

An der Klinik und Poliklinik für Psychiatrie und Psychotherapie  
der Ludwig-Maximilians-Universität München  
Direktor: Prof. Dr. H.-J. Möller

Rückfallprognosen in der forensischen Psychiatrie

Vergleich der prädiktiven Validitäten der Prognoseinstrumente  
ILRV, HCR-20, PCL-R und VRAG

Dissertation zum Erwerb des Doktorgrades der Medizin  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität zu München

vorgelegt von  
Carolin Kröner

München  
2005

Mit Genehmigung der Medizinischen Fakultät der Universität München

Berichterstatter..... Prof. Dr. med. Norbert Nedopil

Mitberichterstatter..... Prof. Dr. W. Eisenmenger  
Priv. Doz. Dr. R. Werth  
Prof. Dr. Dr. W. Bender

Mitbetreuung durch den  
promovierten Mitarbeiter..... Dr. med. Cornelis Stadtland

Dekan..... Prof. Dr. med. Dietrich Reinhardt

Tag der mündlichen Prüfung..... 24.11.2005

# Inhaltsverzeichnis

<b>Zusammenfassung.....</b>	<b>4</b>
<b>1. Einleitung.....</b>	<b>5</b>
<b>1.1. Prognoseinstrumente.....</b>	<b>5</b>
1.1.1. FPDS.....	6
1.1.2. ILRV.....	7
1.1.3. HCR-20.....	10
1.1.4. PCL-R.....	12
1.1.5. VRAG.....	15
<b>2. Fragestellung.....</b>	<b>19</b>
<b>3. Methoden.....</b>	<b>21</b>
<b>3.1. Vorgehen.....</b>	<b>21</b>
<b>3.2. Bundeszentralregistergesetz.....</b>	<b>22</b>
<b>3.3. Definitionen.....</b>	<b>24</b>
<b>3.4. Sonderfälle.....</b>	<b>26</b>
<b>3.5. Statistische Begriffe.....</b>	<b>26</b>
3.5.1. Kreuztabellen.....	26
3.5.2. ROC.....	26
3.5.3. Überlebensanalyse nach Kaplan-Meier.....	27
<b>4. Ergebnisse.....</b>	<b>28</b>
<b>4.1. inter-rater Reliabilität.....</b>	<b>28</b>
<b>4.2. Charakterisierung der Stichprobe.....</b>	<b>28</b>
4.2.1. allgemein.....	28
4.2.2. Rückfälligkeit und psychiatrischer Befund.....	32
4.2.3. Rückfälligkeit und Indexdelikte.....	33
<b>4.3. Vergleich der prädiktiven Validitäten.....</b>	<b>35</b>
4.3.1. ILRV.....	35
4.3.2. HCR-20.....	36
4.3.3. PCL-R Gesamtwert und Kombinationen.....	37
<b>4.4. VRAG.....</b>	<b>38</b>
4.4.1. ROC.....	38
4.4.2. VRAG Summenwerte.....	40
4.4.3. Korrelation vorhergesagte Rückfallwahrscheinlichkeit und beobachtete Rückfallrate.....	42
4.4.4. Kaplan-Meier Überlebensanalyse.....	43
<b>5. Diskussion.....</b>	<b>45</b>
<b>5.1. Stichprobe.....</b>	<b>45</b>
<b>5.2. Vergleich mit einer Voruntersuchung an einer anderen Stichprobe.....</b>	<b>47</b>
5.2.1. Ergebnisse der Voruntersuchung.....	48
5.2.2. Vergleich der Ergebnisse.....	48
<b>5.3. VRAG.....</b>	<b>51</b>
<b>5.4. Ausblick.....</b>	<b>56</b>
<b>6. Anhang.....</b>	<b>59</b>
<b>7. Literaturverzeichnis.....</b>	<b>65</b>
<b>8. Lebenslauf.....</b>	<b>73</b>

## Zusammenfassung

Ziel der Arbeit war, die prädiktive Validität der Prognoseinstrumente ILRV, HCR-20 und PCL-R zu vergleichen und den VRAG erstmals an einer deutschen Stichprobe zu validieren. Hierzu wurde das englische Originalmanual des VRAG und des SORAG ins Deutsche übersetzt. Zu Beginn der Studie lagen die Ergebnisse von Voruntersuchungen vor, in denen unter anderem die prognostische Überlegenheit der modifizierten Basisrate für Rückfälligkeit, des PCL-R Gesamtwertes und der historischen Variablen festgestellt worden war. Diese Ergebnisse wurden an der vorliegenden Stichprobe getestet.

Die Prognoseinstrumente wurden anhand von Gutachten für 136 Straftäter ausgefüllt, die 1994 und 1995 an der Universität München auf Schuldunfähigkeit hin begutachtet worden waren. Das Ergebniskriterium Rückfälligkeit wurde im März 2003 für die Probanden anhand der Einträge in das Bundeszentralregister seit ihrer Begutachtung ermittelt. Die durchschnittliche time-at-risk betrug 58.06 Monate. 43 von 113 Probanden (38.1%) wurden rückfällig, 21 Probanden mit einem gewalttätigen Rückfall, 22 Probanden mit einem nicht-gewalttätigen Rückfall.

Die prädiktive Validität der Prognoseinstrumente wurde mittels der AUC in ROC-Analysen bestimmt. Die erstmalige Validierung des VRAG an einer deutschen Stichprobe wurde durch verschiedene Untersuchungen ergänzt: Korrelation mit der Rückfälligkeit für Probanden mit VRAG-Summenwerten unter oder über dem Mittelwert, ROC-Analysen für die gesamte Stichprobe, für Probanden mit gewalttätigen Indexdelikten und für das Ergebniskriterium Rückfälligkeit allgemein bzw. gewalttätige Rückfälligkeit, Korrelation der beobachteten Rückfallsrate mit der von den Autoren des VRAG vorhergesagten Rückfallswahrscheinlichkeit für die 9 Gruppen der VRAG Summenwerte und die Kaplan-Meier Überlebensanalyse für Probanden mit VRAG Summenwerten über und unter dem Mittelwert.

Der VRAG Gesamtwert (AUC .703), die modifizierte Basisrate (AUC .661) und der PCL-R Gesamtwert (AUC .630) korrelierten am besten mit der Rückfälligkeit, sie sollten standardisiert in die klinische Beurteilung des Rückfallrisikos aufgenommen werden. Der VRAG weist auch für diese deutsche Stichprobe die höchste prädiktive Validität aller Prognoseinstrumente auf. Entwicklungsmethode und Bewertungssystem des VRAG stellen wichtige Grundlagen für die zukünftige Prognoseforschung dar. Die Validierung des im Aufbau dem VRAG ähnlichen SORAG für Sexualstraftäter steht in Deutschland noch aus.

## 1. Einleitung

### 1.1. Prognoseinstrumente

Sachverständigengutachten zur Rückfallprognose psychisch kranker Rechtsbrecher sind derzeit erforderlich bei Einweisung in den Massregelvollzug (§§ 63, 64 StGB in Verbindung mit §246a StPO), Entlassung aus dem Massregelvollzug (§§ 67 d II StGB in Verbindung mit § 463 StPO), Entlassung aus lebenslanger Haft (§ 57a StGB in Verbindung mit § 454 StPO) und bei Aussetzung zur Bewährung von Reststrafen bei Verurteilungen von über zwei Jahren bei Straftätern, die wegen eines Sexualdelikts oder anderer „gefährlicher“ Straftaten verurteilt wurden (§ 57 StGB in Verbindung mit § 454 StPO) [65].

Die Methoden zur Erstellung von Rückfallprognosen werden unterschiedlich beurteilt. Einige Autoren favorisieren eine klinische Beurteilung [83], die meisten Veröffentlichungen empfehlen jedoch, die Rückfallwahrscheinlichkeit auch anhand standardisierter Prognoseinstrumente zu ermitteln (Zusammenfassung bei [64]). Prognoseinstrumente berücksichtigen eine Kombination historischer, klinischer und zukunftsorientierter Risikofaktoren. Ihre prädiktiven Validitäten sind Gegenstand zahlreicher Untersuchungen. Die meisten Validierungen wurden bislang im angloamerikanischen und skandinavischen Raum durchgeführt [9;20;41;44;47;50;78;90], im deutschen Sprachraum liegen nur wenige Untersuchungen vor [45;61;71;86;87]. Insbesondere der VRAG, der Static-99 und der SORAG wurden in Deutschland noch nicht validiert. Aufgabe dieser Arbeit war daher, die Basis für die Anwendung von Prognoseinstrumenten zu erweitern und insbesondere den in Deutschland bislang noch nicht validierten VRAG mit anderen Prognoseinstrumenten zu vergleichen.

Folgende Prognoseinstrumente wurden auf ihre prädiktive Validität hin untersucht: die ILRV, der HCR-20, die PCL-R und der VRAG. Zur Strukturierung der Daten wurde das FPDS verwendet.

### 1.1.1. Forensisch-Psychiatrisches Dokumentationssystem (FPDS)

Das „Forensisch-Psychiatrische Dokumentationssystem“ (FPDS) [66] ist ein Dokumentationssystem für die standardisierte Erfassung psychopathologischer und soziodemographischer Daten forensischer Patienten. Es wurde in den 80-er Jahren in Deutschland entwickelt, um die Datenerhebung zu vereinheitlichen, transparenter zu gestalten [66] und die Qualität von Begutachtung und Therapie im Massregelvollzug zu sichern [21;52;60].

Das FPDS ist ein modulares Dokumentationssystem mit einem Hauptmodul „Allgemeine Basisdokumentation“ und fünf Ergänzungsmodulen:

#### **Ergänzungsmodule**

1. Modul Strafrecht
2. Modul Behandlung (§63 StGB)
3. Modul Sucht
4. Modul Sexualität/Sexualstraftäter
5. Modul Prognose

Die Ergänzungsmodule werden je nach Proband und Fragestellung zusätzlich ausgefüllt. In der vorliegenden Studie wurden das Hauptmodul „Allgemeine Basisdokumentation“ und das Modul „Strafrecht“ verwendet.

Das Hauptmodul „Allgemeine Basisdokumentation“ umfasst 110 Variablen und folgende Abschnitte:

#### **Allgemeine Basisdokumentation**

1. Untersuchungsdurchführung
2. soziodemographische Daten
3. familiärer Hintergrund
4. frühere Störungen
5. Sexualität
6. Untersuchungsbefund
7. Diagnose

Das Modul „Strafrecht“ umfasst 82 Variablen und folgende Abschnitte:

#### **Modul Strafrecht**

1. Allgemeines
2. Aktenlage mit Benennung des Delikts und Fragestellung
3. Anamnese der Delinquenz
4. Tatzeitdiagnose
5. psychosoziale Belastungen zum Tatzeitpunkt

## 6. Tatmerkmale

## 7. multiaxiale Zuordnung des psychischen Status, Einsichts- und Steuerungsfähigkeit und die Feststellung von Schuld

Das FPDS beinhaltet qualitative und quantitative Variablen. Die qualitativen Variablen (z.B. Alkoholmissbrauch bei mindestens einem Elternteil oder Gewalttätigkeit in der Primärfamilie) werden mit „1“ als vorhanden und mit „0“ als nicht vorhanden codiert. Qualitative Variablen wie Religionszugehörigkeit oder Delikt werden mit definierten Ziffern codiert. Quantitative Variablen (z.B.: Untersuchungsbefund Wahnvorstellungen oder Affektstörungen) werden mit „0“ (nicht vorhanden), „1“ (vorhanden in geringer Ausprägung), „2“ (vorhanden in mässiger Ausprägung) und „3“ (vorhanden in sehr starker Ausprägung) codiert. Merkmale, die nicht beurteilbar sind, werden mit 9, 99, oder 9999 codiert.

### 1.1.2. Integrated List of Risk Variables (ILRV)

Die ILRV ist das Prognosemodul des FPDS (Forensisch Psychiatrisches Dokumentationssystem) [66]. Es handelt sich um eine Checkliste zur Beurteilung des Rückfallrisikos forensischer Patienten. Die ILRV geht auf Gutachtenanalysen aus dem Jahr 1986 zurück (Nedopil 1986) und bestand von Anfang an aus vier Teilen:

1. den Besonderheiten des Indexdelikts,
2. den Aspekten der prädeliktischen Persönlichkeit,
3. der postdeliktischen Persönlichkeitsentwicklung
4. den Aspekten des sozialen Empfangsraums.

Bei der ersten Veröffentlichung des HCR-20 im Jahr 1995 [101] ergaben sich Überlappungen mit der ILRV: die historischen Variablen waren nahezu identisch. Um Vergleich und Operationalisierung zu vereinfachen, wurden diese Variablen aus dem HCR-20 übernommen, und mit (H1) bis (H10) bzw. (C2),(C4), (R2),(R4) und (R5) markiert. Die endgültige Version der ILRV wurde 1997 veröffentlicht und besteht aus folgenden Abschnitten:

#### **Abschnitt A (Ausgangsdelikt)**

1. deliktbezogene Basisrate für Rückfälligkeit modifiziert durch persönliche Faktoren der Täter:
2. spielten situationelle Faktoren eine Rolle bei der Tat
3. lag ein Einfluss einer vorübergehenden Krankheit vor
4. wie relevant war die Persönlichkeit
5. gab es eine Motivation durch kriminogene oder sexuell deviante Bedürfnisse

## **Abschnitt B (anamnestische Daten)**

umfasst historische Variablen.

Sie sind entsprechend dem HCR-20 mit (H1) bis (H10) markiert.

## **Abschnitt C (Postdeliktische Persönlichkeitsentwicklung (klinische Variablen))**

1. Krankheitseinsicht
2. Therapiemotivation
3. selbstkritischer Umgang mit bisheriger Delinquenz
4. Besserung psychopathologischer Auffälligkeiten
5. Coping Mechanismen
6. Widerstand gegen Folgeschäden durch Institutionalisierung
7. Antisoziale Lebenseinstellung (C2)
8. fortgesetzte Impulsivität (C4)

## **Abschnitt D (Sozialer Empfangsraum (Risikovariablen))**

1. Arbeit
2. Unterkunft
3. soziale Beziehungen mit Kontrollfunktionen
4. Offizielle Kontrollmöglichkeiten
5. Konfliktbereiche
6. Verfügbarkeit von Opfern
7. Zugangsmöglichkeiten zu Risiken (R2)
8. Compliance (R4)
9. Stressoren (R5)

Zusätzlich wird in der ILRV das Ausgangsdelikt benannt, der PCL-R Summenwert notiert und die Prognose als günstig - ausreichend - nicht ausreichend oder ungünstig bewertet.

Die Variablen der ILRV werden auf einer Punkteskala von 0 bis 2 bewertet. Basisraten unter 3% werden als niedrig (0 Punkte), zwischen 3% und 25% als mittel (1 Punkt) und über 25% als hoch (2 Punkte) definiert.

Die ILRV unterscheidet sich von anderen Prognoseinstrumenten darin, dass in ihrer Entwicklung bereits frühzeitig die Basisraten für Rückfälligkeit berücksichtigt wurden, die von verschiedenen Autoren, wie Dietz (1985) [13] und Hall (1997) [34], als wesentlich für die Einschätzung einer Vorhersage im forensischen Bereich angesehen wurden. Die ILRV ist bislang das einzige veröffentlichte Prognoseinstrument, das die Basisrate für Rückfälligkeit als Einzelvariable berücksichtigt.

Die Basisrate ist definiert als das „Vorkommen von Tätern eines Delikts in einer bestimmten, nach allgemeinen Kriterien definierten Menge von Menschen“ [97]. Basisraten



sind innerhalb eines bestimmten gesellschaftlichen, kulturellen und zeitlichen Zusammenhangs relativ stabile Grössen. Die „Basisrate für Rückfälligkeit“ ist die empirisch ermittelte Rate für Rückfälligkeit bei Tätern einer bestimmten Deliktgruppe. Ihre prognostische Relevanz wurde in mehreren Studien beschrieben [22;27], eine Vernachlässigung von Basisraten kann zu falschen Rückfallprognosen führen [24;25;49]. Einige Autoren nehmen sogar an, dass die Basisrate der wichtigste Faktor für Langzeitprognosen ist [34;64;84].

Unterschieden werden muss zwischen der Basisrate für Kriminalität und der deliktspezifischen Basisrate für Rückfälligkeit. Bei ersterer wird mit der Allgemeinbevölkerung verglichen, bei zweiterer mit der Menge der jeweiligen Delinquenten. Delinquenz in der Allgemeinbevölkerung hat eine Basisrate unter 5%. Gewaltverbrechen haben eine Basisrate von 0.01% (Bundeskriminalstatistik). Die Basisraten für Rückfälligkeit werden mit <2% für Mord bis zu 70% für Drogendelinquenz angegeben. Bei vielen Delikten sind die Basisraten für Rückfälligkeit nicht genau bekannt [67]. Sie sind abhängig von vielen Faktoren, z.B. von der Anzeigen- und Aufklärungsquote eines Delikts, dem Geschlecht und Alter der Täter oder von einer psychischen Erkrankung. Da sie schwierig zu ermitteln sind, gibt es bislang lediglich Schätzgrössen, die in der folgenden Tabelle aufgeführt sind [64]. Detailliertere Aufstellungen gibt es mittlerweile bei Gross 2004 [33], sie lagen bei der Erarbeitung dieser Untersuchung noch nicht vor.

### **Basisraten (Näherungswerte) für Rückfälligkeit**

#### **Delikte mit Rezidivraten über 50%**

Strassenverkehrsdelikte

Drogendelinquenz

Sexualdelikte bei homosexueller Pädophilie

#### **Delikte mit Rezidivraten zwischen 25% und 50%**

Körperverletzung

Eigentumsdelinquenz

Exhibitionismus

Sexualdelikte bei Pädophilie

#### **Delikte mit Rezidivraten zwischen 10% und 25%**

Raub

Brandstiftung

Vergewaltigung und Sexuelle Nötigung

#### **Delikte mit Rezidivraten zwischen 3% und 10%**

Inzest

Gewaltdelikte bei Pädophilie

#### **Delikte mit Rezidivraten zwischen 0% und 3%**

Mord und Totschlag

Der Schweregrad des Delikts ist in vielen Fällen den Rezidivraten invers proportional. Die Basisrate für Rückfälligkeit wird in der Regel nach unten zu korrigieren sein, wenn Delinquenz nur in bestimmten Lebenskrisen oder –situationen auftritt, wie z.B. während der Pubertät, im Klimakterium, bei Ehescheidung oder dem Tod eines Angehörigen. Sie wird auch nach unten korrigiert, wenn Delikte ausschliesslich unter dem Einfluss einer vorübergehenden Krankheit auftreten. Nach oben zu korrigieren ist sie, wenn die Delikte persönlichkeitsgebunden in einer Vielzahl unterschiedlicher Situationen auftreten, oder wenn die Motivationen in kriminogenen Bedürfnissen der Täter liegen. Die Variable „modifizierte Basisrate“ wird in der ILRV anhand dieser Faktoren angepasst.

Die Basisraten hängen mit der Effektivität einer prognostischen Methode in folgender Weise zusammen: Für Basisraten unter 50% gilt: die prognostische Methode ist effektiv, wenn die Basisrate grösser als die Fehlerrate der Methode ist. E wenn  $BR > fp + fn$ . (falsch-positive plus falsch-negative). Für Basisraten über 50% gilt: die prognostische Methode ist effektiv, wenn 1-Basisrate grösser als die Fehlerrate der Methode ist. E wenn  $1 - BR > fp + fn$  [67].

### 1.1.3. HCR-20 (historical–clinical–risk management)

Der HCR-20 (Version 2) [100] wurde als Checkliste für Kliniker entwickelt, um das Risiko von Probanden mit gewalttätigem Verhalten in der Vergangenheit zu bewerten, bei denen der Verdacht auf eine psychische Erkrankung vorliegt.

Der HCR-20 besteht aus 20 Variablen:

#### **Historische Variablen**

---

<b>H1</b>	Frühere Gewaltanwendung
<b>H2</b>	Geringes Alter bei erster Gewaltanwendung
<b>H3</b>	Instabile Beziehungen
<b>H4</b>	Probleme im Arbeitsbereich
<b>H5</b>	Substanzmissbrauch
<b>H6</b>	(gravierende) seelische Störung
<b>H7</b>	PCL-R Gesamtwert
<b>H8</b>	Frühe Fehlanpassung
<b>H9</b>	Persönlichkeitsstörung
<b>H10</b>	Frühere Verstösse gegen Auflagen

### **Klinische Variablen**

---

<b>C1</b>	Mangel an Einsicht
<b>C2</b>	Negative Einstellungen
<b>C3</b>	Aktive Symptome
<b>C4</b>	Impulsivität
<b>C5</b>	Fehlender Behandlungserfolg

### **Risikomanagement**

---

<b>R1</b>	Fehlen realisierbarer Pläne
<b>R2</b>	Destabilisierende Einflüsse
<b>R3</b>	Mangel an Unterstützung
<b>R4</b>	Fehlende Compliance
<b>R5</b>	Stressoren

Die Bewertung der Variablen erfolgt in einem 3 Punkte Schema:

<b>0</b>	Nein	Das Item trifft definitiv nicht zu
<b>1</b>	Möglich	Das Item trifft möglicherweise/teilweise zu.
<b>2</b>	Ja	Das Item trifft sicher zu.
<b>9</b>	unbekannt	Informationen nicht ausreichend für eine Bewertung

Der H-Wert beträgt maximal 20 Punkte, der C-Wert maximal 10 Punkte, der R-Wert maximal 10 Punkte und der Gesamtwert maximal 40 Punkte. Die historischen Variablen entsprechen in ihrer Gewichtung der Summe aus klinischen und Risikomanagement Variablen. Daraus ergibt sich eine drei-stufige Bewertung des Gewalttrisikos: niedriges, mittleres oder hohes Risiko für zukünftige Gewalttaten. Die Autoren des HCR-20 legen keine Punktwerte für die drei Risikobereiche fest, da die Punktwerte nicht immer linear mit dem Risiko korrelieren. Beispielsweise kann bei einem niedrigen Gesamtwert im HCR-20 das Gewalttrisiko dennoch „hoch“ sein, wenn z.B. imperative Stimmen mit Tötungsinhalt oder akut-drängende sadistische Phantasien vorhanden sind. Hingegen kann trotz eines hohen Gewaltpotentials aufgrund einer körperlichen Erkrankung oder eines herabgesetzten Handlungspotentials das tatsächliche Gewalttrisiko niedrig sein.

