

2. Eigene Arbeiten

Zunächst werden die eigenen Studien zur Entwicklung diagnostischer Verfahren vorgestellt. Dabei werden drei exemplarische Studien vertieft, die zentrale Ziele der aktuellen Instrumentenentwicklung verfolgen. Diese Ziele sind: die deutschsprachige Adaptation international etablierter Verfahren, die Validierung und Überprüfung der strukturellen Stabilität von KTT-basierten Instrumenten an verschiedenen Stichproben, die gleichzeitige Kürzung des Umfangs ohne nennenswerten Verlust an Messpräzision, die Anwendung der probabilistischen Testtheorie für computer-adaptives Testen und die Multimodalität der Diagnostik im Rahmen von Evaluationsstudien.

Im anschließenden klinischen Schwerpunkt werden fünf Studien zur Diagnostik autodestruktiver Syndrome dargestellt.

2.1 Diagnostik von Lebensqualität und Belastungserleben

In den eigenen Studien wurden verschiedene Instrumente zur Erfassung des subjektiven Belastungserlebens, der gesundheitsbezogenen Lebensqualität und potentiell vermittelnder Copingprozesse und Ressourcen entwickelt, weiterentwickelt oder für den deutschen Sprachraum adaptiert und validiert. Um die diagnostischen Themen wie auch die Kernzielsetzungen der eigenen Arbeiten zur Instrumentenentwicklung abzustecken, folgt hier ein knapper Überblick über die Arbeiten, die nicht abgedruckt sind.

Zur Erfassung der gesundheitsbezogenen Lebensqualität wurde der Fragebogen Alltagsleben [68], ein frühes krankheitsübergreifendes Lebensqualitätsverfahren der gleichen Autorengruppe, die später im Deutschen den SF-36 [34] vorlegte, weiterentwickelt. Die ursprüngliche Faktorenstruktur war an vergleichsweise kleinen und spezifischen Stichproben ermittelt worden und konnte damit nicht als stabil vorausgesetzt werden. Zudem erwies sich der Umfang mit 42 Items für die hiesige Klinikroutine als zu lang. Basierend auf

eigenen Untersuchungen an 2.105 Patienten mit unterschiedlichen psychischen und körperlichen Krankheitsbildern wurde eine faktorenanalytische Restrukturierung, Validierung und Kürzung auf die Hälfte der Items vorgenommen. Der Fragebogen bewährte sich unter anderem in den Katamnesestudien psychosomatischer Behandlung an der Charité als Evaluationsinstrument [69;70].

Zur Erfassung der krankheitsspezifischen Lebensqualität von Patienten mit Chronisch Entzündlichen Darmerkrankungen (CED) wurde der 10 Items umfassende Short Inflammatory Bowel Disease Questionnaire (SIBDQ) [71] ins Deutsche übertragen, adaptiert und in einer Kooperation mit der CED-Sprechstunde des Virchowklinikums an 120 Patienten validiert. Er konnte in Folgestudien zu Lebensqualität, subjektiven Ursachenannahmen und Coping bei CED-Patienten verwendet werden [72;73].

Im salutogenetischen Ansatz nehmen Erwartungshaltungen der Person und andere gesundheitsrelevante Kognitionen einen zentralen Stellenwert ein [74;75]. Diesem folgend wurde ein Instrument entwickelt, das sich aus einem international validierten Fragebogen zur Selbstwirksamkeit [76] und einem entsprechenden Optimismusfragebogen [77] zusammensetzt [78]. Das 18 Items umfassende Instrument wurde von 726 Patienten aus fünf diagnostischen Gruppen beantwortet. Die Faktorenanalysen ergaben eine dreidimensionale Struktur (Selbstwirksamkeit, Optimismus, Pessimismus). Die Anzahl der Items ließ sich dabei ohne Einbußen an Information oder Messpräzision von 18 auf 9 reduzieren. Der Fragebogen bewährte sich unter anderem für die Prognose langfristig günstiger Verläufe nach stationärer psychosomatischer Behandlung [69;70] und für den Nachweis, dass Zusammenhänge zwischen Krankheitsverarbeitung und Lebensqualität bei CED-Patienten von den Erwartungshaltungen der Patienten vermittelt sind [73].

