

BEITRÄGE ZUR SOZIALEN SICHERHEIT

*Bericht im Rahmen des zweiten mehrjährigen
Forschungsprogramms zur Invalidenversicherung (FoP2-IV)*

Erhebung der formalen Qualität psychiatrischer Gutachten

Forschungsbericht Nr. 2/12



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Département fédéral de l'intérieur DFI
Bundesamt für Sozialversicherungen BSV
Office fédérale des assurances sociales OFAS

Das Bundesamt für Sozialversicherungen veröffentlicht in seiner Reihe "Beiträge zur Sozialen Sicherheit" konzeptionelle Arbeiten sowie Forschungs- und Evaluationsergebnisse zu aktuellen Themen im Bereich der Sozialen Sicherheit, die damit einem breiteren Publikum zugänglich gemacht und zur Diskussion gestellt werden sollen. Die präsentierten Folgerungen und Empfehlungen geben nicht notwendigerweise die Meinung des Bundesamtes für Sozialversicherungen wieder.

Autoren/Autorinnen: Ebner, Gerhard; Dittmann, Volker; Mager, Ralph; Stieglitz, Rolf-Dieter; Träbert, Silke; Bührlen, Bernhard; Herdt, Jörg
Universitäre Psychiatrische Kliniken Basel (UPK)
Willhelm Klein-Strasse 27
4012 Basel
Tel. +41 (0) 61 325 51 16 / Fax +41 (0) 61 325 57 11
E-mail: joerg.herdt@upkbs.ch
Internet: www.upkbs.ch

Auskünfte: Martin Wicki
Forschung und Evaluation
Bundesamt für Sozialversicherungen
Effingerstrasse 20
3003 Bern
Tel. +41 (0) 302 90 02
E-mail: martin.wicki@bsv.admin.ch

ISSN: 1663-4659

Copyright: Bundesamt für Sozialversicherungen, CH-3003 Bern
Auszugsweiser Abdruck – ausser für kommerzielle Nutzung –
unter Quellenangabe und Zustellung eines Belegexemplares
an das Bundesamt für Sozialversicherungen gestattet.

Vertrieb: BBL, Vertrieb Publikationen, CH - 3003 Bern
<http://www.bundespublikationen.admin.ch>

Bestellnummer: 318.010.2/12

**Entwicklung von Leitlinien zur Begutachtung von
psychischen Behinderungen.
Erhebung der formalen Qualität psychiatrischer
Gutachten**

Schlussbericht

15. Dezember 2011

Gerhard Ebner
Volker Dittmann
Ralph Mager
Rolf-Dieter Stieglitz
Silke Träbert
Bernhard Bührlen
Jörg Herdt

unter Mitarbeit von May Helmy und Michael Stiebel



Vorwort des Bundesamtes für Sozialversicherungen

Invalidität aus psychischen Gründen hat in den letzten rund 20 Jahren stark an Bedeutung zugenommen. Personen mit psychischen Störungen haben ein hohes Risiko, ihre Stelle zu verlieren und ihre berufliche Eingliederung, wenn sie einmal aus dem Erwerbsleben ausgetreten sind, gestaltet sich besonders schwierig. Es ist deshalb wichtig, dass die Abklärung dieser versicherten Personen mit grosser Sorgfalt durchgeführt wird.

Der Begutachtungsprozess ist im psychiatrischen oftmals schwieriger als im somatischen Bereich, da die Qualität der klinischen Untersuchung in der Psychiatrie, im Wesentlichen bestehend aus der Anamneseerhebung, der Symptomerfassung und der Verhaltensbeobachtung, stark von der Motivation und der Mitwirkung des Exploranden abhängt. Eine Objektivierung der Befunde, wie dies bei der somatischen Untersuchung mit bildgebenden Verfahren - z.B. Computer- oder Magnetresonanztomografie - und weiteren Methoden möglich ist, ist in der Psychiatrie nicht im selben Ausmass realisierbar. Eine besonders grosse Herausforderung bei psychischen Gesundheitsstörungen ist es, die funktionelle Leistungsfähigkeit einer versicherten Person zu beurteilen, da heute noch keine Leitlinien oder wissenschaftlich validierte Testverfahren vorliegen. Aus diesen Gründen weisen psychiatrische Gutachten oftmals grosse qualitative Mängel auf.

Da der Erfolg einer beruflichen Integration psychisch kranker Versicherter nicht zuletzt auch von der Qualität der versicherungspsychiatrischen Beurteilung abhängig ist, hat das BSV zusammen mit den Schweizerischen Gesellschaften für Psychiatrie und Psychotherapie sowie Versicherungspsychiatrie und der Suva unter dem Patronat der Verbindung der Schweizer Ärztinnen und Ärzte FMH konsensual Leitlinien für psychiatrische Gutachten in der Invalidenversicherung erarbeitet. Während dieses Entwicklungsprozesses gab das BSV im Rahmen des Forschungsprogramms zur Invalidenversicherung (FoP-IV) eine IST-Analyse der Qualität von versicherungspsychiatrischen Gutachten der IV in Auftrag. Die analysierten Gutachten aus den Jahren 2008 und 2009 stammen aus allen Teilen der Schweiz und von verschiedenen Abklärungsstellen (Regionalen Ärztlichen Diensten RAD oder externen medizinischen Abklärungsstellen MEDAS).

Die Resultate der Studie zeigen etliche Lücken bezüglich der Untersuchungen, insbesondere bezüglich der Fragen zum Grad der selbständig verrichtbaren Tätigkeiten sowie zur Wahrnehmung und zum Umgang der Versicherten mit ihrer Erkrankung. Eine versicherungsmedizinische Würdigung der Ressourcen hinsichtlich Funktionen und Fähigkeiten fehlt ebenfalls häufig. Bei den polydisziplinären Gutachten werden die Teilgutachten oft unangemessen berücksichtigt. Angesichts solcher Mängel fokussiert die vorliegende Studie auf die Bewertung des Gutachtenaufbaus als Ausgangsbasis für die Entwicklung formaler Leitlinien, die von Sozialversicherungsgerichten häufig gefordert werden. Für eine fundierte Beurteilung müssten die zugrundeliegenden Kriterien jedoch noch nach deren Bedeutung hinsichtlich ihrer Aussagekraft gewichtet werden.

Die Studie verweist auf grosse Unterschiede hinsichtlich der Berücksichtigung der bereits publizierten und in Fachkreisen breit anerkannten Leitlinien der Schweizerischen Gesellschaft für Versicherungspsychiatrie (SGVP) oder der deutschen Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften.

Es kann gefolgert werden, dass die Leitlinien im Hinblick auf eine Verbesserung der Qualität und der Reproduzierbarkeit der Begutachtungen einem grossen Bedürfnis entsprechen. Wichtig ist, dass die Einführung der Leitlinien in engem Kontakt zwischen der IV und den Ärzten und ihren Standesorganisationen stattfindet. Die Evaluation ist nach einer gewissen Zeit – zu denken ist an etwa drei bis fünf Jahre – zu wiederholen um die Fortschritte auf dem angestrebten Weg dokumentieren und allfälligen weiteren Verbesserungsbedarf ausfindig machen zu können.

Stefan Ritler, Vizedirektor

Leiter Geschäftsfeld Invalidenversicherung

Avant-propos de l'Office fédéral des assurances sociales

Le nombre de cas d'invalidité pour raisons psychiques a fortement augmenté au cours des vingt dernières années. Les assurés atteints de troubles psychiques ont un risque élevé de perdre leur emploi et, lorsqu'ils ont dû interrompre leur activité, leur réinsertion professionnelle peut être très ardue. Il est donc essentiel que leur dossier soit instruit avec le plus grand soin.

Le processus d'expertise pose souvent plus de difficultés dans le domaine de la psychiatrie que pour la médecine somatique, car la qualité de l'examen clinique – comprenant pour l'essentiel l'anamnèse, l'évaluation des symptômes et l'observation du comportement – dépend fortement de la motivation et de la collaboration du sujet. Une objectivation des constats, possible pour les examens somatiques grâce à des procédés d'imagerie (par ex. scanner ou IRM) et à d'autres méthodes, n'est pas réalisable dans une mesure comparable en psychiatrie. Une des principales difficultés rencontrées avec les assurés atteints de troubles psychiques est d'évaluer leur capacité fonctionnelle, car il n'existe pas encore à l'heure actuelle de lignes directrices ou de tests scientifiquement validés. C'est pour ces raisons que la qualité des expertises psychiatriques laisse souvent à désirer.

Convaincu que le succès de la réinsertion professionnelle des assurés concernés dépend dans une mesure non négligeable de la qualité des expertises psychiatriques faites pour l'assurance, l'OFAS a élaboré de concert avec la Société suisse de psychiatrie et psychothérapie, la Société suisse de psychiatrie des assurances et la SUVA, réunies sous le patronat de la Fédération des médecins suisses (FMH), des directives consensuelles relatives à la qualité des expertises psychiatriques dans l'assurance-invalidité. En parallèle, il a demandé à un mandataire, dans le cadre du programme de recherche sur l'assurance-invalidité, d'analyser la qualité actuelle des expertises psychiatriques faites pour l'AI. Les expertises analysées, datant de 2008 ou 2009 et provenant de toutes les régions de Suisse, ont été établies par des centres de deux types : services médicaux régionaux (SMR), à l'interne, et centres d'observation médicale de l'AI (COMAI), à l'externe.

Les résultats de cette étude pointent certaines lacunes des examens, notamment pour ce qui est des questions touchant le pourcentage d'activité que l'assuré peut assumer de façon autonome, ou la perception qu'il a de sa maladie et son attitude à l'égard de celle-ci. Une évaluation médico-assurantielle des atteintes fonctionnelles et de la réduction de capacité qui en découle fait aussi fréquemment défaut. Par ailleurs, les expertises pluridisciplinaires ne tiennent souvent pas suffisamment compte des expertises partielles. Au vu de ce genre de lacunes, la présente étude se concentre sur l'évaluation de la structure des expertises, vue comme base pour l'élaboration de directives de forme, dont les tribunaux des assurances sociales ont souvent déploré l'absence. Pour une évaluation solide, il conviendrait cependant d'apprécier l'importance des critères fondant ces directives sous l'angle de leur force probante.

L'étude s'est bornée à évaluer la structuration formelle des expertises. Elle relève à cet égard de grandes différences dans la prise en compte des directives déjà publiées et largement reconnues par la profession, qu'il s'agisse de celles de la Société suisse de psychiatrie des assurances (SGVP) ou de l'Académie allemande des sciences médicales.

Les nouvelles directives répondent assurément à un grand besoin en vue d'améliorer la qualité et la reproductibilité des expertises. Il est important que leur mise en place se fasse en étroite coopération entre l'AI et les médecins et leurs organisations professionnelles. Il conviendra de répéter l'évaluation après un certain temps – de l'ordre de trois à cinq ans – afin de documenter les progrès réalisés pour atteindre l'objectif visé et de repérer les besoins d'amélioration qui pourraient subsister.

Stefan Ritler, vice-directeur

Chef du domaine Assurance-invalidité

Premessa dell'Ufficio federale delle assicurazioni sociali

Le invalidità causate da motivi psichici sono aumentate in maniera significativa negli ultimi 20 anni. Le persone affette da disturbi psichici sono fortemente soggette al rischio di perdere il proprio posto di lavoro e, una volta uscite dal mondo del lavoro, la loro integrazione professionale risulta particolarmente difficile. È quindi importante che per questi assicurati l'accertamento sia svolto molto accuratamente.

Nell'ambito psichiatrico, la procedura peritale è spesso più difficile che in quello somatico, poiché in psichiatria la qualità dell'esame clinico – che consiste fondamentalmente nella ricostituzione dell'anamnesi, la rilevazione dei sintomi e l'osservazione del comportamento – dipende molto dalla motivazione e dalla collaborazione dell'assicurato. Se negli esami somatici con diagnostica per immagini (p. es. tomografia a risonanza magnetica o computerizzata) e altri metodi è possibile ottenere riscontri oggettivi, in psichiatria ciò non è fattibile nella stessa misura. Un compito particolarmente difficile in presenza di disturbi psichici è la valutazione della capacità funzionale degli assicurati, dato che a tutt'oggi non esistono ancora direttive o test scientificamente validati. Per questi motivi, le perizie psichiatriche presentano spesso gravi lacune qualitative.

Dal momento che la riuscita dell'integrazione professionale degli assicurati con malattie psichiche dipende non da ultimo anche dalla qualità delle valutazioni psichiatrico-assicurative, l'UFAS ha elaborato delle linee direttive per le perizie psichiatriche nell'assicurazione invalidità, d'intesa e in collaborazione con la Società svizzera di psichiatria e psicoterapia, quella di psichiatria assicurativa e la Suva, con il patrocinio della Federazione dei medici svizzeri (FMH). Nel corso di questo processo di sviluppo, nel quadro del Programma di ricerca sull'assicurazione invalidità (PR-AI), l'UFAS ha commissionato un'analisi della situazione attuale della qualità delle perizie psichiatrico-assicurative dell'AI. Le perizie esaminate sono state eseguite nel 2008 e 2009, in tutta la Svizzera e da diversi centri di accertamento (servizi medici regionali [SMR] o servizi di accertamento medico [SAM] esterni).

Dai risultati dello studio emergono diverse lacune negli esami, in particolare per quanto riguarda il grado delle attività che gli assicurati possono svolgere autonomamente e il modo in cui questi percepiscono e affrontano la loro malattia. Inoltre, spesso manca l'apprezzamento medico-assicurativo delle loro risorse funzionali e capacità. Nell'ambito delle perizie pluridisciplinari, sono frequenti i casi in cui spesso non si tiene adeguatamente conto delle perizie dei singoli specialisti. In considerazione di tali lacune, il presente studio si concentra sulla valutazione della struttura delle perizie come base per lo sviluppo di linee direttive formali, spesso richieste dai tribunali delle assicurazioni sociali. Per permettere una valutazione fondata, i criteri delle linee direttive dovrebbero comunque essere ponderati in base alla loro incidenza.

Lo studio ha rilevato notevoli differenze quanto all'osservanza delle linee direttive già pubblicate e ampiamente riconosciute negli ambienti specializzati, come quelle della Società svizzera di psichiatria assicurativa e del gruppo di lavoro delle società medico-scientifiche tedesche (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften).

Si può concludere che le linee direttive rispondono a un forte bisogno di miglioramento della qualità e della riproducibilità delle perizie. L'introduzione di questi strumenti richiederà una stretta collaborazione tra l'AI, i medici e le loro associazioni di categoria. Dopo un certo periodo di tempo (presumibilmente fra tre - cinque anni) si dovrà ripetere la valutazione, in modo da poter documentare i progressi compiuti ed individuare gli eventuali correttivi necessari.

Stefan Ritler, vicedirettore

Responsabile dell'Ambito Assicurazione invalidità

Foreword by the Federal Social Insurance Office

Invalidity for psychological reasons has moved into the spotlight over the last two decades. People with psychiatric disorders are at high risk of losing their jobs. Once they are out of the employment system, it can be very difficult for them to return to work. It is therefore important to examine the situation of these insured people very closely.

It is often more difficult to assess psychiatric disorders than somatic illnesses, as the quality of clinical assessments in psychiatry, which primarily consist of anamnesis, assessing symptoms and observing behaviour, is highly dependent on the patient's motivation and collaboration. Findings of psychiatric assessments cannot be objectified to the same extent as those of somatic examinations using imaging techniques, e.g. X-ray computed tomography or magnetic resonance imaging, and other methods. A particular challenge in the case of psychiatric disorders is to assess the functional capacity of an insured person, as no guidelines or scientifically validated testing procedures are currently available. For these reasons, psychiatric assessments are often seriously lacking in terms of quality.

As the success of integrating the mentally ill into the workforce is also contingent on the quality of psychiatric assessments, the FSIO has developed a set of guidelines for psychiatric assessments for invalidity insurance purposes in collaboration with the Swiss Society for Psychiatry and Psychotherapy (SGPP), the Swiss Insurance Psychiatric Society (SGVP) and the insurer Suva, under the auspices of the Swiss Medical Association (FMH). During this development process, the FSIO also commissioned a status analysis of the quality of psychiatric assessments for invalidity insurance purposes as part of the FOP-IV research programme on invalidity. The assessments from 2008 and 2009 that were analysed originate from all parts of Switzerland and various evaluation services (regional medical services [RAD] and external medical institutes for invalidity assessment [MEDAS]).

The results of the study reveal numerous gaps in the assessments, particularly concerning questions on the extent to which patients are capable of carrying out tasks themselves, and on their perception and management of their condition. An appraisal for medical insurance purposes of the patient's incapacity and functional impairment is also often not carried out. With polydisciplinary assessments, the partial assessments are often not taken adequately into consideration. On the basis of these shortcomings, this study focuses on the evaluation of the assessment structure as the starting point for developing formal guidelines which are often required by social insurance courts. However, to assess the situation properly, it would be necessary to weight the criteria according to their importance with regard to the extent to which they are meaningful.

The study highlights considerable differences in the extent to which the guidelines published by the Swiss Insurance Psychiatric Society (SGVP) and the German Association of Scientific Medical Societies (AWMF), which are widely accepted by specialists, are applied.

In conclusion, the existing guidelines have vast potential for improvement in terms of the quality and reproducibility of assessments. It is therefore important that the guidelines are implemented under close consultation between the invalidity insurance scheme and the physicians and their professional organizations. Following a certain period, say three to five years, it will be necessary to evaluate the guidelines again to document the progress made and to pinpoint any additional areas for improvement.

Stefan Ritler, Deputy Director

Head of Invalidity Insurance

Inhalt

Danksagung	v
Zusammenfassung	vii
Résumé	xi
Riassunto	xv
Summary	xix
1 Einleitung	1
1.1 Ausgangslage	1
1.2 Übersicht zum vorliegenden Bericht	2
2 Fragestellungen	2
3 Methodisches Vorgehen	3
3.1 Studiendesign	3
3.2 Datenerhebung und Datenerfassung	3
3.3 Gutachtenstichprobe	4
3.3.1 Grundgesamtheit	4
3.3.2 Stichprobe	5
3.4 Operationalisierung und Datenstruktur	7
3.5 Auswertung der Gutachten	8
3.5.1 Statistische Verfahren	8
3.6 Interrater-Reliabilität	9
4 Ergebnisse aus der Gesamtstichprobe	10
4.1 Basisdaten	10
4.1.1 Gutachten	10
4.1.2 Explorandinnen und Exploranden	12
4.2 Obligatorische Kriterien	14
Themenbereich 1. „Ausgangslage und Formelles“	15
Themenbereich 2. „Aktenauszug“	18
Themenbereich 3. „Untersuchung“	19
Themenbereich 4. „Befund“	24
Themenbereich 6. „Medizinische und versicherungsmedizinische Beurteilung“	25
4.3 Erfüllung der obligatorischen Kriterien auf Ebene der Bereiche	27
4.4 Qualitätsunterschiede zwischen den einzelnen Gutachtenteilen	32
4.5 Unterschiede zwischen Arten von Gutachten	36
4.6 Fakultative Kriterien	37
4.6.1 Übersicht über alle fakultativen Kriterien	37
4.6.1 Beurteilung bzw. Einordnung der Befunde bei polydisziplinären Gutachten	40
5 Qualitätsunterschiede zwischen den Gutachten verschiedener Institutionen	43
5.1 Unterschiede in einzelnen Themenbereichen	43
Themenbereich 1. „Ausgangslage und Formelles“	43
Themenbereich 2. „Aktenauszug“	44
Themenbereich 3. „Untersuchung“	44
Themenbereich 4. „Befund“	45
Themenbereich 6. „Medizinische und versicherungsmedizinische Beurteilung“	46
5.2 Zusammenfassung über alle Kriterien	46

6	Qualitätsunterschiede zwischen den Gutachten aus verschiedenen Regionen	47
6.1	Unterschiede in einzelnen Themenbereichen	47
	Themenbereich 1. „Ausgangslage und Formelles“	47
	Themenbereich 2. „Aktenauszug“	49
	Themenbereich 3. „Untersuchung“	49
	Themenbereich 4. „Befund“	50
	Themenbereich 6. „Medizinische und versicherungsmedizinische Beurteilung“	51
6.2	Zusammenfassung über alle obligatorischen Kriterien	53
7	Diskussion der Ergebnisse	54
7.1	Bemerkungen zur Stichprobe	54
7.2	Allgemeine Bemerkungen zur Interpretation der Ergebnisse	54
7.3	Aktuelles durchschnittliches Qualitätsniveau hinsichtlich formaler Begutachungskriterien	55
7.4	Durchschnittliche formale Qualität der einzelnen Gutachtenteile und systematische Qualitätsunterschiede zwischen diesen	56
7.5	Qualitätsunterschiede zwischen den Gutachten verschiedener Institutionen	58
7.6	Systematische Qualitätsunterschiede zwischen den Gutachten aus verschiedenen Regionen	58
7.7	Unterschiede zwischen Arten von Gutachten	59
8	Schlussfolgerungen und Empfehlungen	59
8.1	Konsequenzen für die weitere Leitlinienentwicklung	60
8.2	Konsequenzen für die Leitlinienimplementierung	61
8.3	Empfehlungen für die weitere Evaluation der Leitlinienentwicklung und – implementierung	62
Anhang		63
	Anhang 1: Checkliste	63
	Anhang 2: Kodierregeln	69
	Anhang 3: Verwendete Literatur	80

Abbildungsverzeichnis

Abbildung 4-1: Beauftragender RAD/SMR.....	11
Abbildung 4-2: Art des Gutachtens	12
Abbildung 4-3: Altersverteilung	13
Abbildung 4-4: Angaben zum Exploranden.....	16
Abbildung 4-5: Angaben zum Gutachter	16
Abbildung 4-6: Gutachtendaten	17
Abbildung 4-7: Anlass der Begutachtung.....	17
Abbildung 4-8: Eigene Untersuchungen und Befunde	18
Abbildung 4-9: Aktenauszug.....	19
Abbildung 4-10: Jetziges Leiden.....	20
Abbildung 4-11: Anamnesen	21
Abbildung 4-12: Soziale Anamnese	21
Abbildung 4-13: Werdegang.....	22
Abbildung 4-14: Berufliche Tätigkeiten/Arbeitsversuche.....	22
Abbildung 4-15: Alltagsleben	23
Abbildung 4-16: Behandlungen	23
Abbildung 4-17: Zukunft	24
Abbildung 4-18: Verhaltensbeobachtungen.....	24
Abbildung 4-19: Psychiatrischer Befund	25
Abbildung 4-20: Resumé	26
Abbildung 4-21: Diskussion/Würdigung	26
Abbildung 4-22: Bereich Ausgangslage in der Gesamtstichprobe.....	27
Abbildung 4-23: Bereich Aktenauszug in der Gesamtstichprobe	28
Abbildung 4-24: Bereich Untersuchung in der Gesamtstichprobe	29
Abbildung 4-25: Bereich Befund in der Gesamtstichprobe	30
Abbildung 4-26: Bereich Beurteilung in der Gesamtstichprobe.....	31
Abbildung 4-27: Zusammenfassung aller obligatorischen Kriterien in der Gesamtstichprobe.....	32
Abbildung 4-28: Unterschiede zwischen Arten von Gutachten.....	37
Abbildung 4-29: Erfüllung der fakultativen Kriterien.....	38
Abbildung 4-30: Fakultative Kriterien bei polydisziplinären Gutachten.....	41
Abbildung 5-1: Bereich Ausgangslage bei verschiedenen Institutionen.....	44
Abbildung 5-2: Bereich Untersuchung bei verschiedenen Institutionen	45
Abbildung 5-3: Bereich Befund bei verschiedenen Institutionen.....	45
Abbildung 5-4: Beurteilung bei verschiedenen Institutionen.....	46
Abbildung 5-5: Zusammenfassung aller obligatorischen Kriterien bei verschiedenen Institutionen	47
Abbildung 6-1: Bereich Ausgangslage in den verschiedenen Regionen	48
Abbildung 6-2: Bereich Untersuchung in den verschiedenen Regionen.....	50
Abbildung 6-3: Bereich Befund in den verschiedenen Regionen	51
Abbildung 6-4: Bereich Beurteilung in den verschiedenen Regionen.....	52
Abbildung 6-5: Zusammenfassung aller obligatorischen Kriterien in den Regionen	53

Tabellenverzeichnis

Tabelle 3-1: Grundgesamtheit, Stand Nov. 2010	5
Tabelle 3-2: Gesamtzahl aller Gutachten	6
Tabelle 3-3: Verteilung der Gutachten auf RAD/SMR	7
Tabelle 4-1: Institution	11
Tabelle 4-2: Herkunftsland (grösste Gruppen).....	13
Tabelle 4-3: Erstdiagnosen nach ICD-10.....	14
Tabelle 4-4: Erfüllung der obligatorischen Kriterien: Ausgangslage.....	33
Tabelle 4-5: Erfüllung der obligatorischen Kriterien: Aktenauszug	33
Tabelle 4-6: Erfüllung der obligatorischen Kriterien: Untersuchung.....	34
Tabelle 4-7: Erfüllung der obligatorischen Kriterien: Befund.....	34
Tabelle 4-8: Erfüllung der obligatorischen Kriterien: Beurteilung.....	35
Tabelle 4-9: Unterschiede zwischen Gutachtenteilen.....	36
Tabelle 4-10: Erfüllung der fakultativen Kriterien.....	40
Tabelle 4-11: Erfüllung der fakultativen Kriterien für polydisziplinäre Gutachten	41

Danksagung

Die vorliegende Analyse von psychiatrischen Gutachten der Invalidenversicherung beim Bundesamt für Sozialversicherungen (BSV) wäre ohne die tatkräftige Unterstützung von verschiedenen Seiten nicht möglich gewesen. Zu allererst möchten wir den Leiterinnen und Leitern der Regionalärztlichen Dienste (RAD/SMR) und den Verantwortlichen bei den IV-Stellen danken, welche aus der Gesamtzahl ihrer Gutachten für uns die gewünschten Stichproben ermittelt und uns dann die Gutachten zur Auswertung zugesandt haben. Dieser teilweise sicherlich mühsamen Zusatzbelastung haben sich alle Stellen ohne Murren unterzogen und unsere Anfragen jederzeit mit grosser Geduld und Freundlichkeit beantwortet.

Danken wollen wir auch dem BSV für das mit diesem Projektauftrag ausgesprochene Vertrauen und insbesondere den Projektleitern Frau Dr. Rajower und Herrn Wicki für ihre freundliche Unterstützung und ihre fruchtbaren Anregungen.

Wir hoffen, dass wir den RAD/SMR, den IV-Stellen und dem BSV als Gegenleistung für diese Unterstützung nun unsererseits nützliche Ergebnisse liefern können, welche – letztlich vor allem zum Nutzen der Explorandinnen und Exploranden – einen Beitrag zu einer hochstehenden Qualität der psychiatrischen IV-Gutachten leisten können.

Zusammenfassung

Ausgangslage

Vorzeitiges bzw. teilweises Ausscheiden aus der Erwerbstätigkeit und Berentungen aufgrund einer psychischen Belastung haben eine grosse individuelle, gesellschaftliche und volkswirtschaftliche Bedeutung. Die psychiatrische Begutachtung bildet dabei einen zentralen Baustein.