Die Konsequenzen, die aus der Risikobewertung für zukünftige Gewalttaten gezogen werden, hängen vom sozialen Gefüge der Probanden ab, ihrem Gesundheitszustand und von Belastungsfaktoren. Bei einem „mittleren Gewalttrisiko“ wird empfohlen, den künftigen Entwicklungsverlauf im Sinne eines individuellen, gewaltprotektiven Schutzprogrammes regelmässig zu kontrollieren. Ein „hohes Gewalttrisiko“ erfordert hingegen ein striktes Risikomanagement mit verstärkter Supervision, Information der Mitarbeiter und der Auflage für die Probanden, an speziellen Behandlungsprogrammen teilzunehmen [100].

Die prädiktive Validität des HCR-20 wird in verschiedenen Studien mit mittel bis gut angegeben [4;16;17;30;59;68;89;90]. Täter mit überdurchschnittlichen Punktwerten im HCR-20 haben ein 6-fach erhöhtes Risiko, gewalttätig zu werden (Übersicht: [18]). Die Interrater-Reliabilität liegt zwischen .82 - .85 [18].

#### 1.1.4. Psychopathy Checklist Revised - PCL-R

Die „Psychopathy-Checklist“ [38] basiert auf dem klinischen Konstrukt der „psychopathy“. Im deutschen Sprachgebrauch ist der Begriff „Psychopathie“ umstritten und mit abwertenden Konnotationen belastet. Viele Begriffe, die dieses Phänomen beschreiben sollen, sind nicht einheitlich definiert, so ist beispielsweise der Unterschied zwischen „psychopathischer oder abnormer Persönlichkeit“, „Soziopathie“ oder „antisozialer Persönlichkeit“ nicht eindeutig. Im Deutschen wird daher zunehmend der englische Begriff „psychopathy“ verwendet, der bereits seit 1812 einheitlich definiert wurde: Nach Rush (1812) umfasst die „psychopathy“ eine „moral alienation of mind“ mit Verantwortungslosigkeit, Aggressivität und Rücksichtslosigkeit. In Anlehnung daran charakterisierte Checkley 1941 [8] „psychopathic individuals“ als Menschen mit inadäquat motiviertem, antisozialem Verhalten, mit Fehlen einer emotionalen Bindung und der Unfähigkeit zu Schuldgefühl und Bedauern. „Psychopathic individuals“ weisen Persönlichkeitsmerkmale auf, die mit einem erhöhten Deliktrisiko einhergehen: sie zeigen dissoziales oder antisoziales Verhalten, Impulsivität, mangelnde Empathie und einen ausgeprägten Egoismus. „Psychopathic individuals“ erscheinen im interpersonellen Bereich „übertrieben, egozentrisch, manipulativ, dominant, gewaltsam und kaltherzig.“ Affektiv zeigen sie sich „flach und labil, unfähig, längerdauernde Beziehungen zu anderen Menschen zu knüpfen, Prinzipien oder Ziele einzuhalten.“ Sie weisen einen Mangel an Angst, echtem Schuldbewusstsein und Reue auf. Aus ursprünglich 16 Persönlichkeitsmerkmalen nach Checkley entwickelte Hare 1980 die „Psychopathy Checklist“ [36] und 1991 die PCL-R, die „Psychopathy Checklist-Revised“ [38]. Diese umfasst 20 Variablen:

#### **PCL-R**

---

1. Trickreicher sprachgewandter Blender mit oberflächlichem Charme
2. erheblich übersteigertes Selbstwertgefühl
3. Stimulationsbedürfnis (Erlebnishunger), ständiges Gefühl der Langeweile
4. Pathologisches Lügen (Pseudologie)
5. Betrügerisch-manipulatives Verhalten
6. Mangel an Gewissensbissen oder Schuldbewusstsein
7. Oberflächliche Gefühle

8. Gefühlskälte, Mangel an Empathie
9. Parasitärer Lebensstil
10. Unzureichende Verhaltenskontrolle
11. Promiskuität
12. Frühe Verhaltensauffälligkeiten
13. Fehlen von realistischen, langfristigen Zielen
14. Impulsivität
15. Verantwortungslosigkeit
16. Mangelnde Bereitschaft und Fähigkeit, Verantwortung für eigenes Handeln zu übernehmen
17. Viele kurzzeitige ehe(ähn)liche Beziehungen
18. Jugendkriminalität
19. Widerruf einer bedingten Entlassung
20. Polytrope Kriminalität

Anhand von Informationen aus semi-strukturierten Interviews und Gutachten werden die Variablen in einem 3-Punkte System bewertet:

- 0 Das Merkmal trifft auf den Untersuchten nicht zu.
- 1 Das Merkmal trifft bis zu einem gewissen Mass auf den Untersuchten zu, in einigen Punkten bestehen Übereinstimmungen, aber eine gewisse Unsicherheit bleibt.
- 2 Die Eigenschaften des Untersuchten treffen in den meisten wesentlichen Eigenschaften auf das Merkmal zu.

Ergeben sich erhebliche Diskrepanzen aus Interview- und Ergänzungsinformationen, werden diese Variablen ausgelassen. Zur Ermittlung des Gesamtwerts werden die fehlenden Variablen registriert, es ergibt sich ein „prorated“, ein angepasster Gesamtwert zwischen 0 und 40. Er gibt an, in wie weit der Proband dem Prototyp des „psychopaths“ entspricht.

In den USA wird der Untersuchte ab einem Grenzwert von 30 als „psychopath“ klassifiziert [38]. In europäischen Studien neigt man dazu, den empfohlenen Grenzwert von 30 herabzusetzen, da die Prävalenz von „psychopathy“ in Europa niedriger ist. So liegt sie bei schottischen Strafgefangenen bei 3%, in Nordamerika hingegen bei 23% [10]. Unterschiede in der Prävalenz bestehen auch je nach Geschlecht und Diagnose der Probanden: die Basisrate für „psychopathy“ ist bei weiblichen Probanden deutlich niedriger als bei männlichen. Möglicherweise liegt die Prävalenzrate tatsächlich unter der für männliche Probanden oder die PCL-R ist auf weibliche Probanden nur eingeschränkt anwendbar [96]. Die PCL-R Gesamtwerte sind am höchsten für Patienten mit einer Persönlichkeitsstörung, mit Substanzmissbrauch oder einer Komorbidität beider. Probanden mit anderen schweren psychiatrischen Erkrankung weisen niedrigere Werte auf, die niedrigsten Werte erzielen Probanden ohne psychiatrische Diagnose [86].

Die Charakterisierung des „psychopathic“ Täters hat sich empirisch als wichtiger Prognosefaktor bestätigt. Individuen mit hohen PCL-R Gesamtwerten begehen mehr gewaltsame Verbrechen als Individuen mit niedrigen Werten [84], haben ein höheres Risiko, einen gewalttätigen Rückfall zu begehen [46;59;68;96], werden früher rückfällig [42] und sprechen schlechter auf eine Therapie an [72]. Probanden mit hohen PCL-R Werten weisen veränderte physiologische Reaktionen nach emotionalen Stimuli, z.B. Geräuschen auf. So ist z.B. die Leitfähigkeit der Haut im Vergleich zu gesunden Probanden vermindert und die Pulsveränderung verzögert [94]. Nicht jeder nach der PCL-R klassifizierte „psychopath“ handelt unbedingt kriminell. Wenn jedoch kriminelles Verhalten auftritt, besteht empirisch eine gute Korrelation zwischen dem PCL-R Wert und der Rückfallquote dieser Täter [82]. Nicht eindeutig ist, ob die „psychopathic“ Eigenschaften im Verlauf stabil sind [39] oder ob die Ausprägung der „psychopathy“ mit zunehmendem Alter im Sinne eines burn-out abnimmt [93].

Bis vor einigen Jahren wurde die PCL-R als Zwei-Faktoren-Modell aufgefasst: Faktor 1 beschreibt die emotionslose Rücksichtslosigkeit des Probanden in Bezug auf Gefühle oder Rechte anderer, Faktor 2 das Vorhandensein von durchgängig antisozialem Verhalten (einschliesslich jugendlicher Delinquenz und krimineller Vielseitigkeit) [38;40]. Neuerdings wird die PCL-R als 3-Faktoren Modell beschrieben [48]: Faktor 1 umfasst den zwischenmenschlichen Bereich mit täuschenden und manipulativen Facetten, Faktor 2 den affektiven Bereich mit Rücksichts- und Gefühlslosigkeit und der Faktor 3 die Verhaltensstörung mit einer unverantwortlichen und impulsiven Lebensweise [11]. Neuere Untersuchungen belegen die Überlegenheit des 3-Faktoren Modells [35;48].

Die gute Operationalisierung der PCL-R ermöglicht eine reliable und valide Datenerhebung. Die PCL-R gilt als „state-of-the art“ [23] und ist als Variable in anderen Prognoseinstrumenten enthalten (VRAG [43] und HCR-20 [100]). In der ILRV [63] ist der PCL-R Wert nicht integriert, die PCL-R wird jedoch zusätzlich ausgewertet und dokumentiert.

Die prädiktive Validität der PCL-R zur Vorhersage von Rückfälligkeit wird in verschiedenen Studien mit mittel bis gut angegeben [15;29;30;45;54;80;89;91;98;99]. Die Interrater-Reliabilität der PCL-R liegt zwischen .74 und .97 [37].

### 1.1.5. Violence Risk Appraisal Guide - VRAG

Der VRAG unterscheidet sich von den bisher beschriebenen Prognoseinstrumenten in seiner Entwicklungsmethode, der Auswahl seiner Variablen und seinem Bewertungssystem. Der VRAG ist als ein aktuarisches Instrument für Gerichte, Kliniker und Bewährungshelfer konstruiert, das für Probanden mit schwerwiegenden Delikten die Wahrscheinlichkeit eines gewalttätigen Rückfalls angibt. Der VRAG wurde als dimensionales Konstrukt anhand von empirischen Daten entwickelt und 1993 von Harris et al. [43] veröffentlicht.

Ihre Arbeitsgruppe untersuchte eine Stichprobe von 618 psychisch kranken, männlichen Rechtsbrechern aus einer Hochsicherheitsanstalt auf die Wahrscheinlichkeit hin, mit der ein Proband erneut verhaftet oder für ein neues gewalttätiges Delikt angeklagt wurde. Bei einer time-at-risk von 7 Jahren wurden 31% der ursprünglichen Stichprobe gewalttätig rückfällig. 12 statische Risikofaktoren wurden ermittelt, die am besten mit der Rückfälligkeit korrelierten.

Neben dem PCL-R Gesamtwert, der am stärksten gewichtet ist, werden im VRAG folgende anamnestiche Informationen der Probanden in absteigender Gewichtung registriert:

#### **VRAG**

---

1. PCL-R Gesamtwert
2. Probleme des Probanden in der Grundschule
3. das Alter des Probanden zum Zeitpunkt des Indexdelikts
4. die Erfüllung der DSM III Kriterien für eine Persönlichkeitsstörung
5. wuchs der Proband mit beiden biologischen Elternteilen auf
6. liegt ein vorausgegangenes Bewährungsversagen vor
7. der Punktwert für die kriminelle Vorgeschichte nach dem Cormier-Lang Bewertungssystem (siehe Anhang)
8. war er jemals verheiratet
9. die Erfüllung der DSM III Kriterien für Schizophrenie
10. der Verletzungsgrad des Opfers/der Opfer
11. die Alkoholanamnese des Probanden
12. das Geschlecht der Opfer

Alter, Schizophrenie, weibliches Opfer und der Verletzungsgrad des Opfers korrelieren negativ mit dem Rückfallrisiko, alle anderen Variablen korrelieren positiv [42].

Für die fachgerechte Anwendung des VRAG sollte der Bewerter mit der Entwicklungsmethode des VRAG und den zugrundeliegenden Definitionen vertraut sein [74]:

Die 618 männlichen Probanden der ursprünglichen Stichprobe hatten alle mindestens eine schwerwiegende antisoziale Tat verübt. Probanden, die leichtere Delikte verübt hatten, und Personen, die noch kein Delikt verübt hatten, wurden von der Studie ausgeschlossen. Alle Probanden waren mindestens einmal psychiatrisch behandelt worden.

Die **time-at-risk** begann an dem Tag, an dem der Proband die Möglichkeit hatte, rückfällig zu werden und endete mit dem ersten gewalttätigen Rückfall oder im April 1988 (Ende der Studie). Die Zeit, während der der Proband aufgrund eines nicht-gewalttätigen Delikts institutionalisiert war, wurde abgezogen. Der Mittelwert der time-at-risk betrug 81.5 Monate (Standardabweichung 60.6).

Das Ergebnis **gewalttätiger Rückfall** beinhaltete folgende Delikte: Mord, versuchter Mord, Kidnapping, Freiheitsberaubung, Körperverletzung, ein Angriff mit Körperverletzung, Vergewaltigung, bewaffneter Raub, alle sexuellen Vergehen mit Gewalt. Gewertet wurde auch institutionelle Gewalttätigkeit, die eine Verhaftung nach sich gezogen hätte, wenn sie in Freiheit begangen worden wäre.

Die **Möglichkeit**, ein weiteres Gewaltdelikt zu verüben, war gegeben, wenn der Proband in Freiheit entlassen wurde oder in einem psychiatrischen Krankenhaus oder Heim mit nur minimalen Sicherheitsvorkehrungen untergebracht war. Probanden in Haft, die nicht die Möglichkeit hatten, rückfällig zu werden, wurden nicht berücksichtigt. Drei Probanden wurden als gewalttätig rückfällig berücksichtigt, obwohl sie eigentlich nicht die Möglichkeit dazu gehabt hätten (Flucht o.ä.).

Anhand dieser Stichprobe wurden etwa 50 unabhängige Variablen auf ihre Korrelation mit der Rückfälligkeit hin untersucht. Sie umfassten die Bereiche soziodemographische Information, Probleme in der Kindheit, Anpassung als Erwachsene, Charakterisierung des Indexdelikts und psychologische Variablen zur Beurteilung der Verfassung der Probanden nach dem Delikt. Variablen, die nicht korrelierten, wurden gestrichen. Ergab sich eine hohe Korrelation zwischen verschiedenen Variablen (z.B. frühere Anklagen und frühere Verurteilung für ein Gewaltdelikt) wurde die Variable mit der niedrigeren Korrelation gestrichen.

Im zweiten Schritt wurden multiple Regressionsanalysen für verschiedene Gruppen von Variablen (Kindheit, Anpassung im Erwachsenenalter, Indexdelikt etc.) und für



verschiedene Gruppen von Probanden (nur begutachtet, behandelt etc.) angewendet. Regression beschreibt eine Methode zur Ermittlung der funktionalen Abhängigkeit verschiedener Variablen, in diesem Fall der ausgewählten Variablen und der Ergebnisvariablen „Rückfälligkeit“ [102]. Nur Variablen, die in einer Mehrheit von Untergruppen selektiert worden waren, wurden in die abschliessende Regressionsanalyse übernommen. 12 dieser Variablen wurden schliesslich zum VRAG zusammengefasst [43]. Ihre durchschnittliche Interkorrelation beträgt weniger als .3 [42].

Die Gewichtung der 12 selektierten Variablen erfolgte nach einer von Nuffield [74] entwickelten Methode: gewichtet wurde, wie sehr sich das Individuum von der Basisrate unterschied. Ein Unterschied von mindestens 5% von der gewalttätigen Rückfallrate der Gesamtstichprobe (31%) führte zu der Addition oder Subtraktion von je einem Punkt. Verheiratete Probanden beispielsweise hatten eine Rückfallsrate für gewalttätiges Verhalten von 21% - d.h. zwei volle 5% Intervalle unter der Basisrate. Unverheiratete Probanden hatten eine Rückfallsrate von 38%. Somit werden verheiratete Probanden mit einem Punktwert von -2 bewertet, unverheiratete Probanden mit einem Punktwert von +1. Fehlende Variablen werden mit dem Wert 0 codiert.

Die Punkteskala zur Bewertung des VRAG unterscheidet sich von der anderer Prognoseinstrumente, sie umfasst auch negative Werte. Der Summenwert des VRAG kann Werte zwischen -26 und 38 annehmen. Er korreliert positiv mit mindestens einem gewalttätigen Rückfall, mit der Schwere des Rückfalls und mit der Kürze der Zeitspanne in der ein Rückfall auftritt. Der Mittelwert der VRAG-Summenwerte betrug in der Originalstichprobe .91 (Standardabweichung 12.9).

Der VRAG wurde als dimensionales Konstrukt entwickelt: anhand der Summenwerte wurde die Originalstichprobe in 9 gleich grosse Gruppen mit je 7 Punkten eingeteilt. Für diese Gruppen wurden die Rückfallraten innerhalb von 7 und von 10 Jahren berechnet; sie entsprechen den „vorhergesagten Wahrscheinlichkeiten“:

**Rückfallwahrscheinlichkeiten für Gewalttaten über 2 verschiedene Zeiträume als  
Funktion von 9 gleich grossen VRAG Gruppen**

<b>VRAG Kategorie</b>	<b>VRAG Punktwert</b>	<b>7 Jahre</b>	<b>10 Jahre</b>
<b>1</b>	<= -22	0	0.08
<b>2</b>	-21 bis -15	0.08	0.10
<b>3</b>	-14 bis -8	0.12	0.24
<b>4</b>	-7 bis -1	0.17	0.31
<b>5</b>	0 bis +6	0.35	0.48
<b>6</b>	+7 bis +13	0.44	0.58
<b>7</b>	+14 bis +20	0.55	0.64
<b>8</b>	+21 bis +27	0.76	0.82
<b>9</b>	>= +28	1.00	1.00

In der Originalstichprobe wies der VRAG eine hohe prädiktive Validität auf, mit einer AUC (Fläche unter der ROC) von .76 für die Vorhersage von gewalttätigen Rückfällen (95% C.I. = +- .03). Die „common language effect size“ betrug .76: werden zwei Probanden zufällig aus der Gruppe der „Rückfälligen“ und der Gruppe der „Nicht-Rückfälligen“ gezogen, hat der Proband aus der Gruppe der „Rückfälligen“ mit .76 Wahrscheinlichkeit einen höheren VRAG Wert als der Proband aus der anderen Gruppe. Der traditionellere Cohen's *d* Wert (Cohen 1969) betrug 1.06 (nach Cohen ist  $d \geq .8$  gross).

Für die optimale Anwendung des VRAG sollten dem Bewerter folgende Informationen vorliegen: dimensionale Wahrscheinlichkeitsangaben, die operationale Definition des vorhergesagten Ergebnisses, normative Informationen über vergleichbare Risikowerte anderer Individuen und eine kurze Zusammenfassung der Methode, die zur Entwicklung des VRAG geführt hat [74].

In Anlehnung an den VRAG wurde in gleicher Weise ein Instrument für die Beurteilung der Rückfallwahrscheinlichkeit bei Sexualstraftätern entwickelt, der SORAG (Sex Offender Risk Appraisal Guide), der dem VRAG in wesentlichen Variablen entspricht.

## 2. Fragestellung

Die prädiktive Validität eines Prognoseinstruments besagt, mit welcher Wahrscheinlichkeit seine Anwendung zu einer korrekten Prognose führt. In der vorliegenden Studie wurde die prädiktive Validität der beschriebenen Instrumente ermittelt. Es wurden u.a. die Ergebnisse von Voruntersuchungen im Rahmen des Münchner Projekts zur Rückfallprognose getestet [71;89]:

In den Voruntersuchungen waren die prädiktiven Validitäten verschiedener Prognoseinstrumente untersucht worden anhand einer Stichprobe von 185 Probanden, die in den Jahren 1992 und 1993 an der Universität München auf ihre Schuldfähigkeit hin begutachtet worden waren. Die Prognoseinstrumente waren im Jahr 2001 basierend auf den Gutachten ausgewertet worden und die Rückfälle der Probanden im Zeitraum nach ihrer Begutachtung Ende 2001 anhand der Bundeszentralregisterauszüge ermittelt. Die Gutachten umfassten im Mittel über 50 Seiten und enthielten alle erforderlichen Informationen, um -mit wenigen Ausnahmen- die Prognoseinstrumente problemlos auszufüllen.

Neben den prädiktiven Validitäten der Gesamtwerte der Prognoseinstrumente war auch die prognostische Relevanz der Untergruppen der Prognoseinstrumente getestet worden (Basisraten, klinische und historische Variablen). Folgende Ergebnisse der Voruntersuchungen lagen vor:

Signifikante Korrelationen mit der Rückfälligkeit ergaben sich

1. für die historischen Variablen  
(Abschnitt B der ILRV und H1 bis H10 des HCR-20)
2. den PCL-R Gesamtwert und
3. die „modifizierte Basisrate“ (Abschnitt A der ILRV).

Die beste Korrelation mit der Rückfälligkeit wurde

4. durch die Kombination des PCL-R Gesamtwerts mit der modifizierten Basisrate erzielt (AUC der Voruntersuchung .698).