Als ein Kernelement krankheitsbezogener Kognitionen gilt die Frage, auf welche Ursachen die Betroffenen ihre Erkrankung zurückführen [79-82]. Dies gilt insbesondere für Erkrankungen, deren Genese unklar ist [83] oder deren wissenschaftliche Erklärungsmodelle sich wandeln [84] und die chronische Anpassungsleistungen erfordern, wie im Falle der CED. In eigenen Untersuchungen wurde ein Fragebogen entwickelt, der subjektive

Ursachenannahmen von CED-Patienten erfasst. Anhand der Daten von 106 CED-Patienten wurden faktorenanalytisch die Dimensionen „Dysfunktionale Stressregulation“, „Interaktionalität“, „Lebensstil“, „Physiologie“ und „Fatalismus“ ermittelt und der Fragebogen wurde psychometrisch validiert.

2.1.1 Diagnostik des subjektiven Belastungserlebens

In der Stressmessung finden sich zwei unterschiedliche Zugänge. Während der eine auf a priori definierte, objektivierbare, äußere Stressoren, wie Traumata, kritische Lebensereignisse [85] oder die so genannten *daily hassles* [86;87] ausgerichtet ist, legt der zweite Zugang das Augenmerk auf die subjektive Wahrnehmung und Bewertung äußerer Bedingungen durch das Individuum [88].

Levenstein et al. [89] entwickelten ein Verfahren, das dem zweiten Ansatz verpflichtet ist (Perceived Stress Questionnaire – PSQ), mit dem das Ausmaß der subjektiv erlebten aktuellen Belastetheit – unabhängig von a priori definierten, „objektiven“ Stressoren – gemessen werden soll. Die Items sind weitgehend kontextunspezifisch formuliert und sollen damit einen breiten Geltungsbereich haben. Der PSQ war bis dato nur an Patienten mit chronisch entzündlichen Darmerkrankungen eingesetzt worden und lag in keiner deutschen Version vor.

In der eigenen Studie wurde das Instrument in einem Vorwärts-Rückwärts-Verfahren übersetzt und in einer anschließenden strukturierten Konsensusbildung für den deutschen Sprachkontext adaptiert. Es wurde an 650 Patienten unterschiedlicher Krankheitsgruppen teststatistisch überprüft. Faktorenanalytisch wurden vier Skalen ermittelt (Sorgen, Anspannung, Freude, Anforderungen). Der Umfang der Skalen wurde auf je fünf Items gekürzt. Formales Kriterium hierfür war die korrigierte Item-Skalen-Korrelation, mit dem Ziel möglichst hohe Reliabilität der Skalen zu gewährleisten. Die Gütekriterien entlang des Modells der KTT entsprechen den testtheoretischen Anforderungen. Belege für die externe Validität der Stressmessung finden sich unter anderem über Zusammenhänge mit immunologischen Parametern.

In weiteren Schritten wurde mit konfirmatorischen Faktorenanalysen in der Form von linearen Strukturgleichungsmodellen die Stabilität der 4-faktoriellen Skalenlösung gegenüber alternativen Skalenlösungen belegt. Mit Hilfe von Multi-Sample-Analysen, einer Spezifikation von linearen Strukturgleichungsmodellen, wurde die strukturelle Stabilität für unterschiedliche Patientenstichproben und gesunde Erwachsene belegt. Zudem wurden für die einzelnen Krankengruppen und die Gruppe gesunder Erwachsener Referenzwerte ermittelt.

2.1.2 Diagnostik auf Grundlage des probabilistischen Messmodells: Computer-adaptives Testen (CAT)

Während alle bislang vorgestellten Messverfahren auf den Modellannahmen der Klassischen Testtheorie gründen, war ein weiteres Forschungsziel, klinisch psychologische Instrumente auf der Grundlage der Item-Response-Theorie zu entwickeln. Als besonders relevante klinische Konstrukte wurden Depression und Angst fokussiert. Depressive Störungen zählen nicht nur zu den häufigsten psychischen Störungen überhaupt, auch 15% bis 33% somatischer Patienten leiden zusätzlich an einer Depression [90-92]. Die Verbreitung von Angststörungen wird für Patienten in Allgemeinarztpraxen sogar mit bis zu 50% angegeben [93]. Unter den eigenen stationär psychosomatischen Patienten leiden (komorbid) um die 60% an depressiven Störungen und knapp 30% an Angststörungen [69].

