

1 Von der Kriminalprognose zum Risikomanagement - Eine kurze Geschichte der Risikoeinschätzung bei Straftätern

1.1 Historisches

Die Frage nach der Kriminalprognose wird in Deutschland seit Beginn des 20. Jahrhunderts diskutiert und wurde 1933, als der Maßregelvollzug – zunächst unabhängig von den damaligen politischen Verwerfungen – in das Strafgesetzbuch aufgenommen wurde, an die Psychiater gestellt. Die Entscheidung, ob eine Maßregel (psychiatrisches Krankenhaus, Entziehungsanstalt, Sicherungsverwahrung) angeordnet wird oder nicht, oder ob ein Untergebrachter aus dem Maßregelvollzug entlassen wird, hängt nahezu ausschließlich von der Frage ab, ob bei dem Betreffenden ein Rückfall in delinquentes Verhalten zu befürchten ist. Das Gericht hat über die Fortdauer der Unterbringung oder Entlassung zu entscheiden, die primäre Aufgabe der psychiatrischen Sachverständigen ist dabei, dem Gericht empirisches Wissen zu vermitteln, damit es seine normative Entscheidung auf dieser Grundlage fundierter treffen kann.

Empirische Forschung in diesem Bereich begann in den 20er Jahren des 20. Jahrhunderts. Seither gab es verschiedene Phasen, in welchen sich optimistische und pessimistische Einschätzungen zum Sinn von Kriminalprognosen und zur Mitwirkung von forensischen Psychiatern und Psychologen bei der Erstellung von Kriminalprognosen abwechselten. Eine der ersten Merkmalssammlungen wurde von Burgess (1928) in den USA veröffentlicht. Die erste deutsche Prognosetafel stammt von Schiedt aus dem Jahre 1936. Nach dem zweiten Weltkrieg bis in die 70er Jahre wurden große Datenmengen analysiert, um potenzielle Straftäter zu charakterisieren (Böker & Häfner, 1973; Glueck & Glueck, 1950; Hartmann, 1977; Wolfgang et al., 1972). Solche Kriteriensammlungen nahmen ganz unterschiedliche Umfänge an. Sie reichten von drei Merkmalen, nämlich Bettnässen, Tierquälerei und Brandstiftung (Hellman & Blackman, 1966) bis zu Merkmalslisten mit mehr als 80 Fragen (Dietz, 1985). Derartige Merkmalskataloge schienen zunächst für die forensischen Fachleute, die eine Rückfallprognose abgeben sollten, von untergeordneter Bedeutung, da sie vorwiegend kriminologische und psychiatrisch kaum zu beurteilende Merkmale berücksichtigten. Aus heutiger Sicht muss angemerkt werden, dass man mit den damaligen Kriterienlisten eher der Frage nachging, welche Menschen aus einer bestimmten Population einmal Straftäter werden, und nicht der Frage, welche Straftäter rückfällig werden. Der Unterschied zwischen Kriminal- und Rückfallprognose wurde damals noch nicht gemacht. In Deutschland wurden diese Arbeiten auch nur im begrenzten Ausmaß zur Beurteilung der Rückfallprognose angewandt.

Die Literatur der 60er und 70er Jahre des vorigen Jahrhunderts (z.B. Leferenz, 1972) unterschied die folgenden drei Prognose-Methoden:

- Die intuitive Methode: Ihrer bedienten sich die Richter/Richterinnen, die aufgrund ihres theoretischen Allgemeinwissens und ihrer subjektiven Erfahrung in kurzer Zeit entscheiden mussten, welche Strafe oder welche Art der Strafverschöpfung aufgrund des Deliktes und der Persönlichkeit eines Täters gerechtfertigt oder sinnvoll erschien.
- Die statistische Methode: Sie basierte auf empirischen Untersuchungen, die jene (meist wenige) Faktoren ermittelten, die statistisch mit Rückfälligkeit korrelierten oder von Experten als Indikatoren für hohe Rückfallgefahr angesehen wurden.
- Die klinische Methode: Bei ihr wurde aufgrund der sorgfältigen biografischen Anamneseerhebung, einschließlich der Krankheits- und Delinquenzanamnese, von der Vergangenheit über die derzeitige Situation auf die Zukunft extrapoliert.

Damals wurde unter forensischen Psychiatern die sogenannte klinische Kriminalprognose bevorzugt, die eine individualprognostische Erfahrung des Gutachters und eine möglichst sorgfältige individuelle Exploration und Untersuchung in den Vordergrund stellte (Leferenz, 1972). Vor einer Verallgemeinerung prognostischer Kriterien wurde gewarnt. Ein kriteriengeleitetes Vorgehen wurde auch bis zur Jahrtausendwende noch kritisch gesehen (Volckart, 1999), es wurde allerdings nicht als Gegensatz zum hermeneutisch verstehenden und interpretierenden Vorgehen konzipiert, vielmehr wurde beides als Ergänzung zueinander gesehen (Dahle, 1997; Leygraf & Nowara, 1992; Nedopil, 1992, 1995; Nowara, 1995).