Keine signifikante Korrelation ergab sich für

5. die klinischen Variablen (Abschnitt C der ILRV) und für die Variablen betreffend der sozialen Situation der Probanden nach Entlassung (Abschnitt D der ILRV).

Diese Ergebnisse wurden in der vorliegenden Studie an einer neuen Stichprobe getestet.

Der VRAG war in der Vergangenheit nur im angloamerikanischen und skandinavischen Raum validiert worden [1;2;26;28;44;53;56;58;62;69;70;75;76;80;81;85;92] (Übersicht siehe [42]). In dieser Arbeit wurde das englische Originalmanual des VRAG und des SORAG ins Deutsche übersetzt und die prädiktive Validität des VRAG erstmals an einer deutschen Stichprobe ermittelt.

### 3. Methoden

#### 3.1. Vorgehen

136 Probanden waren in den Jahren 1994 und 1995 in der forensischen Abteilung der Klinik für Psychiatrie und Psychosomatik der Universitätsklinik München auf die strafrechtliche Verantwortlichkeit hin nach §§20 und 21 (Schuldfähigkeit) bzw. ggfs. nach §§63 und 64StGB (Unterbringung) begutachtet worden. Im Gegensatz zu den meisten anderen Studien hing bei den Probanden der Münchner Prognosestudie der weitere Verlauf nicht von der psychiatrisch abgegebenen Prognose ab. Die meisten Studien untersuchen Rückfälligkeit nach Entlassungen aus Massregelvollzugseinrichtungen oder Haftanstalten. Entlassungen aus dem Massregelvollzug und aus lebenslanger Haft ebenso wie vorzeitige Entlassungen aus Haftanstalten können nur bei günstiger Prognose erfolgen. Deshalb kann bei ihnen nur die Unrichtigkeit einer günstigen Prognose überprüft werden, d.h. es werden nur die falsch-negativen Prognosen erfasst. In der Münchner Studie wurden hingegen ebenso die falsch-positiven Prognosen erfasst, da die Prognose den Entscheidungsträgern unbekannt blieb und keinen Einfluss auf den weiteren Werdegang der Untersuchten hatte.

Die Stichprobe war repräsentativ und nicht selektiert. Die Schwere der Delikte ging von Beleidigung über Fahren ohne Fahrerlaubnis bis hin zu Mord.

Die Gutachten basierten auf umfangreichen Untersuchungen, die Befunde waren ausführlich dargestellt, so dass ein Ausfüllen der Prognoseinstrumente ohne Schwierigkeiten und ohne dass Informationen fehlten, möglich war. Die Gutachten wurden zwischen Dezember 2002 und März 2003 anhand folgender Prognoseinstrumente ausgewertet:

1. Forensisch Psychiatrisches Dokumentationssystem  
Hauptmodul Basisdokumentation  
Modul Strafrecht
2. ILRV
3. HCR-20
4. PCL-R
5. VRAG

Die Auswertung erfolgte anhand der Manuale der Prognoseinstrumente, nach fachgerechtem Training und unter Supervision. Bei Unklarheiten wurde Rücksprache mit den entsprechenden Gutachtern gehalten. Die ausgewerteten Prognoseinstrumente wurden in SPSS 11.0 für Windows dokumentiert.

11 Gutachten wurden von zwei Bewertern unabhängig anhand der Prognoseinstrumente ausgewertet, um die inter-rater Reliabilität zu testen.

Im März 2003 wurde das Ergebniskriterium „Rückfälligkeit“ für 128 Probanden anhand der Bundeszentralregisterauszüge ermittelt. Für acht Probanden wurde kein Bundeszentralregisterauszug angefordert: ein Gutachten war nicht auffindbar, ein Proband hatte zum Zeitpunkt der Gutachtenerstellung ein Glioblastom und sechs Probanden waren über 80 Jahre alt. Die Bundeszentralregisterauszüge wurden zu wissenschaftlichen Zwecken nach § 42a BZRG (Bundeszentralregistergesetz) bei der Dienststelle Bundeszentralregister des Generalbundesanwaltes beim Bundesgerichtshof in Bonn angefordert.

### 3.2. Bundeszentralregistergesetz

Das Bundeszentralregistergesetz (BZRG) vom 18. März 1971 (BGBl. I S.243) in der Fassung der Bekanntmachung vom 21. September 1984 (BGBl. I S.1229, 1985 I S.195) wurde zuletzt geändert durch Artikel 4 des Gesetzes vom 21. August 2002 (BGBl. I S.3344).

Im Folgenden sind die relevanten Gesetzestexte zusammengefasst, vgl. auch [http://www.bundeszentralregister.de/bzr/bzrg\\_text.htm](http://www.bundeszentralregister.de/bzr/bzrg_text.htm).

Nach § 3 BZRG enthält das Zentralregister folgenden Inhalt:

1. strafgerichtliche Verurteilungen (siehe unten)
2. (aufgehoben)
3. Entscheidungen von Verwaltungsbehörden und Gerichten (§ 10)
4. Vermerke über Schuldunfähigkeit (§ 11)
5. gerichtliche Feststellungen nach § 17 Abs. 2, § 18
6. nachträgliche Entscheidungen und Tatsachen, die sich bis auf eine auf die in den Nummern 1 bis 4 genannten Eintragungen beziehen (§§ 12 bis 16, § 17 Abs. 1).

Nach § 4 BZRG sind rechtskräftige Entscheidungen einzutragen, durch die ein deutsches Gericht im Geltungsbereich des Bundeszentralregistergesetzes wegen einer rechtswidrigen Tat

1. auf Strafe erkannt
2. eine Massregel der Besserung und Sicherung angeordnet
3. nach § 59 des Strafgesetzbuchs mit Strafvorbehalt verwarnt oder

4. nach § 27 des Jugendgerichtsgesetzes die Schuld eines Jugendlichen oder Heranwachsenden festgestellt hat.

Nach § 54 Abs. 1 BZRG werden ausserdem strafrechtliche Verurteilungen, die nicht durch deutsche Gerichte im Geltungsbereich dieses Gesetzes ergangen sind, unter bestimmten Voraussetzungen in das Register eingetragen.

Nach § 5 BZRG haben die Eintragungen folgenden Inhalt:

1. Die Personendaten des Betroffenen: Geburtsname, ein hiervon abweichender Familienname, Vorname, Geschlecht, Geburtsdatum, Geburtsort, Staatsangehörigkeit und die Anschrift sowie abweichende Personendaten.
2. die entscheidende Stelle samt Geschäftsnummer
3. der Tag der (letzten) Tat
4. der Tag des ersten Urteils (...)
5. der Tag der Rechtskraft
6. die rechtliche Bezeichnung der Tat, deren der Verurteilte schuldig gesprochen worden ist, unter Angabe der verwendeten Strafvorschriften,
7. die verhängten Strafen

In den folgenden Paragraphen werden aufzunehmende Informationen genauer definiert:

- § 6 Gesamtstrafe und Einheitsstrafe
- § 7 Aussetzung zur Bewährung
- § 8 Sperre für die Fahrerlaubnis
- § 9 (aufgehoben)
- § 10 Entscheidungen von Verwaltungsbehörden und Gerichten
- § 11 Schuldunfähigkeit
- (...)

Um das Ergebniskriterium **Rückfall** zu bestimmen, wurden folgende Eintragungen der Probanden in das Bundeszentralregister seit dem Zeitpunkt ihrer Begutachtung 1994/1995 dokumentiert:

1. die time-at-risk: der Beobachtungszeitraum, vom Tag der Freilassung bis zum Tag einer erneuten Tat bzw. bis zum Ende des Beobachtungszeitraums
2. die Art des ersten Rückfalls

3. die Gesamtzahl der Rückfälle
4. die Art des schwersten Rückfalls.

Dokumentiert wurde auch, wenn ein Verfahren wegen Schuldunfähigkeit eingestellt worden war, der Proband verstorben war, abgeschoben worden war oder seit dem Indexdelikt die Haftanstalt noch nicht verlassen hatte.

### 3.3. Definitionen

#### **Rückfälle**

Folgende Delikte waren in den Bundeszentralregisterauszügen vermerkt:

1. Verstöße gegen Bewährungs- oder Führungsaufsichtsaufgaben ohne erneute Straftat
2. allgemeine, nicht-gewalttätige Delikte (z.B.: Fahren ohne Fahrerlaubnis, Diebstahl, Betrug)
3. gewalttätige Delikte oder Versuche (z.B.: Körperverletzung, bewaffneter Raub)
4. Sexualdelikte oder Versuche (z.B.: Vergewaltigung, Nötigung)

Kategorien 2 bis 4 wurden als Rückfälle gewertet.

#### **Gewalttätig**

Als gewalttätig wurden gewertet: Mord und Totschlag, Körperverletzung, Vergewaltigung/Nötigung, sonstige Sexualdelikte sowie Erpressung und Entführung. Alle übrigen Delikte wurden als nicht-gewalttätig definiert.

Das **Indexdelikt** bezeichnet das Delikt, das zur ursprünglichen Begutachtung geführt hat.

Das Bundeszentralregister enthielt für 26 Probanden der Stichprobe den Vermerk **keine Eintragung**. Dieser Vermerk kann bedeuten:

- dass der Proband vor Jahren verstorben ist. Nach der neuen Fassung des Bundeszentralregistergesetzes werden Eintragungen drei Jahre nach dem Todestag gelöscht (§24.1)
- dass der Proband über 90 Jahre alt ist (§24(2))



- dass ein Verfahren wegen Schuldunfähigkeit eingestellt wurde. Bei Einstellung eines Verfahrens wegen Schuldunfähigkeit (§11) werden die Eintragungen nach zehn Jahren gelöscht, wegen eines Verbrechens nach 20 Jahren. Sind mehrere Eintragungen nach §11 vorhanden, werden diese erst gelöscht, wenn für alle Eintragungen die Voraussetzungen der Entfernung vorliegen. (§24 (3) (4))
- dass das Geburtsdatum falsch angegeben wurde und der Proband daher nicht eindeutig zuzuordnen war
- dass verschiedene Straftaten nicht einheitlich unter Familien- bzw. Geburtsname registriert sind, und die Straftaten nicht eindeutig zuzuordnen sind
- dass der Eintrag getilgt wurde.

Die Tilgung von Einträgen erfolgt nach § 45 BZRG in folgenden Fällen:

1. Eintragungen über Verurteilungen (§4) werden nach Ablauf einer bestimmten Frist getilgt bzw. nicht mehr aufgenommen (§33(1)) – danach darf über die Eintragung keine Auskunft erteilt werden.
2. Eine zu tilgende Eintragung wird ein Jahr nach Eintritt der Tilgungsreife aus dem Register entfernt. Während dieser Zeit darf über die Eintragung keine Auskunft erteilt werden.
3. Der Absatz 1 gilt nicht
  - bei Verurteilungen zu lebenslanger Freiheitsstrafe
  - bei Anordnung der Unterbringung in der Sicherungsverwahrung oder in einem psychiatrischen Krankenhaus.(§33(2))

§ 46 ff. geben die Tilgungsfristen vor (zwischen fünf und zwanzig Jahren je nach Delikt) und regeln den Ablauf der Tilgung.

Die **time-at-risk** beschreibt den Zeitraum nach dem Index-Delikt, während dem der Proband die Möglichkeit hat, rückfällig zu werden:

- Für inhaftierte Probanden beginnt die time-at-risk am Tag der Entlassung aus Haft, Entziehungsanstalt, Massregelvollzug oder Sicherungsverwahrung.
- Für Probanden, die nicht inhaftiert waren, beginnt sie zum Zeitpunkt des Index Delikts.
- Die time-at risk endet jeweils zum Zeitpunkt des ersten Rückfalls, der einen Eintrag im Bundeszentralregisterauszug erhielt. Erfolgte keine Eintragung, endete sie zum Zeitpunkt der Ausstellung der Registerauszüge.
- Für Probanden mit dem Vermerk „keine Eintragung“ beginnt die time-at-risk am Tag des Gutachtens und endet am Tag der Ausstellung der Bundeszentralregisterauszüge. Diese Probanden wurden als nicht rückfällig gewertet.

### 3.4. Sonderfälle

Drei Probanden der Stichprobe wurden rückfällig, bevor die time-at-risk angelaufen war. Sie begingen im Rahmen einer Lockerungsmassnahme einen Verstoss gegen Bewährungs- oder Führungsaufsichtsaufgaben, aber keine erneute Straftat (technisches Versagen), kehrten z.B. von einer Beurlaubung nicht zurück. Für diese Fälle wurde die time-at-risk mit 0 Monaten bewertet und das Delikt nicht als Rückfall definiert.

### 3.5. Statistische Begriffe

#### 3.5.1. Kreuztabellen

Die Unabhängigkeit zweier Variablen für Kreuztabellen mit einer 2 mal 2 Verteilung wurde mit dem exakten Test nach Fisher geprüft. Als Mass für den linearen Zusammenhang zweier Variablen wurde der Korrelationskoeffizient nach Pearson  $r$  berechnet.

#### 3.5.2. receiver operating characteristics - ROC

Um die Genauigkeit von Prognoseinstrumenten zu prüfen, schlug Mossman 1994 die Verwendung von „receiver operating characteristics“ (ROCs) vor, die ursprünglich im Kontext von Signalübertragung im Maschinenwesen und Psychophysik entwickelt wurden.

Sie wurden bereits im 2. Weltkrieg eingesetzt, um zwischen Zielen und nicht-Zielen zu unterscheiden, beispielsweise zwischen Rauschen und Rauschen plus Signal auf einem Radarempfänger. Der Ausdruck „receiver operating characteristics“ beschreibt, dass die „Charakteristika“ eines Tests analysiert werden können, und der „receiver“ – der „Datenempfänger“ mit jedem beliebigen Kurvenpunkt „operieren“ kann.

Die ROC Kurve ist somit die Funktion der „Trefferrate“ oder der Rate der Richtig-Positiven (Sensitivität) zur „Falsche-Alarm Rate“ (oder 1 minus Spezifität). Die ROC Kurve zieht durch die Koordinaten (0/0) und (1/1). Anhand der Fläche unter der Kurve kann die Validität eines Prognoseinstruments beurteilt werden: je grösser die AUC, die „area-under-the-curve“, desto grösser ist die Differenz an jedem Punkt zwischen der Trefferrate und der Falsche-Alarm-Rate, desto besser ist das Vorhersagemodell. Je grösser die Fläche, desto höher die Wahrscheinlichkeit, dass ein rein zufällig selektionierter gewalttätiger Proband in dem Prognoseinstrument einen höheren Summenwert hat als ein zufällig selektionierter, nicht-gewalttätiger Proband. Die positive Diagonale ist eine Linie, an der die „Trefferrate“ gleich der „Falsche-Alarm-Rate“ ist (Sensitivität = 1-Spezifität). An den Punkten der Diagonale kann der Test nicht zwischen gewalttätig und nicht-gewalttätig unterscheiden, die AUC beträgt .50. Ein perfekter Test hingegen besitzt einen oder mehrere Schnittpunkte, an denen Sensitivität und Spezifität gleichzeitig 1 betragen. Die ROC wäre in diesem Falle eine horizontale Linie, die die Punkte (0/1) und (1/1) verbindet. ROCs sind unabhängig von der Basisrate und vom Niveau des Grenzwertes zur Klassifizierung von Gewalttätigkeit [79].

### 3.5.3. Überlebensanalyse nach Kaplan-Meier

Kaplan-Meier Kurven veranschaulichen die Verteilung von Zeiten zwischen zwei Ereignissen, in diesem Fall zwischen Beginn und Ende der time-at-risk. Die Kaplan-Meier Methode basiert auf der Schätzung bedingter Wahrscheinlichkeiten, in Anwesenheit dieser nicht-ermittelbaren oder „zensierten“ Fälle zu jedem Zeitpunkt eines auftretenden Ereignisses. Durch Bilden des Produktgrenzwerts dieser Wahrscheinlichkeiten werden die Überlebensraten zu jedem Zeitpunkt geschätzt. Kaplan-Meier Kurven veranschaulichen, wie häufig und wie schnell verschiedene Gruppen von Probanden kriminell rückfällig geworden sind. Um die Gleichheit der Überlebensverteilungen für die verschiedenen Stufen des Faktors zu testen, wurden der Log-Rang, der Tarone-Ware Test und der Breslow Test angewandt.

## 4. Ergebnisse

### 4.1. inter-rater Reliabilität

11 Gutachten wurden von zwei Bewertern unabhängig voneinander ausgewertet, um die inter-rater Reliabilität zu bestimmen. Der Intraclass Korrelationskoeffizient betrug für den HCR-20 .89 (95% C.I. .62-.97), für die PCL-R .92 (95% C.I. .70-.98) und für die ILRV .93 (95% C.I. .76-.98).

### 4.2. Charakterisierung der Stichprobe

#### 4.2.1. allgemein

Von 136 begutachteten Probanden wurde für 8 Probanden kein Bundeszentralregisterauszug angefordert: ein Gutachten war nicht auffindbar, ein Proband hatte zum Zeitpunkt der Gutachtenerstellung ein Glioblastom und 6 Probanden waren über 80 Jahre alt.

Tabelle 1

	<b>Anzahl der Probanden</b>
<b>Gesamt</b>	136
<b>Gutachten nicht auffindbar</b>	1
<b>&gt; 80 Jahre bei Begutachtung</b>	6
<b>Glioblastom bei Begutachtung</b>	1
<b>BZR Auszug angefordert</b>	128

Von den 128 Probanden, für die ein Bundeszentralregisterauszug angefordert worden war, waren 6 Probanden verstorben, 5 Probanden noch in Haft und bei 4 Probanden war das Verfahren wegen Schuldunfähigkeit eingestellt worden. Die Stichprobe, für die das Ergebniskriterium Rückfälligkeit ermittelt werden konnte, bestand aus 113 Probanden, die alle die Möglichkeit hatten, rückfällig zu werden.

Tabelle 2

	<b>Anzahl der Probanden</b>
<b>BZR Auszug angefordert</b>	128
<b>Verstorben</b>	6
<b>Noch in Haft</b>	5
<b>Verfahren eingestellt wegen Schuldunfähigkeit</b>	4
<b>Rückfälligkeit ermittelt</b>	113

93 Probanden (82.3%) waren männlich, 20 Probanden (17.7%) weiblich (Tabelle 3). 92 Probanden (81.4%) kamen aus Deutschland, 9 (8.0%) aus Südeuropa, 8 (7.1%) aus der Türkei/dem Orient, ein Proband aus Nordamerika und drei aus anderen Ländern (Tabelle 4). Der Mittelwert der Altersverteilung betrug 33.7 Jahre, der jüngste Proband war zum Zeitpunkt der Gutachtenerstellung 17 Jahre alt, der älteste 71 (Tabelle 5).

Tabelle 3

<b>Geschlecht</b>	<b>Anzahl der Probanden</b>	<b>Prozent</b>
<b>Männlich</b>	93	82.3
<b>Weiblich</b>	20	17.7
<b>Gesamt</b>	113	100

Tabelle 4

<b>Kultureller Hintergrund</b>	<b>Anzahl der Probanden</b>	<b>Prozent</b>
<b>Deutsch</b>	92	81.4
<b>Südeuropa</b>	9	8.0
<b>Türkisch/Orient</b>	8	7.1
<b>Nordamerika</b>	1	0.9
<b>Sonstige Ausländer</b>	3	2.6
<b>Gesamt</b>	113	100

Tabelle 5

<b>Altersgruppen (Jahre)</b>	<b>Anzahl der Probanden</b>	<b>Prozent</b>
<b>17-20</b>	10	8.8
<b>21-30</b>	40	35.4
<b>31-40</b>	36	31.9
<b>41-50</b>	19	16.8
<b>51-60</b>	5	4.4
<b>61-70</b>	2	1.8
<b>71-</b>	1	0.9
<b>Gesamt</b>	113	100

Folgende Tabellen beschreiben Familienstand und Schulbildung der Probanden:

Tabelle 6

<b>Familienstand</b>	<b>Anzahl der Probanden</b>	<b>Prozent</b>
<b>Ledig</b>	60	53.1
<b>Verheiratet zusammenlebend</b>	28	24.8
<b>Verheiratet getrenntlebend</b>	11	9.7
<b>Geschieden</b>	14	12.4
<b>Gesamt</b>	113	100

Tabelle 7

<b>Schulbildung</b>	<b>Häufigkeit</b>	<b>Prozent</b>
<b>Kein Abschluss</b>	25	22.1
<b>Sonderschulabschluss</b>	4	3.5
<b>Qualifizierender Hauptschulabschluss</b>	10	8.9
<b>Sonstiger Hauptschulabschluss</b>	33	29.2
<b>Mittlere Reife</b>	24	21.2
<b>Polytechnische Oberschule</b>	3	2.7
<b>Fachabitur</b>	1	.9
<b>Abitur</b>	11	9.7
<b>Gesamt I</b>	111	98.2
<b>Unklar</b>	2	1.8
<b>Gesamt II</b>	113	100

Zum Zeitpunkt der Begutachtung hatten 40 Probanden (35.4%) keinen psychiatrischen Krankheitsbefund, für 73 Probanden (64.6%) wurde eine Diagnose nach ICD-10 [103] gestellt. Die grösste Diagnosegruppe (26 Probanden, 23%) war die Gruppe F1 „Psychische und Verhaltensstörung durch psychotrope Substanzen“. 21 Probanden (18.6%) litten an einer „Persönlichkeits- und Verhaltensstörung“ (F6). Bei je 8 Probanden (7.1%) wurde eine „Schizophrenie, schizotype oder wahnhaftige Störung“ (F2) diagnostiziert bzw. eine „affektive Störung“ (F3), 6 Probanden (5.3%) litten an einer „organischen, einschliesslich symptomatischen psychischen Störung“ (F0) (Tabelle 8).