Die erste Voraussetzung für die Entwicklung eines computer-adaptiven Tests war die Möglichkeit, Fragebögen computerisiert zu erheben. In unserer Klinik wurde 1995 eine mobile computerisierte Erfassung implementiert und systematisch auf alle standardisierten psychometrischen Verfahren ausgeweitet. Dadurch wird im Vergleich zu herkömmlichen Paper-Pencil-Methoden erreicht, dass die Bearbeitung einzelner Verfahren vollständig erfolgt und eine zeitintensive und fehleranfällige Datenübertragung per Hand durch automatisiertes Einlesen ersetzt wird. Berechnungen zufolge werden dadurch circa 60 Prozent der Kosten eingespart, ohne dass dies zu nennenswerten Veränderungen in den Daten selbst führt [94].

Die Instrumentenentwicklung gliederte sich in folgende Schritte: (1) Itempool-Bildung, (2) Prüfung der Modell-Voraussetzungen (lokale stochastische Unabhängigkeit der Items), (3) Item-Response-Analysen, (4) Festlegung des adaptiven Erhebungsalgorithmus und (5) Vergleich der Ergebnisse mit den Resultaten herkömmlicher Methoden. Das Vorgehen orientiert sich an den Empfehlungen von Embretson & Reise [43] und Ware et al. [46].

Diese werden im Folgenden kurz erläutert:

(1) Itempool-Bildung

Neben den hier validierten Instrumenten zu Belastungserleben, Alltagsfunktion und Selbstwirksamkeit wurden die in der Klinikroutine etablierte sowie zusätzlich herangezogene spezielle Verfahren darauf gesichtet, ob sie Items zu den beiden interessierenden Konstrukten Depression und Angst beinhalteten. Neben dem Kriterium, dass die Items überhaupt repräsentativ für das Konstrukt sein sollten (Konstruktvalidität), sollten sie das Merkmalskontinuum möglichst in seiner gesamten Breite abdecken. Die erste Beurteilung erfolgte durch drei unabhängige Rater und war gefolgt einen Delphi-Prozess für die Items, bei denen die Ratings voneinander abwichen.

Der Umfang der beiden Item-Pools war a priori nicht beschränkt. Er sollte möglichst groß sein, um eine hinreichende Varianz in den Wahrscheinlichkeitsverteilungen der Items zu gewährleisten (s.o.). Ein Pool von mindestens 60 Items war hierfür avisiert. Für Depression ließ sich nach dem Beurteilungsverfahren ein Ausgangsitempool von 144 Items realisieren, für Angst ein Pool von 81 Items.

Für den Itempool zu Depression lagen Datensätze von ca. 1.500 Patienten, zu Angst von etwa 2.300 Patienten zur weiteren Analyse vor.

(2) Prüfung der Modellvoraussetzungen

Voraussetzung für das gewählte Modell (Generalized Partial Credit Model nach Muraki [95]) ist die lokale stochastische Unabhängigkeit der Items. Dies bedeutet, dass die einzelnen Items, nach Auspartialisieren der Ladung auf einem gemeinsamen Faktor keine bedeutsamen Kovariation mehr aufweisen

dürfen. Bei Unidimensionalität der Items wäre diese Voraussetzung gegeben. Überprüft wurde dies also anhand der residualen Korrelationen nach konfirmatorischer Faktorenanalyse (mittels Mplus™).

(3) Item-Response-Analysen

Die Item-Response-Analyse umfasst die grafische Analyse der Item Response Curves (IRC) sowie die Berechnung von Reliabilität und Standardmessfehler der Itembank (mittels Testgraf™). Die Item Response Curves liefern eine Antwortwahrscheinlichkeitsverteilung bezogen auf den latenten Trait für jedes Item und jede Antwortoption. Durch eine mathematische Funktion werden diese Verteilungen etwas geglättet und der in Abbildung 1 veranschaulichten polynomialen Funktion angepasst. Geprüft wird, ob eingipflige, monoton verlaufende Item Response Curves (IRC) vorliegen, welche pro Antwortkategorie in genau einem Messbereich mit ihrem Maximum alle anderen Item Response Curves des jeweiligen Items übersteigen.