Ab den 70er-Jahren wurden die Fähigkeiten der Humanwissenschaftler, insbesondere der Psychiater, Kriminalprognosen zu erstellen, kritisch hinterfragt. Weil der US Supreme Court festgestellt hatte, dass bei Straftätern, die nach Verbüßung ihrer Gefängnisstrafe zur weiteren Sicherung zivilrechtlich in einem „state hospital“ untergebracht worden waren, „Gefährlichkeit“ konkret im Einzelfall beurteilt und nicht aus dem Vorliegen einer Diagnose abgeleitet werden darf, wurden fast 1000 Patienten (die sogenannte „Baxstrom-Population“) aus forensischen Krankenhäusern in die Allgemeinpsychiatrie entlassen. Eine Nachverfolgung dieser gemäß der Psychiater hoch rückfallgefährdeten Straftäter zeigte auf, dass sich Psychiater bei der Abgabe ihrer Prognosen häufiger irrten, als dass sie Recht behielten (Steadman, 1983; ebenso im Fall Dixon: Thornberry & Jacoby, 1979). Vergleichbare Schlussfolgerungen könnten auch aus neueren Untersuchungen aus Deutschland gezogen werden. Das Bundesverfassungsgericht hatte 1993 entschieden, dass eine Unterbringung nach DDR-Recht in den neuen Bundesländern nicht in eine Unterbringung im Maßregelvollzug umgewandelt werden darf. Dadurch wurden Untergebrachte, die zuvor wegen ihrer Gefährlichkeit angeblich nicht entlassen werden konnten, aus gesicherten Krankenhäusern entlassen. Eine Nachuntersuchung nach fünf Jahren zeigte, dass lediglich 5 von 33 der auf dieser Grundlage Entlassenen mit Gewalttaten rückfällig geworden waren. Hier, wie in einer Reihe anderer Untersuchungen, betrug der Anteil der Falsch-Positiven über 80% (siehe auch Nedopil, 2007a).

Ein ähnliches Ergebnis erbrachte die von Müller et al. (2013) durchgeführte Untersuchung. Bis zu ihrer Abschaffung am 01.01.2011 war die 1998 eingeführte nachträgliche unbefristete Sicherungsverwahrung an sog. „Nova“, also neue Anknüpfungstatsachen, die dem ursprünglich urteilenden Gericht nicht bekannt sein konnten, gebunden, wenn eine fortbestehende Gefahr gravierender krimineller Rückfälle von den Betroffenen angenommen wurde. Diese musste von zwei Sachverständigen begründet werden. Trotz deren übereinstimmender ungünstiger Prognosen wurden Strafgefangene entlassen, bei denen Gerichte keine „Nova“

feststellen konnten. Diese Fälle ermöglichten es, Betroffene zu untersuchen, die für gefährlich gehalten und trotzdem in Freiheit entlassen wurden. Die von Müller et al. (2013) durchgeführte Analyse der Legalbewährung umfasste 32 entlassene Straftäter und einen Katamnesezeitraum von 6,5 Jahren. Die entlassenen Straftäter befanden sich durchschnittlich 79 Monate (und mindestens 24 Monate) in Freiheit. Von ihnen wurden 9 (28%) mit Straftaten rückfällig, die Verurteilungen zu Haftstrafen von mehr als sechs Monaten nach sich zogen, eine davon zeitlich unbefristet, 12 Probanden (38%) wurden mit Rechtsverstößen auffällig, die mit Geldstrafe oder Freiheitsstrafe nicht über 6 Monate gehandelt wurden, 11 Probanden (34%) erhielten keinen neuen Eintrag in das Bundeszentralregister (BZR) respektive Vorstrafenregister (VOSTRA). Das Zielkriterium der Rückfälligkeit mit schwersten Delikten, hier definiert über ein Folgeurteil einer mindestens einjährigen Haftstrafe ohne Bewährung, erfüllten 87% nicht.

In den 80er Jahren des vorigen Jahrhunderts wurde zudem kritisiert, dass Psychiater und Psychiaterinnen von falschen Grundlagen für ihre Prognoseentscheidungen ausgingen (Rasch, 1984) oder dass sie es überhaupt mit ihrer professionellen Ethik nicht vereinbaren könnten, Kriminalprognosen abzugeben, die dann wiederum zu Sanktionen der Justiz führen würden (Stone, 1985). Die Entwicklung verlief jedoch anders als sich aus der damaligen skeptischen Einstellung hätte ableiten lassen. Weder waren die Gerichte bereit, auf humanwissenschaftlich begründete Kriminalprognosen zu verzichten, noch fanden sich andere Berufsgruppen, die in der Lage waren, bessere Kriminalprognosen abzugeben als Psychiaterinnen und Psychiater. Vielmehr etablierte sich eine neue Forschungsaktivität, die versuchte, unter Berücksichtigung der methodischen Schwächen der Untersuchungen der 50er- und 60er-Jahre den Zusammenhängen zwischen Gewalttätigkeit und psychischer Krankheit nachzugehen (Menzies, 1995; Monahan, 1996; Monahan & Steadman, 1994; Quinsey, 1995; Swanson, 1994); eine Zusammenfassung findet sich bei Monahan (1996) sowie Nedopil (1998).

Während in Deutschland zunächst eher typologische Ansätze zur Beschreibung rückfallgefährdeter Täter entwickelt wurden (Bischof, 1988; Göppinger, 1983), wurden im angloamerikanischen Raum und insbesondere in Kanada anhand biostatistischer Untersuchungen Korrelate von Rückfälligkeit identifiziert, die damals im deutschsprachigen Raum als „Prognoseinstrumente“ eingeführt wurden und zunehmende praktische Bedeutung gewannen.