Tabelle 8

<b>ICD-10 Achse der Hauptdiagnose</b>	<b>Häufigkeit</b>	<b>Prozent</b>
<b>F 0 organische, einschliesslich symptomatische psychische Störungen</b>	6	5.3
<b>F 1 Psychische und Verhaltensstörungen durch psychotrope Substanzen</b>	26	23.0
<b>F 2 Schizophrenie, schizotype und wahnhaftige Störungen</b>	8	7.1
<b>F 3 Affektive Störungen</b>	8	7.1
<b>F 4 Neurotische, Belastungs- und somatoforme Störungen</b>	3	2.6
<b>F 5 Verhaltensauffälligkeiten in Verbindung mit körperlichen Störungen oder Faktoren</b>	1	0.9
<b>F 6 Persönlichkeits- und Verhaltensstörungen</b>	21	18.6
<b>F 9 Verhaltens- und emotionale Störungen mit Beginn in der Kindheit und Jugend</b>	0	0
<b>Ohne Diagnose</b>	40	35.4
<b>Gesamt</b>	113	100

58 (51.3%) Indexdelikte waren gewalttätig (Mord und Totschlag, Körperverletzung, Vergewaltigung/Nötigung, sonstige Sexualdelikte, Erpressung/Entführung), 55 (48.7%) Indexdelikte waren nicht-gewalttätig (Tabelle 9). Nur sieben Probanden hatten ein Sexualdelikt verübt, so dass die Prognoseinstrumente für Sexualdelikte (z.B. SORAG, Static-99) nicht in die Analyse aufgenommen wurden.

Tabelle 9

<b>Indexdelikte</b>	<b>Häufigkeit</b>	<b>Prozent</b>
<b>Mord und Totschlag</b>	22	19.5
<b>Körperverletzung</b>	17	15.0
<b>Vergewaltigung/Nötigung</b>	5	4.4
<b>Raub</b>	10	8.8
<b>Diebstahl</b>	18	15.9
<b>Betrug</b>	9	8.0
<b>Erpressung/Entführung</b>	2	1.8
<b>Delikte nach dem Betäubungsmittelgesetz</b>	11	9.7
<b>Brandstiftung</b>	6	5.3
<b>Anderes</b>	11	9.7
<b>Sonstige Sexualdelikte</b>	2	1.8
<b>Gesamt</b>	113	100

Das Ergebniskriterium Rückfälligkeit konnte für 113 Probanden ermittelt werden: 43 Probanden (38.1%) wurden rückfällig, 21 Probanden (18.6%) mit einem gewalttätigen Rückfall und 22 Probanden (19.5%) mit einem nicht-gewalttätigen Rückfall. 70 Probanden (61.9%) wurden nicht rückfällig, d.h. sie hatten entweder keinen erneuten Eintrag im Bundeszentralregister oder einen Bewährungswiderruf wegen Nichterfüllung der Bewährungsauflagen (Technisches Versagen) (Tabelle 10).

Tabelle 10

<b>Schwerster Rückfall</b>	<b>Häufigkeit</b>	<b>Prozent</b>
<b>Kein Rückfall</b>	62	54.9
<b>Bewährungswiderruf ohne neues Delikt</b>	8	7.1
<b>Nicht-gewalttätige Rückfälle</b>	22	19.5
<b>Gewalttätige Rückfälle (und Versuche)</b>	19	16.8
<b>Gewalttätige Sexualdelikte (und Versuche)</b>	2	1.8
<b>Gesamt</b>	113	100

Die time-at-risk der Probanden lag zwischen 0 und 115 Monaten, der Mittelwert betrug 58.06 Monate (Standardfehler 3.392).

Männliche Probanden wurden tendenziell häufiger rückfällig als weibliche ( $p = 0.0792$ ): insgesamt wurden 39 männliche Probanden (41.9% der männlichen Probanden) rückfällig, 19 (20.4%) begingen einen gewalttätigen Rückfall und 20 (21.5%) einen nicht-gewalttätigen Rückfall. Insgesamt wurden nur 4 weibliche Probanden (20% der weiblichen Probanden) rückfällig, davon nur eine (5%) mit einem gewalttätigen Rückfall (Tabelle 11).

Tabelle 11

<b>Geschlecht</b>	<b>Männlich</b>	<b>Weiblich</b>
<b>Kein Rückfall</b>	54 58.1%	16 80%
<b>Rückfall insgesamt</b>	39 41.9%	4 20%
<b>RF ohne Gewalt</b>	19 20.4%	3 15%
<b>RF mit Gewalt</b>	20 21.5%	1 5%
<b>Gesamt</b>	93 100%	20 100%

#### 4.2.2. Rückfälligkeit und psychiatrischer Befund

Am häufigsten rückfällig wurden die Probanden der Diagnosegruppen F0 „organische, einschliesslich symptomatische psychische Störungen“ und F4 „neurotische, Belastungs- und somatoforme Störungen“ (je 66.7% aller Probanden mit den Diagnosen F0 und F4). 57.1% der Probanden mit der Diagnose F6 „Persönlichkeits- und Verhaltensstörungen“ wurden rückfällig, 46.2% mit F1 „Psychische und Verhaltensstörungen durch psychotrope Substanzen“. Am seltensten wurden die Probanden mit der Diagnose F2 „Schizophrenie, schizotype und wahnhaftige Störungen“ (12.5%) rückfällig, seltener als die Probanden ohne psychiatrischen Befund (25%). Gewalttätige Rückfälle traten gehäuft in den Diagnosegruppen F0 (50%) und F6 (33.3%) auf (Tabelle 12).

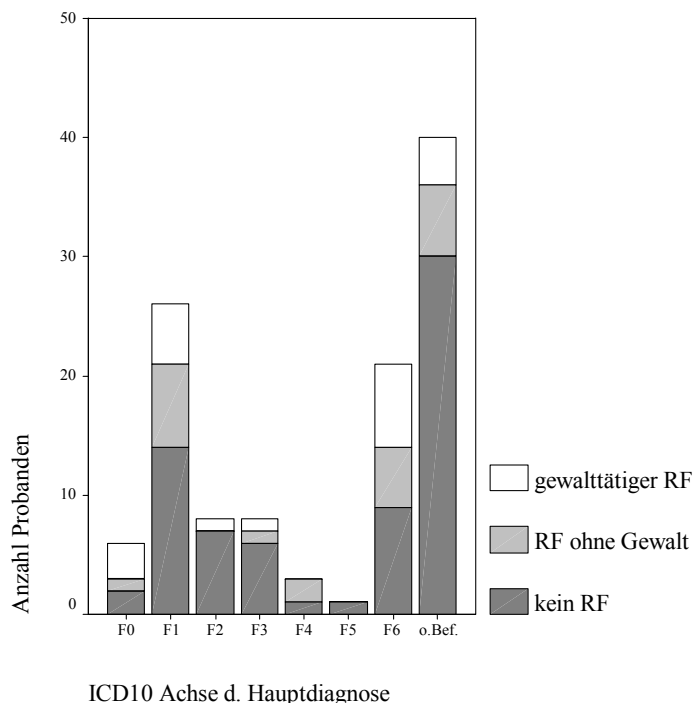
Tabelle 12

<b>ICD10 Hauptdiagnose</b>	<b>F0</b>	<b>F1</b>	<b>F2</b>	<b>F3</b>	<b>F4</b>	<b>F5</b>	<b>F6</b>	<b>Ohne Befund</b>
<b>Kein Rückfall</b>	2 33.3%	14 53.8%	7 87.5%	6 75%	1 33.3%	1 100%	9 42.9%	30 75%
<b>Rückfall insgesamt</b>	4 66.7%	12 46.2%	1 12.5%	2 25%	2 66.7%	0	12 57.1%	10 25%
<b>RF ohne Gewalt</b>	1 16.7%	7 26.9%	0	1 12.5%	2 66.7%	0	5 23.8%	6 15%
<b>RF mit Gewalt</b>	3 50%	5 19.5%	1 12.5%	1 12.5%	0	0	7 33.3%	4 10%
<b>Gesamt</b>	6 100%	26 100%	8 100%	8 100%	3 100%	1 100%	21 100%	40 100%



Bei dieser Analyse ist die unterschiedliche Anzahl der Probanden der einzelnen Diagnosegruppen zu berücksichtigen:

Diagramm 1



#### 4.2.3. Rückfälligkeit und Indexdelikt

Am häufigsten rückfällig wurden die Probanden mit dem Indexdelikt Körperverletzung (76.5% der Probanden dieser Gruppe). Aus den Gruppen Diebstahl, Betrug, Erpressung und Entführung, sonstige Sexualdelikte wurden je etwa die Hälfte der Probanden rückfällig. Am seltensten rückfällig wurden die Probanden mit dem Indexdelikt Mord und Totschlag, nur 4.5% dieser Gruppe. Gewalttätige Rückfälle traten am häufigsten in den Gruppen Erpressung und Entführung (50%), sonstige Sexualdelikte (50%) und Körperverletzung (35.3%) auf (Tabelle 13).

Tabelle 13

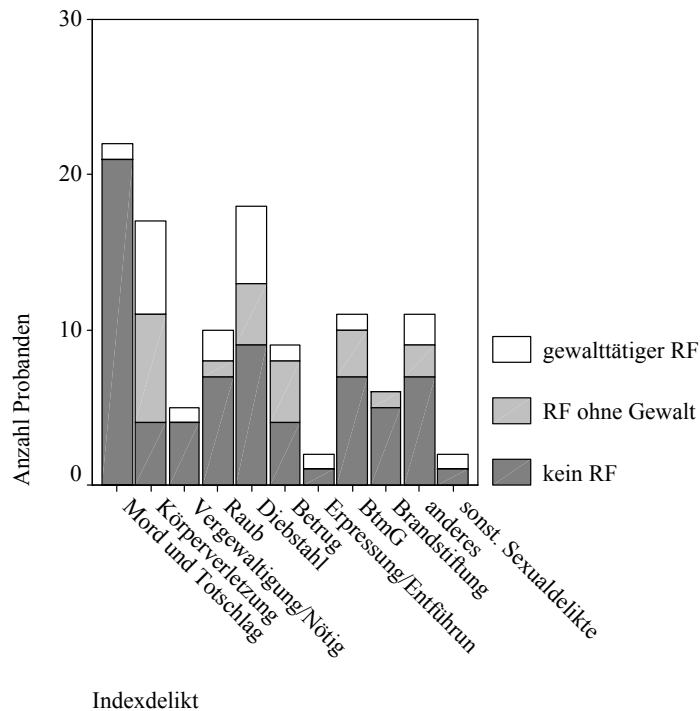
<b>Indexdelikt</b>	<b>Mord und Totschlag</b>	<b>Körperverletzung</b>	<b>Vergewaltigung/Nötigung</b>	<b>Raub</b>	<b>Diebstahl</b>
<b>Kein Rückfall</b>	21 95.5%	4 23.5%	4 80%	7 70%	9 50%
<b>Rückfall insgesamt</b>	1 4.5%	13 76.5%	1 20%	3 30%	9 50%
<b>RF ohne Gewalt</b>	0	7 41.2%	0	1 10%	4 22.2%
<b>RF mit Gewalt</b>	1 4.5%	6 35.3%	1 20%	2 20%	5 27.8%
<b>Gesamt</b>	22 100%	17 100%	5 100%	10 100%	18 100%

Fortsetzung Tabelle 13

<b>Indexdelikt</b>	<b>Betrug</b>	<b>Erpressung/ Entführung</b>	<b>BtmG</b>	<b>Brandstiftung</b>	<b>Anderes</b>	<b>Sonstige Sexualdelikte</b>
<b>Kein Rückfall</b>	4 44.4%	1 50%	7 63.6%	5 83.3%	7 63.6%	1 50%
<b>Rückfall insgesamt</b>	5 55.6%	1 50%	4 36.4%	1 16.7%	4 36.4%	1 50%
<b>RF ohne Gewalt</b>	4 44.4 %	0	3 27.3%	1 16.7%	2 18.2%	0
<b>RF mit Gewalt</b>	1 11.1%	1 50%	1 9.1%	0	2 18.2%	1 50%
<b>Gesamt</b>	9 100%	2 100%	11 100%	6 100%	11 100%	2 100%

Wieder ist in der Analyse die unterschiedliche Anzahl der Probanden in den jeweiligen Gruppen zu berücksichtigen; die Gegenüberstellung gewalttätige versus nicht-gewalttätige Indexdelikte ergab in Bezug auf die Rückfälligkeit keinen statistisch signifikanten Unterschied ( $p = 0.445$ ).

Diagramm 2



### 4.3. Vergleich der prädiktiven Validitäten

#### 4.3.1. ILRV Gesamtwert und Abschnitte A bis D

ROC Analysen wurden für den ILRV Gesamtwert und die Abschnitte A bis D durchgeführt. Die modifizierte Basisrate hatte die höchste AUC (.661). Der Trend für die klinischen Variablen (Abschnitt C, AUC .608) und den ILRV Gesamtwert (AUC .608) wies auf eine zufriedenstellende prädiktive Validität hin. Der Trend für die historischen Variablen (Abschnitt B) wies auf eine schlechtere Korrelation mit der Rückfälligkeit (AUC .596) hin, und die Variablen Sozialer Empfangsraum (Abschnitt D) waren für die Vorhersage dem Zufall kaum überlegen (AUC .565).

Diagramm 3 – ROC Analyse

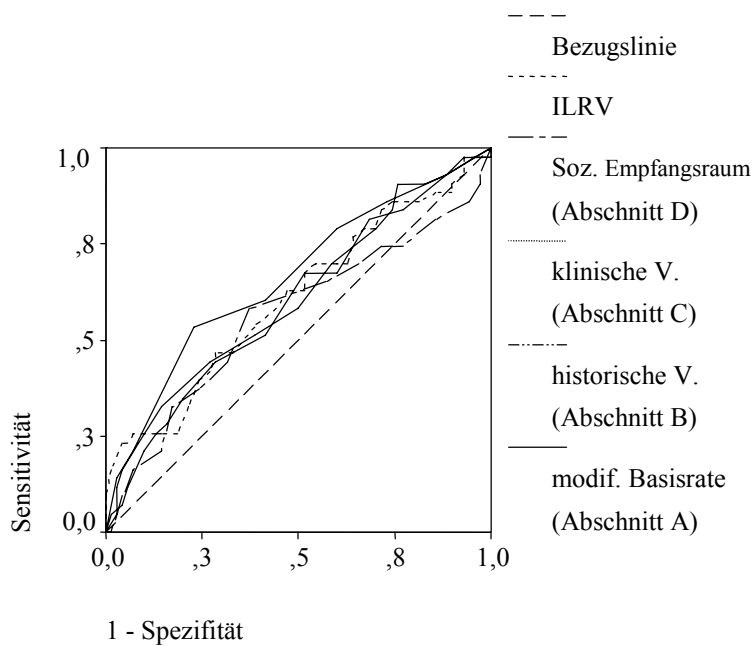


Tabelle 14

	AUC	Standardfehler	Signifikanz
<b>Modifizierte Basisrate (Abschnitt A)</b>	.661	.054	.004
<b>Historische Variablen (Abschnitt B)</b>	.596	.055	.086
<b>Klinische Variablen (Abschnitt C)</b>	.608	.056	.054
<b>Sozialer Empfangsraum (Abschnitt D)</b>	.565	.059	.245
<b>ILRV Gesamtsumme</b>	.608	.056	.055

#### 4.3.2. HCR-20

Die Gesamtsumme des HCR-20 hatte in der ROC-Analyse eine AUC von .618. Die klinischen und die Risikovariablen wiesen für den HCR-20 signifikant zufriedenstellende prädiktive Validitäten auf, wohingegen der Trend für die historischen Variablen auf eine geringere prädiktive Validität hinwies.

Tabelle 15

	AUC	Standardfehler	Signifikanz
<b>H1-10</b>	.601	.055	.072
<b>C1-5</b>	.612	.054	.045
<b>R1-5</b>	.616	.057	.039
<b>Gesamtwert</b>	.618	.056	.035

### 4.3.3. PCL-R Gesamtwert und Kombinationen

In einem nächsten Schritt wurde die ROC-Analyse für den PCL-R Gesamtwert durchgeführt. Darüberhinaus wurde untersucht, ob bestimmte Kombinationen von Variablen verschiedener Instrumente bessere Ergebnisse erzielen würden als die Prognoseinstrumente und ihre ursprünglichen Unterteilungen. Für die Kombination wurde die Anzahl der Variablen der Kombinationspartner durch Division angepasst, um ein Übergewicht durch eine höhere Anzahl an Variablen zu vermeiden. Die Kombination mit dem PCL-R Wert verbesserte die AUC der historischen Variablen und der modifizierten Basisrate signifikant. Die höchste AUC wurde durch die Kombination modifizierte Basisrate und PCL-R Gesamtwert erzielt (.684) (Tabelle 16).

Diagramm 4 – ROC Analyse

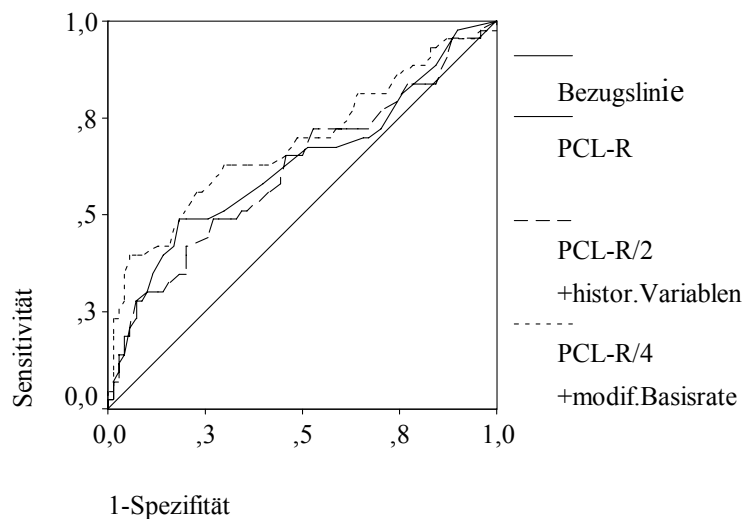


Tabelle 16

	AUC	Standardfehler	Signifikanz
<b>PCL-R Gesamtwert</b>	.630	.057	.021
<b>PCL-R/4 + modifizierte Basisrate</b>	.684	.055	.001
<b>PCL-R/2 + historische Variablen</b>	.616	.056	.040
<b>Modifizierte Basisrate + historische Variablen</b>	.641	.054	.012
<b>PCL-R + modifizierte Basisrate + historische Variablen</b>	.633	.056	.018

#### 4.4. VRAG

##### 4.4.1. ROC Analysen

Für den VRAG Summenwert und für seine Einzelvariablen wurden ROC Analysen durchgeführt. Der VRAG Summenwert erzielte die höchste AUC (.703) aller verwendeten Prognoseinstrumente (Diagramm 5, Tabelle 17), bei Betrachtung der gesamten Stichprobe und dem Zielkriterium „allgemeine Rückfälligkeit“. Bei Betrachtung der gesamten Stichprobe und der Ergebnisvariable „gewalttätige Rückfälligkeit“ ergab die AUC .702 ( $p = 0.001$ ) (Tabelle 17).

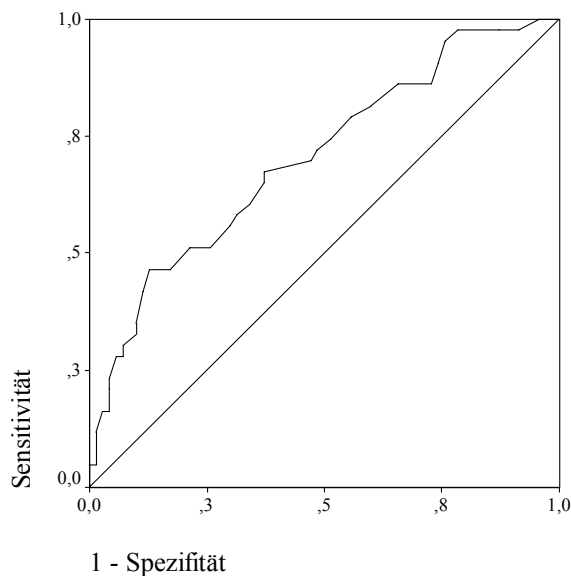
Die AUC erhöhte sich auf .763, wenn nur Probanden mit einem gewalttätigen Indexdelikt analysiert wurden und die Ergebnisvariable „allgemeine Rückfälligkeit“ war. Bei Betrachtung der Probanden mit einem gewalttätigen Indexdelikt und der Ergebnisvariable „gewalttätige Rückfälligkeit“ -am ehesten der Originaluntersuchung des VRAG entsprechend, die nur männliche Probanden einschloss, die alle mindestens einmal psychiatrisch behandelt worden waren- ergab sich eine AUC von .691 ( $p = .066$ ), das Ergebnis für diese Selektion war nicht signifikant.

Tabelle 17 – AUCs für unterschiedliche Probandengruppen und Ergebniskriterien

<b>Probanden</b>	<b>Ergebniskriterium</b>	<b>AUC</b>	<b>Standardfehler</b>	<b>Signifikanz</b>
<b>Gesamte Stichprobe</b>	<b>Rückfälligkeit allgemein</b>	.703	.051	.000
<b>Gesamte Stichprobe</b>	<b>Gewalttätige Rückfälligkeit</b>	.702	.067	.004
<b>Probanden mit gewalttätigen Indexdelikten</b>	<b>Rückfälligkeit allgemein</b>	.763	.068	.001
<b>Probanden mit gewalttätigen Indexdelikten</b>	<b>Gewalttätige Rückfälligkeit</b>	.691	.096	.066

Diagramm 5

ROC Analyse für „allgemeine Rückfälligkeit“ in der gesamten Stichprobe (AUC .703)



Wurden alle Probanden und alle Rückfälle berücksichtigt, so ergab sich die grösste AUC der Einzelvariablen für „Probleme in der Grundschule“ (AUC .635) und für den Punktwert nach Cormier-Lang für die kriminelle Vorgeschichte (.625) (Tabelle 18).