Entsprechende IV-Gutachten werden von den IV-Stellen bei den Regionalärztlichen Diensten (RAD) bzw. Services Médicaux Régionaux/Servizio Medico Regionale (SMR), bei frei praktizierenden PsychiaterInnen, bei den Medizinischen Abklärungsstellen MEDAS oder bei Sonstigen Einrichtungen in Auftrag gegeben. Die RAD/SMR beurteilen sodann aufgrund der vorliegenden Unterlagen einschliesslich vorhandener Gutachten die medizinischen Aspekte der Anspruchsvoraussetzungen bei IV-Leistungsgesuchen. Je nach Fragestellung können die Gutachten mono-, bi- oder polydisziplinär angelegt sein.

Zur Sicherung der Qualität psychiatrischer Gutachten für die IV werden derzeit Leitlinien entwickelt, welche die Gutachterinnen und Gutachter in Zukunft über die Erwartungen, welche an ein optimales Gutachten gestellt werden, informieren sollen.

Im Expertenkonsens wurde bisher eine erste Fassung der Leitlinien für die formale Qualität der Gutachten entwickelt. Sie besteht aus 62 obligatorischen und 25 fakultativen einzelnen Kriterien und ist im Anhang zum Bericht zu finden. Die Kriterien definieren, welche Themenbereiche in einem Gutachten behandelt werden sollen, und sind dementsprechend nach diesen Themenbereichen gegliedert. Zudem befinden sich derzeit inhaltliche Leitlinien in der Entwicklung, welche über die formale Qualität hinaus festlegen werden, in welchem Umfang und in welcher Detaillierung die einzelnen Themen in den Gutachten bearbeitet werden sollen.

Die Einführung der Leitlinien wird wissenschaftlich begleitet, um deren Praxisnutzen zu evaluieren und zu optimieren. Dies geschieht in einem zweistufigen Design, bei dem zunächst die Qualität der Gutachten vor Einführung der Leitlinien erhoben wird, um später die Ergebnisse mit einer zweiten Messung nach der Einführung zu vergleichen. Der vorliegende Bericht stellt die Ergebnisse der ersten Erhebung dar.

Vorgehen

Nachdem das BSV die IV-Stellen über das geplante Projekt informiert hatte, wandte sich die Projektgruppe an alle RAD/SMR mit der Bitte, aus den ca. 15'000 psychiatrischen Gutachten der Jahre 2008 und 2009 per Zufall eine Stichprobe zu ziehen. Angestrebt wurde ein Umfang von 800 Gutachten, die alle RAD/SMR entsprechend deren Gesamtzahl von Gutachten gleichgewichtig repräsentieren. Damit ist gewährleistet, dass die gezogene Stichprobe für alle Regionen der Schweiz repräsentativ ist. Die Zufallsziehung stellt überdies sicher, dass die Stichprobe nicht hinsichtlich Gutachten-erstellender Institution, Sprache des Gutachtens oder etwa nach Diagnose des Exploranden/der Explorandin oder anderen Merkmalen verzerrt ist.

Insgesamt konnten 775 Gutachten ausgewertet werden. Darunter stammten 296 von frei praktizierenden Psychiatern bzw. Psychiaterinnen, 270 von den RAD, 123 von einem MEDAS und 86 von sonstigen Einrichtungen. Etwa die Hälfte der Gutachten war monodisziplinär, 15% bidisziplinär und ein knappes Drittel polydisziplinär.

Die vorgegebene Liste der formalen Qualitätskriterien wurde in eine Checkliste umgewandelt; für jedes Gutachten wurde in der Checkliste markiert, ob sich im Gutachten zumindest minimale Angaben zu den formalen Qualitätskriterien finden lassen konnten. Die Prüfung ergab für das Verfahren, welches eingesetzt wurde, um die Gutachten hinsichtlich der Erfüllung der formalen Kriterien zu

überprüfen, eine ausreichende Zuverlässigkeit. Dadurch ist eine gute Aussagekraft der im Folgenden berichteten Analysen gewährleistet.

Für den Hauptzweck der Studie, die Ausgangsmessung vor Einführung der Leitlinien, sind die Ergebnisse aus der Gesamtstichprobe aller Gutachten am wichtigsten. Ergänzend dazu wurden die Gutachten von den verschiedenen Typen von Institutionen sowie aus den verschiedenen Regionen (sprich: von den verschiedenen RAD/SMR) auf Unterschiede in der Vollständigkeit geprüft. Vollständigkeit ist dabei immer zu verstehen als Erfüllung möglichst vieler der formalen Kriterien, wobei berücksichtigt werden muss, dass die jetzt angewandten Kriterien den Gutachterinnen und Gutachtern beim Erstellen des Gutachtens noch nicht bekannt waren.

Die formalen Leitlinien für die Gutachtenerstellung gliedern sich nach sieben thematischen Einheiten (1. Ausgangslage und Formelles, 2. Aktenauszug, 3. Untersuchung, 4. Befund, 5. Angaben von Drittpersonen, 6. Medizinische und versicherungsmedizinische Beurteilung und 7. Beilagen), dementsprechend wurden die Gutachten auch für die vorliegende Analyse nach diesen sieben Themen bewertet, wobei die fakultativen Bereiche 5 und 7 nicht weiter vertieft wurden. Zudem wurde ein Gesamtwert gebildet, der alle 62 obligatorischen Kriterien einschliesst. Die Aggregation der Kriterien zu solchen Indexwerten ist allerdings aus methodischen und inhaltlichen Gründen nicht unproblematisch, die entsprechenden Ergebnisse müssen deshalb unter Berücksichtigung z.B. der teilweise kleinen Fallzahlen mit Zurückhaltung interpretiert werden. Hierfür wird auf die entsprechenden Kapitel zur Darstellung der Resultate und auf die im Bericht formulierten Schlussfolgerungen und Empfehlungen verwiesen. Die methodisch begründete Validität diene – gemeinsam mit der aus Sicht der Autoren besonderen praktischen Relevanz – auch dazu, die im Folgenden besonders hervorgehobenen Resultate auszuwählen. Eine Bewertung der Befunde erfolgt dann im Abschnitt „Fazit und Empfehlungen“.

Wichtigste Resultate

Die Ergebnisse aus der Gesamtstichprobe aller Gutachten belegen eine sehr breite Verteilung bzgl. der Berücksichtigung der formalen Kriterien: Drei Viertel der untersuchten Gutachten erfüllen jeweils mehr als zwei Drittel der formalen Kriterien, während wenige Gutachten nur die Hälfte der Kriterien oder weniger, ebenfalls wenige (fast) alle Kriterien berücksichtigen.

Bezüglich des Themenbereichs „**Aktenauszug**“ (Darstellung der im Gutachten verwendeten relevanten Vorakten) erreichen fast alle Gutachten hohe Werte.

Im Themenbereich „**Untersuchung**“ wurde vor allem das Krankheitskonzept des Exploranden, der Umgang mit seinen Beschwerden im Alltag, seine Zukunftsvorstellungen, die Sexualanamnese, der soziale Empfangsraum (Wohnung, Finanzen, Beziehungsnetz), das Datum und die Umstände der Arbeitsaufgabe, die Freizeitgestaltung, benötigte Hilfen im Haushalt und im Alltag, die aktuelle Behandlung und die Zukunftsvorstellungen des Exploranden nicht wie vom Kriterienkatalog gefordert dargestellt.

Beim „**Befund**“ wurden in weniger als 90% der Fälle Verhaltensbeobachtungen und die äussere Erscheinung beschrieben. Deutlich vernachlässigt wurde die Beschreibung der Persönlichkeit des Exploranden (22%).

Hinsichtlich der **medizinischen und versicherungsmedizinischen Beurteilung** berücksichtigen nur ca. 7% aller Gutachten alle Kriterien. Während das Kriterium „Beantwortung der Fragen“ in der Regel erfüllt wird, werden beispielsweise unter „Diskussion und versicherungsmedizinische Würdigung“ nur bei ca. zwei Dritteln der Gutachten Funktions- und Fähigkeitsstörungen hergeleitet, ansonsten werden sie häufig nur als blosser Feststellung dokumentiert. Auch die Informationen zu vorhandenen Ressourcen des Exploranden, zu seiner Kooperationsbereitschaft und die Integration aller Einschätzungen und Informationen liegen teilweise weit unter 70%. In lediglich fünf der untersuchten Gutachten wurde dokumentiert, auf welcher Grundlage die Beurteilung der Schwere einer Störung und ihre Auswirkungen auf die „zumutbare“ Arbeitsfähigkeit erfolgten.

Auf der Ebene der **einzelnen Kriterien** wurde das Kriterium „Beschreibung der Persönlichkeit“ am seltensten erfüllt. In weniger als 50% der Gutachten fanden die Kriterien „Eigene Untersuchungen: Untersuchungsdauer“, „Spontane Angaben des Exploranden / offenes Interview“, „Sexualität“, „Zukunftsvorstellungen bzgl. Krankheitsverlauf“, „Zukunftsvorstellungen allgemein“, „Beschreibung der Persönlichkeit“ sowie die „Kritische Würdigung der Selbsteinschätzung des Exploranden“ Erwähnung.

Die **Unterschiede zwischen den begutachtenden Einrichtungen** sind insgesamt nicht gross. Tendenziell, wenn auch nicht in allen Themenbereichen, sind die Gutachten der RAD etwas weniger ausführlich bzw. vollständig als die Gutachten aus anderen Einrichtungstypen. MEDAS-Gutachten sind in der Regel am vollständigsten; bei diesen handelt es sich aber auch besonders häufig um polydisziplinäre Gutachten mit mehreren Teilgutachten, sodass sich ein grösserer Umfang und damit auch eine grössere Vollständigkeit der MEDAS-Gutachten teilweise schon aus der Art der Gutachten ergeben könnte. Weitere Gründe für die Unterschiede lassen sich aus den vorliegenden Daten nicht entnehmen.

Aus den Ergebnissen zu **Gruppenunterschieden zwischen Gutachten aus verschiedenen Regionen** lässt sich kein klares Muster der Gutachtenqualität in den Regionen erkennen. Es lassen sich zwar regionale Unterschiede feststellen, die Rangreihe der Regionen mit den am vollständigsten berücksichtigten Kriterien variiert jedoch je nach betrachtetem Themenbereich.

Bei den verschiedenen **Arten von Gutachten** zeigen sich die polydisziplinären Gutachten gesamthaft als vollständiger als die mono- und bidisziplinären Gutachten. Insbesondere betrifft dies die Themenbereiche „Ausgangslage“ und „Untersuchung“. Dies scheint von daher erwartungsgemäss, als polydisziplinäre Gutachten ohnehin ausführlicher angelegt sind als z.B. monodisziplinäre, insbesondere was die Untersuchung angeht. Monodisziplinäre Gutachten berücksichtigen dagegen im Themenbereich „Befunde“ die meisten Kriterien. Interessanterweise unterscheiden sich die Gutachtenarten nicht in der „Medizinischen und versicherungsmedizinischen Beurteilung“, die grössere Vollständigkeit der polydisziplinären Gutachten bei der Untersuchung schlägt sich also nicht in einer ausführlicheren Beurteilung nieder.

Fazit und Empfehlungen

Ohne das Vorliegen der inhaltlichen Kriterien mit einer Gewichtung der Kriterien nach ihrer Relevanz für die versicherungsmedizinische Beurteilung eines Falles, und ohne die Zweitmessung zur Verlaufsbeurteilung zur Verfügung zu haben, ist die Interpretation der bisherigen Ergebnisse nur der Tendenz nach möglich. Auftragsgemäss wurde prioritär die Erfüllung aller obligatorischen formalen Kriterien als Soll-Wert angesehen. Diese rein formale Sichtweise vernachlässigt aber, dass die Kriterien schon per se eine unterschiedliche Bedeutung für die Beurteilung eines Falls haben, und dass diese Bedeutung auch noch in Abhängigkeit von der Problemkonstellation eines individuellen Falles variieren kann. So ist etwa das vollständige Vorhandensein bzw. Nichtvorhandensein der Adresse für die versicherungs-psychiatrische Beurteilung selbstverständlich weniger bedeutsam als das Kriterium „Beruflicher Werdegang“. Für eine umfassende Bewertung der Qualität und Angemessenheit eines Gutachtens für die individuell vorliegenden Fragestellungen sind also die Gewichtungen bzw. Priorisierungen erforderlich, die sich mit den inhaltlichen Qualitätskriterien erst noch in der Erarbeitung befinden.

Nichtsdestoweniger lässt sich folgendes **Fazit** ziehen:

Die Analyse der Gutachten zeigt, dass einzelne Kriterien, insbesondere solche aus dem psychosozialen Themenspektrum, seltener behandelt und weniger berücksichtigt wurden als die im engeren Sinn „medizinischen“ Kriterien. Verbesserungspotenzial besteht bei der Fundierung der versicherungspsychiatrischen Beurteilung, insbesondere hinsichtlich der Dokumentation und nachvollziehbaren Ableitung des Grades der Arbeitsfähigkeit. Deutlicher Verbesserungsbedarf zeigt sich ausserdem bei polydisziplinären Gutachten hinsichtlich der angemessenen Integration der Teilgutachten.

Aus Unterschieden zwischen begutachtenden Einrichtungen sowie aus regionalen Unterschieden in der Gutachtenqualität können Implikationen für die Umsetzung der Leitlinien gewonnen werden.

Aus den vorliegenden Resultaten lassen sich folgende **Empfehlungen** ableiten:

Die weniger umfassende Berücksichtigung von nicht im engeren Sinn „medizinischen“ Kriterien kann einerseits an der unterschiedlich wahrgenommenen Relevanz, andererseits auch an der teils weniger klaren Operationalisierung oder „Fassbarkeit“ dieser Themen liegen. Die weniger komplett bearbeiteten Kriterien geben Hinweise darauf, welche Themenbereiche von den Gutachtern und Gutachterinnen als weniger relevant oder als schwieriger zu beurteilen angesehen werden. Diese Themen sollten bei der Begründung und Einführung der Leitlinien für die Begutachtung besonders vertieft werden. Für die weitere Leitlinienentwicklung wäre die Präzisierung mancher Kriterien hilfreich, insbesondere derer, die noch nicht standardisiert erfasst werden können.

Bei der nachfolgenden Evaluation von Gutachten wäre dringend anzuraten, bei einzelnen Kriterien eine inhaltlich fundierte Gewichtung oder Priorisierung vorzunehmen, welche den Gutachterinnen und Gutachtern Hinweise darauf gibt, unter welchen Umständen welche Kriterien vertieft exploriert und dargestellt werden sollten. Andererseits könnte mit Hilfe solcher Regeln auch definiert werden, unter welchen Umständen die Angaben zu gewissen Kriterien weggelassen werden können.

Aus der nicht in allen Fällen optimalen Dokumentation der Beurteilung der Arbeitsfähigkeit ergeben sich Ansatzpunkte für die Entwicklung bzw. Verbreitung valider, standardisierter Instrumentarien als Grundlage für die Beurteilung der Arbeitsfähigkeit. Durch die Empfehlung und die Schulung im Einsatz von standardisierten Verfahren (z.B. zur Operationalisierung der ICF, validiert auf Probanden in Rentenverfahren) könnte die Qualität der Beurteilung einerseits und die Transparenz der gewonnenen Einschätzung andererseits deutlich erhöht werden.

Es besteht dringender und hoher Forschungsbedarf insbesondere im Hinblick auf die Fundierung der versicherungspsychiatrischen Beurteilung.

Zur Vertiefung der Befundung wäre bei der Entwicklung der inhaltlichen Leitlinien eine diagnosenabhängige Schwerpunktbildung evtl. mit Empfehlung von störungsspezifischen, normierten Diagnose-Instrumenten zu diskutieren.

Die aufgefundenen Unterschiede zwischen Institutionen und Regionen zeigen, dass eine grössere Vereinheitlichung durch die Einführung der Gutachten-Leitlinien möglich und vielversprechend ist. Weiterbildungsmassnahmen kommt bei der Implementierung der Leitlinien eine wesentliche Funktion zu. Die gefundenen institutionellen und regionalen Unterschiede geben Hinweise für eine allfällige Priorisierung von Schulungsmassnahmen mit einer Differenzierung nach deren Zielgruppe und Zielregion.

Die vorgestellte Untersuchung ist die Prä-Messung vor Einführung der Leitlinien. Für die Abschätzung des Umsetzungsgrades nach Einführung der Leitlinien und damit für die Praxis-Wirkungen der Leitlinien sowie für die Feinsteuerung der Leitlinienimplementierung sind Post-Erhebungen erforderlich. Zur optimalen Vergleichbarkeit der Ergebnisse aus der zweiten Phase mit den vorliegenden Ergebnissen sollte das Verfahren zur Datenerhebung und –analyse möglichst gut mit dem hier beschriebenen Verfahren übereinstimmen.

Die vorliegende Analyse der Erfüllung der formalen Kriterien stösst an vielen Stellen an die Grenzen ihrer Aussagekraft. Dies wurde insbesondere dort deutlich, wo Kriterien eine unterschiedliche Bedeutung für den Begutachtungszweck besitzen. Hier sind in Zukunft die inhaltlichen Kriterien mit einzubeziehen, um differenziertere Aussagen über die Relevanz einzelner Kriterien machen zu können.

Résumé

Situation initiale

Lorsque des personnes cessent prématurément ou partiellement d'exercer une activité lucrative et touchent une rente pour des raisons psychiques, c'est toujours un événement important non seulement pour elles-mêmes, mais aussi pour la société et l'économie.

Les offices AI chargent alors les services médicaux régionaux (SMR ; regionalärztlicher Dienst, servizio medico regionale), des psychiatres indépendants, des centres d'observation médicale (COMAI) ou d'autres institutions d'effectuer des expertises. Sur la base des documents existants et des expertises réalisées, les SMR déterminent si les critères médicaux justifiant l'octroi de prestations de l'AI sont remplis. Selon le type de pathologie, les expertises peuvent impliquer un, deux ou plusieurs spécialistes de différentes disciplines.

Des lignes directrices sur la qualité formelle des expertises psychiatriques sont élaborées pour garantir leur qualité et permettre aux experts de mieux connaître les attentes placées en eux.

Les experts se sont déjà entendus sur une première version des lignes directrices, comprenant 62 critères obligatoires et 25 critères facultatifs (voir annexe du présent rapport). Ces critères définissent les domaines qui doivent faire l'objet d'une expertise et sont donc classifiés selon ces domaines. Parallèlement, des directives sur le contenu sont en voie d'élaboration. Elles indiquent comment traiter les différents sujets, c'est-à-dire quel volume d'information et quel niveau de précision sont requis.

L'introduction des lignes directrices fait l'objet d'un accompagnement scientifique afin d'évaluer et d'optimiser leur utilité pratique. Deux étapes sont prévues : il s'agit d'abord d'analyser la qualité des expertises avant l'introduction des lignes directrices et, ensuite, de comparer les résultats obtenus avec ceux d'une deuxième évaluation effectuée après l'entrée en vigueur du nouvel outil. Le présent rapport contient les résultats de la première étape.

Méthode

L'OFAS a d'abord informé les offices AI. Puis le groupe de projet a demandé à tous les SMR de sélectionner de manière aléatoire un certain nombre d'expertises psychiatriques parmi les 15 000 réalisées dans les années 2008 et 2009. Il s'agissait de rassembler 800 expertises émanant de l'ensemble des SMR, le nombre d'expertises de chaque SMR étant proportionnel à sa contribution à l'ensemble. L'échantillon était ainsi représentatif pour toutes les régions du pays. Le choix étant aléatoire, on s'assurait aussi qu'aucun type d'intervenant, aucune langue, aucun diagnostic ni aucun autre élément ne soit sur- ni sous-représenté.

L'évaluation a finalement porté sur 775 expertises, 296 réalisées par des psychiatres indépendants, 270 par des SMR, 123 par des COMAI et 86 par d'autres institutions. La moitié environ d'entre elles ont été effectuées par un spécialiste, 15 % par deux spécialistes et un petit tiers par un nombre plus élevé de spécialistes.

Une liste de contrôle a été établie sur la base de la liste des critères de qualité. On y a indiqué, pour chaque expertise, si celle-ci contenait au moins des indications minimales sur les critères de qualité. La procédure engagée pour vérifier les expertises au niveau du respect des critères formels s'est avérée fiable. Ainsi, la pertinence des analyses présentées ci-dessous est assurée.

L'étude visant avant tout à réaliser une mesure initiale avant l'introduction des lignes directrices, les résultats concernant l'échantillon global sont les plus importants. A titre complémentaire, les diffé-

rences en termes d'exhaustivité entre les institutions et entre les régions (entre les SMR) ont aussi été examinées. Il s'agissait concrètement d'observer si des institutions ou des régions respectaient davantage de critères, en n'oubliant toutefois pas que ces critères n'étaient pas prévus lorsque les expertises ont été réalisées.

Sept unités thématiques sont définies dans les lignes directrices formelles régissant les expertises (1. Situation initiale et aspects formels, 2. Extrait de dossier, 3. Examen, 4. Constats, 5. Informations fournies par des tiers, 6. Evaluation médicale et médico-assurantielle, 7. Annexes). Elles ont aussi été utilisées pour la présente analyse des expertises, même si les points 5 et 7, facultatifs, n'ont pas été traités de manière approfondie. Un score d'ensemble pour les 62 critères obligatoires a aussi été calculé. Pour des raisons de méthode et de contenu, le passage des critères aux valeurs indicielles n'est pas sans poser de problèmes. Il faut donc interpréter les résultats obtenus avec une certaine réserve, en tenant compte par ex. du fait que les cas sont parfois peu nombreux. Nous renvoyons ici aux chapitres sur la présentation des résultats et aux conclusions et recommandations formulées dans le rapport. Une méthode éprouvée a permis de dégager les résultats mis tout particulièrement en évidence ci-dessous, qui ont aussi une grande importance pratique. Une évaluation des résultats est ensuite présentée dans le paragraphe « Conclusion et recommandations ».

Principaux résultats

L'analyse de l'échantillon global montre qu'il existe de grandes disparités dans le respect des critères formels. Pour les trois quarts des expertises examinées, plus des deux tiers des critères formels sont respectés alors qu'un petit nombre d'expertises ne respecte que la moitié des critères ou moins, ou, au contraire, respectent (presque) tous les critères.

Presque toutes les expertises obtiennent de bonnes notes pour l'« extrait de dossier » (présentation des documents pertinents utilisés dans l'expertise).

En ce qui concerne l'« **examen** », voici les principaux points qui n'ont pas été présentés comme le demande le catalogue de critères : l'idée que l'assuré se fait de sa maladie, la manière dont il la vit au quotidien, la perception qu'il a de l'avenir, l'anamnèse sexuelle, le contexte social (logement, ressources financières, relations), la date et les circonstances de la cessation de l'activité, les loisirs, les aides nécessaires pour le ménage et les activités quotidiennes, ainsi que le traitement au moment de l'expertise.

Dans la rubrique des « **constats** », des observations relatives au comportement et à l'apparence extérieure étaient faites dans moins de 90 % des cas. La personnalité de l'assuré était très rarement décrite (22 %).

Pour ce qui est de l'« **évaluation médicale et médico-assurantielle** », seuls 7 % des expertises traitent tous les sous-critères. Alors que le critère « Réponse aux questions » est abordé en principe, des atteintes fonctionnelles et des incapacités n'apparaissent que dans 2/3 des expertises sous la rubrique « **discussion et appréciation médico-assurantielle** », pour ne figurer qu'à titre de constat dans les autres cas. Pour les informations sur les ressources de l'assuré et sa volonté de coopérer, ainsi que pour l'intégration de toutes les estimations et informations, le taux est parfois inférieur à 70 %. Seulement cinq expertises indiquent sur quelle base a été porté le jugement sur la gravité du trouble et ses conséquences concernant la capacité de travail raisonnablement exigible.

Au niveau des **critères individuels**, c'est le critère « Description de la personnalité » qui est le moins souvent abordé. Les critères suivants sont mentionnés dans moins de 50 % des expertises : « Examens propres : durée de l'examen », « Indications spontanées de l'assuré / entretien ouvert », « Sexualité », « Perspectives concernant l'évolution de la maladie », « Perspectives générales », « Description de la personnalité » et « Appréciation critique de l'auto-évaluation de l'assuré ».

Dans l'ensemble, **il n'y a pas de grandes différences entre les intervenants qui font des expertises**. Toutefois, les expertises des SMR sont souvent un peu moins détaillées et complètes que les autres, même si ce n'est pas vrai pour tous les thèmes. Les expertises des COMAI sont en général

les plus complètes. Mais, très souvent, elles impliquent plus d'un spécialiste et comprennent plusieurs expertises partielles, ce qui pourrait expliquer pourquoi elles sont plus volumineuses et exhaustives. Au vu des données existantes, il n'est pas possible d'expliquer les différences d'une autre manière.

Les résultats ne permettent pas de faire des **distinctions tranchées entre les régions** en ce qui concerne la qualité. Il existe certes des différences régionales, mais le classement des régions varie selon le thème pris en compte.

En ce qui concerne les **types d'expertises**, les expertises pluridisciplinaires sont dans l'ensemble plus complètes que les expertises monodisciplinaires et bidisciplinaires. Cela est particulièrement vrai pour les points « situation initiale » et « examen ». On pouvait s'y attendre, car une approche pluridisciplinaire est plus détaillée qu'une approche monodisciplinaire, en particulier pour ce qui est de l'examen. Par contre, les expertises monodisciplinaires traitent la majorité des thèmes dans le chapitre des « Constats ». Fait intéressant, il n'y a pas de différences entre les types d'expertises en ce qui concerne l'« évaluation médicale et médico-assurantielle ». Autrement dit, si les expertises pluridisciplinaires vont plus loin dans la présentation de l'examen, l'évaluation qu'elles proposent n'est pas plus détaillée.

Conclusion et recommandations

Faute d'une pondération des critères relatifs au contenu en fonction de leur importance pour l'évaluation médico-assurantielle du cas, et en l'absence d'une deuxième mesure permettant de juger l'évolution, une évaluation fine des résultats obtenus à ce stade n'est pas possible. Seule une tendance peut être dégagée. Le mandat demandait avant tout de juger si tous les critères formels obligatoires étaient remplis. Cette approche purement formelle néglige le fait que tous les critères n'ont pas la même importance pour l'évaluation d'un cas, leur importance pouvant d'ailleurs varier en fonction de la situation particulière des personnes. Ainsi la mention ou non de l'adresse complète joue un rôle moins important pour l'évaluation médico-assurantielle que le critère « Expérience professionnelle ». Pour évaluer la qualité d'une expertise et juger si elle est adaptée à la situation de la personne, il faut donc faire des pondérations et établir des priorités. des opérations qui sont en cours avec la définition des critères de qualités portant sur le contenu.

Il est toutefois possible de tirer les **conclusions** suivantes :

L'analyse des expertises montre que certains critères, en particulier ceux concernant les éléments psychosociaux, sont moins souvent traités et pris en compte que les critères « médicaux » au sens étroit du terme. Les bases de l'évaluation en psychiatrie des assurances peuvent être améliorées, en particulier en ce qui concerne la documentation et le calcul de la capacité de travail. Il y aurait aussi beaucoup à faire pour mieux intégrer les expertises partielles dans les expertises pluridisciplinaires. Une étude des différences de qualité entre les institutions qui procèdent aux expertises et entre les régions permettrait également d'optimiser l'application des lignes directrices.

Sur la base des résultats obtenus, on peut formuler les **recommandations** suivantes :

Si les critères qui ne sont pas strictement « médicaux » sont moins bien pris en compte, cela vient peut-être du fait que tous les experts n'accordent pas la même importance à ces éléments ou qu'il est plus difficile d'en saisir le sens concret. Si un critère est traité moins à fond, c'est parce que les experts le jugent moins important ou plus difficile à évaluer. Ces thèmes devraient être approfondis lors de la mise au point et de l'introduction des lignes directrices. Lors de la poursuite des travaux, il serait utile de préciser plusieurs critères, en particulier ceux qui ne peuvent pas encore être saisis d'une manière standardisée.