Schon in der Erstauflage dieses Buches (Nedopil, 2005) wurde darauf hingewiesen, dass der Versuch, eine möglichst exakte Vorhersage eines Rückfalls vorzunehmen, die tatsächlichen Gegebenheiten der klinischen Realität verkennt: Klinische Prognosen werden abgegeben, um das vorhergesagte Verhalten zu verhindern. Die mit Hilfe von Instrumenten zur Risikoeinschätzung identifizierten mutmaßlichen Risikomerkmale dienen im klinischen Kontext der Indikationsstellung für Interventionen und können deshalb nach der Intervention nicht mehr mit Rückfälligkeit korrelieren. Es war zu erwarten, dass Merkmale, die zunächst (in der damaligen Diktion) als „aktuarische Risikofaktoren“ in Prognosebeurteilungen einfließen und als Indikatoren für Interventionsbedarf angesehen wurden, in späteren Untersuchungen nicht mehr mit Rückfällen oder Zwischenfällen korrelieren, weil durch ihre Beachtung Rückfälle aktiv verhindert wurden. Möglicherweise ist dieser Prozess, der von einem „aktuarischen Risikofaktor“ zu einem Interventionsindikator führt, ein Zeichen dafür, dass es sinnvoll ist, das betreffende Merkmal als mögliches Risikomerkmale zu betrachten. Diese Überlegungen hätten jedoch zur Vorsicht mahnen sollen: Klinische Risikomerkmale sollten solange nicht vernachlässigt werden, als nicht nachgewiesen ist, dass ihr Außer-Acht-Lassen zu keiner Erhöhung der Rückfallhäufigkeit führt.

Andere Autoren setzten sich kritisch mit der Überbetonung aktuarischer Risikoeinschätzungen auseinander. Leygraf (2004) warnte vor dem Gedanken, „*man könne die individuelle Gefährlichkeitsprognose allein oder auch nur überwiegend mit einem solchen Instrument beurteilen*“. Als ein Vorteil der klinischen Einschätzung wurde deren Dynamik und die Möglichkeit der qualitativen Erfassung angeführt (Davison, 1997). Die meisten Autoren sahen in der Folge eine professionell angewandte Kombination aktuarischer und klinischer Vorgehensweisen als sinnvoll an (Douglas et al., 2003; Lewis & Webster, 2004), wenngleich spätere Untersuchungen auch diese Auffassung in Frage stellten (z.B. Hanson & Morton-Bourgon, 2009).

In einigen Ländern wurden Gutachtergremien als sinnvoll erachtet, um Prognosen zum Zeitpunkt der Entlassung aus Institutionen vorzubereiten. In der Schweiz handelte es sich dabei um eine interdisziplinäre Fachkommission, welche aufgrund von Akteninformationen Empfehlungen an die Entscheidungsbehörden abgab (Dittmann, 2000; Ermer & Dittmann, 2001). Auch Niedersachsen hat Prognosekommissionen eingerichtet (MRVZN Moringen, 2009). Allerdings wurde bislang empirisch nie untersucht, ob sich durch die Einrichtung solcher Gremien die Risikoeinschätzung und das Risikomanagement verbessert oder ob sie möglicherweise sogar das Risikomanagement behindert haben.

1.2 Prognoseinstrumente und Risikofaktoren

Die Merkmalskataloge, aufgrund derer delinquentes und insbesondere aggressives oder sexuell deviantes Verhalten prognostisch besser eingeschätzt werden sollte, sind kaum mehr überschaubar. Bis 2008 wurden über 450 solcher Kataloge respektive „Prognoseinstrumente“ entwickelt (Chambers et al., 2009; Singh et al., 2014). Viele von ihnen haben einen großen Überlappungsbereich. Es zeigt sich somit, dass klinische Erfahrungen und empirische Untersuchungen zumindest in weiten Bereichen zu vergleichbaren Merkmalen gelangen. In einigen „Prognoseinstrumenten“ war die PCL-R von Hare (Hare, 1990, 1991, 2003) enthalten. In vielen Instrumenten hatten die anamnестischen und die objektiv feststellbaren Daten ein Übergewicht.

Bemerkenswert ist der Umstand, dass es sich bei den aktuarischen Instrumenten zur Risikoeinschätzung um Instrumente handelt, deren Algorithmus einfach gehalten ist: Der Algorithmus beinhaltet meist die Addition von denjenigen Merkmalen, die aufgrund einfacher Korrelationen mit Rückfälligkeit als Risikomerkmale eingestuft wurden. Diese Merkmale sind offenbar so relevant, dass sie austauschbar in den unterschiedlichen Instrumenten zur Risikoeinschätzung berücksichtigt werden können. Kroner et al. (2005) haben in einer sehr breit rezipierten Arbeit die Trennschärfe mehrerer Instrumente, namentlich der „Psychopathy-Checklist“ (PCL-R), dem Level of Service Inventory (LSI) und dem Violence Risk Appraisal Guide (VRAG), untersucht. Ergänzend haben die Autoren die Trennschärfe künstlicher Skalen berücksichtigt. Diese künstlichen Skalen wurden nach dem Zufallsprinzip erzeugt: Alle Merkmale der untersuchten Instrumente wurden in einen Topf geworfen. Anschließend wurden zufällig Kriterien aus dem Topf gezogen und daraus eine künstliche Skala geformt. Bemerkenswert war zweierlei: Erstens war die Trennschärfe dieser künstlichen Skalen signifikant besser als der Zufall. Zweitens waren die künstlichen Skalen nicht schlechter als die etablierten Instrumente zur Risikoeinschätzung. Das heißt, dass es empirisch keine Rolle spielt, ob man die PCL-R allein, den VRAG allein oder eine bunte Mischung aus beidem anwendet.