Tabelle 18

	<b>AUC</b>	<b>Standardfehler</b>	<b>Signifikanz</b>
<b>VRAG Summenwert</b>	.703	.051	.000
<b>1. PCL-R</b>	.575	.058	.182
<b>2. Probl. Grundschule</b>	.635	.056	.016
<b>3. Persönlichkeitsst.</b>	.554	.056	.337
<b>4. Alter</b>	.611	.053	.049
<b>5. lebte mit Eltern</b>	.615	.056	.040

<b>6. früher.</b>	.535	.056	.537
<b>Bewährungsversagen</b>			
<b>7. kriminelle Vorgesch.</b>	.625	.055	.026
<b>8. jemals verheiratet</b>	.589	.055	.113
<b>9. Schizophrenie</b>	.525	.056	.662
<b>10. Verletzung Opfer</b>	.614	.053	.043
<b>11. Alkoholmissbrauch</b>	.532	.057	.566
<b>12. weibliches Opfer</b>	.587	.054	.122

#### 4.4.2. VRAG Summenwerte

Der Mittelwert der VRAG-Summenwerte für die gesamte Stichprobe betrug -0.84, der Median -3.00 (Standardabweichung 11.013). Der kleinste absolute VRAG-Summenwert betrug -22, der höchste Wert 29. Für nicht-rückfällige Probanden betrug der Mittelwert -3.94, der Median -5.00 (Standardabweichung 9.413). Für rückfällige Probanden betrug der Mittelwert +4.21, der Median +4.00 (Standardabweichung 11.651) (Tabelle 19).

Tabelle 19

	<b>Mittelwert</b>	<b>Median</b>	<b>Standardabweichung</b>
<b>kein Rückfall (n=70)</b>	-3.94	-5.00	9.413
<b>Rückfall insgesamt (n=43)</b>	+4.21	+4.00	11.651
<b>gewalttätiger Rückfall (n=21)</b>	+5.95	+6.00	12.143
<b>nicht-gewalttätiger Rückfall (n=22)</b>	+2.55	+7.00	11.186
<b>Gesamte Stichprobe (n=113)</b>	-0.84	-3.00	11.013

Der VRAG-Summenwert korrelierte mit der Wahrscheinlichkeit des Probanden, rückfällig zu werden. Die Stichprobe wurde unterteilt in eine Gruppe mit VRAG-Werten oberhalb des Mittelwertes (-.84) und eine Gruppe mit Werten unterhalb des Mittelwertes. Probanden mit Werten oberhalb des Mittelwertes wurden häufiger rückfällig.

73% der Probanden mit VRAG-Summenwerten unter dem Mittelwert für die gesamte Stichprobe wurden nicht rückfällig, 27% der Probanden mit VRAG-Summenwerten unter dem Mittelwert wurden rückfällig, 9.5% mit einem gewalttätigen Rückfall und 17.5% mit einem nicht-gewalttätigen Rückfall. 52% der Probanden mit VRAG-Summenwerten oberhalb



des Mittelwerts wurden rückfällig, 30% mit einem gewalttätigen und 22% mit einem nicht-gewalttätigen Rückfall. In der Gruppe mit VRAG-Summenwerten über dem Mittelwert kam es signifikant häufiger zu einem Rückfall als in der Gruppe mit VRAG-Summenwerten unter dem Mittelwert (52% vs 27%,  $p = 0.011$ , vgl. Tabelle 20). Es ergab sich kein signifikanter Unterschied bei separater Betrachtung der männlichen Probanden oder bei Verwendung des Medians als Bezugsgrösse. Die Wahrscheinlichkeit eines gewalttätigen oder nicht-gewalttätigen Rückfalls kann somit anhand der VRAG-Summenwerte eingeschätzt werden (Tabelle 20).

Tabelle 20

	<b>VRAG Summenwert unter Mittelwert</b>	<b>VRAG Summenwert über Mittelwert</b>	Bezugsgrösse: Mittelwert für gesamte Stichprobe (- 0.84) Unterschied Rückfall vs kein Rückfall 2-seitig signifikant ( $p=0.011$ )
<b>Kein Rückfall (n=70)</b>	46 73.0%	24 48.0%	
<b>Rückfall insgesamt (n=43)</b>	17 27.0%	26 52.0%	
<b>Gewalttätiger Rückfall (n=21)</b>	6 9.5%	15 30%	Die 9 Risikogruppen des VRAG wiesen folgende Häufigkeitsverteilun
<b>Nicht-gewalttätiger Rückfall (n=22)</b>	11 17.5%	11 22.0%	
<b>Gesamte Stichprobe (n=113)</b>	63 100%	50 100%	

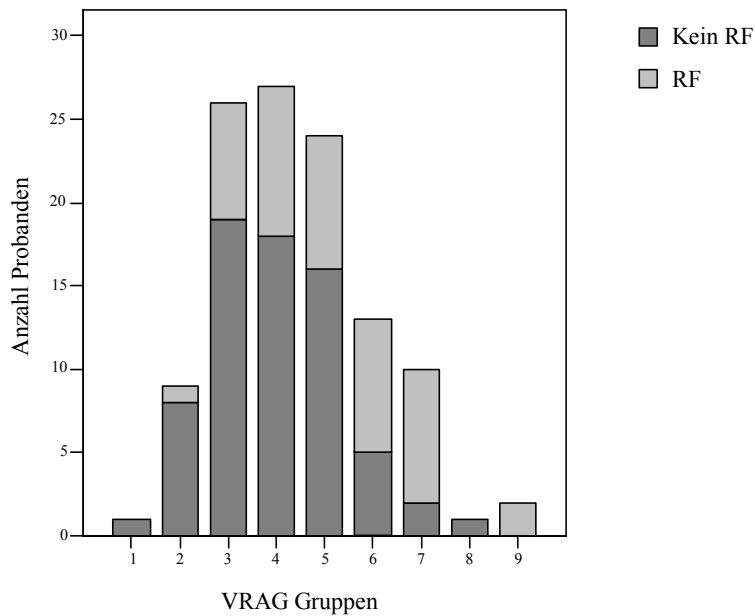
g auf:

Tabelle 21

<b>VRAG Kategorie</b>	<b>VRAG Punktwert</b>	<b>Anzahl Probanden</b>
<b>1</b>	<=-22	1
<b>2</b>	-21 bis -15	9
<b>3</b>	-14 bis -8	26
<b>4</b>	-7 bis -1	27
<b>5</b>	0 bis +6	24
<b>6</b>	+7 bis +13	13
<b>7</b>	+14 bis +20	10
<b>8</b>	+21 bis +27	1
<b>9</b>	>= +28	2

Die Verteilung der Rückfälligkeit innerhalb der 9 Risikokategorien ist in folgendem Diagramm dargestellt:

Diagramm 6



#### 4.4.3. Korrelation vorhergesagte Rückfallwahrscheinlichkeit und beobachtete Rückfallrate

Die Autoren des VRAG [43] geben für jede der 9 Gruppen eine empirisch ermittelte Wahrscheinlichkeit an, mit der ein Proband innerhalb von 7 oder von 10 Jahren gewalttätig rückfällig wird. Dieser vorhergesagte Wert wurde mit der tatsächlichen Rückfallrate unserer Stichprobe in den 9 Summenwertkategorien verglichen. Für 92 Probanden der Stichprobe endete die time-at-risk innerhalb von 7 Jahren nach Begutachtung. Vorhergesagte Wahrscheinlichkeit und beobachtete Rückfälle der einzelnen Gruppen korrelierten mit einem Koeffizienten nach Pearson  $r = .941$  (Signifikanz  $.005^{**}$ ).

Tabelle 22

<b>VRAG Kategorie</b>	<b>VRAG Punktwert</b>	<b>Anzahl Probanden</b>	<b>vorhergesagte p</b>	<b>beobachtete p</b>
<b>1</b>	<= -22	1	0	(0)*
<b>2</b>	-21 bis -15	9	.08	.11
<b>3</b>	-14 bis -8	26	.12	.27
<b>4</b>	-7 bis -1	27	.17	.30
<b>5</b>	0 bis +6	24	.35	.33
<b>6</b>	+7 bis +13	13	.44	.61
<b>7</b>	+14 bis +20	10	.55	.80
<b>8</b>	+21 bis +27	1	.76	(0)*
<b>9</b>	>= +28	2	1.00	(1.00)*

\*\* Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant

( )\* die Anzahl der Probanden in der jeweiligen Risikogruppe reicht für eine Wahrscheinlichkeitsberechnung nicht aus

#### 4.4.4. Kaplan-Meier Überlebensanalyse

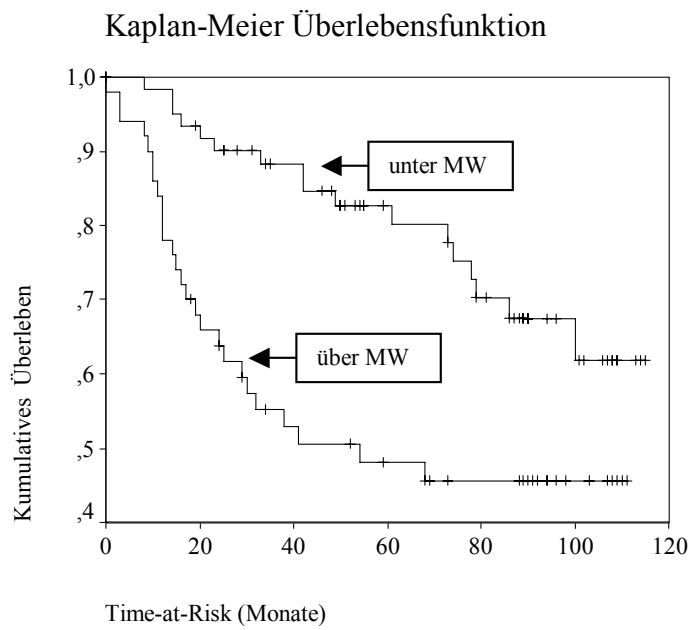
Die Überlebensanalyse nach Kaplan-Meier für Probanden mit VRAG-Summenwerten über und unter dem Mittelwert für die gesamte Stichprobe veranschaulicht, dass Probanden mit höheren VRAG Summenwerten häufiger und schneller rückfällig wurden. Nach einer time-at-risk von 5 Jahren beispielsweise sind erst ca. 20% der Probanden mit VRAG-Summenwerten unter dem Mittelwert rückfällig geworden, hingegen bereits über 50% der Probanden mit VRAG Summenwerten über dem Mittelwert.

Das kumulative Überleben der 63 Probanden (55.8%) mit VRAG-Summenwerten unter dem Mittelwert für die gesamte Stichprobe (-0.84) ist in der oberen Kurve dargestellt. Der Mittelwert ihrer time-at-risk beträgt 65.44 Monate (Standardabweichung 33.27).

Das kumulative Überleben der 50 Probanden (44.2%) mit VRAG-Summenwerten über dem Mittelwert für die gesamte Stichprobe (-0.84) ist in der unteren Kurve dargestellt. Der Mittelwert ihrer time-at-risk beträgt 48.76 (Standardabweichung 37.58).

Der Log-Rank beträgt 9.55 ( $p=0.002$ ), der Breslow-Index 14.02 ( $p=0.002$ ) und der Tarone-Ware Index 12.26 ( $p=0.005$ ).

Diagramm 7



+ → time-at-risk endet ohne Rückfall

## 5. Diskussion

### 5.1. Stichprobe

Im Rahmen des Münchner Projekts zur Rückfallprognose wurden die prädiktiven Validitäten der Prognoseinstrumente ILRV, HCR-20, PCL-R und VRAG untersucht. In einer Voruntersuchung [71] war unter anderem die prognostische Überlegenheit der modifizierten Basisrate für Rückfälligkeit, des PCL-R Gesamtwertes und der historischen Variablen festgestellt worden. Die Ergebnisse dieser Voruntersuchung wurde an der vorliegenden Stichprobe überprüft. Darüberhinaus wurde der VRAG [43] erstmals an einer deutschen Stichprobe validiert und seine Anwendung getestet. Hierzu wurde das englische Originalmanual ins Deutsche übersetzt.

Die Stichprobe bestand aus 136 Straftätern, die in den Jahren 1994 und 1995 in der forensischen Abteilung der Klinik für Psychiatrie und Psychotherapie der Universität München auf ihre strafrechtliche Verantwortlichkeit hin begutachtet worden waren. Ihr Risikoprofil wurde 9 Jahre nach Begutachtung entsprechend der Prognoseinstrumente anhand von Aktenmaterial, Gutachten und klinischer Dokumentation retrospektiv erhoben. Die Validität dieser Methode ist durch frühere Untersuchungen belegt [45;88;89]. Die inter-rater Reliabilitäts-Koeffizienten für die Prognoseinstrumente betrug in dieser Studie .89 für den HCR-20, .92 für die PCL-R und .93 für die ILRV.

Das Ergebniskriterium der Arbeit war Rückfälligkeit allgemein sowie eine Aufteilung in gewalttätige und nicht-gewalttätige Rückfälligkeit. Das Ergebniskriterium wurde für 113 Probanden anhand ihrer Eintragungen in das Bundeszentralregister nach Begutachtung ermittelt. Die time-at-risk der Probanden begann mit der Möglichkeit, rückfällig zu werden, d.h. vom Zeitpunkt der Entlassung in Freiheit bis zum ersten Rückfall bzw. bis zum Ende des Beobachtungszeitraums, zum Zeitpunkt der Anforderung der Bundeszentralregisterauszüge im März/April 2003.

Zunächst wurde geprüft, in wie weit die vorliegende Stichprobe mit der anderer Studien vergleichbar ist. Vielen Studien über kriminelle Rückfälligkeit liegen selektierte Stichproben mit Probanden nach Entlassung aus Haftanstalten oder aus forensischen Krankenhäusern zugrunde [3;5;12;14;19;31;32;42;53;56;57;67;73;104]. In den meisten Fällen war die Entlassung abhängig von einer günstigen, d.h. negativen Rückfallprognose. Die Ergebnisse dieser Studien werden durch die selektive Betrachtung falsch-negativer

Vorhersagen verzerrt, d.h. durch Prognosen für Straftäter, die trotz negativer Rückfallprognose ein erneutes Delikt begehen. Die Anzahl der falsch-positiven Vorhersagen hingegen bleibt unbekannt, d.h. die Anzahl der Prognosen für Probanden mit einer schlechten Rückfallprognose, die nicht in Freiheit entlassen werden, obwohl sie nicht rückfällig geworden wären. Um diese Verzerrung zu vermeiden, wurde in der vorliegenden Stichprobe die Risikoeinschätzung bei Probanden vorgenommen, die zu ihrer Schuldfähigkeit begutachtet wurden. Ihre Zukunft hing damals nicht von der Vorhersage künftiger Kriminalität ab, ihre Rückfallprognose hatte keinen Einfluss auf ihren weiteren Werdegang. Die Stichprobe war nicht selektiert.

43 der 113 Probanden wurden rückfällig. Das entspricht einer Rate für Rückfälligkeit von 38.1%. Basisraten für Rückfälligkeit werden im allgemeinen um diesen Wert angegeben, so z.B. in der Statistik des deutschen Bundeskriminalamts (erster periodischer Sicherheitsbericht) [7]. Bei der Entwicklung des VRAG lag die Basisrate für Rückfälligkeit bei 31% [74], in der Voruntersuchung des Münchner Projekts zur Rückfallprognose bei fast 39% [71;89].

Deliktspezifische Basisraten für Rückfälligkeit sind schwer zu ermitteln. Dies ist möglicherweise ein Grund, warum sie in den meisten Prognoseinstrumenten nicht berücksichtigt werden, auch wenn ihr prognostischer Wert in zahlreichen Studien belegt wurde [22;24;27]. Bislang bestanden nur grobe Angaben für deliktspezifische Basisraten für Rückfälligkeit [64], genauere Basisraten wurden erst vor kurzem zusammengefasst, sie lagen bei der Erstellung der Arbeit noch nicht vor [33]. Die ILRV ist das einzige Prognoseinstrument, das Basisraten in der Risikoevaluation berücksichtigt.

Zwischen psychischer Erkrankung und Rückfälligkeit konnte ein mit der Voruntersuchung [71] und anderen Studien [6;51;77;78] übereinstimmender Zusammenhang beobachtet werden. Überdurchschnittlich viele Rückfälle wurden in der Diagnosegruppe F1 „Psychische und Verhaltensstörungen durch psychotrope Substanzen“ und F6 „Persönlichkeits- und Verhaltensstörung“ beobachtet: 46.2% der Probanden der Diagnosegruppe F1 und 57.1% der Gruppe F6 wurden rückfällig. Ebenfalls häufig rückfällig wurden Probanden der Gruppen F0 „organische, einschliesslich symptomatische psychische Störungen“ und F4 „Neurotische, Belastungs- und somatoforme Störungen“ (je 66.7%), wobei hier die geringe Anzahl der Probanden dieser Gruppen zu berücksichtigen ist (6 Probanden mit F0, 3 Probanden mit F4). Je ein Viertel (25%) aller Probanden mit F3

„affektive Störungen“ und der Probanden ohne psychiatrischen Befund wurden rückfällig. Am wenigsten häufig wurden die Probanden der Diagnosegruppe F2 „Schizophrenie, schizotype und wahnhaftige Störungen“ rückfällig, insgesamt 12.5%. Das Rückfallrisiko dieser Gruppe war dann leicht erhöht, wenn Rückfälle im Zustand der Schuldunfähigkeit einbezogen wurden.

Auch die Korrelation des Indexdelikts mit der Rückfälligkeit stimmte mit früheren Untersuchungen überein: 76.5% der Probanden mit dem Indexdelikt „Körperverletzung“ wurden rückfällig. Am seltensten rückfällig wurden Probanden mit dem Indexdelikt „Mord und Totschlag“, nur 4.5%. Dieser Umstand kann u.a. dadurch erklärt werden, dass diese Probanden lange Haftzeiten aufweisen, auf Bewährung entlassen werden und Kontroll- und Sicherheitsvorkehrungen unterworfen sind und so kaum die Möglichkeit haben, rückfällig zu werden. Die gewalttätigen Rückfälle traten am häufigsten in den Deliktgruppen Körperverletzung auf (35.3%), was auf eine bleibende Gewaltbereitschaft dieser Deliktgruppe hinweist. Auch Probanden mit nicht-gewalttätigen Indexdelikten weisen ein erhöhtes Gewaltpotential für die Zukunft auf: 27.8% der Probanden der Deliktgruppen Diebstahl wurden gewalttätig rückfällig.

In der vorliegenden Stichprobe unterschied sich die time-at-risk für jeden Probanden und das Ergebniskriterium war allgemeine Rückfälligkeit einschliesslich gewalttätiger und nicht-gewalttätiger Rückfälle, nicht aber die spezifische Rückfälligkeit mit einem dem Indexdelikt vergleichbaren Delikt. So wurde keiner der Mörder und Totschläger mit einer erneuten Tötung und keiner der Sexualdelinquenten mit einem erneuten Sexualdelikt rückfällig. Diese Umstände sind bei einem Vergleich mit anderen Studien zu berücksichtigen.

## 5.2. Vergleich mit Voruntersuchung an einer anderen Stichprobe

Zunächst wurden die Ergebnisse einer Voruntersuchung [71;89] an der neuen Stichprobe getestet. Es wurden ROC-Analysen für die Gesamtwerte der ILRV, des HCR-20, der PCL-R und ihrer Untergruppen durchgeführt. Separat wurden historische, klinische und Zukunftsvariablen auf ihre prädiktive Validität hin untersucht, um die Untergruppen oder die Kombination von Untergruppen zu ermitteln, die am besten mit der Ergebnisvariable Rückfälligkeit korrelieren.

### 5.2.1. Ergebnisse der Voruntersuchung

In der Voruntersuchung war die prädiktive Validität dieser Prognoseinstrumente explorativ untersucht worden anhand einer Stichprobe von 185 Probanden, die in den Jahren 1992 und 1993 auf ihre Schuldfähigkeit hin begutachtet worden waren. Die ILRV, der HCR-20 und die PCL-R waren auf Gutachten basierend ausgewertet worden und das Risikoprofil der Probanden mit ihrer tatsächlichen Rückfallrate korreliert worden. Das Ergebniskriterium Rückfälligkeit war im Dezember 2001 anhand der Eintragungen in das Bundeszentralregister ermittelt worden. Die Voruntersuchung ergab folgende Ergebnisse:

Signifikante Korrelationen mit der Rückfälligkeit ergaben sich

1. für die historischen Variablen (Abschnitt B der ILRV und H1 bis H10 des HCR-20, AUC der Voruntersuchung .603)
2. den PCL-R Gesamtwert (AUC der Voruntersuchung .768) und die „modifizierte Basisrate“ (Abschnitt A der ILRV, AUC der Voruntersuchung .605)

Die beste Korrelation mit der Rückfälligkeit wurde

4. durch die Kombination des PCL-R Gesamtwerts mit der modifizierten Basisrate erzielt (AUC der Voruntersuchung .698).

Keine signifikante Korrelation ergab sich für

5. die klinischen Variablen (Abschnitt C der ILRV, AUC der Voruntersuchung .487) und für die Variablen betreffend der sozialen Situation der Probanden nach Entlassung (Abschnitt D der ILRV, AUC der Voruntersuchung .544).