Die Schnittpunkte der IRCs, also die Schwellen zwischen den Antwortkategorien nennen sich „Thresholds“ (Schwellenparameter). Der Mittelwert der Thresholds wird „Location Parameter“ genannt. Er dient der Lokalisation der Items auf dem latenten Merkmalskontinuum und gilt quasi als mittlerer Messbereich des Items. Die durchschnittliche Steigung („Slope Parameter“) beinhaltet die Information darüber, wie gut das Item zwischen Testpersonen unterschiedlicher Merkmalsausprägungen differenziert (höhere Steigung entspricht besserer Differenzierung). Location Parameter, Item Response Thresholds und Slope Parameter werden mittels Parscale™ überprüft.

Die beschriebenen Prüfungen erfolgten nach festgelegten Cut-Off-Kriterien, für die es kaum mathematische Vorgaben, teilweise aber Modelle in der Literatur gab. Nachdem die Items diese Selektion vollständig durchliefen, verblieben im Itempool für Depression 64 Items, für Angst 50 Items.

(4) Festlegung des Algorithmus

Der Steueralgorithmus wurde von Otto B. Walter in unserer Klinik entwickelt und mit Hilfe des Programmes C++ programmiert.

Der Algorithmus, der die Abfolge der Items reguliert, sieht folgendes Procedere vor: Das zuerst präsentierte Item sollte einen möglichst hohen Informationsgehalt aufweisen, indem seine Wahrscheinlichkeitsverteilungen über die Antwortalternativen einen möglichst großen Bereich des gesamten Merkmalskontinuums abdecken. Der Patient wählt eine Antwortalternative aus. Mit dieser Wahl lässt sich das Intervall des Merkmalsbereichs, in dem die Merkmalsausprägung des Patienten mit einer bestimmten Wahrscheinlichkeit liegt, näher einengen. Die Berechnung folgt der Expected A Posteriori (EAP) Methode. Eine Suchroutine wählt für die nachfolgende Itempräsentation dasjenige Item aus, das wiederum dieses Intervall am besten differenziert. Dies wird solange fortgesetzt, bis die Stoppfunktion ausgeführt wird. Die Stoppfunktion wird im ungünstigsten Fall bei maximaler Itemzahl, im günstigeren Fall dann ausgeführt, wenn die vorab festgelegte Präzision der Messung erreicht ist. Hierzu wird nach jeder Beantwortung eines Items ein Konfidenzintervall für die Merkmalsausprägung auf der Gesamtskala berechnet. Dieses wurde mit einem Standardfehler von SE kleiner/gleich 0.32 festgelegt, was einer Reliabilität von $r=0.90$ gleichkommt. Ist die zuvor erwartete Genauigkeit der Messung erreicht oder die maximale Anzahl von Items beantwortet, ist die Stoppfunktion erfüllt.

(5) Ergebnisvergleich

In Simulationsstudien wurde über den jeweiligen Vergleich mit etablierten Fragebögen zu den Konstrukten Depression und Angst, die nach den Annahmen der KTT konstruiert sind, belegt, dass die neu entwickelten CATs konstruktvalide sind, keine Einbußen an Messpräzision mit sich bringen, in den Extrembereichen der Merkmalsausprägung besser differenzieren und einen deutlich verringerten durchschnittlichen Erhebungsaufwand mit sich bringen.

Die Publikationen berichten über die Instrumentenentwicklung. In die Übersicht wurde die Arbeit zur Messung von Depression aufgenommen. In der Folge wurden prospektive Validierungsstudien an weiteren Patientenkollektiven durchgeführt, die aktuell in der Abschlussphase der Datenerhebung sind.

2.1.3 Prognose, Evaluation und Qualitätssicherung psychosomatischer Behandlung

Eine wesentliche Funktion psychologischer Diagnostik in der Psychosomatik besteht darin, die Effekte psychosomatischer Behandlung zu evaluieren und prognostische Faktoren für den Behandlungserfolg zu ermitteln.