Ceux qui procéderont à la prochaine évaluation des expertises devraient impérativement pondérer certains critères en fonction de leur contenu et établir un ordre de priorité, afin que les experts sachent dans quels cas ils doivent analyser et présenter de manière plus détaillée des critères, et les-

quels. S'il y avait des règles sur ce point, les experts sauraient dans quels cas ils ne seraient pas tenus de fournir les données relatives à certains critères.

Des instruments standardisés pourraient être développés et diffusés pour fonder l'évaluation de la capacité de travail, dont la présentation n'est parfois pas optimale. La qualité et la transparence de l'évaluation seraient nettement meilleures si les experts étaient encouragés à recourir à des procédures standardisées (par ex. pour opérationnaliser la classification CIF, avec validation auprès d'assurés dans le processus de rente) et formés à leur utilisation.

Il faut mener sans tarder des recherches en particulier sur les bases de l'évaluation médico-assurantielle.

Pour approfondir les résultats, il conviendrait, lors de l'élaboration des lignes directrices sur le contenu, de définir des points principaux liés au diagnostic et, le cas échéant, de faire des recommandations sur les instruments diagnostics normés propres au trouble.

Les différences décelées entre intervenants et régions montrent que l'introduction des lignes directrices pour les expertises favoriserait une uniformisation bienvenue des pratiques, mais qu'il faudrait veiller à ce que leur mise en œuvre s'accompagne de mesures de formation différentes selon la région ou le groupe visé.

Il faut rappeler que l'enquête présentée ici a un caractère préliminaire, puisqu'elle a été menée avant l'introduction des lignes directrices. D'autres enquêtes doivent être réalisées dans une seconde étape, pour évaluer l'application des lignes directrices, leur effet pratique et les mesures correctrices à prendre. La procédure de relevé et d'analyse des données adoptée pour réaliser ces enquêtes devrait concorder aussi étroitement que possible avec celle qui a été adoptée ici, afin que les résultats obtenus puissent être comparés de manière optimale avec ceux qui sont présentés ici.

Dans un certain nombre de domaines, la présente analyse sur le respect des critères formels ne permet pas de tirer de conclusions. Cela est vrai en particulier pour les domaines où certains critères sont plus importants que d'autres pour le but de l'expertise. Dans ce cas, les critères de contenu devront être pris en compte afin de pouvoir se prononcer plus clairement sur l'importance de certains critères.

Riassunto

Contesto

Il pensionamento anticipato o parziale e la concessione di una rendita per motivi psichici sono di grande rilevanza individuale, sociale ed economica. In questo contesto spetta un ruolo centrale alle perizie psichiatriche.

Le perizie sono commissionate dagli uffici AI ai Servizi Medici Regionali (SMR), a psichiatri indipendenti, ai servizi di accertamento medico (SAM) o a istituzioni simili. In base alla documentazione disponibile, incluse le perizie, i SMR valutano successivamente gli aspetti medici del diritto alle prestazioni AI se   stata presentata una richiesta di prestazioni. A seconda del problema in esame le perizie possono essere d'impostazione mono-, bi- o multidisciplinare.

Per garantirne la qualit  ai fini dell'AI sono in fase di elaborazione le linee direttive che in futuro dovranno fornire ai periti i requisiti per una perizia a regola d'arte.

Gli esperti incaricati hanno concordato ora una prima versione delle linee direttive sulla qualit  formale delle perizie, riportata nell'allegato al presente rapporto. Definendo 62 criteri obbligatori e 25 facoltativi, hanno stabilito gli ambiti tematici, basandovi anche l'articolazione stessa dei criteri, che dovranno essere trattati nelle perizie. Contemporaneamente sono in fase di sviluppo linee direttive che, al di l  della qualit  formale, vertono sul contenuto delle perizie, fissando l'ampiezza e il grado di dettaglio richiesti per il trattamento dei singoli temi.

L'introduzione delle linee direttive   seguita scientificamente per permettere di valutare e ottimizzare la loro utilit  nella prassi. Per questo scopo viene applicata una procedura a due fasi: la qualit  delle perizie sar  rilevata prima e dopo l'introduzione delle linee direttive per consentire di metterle a confronto i risultati. I risultati della prima rilevazione sono oggetto del presente rapporto.

Procedimento

Dopo che l'UFAS aveva informato tutti gli uffici AI del progetto in programma, il gruppo di lavoro incaricato ha invitato i SMR a tirare a sorte un campione di 800 fra le circa 15 000 perizie psichiatriche emesse tra il 2008 e il 2009. Con tale campione si   cercato di ottenere una rappresentanza equa di tutti i SMR in rapporto al numero complessivo delle perizie da loro emesse, garantendo cos  che tutte le regioni della Svizzera fossero adeguatamente considerate. Il metodo del sorteggio assicura inoltre che il campione non presenti distorsioni riguardo all'istituzione emittente, la lingua della perizia, la diagnosi o altre caratteristiche.

Si   potuto analizzare un campione complessivo di 775 perizie, di cui 296 emesse da psichiatri indipendenti, 270 dai SMR, 123 dai SAM e 86 da altre istituzioni. Circa la met  delle perizie era monodisciplinare, il 15% bidisciplinare e un terzo scarso multidisciplinare.

A partire dalla lista predefinita dei criteri di qualit    stata allestita una *check list* su cui   stato segnato se nella perizia in esame erano riscontrabili almeno i dati minimi relativi ai singoli criteri di qualit  formale. Dall'esame del metodo di valutazione adottato per verificare l'adempimento dei criteri formali   risultata una sufficiente affidabilit . Questo garantisce un buon grado di attendibilit  dei risultati qui di seguito esposti.

Per lo scopo principale dello studio, vale a dire la valutazione iniziale prima dell'introduzione delle linee direttive, sono di preminente importanza i risultati del campione complessivo di tutte le perizie.   stata inoltre esaminata la completezza delle perizie per individuare eventuali differenze fra i diversi tipi di istituzione emittente e le diverse regioni di provenienza (ossia i diversi SMR). Per

completezza si intende in questo contesto la considerazione del maggior numero possibile di criteri formali, tenendo conto, per , che al momento dell'allestimento della perizia i nuovi criteri non erano ancora noti agli esperti.

Le linee direttive per l'allestimento di perizie sono raggruppate in sette unit  tematiche: 1. Situazione di partenza e aspetti formali, 2. Estratto dall'incartamento, 3. Visita medica, 4. Referto medico, 5. Dati forniti da terzi, 6. Valutazione medica e medico-assicurativa, 7. Allegati. Questi temi sono stati anche alla base dell'analisi delle perizie nell'ambito del presente studio, ad eccezione dei temi facoltativi 5 e 7 che non sono stati approfonditi. Si   inoltre creato un punteggio complessivo che comprende tutti i 62 criteri obbligatori. Aggregare i criteri per ottenere siffatti valori indice pu , tuttavia, causare problemi a livello sia di metodo che di contenuto; i risultati cos  ottenuti vanno pertanto interpretati con cautela, tenendo presente, p.es., l'esiguo numero di casi. A questo proposito si rinvia ai singoli capitoli in cui sono esposti i risultati e alle conclusioni e raccomandazioni formulate nel rapporto. Insieme alla particolare rilevanza dei criteri per la prassi, la validit  metodologica dei risultati   servita anche da base per selezionare i risultati qui di seguito rilevati. La valutazione dei referti   svolta nella sezione "Conclusioni e raccomandazioni".

Principali risultati

I risultati del campione complessivo dimostrano un ampio scarto nella considerazione dei criteri formali: tre quarti delle perizie esaminate soddisfano almeno due terzi dei criteri formali, mentre sono poche le perizie che considerano solo la met  o addirittura meno della met  dei criteri e quelle che li considerano (quasi) tutti.

Quasi tutte le perizie raggiungono valori elevati nell'ambito del tema "**Estratto dall'incartamento**" (esposizione degli atti preliminari utilizzati per la stesura della perizia).

Nell'ambito del tema "**Visita medica**", invece, non sono stati descritti, come richiesto dal catalogo dei criteri, l'idea che l'assicurato si   fatto della sua malattia, il suo atteggiamento nei confronti dei disturbi nella vita quotidiana, l'anamnesi sessuale, la sfera sociale (casa, finanze, rete relazionale), la data e le circostanze dell'abbandono del lavoro, le attivit  nel tempo libero, l'aiuto necessario in casa e nella vita quotidiana, la terapia attuale e la visione del futuro dell'assicurato.

Nell'ambito del tema "**Referto medico**" in meno del 90% dei casi sono stati descritti il comportamento e l'aspetto esteriore dell'assicurato, mentre   stata trascurata nettamente la descrizione della personalit  (22%).

Riguardo alla **valutazione medica e medico-assicurativa** solo il 7% circa delle perizie considera tutti i criteri. Mentre il criterio "Risposte alle domande"   di regola soddisfatto, nell'ambito del criterio "Discussione e valutazione medico-assicurativa" si accertano disturbi funzionali e delle capacit  soltanto in due terzi delle perizie; nelle altre, invece, i disturbi sono sovente solo menzionati senza essere descritti. Anche le informazioni sulle risorse disponibili dell'assicurato, sulla sua disponibilit  a cooperare e sulla combinazione di tutte le valutazioni e informazioni insieme spesso si situano ben al di sotto del 70%. Soltanto in cinque delle perizie esaminate sono stati documentati i criteri in base ai quali si   giunto alla valutazione del grado del disturbo e delle ripercussioni sulla capacit  lavorativa esigibile.

Per quanto riguarda i **singoli criteri**, il criterio "Descrizione della personalit "   stato quello di cui si   tenuto meno conto in assoluto. In meno del 50% delle perizie sono stati menzionati i criteri "Visite eseguite dal medico stesso: durata della visita", "Dati forniti spontaneamente dall'assicurato / intervista aperta", "Sessualit ", "Aspettative riguardo al decorso della malattia", "Visione del futuro in generale", "Descrizione della personalit " e "Valutazione critica dell'autovalutazione dell'assicurato".

Complessivamente, le differenze tra le **istituzioni incaricate di eseguire perizie** non sono enormi. Le perizie dei SMR sono tendenzialmente meno esaurienti o complete di quelle emesse da

altri tipi di istituzioni, anche se ci  non vale per tutti gli ambiti tematici. Di regola, le perizie dei SAM sono dotate di maggiore completezza, qualit  che pu  tuttavia essere dovuta alla loro stessa natura, dato che sono sovente multidisciplinari e composte da pi  perizie parziali, e quindi automaticamente pi  ampie e dettagliate. Sulla base dei dati disponibili non   possibile risalire ad altri motivi.

Dai risultati sui **gruppi di perizie provenienti dalle diverse regioni** non   possibile dedurre differenze sistematiche tra la qualit  delle medesime. Si possono constatare alcune differenze regionali, che variano per  a seconda dei singoli criteri; pertanto variano anche le classifiche delle regioni stilate in base alla completezza per ambito tematico.

Fra i **vari tipi**, le perizie multidisciplinari si sono dimostrate complessivamente pi  dettagliate di quelle mono- o bidisciplinari. Ci  vale soprattutto per gli ambiti tematici "Situazione iniziale" e "Visita medica", il che sembra corrispondere alle aspettative, in quanto le perizie multidisciplinari sono in ogni caso pi  dettagliate di quelle monodisciplinari, in particolare nell'ambito della visita medica. Le perizie monodisciplinari, invece, considerano il maggior numero di criteri nell'ambito tematico "Referto medico".   molto interessante notare che nell'ambito della "Valutazione medica e medico-assicurativa" non vi sono differenze fra i diversi tipi di perizia; la maggiore completezza delle perizie multidisciplinari a livello di visita medica non si traduce dunque in una valutazione pi  dettagliata in quest'ambito.

Conclusioni e raccomandazioni

In assenza di criteri di qualit  applicabili al contenuto, inclusa la ponderazione dei criteri secondo la loro rilevanza per la valutazione medico-assicurativa di un caso, e senza seconda misurazione per la valutazione del decorso, i risultati finora ottenuti sono interpretabili soltanto tracciando una tendenza generale. Conformemente ai termini del mandato, come valore di riferimento   stato assunto innanzitutto l'adempimento di tutti i criteri obbligatori. Questa visione prettamente formale trascura tuttavia che i criteri rivestono di per s  una rilevanza diversa per la valutazione dei casi e che la valutazione pu  variare a seconda dei problemi del singolo caso; la completezza o meno dell'indirizzo, ad esempio,   ovviamente meno rilevante per la valutazione psichiatrica in campo assicurativo del criterio "Percorso professionale". Per una valutazione complessiva della qualit  e dell'adeguatezza delle perizie riguardo alle questioni sollevate dal singolo caso   dunque necessario ponderare o organizzare per priorit  i criteri, un compito questo che   in fase di elaborazione insieme ai criteri di qualit  per il contenuto delle perizie.

Ciononostante   possibile trarre le seguenti **conclusioni**:

L'analisi delle perizie dimostra che singoli criteri, in particolare quelli concernenti l'ambito psicosociale, sono stati trattati e considerati pi  raramente dei criteri "medici" in senso stretto. Eventuali migliorie possono essere apportate alle valutazioni medico-assicurative, in particolare a livello di documentazione e derivazione plausibile del grado di capacit  al lavoro. Le perizie multidisciplinari invece necessitano di migliorie pi  incisive per quel che riguarda l'integrazione delle perizie parziali. Dalle differenze riscontrate tra le istituzioni incaricate di emettere perizie e tra le regioni per quanto riguarda la qualit  delle perizie si possono trarre alcune conclusioni utili per l'applicazione delle linee direttive.

Dai presenti risultati si possono desumere le seguenti **raccomandazioni**:

La scarsa considerazione riservata ai criteri non "medici" in senso stretto pu  essere dovuta al fatto che questi ambiti tematici hanno, da un lato, una rilevanza percepita diversa e, dall'altro, sono meno operazionalizzabili e afferrabili. I criteri trattati in minor dettaglio forniscono indicazioni su quali ambiti sono stati considerati meno rilevanti o pi  difficili da valutare dai periti. Questi temi andranno ulteriormente approfonditi in occasione della motivazione e dell'introduzione delle linee direttive per le perizie. Per lo sviluppo delle linee direttive sarebbe utile precisare alcuni criteri, in particolare quelli che per ora non   possibile standardizzare.

Nell'ambito della seconda valutazione delle perizie si raccomanda di procedere a una ponderazione basata sul contenuto di alcuni criteri, il che fornirebbe ai periti indicazioni in merito alle circostanze in cui singoli criteri andranno approfonditi e illustrati in maggior dettaglio. Siffatte regole permetterebbero di definire tra l'altro anche in quali circostanze è possibile trascurare determinati criteri.

Dalla documentazione, non sempre ottima, dei casi di valutazione della capacità lavorativa si evincono spunti per lo sviluppo e la diffusione di strumenti standardizzati validi quale base per questo compito. Grazie alla raccomandazione e alla formazione ai fini dell'applicazione di procedure standardizzate (ad es. per rendere operativa l'ICF – acronimo inglese per la classificazione del funzionamento, della disabilità e della salute – e validarla per gli assicurati nell'ambito di procedure di concessione di rendite) sarebbe possibile migliorare nettamente la qualità e la trasparenza della valutazione.

Sussiste un urgente ed elevato bisogno di ricerca per creare le basi della valutazione psichiatrica in campo assicurativo.

Per approfondire la valutazione nell'ambito dello sviluppo delle linee direttive per il contenuto sarebbe utile discutere dell'opportunità di creare dei punti focali in funzione della diagnosi, integrati eventualmente dalla raccomandazione di strumenti di diagnosi specifici per il disturbo e standardizzati.

Le differenze riscontrate tra le diverse istituzioni e regioni dimostrano che grazie all'introduzione delle linee direttive per le perizie sarà possibile ottenere una maggiore armonizzazione. Ciò è molto promettente. Nell'applicazione delle linee direttive spetta un ruolo chiave alle misure di perfezionamento. Le differenze fra istituzioni e regioni forniscono indicazioni circa la necessità di fissare priorità fra le misure di formazione distinguendo tra gruppi di destinatari e regioni diversi.

L'analisi presentata rappresenta la valutazione delle perizie eseguita prima dell'introduzione delle linee direttive. Per valutare il grado di attuazione e l'efficacia nella prassi, una volta introdotte le linee direttive, e per consentire la gestione dettagliata della loro introduzione, è necessario eseguire un'indagine a posteriori. Per garantire che i risultati della seconda fase siano comparabili con quelli del presente studio, occorre che la procedura di rilevazione e analisi dei dati corrisponda alla procedura qui presentata.

In molti punti la presente analisi dell'adempimento dei criteri formali è andata vicina al limite di attendibilità. Ciò si è verificato soprattutto là dove i criteri rivestono un'importanza diversa a seconda dello scopo della perizia. In questi casi occorrerà in futuro addurre anche criteri di contenuto per poter fare affermazioni differenziate sulla rilevanza dei singoli criteri.

Summary

Background

Early or partial withdrawal from the labour market and taking early retirement because of mental health problems have far-reaching consequences not only for the individual, but also for the economy and society as a whole. Psychiatric assessments are therefore one of the most important instruments at the disposal of the Swiss invalidity insurance system (IV).

IV psychiatric assessments are the responsibility of the Regional Medical Services (Regionalärztliche Dienste [RAD], Services Médicaux Régionaux/Servizio Medico Regionale [SMR]). They can either outsource this work to independent psychiatrists or to medical observation centres (MAS), or carry out the assessments themselves. The RAD/SMR examines the case file, including any existing assessments, to ascertain the medical grounds for IV pension eligibility. Depending on the subject at hand, these assessments can be monodisciplinary, bidisciplinary or multidisciplinary in design.

Given the importance of these psychiatric assessments for the IV, guidelines are currently being developed to assure the future quality of this work.

The draft guidelines put forward unanimously by a panel of experts focus on the formal quality of IV psychiatric assessments. They set out 62 mandatory and 25 optional individual criteria which define the key points that these assessments should cover. Consequently, each criterion is classified according to the subject area to which it relates. Work is currently under way on the formulation of content-specific guidelines setting out the scope and level of detail that assessors should afford each point.

Scientific research will be called upon to evaluate and optimise the practical benefits of these guidelines upon their introduction. This process is to be divided into two phases. Phase 1 has evaluated the quality of psychiatric assessments prior to the adoption of the guidelines. The findings from this research will be compared in Phase 2 to the quality of assessments after the introduction of the guidelines. The present report describes the results of the Phase 1 survey.

Method

Once the FSIO had informed the IV offices about the project, the project team asked every RAD/SMR in Switzerland to provide a random sample from the 15,000 assessments carried out in 2008 and 2009. The authors sought to achieve a sample size of 800 to ensure that the sample was representative of the regions, of all the RAD, and of the total number of assessments each RAD had carried out during the observation period. Random sampling also ensured that there was no selection bias in terms of the individual/institutional assessors, language or diagnosis.

The final sample contained 775 assessments, of which 296 were carried out by independent psychiatrists, 270 by the RAD/SMR, 123 by MEDAS, and 86 by other assessors. Roughly half of the assessments were monodisciplinary, 15% bidisciplinary and nearly one-third multidisciplinary.

To facilitate the evaluation, the authors converted the formal quality criteria into a checklist, allowing them to note whether each assessment provided a minimum of information on each of the formal quality criteria. The method used to evaluate how well the assessments satisfied the formal criteria was found to be wholly reliable. This ensures a high degree of validity as regards the analytical findings described herein.

Given that the primary aim of the research was to establish baseline measurements prior to the introduction of guidelines, the most important findings were those on the total sample. Nevertheless, the authors also checked whether the completeness of assessments varied according to the type of assessor and according to the region (i.e. the RAD) where they were carried out. For the purpose of

this study, completeness means that the assessment satisfied a maximum of the formal criteria. However, allowances were made because the criteria that are now applied were not known to the assessors at the time of the assessment.

The formal assessment guidelines break down into seven distinct subject areas (1. Background and Formal Requirements; 2. Case-File Documentation; 3. Consultation; 4. Diagnostic Findings; 5. Third-Party Information; 6. Medical and Insurance-Medicine Assessment; and 7. Accompanying Documentation). The researchers retained these categories for their evaluation of the psychiatric assessments. Having said that, Categories 5 and 7 are optional and were therefore not explored in detail by the research team. A total score was derived across the 62 mandatory criteria. Given the methodological and substantive problems of aggregating criteria into indices and the sometimes small sample sizes, caution had to be exercised when interpreting the results. This point is reiterated in the Results section as well as in the Conclusions and Recommendations section of the report. The validity of the test methodology, which the authors considered to have a high practical relevance, also facilitated the identification of the most important findings yielded by this research. These results are evaluated in the “Conclusions and Recommendations” Section below.

Main findings

Analysis of the total sample shows a wide distribution of the inclusion incidence of the formal criteria. Three-quarters of the assessments satisfied over two-thirds of the formal criteria. A handful of assessments satisfied half of the criteria or less, and a few satisfied (almost) all criteria.

Almost all assessments scored well on the category “**Case-File Documentation**” (i.e. inclusion of all earlier case notes and reports which are relevant to the present assessment).

With regard to the “**Consultation**” category, many assessments did not mention the following points contained in the criteria: subjective health; how the subject copes with his/her health problems on a day-to-day basis; self-reported prospects; sexual history; general circumstances (housing, finances, social network); when and why (s)he left work; leisure activities; help with household chores and routine activities.

With regard to the category “**Diagnostic Findings**”, less than 90% of assessments contained references to the observed behaviour and general appearance of the subject. Only 22% offered information on the personality type of the subject.

With regard to the category “**Medical and Insurance-Medicine Assessment**”, a mere 7% of assessments considered the entire set of sub-criteria. Although most assessments satisfied the criterion “response to questions”, only two-thirds explicitly identified dysfunction and incapacity from the “discussion and insurance-medicine appraisal”. When these two points were mentioned in the remaining cases, they often amounted to nothing more than a simple statement of fact. Less than 70% of assessments detailed the resources available to the subject and his willingness to cooperate, and integrated all estimates and information. Only five assessments documented the bases for the assessment of the severity of the illness and its effect on the capacity to work which could be reasonably expected from the subject.

With regard to the **individual criteria**, the criterion “Description of personality” was overlooked the most. Less than 50% of the assessments referred to the following criteria: own consultations (incl. duration); spontaneous information offered by the subject/open interview; sexuality; progression and outlook of illness; general prospects; description of personality; and the critical appraisal of self-reported health status.

Generally speaking, no major **differences were observed across assessors**. However, the study identified that the information contained under each category tended to be less detailed and complete in the RAD assessments. Overall, assessments conducted by MEDAS tended to be the most complete and had the widest scope. Many of these assessments were multidisciplinary in design,

which meant that they were based on several reports. This could explain some of their more complete nature. We were unable to identify from the data any additional reasons for these differences.

With regard to group differences across regions, we did observe certain variations. However, no clear pattern in terms of quality emerged due to the fact that the criteria completeness scores of each region varied depending on the criteria category under assessment.

With regard to the **type of assessment**, the study found that the completeness of multidisciplinary assessments was generally better, particularly in the categories “Background” and “Consultation”. This is to be expected given that a multidisciplinary design provides for greater scope and detail than a monodisciplinary study, particularly in relation to the “Consultation” category. However, monodisciplinary assessments covered most of the criteria in the “Diagnostic Findings” category. Interestingly, as regards the category “Medical and Insurance-Medicine Assessment”, no differences were observed across the various types of assessment. We can therefore conclude that although multidisciplinary assessments performed better in the “Consultation” category, this did not automatically mean that these were the most detailed assessments overall.

Conclusion and recommendations

Given that the formulation of content-related criteria weighted according to their relevance for an insurance-medicine assessment is still ongoing and that the second round of evaluations has yet to be undertaken, the present study only allows for broad and preliminary inference. In keeping with the project specifications, our priority was to evaluate the fulfilment of all mandatory formal criteria. However, this approach overlooks the fact that the relevance of each criterion differs not only for the psychiatric assessment but also according to the specific set of problems inherent to a given case. For example, the presence or absence of the address for the insurance-related psychiatric assessment is less important than the provision or lack of information on the person’s “employment history”. For a comprehensive assessment of the quality and adequacy of an assessment with regard to individual problems, the formal criteria need to be weighted and ranked according to priority. This work is currently being carried out in conjunction with the formulation of content-related criteria.

These caveats notwithstanding, the following **conclusions** can be drawn:

The findings of the assessment evaluation show that individual criteria, particularly those that fall into the psychosocial spectrum, tended to be considered and dealt with less frequently than the narrower “medical” criteria. In order to remedy this situation, the fundamentals of insurance-related psychiatric assessments could be improved, particularly as regards documentation and the transparent evaluation of the person’s degree of incapacity. Another evident area for improvement is the integration of partial reports in multidisciplinary assessments. The future implementation of the guidelines could benefit from pointers derived from analysing the differences across the assessment institutions as well as regional variations in the quality of assessments.

Based on the evaluation findings, the authors put forward the following **recommendations**:

The fact that the narrower “medical” criteria were dealt with less thoroughly than other criteria may be due to differing perceptions of their relevance for the assessment, and to the fact that some psychosocial variables may be more difficult to operationalise or are simply harder to grasp. The list of criteria that were not covered as well as they should have been point to the areas that assessors considered of lesser relevance or more problematic. More work must be done on these areas as part of the process to formulate and introduce the new psychiatric assessment guidelines. In addition, a number of criteria could be refined, particularly those that do not lend themselves to standardised documentation.

The next stage of this research should weight/prioritise individual criteria according to their content. This will help assessors understand when they should include certain criteria and how much attention they should give to them. Likewise, the guidelines could help assessors decide when they could omit specific criteria.

Given that the study found that many of the assessments fell short on the matter of work capacity evaluations, validated and standardised evaluation tools should be developed and disseminated. The quality of the assessment and the transparency of its outcomes could be greatly improved by recommending the use of standardised procedures (e.g. operationalisation of the ICF, which have already been validated on IV pension claimants) and providing assessors with the necessary training to apply them.

There is a considerable and urgent need for research on the fundamentals of insurance-related psychiatric assessments.

In the interests of more precise appraisals/diagnoses, future content-related guidelines should have a diagnostic focus, which could possibly extend to including recommendations on the use of standardised, disorder-specific diagnostic tools.

Differences observed across assessors/assessing institutions and across regions show that the introduction of assessment guidelines could greatly help standardise procedures. Providing assessors with the necessary training to apply these criteria will be particularly important. Training should be tailored to target groups and target regions.

The present study provides a detailed description of assessment indices prior to the introduction of the guidelines. Follow-up studies will be needed to evaluate implementation as well as the effect of these guidelines in practice, and to fine-tune their application. In the interests of comparability, the methods used to collect and analyse data in Phase II should be the same as, or at least compatible with, the methods used in Phase I.

In many instances, the validity of our findings is inevitably limited, particularly regarding criteria that are of unequal relevance depending on the purpose of the assessment. Content-related criteria therefore are needed to ensure that more rigorous conclusions can be drawn on the relevance of individual criteria.

1 Einleitung

In den letzten Jahren hat die Zahl der IV-RentnerInnen in der Schweiz sehr stark zugenommen. 40% der Neuberentungen erfolgen aufgrund psychischer Störungen (Rajower, 2007). Hierfür sind insbesondere solche Störungen verantwortlich, die im Grenzbereich zwischen somatischen und psychischen Erkrankungen angesiedelt sind und die nicht oder nicht ausreichend durch somatische Befunde oder „objektivierbare“ psychische Störungen erklärbar sind.