Aufgrund dieser Untersuchungen und der dabei entwickelten Instrumente wurden die Vorhersagetechniken operationalisiert und es wurde versucht die Methodik an jene Risikovorhersagen anzugleichen, die z.B. im Versicherungswesen Anwendung finden. Sie hießen aus diesem Grunde auch „actuarial predictions“. Damit waren kriterienorientierte Vorhersagetechniken gemeint, bei denen nicht nur die einzelnen Merkmale zuverlässig erhebbar sein müssen, sondern darüber hinaus die Abschätzung der verschiedenen Merkmale einem spezifischen Algorithmus unterworfen werden.

Eine Vielzahl von Veröffentlichungen wies darauf hin, dass durch die Verwendung standardisierter Instrumente zur Risikoeinschätzung eine größere Trennschärfe bezüglich krimineller Rückfälle erreicht werden kann als mit klinischen Methoden (z.B. Steadman et al., 2000). Dieser Befund steht im Einklang mit 50 Jahren psychologischer Forschung zur Urteilsbildung: In den 1950er Jahren publizierte Paul Meehl eine systematische Literaturanalyse und zeigte, dass die hochstandardisierte Urteilsfindung dem freien Expertenurteil überlegen ist (Meehl, 1954). Knapp fünfzig Jahre später zeigten Ægisdóttir et al. (2006) in einer sehr aufwendigen Meta-Analyse erneut, dass die klinische Urteilsfindung der algorithmusgestützten Beurteilung deutlich unterlegen ist. Die Autoren berücksichtigten dabei Daten aus dem ganzen Spektrum der Psychologie. Die Wahrscheinlichkeit, ein zutreffendes Urteil abzugeben, war bei Anwendung der standardisierten Methode 13% höher als bei Verwendung der klinischen Methode. Von allen Bereichen psychologischer Urteilsbildung war der gefundene Effekt der Überlegenheit der standardisierten Herangehensweise insbesondere für die Beurteilung des Gewalttrisikos deutlich.

Neben den Kriterienkatalogen für die Risikoeinschätzung bei Gewalt- und Sexualstraftätern wurden für bestimmte Populationen, z.B. Jugendliche (SAVRY - Structured Assessment of Violence Risk in Youth; Rieger et al., 2009) oder Frauen (Kjelsberg et al. 2009), und für bestimmte Fallkonstellationen, z.B. Gewalt in Partnerschaften (SARA - Spousal Assault Risk Assessment; Kropp et al., 1999), eigens konstruierte Instrumente zur Risikoeinschätzung entwickelt.

Weitere Überlegungen haben gezeigt, dass aktuarische Instrumente zur Risikoeinschätzung zwar im statistischen Schnitt trennschärfer zwischen rückfälligen und nicht rückfälligen Tätern unterscheiden können als andere Herangehensweisen, aber an der Anforderung scheitern, das Rückfallrisiko im Einzelfall (und nicht nur gruppenstatistisch) auszuweisen. Auf diese Schwierigkeiten haben schon in den 80er Jahren des vorigen Jahrhunderts verschiedene Fachleute (z.B. Monahan, 1981) hingewiesen und daraus gefolgert, dass Humanwissenschaftler zwar Risikofaktoren benennen können, dass aber individuelle Vorhersagen mit unzulässig großen Irrtumsmöglichkeiten verbunden sind. Die praktische Prognoseforschung hat aus diesen theoretischen Überlegungen Konsequenzen gezogen. Demnach ist eine psychiatrische und psychologische Risikoeinschätzung im konkreten Einzelfall nur dann sinnvoll, wenn sie mit einem adäquaten Risikomanagement verbunden ist (Bjorkly, 2004; Freese, 2003; Green et al. 2004; Hodgins & Müller-Isberner, 2004; Monahan & Silver, 2003; Nedopil, 2000b).

1.3 Zwischenbilanz

Diese Erkenntnisse haben zu einer gewissen Neuorientierung geführt. Aus dem Wissen, welches bis etwa 2005 angesammelt wurde, lassen sich folgende Schlussfolgerungen ableiten:

1. Es gibt unterschiedliche Konzepte der Risikoeinschätzung, die nicht isoliert angewandt werden sollten, von denen aber je nach der individuellen Fallkonstellation eines im Vordergrund stehen kann.
 - Ein **idiografisches Konzept**, bei welchem eingeschliffene individuelle Verhaltensmuster, die ein Wiederauftreten des Verhaltens wahrscheinlich machen, zur Grundlage der Beurteilung gemacht werden. Dabei sind auch statistische Einzelfallanalysen denkbar. Derartige eingeschliffene Verhaltensweisen, die zu oft wiederkehrendem Fehlverhalten führen, sind allerdings selten und die Datengrundlage würde nur in seltenen Fällen für inferenzstatistische bzw. empirisch abgestützte Aussagen reichen.
 - Häufig wird deswegen ein **nomothetisches Konzept** verfolgt, bei dem empirische Erkenntnisse aus einer Vielzahl von Untersuchungen auf den Einzelfall angewandt werden. Dieses Konzept war die Grundlage der später gängigen, empirisch begründeten Instrumente zur Risikoeinschätzung. Dieses Konzept allein reicht jedoch häufig nicht aus und ermöglicht kaum eine einzelfallbezogene Risikoeinschätzung.
 - Hierzu kann ein **hypothesengeleitetes Konzept** dienen, das auf der Entwicklung einer individuellen Hypothese zur Delinquenzgenese beruht. Dabei müssen die spezifischen Risikomerkmale, die der Hypothese zugrunde liegen, identifiziert werden. Hierzu bieten jene Instrumente zur Risikoeinschätzung eine wertvolle Hilfe, die eine hohe Inhaltsvalidität aufweisen. Nach diesem Konzept muss überprüft werden, ob die im Einzelfall identifizierten Risikomerkmale fortbestehen, wie relevant sie aktuell sind und ggf. ob sie durch protektive Faktoren kompensiert wurden. Damit wird die Risikoeinschätzung zu einem Prozess, der auch die Anwendung empirischen Wissens für den Einzelfall möglich macht.
2. Eine Reihe von Forschern hat relativ unabhängig voneinander die in den Instrumenten der Risikoeinschätzung enthaltenen Parameter in statische und dynamische Merkmale aufgeteilt (Andrews & Bonta, 1994; Nedopil, 1997; Quinsey et al., 1998). Die dynamischen Merkmale wurden weiter in "fixierte dynamische" und "akut dynamische" differenziert (Hanson & Bussière, 1998). Etwas plakativ kann man die Risikomerkmale folgendermaßen einteilen (s. a. Nedopil, 2000a):
 - **Statische Risikomerkmale:** Sie setzen sich aus anamnestischen Daten, stabilen persönlichkeitsgebundenen Dispositionen und kriminologischen Faktoren zusammen. Sie können sich im weiteren Verlauf des Lebens kaum noch ändern, sondern allenfalls korrigiert werden, wenn sie sich als Fehleinschätzung erwiesen haben. Sie bilden die Grundlage der aktuarischen Risikoeinschätzung und sagen, um wen man sich grundsätzlich Sorgen machen muss.

- **Dynamische Risikomerkmale:** Diese werden weiter unterschieden in „fixierte dynamische (aber veränderbare) Risikofaktoren“ und „aktuelle, sich ändernde Risikofaktoren“. Erstere beinhalten Fehlhaltungen und -einstellungen sowie risikoträchtige Reaktionsmuster. Sie erlauben eine Einschätzung der Behandlungsmöglichkeit und besagen, bei wem Änderungen möglich und erreichbar sind. Letztere bestehen aus klinischer Symptomatik, Einstellung und Verhalten in verschiedenen Situationen z.B. dissoziales Verhalten in einer Einrichtung. Sie ermöglichen eine klinische Risikoeinschätzung und sagen, wann man sich Sorgen machen muss.
3. Schließlich wurden die Instrumente zur Risikoeinschätzung sinnvollerweise nach unterschiedlichen Entwicklungsansätzen in aktuarische Risk-Assessment-Instrumente und in Instrumente des „Structured Professional Judgement“ unterteilt (Webster et al., 2002). Daneben gibt es noch Instrumente, die einem psychometrischen Ansatz folgen (vgl. Kapitel 10.7).
- Die **aktuarischen Instrumente** zur Risikoeinschätzung basieren auf biostatistischen Modellen. Die Bewertung der Merkmale erfolgt in einem geschlossenen Fragenformat, und die Integration der Merkmalsausprägungen zu einem Gesamtergebnis folgt einem festgesetzten Algorithmus. Für den Summenwert gilt, dass je höher der Punktwert ausfällt, desto höher fallen auch die Rückfallraten aus. Der erzielte Punktwert kann zudem in eine Risikokategorie überführt werden, für die erwartete Rückfallraten hinterlegt sind (die den in der Entwicklungsstichprobe beobachteten Rückfallraten entsprechen). Sie enthalten vornehmlich statische Risikomerkmale.
 - Die **Instrumente des Structured Professional Judgement (SPJ)** fassen die Merkmale zusammen, die sich aufgrund verschiedener Methoden als häufige und bedeutsame Risiko- oder protektive Merkmale ergeben haben. Sie dienen der Erstellung eines Risikoprofils im Einzelfall, bei dem üblicherweise wichtige Merkmale nicht übersehen und die Bedeutung der individuellen Merkmale hervorgehoben werden sollen. Eine Verrechnung der Einzelmerkmale ist nicht vorgesehen und wäre wissenschaftlich unsauber. Im Einzelfall kann u.U. ein Merkmal so gewichtig sein, dass sich der Fokus auf dieses Merkmal einschränkt. Das erstellte Risikoprofil sollte auch Hinweise für das Risikomanagement geben, es ist somit nicht nur prognostisch, sondern auch präventiv relevant.
 - Instrumente, die einem **psychometrischen Ansatz** folgen, bilden auf der Grundlage testtheoretischer Methoden Skalen, die indirekt das Rückfallrisiko erfassen. Grundlage dafür bildet ein theoretisches Konzept, ein Konstrukt. So kann z.B. die Ausprägung der „Psychopathy“ in der PCL-R oder die „criminogenic needs“ im LSI-R erfasst werden.

1.4 Von der Prognose zur Risikoeinschätzung

Bis um die Jahrtausendwende verkannte die Forschung häufig die praktische Aufgabe beim humanwissenschaftlichen Umgang mit psychisch gestörten Straftätern. Sie suchte in der Praxis nach der Bestätigung ihres prognostischen Könnens. Klinische Prognosen werden hingegen abgegeben, um das vorhergesagte Verhalten zu verhindern. Die Risikomerkmale dienen im klinischen Kontext zur Indikationsstellung für Interventionen und können deshalb nicht mehr mit der vorhergesagten Gefährlichkeit korrelieren.

