Sexualdelikts, wohl der Gruppe C zugewiesen worden (EB, 26). Bei Gewaltdelikten erfolgt die Zuteilung weniger systematisch, wobei ein Fall mit einem schwereren Gewaltdelikt als Anlasstat eher der Gruppe C zugeteilt wurde. *Wir haben in den Berichten keine Angaben darüber gefunden, wie viele Personen mit einer Vorverurteilung wegen eines Gewalt- und/oder Sexualdelikts dennoch den Gruppen A oder B zugewiesen wurden. Dies wäre wünschbar, weil nur so beurteilt werden kann, inwiefern FaST an bekannten Vorstrafen anknüpft. Die Kriterien der Gruppeneinteilung sind daher nicht vollständig transparent, weder in theoretischer noch in empirisch-praktischer Hinsicht.*

Um die **Treffsicherheit** des FaST-Instrumentes zu ermitteln werden, dem Vorbild der Systematic Review von Singh, Grann und Fazel (2011)⁶ folgend, die sog. Positive Predictive Value (PPV) und die Negative Predictive Value (NPV) berechnet. PPV bezieht sich auf die positive Prognose der Rückfälligkeit. Sie ist ein Mass (d.h. der Prozentwert), in dem eine vorausgesagte Rückfälligkeit tatsächlich eintritt. Vorliegend kann man FaST derart interpretieren, dass die Zuteilung in die Gruppe C einer Vorhersage eines erhöhten Risikos im G/S-Bereich gleichkommt. Wie aus der vorstehenden Tabelle 2 (= Tab. 12 im EB) hervorgeht, hat sich bei 13 von 52 Personen in der Gruppe C diese Einschätzung bewahrheitet, da sie erneut einschlägig rückfällig geworden sind. Das entspricht einem PPV von 25%. Man kann Prognose-Instrumente indessen auch danach beurteilen, mit welcher Präzision sie das Fehlen einer solchen Gefährdung voraussagen. Im Hinblick auf die grosse praktische Bedeutung der „falsch-positiven“ Prognosen für die Überlastung der Gefängnisse⁷ ist auch der NPV-Wert eines Prognose-Instrumentes wichtig. Vorliegend kann man die Einteilung in die Gruppe A dahingehend interpretieren, dass nur ein geringes Risiko künftiger Delinquenz im G/S-Bereich vorliegt, und dass bei der Gruppe B dieses Risiko nur mässig hoch ist. Vorliegend ist der Tabelle 2 (vorstehend) zu entnehmen, dass 430 von 478 Personen in der Gruppe A nicht mit einem G/S-Delikt rückfällig geworden sind, was 90% entspricht. Fast gleich hoch, nämlich 86%, ist der NPV-Wert bei der Gruppe B (mit 86 von 100 Fällen). **Insgesamt erweist sich FaST also als recht verlässliches Instrument, wenn es um das Fehlen einer Prognose im Bereich von Gewalt- und/oder Sexualdelikten geht. Umgekehrt ist die Nützlichkeit von FaST indessen sehr**

⁶ Singh, M. Grann & S. Fazel (2011). A comparative study of violence and risk assessment tools: A systematic review and meta-regression analysis of 68 studies involving 25,980 participants, *Clinical Psychology Review*, 31(3): 499-513.

⁷ M. Killias, A. Kuhn, M. Aebi, Grundriss der Kriminologie – eine europäische Perspektive, Bern: Stämpfli, 2011 Rz 1170 (und dort zitierte Literatur).

beschränkt, wenn es um die Prognose eines real gegebenen Risikos im Gewalt- und/oder Sexualbereich geht, werden doch gerade mal 25% der wirklichen – gefährlichen – Verläufe korrekt vorhergesagt.

Wie sind diese Ergebnisse vor dem Hintergrund vergleichbarer Prognose-Instrumente einzuschätzen? In einer systematischen Literaturübersicht haben Singh, Grann und Fazel (2011) die neun heute in der forensischen Praxis am häufigsten verwendeten Prognose-Instrumente miteinander verglichen. Bei der Unterscheidung hoch vs. mittel/wenig gefährlich erzielten diese Instrumente PPV-Werte von 33 bis 76 Prozent. Vor diesem Hintergrund ist das Ergebnis von FaST mit 25 Prozent richtigen Zuweisungen eher bescheiden. Beim NPV sind die Ergebnisse von FaST im Vergleich zu denjenigen anderer Prognose-Instrumente dagegen besser (86-90% für FaST, 53-82% bei den anderen).