In den eigenen Studien wurde überprüft, inwieweit sich Körperbeschwerden, Stimmungsmerkmale, Lebenszufriedenheit, Alltagsfunktion und soziale Beziehungen unter stationärer psychosomatischer Therapie verbessern und inwieweit entsprechende Verbesserungen katamnestisch stabil bleiben. Hierzu wurden Statuserhebungen – teilweise mit Hilfe eigens neu validierter Instrumente [96] – jeweils zur stationären Aufnahme, zur Entlassung und zur Katamnese durchgeführt, über die Messwiederholung in einen zeitlichen Bezug zueinander gesetzt und um retrospektive Einschätzungen der Patienten zum Katamnesezeitpunkt ergänzt [97]. Die Auswahl der Instrumente war inhaltlich an den zentralen Dimensionen gesundheitsbezogener Lebensqualität orientiert. Sie sollte möglichst ökonomisch sein und Redundanzen vermeiden, um die Patientenbelastung gering zu halten und vor allem die Bereitschaft zur Mitarbeit bei der katamnestischen Nachbefragung nicht zu gefährden. Zudem sollten die Verfahren störungsübergreifend sein, um differentielle Veränderungen zwischen Gruppen vergleichen zu können. Um die Daten so aufzubereiten, dass sie auch geeignet sind, in Metaanalysen zu Therapieeffekten aufgenommen zu werden, wurden nicht nur absolute Veränderungen, sondern auch Verteilungswerte und Effektstärken für die einzelnen Änderungen mitgeteilt.

Als prognostisch relevante Merkmale der Person erwiesen sich erwartungsgemäß generalisierte Erwartungshaltungen – gemessen mit dem SWOP [78] –, insbesondere Pessimismus als prognostisch ungünstiger Faktor, in geringerem Umfang auch Optimismus und Selbstwirksamkeit als prognostisch günstige Faktoren.

In der Evaluation psychosomatischer, psychiatrischer und psychotherapeutischer Behandlung gilt grundsätzlich das Primat multimodaler Diagnostik [19]. Neben der Erfassung unterschiedlicher Datenebenen

(körperliche, psychische und soziale Dimension sowie Alltagsbewältigung) erstreckt sich diese Forderung auch darauf, auf verschiedene Datenquellen zuzugreifen [25;26]. Dabei spielen insbesondere unter ökonomischen Gesichtspunkten kostenrelevante Daten der medizinischen Inanspruchnahme und des Arbeitsausfalls durch Krankheit eine Rolle. Selbstauskünfte von Patienten gelten in dieser Hinsicht als eingeschränkt verlässlich, so dass eine Erhebung solchermaßen objektivierbarer Parameter auf andere Datenquellen, etwa die Kostenträger der Krankenbehandlung zugreifen sollte.

In einem weiteren Schritt der eigenen Katamnesestudien wurden als kostenrelevante Parameter Krankenhausbehandlungstage und Arbeitsunfähigkeitszeiten erhoben und es wurde ermittelt, inwieweit sich in der Folge stationärer psychosomatischer Therapie günstige Effekte auf diese Parameter zeigen. Unter methodischen Gesichtspunkten war zudem die Frage von Interesse, inwieweit subjektive Angaben der Patienten zu Körperbeschwerden und anderen Parametern ihrer gesundheitsbezogenen Lebensqualität Korrespondenzen zu den objektiven Parametern (Krankenhaustage und Arbeitsunfähigkeitszeiten) aufweisen.

Tatsächlich zeigen sich kaum Korrespondenzen zwischen subjektiven und objektiven Erfolgsparametern psychosomatischer Therapie. Zwar lassen sich auf allen Ebenen Verbesserungen nachweisen, diese erfolgen jedoch nicht nur für subjektive und objektive Indikatoren, sondern auch für die beiden objektiven Indikatoren Krankenhausstage und Arbeitsunfähigkeitstage zeitversetzt.

Generalisierte Erwartungshaltungen erwiesen sich auch in Hinblick auf die objektiven Erfolgsparameter als relevante Prädiktoren. Hier war insbesondere die Selbstwirksamkeit prognostisch günstig für abnehmende Krankenhausstage.

In der Übersicht wird die zweite Arbeit dargestellt, da diese der Forderung nach einer multimodalen Diagnostik in der Therapieevaluation nachkommt und zugleich prognostisch relevante Faktoren für die hinzugenommenen Datenebenen ermittelt.