Erst in den letzten zehn Jahren gelangte es immer mehr in das Bewusstsein, dass die Risikoeinschätzung in der forensisch-psychiatrischen/psychologischen Praxis nicht dazu dient, möglichst treffsichere Prognosen zu erstellen, sondern dazu beitragen soll, dass sich ungünstige Prognosen falsifizieren und günstige bewähren. Dieser Gedanke hat in anderen Disziplinen schon länger zu einem Paradigmenwechsel geführt. Der Historiker Yuval Noah Harari schrieb in seinem Buch „Homo Deus“ über den Sinn der Prognosen menschlichen Handelns (2017): *„Welchen Sinn macht es Prognosen abzugeben, wenn diese nichts ändern? Einige komplexe Systeme, wie z.B. das Wetter, sind sich der Prognosen nicht bewusst. Der Prozess der menschlichen Entwicklung reagiert im Gegensatz dazu aber auf Prognosen. In der Tat, je besser die Vorhersagen sind, desto mehr Reaktionen rufen sie hervor. Es wird mehr Wissen geschaffen und Menschen ändern ihr Verhalten und damit werden die Grundlagen unserer Prognosen und auch die Prognosen selbst überholt und hinfällig.“*

Dies entspricht auch medizinischem und klinischem Handeln: Die Analyse einer Symptomatik und die Diagnose einer Störung führen zwar zur Vorhersage des Verlaufs einer Krankheit, aber gleichzeitig zur Intervention, um einen ungünstigen Verlauf zu verhindern und die Prognose zu falsifizieren.

Es wurde bereits erwähnt, dass Merkmale, die zunächst als Risikomerkmale in Risikoeinschätzungen einfließen, im Lauf der Zeit als Indikatoren für Interventionsbedarf angesehen werden und dann nicht mehr mit Rückfällen oder Zwischenfällen korrelieren. Dieser Prozess, der von einem Prognosemerkmal zu einem Interventionsindikator führt, macht es sinnvoll, das betreffende Merkmal als Risikomerkmale, aber nicht als Prädiktor, zu betrachten.

Dieser Paradigmenwechsel hat in der forensischen Psychiatrie und Psychologie insofern zu einem Umdenken geführt, als es zu einer Konvergenz von Risikoeinschätzung und Behandlung kam und gleichzeitig die kontinuierliche und individuelle Risikoeinschätzung zu einem wesentlichen Faktor des professionellen Risikomanagements wurde.

Auch die Begrifflichkeiten wurden den neuen Paradigmen angepasst. Es geht nicht mehr um die Prognose, sondern um eine Risikoeinschätzung. Der grundsätzliche Unterschied besteht darin, dass Prognosen voraussagen sollen, was geschehen wird. Ihre Richtigkeit erweist sich, wenn das eintritt, was vorausgesagt wurde. Risikoeinschätzungen hingegen dienen dazu aufzuzeigen, was sich ändern muss, damit eine Gefahr nicht zur Realität wird, sondern abgewendet werden kann. Sie sind Voraussetzungen für ein optimales oder zumindest adäquates Risikomanagement. Ihre Richtigkeit erweist sich, wenn die Gefahr, die befürchtet wurde, vermieden oder verhindert wurde.

Um die Qualität einer Risikoeinschätzung zu erfassen, ist es erforderlich, eine große Anzahl von Schätzungen zu überprüfen und zu analysieren, ob die Interventionen das identifizierte Risiko adäquat und erfolgreich beseitigt haben oder ob sich bei Ausbleiben von Interventionen überzufällig häufig das Risiko verwirklicht hat (Silver, 2013). Risikoeinschätzungen sind

auch fehlerhaft, wenn die Kriterien oder die Methodik nicht sachgerecht angewandt wurden, z.B. wenn Risikomerkmale, deren Gültigkeit für Männer nachgewiesen wurde, unreflektiert bei Frauen angewandt werden oder wenn Parameter, die eine kurzfristige Risikoeinschätzung ermöglichen (z.B. die Entwertung der behandelnden Einrichtung), für langfristige Risikoeinschätzungen verwendet werden. Das Gleiche gilt auch umgekehrt, wenn z.B. statische Risikomerkmale (z.B. „Broken-Home-Situation“ in der Kindheit) bei kurzfristigen Risikoeinschätzungen Bedeutung erhalten. Wie häufig solche unsachgerechten Risikoeinschätzungen tatsächlich vorkommen, haben Rossegger et al. (2013) in einer umfangreichen Analyse von wissenschaftlichen Arbeiten zur Risikoeinschätzung nachgewiesen. Auf die Irrtümer und Fehler, die viele Menschen und auch Fachleute bei Vorhersagen und Risikoeinschätzungen machen, haben auch Tetlock und Gardner (2016) hingewiesen (vgl. **Kapitel 7**).

Auch in der Landschaft, in der Gutachten zur Risikoeinschätzung für Gerichte abgegeben werden, haben sich in den letzten 20 Jahren erhebliche Änderungen ergeben. Von juristischer Seite sind Gesetze zu berücksichtigen, die neue Entwicklungen ermöglichten.