Im Schlussbericht (wohl aber im EB, 26 und 74) wird nicht darauf eingegangen, dass von den 75 Rückfälligen durch ein G/S-Delikt lediglich 13 der C-Gruppe angehörten. Es wurden also lediglich 17 % der im G/S-Bereich Rückfälligen in die richtige Fall-Gruppe eingeteilt, und 83% dieser Rückfälle wurden nicht erkannt bzw. nicht in die C-Gruppe eingeteilt. Umgekehrt wurden in der Gruppe C 39 von 52 Personen, also 75%, innerhalb der Beobachtungsperiode von 6 Jahren nicht einschlägig rückfällig. Dies wird im EB ausdrücklich eingeräumt (S. 26, am Ende). Erschwert wird die Evaluation von FaST bezüglich sehr schwerer Verbrechen (vorsätzliche Tötung, schwere Körperverletzung oder qualifizierter Raub) durch die geringen Fallzahlen, wie im Bericht festgehalten wird (EB, Tab. 14, 28). Von den fünf Fällen in dieser Schwerekatgorie wurden vier der Gruppe C zugewiesen. Dies deutet darauf hin, dass Prognosen bei hochgradig selektiven Populationen (wie Verurteilten wegen sehr schwerer Verbrechen) tendenziell präziser werden (in diesem Sinne auch Singh, Grann & Fazel 2011).

Richtig ist, dass ROS vorsieht, dass im Zuge des weiteren Abklärungsbedarfs weitere Prognose-Instrumente (wie VRAG, PCL-R, Static-99 u.a.) eingesetzt werden sollen (SB, 27). Diese Instrumente weisen gemäss der Literaturübersicht von Singh, Grann & Fazel (2011) im Vergleich zu FaST bessere PPV auf (nämlich 66, 52 und 33 Prozent). Allerdings ist FaST als „Eingangstor“ konzipiert: Wer in die Gruppe A gelangt, erfährt kaum eine nähere Abklärung. Insofern können die an sich treffgenauerer Zusatz-Instrumente die gesamte Trefferquote nachträglich nicht verbessern.

*Bezüglich der **allgemeinen Rückfälligkeit** nach 6 Jahren, sind 75% der C-Fälle mindestens 1 Mal Rückfällig geworden, gegenüber den B-Fällen mit 60% und den A-Fällen mit 59%. Die C-Fälle wurden im Verlaufe der 6 Jahren schneller rückfällig als die übrigen Fälle, nicht jedoch im G/S-Bereich (EB, 30). Wie der folgenden Tabelle 2 zu entnehmen ist, werden 23% der C-Fälle mit einem G/S-Delikt rückfällig, also häufiger als Personen in den Gruppen A und B, von denen aber immerhin auch 10% bzw. 14% später wegen eines G/S-Delikt verurteilt werden. Im Bereich der nicht-spezifischen Rückfälligkeit lassen sich keine Unterschiede zwischen den Gruppen A und B feststellen (mit allgemeinen Rückfallraten von 48 und 45%), sogar in der Gruppe C beträgt die Wiederverurteilungsrate wegen unspezifischer Delikte 52%. Diese sehr geringen, wenn auch signifikanten Unterschiede zwischen den drei Gruppen wie auch zwischen den Gruppen C und A/B weisen darauf hin, **dass FaST nicht dazu geeignet ist, den Abklärungsbedarf bei allgemeiner Delinquenz zu ermitteln.***

Tabelle 2: Rückfälligkeit, differenziert nach FaST-Kategorien, 6 Jahre Time-at-risk (übernommen aus EB, Tab. 19, S. 32)

			A	B	C	Gesamt
6 Jahre Time-at-risk	Kein Rückfall	Absolute Häufigkeiten:	197	39	12	248
		Spaltenprozente:	41.4%	40.2%	25.0%	39.8%
		Residualwerte:	0.5	0.0	-1.6	
	Rückfall allgemeine Delinquenz	Absolute Häufigkeiten:	231	44	25	300
		Spaltenprozente:	48.5%	45.4%	52.1%	48.3%
		Residualwerte:	0.1	-0.4	0.4	
	Rückfall G/S-Delinquenz	Absolute Häufigkeiten:	48	14	11	73
		Spaltenprozente:	10.1%	14.4%	22.9%	11.8%
		Residualwerte:	-1.1	0.8	2.3	
	Gesamt	Absolute Häufigkeiten:	476	97	48	621
		Spaltenprozente:	100.0%	100.0%	100.0%	100.0%

$\chi^2(4) = 10.06, p \leq .05$

Aus dieser Tabelle lassen sich erneut die PPV- und die NPV-Werte ableiten. Der PPV-Wert beträgt 45% für die Prognose allgemeiner Delinquenz in der Gruppe B und 23% für G/S-Delinquenz in der Gruppe C. Der NPV-Wert für die Nicht-Rückfälligkeit (in die Gruppe A eingeteilte Personen) beträgt 41%. Diese Werte sind wiederum nicht sehr erfreulich, dies auch im Blick auf andere verbreitete Prognose-Instrumente (s. oben).