Versicherungspsychiatrischen Gutachten zur Beurteilung einer dauerhaften Einschränkung der Arbeitsfähigkeit aus psychischen Gründen kommt eine entscheidende Bedeutung zu. Trotz der zentralen Wichtigkeit der Begutachtung wird die Qualität versicherungspsychiatrischer Gutachten in der Schweiz von den Auftraggebern als heterogen und oftmals als unzureichend angesehen (Mosimann & Ebner, 2008). Eine Dossieranalyse des BSV zur Invalidität aus psychischen Gründen (Baer et al. 2009) ergab insbesondere, dass in der Hälfte der medizinischen Dokumente, auf deren Grundlage eine Rente zugesprochen wurde, nicht ausreichend dokumentiert ist, wie sich die Erkrankung auf die Arbeitsfähigkeit auswirkt.

1.1 Ausgangslage

Vor diesem Hintergrund besteht insbesondere Handlungsbedarf bezüglich der Entwicklung von Leitlinien, die sowohl formale als auch inhaltliche Kriterien einer qualitativ hoch stehenden Begutachtung definieren. Aus diesem Grund initiierte das BSV zwei Arbeitsgruppen, die solche formalen und inhaltlichen Kriterien erarbeiten. Parallel zu deren Tätigkeit beauftragte das BSV die UPK-Arbeitsgruppe in einem ersten Schritt mit der Durchführung einer „Literaturstudie als Grundlage zur Entwicklung von evidenzbasierten Gütekriterien zur Beurteilung von psychischen Behinderungen“.

Das Ergebnis dieser systematischen Literaturrecherche ist ernüchternd: Keine der als relevant ermittelten Publikationen weist ein Evidenzniveau von I (Meta-Analysen, randomisierte kontrollierte Studien) auf. Alle gängigen Beurteilungskriterien sind dem Evidenzniveau IV zuzuordnen (Expertenmeinungen, klinische Erfahrungen ausgewiesener Autoritäten). Der Ist-Stand ist unbefriedigend, der dringende Bedarf an evidenzbasierten, validierten Kriterien konnte deutlich herausgearbeitet werden (siehe Dittmann et al., 2009).

Zwischenzeitlich haben Linden et al. (2005, 2009) ein Instrument (Mini ICF APP) entwickelt, das die Komponente „Aktivitäten und Partizipation“ der ICF abbildet, soweit sie im Rahmen psychischer Erkrankungen beeinträchtigt sein können. Mit dem Mini-ICF-APP-Rating kann eingeschätzt werden, in welchem Ausmaß ein Patient in seinen Fähigkeiten bei der Durchführung von Aktivitäten beeinträchtigt ist. Das System wurde bereits für Rehabilitationspatienten validiert und wird aktuell in der Schweiz für Personen in Rentenverfahren validiert.

Für die weitere Entwicklung und Implementierung von Leitlinien für die Begutachtung psychischer Störungen spielt die begleitende (formative) Evaluation eine wesentliche Rolle. Ein umfassendes Evaluationskonzept erlaubt es, sowohl den Implementierungsgrad der Leitlinien wie auch deren Wirksamkeit zu beurteilen und durch entsprechende Massnahmen zu optimieren. Ausgangspunkt ist dabei die Erhebung der heutigen Gutachtenqualität. Als Basis für alle weiteren Vergleichserhebungen dient eine Baseline-Erhebung („Null-Messung“), die die Qualität der Gutachten zum aktuellen Zeitpunkt, also vor Einführung von Leitlinien, erfasst. Nachfolgende Erhebungen, z.B. im Rahmen einer pilotmässigen Leitlinienimplementierung, können dann vor dem Hintergrund dieser Baseline-Erhebung interpretiert werden und dadurch Hinweise zur Implementierungssteuerung liefern.

Die hier dargestellten Ergebnisse der Analyse von Gutachten aus den Jahren 2008 und 2009 stellt eine solche Baseline-Erhebung dar. Die Beurteilung der Gutachtenqualität hinsichtlich formaler Gutachtkriterien erfolgt dabei anhand der in der Arbeitsgruppe „Formale Leitlinienentwicklung“ des BSV-Projekts zur Erarbeitung von Qualitätsleitlinien entwickelten Kriterien (siehe Anhang 1).

1.2 Übersicht zum vorliegenden Bericht

Der vorliegende Bericht stellt den Schlussbericht zur Erhebung der formalen Qualität psychiatrischer Gutachten dar. Dem Auftraggeber wurden im Projektzeitraum drei Zwischenberichte vorgelegt. Um die Lesbarkeit dieses Schlussberichts zu erhöhen, wird auf Verweise auf diese Zwischenberichte verzichtet. Entwicklungen und Anpassungen gegenüber dem ursprünglich vorgesehenen Vorgehen im Projektverlauf betreffen das Vorgehen bei der Stichprobengewinnung und werden im entsprechenden Abschnitt (3.2) behandelt.

Nach der Darstellung der bearbeiteten Fragestellungen in Abschnitt 2 und der Methodik in Abschnitt 3 werden Ergebnisse auf der Basis der Gesamtstichprobe in Abschnitt 4 behandelt. Im Anschluss werden separat die Analysen bzgl. Qualitätsunterschieden zwischen Gutachten verschiedener Institutionen (Abschnitt 5) und zwischen Gutachten aus verschiedenen Regionen (Abschnitt 6) dargestellt. In den abschliessenden Berichtsteilen werden die gefundenen Ergebnisse diskutiert (Abschnitt 7) und daraus Schlussfolgerungen und Empfehlungen (Abschnitt 8) abgeleitet.

2 Fragestellungen

Die vorliegende Arbeit geht folgender übergeordneter Fragestellung nach:

Wie ist das aktuelle durchschnittliche Qualitätsniveau der zum jetzigen Zeitpunkt (vor Einführung von Leitlinien) vorliegenden Gutachten hinsichtlich formaler Begutachungskriterien?

Dieser Hauptfragestellung wird durch die Bearbeitung der folgenden detaillierten Fragestellungen nachgegangen:

1. Wie stellt sich die durchschnittliche formale Qualität bezogen auf die einzelnen Gutachtenteile (Ausgangslage, Zusammenfassung vorliegender Aktenstücke, Anamnese, Untersuchung, Verwendung anerkannter Diagnosesysteme, Art und Umfang der Begründung gutachterlicher Schlussfolgerungen, Literaturbelege etc.) dar (siehe Abschnitt 4.3)?
2. Bestehen systematische Qualitätsunterschiede zwischen den einzelnen Gutachtenteilen (siehe Abschnitt 4.4)?
3. Bestehen systematische Qualitätsunterschiede zwischen den Gutachten verschiedener Institutionen (z.B. Freipraktizierende, Kliniken, RAD, Medas; siehe Abschnitte 4.5 und 5)?
4. Bestehen systematische Qualitätsunterschiede zwischen den Gutachten aus verschiedenen Regionen (siehe Abschnitt 6)?
5. Welche Konsequenzen ergeben sich aus der ermittelten formalen Gutachtenqualität für a.) die weitere Leitlinienentwicklung (siehe Abschnitt 8) und b.) die Leitlinienimplementierung (siehe Abschnitt 8.2)?
6. Welche Empfehlungen lassen sich für die weitere Evaluation der Leitlinienentwicklung und -implementierung ableiten (siehe Abschnitt 8.3)?

3 Methodisches Vorgehen

Nachfolgend werden Studiendesign, Gutachtenstichprobe, Operationalisierung und Datenstruktur, Datenerhebung und Datenerfassung sowie das Vorgehen bei der Auswertung der Gutachten vorgestellt.

3.1 Studiendesign

Die Studie wurde als Querschnittsuntersuchung angelegt. Berücksichtigt wurden dabei Gutachten, welche in den Jahren 2008 und 2009 fertiggestellt wurden, mit dem Ziel, den aktuellen *Ist*-Stand hinsichtlich der Erfüllung formaler Qualitätskriterien anhand einer Stichprobe von $n=800$ regional und institutionell repräsentativen Gutachten verschiedener Regionen und Institutionen der Schweiz zu ermitteln.

Um die (sprach-)regionale Repräsentativität, sprich die angemessene Berücksichtigung der drei grossen Sprachgebiete (Deutsch, Französisch & Italienisch) zu gewährleisten, wurden alle Gutachten ihrer Verteilung entsprechend aus allen Kantonen gezogen, so dass eine Schichtung nach Deutschschweiz, Romandie und Tessin nicht erforderlich war. Regionale Repräsentativität wurde hergestellt, indem sämtliche Gutachten eines Kantons resp. eines RAD auf einer Liste erfasst wurden und die Stichprobenziehung anhand dieser Liste per Zufall erfolgte.

Die institutionelle Repräsentativität wurde sichergestellt, indem die Gesamtzahl aller auf RAD und MEDAS entfallenden Gutachten erhoben wurde. Die Ziehung der Gutachtenstichprobe (einschliesslich der MEDAS-Gutachten) erfolgte ausschliesslich über die RAD bzw. IV-Stellen.

Da die Stichproben aller RAD ein genaues Abbild der Gesamtzahl ihrer Gutachten sind, und die RAD die Auswahl der Gutachten für die Auswertung jeweils per Zufall trafen, sind die Stichproben repräsentativ für die RAD und – da die RAD alle Regionen der Schweiz abdecken – auch für die Sprachregionen und für die Gutachten-erstellenden Institutionen.

3.2 Datenerhebung und Datenerfassung

Zur Unterstützung der Datenerhebung wurde der Projektgruppe ein Begleitschreiben des BSV in französischer und deutscher Sprache zur Verfügung gestellt. Zudem kündigte das BSV den beteiligten RAD (und MEDAS) das Evaluationsprojekt per E-Mail an. Daraufhin nahmen die Projektmitarbeitenden Kontakt mit den RAD auf, die mit einem detaillierten Anschreiben zum Ziel und geplanten Ablauf der Studie in deutscher, italienischer und französischer Sprache informiert wurden.

Die genaue Grösse N der Gutachtenpopulation, also die Anzahl aller im Untersuchungszeitraum in der Schweiz angefertigten psychiatrischen IV-Gutachten, war nicht bekannt und musste in einem ersten Schritt durch Befragung aller RAD und MEDAS ermittelt werden. Zur Klärung möglicher Fragen bzw. Unklarheiten erfolgte nach etwa 7 Tagen eine weitere Kontaktaufnahme durch die Projektmitarbeiter, die jeweils telefonisch bzw. zu RAD/SMR 5 schriftlich durchgeführt wurde.

Aus der ermittelten Gesamtzahl aller bei den RAD vorliegenden psychiatrischen Gutachten einerseits und der Zielvorgabe für das Projekt, eine Stichprobe von 800 Gutachten zu analysieren, andererseits, wurde errechnet, wie viele Gutachten die einzelnen RAD beisteuern sollten. Da die Erhebung Schweiz weit bei allen RAD/SMR erfolgte, konnte auf eine Schichtung nach Regionen etc. verzichtet werden. Der Faktor, der zur Gewinnung der Stichprobe von $n=800$ Gutachten anzuwenden war, war deshalb der exakte ganzzahlige Teiler aus der Summe aller GA dividiert durch die vorgese-

hene Stichprobengrösse von $n=800$ multipliziert mit dem Anteil, den jeder RAD an der Gesamtzahl hatte.

Die Einhaltung des Datenschutzes bei den Gutachtenanalysen wurde durch einen entsprechenden Vertrag zwischen Auftraggeber und Auftragnehmer und entsprechende Dienstvorschriften und Verfahrensabläufe beim Auftragnehmer sichergestellt.

Wie in den dem Auftraggeber vorgelegten Zwischenberichten (12.08.2010, 1.12.2010 und 31.3.2011) dokumentiert, konnte die Analyse der Gutachten nicht wie im ursprünglichen Zeitplan vorgesehen im September 2010, sondern erst im März 2011 begonnen werden. Diese erhebliche zeitliche Verzögerung entstand aufgrund von zunächst mangelhaften Angaben der befragten Stellen: Die für alle weiteren Arbeitsschritte erforderlichen Angaben zur Grundgesamtheit aller im Zeitraum 2008-2009 erstellten Gutachten konnte final erst im Februar 2011 aus den dann vorliegenden Angaben ermittelt werden, da die RAD teilweise ihre zunächst gemeldeten Zahlen noch korrigieren mussten. Die Zusage der Gutachten selbst zog sich dann bis zum 02. August 2011 hin, an dem die letzte Tranche von Gutachten für die Auswertung eintraf.

Der gegenüber der ursprünglichen Kalkulation erhebliche Mehraufwand durch die Gutachtenbeschaffung und Nacherhebungen bei den beteiligten Stellen konnte teilweise durch die intensiviertere Gutachtenanalyse ab März 2011 kompensiert werden, wodurch sich der Zeitraum der Gutachtenanalyse von 7 Monaten auf 6 Monate reduzierte. Trotzdem verlängerte sich die Projektlaufzeit gegenüber dem ursprünglichen Zeitplan um fünf Monate.

Die Gutachten wurden der Projektgruppe durch die RAD in der Regel als PDF-Dateien auf CDs zugeschickt. Die (deskriptiven) Basisdaten der einzelnen Gutachten (Datum, beauftragende Institution etc.) wurden in einer Datenbank erfasst und dann hinsichtlich der formalen Kriterien ausgewertet.

3.3 Gutachtenstichprobe

3.3.1 Grundgesamtheit

Auf die Anfrage der Arbeitsgruppe an die RAD bzw. IV-Stellen, die Gesamtzahl der psychiatrischen, mono-, bi- bzw. polydisziplinären Gutachten der Jahre 2008 und 2009 zu ermitteln, wurden dem Projektteam im November 2010 folgende Werte mitgeteilt (siehe Tabelle 3-1).

Tabelle 3-1: Grundgesamtheit, Stand Nov. 2010

Herkunft	Gesamt		
	2008	2009	gesamt
RAD/SMR 3			1'756
RAD/SMR 8			2'666
RAD/SMR 5	718	784	1'502
RAD/SMR 4	821	720	1'541
RAD/SMR 7			1'942
RAD/SMR 2			3'387
RAD/SMR 9	297	415	712
RAD/SMR 6	339	284	623
RAD/SMR 1	823	560	1'383
RAD/SMR 10			792
Summen			16'304

* Diese Zahlen mussten nachträglich noch korrigiert werden, die finalen Zahlen finden sich in Tabelle 3-2.

Als Rate für die Stichprobenziehung, errechnet aus der Gesamtzahl von 16'304 Gutachten dividiert durch die angestrebte Stichprobengrösse (n=800) wurde ein Faktor 20.38 ermittelt, es musste von allen Stellen also jedes 20. der vorliegenden Gutachten per Zufall ausgewählt und in die Bewertung eingeschlossen werden. Damit hätte sich eine Stichprobengrösse von n=820 Gutachten ergeben. Weil aber damit gerechnet werden musste, dass einige der ausgewählten Gutachten sich letztlich als nicht auswertbar herausstellen könnten, wurden die RAD bzw. IV-Stellen gebeten, eine etwas grössere Zahl auszuwählen (per Zufall jedes 16. der vorliegenden Gutachten). Diese Massnahme hat sich bewährt, da zwei RAD ihre Angaben zur Grundgesamtheit nachträglich korrigieren mussten, woraus für sie ein deutlich geringerer Stichprobenumfang resultierte als erwartet. Zudem erwiesen sich einige der zugesandten Gutachten aus verschiedenen Gründen als nicht auswertbar.

Nach der Meldung abschliessender Zahlen konnten als Gesamtzahl psychiatrischer Gutachten der Jahre 2008 und 2009 insgesamt 14'644 Gutachten ermittelt werden (vgl. Tabelle 3-2).

3.3.2 Stichprobe

Aus den genannten Zahlen zogen die IV-Stellen bzw. RAD per Zufall jedes 16. Gutachten. Ansprechpartner dafür waren nahezu überall die RAD selbst. In Basel werden die Gutachten getrennt von den beiden IV-Stellen Basel-Land und Basel-Stadt verwaltet und deshalb auch getrennt von beiden IV-Stellen an das Projektteam übersandt. Zur Ziehung wurden Listen mit allen psychiatrischen Gutachten der Jahre 2008 und 2009 erstellt, diese nach Nachname und Vorname des Exploranden bzw. der Explorandin aufsteigend sortiert und dann jeder 16. Eintrag markiert. Das entsprechende Gutachten wurde in Kopie, in der Regel auf einer CD, an die Projektgruppe geschickt.

Es ergaben sich die in Tabelle 3-2 genannten Stichprobengrössen. Zwischen der Anzahl angeforderter und tatsächlich erhaltener Gutachten ergaben sich kleinere Abweichungen, welche sich aber lediglich beim RAD/SMR 2 auf etwas höherem Niveau bewegen, und die mit insgesamt 16 Gutachten weniger als erwartet in einem Bereich liegen, welcher angesichts der Gesamtzahl keine relevanten Verzerrungen der Stichprobe bedeutet.

Tabelle 3-2: Gesamtzahl aller Gutachten

RAD/SMR	Gesamtzahl aller GA	Anzahl erhaltener GA	Anzahl auswertbarer GA
RAD/SMR 2	3'387	195	161
RAD/SMR 3	2'874	182	179
RAD/SMR 7	1'942	121	83
RAD/SMR 4	1'541	97	86
RAD/SMR 5	1'502	100	96
RAD/SMR 8	1'370	81	63
RAD/SMR 9	712	44	41
RAD/SMR 6	623	37	33
RAD/SMR 1	440	27	19
RAD/SMR 10	253	15	14
Summen	14'644	899	775

Insgesamt wurden der Projektgruppe 899 Gutachten zugesandt. Diese wurden in einer Datenbank registriert und auf Auswertbarkeit gesichtet.

Tatsächlich mussten aus dieser Roh-Stichprobe noch einige der Gutachten von der Auswertung ausgeschlossen werden. Einige der zugesandten Gutachten stammten dem Datum ihrer Abfassung nach nicht von 2008 oder 2009, sondern von 2007 oder 2010. Um bei den RAD, welche etwas zu wenige Gutachten liefern konnten, die letztlich auswertbar waren, noch eine möglichst ausreichende Stichprobengrösse zu sichern, wurden noch diejenigen Gutachten eingeschlossen, die zwar im Jahr 2010 verfasst wurden, bei denen die Exploration aber noch im Jahr 2009 stattgefunden hatte. Einzelne weitere Gutachten wiesen bei der Durchsicht teilweise eine rein somatische Indikation auf. Dies ist v.a. damit zu erklären, dass es einigen RAD nicht möglich war, ihre polydisziplinären Gutachten vor der Stichprobenziehung vorzuselektieren und dabei rein somatische Gutachten a priori auszuschliessen. Wenige andere Dokumente bestanden nur aus Arztberichten ohne das eigentliche Gutachten, zwei der Dokumente waren nur briefliche Hinweise, der Explorand sei nicht zur Untersuchung erschienen, weshalb das Gutachten nicht habe erstellt werden können.

Diese Gutachten mussten ausgeschlossen werden, sodass schliesslich 775 auswertbare Gutachten resultierten. Sie verteilen sich wie in Tabelle 3-2 dargestellt auf die Regionen.

Aus den genannten Gründen wich die Zahl gesendeter und auswertbarer Gutachten leicht von der Zahl der Gutachten ab, die aufgrund der Zahl aller Gutachten erwartet worden war. Die folgende Tabelle zeigt diese Abweichungen. An der Gesamtzahl aller psychiatrischen Gutachten der Jahre 2008 und 2009 hatte der RAD/SMR 3 einen Anteil von 19.6%. Kleinere Abweichungen bei der Stichprobenziehung führten dazu, dass dem Projektteam vom RAD/SMR 3 182 Gutachten zugesandt wurden, was einen Anteil von 20.2% an allen zugesandten Gutachten ausmacht. Nach der Sichtung mussten drei Gutachten ausgeschlossen werden, die resultierenden 179 auswertbaren Gutachten machen einen Anteil an den ausgewerteten Gutachten von 23.1% aus. RAD/SMR 3 ist somit in der Stichprobe, die den später folgenden inhaltlichen Ergebnissen zugrunde liegt, leicht mit 3.4% überrepräsentiert. Die deutlichste Abweichung nach unten hin liegt beim RAD/SMR 7 vor, welcher in der Stichprobe mit 2.6% fehlenden Gutachten leicht unterrepräsentiert ist, in nahezu gleichem Umfang der RAD/SMR 2.

Tabelle 3-3: Verteilung der Gutachten auf RAD/SMR

RAD/SMR	Anteil an Gesamtzahl GA	Anteil an erhaltenen GA	Anteil an auswertbaren GA	Differenz Anteil gesamt – Anteil auswertbar
RAD/SMR 2	23.1%	21.7%	20.7%	-2.4%
RAD/SMR 3	19.6%	20.2%	23.1%	+3.4%
RAD/SMR 7	13.3%	13.5%	10.7%	-2.6%
RAD/SMR 4	10.5%	10.8%	11.1%	+0.6%
RAD/SMR 5	10.3%	11.1%	12.4%	+2.1%
RAD/SMR 8	9.4%	9.0%	8.1%	-1.2%
RAD/SMR 9	4.9%	4.9%	5.3%	+0.4%
RAD/SMR 6	4.3%	4.1%	4.3%	±0.0%
RAD/SMR 1	3.0%	3.0%	2.4%	-0.6%
RAD/SMR 10	1.7%	1.7%	1.8%	+0.1%
Summen	100.0%	100.0%	100.0%	±0.0%

Nach der für die IV-Stellen bereits aufwändigen Ermittlung ihrer Gesamtzahlen mit anschließender Stichprobenziehung wäre es ihnen nicht mehr zumutbar gewesen, die (wenigen) fehlenden Gutachten noch einmal ermitteln und nachliefern zu müssen. Andererseits sollten von den Gutachten aus den bis dato leicht überrepräsentierten Regionen aber auch keine ausgeschlossen werden, um wieder völlige Parität zu erreichen, da damit auch die Gesamtzahl ausgewerteter Gutachten hätte weiter sinken müssen. Die insgesamt minimalen Abweichungen vom Plan wurden deshalb akzeptiert im Bewusstsein, dass sie sich höchstens unwesentlich auf die Ergebnisse auswirken würden.

3.4 Operationalisierung und Datenstruktur

Ausgehend von den von der BSV-Arbeitsgruppe entwickelten Formalen Kriterien psychiatrischer Gutachten wurde ein Kriterienkatalog und eine Checkliste für die Erhebung der formalen Qualität der zu untersuchenden psychiatrischen Gutachten erarbeitet (siehe Anhang 1).

Im Rahmen dieser Checklistenentwicklung wurden die im Kriterienkatalog zusammengestellten Kriterien operationalisiert und mit einer Expertengruppe abgestimmt. Die Operationalisierung der einzelnen Kriterien ist vom jeweiligen Kriterium abhängig: Kriterien wie „Vorliegen einer Diagnose nach ICD-10 respektive DSM IV“ können dichotom geratet werden (vorhanden vs. nicht vorhanden), während komplexe Kriterien wie die Angabe subjektiver Einschätzungen im Rahmen der objektiven Befunddarstellung abgestuft erfasst wurden. Für jedes Kriterium wurden maximale Punktwerte (Scores) vergeben. Der je nachdem unterschiedlich zu gewichtende Beitrag eines einzelnen Kriteriums an der gesamthaften Qualität des Gutachtens konnte so durch die Höhe des jeweiligen Scores abgebildet werden. Der maximal erreichbare Gesamtscore ist definiert über die Summe aller Sub-Scores bei vollumfänglicher Berücksichtigung aller Kriterien.

Die formalen Kriterien beinhalten auch eine Vielzahl von Kriterien und Subkriterien, die vorhanden sein können, aber nicht zwingend vom Gutachter bearbeitet werden müssen wie z.B. „einschneidende Erlebnisse“ (Punkt 3.2.9). Diese Kriterien wurden nachfolgend als „fakultative formale Kriterien“ bezeichnet (siehe Anhang 1) und wurden bei der Auswertung ebenfalls berücksichtigt und gesondert berechnet.

Kodierregeln, welche die Anwendung der Checkliste so präzise wie möglich beschreiben, sind entscheidend dafür, *wie* geratet wird, und bilden die Basis für die Erfüllung der Gütekriterien Objektiv-

tät, Validität und (Inter-Rater-) Reliabilität. Es wurde deshalb viel Sorgfalt auf die Entwicklung der Kodierregeln gelegt.

Im Rahmen des Entwicklungsprozesses dieser Kodierregeln wurde von der Prämisse ausgegangen, dass für die Bewertung der formalen Kriterien ausschliesslich der Erfüllungsstand bezüglich formaler Kriterien unabhängig von (inhaltlichem) Umfang und Angemessenheit der Darstellung für den individuellen Fall beurteilt wird. So wurde bspw. festgelegt, dass Kriterien wie „Familianamnese / Heredität“ (Punkt 3.2.3) als erfüllt gelten, wenn aus dem Gutachtentext mindestens eine Angabe zu Eltern und / oder Geschwistern des Exploranden (z.B. Gesundheit oder Berufe der Eltern oder früheres Familienleben) entnommen werden konnte.

Um die formale Vollständigkeit des Punktes „Beantwortung der Fragen (6.4)“ beurteilen zu können, wurde in Abstimmung mit Experten als Mindestanforderung für die Vollständigkeit folgende Inhalte definiert: Angabe zu Arbeitsfähigkeit in h/%, Differenzierung bisherige und adaptierte Tätigkeit, Aussagen zur Diagnose (ICD-10 bzw. DSM IV), Auswirkungen auf die Arbeitsfähigkeit, Prognose und Therapieempfehlung.

Weiteres Augenmerk galt der Beurteilung des Punktes „Herleitung und Darstellung der Funktions- und Fähigkeitsstörungen“ (6.3 a). Bei diesem Punkt wurde zusätzlich erfasst, ob die Herleitung bzw. Begründung mit (halb-)standardisierten Hilfsmitteln wie Checklisten oder bspw. ICF erfolgte.

Diese im Anhang 1 dargestellte Checkliste dient zusammen mit den detaillierten Kodierregeln (Anhang 2) als Beurteilungsgrundlage für jedes Gutachten bzw. als Analyseinstrument zur Erhebung der formalen Gutachtenqualität.

3.5 Auswertung der Gutachten

Die Auswertung bzw. Bewertung der Gutachten erfolgte durch fünf erfahrene und speziell geschulte ärztliche bzw. psychologische Fachpersonen. Für speziell schwierig zu beurteilende Gutachten standen zusätzlich die Gutachtenexperten der Projektgruppe zur Verfügung, um auch hier die verlässliche Beurteilung sicher zu stellen.

Je Gutachten wurde ein Datensatz generiert. Die Daten wurden anschliessend in ein Tabellenkalkulationsprogramm eingegeben und von dort in das Statistik-Programm IBM SPSS Statistics 19 übernommen und mit diesem Programm analysiert.

3.5.1 Statistische Verfahren

Die Analysen dienen in erster Linie der Beschreibung des Ausgangsstandes vor Einführung der Leitlinien zur Begutachtung, weshalb deskriptive Verfahren wie Häufigkeitsauszählungen angemessen sind.

Fragestellung 1 wurde durch die Analyse der Gesamtstichprobe, die Fragestellungen 2 bis 4 wurden durch Subgruppenanalysen bearbeitet.