2.2 Diagnostik autodestruktiver Syndrome

Die Messung autodestruktiver Phänomene ist noch in einem Entwicklungsstadium begriffen, in dem die Aufgaben der Deskription und der Klassifikation zu lösen sind. Autodestruktive Phänomene werden in der einschlägigen Literatur bislang nicht einheitlich definiert und nicht befriedigend klassifiziert. Insbesondere die etablierten Klassifikationsschemata wie die ICD-10 oder das DSM lösen diese Aufgabe nicht (siehe oben). In der Literatur zur Klassifikation autodestruktiver Phänomene finden sich mehrere potentiell relevante Alternativen [63;98-100].

Verschiedene Autoren verlassen den Ansatz, der von exklusiven Klassen, also eindeutig voneinander abgrenzbaren Störungs-Entitäten ausgeht. Alternativ werden Kontinuumsmodelle und Typenmodelle vorgeschlagen.

Nadelson [99] formuliert ein breites Kontinuum, auf dem sich sämtliche Krankheiten einordnen lassen sollen und welches abbilden soll, wie real bzw. fiktiv („real / unreal“) eine Erkrankung ist. Zwischen den Endpunkten realer körperlicher Erkrankung auf der einen Seite und Simulation auf der anderen liegen dem Modell zufolge nur graduelle Abstufungen, die von real vorhandenen, organpathologisch erklärbaren und adäquat wiedergegebenen Beschwerden über aggravierende Formen des Umgangs mit Erkrankungssymptomen, weiter über das Spektrum somatoformer Beschwerden mit mehr oder weniger Ausmaß an organpathologischem Korrelat hin zu artifiziellen Symptomen führen. Die Abstufungen beziehen sich auf den Gesichtspunkt des Gemachten bzw. Fiktiven. Kritisch anzumerken ist daran, dass die Zuordnung zu der Dimension des „Fiktiven“ (im amerikanischen Orig. „unreal“) missverständlich sein könnte. Denn sofern der „fiktiven“ körperlichen Störung eine psychische Störung zugrunde liegt, ist diese ja wiederum real, auch wenn die Körpersymptome fiktiv im Sinne von vorgegeben oder selbst herbeigeführt sein mögen.

Eisendrath [63] sowie Feldman, Hamilton und Deemer [98] sprechen sich für ein Kontinuum artifiziellen Krankheitsverhaltens aus. Ihr Ansatz bezieht sich auf das Spektrum organpathologisch unerklärter Krankheitssymptome und reicht von somatoformen Symptomen am einen Ende des Kontinuums über

selbstschädigendes Verhalten hin zu „artifiziellen Störungen“ am anderen Ende. Der Vorteil dieses Ansatzes ist vor allem, dass die Nähe „artifizieller Störungen“ und somatoformer Störungen zueinander erhalten bleibt. Patienten beider Störungen werden stark durch ihren Umgang mit der Symptomatik im Rahmen der Interaktion mit dem medizinischen System definiert. Die Annahme fließender Übergänge zwischen diesen Phänomenen dürfte daher der klinischen Realität eher gerecht werden als die Annahme disjunkter Klassen.

Kontinuumsansätze werden den fließenden Übergängen klinischer Phänomene gerecht, sie sind allerdings überaus schwierig messtheoretisch umzusetzen, da sie nicht nur Annahmen über graduelle Abstufungen, also quantitative Aspekte, sondern auch über inhaltliche Akzentverschiebungen, also qualitative Aspekte der klinischen Phänomene beinhalten. Ein Messinstrument, das einen solchen Ansatz umsetzen würde, ist bislang nicht verfügbar.

Dagegen entwickelte die „Arbeitsgruppe Artifizielle Störungen“ des Deutschen Kollegiums für Psychosomatische Medizin (DKPM) unter der Leitung von Hans Willenberg und Annegret Eckhardt-Henn einen typologischen Zugang [55;56;100;101]. Er orientiert sich an klinisch relevanten Typen offenen und heimlichen, direkten und indirekten selbstschädigenden Verhaltens. Unterschieden werden drei Prägnanztypen pathologisch selbstschädigenden Verhaltens. Diese sind (1) die direkte Selbstschädigung, bei der sich die Person selbst unmittelbaren äußerlichen körperlichen Schaden zufügt, indem sie sich beispielsweise schneidet, brennt, schlägt oder Haar ausreißt, (2) die artifiziell induzierte Erkrankung, bei der sich die Person eine Krankheit zufügt, indem sie ihrem Körper beispielsweise toxische Substanzen, infektiöses Material oder kontraindizierte Medikamente zuführt, und (3) die delegierte Selbstschädigung, bei der die Person durch Vortäuschung von Beschwerden oder Symptomen das Medizinalsystem zu Risiko behafteten, meist invasiven Maßnahmen veranlasst, die schädlich sein können, wobei die schädigende Absicht der Person an den Behandler ohne dessen Wissen delegiert wird.