Zwischen dem Einkommen und der Einteilung einer Person in die Gruppe A, B oder C sind keine Zusammenhänge erkennbar (EB, 41). Wir diskutieren diesen Aspekt daher hier nicht weiter.

Die **zweite Leitfrage** (EB, 42-47) ist die vergleichende Beurteilung der Validität von FaST und VRAG. Die Einteilung in drei Risikogruppen gemäss FaST zeigte indessen eine geringe Übereinstimmung mit den VRAG-Werten, wie sich der folgenden Tabelle 3 (aus EB, Tab. 27, s. 45) entnehmen lässt. Bei guter Validität wären die VRAG-Werte in der Gruppe B höher als in der Gruppe A und tiefer als in der Gruppe C. Weiter müssten die Personen, auf die das VRAG anwendbar ist, grundsätzlich in der C-Gruppe figurieren, da dieses Instrument auf schwere Gewalt- und Sexualdelinquenz zugeschnitten ist (EB, 15). In Wirklichkeit aber waren von 95 Personen, auf die VRAG angewendet werden konnte, nur 58 der C-Gruppe zugeteilt worden (EB, 42). Von der ursprünglichen Stichprobe von 733 Personen wiesen 519 kein VRAG-relevantes Anlassdelikt oder zu wenig Informationen (119) auf, weshalb nur 95 (oder 13% der Stichprobe) übrig blieben. Im Schlussbericht wurde daher gefolgert, dass „VRAG kein optimal geeignetes Instrument zur Validitätsbestimmung des FaST ist“ (SB, S. 55). *Die Frage bleibt, ob dies am VRAG liegt oder nicht auch mit dem FaST zu tun haben könnte. Nicht völlig nachvollziehbar ist für uns daher, weshalb nicht versucht wurde, die Gruppenzuteilung sowohl anhand des FaST wie auch des VRAG vorzunehmen und die Ergebnisse (bezüglich Rückfall) miteinander zu vergleichen.*

Tabelle 3: Statistische Verurteilung der VRAG-Werte, differenziert nach FaST-Kategorien (übernommen aus EB, Tab. 27, S. 45)

		A	B	C	Gesamt
VRAG	Anzahl Fälle:	15	22	58	95
	Mittelwert:	4.47	5.82	5.60	5.47
	Median:	5.00	6.00	5.50	5.00
	Standardabweichung:	1.36	1.47	1.64	1.61
	Minimum:	2.00	3.00	3.00	2.00
	Maximum:	7.00	9.00	9.00	9.00

Eine **Prospektive Studie** soll die Rückfälligkeit nach Einführung von ROS mit den Prä-ROS-Fällen aufzeigen. Hier geht es somit um die Frage, ob die rückfallpräventive Arbeit durch ROS gegenüber dem früheren Zustand verbessert werden konnte. Diese Studie wird erst nach Abschluss des Modellversuches durchgeführt und liegt noch nicht vor.

Evaluation ROS-Verlaufslisten

Die **Verlaufslisten** dienen der standardisierten Überprüfung des Therapieverlaufs und liegen in Form computergestützter „Checklists“ vor (EB, S. 48). Bisher wurde anscheinend lediglich die Verlaufsliste Therapie, die mit der VL Alarm verknüpft ist, auf ihre Validität überprüft. Anhand von 30 Fällen wurden 15 Berichte über Personen erstellt, die eine ambulante Therapie besuchen. Dabei zeigte sich eine hohe Interrater-Reliabilität, aber auch eine gute Übereinstimmung zwischen den theoretisch abgeleiteten Muster- und den realen Verläufen (EB, 60ff.). Diese Ergebnisse deuten wiederum darauf hin, dass ROS in organisatorischer Hinsicht gut funktioniert.