Für die Vergleiche zwischen Gruppen von Gutachten je nach Institution und Region werden non-parametrische Verfahren angewandt, d.h. solche, die hinsichtlich ihrer Verteilungsannahmen keine hohen Voraussetzungen besitzen. Die abhängigen Variablen, z.B. zusammenfassende Indizes, die beinhalten, wie viele Kriterien eines Themenbereichs erfüllt sind, können theoretisch als intervallskaliert angesehen werden. Jedoch setzten sich die Indizes teilweise nur aus sehr wenigen Einzelwerten zusammen (im Bereich „Aktenauszug“ wurden z.B. 3 Kriterien vorgegeben, der entsprechende Index kann also lediglich Werte zwischen 0 und 3 annehmen). Auch die Verteilungsvoraussetzungen für die Anwendung von parametrischen Testverfahren (Normalverteilung, Varianzhomogenität etc.) sind verletzt. Daher werden non-parametrische Tests für ordinalskalierte Daten eingesetzt. Für den Vergleich der zentralen Tendenz einer ordinalskalierten Variablen, die in zwei Teilstichproben erho-

ben wurde, eignet sich der Mediantest. Er ist zudem robust gegenüber Ausreissern. Wurde die ordinalskalierte Variable in mehr als zwei Stichproben erhoben, kann der H-Test von Kruskal & Wallis (Rangvarianzanalyse) eingesetzt werden, welcher eine Erweiterung des Mediantests darstellt. Aufgrund der Grösse der untersuchten Stichproben und der mehrfachen Testung derselben Subgruppen wurden Unterschiede zwischen Subgruppen auf dem 1%-Fehlerniveau getestet.

Eine übliche Darstellung für Unterschiede zwischen verschiedenen Gruppen (z.B. den Arten von Gutachten, den Gutachten-erstellenden Institutionen oder den Regionen) in der Verteilung von Werten mit dem vorliegenden Datenniveau sind Boxplots (vgl. Abbildung 4-28). Für jede Gruppe von Gutachten zeigt ein solches Diagramm die schematisierte Verteilung der Werte, wobei ein rechteckiger Kasten den Wertebereich darstellt, in dem 50% aller Gutachten zu finden sind (also die Werte zwischen dem 25. und dem 75. Perzentil). Innerhalb dieses Bereichs liegt der Median aller Werte aus dieser Gruppe, der durch eine dickere waagrechte Linie markiert wird. Die über dem 75. bzw. unter dem 25. Perzentil liegenden Werte werden durch senkrechte Linien dargestellt, deren jeweilige Enden also den gesamten Wertebereich (Range) dieser Gruppe anzeigen. Ausgenommen in dieser Darstellung sind einzelne Ausreisser- und Extremwerte, also Gutachten, deren Beurteilung übermässig von den anderen Gutachten in dieser Gruppe abweicht. Auf die Darstellung dieser insgesamt sehr wenigen Werte wird hier verzichtet, weil ihr Einfluss auf das Gesamtergebnis aufgrund der hohen Fallzahlen irrelevant ist und ihre Darstellung die Diagramme unnötig verkomplizieren würde¹.

Der Katalog formaler Kriterien, welcher den vorliegenden Analysen zugrunde liegt, gliedert sich in sechs Themenbereiche (z.B. Themenbereich 1. „Ausgangslage und Formelles“), davon einer fakultativ, die sich wiederum in Abschnitte und Unterabschnitte gliedern (z.B. Abschnitt 1.1 „Abwicklung des Gutachtensauftrages/ Formelles“, Unterabschnitt 1.1.1 „Angaben zum Auftraggeber“). Insgesamt wurden gegliedert nach diesen inhaltlichen Bereichen von der Expertengruppe 62 obligatorische und 25 fakultative Kriterien formuliert.

Für die Auswertungen des Status quo der formalen Qualität der Gutachten, dem Hauptzweck der vorliegenden Studie, werden die Daten aus der Gesamtstichprobe zu Grunde gelegt. Hierfür wird in Kapitel 4 die Vollständigkeit der Gutachten bezogen auf jedes einzelne Kriterium dargestellt. Für die zusätzlichen Vergleiche zwischen Subgruppen der Gutachten wäre eine so feine Gliederung allerdings nicht übersichtlich darstellbar. Deshalb wurden für die Gruppenvergleiche die Kriterien in den fünf obligaten Themenbereichen jeweils zusammengefasst; zudem wurde eine Gesamtsumme über alle 62 obligatorischen Kriterien berechnet. Diese aggregierten Werte sind allerdings aus methodischer Sicht nur eingeschränkt interpretierbar, da dabei grossteils inhaltlich sehr unterschiedliche Kriterien, die inhaltlich verschiedene Bedeutungen haben und von ihrer Relevanz für die Begutachtung her eigentlich unterschiedlich gewichtet werden müssten. Zudem sind die Fallzahlen insgesamt sehr hoch, so dass selbst zahlenmässig kleine Unterschiede schnell statistische Signifikanz erzielen. Dieser Effekt wird noch dadurch verstärkt, dass die verglichenen Gruppen, z.B. von Gutachten aus verschiedenen Regionen, sehr unterschiedlich stark besetzt sind.

3.6 Interrater-Reliabilität

Zur Bestimmung der Beurteilerübereinstimmung (Interrater-Reliabilität) wurde zufallsbasiert ein Subsample von n= 50 Gutachten gezogen. Nach Ausschluss von zwei Gutachten, welche aus den genannten Gründen auch aus der Hauptstichprobe ausgeschlossen werden mussten, blieben 48 Gutachten, welche von einem zweiten Rater anhand der Checkliste ausgewertet wurden.

¹ Werte, die mehr als das Eineinhalbfache des Interquartilabstandes vom Ende der Box entfernt liegen, werden als Ausreisser klassifiziert, und Werte mit einem Abstand vom Ende der Box von mehr als dem Dreifachen des Interquartilabstandes als Extremwerte. Pro Diagramm handelt es sich dabei in der Regel um weniger als 10 Gutachten, die in der graphischen Darstellung ausgelassen werden.

Die 48 Paare von Datensätzen wurden statistisch ausgewertet. Das am häufigsten angewandte Mass zur Bestimmung der Übereinstimmung bei kategorialen Daten ist Cohens κ (kappa) (Wirtz & Caspar, 2002). Es wurde auch hier in seiner Variante für zwei Rater angewandt. Cohen ist zufallskorrigiert, d.h. es berechnet den Grad an Übereinstimmung zwischen den Beurteilern, der über rein zufällige gleiche Einschätzungen hinaus geht. Damit ist es ein konservativeres Mass als die prozentuale Übereinstimmung: Auch bei relativ grossem prozentuaalem Anteil an gleichen Bewertungen der beiden Rater bzw. relativ wenigen Abweichungen in den Beurteilungen kann das κ bereits geringere Werte annehmen. Die Interpretation der Höhe des Wertes kann nicht schematisch erfolgen. Sie ist einerseits abhängig von den vorliegenden Daten: Wenn wie im vorliegenden Fall für die Frage, ob im Gutachten Informationen zu einem bestimmten Kriterium vorhanden sind oder nicht, also nur zwei Antwortmöglichkeiten vorliegen (vorhanden/nicht vorhanden), fällt sein Wert systematisch niedriger aus als bei einer grösseren Zahl von Antwortkategorien. Andererseits muss bei der Bewertung seiner Höhe auch die Schwere der Beurteilungsaufgabe berücksichtigt werden. Beim Kriterium 1.1.2 a) „Angaben zum Exploranden: Name“ wird man eine sehr grosse Übereinstimmung verlangen, um von einer akzeptablen Qualität des Ratingsystems zu sprechen, während der Punkt 3.2.1 f) „Zukunftsvorstellungen bzgl. Krankheitsverlauf“ schon an sich schwerer zu beurteilen ist.

Dementsprechend variieren die κ -Werte für die 62 obligatorischen Kriterien von $\kappa=-0.121$ für das Item 4.3.1b) „Beschreibung der Persönlichkeit“ bis $\kappa=1.000$ für die soziodemographischen Angaben zum Exploranden und die Deskription des Auftraggebers des Gutachtens.

Der Median liegt bei $\kappa=0.487$, was angesichts der Heterogenität der Gutachten in Sprache und Stil als für den vorliegenden Zweck völlig ausreichend angesehen werden kann.

4 Ergebnisse aus der Gesamtstichprobe

Im Folgenden werden zunächst die ausgewerteten Gutachten hinsichtlich verschiedener Basisdaten beschrieben, bevor dann ihre formale Qualität im Sinn der Kriterien aus dem Kodierschema² dargestellt wird. Die obligatorischen Kriterien werden vertieft behandelt und in den folgenden Kapiteln auch auf Unterschiede zwischen den Gutachtenteilen, zwischen Institutionen und Regionen geprüft.

Die fakultativen Kriterien (Nummern 1.3.2, 1.3.4, 1.3.5, 2.d), 3.2.6c), 3.2.9, 4.2, 4.3, 4.4, 5, 6.2, 7) eignen sich aufgrund ihrer Heterogenität und der kleinen Zahl von Gutachten, in denen diese Kriterien verwendet wurden, nicht für systematische Gruppenvergleiche. Die fakultativen Kriterien wurden deshalb lediglich deskriptiv-statistisch ausgewertet, die entsprechenden Ergebnisse finden sich in Abschnitt 4.4.

4.1 Basisdaten

4.1.1 Gutachten

Die Gutachten stammen jeweils etwa zur Hälfte aus den beiden Erhebungsjahren (2008: $n=327$; 42.2%; 2009: $n=448$; 57.8%)³.

² Eine Übersicht zu allen Kriterien findet sich in der Checkliste für die Kodierung der formalen Qualitätskriterien im Anhang 1.

³ Insgesamt $n=18$ Gutachten, die Anfang 2010 beim RAD eingingen, bei denen die Exploration aber noch im Jahr 2009 stattgefunden hatte, wurden dem Jahr 2009 zugeschlagen.

Von den 775 ausgewerteten Gutachten stammt der grösste Anteil – der Verteilung aller Gutachten auf die Gesamtschweiz entsprechend – vom RAD/SMR 3, gefolgt vom RAD/SMR 2 und mit grösserem Abstand von RAD/SMR 5 und RAD/SMR 4 (Abbildung 4-1).

Abbildung 4-1: Beauftragender RAD/SMR

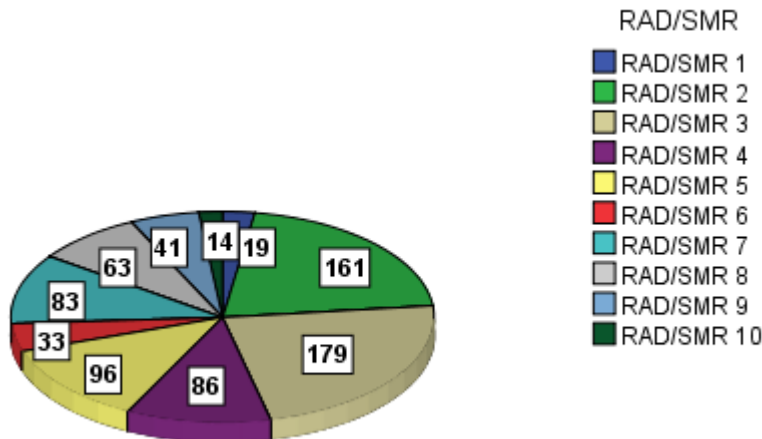


Tabelle 4-1 gibt einen Überblick zu den Institutionen, in denen die Gutachten angefertigt wurden.

Tabelle 4-1: Institution

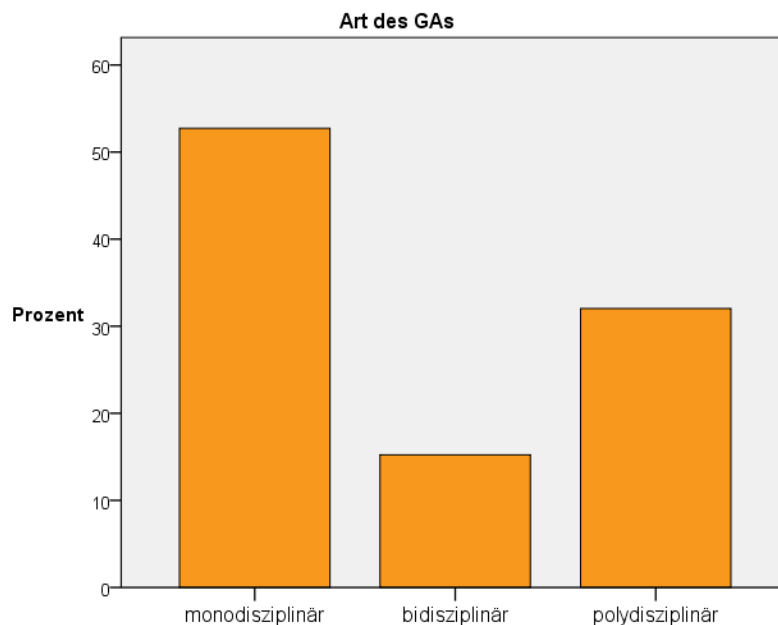
	MEDAS		RAD		Frei praktizierende/r Psychiater/in		Sonstige	
	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
RAD/SMR 1	2	10,5%	0	0%	12	63,2%	5	26,3%
RAD/SMR 2	40	24,8%	33	20,5%	76	47,2%	12	7,5%
RAD/SMR 3	20	11,2%	74	41,3%	75	41,9%	10	5,6%
RAD/SMR 4	29	33,7%	45	52,3%	1	1,2%	11	12,8%
RAD/SMR 5	0	0%	79	82,3%	2	2,1%	15	15,6%
RAD/SMR 6	0	0%	13	39,4%	18	54,5%	2	6,1%
RAD/SMR 7	19	22,9%	1	1,2%	63	75,9%	0	0%
RAD/SMR 8	0	0%	22	34,9%	29	46,0%	12	19,0%
RAD/SMR 9	13	31,7%	3	7,3%	10	24,4%	15	36,6%
RAD/SMR 10	0	0%	0	0%	10	71,4%	4	28,6%
Gesamt	123	100.0	270	100.0	296	100.0	86	100.0

Knapp 40% der Gutachten wurden von frei praktizierenden Psychiater/innen, ca. 35% von einem RAD, 16% von einem MEDAS und 11% von anderen Einrichtungen, darunter v.a. Spitälern, erstellt.

Die verschiedenen Institutionen werden von den RAD/SMR in unterschiedlichem Ausmass mit Gutachten beauftragt. Vier der zehn RAD/SMR vergeben keine Gutachtaufträge an MEDAS, andere RAD/SMR lassen bis zu ein Drittel ihrer Gutachten von einem MEDAS anfertigen. Der Anteil der Gutachten, die vom RAD selbst angefertigt werden, variiert gar von 0 bis 80 Prozent. Abgesehen von drei RAD/SMR, bei denen nur ein sehr kleiner oder kleinerer Anteil der Gutachten von frei praktizierenden Psychiaterinnen oder Psychiatern kommt, wird in der Regel von dieser Gruppe der Hauptteil der Gutachten verfasst. Bei drei RAD/SMR besitzen auch die sonstigen Institutionen für das Abfassen von Gutachten eine grosse Bedeutung.

Grob die Hälfte der Gutachten (52.3%) war monodisziplinär, 15.1 % bidisziplinär und ein knappes Drittel (31.7%) polydisziplinär (Abbildung 4-2). N=7 Gutachten (0.9%) liessen sich nicht eindeutig bestimmen (Abbildung 4-2).

Abbildung 4-2: Art des Gutachtens



Zwischen Art des Gutachtens und Institution besteht ein enger Zusammenhang in dem Sinn, dass Gutachten von MEDAS nahezu immer polydisziplinär abgefasst werden und Gutachten von frei Praktizierenden zu ca. 80% monodisziplinär sind. Bei den RAD/SMR sind 58% der GA monodisziplinär, der Rest verteilt sich etwa gleichmässig auf bi- und polydisziplinäre Gutachten. Sonstige Gutachtenstellen fertigen zu 60% polydisziplinäre GA an, zu ca. 30% mono- und zu ca. 10% bidisziplinäre GA.

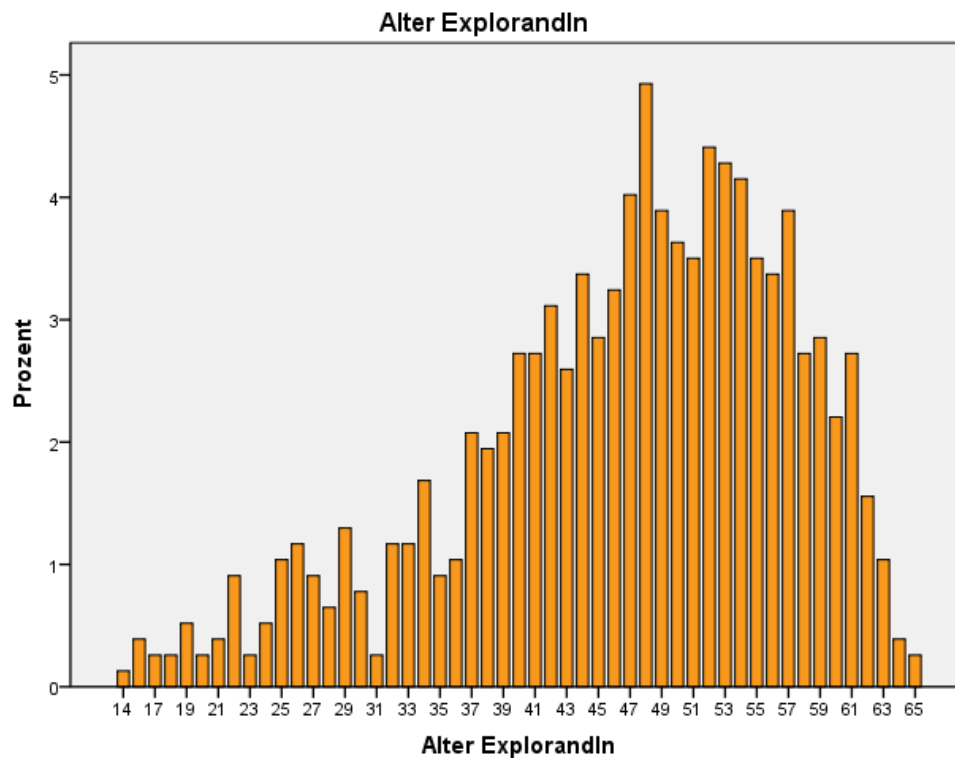
4.1.2 Explorandinnen und Exploranden

4.1.2.1 Soziodemographie

405 der untersuchten Personen waren männlich (52.3%), 370 Personen weiblich (47.7%).

Das Alter der untersuchten Personen lag im Mittel bei 46.6 Jahren (Standardabweichung SD=10.483 Jahre) mit einem Minimum von 14 und einem Maximum von 65 Jahren. Die Altersverteilung dieser Stichprobe wird in Abbildung 4-3 dargestellt.

Abbildung 4-3: Altersverteilung



Die Untersuchten hatten im Mittel einen Grad der Arbeitsfähigkeit von 59.7%, allerdings mit einer sehr grossen Variationsbreite (SD=34.610 %). Es wurde die gesamte Bandbreite von vollständig erhaltener bis zu völlig entfallener Arbeitsfähigkeit ausgeschöpft.

Die ExplorandInnen sind ganz überwiegend Schweizer StaatsbürgerInnen. Die nach Grösse folgenden Herkunftsländer sind Italien und die Türkei. Die anderen Nationalitäten – vereinzelt auch mit doppelter Staatsbürgerschaft – verteilen sich über eine grosse Zahl von Ländern vorwiegend aus dem südeuropäischen Raum (Tabelle 4-2).

Tabelle 4-2: Herkunftsland (grösste Gruppen)

Herkunftsland	Häufigkeit	Prozent
Schweiz	382	49.3
Italien	65	8.4
Türkei	61	7.9
Serbien	32	4.1
Kosovo	32	4.1
Bosnien	28	3.6
Mazedonien	18	2.3
Portugal	18	2.3
Deutschland	13	1.7
unbekannt	25	3.2

Insgesamt betrug der Anteil von begutachteten ImmigrantInnen also 50.7%. In den Jahren 2008 und 2009 lag der Anteil von Personen, die in der Schweiz leben, aber ein anderes Herkunftsland haben, bei 21.68% bzw. 22.02%⁴. Baer et al. (2009) hatten festgestellt, dass 34% der aufgrund psychischer Leiden

⁴ Quelle: Bundesamt für Statistik, eigene Berechnungen.

Berenteten in die Schweiz immigriert sind. Vor diesem Hintergrund sind bei den ausgewerteten Gutachten Explorandinnen und Exploranden mit einem Herkunftsland ausserhalb der Schweiz sowohl gegenüber ihrem Anteil an der Wohnbevölkerung und an den Erwerbstätigen, aber auch an den Berenteten deutlich überrepräsentiert.

4.1.2.2 Diagnosen

Bei 737 (95.1%) der 775 ausgewerteten Gutachten war mindestens eine Diagnose angegeben, bei 38 Gutachten fehlte die Angabe zur Erstdiagnose völlig.

Bei 39 Gutachten (5.3%) war angegeben, dass keine psychiatrische Diagnose vorliege.

Bei 415 Gutachten war auch eine zweite Diagnose gestellt, bei 184 davon auch eine dritte.

Die psychiatrischen Erstdiagnosen entstammen einem weiten Bereich des Kapitels V der ICD-10, „Psychische und Verhaltensstörungen“, vereinzelt wurden aber auch Erstdiagnosen aus anderen Gruppen vergeben. Die grösste Gruppe machen die PatientInnen mit Rezidivierenden depressiven Störungen (F33) aus, gefolgt von Depressiven Episoden (F32) und Somatoformen Störungen (F45); Tabelle 4-3).

Tabelle 4-3: Erstdiagnosen nach ICD-10

1. Diagnose gruppiert	Häufigkeit	Prozent
F33 - Rezidivierende depressive Störung	129	17.5
F32 - Depressive Episode	110	14.9
F45 - Somatoforme Störungen	88	11.9
F60 - Spezifische Persönlichkeitsstörungen	49	6.6
F43 - Reaktionen auf schwere Belastungen und Anpassungsstörungen	45	6.1
<i>keine psychiatrische Diagnose</i>	39	5.3
F34 - Anhaltende affektive Störungen	31	4.2
F41 - Andere Angststörungen	30	4.1
F61 - Kombinierte und andere Persönlichkeitsstörungen	29	3.9
F54 - Psychologische Faktoren oder Verhaltensfaktoren bei anderenorts klassifizierten Krankheiten	17	2.3
F10 - Psychische und Verhaltensstörungen durch Alkohol	15	2.0
F48 - Andere neurotische Störungen	15	2.0
F20 - Schizophrenie	13	1.8
F31 - Bipolare affektive Störung	13	1.8
F40 - Phobische Störungen	12	1.6
F07 - Persönlichkeits- und Verhaltensstörung aufgrund einer Krankheit, Schädigung oder Funktionsstörung des Gehirns'	10	1.4
F19 - Psychische und Verhaltensstörungen durch multiplen Substanzgebrauch und Konsum anderer psychotroper Substanzen	8	1.1

Nur Diagnosegruppen > 1.0% der Gutachten sind aufgeführt.

4.2 Obligatorische Kriterien

Im Folgenden wird berichtet, inwiefern die 775 Gutachten die Kriterien der Checkliste erfüllt haben. Dabei wird jeweils dargestellt, bei welchem Prozentanteil der Gutachten die von der Checkliste geforderte Information enthalten war.

Die Kriterien gliedern sich in die folgenden Themenbereiche und Abschnitte:

1. Ausgangslage und Formelles
 - 1.1 Abwicklung des Gutachtensauftrages/ Formelles
 - 1.2 Anlass der Begutachtung
 - 1.3 Grundlagen für das Gutachten
2. Aktenauszug
3. Untersuchung
 - 3.1 Spontane Angaben des Expl. / offenes Interview
 - 3.2 Vertiefendes Interview - je nach Fragestellung –
4. Befund
 - 4.1 Verhaltensbeobachtungen / äussere Erscheinung
 - 4.2 Sprachliche Verständigung
 - 4.3 Objektive Untersuchungsbefunde
 - 4.4 Zusatzbefunde
5. Angaben von Drittpersonen
6. Medizinische und versicherungsmedizinische Beurteilung
 - 6.1 Resumé
 - 6.2 Beurteilung bzw. Einordnung der Befunde
 - 6.3 Diskussion und versicherungsmedizinische Würdigung
- 6.4 Beantwortung der Fragen
7. Beilagen

Die Themenbereiche 5 und 7 sowie einzelne Kriterien in den anderen Themen sind fakultativ und werden deshalb für die detaillierte Beschreibung der formalen Qualität und auch für die Gruppenvergleiche nicht verwendet.

Die Nummerierung der folgenden Abschnitte und der Abbildungsüberschriften entspricht den Nummern der Kriterien in der Checkliste.

Themenbereich 1. „Ausgangslage und Formelles“

Abschnitt 1.1 „Abwicklung des Gutachtensauftrages/ Formelles“

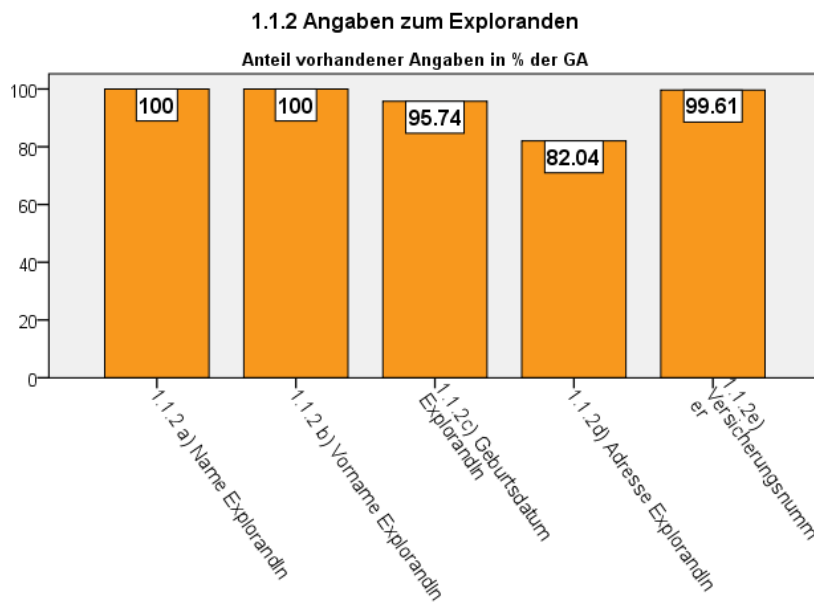
Unterabschnitt 1.1.1 „Angaben zum Auftraggeber“

Angaben zum Auftraggeber waren bei 764 der 775 Gutachten (98.6 %) vorhanden.

Unterabschnitt 1.1.2 „Angaben zum Exploranden“

Die Angaben zum Exploranden waren in nahezu allen Gutachten vorhanden, abgesehen von deren Adresse, welche in etwa einem Fünftel der Gutachten nicht enthalten war.

Abbildung 4-4: Angaben zum Exploranden



Unterabschnitt 1.1.3 „Angaben zum Gutachter“

Die geforderten Angaben zum Gutachter bzw. der Gutachterin waren in nahezu allen Fällen vorhanden.