1. Die Neuregelung der Führungsaufsicht eröffnet neue rechtliche Möglichkeiten der Nachsorge und des Risikomanagements nach einer Entlassung aus gesicherten Einrichtungen (§§ 68a bis 68c, 68e, 68f StGB).
2. Die elektronische Aufenthaltsüberwachung ermöglicht Kontrolle, welche sich ebenso wie die Führungsaufsicht und deren Möglichkeiten auf die Effektivität des Risikomanagements auswirken (§ 68b Abs.1 Nr. 12 StGB).
3. Die Gesetzesänderungen zum psychiatrischen Maßregelvollzug und zur Sicherungsverwahrung erfordern eine sachverständige Beurteilung der bislang durchgeführten therapeutischen Maßnahmen und des Therapieangebots im Prognosegutachten (§§ 66c, 67d StGB, § 119a StVollzG).
4. Die Gesetzesrevisionen zum Maßregelvollzug enthalten differenziertere prognostische Formulierungen, die je nach Ausgangsdelikt letztendlich (unterschiedliche) Antworten zur Zumutbarkeit eines Restrisikos für die Allgemeinheit erfordern (§§ 63, 64, 66, 66a, 66b 67, 67d StGB sowie § 463 Abs. 3 und 4 StPO).

Als weitere tatsächliche Änderung der vergangenen 15 Jahre wurde die ambulante Nachsorge für psychisch kranke Rechtsbrecher und in Bayern und z.T. in anderen Bundesländern auch für Strafgefangene, die aus Sozialtherapeutischen Abteilungen der Justizvollzugsanstalten entlassen werden, flächendeckend eingeführt. Gleichzeitig wurde die Kooperation der an der Nachsorge Beteiligten gefördert. Die beteiligten Organe arbeiten zusammen und nur noch selten gegeneinander: Ambulante Nachsorge, Führungsaufsicht, Bewährungshilfe, Wohneinrichtungen, Spezialabteilungen der Polizei, wie HEADS in Bayern, und die Strafvollstreckungskammern.

Zusammenfassend ist es heute viel deutlicher, dass das Risikomanagement im Vordergrund des Umgangs mit behandlungsbedürftigen Rechtsbrechern steht. Risikoeinschätzung, Risikokommunikation und adäquate Interventionsstrategien sind die Techniken, mit denen dieses Ziel angestrebt wird. Die Fragen, die an die Sachverständigen gestellt werden, der eine Risikoeinschätzung vornehmen soll, sind heute deutlich umfangreicher und differenzierter geworden, als dies z.B. bei der Erarbeitung der Mindestanforderungen 2006 (Boetticher et al., 2006) der Fall war. Sie wurden deshalb überarbeitet und 2019 als Empfehlungen den heutigen Anforderungen angepasst (Boetticher et al., 2019a; Kröber et al., 2019a).

2 Begriffsbestimmungen und Klarstellungen

2.1 Definitionen

In einem Buch, das von Autoren aus verschiedenen Fachdisziplinen geschrieben wurde und welches sich an eine Leserschaft aus unterschiedlichen Fachrichtungen wendet, ist eine möglichst korrekte und einverständliche Klärung der Terminologie erforderlich, um Missverständnisse zu vermeiden. Von einer solchen Klärung ist man aber im interdisziplinären Dialog weit entfernt. Begriffe wie Kriminalprognose, Legalprognose, Sozialprognose, Rückfallprognose oder Gefährlichkeitsprognose stehen immer noch oft austauschbar nebeneinander, obwohl sie durchaus etwas Unterschiedliches bedeuten. Deshalb sollen, wie schon in der Voraufgabe, einige Begriffe vorweg definiert werden.

Wie bereits dargestellt, hat sich die Hauptrichtung des Buches entsprechend der Entwicklung der Forschung und Praxis von der „Prognose“ zur „Risikoeinschätzung“ gewandelt. Die Fortschreibung der Mindestanforderungen für die Prognosebegutachtung, die 2019 als Empfehlungen veröffentlicht wurden (Boetticher et al., 2019b; Kröber et al., 2019b), haben den Begriff „Prognosen“ wohl aus Tradition beibehalten, zumal die von den Gerichten in Auftrag gegebenen Gutachten (irrtümlich) immer noch Prognosegutachten heißen, obwohl die ihnen zugrunde liegenden Gesetze gerade dazu dienen sollen, die prognostizierten Ereignisse abzuwenden.

2.2 Prognose oder Risikoeinschätzung

Prognosen sagen voraus, was geschehen wird, ihre Richtigkeit erweist sich, wenn das eintrifft, was vorausgesagt wurde.

Risikoeinschätzungen dienen dazu aufzuzeigen, was sich ändern muss, damit eine Gefahr nicht zur Realität wird, sondern abgewendet werden kann. Sie sind Voraussetzungen für ein optimales oder zumindest adäquates Risikomanagement. Ihre Richtigkeit erweist sich, wenn die Gefahr, die befürchtet wurde, vermieden oder verhindert wurde.

Von den verschiedenen Begriffsmöglichkeiten (Risikoerfassung, Risikobeurteilung, Risikoeinschätzung etc.) haben wir den Begriff Risikoeinschätzung gewählt, weil er dem seit vielen Jahren gebräuchlichen englischen Fachausdruck „risk assessment“ am nächsten kommt und weil er sich auf die empirische Vorgehensweise beschränkt und keine Wertung und kein „Urteil“ impliziert. Auch die Risikoeinschätzung enthält eine Annahme über den Zusammenhang zwischen gegenwärtigen Merkmalen und künftigen Verhalten und Ereignissen. Dieser Zusammenhang bezieht sich aber weit mehr auf empirisch statistische Korrelationen und fokussiert auf die Merkmale, die ein Risiko für gefährliches oder normüberschreitendes Verhalten nahelegen und die beim Risikomanagement besonders beachtet werden müssen, während Prognosen konkrete Vorhersagen individuellen Verhaltens implizieren.