Gesamteinschätzung

Folgende Feststellungen scheinen uns aufgrund der zugänglichen Dokumentation möglich:

1. Die Evaluation der Rückfälligkeit ist sehr aufschlussreich und wurde kompetent durchgeführt. Das vorgelegte Material ist in methodischer Hinsicht sehr überzeugend.
2. Das Problem von ROS scheint zu sein, dass es auf einem Fall-Screening-Tool (FaST) beruht, das sich nicht besonders gut eignet, um die Gefährlichkeit der in den Straf- und Massnahmenvollzug eingewiesenen Personen zu prognostizieren. Nur jeder vierte Straftäter, der in die „Hochrisiko“-Gruppe C eingeteilt wurde, wird im Laufe der Beobachtungsperiode von sechs Jahren im Bereiche der Gewalt- und/oder Sexualdelikte erneut verurteilt. Von allen in der Stichprobe figurierenden Personen, die wegen eines Gewalt- und/oder Sexualdelikts erneut verurteilt wurden, gelangten nur gerade 17% in die Hochrisikogruppe C. Bei nicht weniger als 83% wurde somit der objektiv gegebene erhöhte Abklärungsbedarf nicht erkannt.
3. Diese Werte liegen deutlich unter denjenigen, die von den heute in der forensischen Praxis am häufigsten verwendeten Prognose-Instrumenten erzielt werden.
4. Soweit es um die Prognose eines günstigen Verlaufs (kein Rückfall, oder zumindest keiner im Bereich der Gewalt- und/oder Sexualdelikte) geht, erzielt FaST hohe Trefferquoten (NPV-Werte von 86% und mehr). Diese sind mit denjenigen, die mittels anderer heute gebräuchlicher Instrumente erzielt werden, vergleichbar.
5. Zur Prognose der allgemeinen Rückfälligkeit (ausserhalb des Bereichs der Gewalt- und/oder Sexualdelikte) erweist sich FaST als ungeeignet. In der B-Gruppe, welcher vor allem Personen mit allgemeiner, d.h. unspezifischer Rückfallgefährdung zugewiesen wurden, war ursprünglich eine vertiefte Abklärung in einzelnen Fällen vorgesehen. Sie fand dann aber aus Ressourcengründen nicht statt (SB, 42). Ob bei einer landesweiten Umsetzung von ROS die dafür nötigen Ressourcen zur Verfügung stünden, darf bezweifelt werden.
6. Relativ gut schneidet FaST bei schwersten Gewaltdelikten ab, allerdings wurden von allen Angehörigen der Stichprobe während des Beobachtungszeitraums von sechs Jahren nur gerade fünf Verbrechen dieser Kategorie (vorsätzliche Tötung, schwere Körperverletzung oder qualifizierter Raub) begangen. Dies macht eine sinnvolle Analyse unmöglich, wie auch im Bericht zutreffend bemerkt wird. Dass Prognosen bei hochgradig selektiven, extremen Gruppen indessen präziser ausfallen, gilt aber auch für andere Prognose-Instrumente.
7. Da FaST als Eingangspforte für vertiefte Abklärungen konzipiert ist, werden die Vorhersagen künftiger Gefährlichkeit (PPV) nicht präziser ausfallen, indem zusätzlich weitere Prognose-Instrumente – vor allem VRAG, Static-99, SORAG, ODARA/DVRAG, FOTRES und PCL-R – eingesetzt werden. Ob es zum Einsatz eines solchen Instruments kommt, hängt nämlich davon ab, ob FaST einen vertieften Abklärungsbedarf erkennt. Wo es dies nicht tut – wie gezeigt ist dies in über 80% der gefährlichen Fälle der Fall – wird auch ein weiteres Screening die

PPV-Werte nicht verbessern können. Diese zusätzlich eingesetzten Instrumente werden dazu beitragen, die falsch-positiven Prognose zu reduzieren. Von den 75% C-Personen, die in den sechs folgenden Jahren kein Gewalt- und/oder Sexualdelikt begangen haben, werden dank zusätzlicher Abklärungen wohl viele als ungefährlich erkannt werden. Verbessert wird damit der NPV-, nicht der PPV-Wert.

8. Im Hinblick auf die Ausdehnung von ROS auf die gesamte Schweiz stellt sich die Frage, ob ein Instrument mit insgesamt doch recht begrenzter Treffsicherheit zum künftigen Standard werden soll. Föderalistische Systeme haben den Vorteil der erleichterten Lernfähigkeit, da an einzelnen Orten erprobte Lösungen nach adäquater Evaluation anderswo übernommen und weiterentwickelt werden können. Zentralistische Systeme neigen dagegen zur Erstarrung einmal eingeführter und implementierter Verfahren. Es fällt jeweils sehr schwer, neue Lösungen zu erproben, weil vom Ansatz her diese - anstatt schrittweise erprobt zu werden - immer gleich flächendeckend implementiert werden müssen.