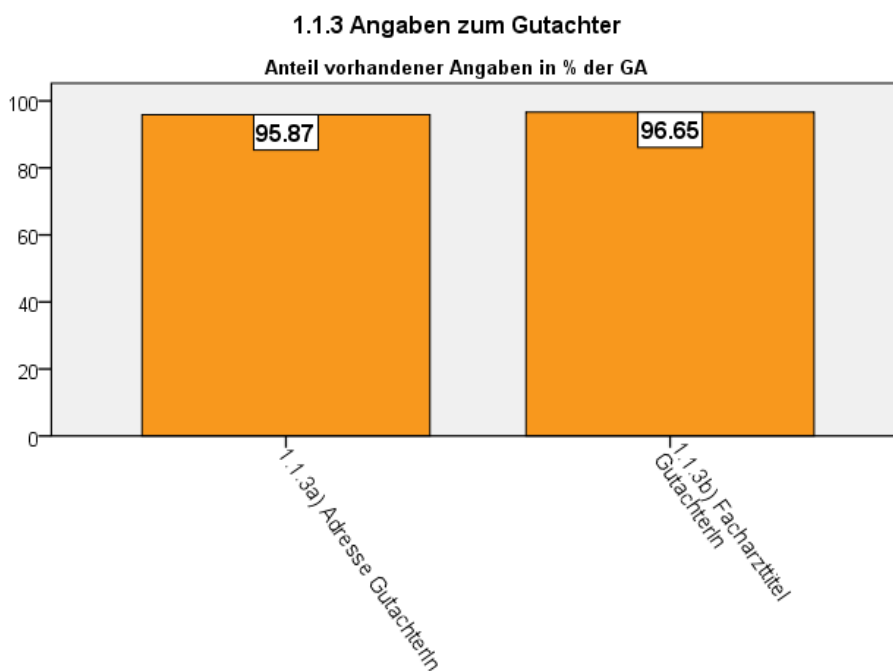
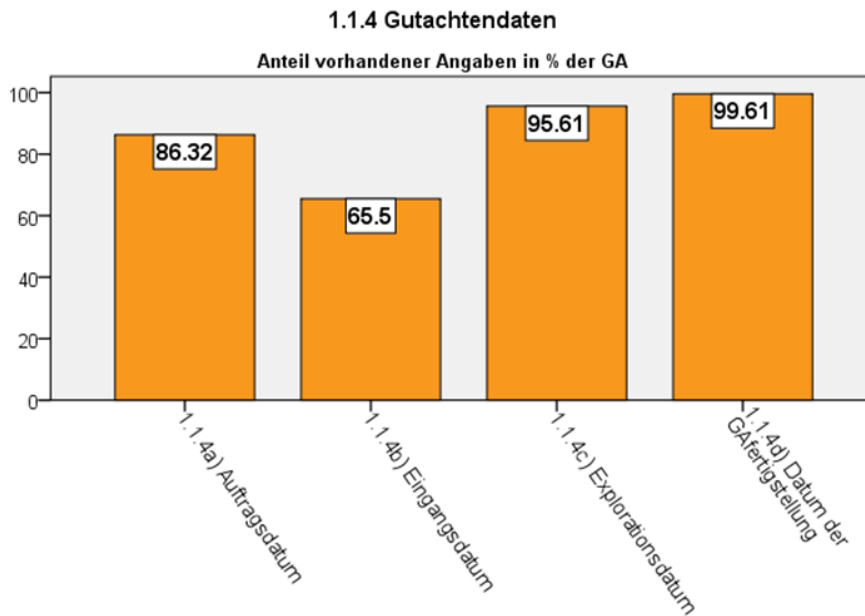


Abbildung 4-5: Angaben zum Gutachter

Unterabschnitt 1.1.4 „Gutachtendaten“

Wie die folgende Abbildung zeigt, gab es bei den Beschreibungsmerkmalen für die Gutachten teils leichte Abweichungen vom „Soll“, insbesondere was deren Eingangsdatum angeht.

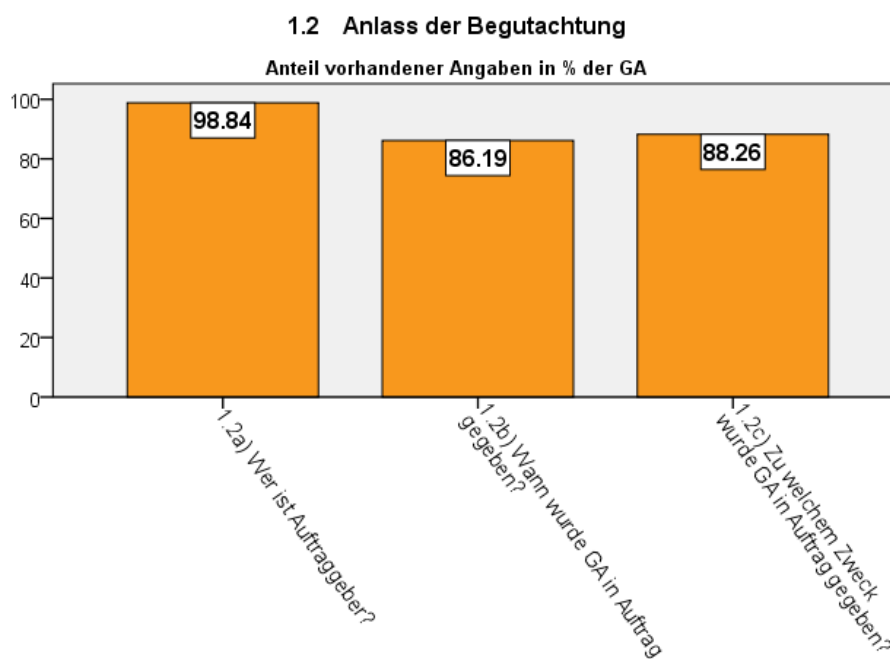
Abbildung 4-6: Gutachtendaten



Abschnitt 1.2 „Anlass der Begutachtung“

Beim Anlass wurde in nahezu allen Fällen der Auftraggeber des Gutachtens genannt, in wenigen Fällen fehlten das Datum des Auftrages und dessen Zweck.

Abbildung 4-7: Anlass der Begutachtung



Abschnitt 1.3 „Grundlagen für das Gutachten“

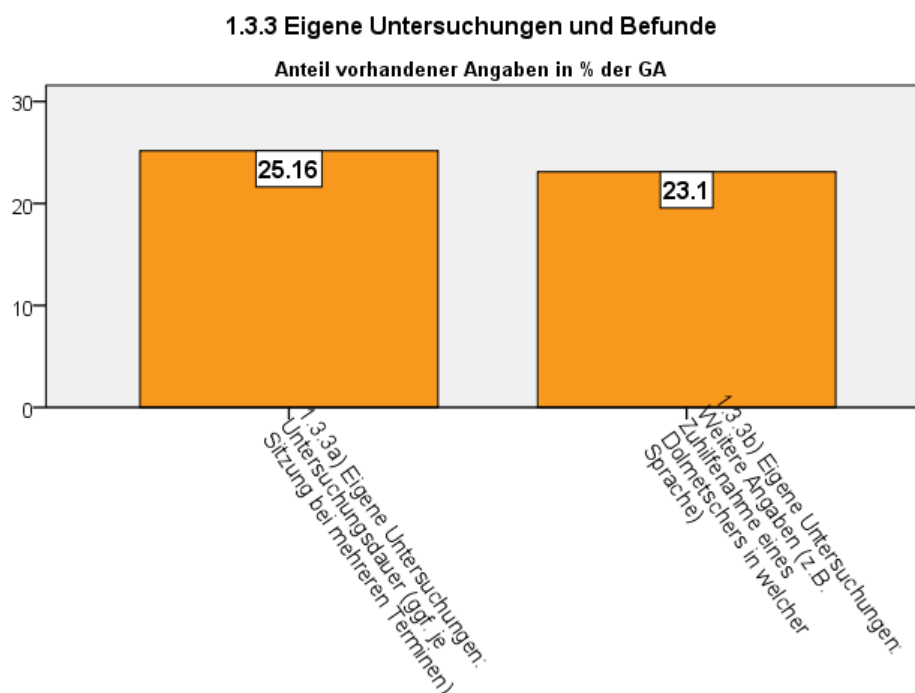
Unterabschnitt 1.3.1 „Aufzählung der vom Auftraggeber zur Verfügung gestellten Aktendossiers“

Die vom Auftraggeber zur Verfügung gestellten Aktendossiers wurden in 741 der 775 Gutachten (95.6%) erwähnt.

Unterabschnitt 1.3.3 „Eigene Untersuchungen und Befunde“

Hinsichtlich der eigenen, d.h. durch den Gutachter bzw. die Gutachterin durchgeführten Untersuchungen sieht die Checkliste vor, dass die Dauer der Untersuchungen und – falls relevant – weitere Angaben, z.B. zur Zuhilfenahme eines Dolmetschers etc., erfasst werden. Angaben zur Untersuchungsdauer wurden nur in der Minderzahl der Gutachten gemacht (s. Abbildung 4-8). Dies kann allerdings grossteils darauf zurückgeführt werden, dass viele der Gutachten allein auf Basis der Aktenlage und einer Exploration, nicht aber auf zusätzlichen medizinischen Untersuchungen beruhen und diese Angabe dann als überflüssig erachtet wurde. Auch die weiteren Angaben sind relativ selten, als Hypothese, weil es in vielen Fällen nichts Aussergewöhnliches mehr zu den Untersuchungen zu berichten gab.

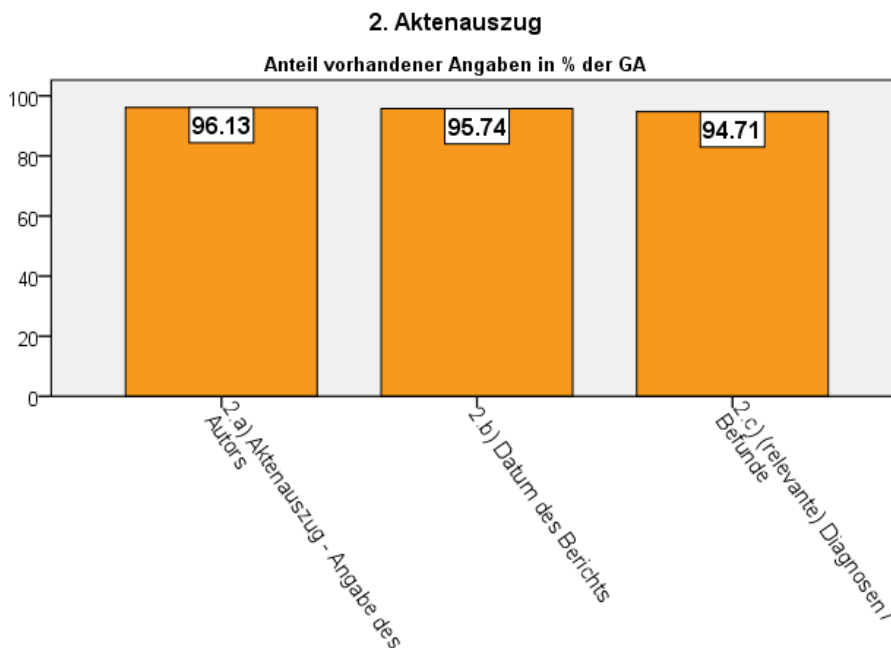
Abbildung 4-8: Eigene Untersuchungen und Befunde



Themenbereich 2. „Aktenauszug“

Beim Aktenauszug wurden in nahezu allen Fällen die AutorInnen der Akten angegeben, etwas seltener das Datum der Abfassung und die darin aufgeführten Diagnosen und Befunde.

Abbildung 4-9: Aktenauszug



Themenbereich 3. „Untersuchung“

Abschnitt 3.1 „Spontane Angaben des Expl. / offenes Interview“

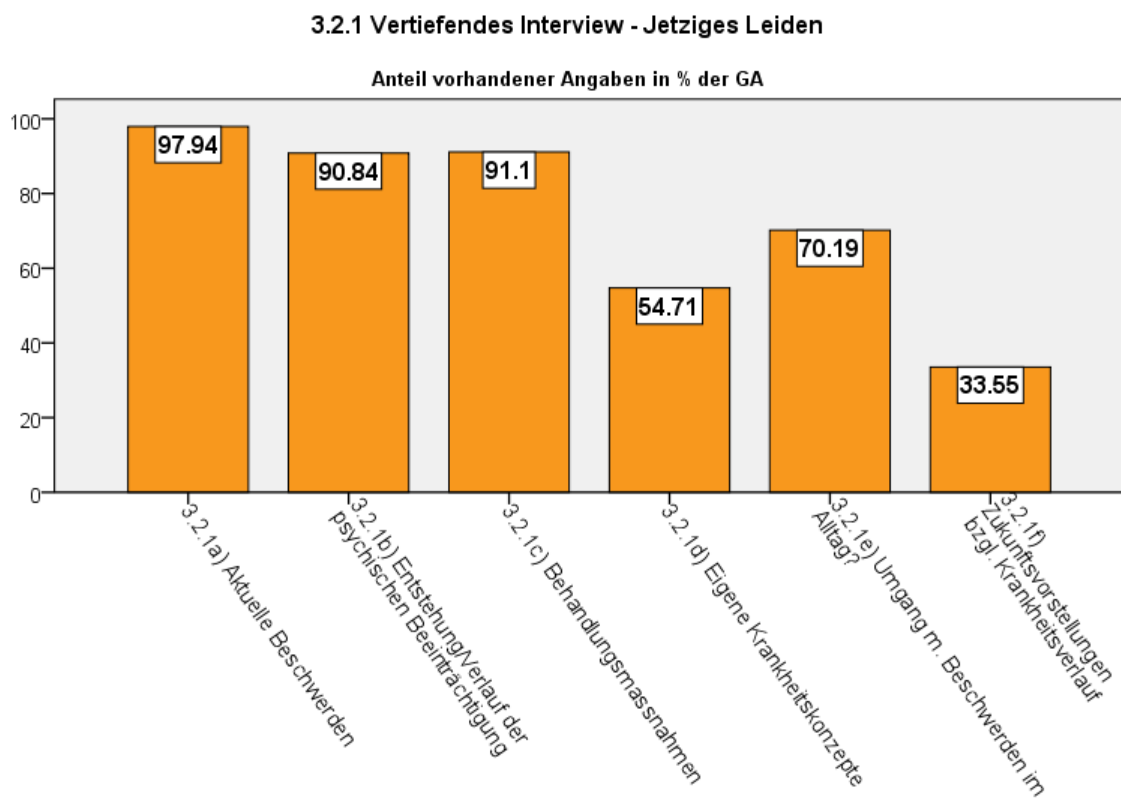
Nur in etwas mehr als einem Drittel der Gutachten wurden spontane Angaben des Exploranden aus dem offenen Interview dokumentiert. In welchem Umfang die Explorierten spontane Angaben machen, ist stark abhängig von der Art und Weise, wie offen bzw. direktiv der Gutachter die Exploration durchführt. Zudem kann eine systematische Exploration aus Sicht des Gutachters bereits alle aus seiner Sicht relevanten Themenbereiche abdecken, so dass spontane Äusserungen des Pat. nur einen kleinen zusätzlichen Informationsgewinn bringen, der evtl. nicht als berichtenswert angesehen wurde. Eine grosse Variabilität ist also schon in der Sache angelegt.

Abschnitt 3.2 „Vertiefendes Interview – je nach Fragestellung“

Unterabschnitt 3.2.1 „Jetziges Leiden“

Im Teil zum vertiefenden Interview wurden in der weit überwiegenden Zahl der Gutachten Angaben zu den aktuellen Beschwerden, Entstehung und Verlauf der psychischen Beeinträchtigungen und Behandlungsmassnahmen gemacht, etwas seltener zum Umgang mit den Beschwerden im Alltag. Nur in der Hälfte der Gutachten waren Angaben zum Krankheitskonzept aus Sicht des Exploranden zu finden, zu gut zwei Dritteln der Umgang des Exploranden mit seinen Beschwerden im Alltag und in einem Drittel zu dessen diesbezüglichen Zukunftsvorstellungen.

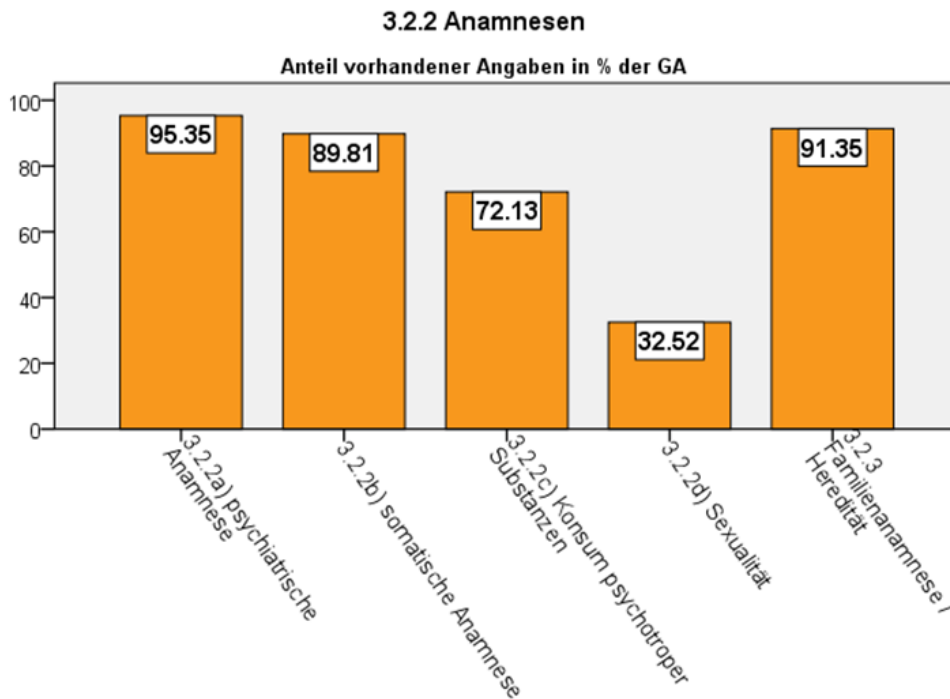
Abbildung 4-10: Jetziges Leiden



Unterabschnitt 3.2.2 „Anamnesen“

Hinsichtlich der Anamnese wurden in der Regel die psychiatrische, die somatische und die Familienanamnese aufgeführt. Drei Viertel der Gutachten enthielten die Vorgeschichte hinsichtlich des Konsums psychotroper Substanzen, nur ein Drittel zur Sexualität. Je nach Vorgeschichte und Fragestellung scheint insbesondere der letzte Punkt den Gutachtenden also häufig nicht als relevant erschienen sein.

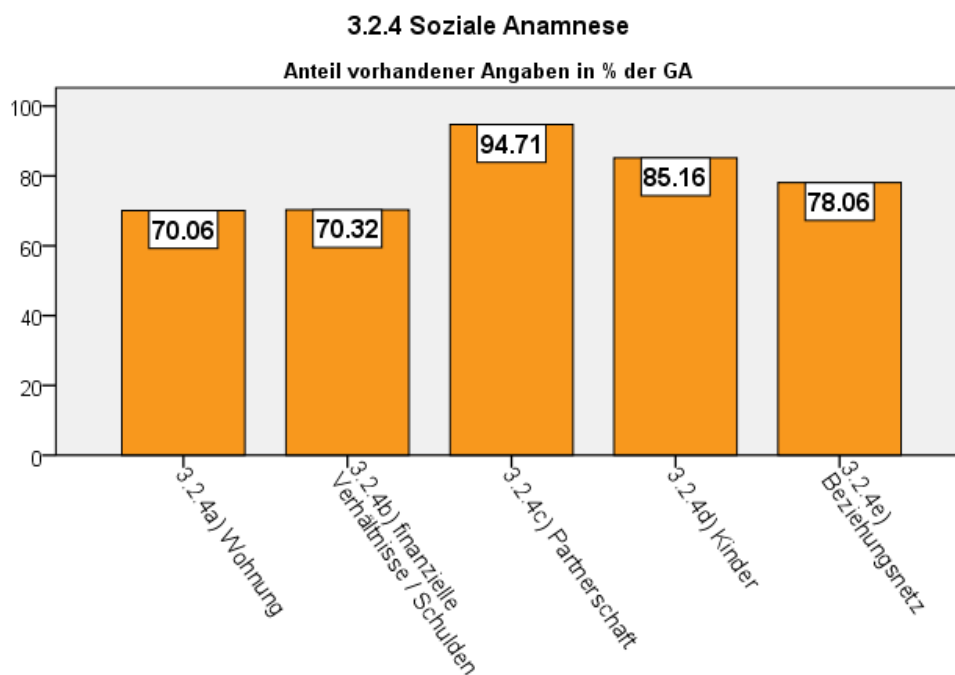
Abbildung 4-11: Anamnesen



Unterabschnitt 3.2.4 „Soziale Anamnese“

Als Aspekte der Sozialen Anamnese finden sich in mindestens 70% der ausgewerteten Gutachten, v.a. die Aspekte Partnerschaft und Kinder, seltener Angaben zum Beziehungsnetz, zu den finanziellen Verhältnissen und zur Wohnsituation, obwohl gerade bei einer Begutachtung diese Angaben zur Beurteilung der vorhandenen Ressourcen und der Arbeitsfähigkeit sehr bedeutsam wären.

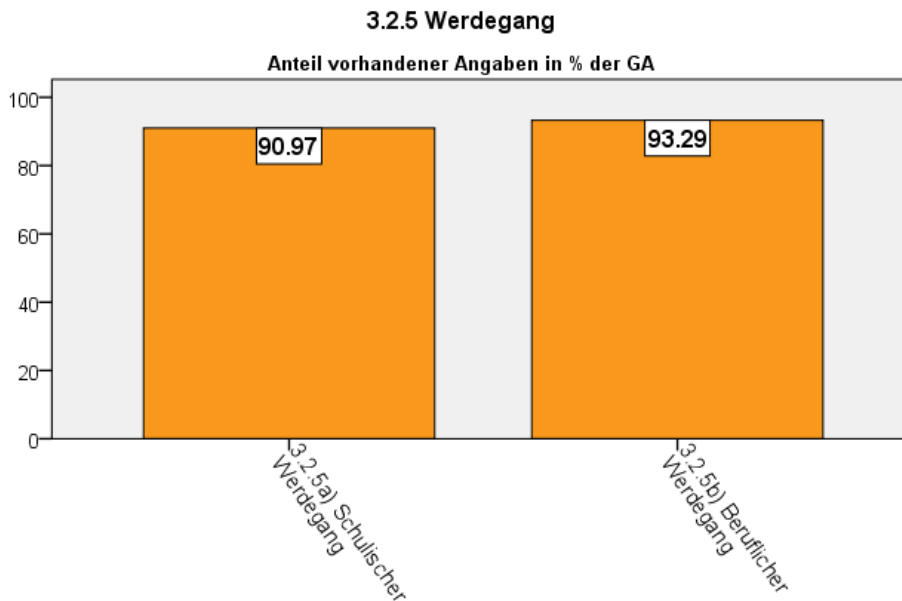
Abbildung 4-12: Soziale Anamnese



Unterabschnitt 3.2.5 „Werdegang“

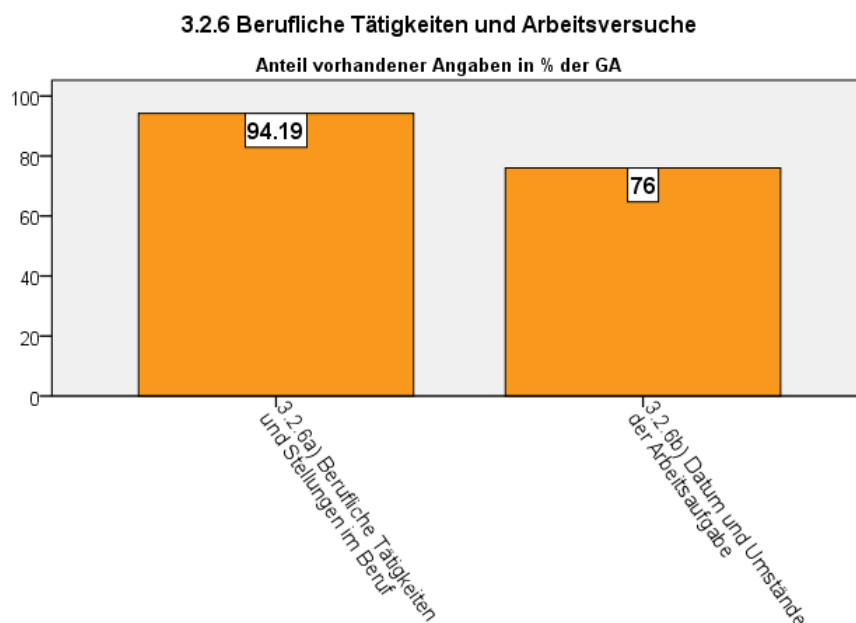
Der schulische Werdegang wurde in knapp 91% der Gutachten dargestellt, der berufliche Werdegang in 93%.

Abbildung 4-13: Werdegang



Unterabschnitt 3.2.6 „Berufliche Tätigkeiten und Arbeitsversuche“

Abbildung 4-14: Berufliche Tätigkeiten/Arbeitsversuche

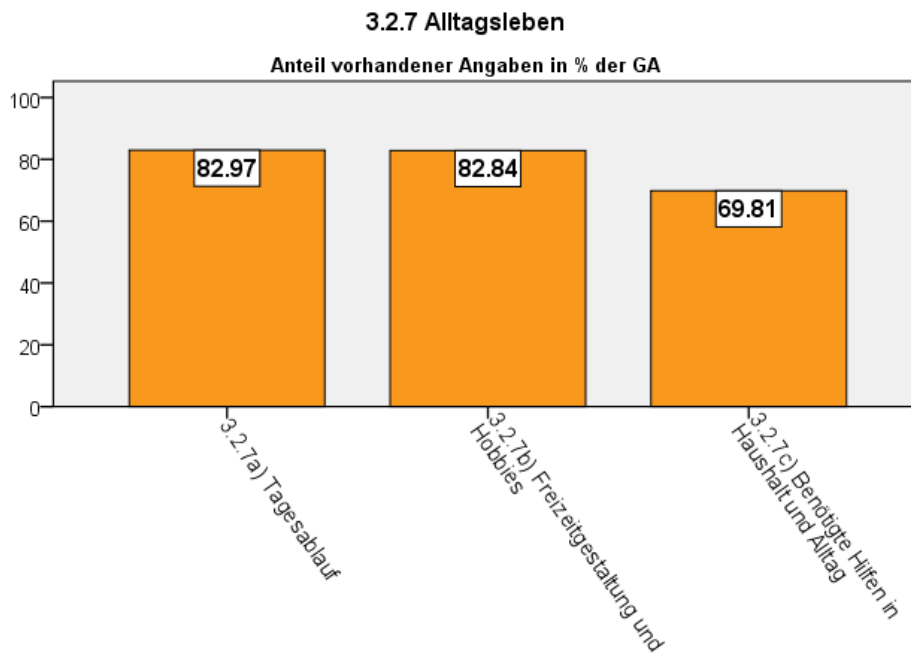


Die meisten Gutachten (94%) enthalten Angaben zu den beruflichen Tätigkeiten der Exploranden, über die Umstände der Arbeitsaufgabe wird in 76% der Gutachten berichtet.

Unterabschnitt 3.2.7 „Alltagsleben“

In etwa vier Fünftel der Gutachten findet sich eine der vorgesehenen Angaben zum Alltagsleben. Dies sind der Tagesablauf (83%), die Freizeitgestaltung (83%) sowie evtl. benötigte Hilfen im Alltag (70%), wobei die letzte Angabe sicherlich in einigen Fällen aus sachlichen Gründen bei der Exploration gar nicht thematisiert werden musste.