Der Vorteil dieser Systematik liegt darin, dass sie eine Überschneidung zwischen den drei Typen zulässt und dass sie sämtliche klinisch relevanten Formen autodestruktiven Verhaltens, also sowohl offene als auch heimliche abzubilden versucht.

In einer eigenen Arbeit, die einen Überblick über verhaltenstherapeutische Ansätze zu diesem Störungsspektrum gibt, wurden anhand von zwei Falldarstellungen aus der eigenen Behandlung zwei unterschiedliche Prägnanztypen exemplarisch dargestellt, eine Patientin mit überwiegend direktem selbstschädigenden Verhalten und eine Patientin mit einer Kombination aus überwiegend delegierter Selbstschädigung mit heimlich induzierten Erkrankungen [102]. Dabei wurden die Gegensätze wie auch die Überschneidungen zwischen diesen Typen verdeutlicht. Über die Wendung aggressiver Impulse gegen den eigenen Körper weisen die Fälle gemeinsame Merkmale auf. Daneben finden sich einige Gemeinsamkeiten in der zugrunde liegenden Psychopathologie. Ein maßgeblicher Unterschied für den therapeutischen Zugang ist das Ausmaß der Verheimlichungstendenz, worauf in der Arbeit näher eingegangen wird.

Der beschriebene typologische Zugang der DKPM-Arbeitsgruppe war auch die Grundlage für die Entwicklung eines Basisdokumentationssystems [56]. Dabei handelt es sich um ein Fremdbeurteilungsverfahren, das durch den Diagnostiker nach Anamnese, Akteneinsicht und klinischer Untersuchung auszufüllen ist. In dem Instrument werden die drei beschriebenen Typen autodestruktiven Verhaltens dokumentiert und es wird erfasst, ob das Verhalten offen thematisiert oder geheim gehalten wird. Quantitative Aspekte werden über das Ausmaß der Schädigung separat berücksichtigt, indem Schweregrade und Folgen eingestuft werden. Zusätzlich werden frühe Verhaltensvorläufer und quantitative Auffälligkeiten in der medizinischen Inanspruchnahme dokumentiert. Das im Expertenkonsens entwickelte klinische Fremdbeurteilungsverfahren war bis dato nicht an klinischen Stichproben eingesetzt und validiert worden.

In den eigenen Arbeiten wurde dieses Fremdbeurteilungsverfahren eingesetzt. Es wurden Inter-Rater-Übereinstimmungen und Test-Retest-Übereinstimmungen ermittelt. Die Inter-Rater-Übereinstimmungen liegen in einem zufrieden stellenden Bereich, die Test-Retest-Übereinstimmungen in einem noch akzeptablen. Für die Werte ist in Rechnung zu stellen, dass es sich um ein Phänomen mit niedrigen Basisraten handelt, was Inter-Rater-Übereinstimmungen grundsätzlich verringert [103], und dass in Studien zu den

Forschungskriterien der ICD-10 Klassifikation eine Fallvignette eines Patienten mit artifizieller Störung auch „nur“ mit 53% Übereinstimmung richtig diagnostiziert wurde [104].

Über Zusammenhänge mit Selbstauskunftsverfahren zu der Thematik wurden erste Belege für die Validität des Fremdbeurteilungsverfahrens ermittelt [70].

Das Instrument wurde – unter kontinuierlicher Supervision der Diagnostiker – in einer unselektierten Klinikstichprobe psychosomatischer Patienten über 1 Jahr eingesetzt und es wurden Raten für die Häufigkeit autodestruktiven Verhaltens in diesem Patientenkollektiv ermittelt [59]. Um die eigenen Häufigkeitsquoten in den Kontext anderer Befunde zu der Thematik einordnen zu können, wurde in einer separaten Arbeit eine systematische Literaturübersicht zum Vorkommen autodestruktiver Syndrome erstellt [105].