### 5.2.2. Vergleich der Ergebnisse der Voruntersuchung mit den Ergebnissen der vorgelegten Studie

Zu 1:

In der vorliegenden Studie lag die AUC der historischen Variablen der ILRV (.596) tendenziell -im Gegensatz zu den Voruntersuchungen- unter der AUC für die modifizierte Basisrate der ILRV (.661) und für die klinischen Variablen des HCR-20 (.612), wobei die AUCs der historischen und der klinischen Variablen der ILRV und die AUC der historischen Variablen des HCR-20 nicht signifikant waren. Sie sind daher nur als Trend mit der



Voruntersuchung vergleichbar. Die im Trend niedrigere prädiktive Validität der historischen Variablen in der vorliegenden Stichprobe kann durch den unterschiedlichen Aufbau der beiden Stichproben erklärt werden. Sie glichen sich zwar grösstenteils in der Verteilung von Geschlecht, Alter, Diagnosen und der Basisrate für Rückfälligkeit (38.1% bzw. fast 39%), Unterschiede bestanden jedoch in der Verteilung der Indexdelikte, die zur Begutachtung geführt haben und in der Verteilung der Rückfälle. In der vorliegenden Stichprobe wurden tendenziell leichtere Indexdelikte begutachtet: beispielsweise waren in der vorliegenden Studie 15.9% der Indexdelikte Diebstahl, gegenüber nur 5% in der Voruntersuchung. Dementsprechend sind nur 4 Probanden der vorliegenden Stichprobe zum Zeitpunkt der Anforderung der Bundeszentralregisterauszüge noch in Haft, gegenüber 17 Probanden der früheren Studie. Die Verteilung der Rückfälle unterschied sich ebenfalls: in der vorliegenden Stichprobe kamen gewalttätige Rückfälle (bei 16.8% der Probanden) häufiger vor als in der Voruntersuchung (5.8% der Probanden). Probanden mit schwerwiegenderen Indexdelikten haben längere Haftstrafen und damit bei vorgegebener Zeitbegrenzung nach Begutachtung seltener die Möglichkeit, rückfällig zu werden. Delinquenten mit weniger gravierenden Indexdelikten hingegen haben länger die Möglichkeit, Delikte, auch schwerwiegendere Delikte, zu begehen.

Statische, d.h. historische Risikofaktoren sind demnach möglicherweise für Probanden, die schwerwiegendere Indexdelikte verübt haben, in der Risikoevaluation relevanter. Dies erklärt auch, dass der VRAG in seiner Originalstichprobe eine höhere AUC aufwies als in der vorliegenden: die Originalstichprobe für den VRAG bestand ausschliesslich aus gewalttätigen Probanden.

Die Unterschiede in der prädiktiven Validität dynamischer und historischer Risikofaktoren sind in der Literatur mehrfach beschrieben worden: dynamische Risikofaktoren erweisen sich z.B. nützlicher für Kurzzeitprognosen und für die Vorhersage institutioneller Gewalt [74] und bei selektiver Anwendung auf psychisch kranke Straftäter [67]. Vorliegende Daten deuten darauf hin, dass statische Risikofaktoren bei schwerwiegenderen Delinquenten mit Indexdelikten wie Körperverletzung oder Mord und Totschlag eine höhere Validität für die Rückfallprognose haben. Dies hat sich auch in Untersuchungen gezeigt, in welchen die ROC-Analysen für gewalttätige und nicht-gewalttätige Indexdelikte getrennt berechnet wurden [89].

zu 2 und 3:

Die ROC-Analyse der Untergruppen ergab in Übereinstimmung mit der Voruntersuchung, dass die modifizierte Basisrate (AUC .661) und der PCL-R Gesamtwert (AUC .630) die höchsten prädiktiven Validitäten aller Untergruppen aufwiesen. Diese Ergebnisse werden auch in anderen Studien beschrieben, in welchen Basisrate und PCL-R Wert als wichtige Prognosefaktoren identifiziert wurden [22;24;25;27;34;49;64;84]. Die PCL-R wird allgemein sogar als „state-of-the-art“ für Rückfallprognosen anerkannt [23] mit AUCs zwischen .61 und .70, bei längerem Beobachtungszeitraum [29] auch bis .75 [91] angegeben.

Zu 4.

Im nächsten Schritt wurde die Kombination von Untergruppen mit der höchsten prädiktiven Validität ermittelt. Auch hier stimmen die Ergebnisse mit den Voruntersuchungen überein: die AUC aller Untergruppen erhöhte sich, wenn sie mit dem PCL-R Gesamtwert kombiniert wurden. Die AUC der historischen Variablen stieg bei Kombination mit dem PCL-R Gesamtwert von .596 auf .616 an, die AUC der modifizierten Basisrate von .661 auf .684 und die Kombination aller drei Werte ergab eine AUC von .633. Die höchste Korrelation mit der Rückfälligkeit wurde -wie in der Voruntersuchung- durch die Kombination modifizierte Basisrate und PCL-R Wert erzielt. Wieder wurde der PCL-R Gesamtwert als wichtigster Parameter in der Kombination identifiziert [63;74].

Zu 5:

Der Abschnitt D der ILRV (sozialer Empfangsraum der Probanden) wies in Übereinstimmung mit der Voruntersuchung die niedrigste AUC (.565) aller Untergruppen auf (nicht signifikant). Dynamische Risikofaktoren sind für Langzeitprognosen einer unselektierten Stichprobe prognostisch weniger relevant. Sie sind eher nützlich für Kurzzeitprognosen, das Management von Zwischenfällen in forensischen Einrichtungen oder zur Einschätzung des Risikos während der Bewährungszeit [74;104]. Möglicherweise steigt die prädiktive Validität dynamischer Risikofaktoren bei selektiver Anwendung auf bestimmte Tätergruppen; die vorliegende Stichprobe war jedoch zu klein, um eine Analyse für verschiedene Tätergruppen durchzuführen.

Zusammenfassend wurden die Hypothesen der Voruntersuchungen in der vorliegenden Studie grösstenteils bestätigt: die höchsten AUC Werte ergaben sich in beiden Studien für die modifizierte Basisrate und den PCL-R Wert. In der Kombination verbesserte der PCL-R Gesamtwert die AUC aller Untergruppen signifikant, die höchste AUC wurde für

die Kombination modifizierte Basisrate und PCL-R Gesamtwert erzielt. In Übereinstimmung mit den Voruntersuchungen zeigte der Abschnitt D der ILRV (sozialer Empfangsraum der Probanden) im Trend eine schlechtere Korrelation mit der Rückfälligkeit. Lediglich bezüglich der klinischen und der historischen Variablen ergaben sich Unterschiede: während in der Voruntersuchung die historischen Variablen signifikant mit der Rückfälligkeit korrelierten, nicht aber die klinischen, war der Trend in der vorliegenden Studie umgekehrt.

### 5.3. VRAG

Der VRAG erzielte in der vorliegenden Stichprobe die höchste AUC aller Instrumente und Untergruppen (.703). Diese hohe prädiktive Validität des VRAG wurde bereits in zahlreichen internationalen Studien gezeigt [1;2;26;28;44;53;58;62;69;70;75;76;80;81] [56;85;92] (Übersicht siehe [42]) und konnte an der vorliegenden Stichprobe erstmals im deutschen Sprachraum reproduziert werden.

Der Zusammenhang zwischen VRAG-Summenwert und Rückfälligkeit bzw. gewalttätiger Rückfälligkeit der Probanden kann anhand des Mittelwerts beschrieben werden: Der Mittelwert für rückfällige Probanden (Mittelwert 4.21, Median 4.00, Standardabweichung 11.651) lag -in Übereinstimmung mit früheren Studien [42]- über dem Mittelwert für nicht-rückfällige Probanden (Mittelwert -3.94, Median -5.00, Standardabweichung 9.413). Der Mittelwert für Probanden mit einem gewalttätigen Rückfall (Mittelwert 5.95, Median 6.00, Standardabweichung 12.143) lag über dem Mittelwert für Probanden mit einem nicht-gewalttätigen Rückfall (Mittelwert 2.55, Median 7.00, Standardabweichung 11.186). Die Analyse der Gesamtstichprobe und die separate Analyse der männlichen Probanden ergab keinen statistisch signifikanten Unterschied, ebenso wenig die Differenzierung zwischen Median und Mittelwert.

Die Stichprobe wurde unterteilt in eine Gruppe mit VRAG-Werten oberhalb des Mittelwertes (-.84) und eine Gruppe mit Werten unterhalb des Mittelwertes. In der Gruppe mit VRAG-Summenwerten über dem Mittelwert kam es signifikant häufiger zu einem Rückfall als in der Gruppe mit VRAG Summenwerten unter dem Mittelwert (52% vs 27%,  $p=0.011$ ). Es ergab sich kein signifikanter Unterschied bei separater Betrachtung der männlichen Probanden oder bei Verwendung des Medians als Bezugsgröße.

Die Gruppen der Probanden mit niedrigeren VRAG-Summenwerten haben einen geringeren Anteil an den Rückfällen als die Gruppen mit höheren VRAG-Summenwerten. Die

Anwendbarkeit des VRAG als dimensionales Konstrukt wurde auch in dieser Untersuchung bestätigt.

Für 92 Probanden konnte eine konstante time-at-risk  $\leq 7$  Jahren bestimmt werden. Die in der Arbeit von Harris et al. [74] vorhergesagten Wahrscheinlichkeitswerte für die 9 VRAG-Summenwert-Kategorien korrelierten mit den beobachteten Rückfallraten mit einem Koeffizienten nach Pearson  $r = .941$  (Signifikanz  $.005$ ). Die beobachtete Rückfallrate der vorliegenden Stichprobe lag leicht über den vorausgesagten Wahrscheinlichkeitswerten, die in [74] aufgeführt sind. Der Korrelationskoeffizient nach Pearson ( $r = .941$ ) lag etwas unter dem von Harris 2002 ermittelten (nach Pearson  $r=.98$ ) [42]. Dies erklärt sich durch den Unterschied der Ergebniskriterien der Arbeiten: in der vorliegenden Arbeit war das Ergebniskriterium „Rückfälligkeit“, in den Arbeiten von Quinsey und Harris jedoch „gewalttätige Rückfälligkeit“ [42].

Der VRAG wies in der vorliegenden Studie mit einer AUC von  $.703$  -wie auch in anderen internationalen Studien [41;42;44;74]- eine höhere prädiktive Validität auf als die übrigen Prognoseinstrumente. Im Folgenden wurden mögliche Ursachen für die prädiktive Überlegenheit des VRAG untersucht.

Zunächst wurden ROC-Analysen der Einzelvariablen des VRAG durchgeführt, um zu ermitteln, ob eine spezifische Einzelvariable für die hohe prädiktive Validität des VRAG verantwortlich ist. Vier der fünf Variablen mit der höchsten AUC sind nicht in anderen gängigen Prognoseinstrumenten enthalten: das „Alter des Probanden zum Zeitpunkt des Indexdelikts“ (AUC  $.611$ ), „lebte der Proband mit beiden biologischen Elternteilen bis zum Alter von 16 Jahren“ (AUC  $.615$ ), die „Verletzung des Opfers“ (AUC  $.614$ ) und der „Punktwert für die kriminelle Vorgeschichte für Delikte ohne Gewaltanwendung vor dem Indexdelikt (aus dem Cormier-Lang System ermittelt)“ (AUC  $.625$ ). Lediglich „schlechte Anpassung in der Grundschule“ (AUC  $.635$ ) taucht in ähnlicher Weise im HCR-20, der ILRV und der PCL-R auf.

Der VRAG unterscheidet sich darüberhinaus in seiner Entstehungsweise von anderen Prognoseinstrumenten: er basiert als einziges Instrument auf einem empirisch ermittelten Datensatz von 618 Straftätern. Seine Variablen wurden retrospektiv mittels multipler Regressionsanalysen ermittelt. Die übrigen Prognoseinstrumente wurden mit anderen Methoden und Zielsetzungen entwickelt. Die Variablen des HCR-20 und der ILRV wurden

z.B. anhand umfassender Literaturquellen klinisch und konzeptionell zusammengestellt als Liste wichtiger Variablen für die Risikoevaluation in Klinik und Forschung. Die PCL-R wurde mittels Faktorenanalyse, basierend auf dem klinischen Konstrukt der „psychopathy“ entwickelt. Es liegt nahe, dass Unterschiede in den Entstehungsweisen der Instrumente einen Einfluss auf ihre prädiktiven Validitäten haben. Empirisch ermittelte, durch Regressionsanalysen selektierte Variablen erzielen die höchste prädiktive Validität, da sich die Methodik der Entwicklung und der Anwendung gleichen.

Darüberhinaus verfügt der VRAG über ein ausgefeilteres Bewertungssystem als andere Prognoseinstrumente: die Bewertung der Einzelvariablen erfolgt in Abhängigkeit von der Basisrate für gewalttätige Rückfälle der Gesamtstichprobe der Originalpopulation (31%). Die Einzelvariablen werden nicht, wie in den anderen Instrumenten, in 3-er Schritten bewertet, sondern können Werte zwischen -5 und +12 annehmen. Die Gewichtung der Einzelvariablen erfolgte in der Originalstichprobe empirisch anhand des Systems nach Nuffield [74]. Ein Punkt entsprach einer Basisrate für Rückfälligkeit für die Probanden, auf die diese spezifische Variable zutrifft, die sich um 5% von der Basisrate der Rückfälligkeit der Originalstichprobe unterschied: ein Unterschied von mindestens 5% von der gewalttätigen Rückfallrate der Gesamtstichprobe (31%) führte zu der Addition oder Subtraktion von je einem Punkt. Verheiratete Probanden beispielsweise hatten eine Rückfallsrate für gewalttätiges Verhalten von 21% - d.h. zwei volle 5% Intervalle unter der Basisrate. Unverheiratete Probanden hatten eine Rückfallsrate von 38%. Somit werden verheiratete Probanden mit einem Punktwert von -2 bewertet, unverheiratete Probanden mit einem Punktwert von +1. Die Bewertung der kriminellen Vorgeschichte für nicht-gewalttätige Delikte (AUC .625) wird anhand eines nochmals differenzierteren Bewertungssystems vorgenommen, des Cormier-Lang Crime Index [74]. Das empirisch ermittelte Gewichtungssystem des VRAG ermöglicht eine differenziertere Bewertung der Einzelitems. Die Unterteilung in 3-er Schritte der übrigen Prognoseinstrumente ist gröber, nicht empirisch ermittelt und führt möglicherweise zu einem Informationsverlust. Die Punktwerte dieser Instrumente korrelieren etwas schlechter mit der Rückfälligkeit als der VRAG-Gesamtwert.

Im Gegensatz zu anderen Instrumenten korreliert die Einzelvariable „Schizophrenie“ im VRAG negativ mit der Rückfälligkeit der Probanden. Ob Schizophrenie ein protektiver Faktor oder ein Risikofaktor für zukünftige Straftaten ist, bleibt bis heute umstritten (Überblick [88]). Der Trend der ROC-Analyse für diese Einzelvariable ergab jedoch keine

überdurchschnittlich hohe AUC (.525), so dass die hohe prädiktive Validität des VRAG nicht auf diese negative Korrelation zurückgeführt werden kann.

Die prädiktive Überlegenheit des VRAG beruht somit auf der Kombination seiner 12 Variablen und des abgestimmten Bewertungssystems, da keine der Einzelvariablen für sich einen höheren Wert aufweist als die Einzelvariablen anderer Instrumente. Die höchste AUC war .635 (schlechte Anpassung in der Grundschule), wohingegen im HCR-20 die höchste AUC der Einzelvariablen sogar .656 (C1 Mangel an Einsicht) und im ILRV .649 (A1 statistische Basisrate der Rückfälligkeit) betrug. Auch die Anzahl der Einzelvariablen mit hohen AUC-Werten im VRAG übersteigt nicht die anderer Instrumente. Die prädiktive Überlegenheit des VRAG-Summenwerts basiert nicht auf einer oder mehreren überlegenen Einzelvariablen, sondern resultiert aus der Kombination seiner Variablen und dem differenzierteren Bewertungssystem, das die Variablen in Abhängigkeit von der Basisrate für Rückfälligkeit gewichtet.

Folgender Umstand spricht ebenfalls für die Relevanz des ausgefeilten Bewertungssystems: die Variablen des VRAG bestehen ausschliesslich aus statischen Risikofaktoren. Diese Tatsache steht im Gegensatz zu der niedrigen prädiktiven Validität der historischen Variablen in der ILRV und der HCR-20 in dieser Studie, wobei beide Ergebnisse nicht signifikant waren. Analysiert man dennoch den Trend der schlechten Korrelation von historischen Variablen mit der Rückfälligkeit, muss die hohe AUC des VRAG an der Kombination seiner Variablen, sowie insbesondere an seinem Bewertungssystem liegen. Möglicherweise ist die Diskussion über die Überlegenheit von klinischen versus historischen Variablen für die prädiktive Validität weniger relevant als die Bewertung und Gewichtung der Variablen. Endgültige Aussagen lassen sich jedoch aufgrund der mangelnden Signifikanzen der historischen Variablen nicht treffen. Die klinischen Variablen erzielten in der vorliegenden Studie eine zufriedenstellende prädiktive Validität, so dass sie in der Prognoseermittlung zumindest mitberücksichtigt werden sollten.

Zur prädiktiven Überlegenheit des VRAG trägt auch die Tatsache bei, dass der VRAG den PCL-R Gesamtwert als am stärksten gewichteten Parameter inkorporiert. Der PCL-R Gesamtwert ist mit einer AUC von .630 nach dem Gesamtwert des VRAG und der modifizierten Basisrate der prognostisch wichtigste Parameter aller Prognoseinstrumente der untersuchten Stichprobe. Die Ergebnisse der ROC-Analysen aller Untergruppen der ILRV verbessern sich durch Kombination mit dem PCL-R Gesamtwert signifikant. Der PCL-R Wert

wird in zahlreichen Studien als wichtiger Prognoseparameter beschrieben [15;29;30;30;37;39;45;54;62;80;82;89;91]. Es liegt nahe, dass der PCL-R Gesamtwert als am stärksten gewichteter Parameter auch die prädiktive Validität des VRAG signifikant bestimmt. Um diese Hypothese zu testen, wurde eine ROC-Analyse des VRAG unter Ausschluss des PCL-R Gesamtwerts durchgeführt. Der Summenwert der 11 übrigen Variablen ergab jedoch eine AUC von .720, die sogar über der bereits beschriebenen AUC des VRAG Gesamtwerts von .703 liegt; diese Hypothese wurde somit nicht bestätigt. Dieses Ergebnis bedeutet jedoch nicht, dass der VRAG ohne die Variable „PCL-R Gesamtwert“ eine höhere prädiktive Validität aufweisen würde. Auch hier ist der Unterschied von Entwicklung und Bewertungssystem des VRAG zu anderen Instrumenten relevant: der PCL-R Wert wurde anhand des Systems nach Nuffield gewichtet und fügt sich in das numerische Bewertungssystem des VRAG ein, das die prozentualen Abweichungen der Probanden von der Basisrate für Rückfälligkeit von 31% widerspiegelt. Die exakt definierten Schwellenwerte für die PCL-Gesamtwerte (zwischen -5 und +12) führen somit dazu, dass eine einfache Subtraktion der Variable „PCL-R Gesamtwert“ vom VRAG-Summenwert zu einer Verschiebung des abgestimmten Punktesystems führt. Die höhere AUC der Summe der 11 übrigen Variablen entspricht daher nicht einer besseren prädiktiven Validität, sondern ergibt sich aus der Verzerrung des Ergebnisses durch die Störung des ausgearbeiteten Bewertungssystems. Andere Autoren fanden keinen Unterschied in der prädiktiven Validität, wenn der PCL-R Wert ausgeschlossen wurde; sie führten diesen Umstand darauf zurück, dass die Ergebniskriterien sich jeweils unterschieden: gewalttätige Rückfälligkeit oder Rückfälligkeit allgemein [43;53;55].

Auch in der vorliegenden Studie unterschied sich das Ergebniskriterium von dem der Originalstichprobe des VRAG: beurteilt wurde Rückfälligkeit allgemein, nicht gewalttätige Rückfälligkeit wie in der Originalstichprobe. Die AUC des VRAG für Probanden mit gewalttätigen Indexdelikten -wie in der Originalstichprobe- betrug in der vorliegenden Stichprobe .763 (Signifikanz .001) und lag somit über der AUC für die gesamte Stichprobe (.703). Bei Definition des Ergebniskriteriums als gewalttätige Rückfälligkeit -wie in der Originalstichprobe- betrug die AUC .702 (Signifikanz .004), unterschied sich somit nur leicht von dem Ergebniskriterium Rückfälligkeit allgemein. Bei der Analyse von Probanden mit gewalttätigen Indexdelikten und dem Ergebniskriterium gewalttätige Rückfälligkeit (dies kommt der Originalstichprobe des VRAG am nächsten, in der ausschliesslich männliche Probanden berücksichtigt wurden, die alle mindestens einmal psychiatrisch behandelt worden

waren) wurde kein signifikantes Ergebnis erzielt, so dass nicht abschliessend beurteilt werden kann, ob die prädiktive Validität mit dieser Selektion noch grösser wäre.