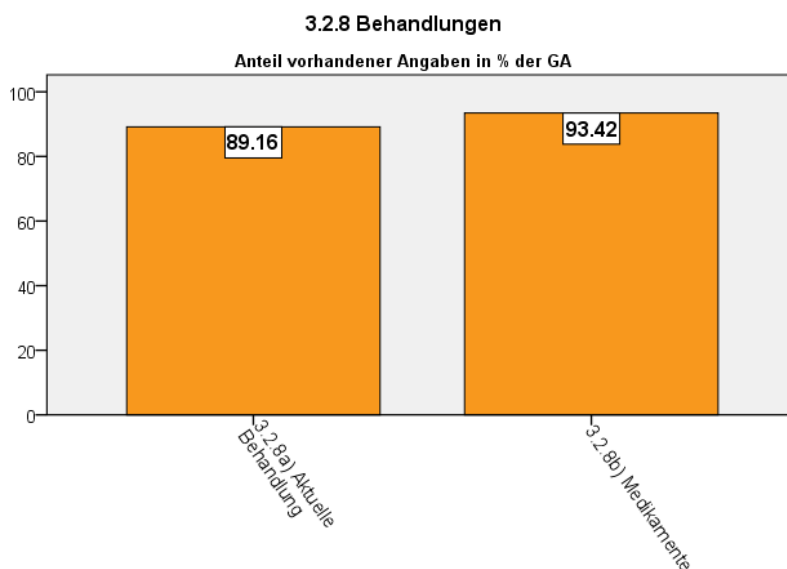
Abbildung 4-15: Alltagsleben



Unterabschnitt 3.2.8 „Behandlungen“

Eine Darstellung der aktuellen Behandlung war in 89% der untersuchten Gutachten enthalten; die medikamentöse Behandlung wurde jedoch nur in 93% der Gutachten dargestellt. Die Behandlungsanamnese wird in der Checkliste nicht eigens erfasst, spiegelt sich jedoch in der Auflistung der vom Auftraggeber zur Verfügung gestellten Akten dossiers (Kriterium 1.3.1) sowie im fakultativen Kriterium 1.3.2 a) „Summarische Auflistung der vom Gutachter zusätzlich beigebrachten (oder anderweitig bei ihm eingegangenen) Aktenstücke respektive telefonischer Nachfragen z. B. beim Hausarzt“ wieder.

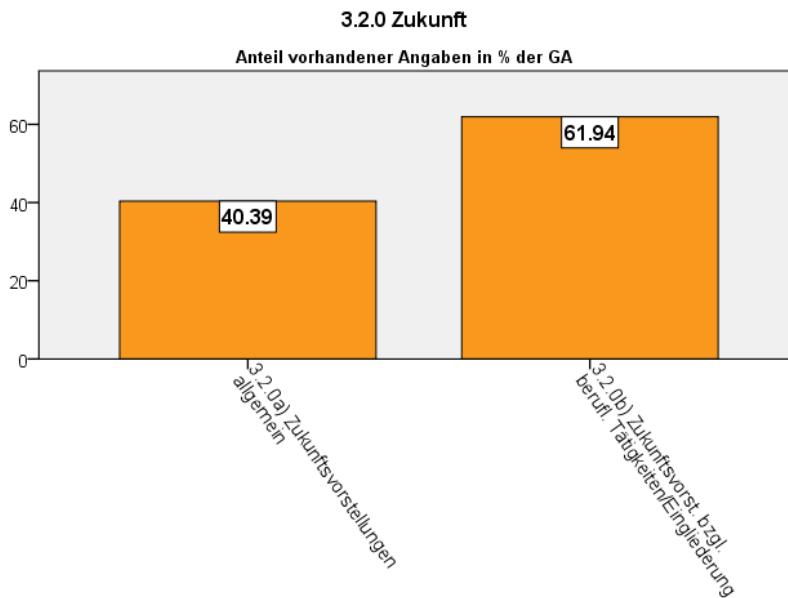
Abbildung 4-16: Behandlungen



Unterabschnitt 3.2.0 „Zukunft“

Allgemeine Zukunftsvorstellungen der Probanden wurden in 40% der Gutachten dargestellt, etwas häufiger (62% aller Gutachten) die Prognosen bezüglich der beruflichen Tätigkeiten bzw. der Eingliederung in das Berufsleben.

Abbildung 4-17: Zukunft

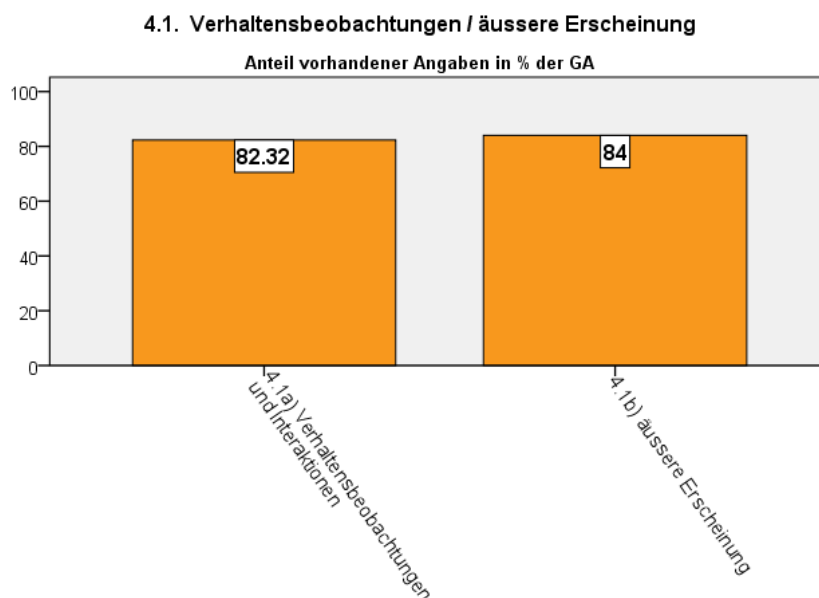


Themenbereich 4. „Befund“

Abschnitt 4.1. „Verhaltensbeobachtungen / äussere Erscheinung“

Verhaltensbeobachtungen wurden in 82% der Gutachten geschildert, das äussere Erscheinungsbild in 84% der Gutachten dargestellt.

Abbildung 4-18: Verhaltensbeobachtungen

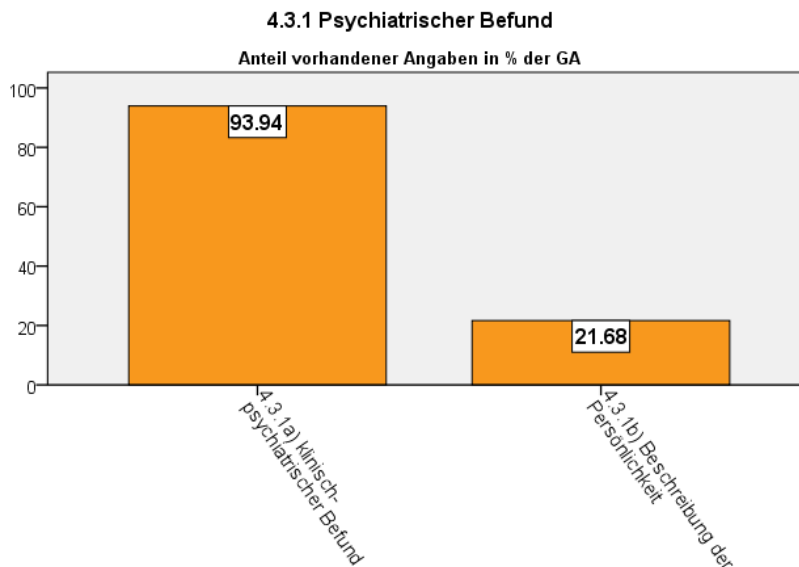


Abschnitt 4.3. „Objektive Untersuchungsbefunde“

Unterabschnitt 4.3.1 „Psychiatrischer Befund“

In 94% der 775 Gutachten wurde der klinisch-psychiatrische Befund dargestellt. Nur ein Fünftel der Gutachten enthält auch eine Beschreibung der Persönlichkeit der Explorandin bzw. des Exploranden.

Abbildung 4-19: Psychiatrischer Befund

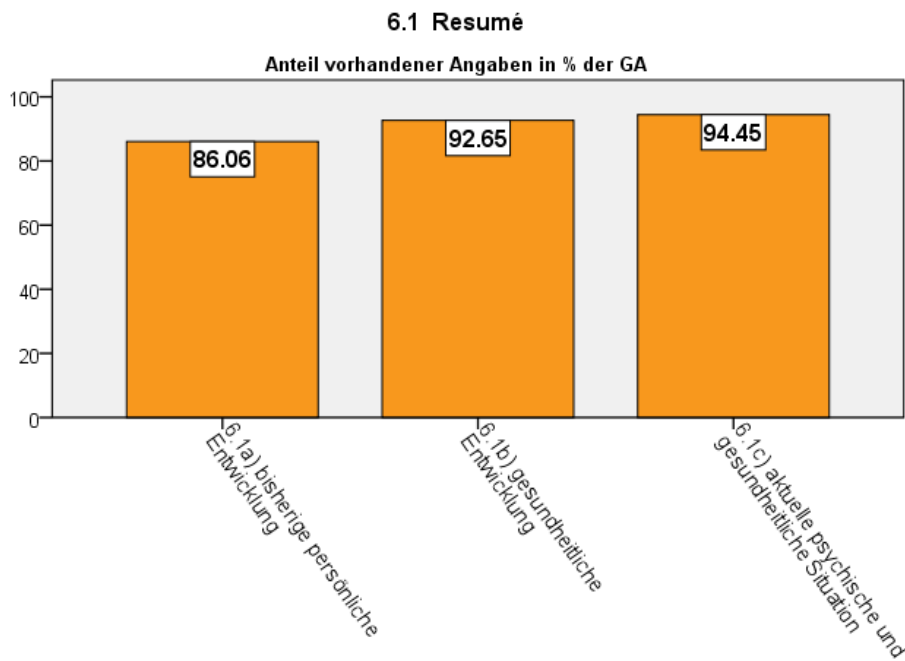


Themenbereich 6. „Medizinische und versicherungsmedizinische Beurteilung“

Abschnitt 6.1 „Resumé“

Die Angaben zum Resumé sind in der folgenden Abbildung dargestellt. Dieses enthält in aller Regel die aktuelle psychische und gesundheitliche Situation und die gesundheitliche Entwicklung, seltener die bisherige persönliche Entwicklung.

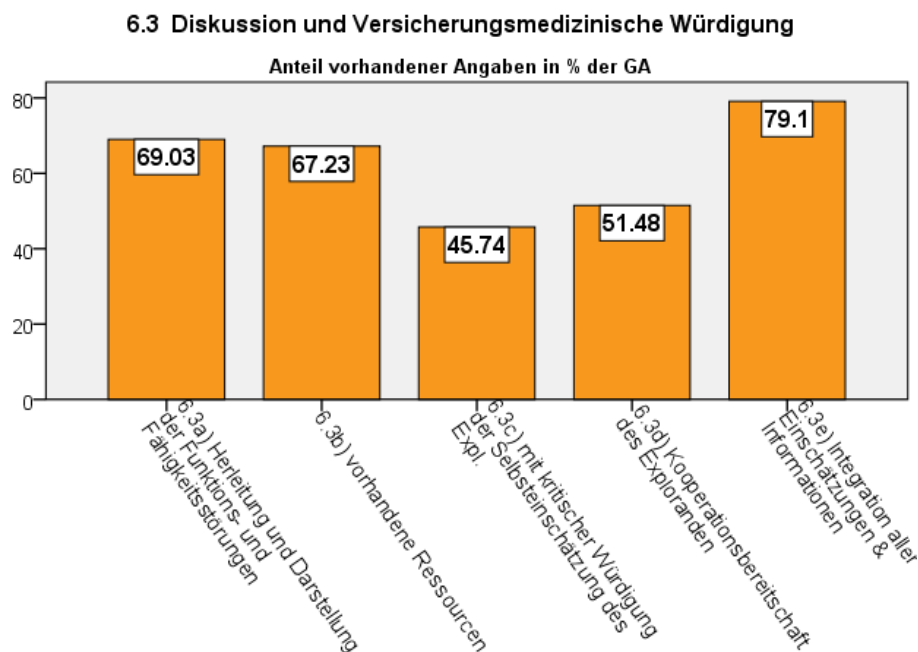
Abbildung 4-20: Resumé



Abschnitt 6.3 „Diskussion und versicherungsmedizinische Würdigung“

Die abschliessende Diskussion und versicherungsmedizinische Würdigung des Falles zeigt ein heterogenes Bild: Die Funktions- und Fähigkeitsstörungen werden in 69% der Gutachten hergeleitet, in 67% werden die vorhandenen Ressourcen des Exploranden dargelegt. Nur in 46% wird die Selbsteinschätzung des Exploranden kritisch gewürdigt und in 51% dessen Kooperationsbereitschaft diskutiert. In 79% der Gutachten wurden die eigenen Einschätzungen und sonstigen vorhandenen Informationen integriert.

Abbildung 4-21: Diskussion/Würdigung



Abschnitt 6.4 „Beantwortung der Fragen“

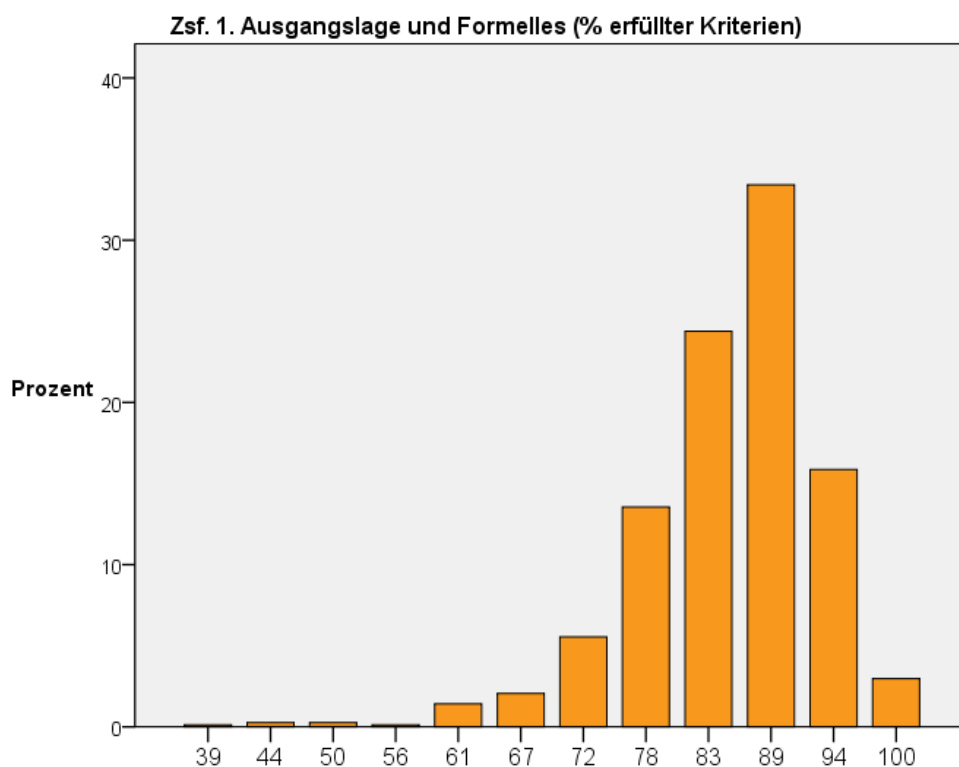
Letzter Punkt der Checkliste für die formale Bewertung der Gutachten ist die Beantwortung der Fragen, wegen derer das Gutachten in Auftrag gegeben worden war. Die Fragen oder zumindest ein Teil davon wurden in 702 der 775 Gutachten beantwortet, was einem Anteil von 91% aller Gutachten entspricht.

4.3 Erfüllung der obligatorischen Kriterien auf Ebene der Bereiche

Im Folgenden wird untersucht, inwiefern die obligatorischen Kriterien in den thematischen Bereichen der Gutachten vollständig behandelt wurden bzw. welcher Anteil an Kriterien noch gefehlt hat. Die Bereiche bestehen aus unterschiedlich vielen Kriterien; um dies auszugleichen, wurde die Anzahl erfüllter Kriterien auf die Gesamtzahl der Kriterien in jedem Bereich bezogen und daraus ein Prozentwert erfüllter Kriterien errechnet. Dieser Anteil erfüllter Kriterien ist in den folgenden Abbildungen jeweils auf der Abszisse abgetragen, die Häufigkeit, mit der jeder Wert bei den 80 Gutachten auftrat, wird auf der Ordinate wiedergegeben.

Im Themenbereich „Ausgangslage und Formelles“ wurden, wie Abbildung 4-22 zeigt, nur in seltenen Fällen weniger als 78% der dort angesiedelten insgesamt 18 Kriterien (z.B. Angaben zum Auftraggeber oder zum Exploranden) bearbeitet. Bei knapp einem Viertel der Gutachten wurden 83% der Kriterien bearbeitet, bei einem Drittel 89%. 94% der Kriterien, d.h. 17 der 18 Kriterien, wurden von 123 Gutachten oder 16% der Gutachten erfüllt. Alle 18 Kriterien (100%) wurden in 23 Gutachten dargestellt.

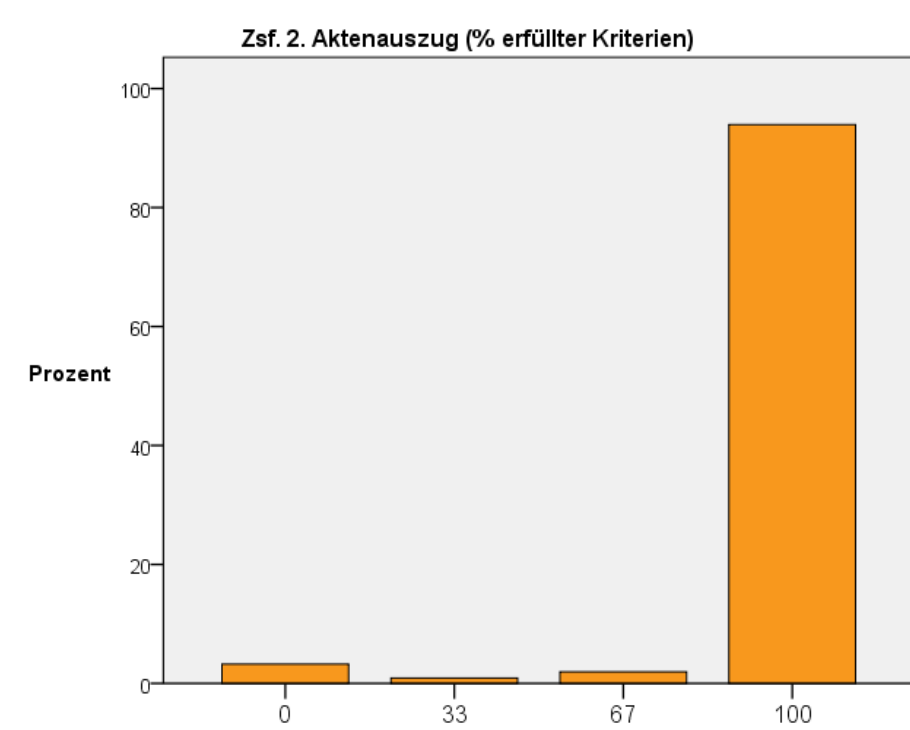
Abbildung 4-22: Bereich Ausgangslage in der Gesamtstichprobe



Im zweiten Bereich, „Aktenauszug“, stellt sich die Verteilung der erfüllten Kriterien wie folgt dar. Hier sind nur drei Kriterien vorgegeben: 2 a) Angabe des Autors, 2 b) Datum des Berichts und 2 c) (relevante) Diagnosen / Befunde.

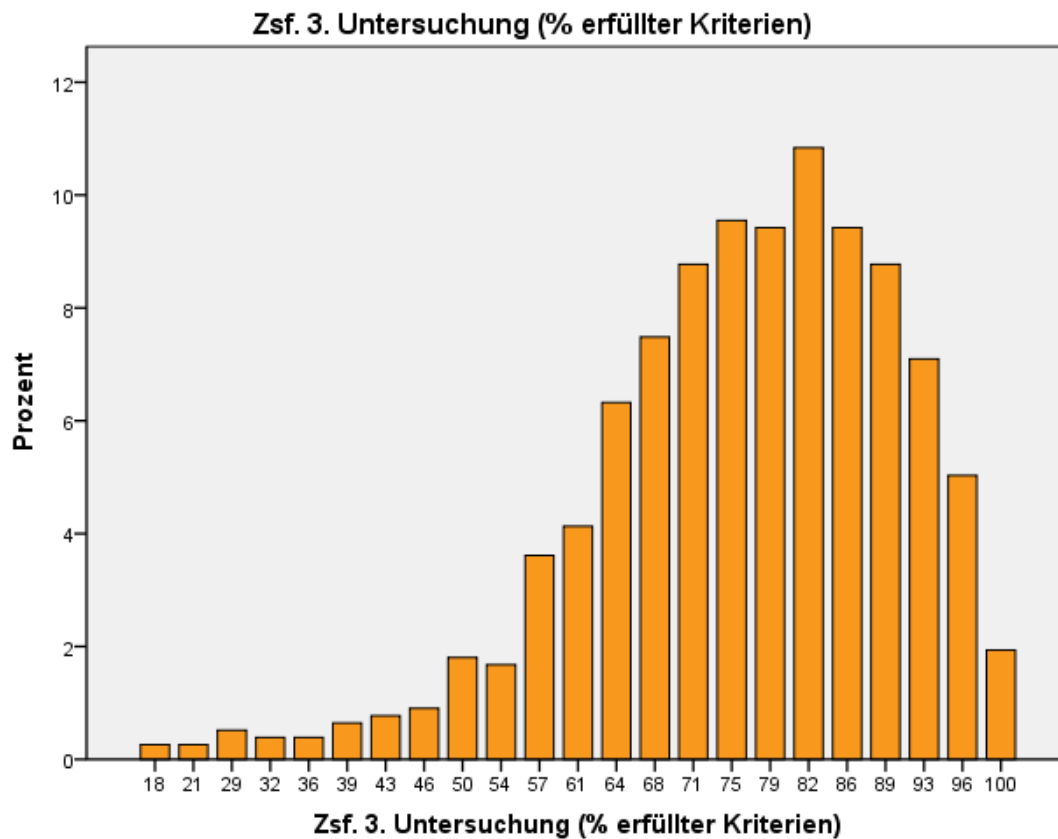
25 Gutachten erwähnten keines der Kriterien, sieben Gutachten eines von drei Kriterien, 15 Gutachten erwähnten zwei von den drei Kriterien, und 94% der Gutachten erwähnten alle Kriterien.

Abbildung 4-23: Bereich Aktenauszug in der Gesamtstichprobe



Der Themenbereich „Untersuchung“ ist mit 28 obligatorischen Kriterien relativ fein differenziert. Wie Abbildung 4-24 verdeutlicht, erfüllten die meisten Gutachten mehr als die Hälfte der Kriterien. Nur 30% der Gutachten erfüllten weniger als zwei Drittel der Kriterien, drei Viertel aller Gutachten nahmen zu mindestens 24 der 28 Kriterien Stellung. 39 Gutachten bearbeiteten 27 der 28 Kriterien, 15 Gutachten enthielten alle „verlangten“ Angaben.

Abbildung 4-24: Bereich Untersuchung in der Gesamtstichprobe

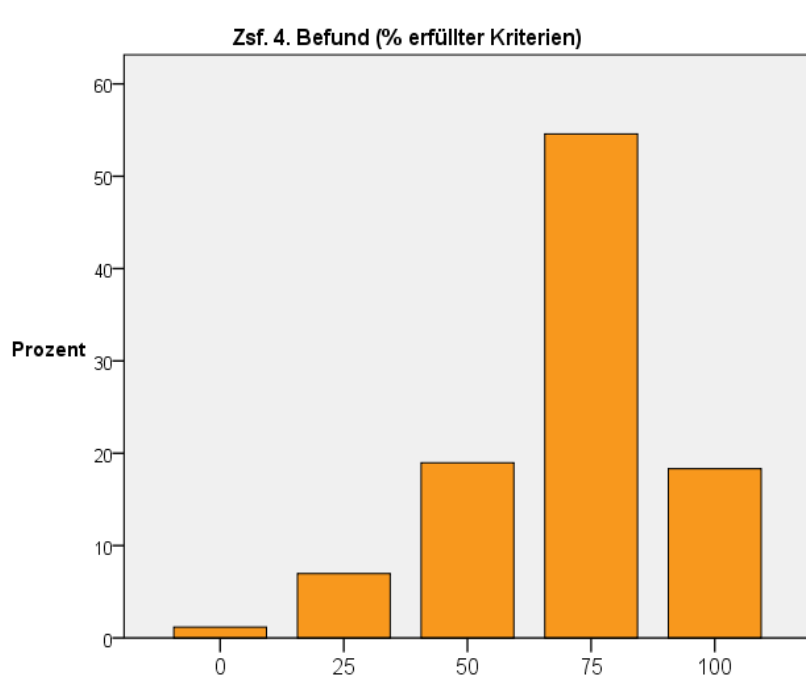


Im Bereich „Befund“ sind vier Kriterien obligatorisch:

- Verhaltensbeobachtungen / äussere Erscheinung:
 - a) Verhaltensbeobachtungen und Interaktionen
 - b) äussere Erscheinung
- Psychiatrischer Befund:
 - a) klinisch-psychiatrischer Befund
 - b) Beschreibung der Persönlichkeit

Neun Gutachten nahmen zu keinem der Kriterien Stellung, 54 nur zu einem und ein Fünftel der Gutachten zu zwei der Kriterien. Über 50% der Gutachten enthielten Angaben zu drei der Kriterien, 18.3% aller Gutachten bearbeiteten alle vier Kriterien (Abbildung 4-25).

Abbildung 4-25: Bereich Befund in der Gesamtstichprobe



Der Bereich der Medizinischen und versicherungsmedizinischen Beurteilung enthält neun Kriterien zu den Themen „Resumé“, „Diskussion und Versicherungsmedizinische Würdigung“ sowie „Beantwortung der Fragen“.

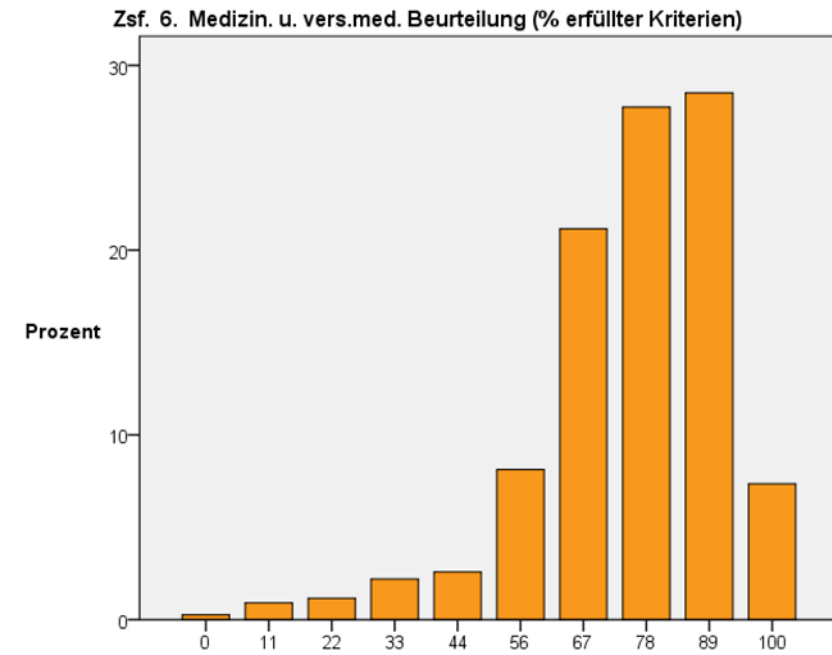
Kriterium 6.1 „Resumé“ gliedert sich in die Unterpunkte a) bisherige persönliche Entwicklung, b) gesundheitliche Entwicklung und c) aktuelle psychische und gesundheitliche Situation.