Neben dem hohen Anteil autodestruktiver Patienten in der eigenen Stichprobe psychosomatischer Patienten (7,5%; davon 40% Verdachtsfälle) war ein relevantes Ergebnis, dass sich in knapp 40% dieser Fälle Überschneidungen zwischen den Typen autodestruktiven Verhaltens fanden [59]. Dies lässt sich als Argument dafür werten, die verschiedenen Formen pathologisch autodestruktiven Verhaltens nicht als trennscharf voneinander abgegrenzte Störungskategorien zu konzipieren. Für diesen Ansatz spricht auch, dass es Fälle gibt, in denen sowohl offenes als auch heimliches autodestruktives Verhalten dokumentiert wurde. Auch fiel die Differenzierung in direkte Selbstschädigung und selbstinduzierte Erkrankung nicht notwendigerweise mit der Differenzierung in offenes und heimliches Verhalten zusammen. So gab es sowohl Fälle, in denen direkte Selbstschädigung verheimlicht wurde, als auch die Fälle, in denen eine selbstinduzierte Erkrankung offen kommuniziert wurde.

Im nächsten Schritt wurden psychopathologische Auffälligkeiten der Gruppe selbstschädigender Patienten gegenüber der Gruppe psychosomatischer Patienten ohne Hinweise auf Autodestruktion ermittelt [106;107]. Vor dem Hintergrund von Modellen zur Selbststeuerung und Selbstwertregulation wurden verschiedene Hypothesen formuliert und überprüft, die Störungen in diesen Bereichen für autodestruktive Patienten postulieren. Es wurde zudem postuliert, dass sich eine verringerte Selbstwirksamkeitserwartung sowie

Störungen in der Selbstwertregulation zwar bei Patienten mit offenkommunizierter Selbstschädigung zeigen, sich aber nicht bei Patienten mit einer Verheimlichungstendenz abbilden lassen. Die eigenen empirischen Befunde bestätigten diese Annahmen.

Unter diagnostischen Gesichtspunkten war zudem bedeutsam, inwieweit sich die gefundenen Häufigkeitsraten replizieren ließen. Hierzu wurde der gleiche Fremdbeurteilungsbogen in der gleichen Klinik ein weiteres Jahr in der Routinediagnostik eingesetzt. Allerdings wurde die themenspezifische Supervision für die Diagnostiker abgesetzt, so dass anzunehmen war, dass die zuvor erhöhte diagnostische Aufmerksamkeit wieder etwas absank. Tatsächlich wurden Fälle autodestruktiven Verhaltens in der zweiten Studienphase nur noch halb so oft diagnostiziert.

Für die Diagnostik direkt selbstschädigenden Verhaltens („deliberate self-harm behavior“) wurden im angloamerikanischen Raum verschiedene Selbstauskunftsverfahren publiziert, die bislang nicht für den deutschen Sprachraum verfügbar waren [108-111]. Aus diesen Verfahren wurden diejenigen ausgewählt, die – im Einklang mit dem hier vertretenen Ansatz – ein möglichst breites Spektrum pathologisch autodestruktiven Verhaltens abdecken sollten und die im Original zufrieden stellend teststatistisch überprüft waren [108;111].

Die beiden Instrumente wurden – unter Mitarbeit der Originalautoren – für den deutschen Sprachraum adaptiert und an Stichproben psychosomatischer Patienten validiert. Die Konstruktvalidität wurde aneinander und an Ergebnissen in der soeben beschriebenen Fremdbeurteilung belegt. Es finden sich insgesamt gute psychometrische Kennwerte.

Die Häufigkeitsraten, die mit den Selbstauskunftsverfahren in der Stichprobe psychosomatischer Patienten ermittelt wurden, lagen höher als in ähnlichen Untersuchungen an psychiatrischen Patienten. Es fanden sich keine Geschlechtsunterschiede in der Häufigkeit selbstschädigenden Verhaltens, lediglich in der Häufigkeit spezifischer Formen.

Es wurden verschiedene Hypothesen zu psychopathologischen Hintergründen selbstschädigenden Verhaltens geprüft. Erwartungsgemäß zeigte sich, dass Patienten mit selbstschädigendem Verhalten gegenüber solchen ohne dieses Verhalten höhere Werte für Angst, Depression, Feindseligkeit, Paranoidität und Stresserleben aufweisen, während ihre Selbstwirksamkeitserwartung und ihr Optimismus geringer ausgeprägt sind.