#### 5.4. Ausblick

Insgesamt wurden in der vorliegenden Studie, wie auch in den Voruntersuchungen, etwas niedrigere AUC-Werte als in anderen Studien erzielt. Beispielsweise betrug die AUC für den VRAG in der vorliegenden Studie .703 im Vergleich zu .76 für die Originalpopulation [74] oder .80 in anderen Studien [42]. Dies ist möglicherweise auf unterschiedliche forensische Rahmenbedingungen in verschiedenen Studien und Ländern zurückzuführen. AUC Werte liegen höher, wenn die Sicherheit, mit der die Prognosen für die Probanden einer Stichprobe gestellt werden kann, möglichst gross ist. Dies kann dadurch erreicht werden, dass Probanden mit einer unsicheren Prognose aus der Stichprobe ausgeschlossen werden, so beispielsweise auch Probanden mit schwerwiegenden Indexdelikten, die aufgrund langer Haftzeiten nicht in die Stichprobe aufgenommen werden. Wenn „unsichere Prognosen“ in verschiedenen Ländern unterschiedlich behandelt werden, führt dies zu unterschiedlichen AUC Werten. Je mehr „unsichere Prognosen“ in eine Stichprobe aufgenommen werden, desto grösser ist die Wahrscheinlichkeit, dass das getestete Prognoseinstrument falsch prognostiziert, desto niedriger die AUC.

Um die Ergebnisse verschiedener Studien zu vergleichen ist essentiell, dass Stichproben, Methoden und Definitionen der Ergebnisvariable vergleichbar sind. So wurde der VRAG als Instrument zur Beurteilung von gewalttätiger Rückfälligkeit entwickelt, anhand einer Stichprobe von männlichen, gewalttätigen Verbrechern, die alle mindestens einmal psychiatrisch behandelt worden waren. In der Literatur wird mehrfach darauf hingewiesen [56;74], dass die Instrumente in verschiedenen Stichproben unterschiedliche Ergebnisse erzielen. Auf einheitliche Definitionen des Ergebniskriteriums sollte geachtet werden, ebenso wie auf vergleichbare Beobachtungszeiträume. So ist die prädiktive Validität der PCL-R für verschiedene Stichproben, beispielsweise für männliche und weibliche Probanden ebenso wenig einheitlich [82;96], wie die prädiktive Validität des VRAG [41;42;56;74]. Unsere Stichprobe unterscheidet sich von der Originalpopulation des VRAG dahingehend, dass nicht ausschliesslich männliche Probanden eingeschlossen waren, nicht alle mindestens einmal psychiatrisch behandelt worden waren, und nicht alle aufgrund eines schwerwiegenden Indexdelikts begutachtet worden waren. Bei spezifischer Anwendung des VRAG auf selektive Täterpopulationen mit hohen Basisraten für Rückfälligkeit,



beispielsweise bei Patienten aus einer Erziehungsanstalt oder einem psychiatrischen Krankenhaus, werden höhere AUC Werte erzielt [95]. Auch die AUC der vorliegenden Stichprobe steigt von .703 auf .763 an, wenn nur Probanden mit gewalttätigen Indexdelikten berücksichtigt werden. Das Ergebnis für gewalttätige Indexdelikte und gewalttätige Rückfälle, das den Stichproben der Originaluntersuchung des VRAG am nächsten kommt, war nicht signifikant.

In diesem Kontext ist auch zu beachten, dass bei der Risikobeurteilung die Differenzierung zwischen sehr niedrigen und sehr hohen Basisraten von Bedeutung ist [67]. Die Gesetzmässigkeiten der Basisraten (vgl. Einleitung) lassen vermuten, dass sie für sehr hohe oder sehr niedrige Basisraten aussagekräftiger sind. Dies wiederum entspricht dem bereits erwähnten Umstand, dass Täter, die für schwerwiegende Delikte mit niedrigen Basisraten verurteilt wurden, nicht die Möglichkeit haben, rückfällig zu werden. In einer Studie von Loza et al. [56] deuten die Ergebnisse darauf hin, dass der VRAG gewalttätiges Verhalten in 12.9% der Fälle zu häufig vorhersagt. Bei einer Unterteilung der Stichprobe in drei Risikokategorien (niedrig, mittel und hoch) beobachtet Loza, dass Probanden der mittleren und hohen Risikokategorien weniger häufig rückfällig wurden als dies von den Autoren des VRAG vorhergesagt wurde. Dies entspricht der verminderten Möglichkeit, rückfällig zu werden.

Eine weitere Voraussetzung für eine reproduzierbar hohe prädiktive Validität des VRAG ist, dass die time-at-risk Periode in etwa der Periode gleicht, die zur Etablierung des VRAG verwendet wurde (7 oder 10 Jahre). Werden längere time-at-risk Perioden gewählt, ist nicht auszuschliessen, dass die Basisrate für Rückfälligkeit steigt und mit ihr die prädiktive Validität. Für die praktische Anwendung des VRAG in der Klinik ändert diese durch längere Beobachtungszeiträume scheinbar verbesserte Validität jedoch nichts an den Konsequenzen, die aus der Risikoevaluation gezogen werden. In der vorliegenden Stichprobe wurde die time-at-risk für jeden Probanden individuell berechnet, der Mittelwert betrug 58.06 Monate.

Loza kommt zu dem Schluss, dass die Anwendung aktuarischer Methoden möglicherweise nicht vertretbar ist, wenn bekannt ist, dass die Grundlagen (Zeitraum und Basisrate) von den ursprünglich vorhandenen Grundlagen des empirisch ermittelten Instruments deutlich abweichen. Unsere Ergebnisse bestätigen diese Hypothese nur partiell, da trotz beträchtlicher Unterschiede in der Stichprobe und der time-at-risk Periode der VRAG als Prognoseinstrument die höchste prädiktive Validität aller verwendeten Instrumente

aufweist und beobachtete Rückfallraten gut mit den vorhergesagten Wahrscheinlichkeiten korrelieren. Unterschiede in der Stichprobe, Ergebnisvariablen und time-at-risk stehen der klinischen Anwendung des VRAG nicht entgegen. Dies entspricht auch den Ergebnissen einer Studie von Harris et al. von 2004 [41], in der der VRAG auch bei Anwendung auf weibliche und auf nicht-forensische Probanden eine hohe Validität aufweist. Die Unterschiede im Studienaufbau sind insbesondere bei wissenschaftlichen Fragestellungen relevant, im klinischen Alltag empfiehlt sich die Anwendung von Prognoseinstrumenten trotz Unterschieden der Probanden, Ergebniskriterien oder der time-at-risk Perioden.

Zusammenfassend waren in dieser Studie die modifizierte, deliktspezifische Basisrate und der PCL-R Wert -wie in der Voruntersuchung- die Faktoren mit der höchsten prognostischen Relevanz. Die höchste prädiktive Validität aller Prognoseinstrumente wurde auch in dieser deutschen Stichprobe für den VRAG ermittelt. Nach seiner Validierung an einer deutschen Stichprobe sollte der VRAG auch im deutschen Sprachraum standardisiert in die Erstellung von Rückfallprognosen aufgenommen werden. Ähnlich gute Ergebnisse können für den SORAG, dem Äquivalent des VRAG für Sexualstraftäter, erwartet werden. Seine Validierung im deutschen Sprachraum steht noch aus.

Basisraten, PCL-R und VRAG sind im klinischen Alltag essentielle Elemente für die Erstellung von Rückfallprognosen. Insbesondere die Basisrate wurde bislang in den meisten Studien und Prognoseinstrumenten nicht berücksichtigt, obwohl ihr prognostischer Wert mehrfach belegt wurde. Die hohe prädiktive Validität des VRAG steht in engem Zusammenhang mit seinem Bewertungssystem, das auf einem Konstrukt aus empirisch ermittelten Variablen basiert. Entwicklungsmethode und Bewertungssystem sind bestimmende Elemente für eine hohe prädiktive Validität eines Prognoseinstruments, ihre Bedeutung und Modifizierung stellen neue Herausforderungen an die zukünftige Prognoseforschung.

## 6. Anhang

### 6.1. Übersetzung des VRAG ins Deutsche

APPENDIX A  
Handbuch zur Einschätzung des Gewaltrisikos  
Violence Risk Appraisal Guide (VRAG)

Bewertung des VRAG

1. Lebte mit beiden biologischen Elternteilen bis zum Alter von 16 (Ausnahme: Tod eines Elternteils)  
Ja = -2, Nein = +3
  
2. schlechte Anpassung in der Grundschule  
keine Probleme = -1  
leichte (leichte Disziplinar- oder Anwesenheitsprobleme) = +2  
oder mittelschwere Probleme  
schwere Probleme (häufiges störendes Benehmen und/oder = +5  
schwere Anwesenheitsprobleme oder Benehmen, das eine  
Schulverweisung oder eine ernsthafte Suspendierung  
vom Unterricht zur Folge hatte)
  
3. Alkoholanamnese  
Für jeden der folgenden Punkte wird ein Punkt vergeben: Alkoholmissbrauch bei  
einem biologischen Elternteil, Alkoholprobleme als Teenager, Alkoholprobleme als  
Erwachsener, Alkoholeinfluss bei einem früheren Delikt, Indexdelikt unter  
Alkoholeinfluss  
0 = -1, 1 oder 2 = 0, 3 = +1, 4 oder 5 = +2
  
4. Familienstand  
Jemals verheiratet (oder Zusammenleben in einer eheähnlichen = -2  
Gemeinschaft für mindestens 6 Monate)  
Nie verheiratet = +1
  
5. Punktwert für die kriminelle Vorgeschichte für Delikte ohne Gewaltanwendung vor  
dem Indexdelikt (aus dem Cormier-Lang System ermittelt, wie im Folgenden  
beschrieben)  
  
Punktwert 0 = -2  
Punktwert 1 oder 2 = 0  
Punktwert 3 oder mehr = +3
  
6. Bewährungsversagen (einschliesslich Verstösse bei bedingter Entlassung oder  
während einer Bewährungszeit oder Widerruf der Bewährung, nicht Einhalten von  
Vorschriften und Auflagen, Verfallenlassen der Kaution zur Haftverschonung und  
erneute Verhaftung während der Bewährung)  
  
Nein = 0, Ja = +3

7. Alter zum Zeitpunkt des Indexdelikts (vollendete Jahre)
- $\geq 39 = -5$
  - 34-38 = -2
  - 28-33 = -1
  - 27 = 0
  - $\leq 26 = +2$
8. Verletzung des Opfers (beim Indexdelikt; die schwerwiegendste Verletzung wird bewertet)
- Tod = -2
  - Stationäre Behandlung erforderlich = 0
  - Ambulante Behandlung erforderlich = +1
  - Keine oder leichte Verletzung (oder kein Opfer) = +2
9. weibliche Opfer (beim Indexdelikt)
- Ja = -1, Nein (oder kein Opfer) = +1
10. erfüllt die DSM-III Kriterien für eine Persönlichkeitsstörung
- Nein = -2, Ja = +3
11. Erfüllt die DSM-III Kriterien für Schizophrenie
- Ja = -3, Nein = +1
12. Punktwert in der „Psychopathy Checklist“
- $\leq 4 = -5$
  - 5-9 = -3
  - 10-14 = -1
  - 15-24 = 0
  - 25-34 = +4
  - $\geq 35 = +12$

**Cormier-Lang Punktwerte für die kriminelle Vorgeschichte für Delikte ohne Gewaltanwendung (nach Akman und Normandeau, 1967)**

Raub (Bank, Geschäft)	7
Raub (Taschendiebstahl)	3
Brandstiftung (Kirche, Haus, Scheune)	5
Brandstiftung (Mülltonne)	1
Drohung mit einer Waffe	3
Drohung (Äusserung einer Drohung)	2
Schwerer Diebstahl* (einschliesslich Autodiebstahl und Besitz gestohlenen Eigentums)	5
Schwere Beschädigung von öffentlichem oder privatem Eigentum	5
Einbruch (Einbruchdiebstahl)	2
Leichter Diebstahl* (einschliesslich Besitz kleinerer gestohlener Güter)	1
Leichte Beschädigung öffentlichen oder privaten Eigentums(auch Aufsehererregen in der Öffentlichkeit)	1
Einbruch (einschliesslich Einbruch mit anderer Intention)	1

Betrug (Erpressung, Untreue)	5
Betrug (Scheckfälschung, Vortäuschen einer falschen Identität)	1
Besitz einer verbotenen oder im Gebrauch eingeschränkten Waffe	1
Zuhälterei oder Finanzierung des Lebensunterhalts durch Prostitution	1
Handel mit Betäubungsmitteln	1
Verkehrsgefährdendes Fahren, Fahren unter Beeinträchtigung von Alkohol oder Drogen (Trunkenheit am Steuer)	1
Behinderung eines Polizisten/Beamten (einschliesslich Widerstand gegen Verhaftung)	1
Störung der öffentlichen Ordnung	1
eine Verkleidung tragen mit der Absicht, ein Delikt zu begehen	1
unsittliche Entblössung	1

\*in etwa gleichbedeutend mit einfachem und schwerem Diebstahl, basierend auf dem Wert des gestohlenen Eigentums. 1997 lag die Grenze bei USD 1000,00.

Alle Verhaftungen für alle Anklagepunkte für Delikte ohne Gewaltanwendung (einschliesslich Jugend-Vorstrafen) werden bewertet und addiert. Wenn sich ein Strafgesetz vom kanadischen Strafgesetz unterscheidet, wird eine gewisse Beurteilungsfähigkeit vorausgesetzt, um die Delikte in den anderen Zuständigkeitsbereichen einzuschätzen. Beispielsweise erscheint „larcency“ (Diebstahl) nicht im kanadischen Strafgesetzbuch, aber ist in etwa gleichbedeutend wie „theft“ (Diebstahl). *Weitere Details der Bewertung siehe Appendix E.*

#### **Rückfallwahrscheinlichkeiten für Gewalttaten über 2 verschiedene Zeiträume als Funktion von 9 gleich grossen VRAG Gruppen**

<b>VRAG Kategorie</b>	<b>VRAG Punktwert</b>	<b>7 Jahre</b>	<b>10 Jahre</b>
<b>1</b>	<= -22	0	0.08
<b>2</b>	-21 bis -15	0.08	0.10
<b>3</b>	-14 bis -8	0.12	0.24
<b>4</b>	-7 bis -1	0.17	0.31
<b>5</b>	0 bis +6	0.35	0.48
<b>6</b>	+7 bis +13	0.44	0.58
<b>7</b>	+14 bis +20	0.55	0.64
<b>8</b>	+21 bis +27	0.76	0.82
<b>9</b>	>= +28	1.00	1.00

## 6.2. Übersetzung des SORAG ins Deutsche

### APPENDIX B

#### Handbuch zur Risikoeinschätzung bei Sexualtätern Sex Offender Risk Appraisal Guide (SORAG)

#### Bewertung des SORAG

1. Lebte mit beiden biologischen Elternteilen bis zum Alter von 16 (Ausnahme: Tod eines Elternteils)  
  
Ja = -2, Nein = +3
  
2. schlechte Anpassung in der Grundschule  
keine Probleme = -1  
leichte (leichte Disziplinar- oder Anwesenheitsprobleme) = +2  
oder mittelschwere Probleme  
schwere Probleme (häufiges störendes Benehmen und/oder = +5  
schwere Anwesenheitsprobleme oder Benehmen, das eine  
Schulverweisung oder eine ernsthafte Suspendierung  
vom Unterricht zur Folge hatten)
  
3. Alkoholanamnese  
Für jeden der folgenden Punkte wird ein Punkt vergeben: Alkoholmissbrauch bei  
einem biologischen Elternteil, Alkoholprobleme als Teenager, Alkoholprobleme als  
Erwachsener, Alkoholeinfluss bei einem früheren Delikt, Indexdelikt unter  
Alkoholeinfluss  
0 = -1, 1 oder 2 = 0, 3 = +1, 4 oder 5 = +2
  
4. Familienstand  
Jemals verheiratet (oder Zusammenleben in einer eheähnlichen = -2  
Gemeinschaft für mindestens 6 Monate)  
Nie verheiratet = +1
  
5. Punktwert für die kriminelle Vorgeschichte für Delikte ohne Gewaltanwendung vor  
dem Indexdelikt (aus dem Cormier-Lang System ermittelt, vgl. Appendix A)  
  
Punktwert 0 = -2  
Punktwert 1 oder 2 = 0  
Punktwert 3 oder mehr = +3
  
6. Punktwert für die kriminelle Vorgeschichte für gewalttätige Delikte vor dem  
Indexdelikt (aus dem Cormier-Lang System ermittelt, wie im Folgenden beschrieben).  
Alle gewalttätigen Sexualdelikte müssen aufgenommen werden.  
  
Punktwert 0 = -2  
Punktwert 2 = 0  
Punktwert 3 oder mehr = +6

7. Zahl der Vorverurteilungen für Sexualdelikte (betrifft alle Urteile, die aus vorhandener Dokumentation bzgl. Sexualdelikte vor dem Indexdelikt hervorgehen.) Gezählt werden alle sexuellen Delikte, einschliesslich beispielweise unsittliche Entblössung.  
 0 = -1  
 1 oder 2 = +1  
 >=3 = +5

8. Frühere Sexualstraftaten gegen Mädchen unter 14 Jahren (einschliesslich des Indexdelikts; wenn der Täter weniger als 5 Jahre älter war als das Opfer, wird immer mit +4 bewertet)

9. Bewährungsversagen (einschliesslich Verstössen bei bedingter Entlassung oder während einer Bewährungszeit oder Widerruf der Bewährung, nicht Einhalten von Vorschriften und Auflagen, Verfallenlassen der Kautions zur Haftverschonung und erneute Verhaftung während der Bewährung)

Nein = 0, Ja = +3

10. Alter zum Zeitpunkt des Indexdelikts (vollendete Jahre)

>= 39 = -5  
 34-38 = -2  
 28-33 = -1  
 28 = 0  
 <= 26 = +2

11. Erfüllt die DSM-III Kriterien für eine Persönlichkeitsstörung

Nein = -2, Ja = +3

12. Erfüllt die DSM-III Kriterien für Schizophrenie

Ja = -3, Nein = +1

Ergebnisse von phallometrischen Tests

Kein Test weist auf abweichende sexuelle Präferenzen hin = -1

Mindestens ein Test weist auf abweichende sexuelle Präferenzen hin = +1

13. Punktwert in der „Psychopathy Checklist“

<= 4 = -5  
 5-9 = -3  
 10-15 = -1  
 15-24 = 0  
 25-35 = +4  
 >= 35 = +12

### **Cormier-Lang Punktwerte für die kriminelle Vorgeschichte für gewalttätige Delikte**

Totschlag (Mord, Totschlag, grobe Fahrlässigkeit mit Todesfolge)	28
Versuchter Mord, beabsichtigte Körperverletzung	7
Menschenraub, Entführung und gewaltsame Freiheitsberaubung	6
Schwere Körperverletzung, Würgen, Verabreichen einer schädlichen Substanz	6
Angriff mit Körperverletzung	5

Angriff (oder gewaltsame Drohung) mit einer Waffe	3
Angriff (oder gewaltsame Drohung), Angriff auf einen Polizisten/Beamten	2
Schweres sexuelles Vergehen, sexuelles Vergehen mit Körperverletzung	15
Sexuelles Vergehen mit einer Waffe	
Sexuelles Vergehen, schwere Unzucht (vaginale, anale oder orale Penetration)	10
Sexuelles Vergehen (versuchte Vergewaltigung, Nötigung zur Unzucht)	6
Schwere Unzucht (Täter führt den Oralverkehr beim Opfer aus)	6
Sexuelles Vergehen (sexueller Eingriff, Einladung zu sexueller Berührung)	2
Bewaffneter Raubüberfall (Bank, Geschäft)	8
Gewalttätiger Raubüberfall	5
Bewaffneter Raubüberfall (keine Bank und kein Geschäft)	4

Alle Verhaftungen für alle Anklagepunkte von gewalttätigen Delikten (einschliesslich Jugendstrafen) werden bewertet und addiert. Dies ergibt einen Gesamtwert für die gewalttätige Vorgeschichte. Mit Ausnahme von versuchtem Mord, werden „versuchte“ Gewalttaten, z.B. versuchter bewaffneter Raubüberfall gleich bewertet wie die Gewalttaten selbst. Dokumentation mit Details zu den Gewalttaten können (und sollten, wann immer möglich) für die Bewertung herangezogen werden. Im allgemeinen werden z.B. bewaffneter Raubüberfall und gewalttätiger Raubüberfall als Gewalttaten bewertet, aber ein Raubüberfall an sich wird als nicht-gewalttätig bewertet. Wenn jedoch die Berichte der Untersuchungsbehörden darauf hinweisen, dass eine Verhaftung für einen Raubüberfall mit einem gewalttätigen Verhalten assoziiert war (z.B. das Opfer wurde verletzt), würde dieses Vergehen als gewalttätig kodiert werden. *Für weitere Details der Bewertung vgl. Bemerkungen in Appendix E über das vollständige Cormier-Lang System.*



## 7. Literaturverzeichnis

1. Barbaree, H., M. Seto, C. Langton, und E. Peacock. 2001. Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior* 28: 490-521.
2. Belanger, H. und C. Earls. 1996. Sex offender recidivism prediction. *Forum on Correctional Research* 8: 22-24.
3. Belfrage, H. 1998. Making risk predictions without an instrument. Three years' experience of the new Swedish law on mentally disordered offenders. *International Journal of Law and Psychiatry* 21: 59-64.
4. Belfrage, H. und K. S. Douglas. 2002. Treatment effects on forensic psychiatric patients measured with the HCR-20 Violence Risk Assessment Scheme. *International Journal of Forensic Mental Health* 1: 25-36.
5. Belfrage, H., G. Fransson, und S. Strand. 2000. Prediction of violence using the HCR-20: a prospective study in two maximum security correctional institutions. *Journal of Forensic Psychiatry* 11: 167-175.
6. Bieber, S. L., R. A. Pasewark, K. Bosten, und H. J. Steadman. 1988. Predicting criminal recidivism of insanity acquittees. *International Journal of Law and Psychiatry* 11: 105-112.
7. Bundesministerium des Inneren. 2001. Erster periodischer Sicherheitsbericht. Seitz, N., Kiel, R., von Kiedrowski, U., Lehmann, C., and Blath, R. Elz, Druckerei Ruster und Partner (Bericht).
8. Checkley, H. 1941. *The mask of sanity*. Mosby, St. Louis, MO.
9. Claix, A. und T. H. Pham. 2004. Evaluation of the HCR-20 Violence Risk Assessment Scheme in a Belgian forensic population. *Encephale* 30: 447-453.
10. Cooke, D. J., D. S. Kosson, und C. Michie. 2001. Psychopathy and ethnicity: structural, item, and test generalizability of the Psychopathy Checklist-Revised (PCL-R) in Caucasian and African American participants. *Psychological Assessment* 13: 531-542.
11. Cooke, D. J. und C. Michie. 2001. Refining the construct of psychopathy: towards a hierarchical model. *Psychological Assessment* 13: 171-188.
12. Davison, S., E. Jamieson, und P. J. Taylor. 1999. Route of discharge for special (high-security) hospital patients with personality disorder - relationship with reconviction. *British Journal of Psychiatry* 175: 224-227.
13. Dietz, P. E. 1985. Hypothetical criteria for the prediction of individual criminality. In: Webster, C. D., M. H. Ben-Aron, and S. J. Hucker (Hrsg.) *Dangerousness*. Cambridge University Press, Cambridge, London: 87-102.
14. Dimmek, P. 1999. Zur Rückfälligkeit von entlassenen Patienten des Massregelvollzugs gemäss § 63 StGB. In: Egg, R. (Hrsg.) *Sexueller Missbrauch von Kindern*. Kriminologie und Praxis, Wiesbaden: 89-102.