In der klinischen Alltagssprache haben sich jedoch einige Begriffe etabliert, die in Zusammenhang mit diesem Buch als tatsächliche Vorhersagen verstanden werden, dennoch aber einer Abgrenzung bedürfen:

Die **Krankheitsprognose** beantwortet die Frage, ob und mit welcher Wahrscheinlichkeit Patienten mit bestimmten Krankheiten wieder genesen werden oder ob und mit welcher Wahrscheinlichkeit sie krank bleiben oder noch kränker werden.

Die **Behandlungsprognose** besagt, ob und mit welcher Wahrscheinlichkeit Symptome von Patienten mit einer bestimmten Störung durch eine Behandlung gebessert werden. In der forensischen Psychiatrie bezieht sich die Behandlungsprognose vorwiegend auf die Frage, ob künftiges kriminelles oder gefährliches Verhalten durch die Behandlung vermieden werden kann.

Die **Sozialprognose** ist der umfassendste Begriff, der im Zusammenhang mit prognostischen Fragestellungen in der forensischen Psychiatrie verwendet wird. Sie beantwortet die Frage, ob und mit welcher Wahrscheinlichkeit sich Menschen sozial integrieren und in welchem sozialen Umfeld sie sich in Zukunft bewegen werden. Die Sozialprognose bezieht sich sowohl auf die Arbeitswelt, die wirtschaftlichen und gesellschaftlichen Verhältnisse als auch auf Konflikte mit den gesellschaftlichen Normen oder deren Ausbleiben.

Die **Legalprognose** beantwortet die Frage, ob und mit welcher Wahrscheinlichkeit Menschen sich in Zukunft innerhalb des gesetzlichen Rahmens konfliktfrei bewegen oder mit den Gesetzen in Konflikt geraten werden. Kriminalprognose und Legalprognose sind weitgehend identisch, der Begriff der Legalprognose ist allerdings etwas weiter, weil er auch Konflikte mit dem Gesetz, die sich nicht als Kriminalität auswirken, umfasst, und die z.B. zu einer präventiven Unterbringung oder dem Widerruf einer Bewährung führen.

Die **Kriminalprognose** beantwortet die Frage, ob ein Mensch, eine bestimmte Gruppe von Menschen oder bestimmte Bevölkerungsanteile in Zukunft überhaupt kriminell werden – und zwar unabhängig von der Frage, ob sie bislang kriminell waren.

Die **Gefährlichkeitsprognose** beantwortet die Frage, ob ein Mensch, eine bestimmte Gruppe von Menschen oder ein bestimmter Bevölkerungsanteil in Zukunft eine Gefahr für andere darstellen werden – und zwar je nachdem, wie Gefahr definiert wird, und wiederum unabhängig von der Frage, ob sie bislang schon eine Gefahr verwirklicht haben. Üblicherweise wird in der forensischen Psychiatrie unter Gefährlichkeit eine Gewalttätigkeit oder ein sexueller Übergriff verstanden. Gefährlichkeitsprognosen müssen im klinischen Alltag der Psychiatrie häufig abgegeben werden, z.B. bei Unterbringungen auf geschlossenen Stationen oder bei Anwendungen der Unterbringungsgesetze oder PsychKGs bei Selbst- oder Fremdgefährdung.

Die **Rückfallprognose** beantwortet die Frage, ob Menschen, die bereits Delikte begangen haben, weitere Straftaten begehen werden. Die Vergleichspopulation für Rückfallprognosen sind somit alle Straftäter. Die Definition und Erfassung des Rückfalls werden in der Wissenschaft unterschiedlich vorgenommen. Mit unterschiedlichen Begründungen werden je nach Untersuchung erneute polizeiliche Festnahmen, erneute Verurteilungen oder erneute strafrechtliche Inhaftierungen als Rückfall definiert. In Deutschland und der Schweiz wird meist ein erneuter Eintrag in das BZR respektive VOSTRA, welcher in der Regel einer Wiederverurteilung entspricht (Ausnahme: Widerruf einer zur Bewährung ausgesetzten Maßregel), als Rückfallkriterium verwendet.

Unterschieden werden muss aber in jedem Fall zwischen einem Wiedereintrag ganz allgemein (allgemeine Rückfälligkeit, die Kriterium ist bei der **allgemeinen Rückfallprognose**) und einem Wiedereintrag wegen einer der früheren Delinquenz vergleichbaren Straftat (spezifische Rückfälligkeit und **spezifische Rückfallprognose**).

Kriminalprognose, Gefährlichkeitsprognose und Rückfallprognose beantworten unterschiedliche Fragen und müssen von unterschiedlichen empirischen Daten und Vergleichen ausgehen. Die Begriffe sind somit nicht austauschbar und sollten auch in Gutachten richtig angewandt werden. Bei Begutachtungen wird in der Regel nach der Rückfallprognose gefragt, weil das Fach sich zumeist mit Menschen befasst, die bereits straffällig geworden sind. Insofern ist auch der Untertitel in der Neufassung der Mindestanforderungen „Erfahrungswissenschaftliche Empfehlungen für kriminalprognostische Gutachten“ (Kröber et al., 2019b) irreführend.