15. Dolan, M. und M. Doyle. 2000. Violence risk prediction. Clinical and actuarial measures and the role of the Psychopathy Checklist. *British Journal of Psychiatry* 177: 303-311.
16. Douglas, K.S., und C.D. Webster. 1998. The HCR-20 violence risk assessment scheme: concurrent validity in a sample of incarcerated offenders. *Criminal Justice and Behavior* 26: 3-9.
17. Douglas, K.S., J.R. Ogloff, und S.D. Hart. 2003. Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services* 54: 1372-1379.
18. Douglas, K.S., J.R. Ogloff, T.L. Nicholls, und I. Grant. 1999. Assessing risk for violence among psychiatric patients: the HCR-20 violence risk assessment scheme and the Psychopathy Checklist: Screening Version. *Journal of Consulting and Clinical Psychology* 67: 917-930.
19. Edwards, J., P. Steed, und K. Murray. 2002. Clinical and forensic outcome 2 years and 5 years after admission to a medium secure unit. *The Journal of Forensic Psychiatry* 13: 68-87.
20. Eronen, M., P. Hakola, und J. Tiihonen. 1996. Factors associated with homicide recidivism in a 13-years sample of homicide offenders in Finland. *Psychiatric Services* 47: 403-406.
21. Eucker, S., R. Müller-Isberner, und L.G. 1992. Strukturierte Krankenblattokumentation im Massregelvollzug. *Recht und Psychiatrie* 10: 20-26.
22. Faust, D. und B. Nurcombe. 1989. Improving the accuracy of clinical judgement. *Psychiatry* 52: 197-208.
23. Fulero, S.M. 1995. Review of the Hare-Psychopathy Checklist-Revised. In: Conoley, J.C. and J.C. Impara (Hrsg.) *Twelfth Mental Measurements Yearbook*. Buros Institute, Lincoln, N.E.: 453-454.
24. Gigerezer, G. 2000. *Adaptive Thinking - Rationality in the Real World*. Oxford University Press, Oxford, New York.
25. Ginossar, Z. und Y. Trope. 1980. The effects of base rate and individuating information on judgements about another person. *Journal of Experimental Social Psychology* 16: 228-242.
26. Glover, A.J.J., D.E. Nicholson, T. Hemmati, G.A. Bernfeld, und V.L. Quinsey. 2002. A comparison of predictors of general and violent recidivism among high risk federal offenders. *Criminal Justice and Behavior* 29: 235-249.
27. Gouvier, W.D. 1998. Baserates and clinical decision making in neuropsychology. In: Sweet, J.J. (Hrsg.) *Forensic Neuropsychology*. Swets and Zeitlinger, Lisse, Abingdon, Exton, Tokyo: 27-37.
28. Grann, M., H. Belfrage, und A. Tengstrom. 2000. Actuarial assessment of risk for violence: predictive validity of the VRAG and historical part of the HCR-20. *Criminal Justice and Behavior* 27: 97-114.

29. Grann,M., N.Langström, A.Tengström, und E.G.Stalenheim. 1999. Psychopathy (PCL-R) predicts violent recidivism among criminal offenders with personality disorders in Sweden. *Law and Human Behavior* 23: 203-215.
30. Gray,N.S., C.Hill, A.McGleish, D.Timmons, M.J.MacCulloch, und R.J.Snowden. 2003. Prediction of violence and self-harm in mentally disordered offenders: a prospective study of the efficacy of HCR-20, PCL-R, and psychiatric symptomatology. *Journal of Consulting and Clinical Psychology* 71: 443-451.
31. Gretenkord,L. 1994. Gewalttaten nach dem Massregelvollzug (§ 63 StGB). In: Steller,M., K.P.Dahle, and M.Basqué (Hrsg.) *Straftäterbehandlung*. Centaurus, Pfaffenweiler: 75-89.
32. Gretton,H.M., R.D.Hare, R.O'Shaughnessy, und G.Kumka. 2001. Psychopathy and recidivism in adolescent sex offenders. *Criminal Justice and Behavior* 28: 427-449.
33. Gross, G. 2004. *Deliktspezifische Rezidivraten im internationalen Vergleich*. Klinik und Poliklinik für Psychiatrie und Psychotherapie der Ludwig-Maximilians Universität München (Dissertation).
34. Hall,H.V. 1987. *Violence Prediction: Guidelines for the Forensic Practitioner*. Charles C. Thomas, Springfield, IL.
35. Hall,J.-R., S.D.Benning, und C.J.Patrick. 2004. Criterion-related validity of the three-factor model of psychopathy: personality, behavior, and adaptive functioning. *Assessment* 11: 4-16.
36. Hare,R.D. 1980. A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences* 1: 111-119.
37. Hare,R.D. 1990. *The Hare Psychopathy Checklist-Revised*. Multi-Health Systems, Niagara Falls, Toronto, Canada.
38. Hare,R.D. 1991. *The Hare PCL-R: Rating booklet*. Multi-Health Systems, Toronto, Canada.
39. Hare,R.D., D.Clark, M.Grann, und D.Thorton. 2000. Psychopathy and the predictive validity of the PCL-R: an international perspective. *Behavioral Science and the Law* 18: 623-645.
40. Hare,R.D., T.J.Harpur, A.R.Hakistan, A.E.Forth, S.D.Hart, und J.P.Newman. 1990. The Revised Psychopathy Checklist: Reliability and factor structure. *Psychological Assessment: A journal of Consulting and Clinical Psychology* 2: 338-341.
41. Harris,G.T., M.E.Rice, und J.A.Camilleri. 2004. Applying a forensic actuarial assessment (the Violence Risk Appraisal Guide) to nonforensic patients. *Journal of Interpersonal Violence* 19: 1063-1074.
42. Harris,G.T., M.E.Rice, und C.A.Cormier. 2002. Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior* 26: 377-394.

43. Harris,G.T., M.E.Rice, und V.L.Quinsey. 1993. Violent recidivism of mentally disordered offenders: the development of a statistical prediction instrument. *Criminal Justice and Behavior* 20: 315-335.
44. Harris,G.T., M.E.Rice, V.L.Quinsey, M.L.Lalumiere, D.Boer, und C.Lang. 2003. A multisite comparison of actuarial risk instruments for sex offenders. *Psychological Assessment* 15: 413-425.
45. Hartmann,J., M.Hollweg, und N.Nedopil. 2001. Quantitative Erfassung dissozialer und psychopathischer Persönlichkeiten bei der strafrechtlichen Begutachtung. *Nervenarzt* 72: 365-370.
46. Hildebrand,M., C.deRuiter, und V.deVogel. 2004. Psychopathy and sexual deviance in treated rapists: association with sexual and nonsexual recidivism. *Sex Abuse* 16: 1-24.
47. Jamieson,E. und P.J.Taylor. 2002. Follow-up of serious offender patients in the community: multiple methods of tracing. *International Journals of Methods in Psychiatric Research* 11: 112-124.
48. Johansson,E., H.Andershed, M.Kerr, und S.Levander. 2002. On the operationalization of psychopathy: further support for a three-faceted personality oriented model. *Acta Psychiatrica Scandinavia* 106: 81-85.
49. Kahneman,D. und A.Tversky. 1973. On the psychology of prediction. *Psychological Review* 80: 237-251.
50. Kjelsberg,E. 1999. A long-term follow-up study of adolescent psychiatric in-patients.Part IV.Predictors of a non-negative outcome. *Acta Psychiatrica Scandinavia* 99: 247-251.
51. Krakowski,M., J.Volavka, und D.Brizer. 1986. Psychopathology and violence: a review of literature. *Comprehensive Psychiatry* 27: 131-148.
52. Kröber,H.L., U.Faller, und J.Wolf. 1994. Nutzen und Grenzen standardisierter Schuldfähigkeitsbegutachtung. Eine Überprüfung des forensisch-psychiatrischen Dokumentationssystems. *Monatsschrift für Kriminologie und Strafrechtsreform* 77: 339-352.
53. Kroner,D.G. und J.F.Mills. 2001. The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior* 28: 471-489.
54. Langstrom,N. und M.Grann. 2002. Psychopathy and violent recidivism among young criminal offenders. *Acta Psychiatrica Scandinavia Supplement*: 86-92.
55. Loza,W. und G.K.Dhaliwal. 1997. Psychometric evaluation of the Risk Appraisal Guide (RAG): a tool for assessing violent recidivism. *Journal of Interpersonal Violence* 12: 779-793.
56. Loza,W., D.B.Villeneuve, und A.Loza-Fanous. 2002. Predictive validity of the Violence Risk Appraisal Guide: a tool for assessing violent offender's recidivism. *International Journal of Law and Psychiatry* 25: 85-92.

57. Maden,A., S.Rutter, T.McClintock, C.Friendship, und J.Gunn. 1999. Outcome of admission to a medium secure psychiatric unit - I. short- and long-term outcome. *British Journal of Psychiatry* 175: 313-316.
58. McBride,M. 1999. Predicting violence among federal inmates. *Corrections Research Forum*, Toronto, ON.
59. McNiel,D.E., A.L.Gregory, J.N.Lam, R.L.Binder, und G.R.Sullivan. 2003. Utility of decision support tools for assessing acute risk of violence. *Journal of Consulting and Clinical Psychology* 71: 945-953.
60. Mende,W. 1983. Zur Frage der Quantifizierung in der forensischen Psychiatrie. *Monatsschrift für Kriminologie* 66: 343-354.
61. Müller-Isberner,R., R.Freese, und D.Jockel. 2000. Forensic Psychiatric Assessment and Treatment in Germany. Legal Framework, Recent Developments, and Current Practice. *International Journal of Law and Psychiatry* 23: 467-480.
62. Nadeau, J., Nadeau, B., Smiley, W. C., und McHattie, L. 1999. The PCL-R and the VRAG as predictors for institutional behavior. Vancouver, BC.
63. Nedopil,N. 1997. Die Bedeutung von Persönlichkeitsstörungen für die Prognose künftiger Delinquenz. *Monatsschrift für Kriminologie und Strafrechtsreform* 80: 79-92.
64. Nedopil,N. 2000. *Forensische Psychiatrie*, 2 ed. Thieme, Stuttgart, New York.
65. Nedopil,N. 2001. Rückfallprognose bei Straftätern - eine wachsende Aufgabe für die forensische Psychiatrie. *Psycho* 70: 363-369.
66. Nedopil,N. und P.Grassl. 1988. Das Forensisch-Psychiatrische Dokumentationssystem (FPDS). *Forensia* 9: 139-147.
67. Nedopil, N. und Stadtland, C. 2003. The impact of base-rates of criminal recidivism on long-term outcome - results of a follow-up study of unselected group of criminal offenders (unveröffentlichtes Manuskript).
68. Nicholls,T.L., J.R.Ogloff, und K.S.Douglas. 2004. Assessing risk for violence among male and female civil psychiatric patients: the HCR-20, PCL:SV and VSC. *Behavioral Sciences and the Law* 22: 127-158.
69. Nugent, P. 2001. The use of detention legislation: factors affecting detention decision and recidivism among high risk offenders in Ontario. *Queens University at Kingston (Dissertation)*.
70. Nunes,K.L., P.Firestone, J.M.Bradford, D.M.Greenberg, und I.Broom. 2002. A comparison of modified versions of the Static-99 and the Sex Offender Risk Appraisal Guide. *Sex Abuse* 14: 253-269.
71. Nygaard, D. 2005. Die prädiktive Validität des HCR 20/2, der PCL-R und der ILRV für die Vorhersage gewalttätiger krimineller Rückfälle - eine Langzeituntersuchung von 185 Gutachtenprobanden . *Klinik und Poliklinik für Psychiatrie und Psychotherapie der Ludwig-Maximilians-Universität München (Dissertation)*.

72. Ogloff, J.R.P., S.Wong, und A.Greenwood. 1990. Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences and the Law* 8: 181-190.
73. Pierschke, R. 2001. Tötungsdelikte nach -scheinbar- günstiger Kriminalprognose - eine Analyse von Fehlprognosen. *Monatsschrift für Kriminologie und Strafrechtsreform* 84: 249-259.
74. Quinsey, V.L. 1998. *Violent Offenders - Appraising and managing risk*. American Psychological Association, Washington, D.C.
75. Quinsey, V. L., Book, A., und Skilling, T. A. 2001. A follow-up of deinstitutionalized developmentally handicapped men with histories of antisocial behavior (unveröffentlichtes Manuskript)
76. Quinsey, V.L., G.Coleman, B.Jones, und I.Altrows. 1997. Proximal antecedents of eloping and reoffending among mentally disordered offenders. *Journal of Interpersonal Violence* 12: 794-813.
77. Quinsey, V.L., M.Pruesse, und R.Fernley. 1975. A follow-up of patients found not guilty by reason of insanity and convicted felons. *Canadian Psychiatric Association Journal* 20: 461-467.
78. Quinsey, V.L., A.Warneford, M.Pruesse, und N.Link. 1975. Released Oak Ridge patients: a follow-up of review board discharges. *British Journal of Criminology* 15: 264-270.
79. Rice, M.E. und G.T.Harris. 1995. Methodological development. Violent recidivism: assessing predictive validity. *Journal of Consulting and Clinical Psychology* 63: 737-748.
80. Rice, M.E. und G.T.Harris. 1995. Violent recidivism: assessing predictive validity. *Journal of Consulting and Clinical Psychology* 63: 737-748.
81. Rice, M.E. und G.T.Harris. 1997. Cross-validation and extension of the violence risk appraisal guide for child molesters and rapists. *Law and Human Behavior* 21: 231-241.
82. Salekin, R. T., R.Rogers, und K.W.Sewell. 1997. Construct validity of psychopathy in a female offender sample: a multitrait-multimethod evaluation. *Journal of Abnormal Psychology* 106: 576-585.
83. Seifert, D., K.Jahn, S.Bolten, und M.Wirtz. 2002. Prediction of dangerousness in mentally disordered offenders in Germany. *International Journal of Law and Psychiatry* 25: 51-66.
84. Serin, R.C. und N.L.Amos. 1995. The role of psychopathy in the assessment of dangerousness. *International Journal of Law and Psychiatry* 18: 231-238.
85. Sjöstedt, G. und N.Langström. 2002. Assessment of risk for criminal recidivism among rapists: a comparison of four different measures. *Psychology, Crime and the Law* 8: 25-40.

86. Stadtland,C., N.Kleindienst, M.Eidt, C.Kröner und N.Nedopil. 2004. Risk assessment during pre-trial examinations. The impact of mental disorders on criminal recidivism (eingereicht).
87. Stadtland,C. und N.Nedopil. 2003. Alkohol und Drogen als Risikofaktoren für kriminelle Rückfälligkeit. Fortschritte der Neurologie und Psychiatrie 71: 654-660.
88. Stadtland, C. und Nedopil, N. 2004. Psychiatrische Erkrankungen und die Prognose krimineller Rückfälligkeit. Der Nervenarzt, Online-Publikation, 1-19.
89. Stadtland,C. und N.Nedopil. 2004. Vergleichende Anwendung heutiger Prognoseinstrumente zur Vorhersage krimineller Rückfälle bei psychiatrisch begutachteten Probanden. Monatsschrift für Kriminologie und Strafrechtsreform 87: 78-85.
90. Tengstrom,A. 2001. Long-term predictive validity of historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. Nordic Journal of Psychiatry 55: 243-249.
91. Tengström,A., M.Grann, N.Langström, und G.Kullgren. 2000. Psychopathy (PCL-R) as a predictor of violent recidivism among criminal offenders with schizophrenia. Law and Human Behavior 24: 45-58.
92. Thorton,D., R.Mann, S.Webster, L.Blud, R.Travers, C.Friendship, und M.Erikson. 2003. Distinguishing and combining risks for sexual and violent recidivism. Annals of N. Y. Academic Sciences 989: 225-235.
93. Ullrich,S., P.Borkenau, und A.Marneros. 2001. Personality disorders in offenders: categorical versus dimensional approaches. Journal of Personality Disorders 15: 442-449.
94. Verona,E., C.J.Patrick, J.J.Curtin, M.M.Bradley, und P.J.Lang. 2004. Psychopathy and physiological response to emotionally evocated sounds. Journal of Abnormal Psychology 113: 99-108.
95. Villeneuve,D.B. und V.L.Quinsey. 1995. Predictors of general and violent recidivism among mentally disordered inmates. Criminal Justice and Behavior 22: 397-410.
96. Vitale,J.E. und J.P.Newman. 2001. Using the Psychopathy Checklist Revised - with female samples: reliability, validity, and implications for clinical utility. Clinical Psychology: Science and Practice 8: 117-132.
97. Volckart,B. 2002. Zur Bedeutung der Basisrate in der Kriminalprognose. Recht und Psychiatrie 20: 105-114.
98. Walters,G.D. 2003. Predicting institutional adjustment and recidivism with the psychopathy checklist factor scores: a meta-analysis. Law and Human Behavior 27: 541-558.
99. Walters,G.D. 2003. Predicting criminal justice outcomes with the Psychopathy Checklist and Lifestyle Criminality Screening Form: a meta-analytic comparison. Behavioral Science and the Law 21: 89-102.

100. Webster,C.D. und C.Eaves. 1997. The HCR-20 Scheme. The Assessment of Dangerousness and Risk, 2 ed. Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia, Vancouver.
101. Webster,C.D. und D.Eaves. 1995. The HCR-20 Scheme. The Assessment of Dangerousness and Risk. Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia, Vancouver.
102. Werner,J. 1992. Biomathematik und medizinische Statistik, 2 ed. Urban und Schwarzenberg, Wien,München,Baltimore.
103. WHO. 1992. The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research. Geneva (Bericht).
104. Wormith,J.S. und M.E.Olver. 2002. Offender treatment attrition and its relationship with risk, responsivity, and recidivism. Criminal Justice and Behavior 29: 447-471.



## 8. Lebenslauf

Carolin Kröner  
Poschingerstr. 9  
81679 München

Tel.: 089-500 96 800  
email: [carolin.kroner@gmx.de](mailto:carolin.kroner@gmx.de)

### **Persönliche Daten**

---

Geburtsdatum und -ort	6. April 1979 in München
Familienstand	ledig
Staatsangehörigkeit	deutsch
Konfession	evangelisch

### **Schulausbildung**

---

September 1985 - Juli 1989	Georg-Schütte Grundschule, Garmisch-Partenkirchen
September 1989 - Juli 1997	Werdenfels Gymnasium, Garmisch-Partenkirchen
September 1994	Überspringen der 10. Jahrgangsstufe
Juli 1997	Abitur (Note: 1,0)

### **Studium**

---

September 1997 - Mai 2000	Yale University, New Haven, USA Studium der freien Künste (liberal arts, pre-medical program)
Mai 2000	Bachelor of Arts: Politikwissenschaften (Schwerpunkt internationale Beziehungen)
April 1999 - Juli 2000	Ludwig-Maximilians Universität, München Studium der Humanmedizin, vorklinischer Abschnitt, Semester 2-4 (Semester 1: Anerkennung der Studienleistungen aus den USA)
April 1999 – Mai 2000	Doppelimmatrikulation USA und Deutschland
August 2000	Ärztliche Vorprüfung
Oktober 2000 - Juli 2001	Technische Universität, München Studium der Humanmedizin, klinischer Abschnitt, Semester 5-6
August 2001	1. Staatsexamen
Oktober 2001- Juli 2002	Universität von Lausanne, Lausanne, Schweiz Studium der Humanmedizin, klinischer Abschnitt, Semester 7-8
September 2002 - März 2004	Technische Universität, München Studium der Humanmedizin, klinischer Abschnitt, Semester 9-10
März 2004	2. Staatsexamen
April 2005	3. Staatsexamen (Gesamtnote: 2,16)

## **Praktisches Jahr**

---

April 2004- August 2004	Chirurgie Clinico San Cecilio, Universität Granada, Granada, Spanien
August 2004 – Oktober 2004	Innere Medizin Intensivstation 2/11, Klinikum Rechts der Isar, TU München
Oktober 2004 – Dezember 2004	Western General Hospital/New Royal Infirmary, Universität Edinburgh, Grossbritannien
Dezember 2004 – März 2005	Neurologie Neurologische Klinik Rechts der Isar, TU München