Kriterium 6.2 „Beurteilung bzw. Einordnung der Beschwerden und Befunde in Bezug auf aktuelle Diagnosesysteme und in Bezug auf frühere, ggf. anderslautende Beurteilungen“ ist optional und nur bei polydisziplinären Gutachten anzuwenden.

Kriterium 6.3 ist die „Diskussion und Versicherungsmedizinische Würdigung“. Es umfasst die a) Herleitung und Darstellung der Funktions- und Fähigkeitsstörungen, b) vorhandene Ressourcen, c) die kritische Würdigung der Selbsteinschätzung des Exploranden, d) die Kooperationsbereitschaft des Exploranden und e) die Integration aller vorhandenen früheren fachlichen Einschätzungen sowie anderer aktenkundiger Informationen. Eine explizite, ausführliche Bewertung der Arbeitsfähigkeit war an dieser Stelle nicht verlangt. Diese ist gleichwohl ja in aller Regel eines der Grundanliegen jedes IV-Gutachtens und war meist auch als Frage an das Gutachten formuliert. Somit treten die Einschätzungen zur Arbeitsfähigkeit meist unter Kriterium 6.4 „Beantwortung der Fragen“ auf mit Angaben zur Arbeitsfähigkeit in Stunden pro Tag, pro Woche oder in %, teilweise zusätzlich differenziert nach der bisherigen Tätigkeit und einer adaptierten Tätigkeit, teilweise auch mit den Auswirkungen der Störungen auf die Arbeitsfähigkeit und mit einer Prognose dazu.

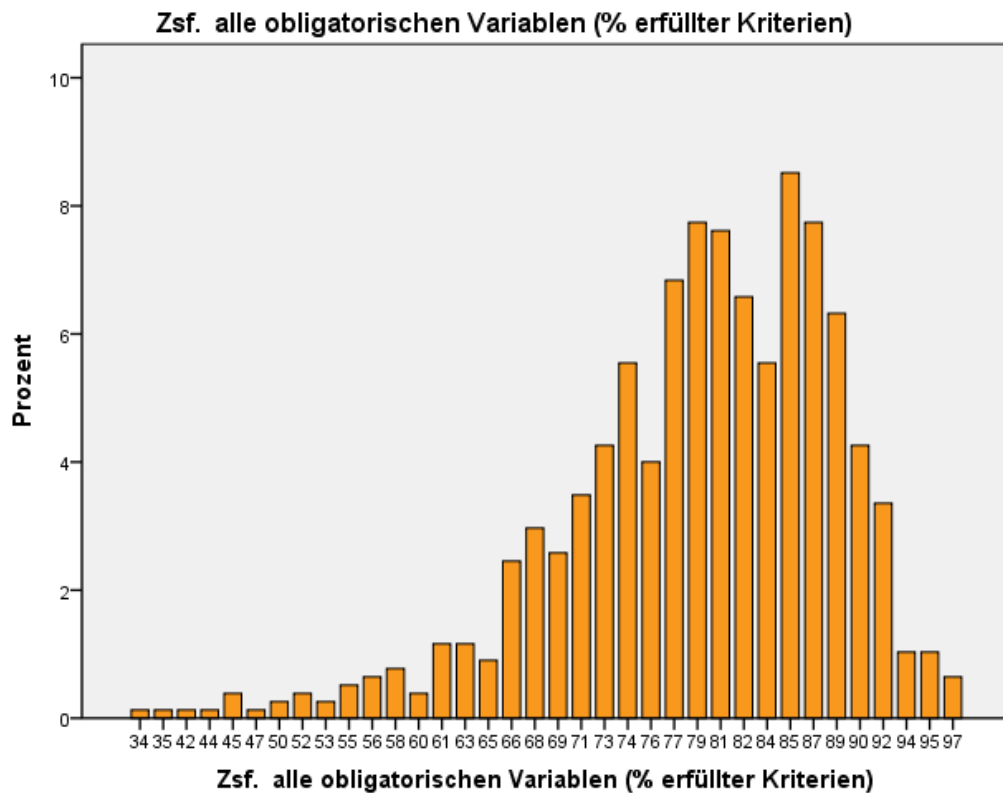
Ohne zumindest eine numerische Angabe zur Arbeitsfähigkeit blieben nur 20 der 775 ausgewerteten Gutachten (2.6%). Weiter gehende Aussagen zur qualitativen Bewertung der Arbeitsfähigkeit und ihrer Darstellung in den Gutachten können mit den vorliegenden Daten zur formalen Qualität nicht gemacht werden.

Abbildung 4-26 zeigt, dass nur ein kleiner Teil der Gutachten weniger als zwei Drittel der Kriterien im Bereich der Medizinischen und versicherungsmedizinischen Beurteilung erfüllt. Im Bereich zwischen einem Erfüllungsgrad von 66 bis 90% liegen 600 der 775 Gutachten, also 77% der Gutachten. Immerhin 57 Gutachten oder 7.4% thematisierten alle neun, also 100% der Kriterien.

Abbildung 4-26: Bereich Beurteilung in der Gesamtstichprobe

Die letzte Abbildung fasst den Erfüllungsgrad über alle obligaten Kriterien zusammen. Hierfür wurde für jedes Gutachten berechnet, zu wie vielen der 62 obligaten Kriterien es Stellung nimmt. Es zeigt sich eine sehr breite Verteilung des Anteils an erfüllten Kriterien: Es sind Gutachten vorhanden, welche nur die Hälfte aller Kriterien erwähnen. Der breite mittlere Bereich von Gutachten enthält Angaben zu zwei Dritteln der Kriterien bis zu 90% der Kriterien. Nur wenige Gutachten nehmen Stellung zu nahezu allen in der Checkliste als obligatorisch vorgeschlagenen Kriterien.

Abbildung 4-27: Zusammenfassung aller obligatorischen Kriterien in der Gesamtstichprobe



4.4 Qualitätsunterschiede zwischen den einzelnen Gutachtenteilen

Die folgenden Tabellen geben eine Übersicht dazu, wie häufig die als obligatorische Einzelkriterien festgelegten Themen in den analysierten Gutachten Erwähnung fanden.

Im Mittel (Median) über alle Einzelkriterien hinweg wurden die Kriterien in 87.3% aller Gutachten erfüllt.

Tabelle 4-4: Erfüllung der obligatorischen Kriterien: Ausgangslage

Kriterium	% erfüllt
1. Ausgangslage und Formelles	
1.1 Abwicklung des Gutachtensauftrages/ Formelles	
1.1.1 Angaben zum Auftraggeber	98.6
1.1.2 a) Name ExplorandIn	100.0
1.1.2 b) Vorname ExplorandIn	100.0
1.1.2c) Geburtsdatum ExplorandIn	95.7
1.1.2d) Adresse ExplorandIn	81.9
1.1.2e) Versicherungsnummer	99.6
1.1.3a) Adresse GutachterIn	95.9
1.1.3b) Facharzttitel GutachterIn	96.6
1.1.4a) Auftragsdatum	86.3
1.1.4b) Eingangsdatum	65.4
1.1.4c) Explorationsdatum	95.6
1.1.4d) Datum der GA-Fertigstellung	99.6
1.2 Anlass der Begutachtung	
1.2a) Wer ist Auftraggeber?	98.8
1.2b) Wann wurde GA in Auftrag gegeben?	86.2
1.2c) Zu welchem Zweck wurde GA in Auftrag gegeben?	88.3
1.3 Grundlagen für das Gutachten	
1.3.1 Auflistung der zur Verfügung gestellten Aktendossiers	95.6
1.3.3a) Eigene Untersuchungen: Untersuchungsdauer	25.2
1.3.3b) Eigene Untersuchungen: Weitere Angaben	23.1

Tabelle 4-5: Erfüllung der obligatorischen Kriterien: Aktenauszug

Kriterium	% erfüllt
2. Aktenauszug	
2.a) Aktenauszug - Angabe des Autors	96.1
2.b) Datum des Berichts	95.7
2.c) (relevante) Diagnosen / Befunde	94.7

Tabelle 4-6: Erfüllung der obligatorischen Kriterien: Untersuchung

Kriterium	% erfüllt
3. Untersuchung	
3.1 Spontane Angaben des Expl. / offenes Interview	36.6
3.2 Vertiefendes Interview - je nach Fragestellung -	
3.2.1a) Aktuelle Beschwerden	97.9
3.2.1b) Entstehung/Verlauf der psychischen Beeinträchtigung	90.8
3.2.1c) Behandlungsmassnahmen	91.1
3.2.1d) Eigene Krankheitskonzepte	54.7
3.2.1e) Umgang m. Beschwerden im Alltag?	70.2
3.2.1f) Zukunftsvorstellungen bzgl. Krankheitsverlauf	33.5
3.2.2a) psychiatrische Anamnese	95.4
3.2.2b) somatische Anamnese	89.8
3.2.2c) Konsum psychotroper Substanzen	72.1
3.2.2d) Sexualität	32.5
3.2.3 Familienanamnese / Heredität	91.4
3.2.4a) Wohnung	70.1
3.2.4b) finanzielle Verhältnisse / Schulden	70.3
3.2.4c) Partnerschaft	94.7
3.2.4d) Kinder	85.2
3.2.4e) Beziehungsnetz	78.1
3.2.5a) Schulischer Werdegang	91.0
3.2.5b) Beruflicher Werdegang	93.3
3.2.6a) Berufliche Tätigkeiten und Stellungen im Beruf	94.2
3.2.6b) Datum und Umstände der Arbeitsaufgabe	76.0
3.2.7a) Tagesablauf	83.0
3.2.7b) Freizeitgestaltung und Hobbies	82.8
3.2.7c) Benötigte Hilfen in Haushalt und Alltag	69.8
3.2.8a) Aktuelle Behandlung	89.2
3.2.8b) Medikamente	93.4
3.2.0a) Zukunftsvorstellungen allgemein	40.4
3.2.0b) Zukunftsvorst. bzgl. berufl. Tätigkeiten/Eingliederung	61.9

Tabelle 4-7: Erfüllung der obligatorischen Kriterien: Befund

Kriterium	% erfüllt
4. Befund	
4.1. Verhaltensbeobachtungen und äussere Erscheinung	
4.1a) Verhaltensbeobachtungen und Interaktionen	82.3
4.1b) äussere Erscheinung	84.0
4.3 Objektive Untersuchungsbefunde	
4.3.1a) klinisch-psychiatrischer Befund	93.9
4.3.1b) Beschreibung der Persönlichkeit	21.7

Tabelle 4-8: Erfüllung der obligatorischen Kriterien: Beurteilung

Kriterium	% erfüllt
6. Medizin. u. versicherungsmediz. Beurteilung	
6.1 Resumé	
6.1a) bisherige persönliche Entwicklung	86.1
6.1b) gesundheitliche Entwicklung	92.6
6.1c) aktuelle psychische und gesundheitliche Situation	94.5
6.3 Diskussion und Versicherungsmedizinische Würdigung	
6.3a) Herleitung und Darstellung der Funktions- und Fähigkeitsstörungen	69.0
6.3b) vorhandene Ressourcen	67.2
6.3c) Kritische Würdigung der Selbsteinschätzung des Expl.	45.7
6.3d) Kooperationsbereitschaft des Exploranden	51.5
6.3e) Integration aller Einschätzungen & Informationen	79.1
6.4 Beantwortung der Fragen	90.6

Kriterium 4.3.1b) „Beschreibung der Persönlichkeit“ war am seltensten erfüllt. Nur 21.7% aller Gutachten machten dazu Angaben. In weniger als 50% der Gutachten fanden die folgenden Kriterien Erwähnung:

- 4.3.1b) Beschreibung der Persönlichkeit
- 1.3.3b) Eigene Untersuchungen: Weitere Angaben (z.B. Zuhilfenahme eines Dolmetschers in welcher Sprache)
- 1.3.3a) Eigene Untersuchungen: Untersuchungsdauer (ggf. je Sitzung bei mehreren Terminen)
- 3.2.2d) Sexualität
- 3.2.1f) Zukunftsvorstellungen bzgl. Krankheitsverlauf
- 3.1 Spontane Angaben des Expl. / offenes Interview
- 3.2.0a) Zukunftsvorstellungen allgemein
- 6.3c) Kritische Würdigung der Selbsteinschätzung des Expl.

Am vollständigsten waren die deskriptiven Angaben zum Exploranden, zum Gutachter und Auftraggeber. Ebenfalls relativ häufig waren Angaben zur aktuellen Krankheitssituation mit den Kriterien 3.2.1a) „Aktuelle Beschwerden“, das von 97.9% der Gutachten erfüllt wurde, 3.2.2a) „psychiatrische Anamnese“ von 95.4% sowie 3.2.4c) „Partnerschaft“ und 2.c) „(relevante) Diagnosen / Befunde“ von je 94.7%.

Die folgende Tabelle 4-9 gibt eine Übersicht zu den thematischen Bereichen, nach denen die Kriterien sich gliedern.

Am vollständigsten sind die Angaben im Bereich „Aktenauszug“. Zwar gibt es einige Gutachten, in denen dieser Punkt überhaupt nicht erwähnt wurde (Minimum: 0% der Kriterien dieses Bereichs erfüllt), bei einer grossen Zahl der Gutachten sind aber alle Kriterien dieses Bereichs erfüllt (Maximum: 100% der Kriterien dieses Bereichs erfüllt). Die Zahl der hier vollständigen Gutachten ist so gross, dass selbst der Median des Prozentsatzes den Maximalwert von 100% Erfüllung annimmt.

Weiterhin relativ vollständig werden die Ausgangslage und die Formalia dargestellt (Median aller Gutachten: 89% der Kriterien erfüllt). Etwas weniger komplett beschrieben werden die Bereiche „Untersuchung“, „Beurteilung“ und „Befund“.

Werden alle 62 obligatorischen Kriterien betrachtet, so wird im Mittel der Gutachten nur ca. zur Hälfte aller Kriterien Stellung genommen.

Tabelle 4-9: Unterschiede zwischen Gutachtenteilen

Themenbereich	Median	Minimum	Maximum
% erfüllter Kriterien			
2. Aktenauszug	100.00	0	100
1. Ausgangslage und Formelles	88.89	39	100
3. Untersuchung	80.00	17	100
6. Medizin. u. vers.med. Beurteilung	77.78	0	100
4. Befund	75.00	0	100
Zsf. alle obligatorischen Variablen	52.00	32	95

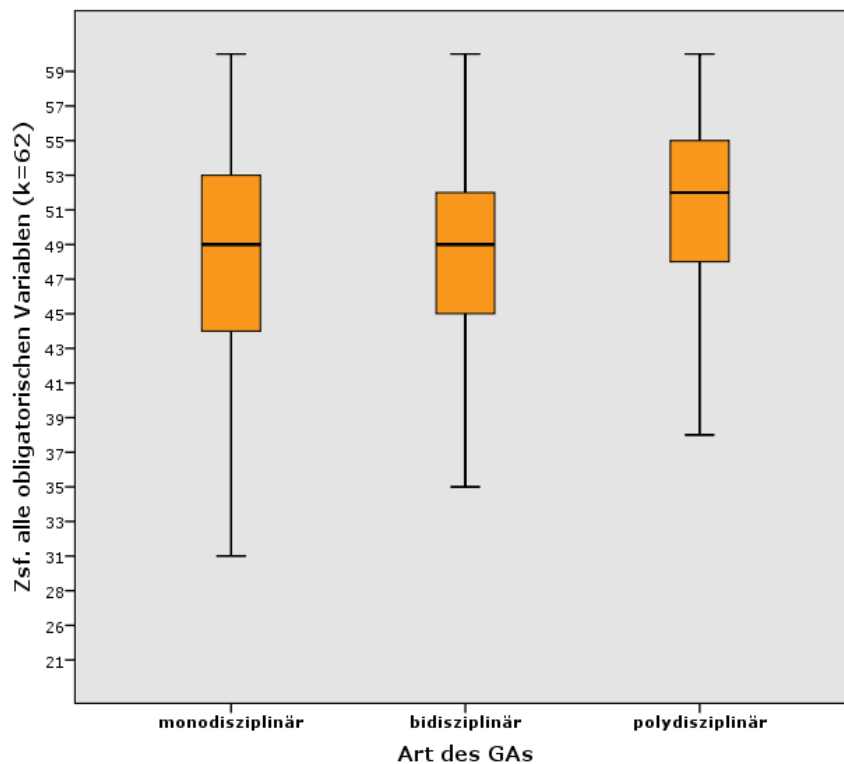
Vollständige Angaben auf allen Kriterien fanden sich

- im Themenbereich 2. Aktenauszug bei 93.9% aller Gutachten,
- im Themenbereich 4. Befund bei 18.3% aller Gutachten,
- im Themenbereich 6. Medizin. u. vers.med. Beurteilung bei 7.4% aller Gutachten,
- im Themenbereich 1. Ausgangslage und Formelles bei 3.0% aller Gutachten,
- im Themenbereich 3. Untersuchung bei keinem Gutachten.

4.5 Unterschiede zwischen Arten von Gutachten

Wie vielleicht auch zu erwarten, unterscheiden sich die Arten von Gutachten. Polydisziplinäre Gutachten sind hinsichtlich der Kriterien aus der Checkliste vollständiger als mono- und bidisziplinäre Gutachten (Abbildung 4-28). Diese Unterschiede sind hoch signifikant (H-Test, $\chi^2=35.968$, $df=2$, $p<.001$). Die Unterschiede gehen v.a. auf bessere Bewertungen der polydisziplinären Gutachten in den Bereichen „Ausgangslage“ und „Untersuchung“ zurück, während die Befunde bei den monodisziplinären Gutachten am vollständigsten geschildert werden (H-Test, $\chi^2=25.755$, $df=2$, $p<.001$). Die Medizinische und versicherungsmedizinische Beurteilung unterscheidet sich nicht zwischen den Gutachtentypen (H-Test, $\chi^2=6.854$, $df=2$, $p=.032$).

Abbildung 4-28: Unterschiede zwischen Arten von Gutachten



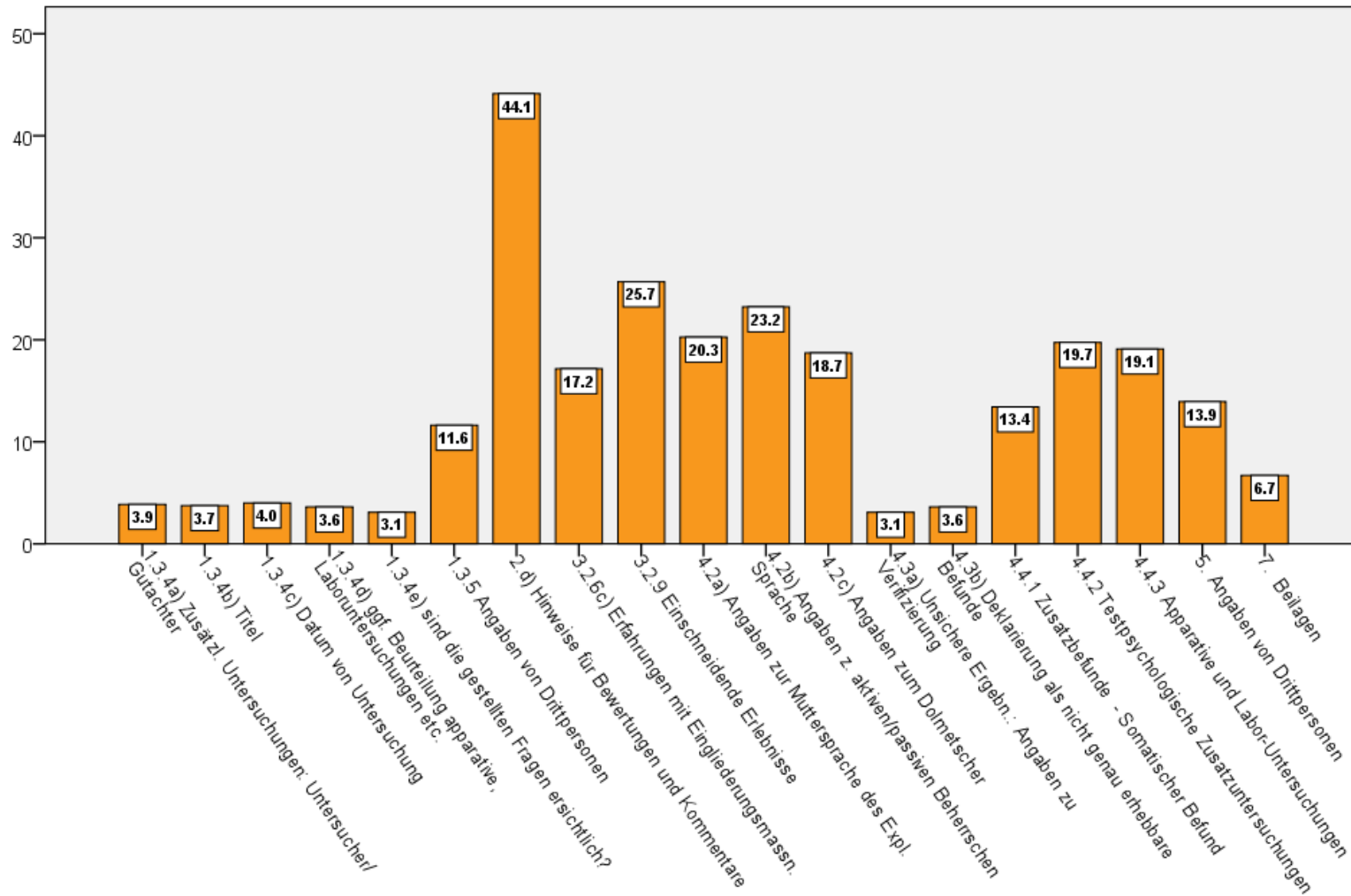
4.6 Fakultative Kriterien

4.6.1 Übersicht über alle fakultativen Kriterien

Wie die Übersicht zu den fakultativen Kriterien auf der nächsten Seite zeigt (Abbildung 4-29 und Tabelle 4-10), sind Angaben zu diesen Punkten nur in der Minderheit der Gutachten zu finden. Allerdings schwankt der Prozentanteil in Abhängigkeit vom jeweiligen Bereich und zwischen einzelnen Kriterien stark.

Abbildung 4-29: Erfüllung der fakultativen Kriterien

Anteil vorhandener Angaben in % der GA



Das am häufigsten erfüllte Zusatzkriterium findet sich im *Themenbereich* „Aktenauszug“, 44.1% der Gutachten wiesen dort Hinweise für Bewertungen und Kommentare auf. Alle anderen optionalen Kriterien wurden in weniger als einem Viertel der Gutachten beantwortet, vermutlich, da sie nicht auf alle Gutachten zuträfen, sondern nur in besonderen Fällen relevant wurden. Dies gilt auch für die anderen fakultativen Kriterien, sodass die Ergebnisse in Abbildung 4-29 weniger als Indiz für die Vollständigkeit der Berichterstattung in einem Gutachten, sondern mehr als Hinweis darauf verstanden werden müssen, wie häufig zusätzliche Aktenstücke benutzt (Kriterium 1.3.2) und zusätzliche Untersuchungen durchgeführt (Kriterium 1.3.4) wurden, in wie vielen Fällen Erfahrungen mit Eingliederungsmassnahmen vorlagen (Kriterium 3.2.6c), einschneidende Erlebnisse zu berichten (Kriterium 3.2.9), die Sprachliche Verständigung ein relevantes Thema (Kriterium 4.2), Untersuchungsergebnisse unsicher (Kriterium 4.3), Zusatzbefunde (Kriterium 4.4) oder Angaben von Drittpersonen vorlagen (Kriterium 5.) oder Beilagen erforderlich waren (Kriterium 7.).

Als einziges fakultatives Kriterium konnte 2.d) „Hinweise für Bewertungen und Kommentare“ aus dem Aktenauszug bei allen Gutachten gleichermaßen gelten und war entsprechend häufig erfüllt, wobei Bewertungen sich in der Darstellung des Aktenauszugs optimaler Weise nicht finden sollten.

Im *Themenbereich* „Untersuchung“ enthielten 17.2% der Gutachten Angaben zu Eingliederungsmassnahmen, 25.7 % der Gutachten berichteten Einschneidende Ereignisse der Exploranden.

Der *Themenbereich* „Befund“ zeigt bezüglich der Beantwortung der Zusatzkriterien ein heterogenes Bild. 23% der Gutachten machen Angaben zur aktiven / passiven Beherrschung der Sprache, jeweils rund 20% zur Muttersprache des Exploranden, zu Testpsychologischen Zusatzuntersuchungen und zu Apparativen und Laboruntersuchungen. Angaben zum Dolmetscher enthalten 18.7%, 13.4% berichten Somatische Zusatzbefunde. Lediglich 3.1% bzw. 3.6% der Gutachten enthalten Hinweise auf Unsichere Ergebnisse / Angaben zur Verifizierung oder nicht genau behebbar Befunde.

Wenig fakultative Angaben finden sich im *Themenbereich* „Ausgangslage und Formelles“. Während jeweils 11.6% der Gutachten eine summarische Auflistung zusätzlich beigebrachter Aktenstücke enthält und Angaben zu Drittpersonen aufweisen, liegt der Prozentsatz der Gutachten, die die anderen fakultativen Kriterien dieses Bereichs erfüllt bei unter 5.5%. Allerdings beziehen sich die Kriterien 1.3.4 a) – e) nur auf zusätzliche Untersuchungen, zusätzlich angeforderte Gutachten oder Berichte, welche über die vom Auftraggeber zur Verfügung gestellten Aktenstücke und die eigenen Untersuchungen hinaus noch vom Gutachter verwendet wurden, was nur in einem kleineren Teil der Fälle aufgetreten sein dürfte.

Tabelle 4-10: Erfüllung der fakultativen Kriterien

Kriterium erfüllt	nein		ja	
	Anzahl	% der GA	Anzahl	% der GA
1. Ausgangslage und Formelles				
1.3.2a) Summarische Auflistung von zusätzlich beigebrachten Aktenstücke/Nachfragen	685	88.4%	90	11.6%
1.3.2b) Aktenstücke mit dem GA dem Auftrag. in Kopie zugestellt?	733	94.6%	42	5.4%
1.3.4a) Zusätzl. Untersuchungen: Untersucher/ Gutachter	745	96.1%	30	3.9%
1.3.4b) Zusätzl. Untersuchungen: Titel	746	96.3%	29	3.7%
1.3.4c) Zusätzl. Untersuchungen: Datum von Untersuchung	744	96.0%	31	4.0%
1.3.4d) Zusätzl. Untersuchungen: ggf. Beurteilung apparative, Laboruntersuchungen etc.	747	96.4%	28	3.6%
1.3.4e) Zusätzl. Untersuchungen: sind die gestellten Fragen ersichtlich?	751	96.9%	24	3.1%
1.3.5 Angaben von Drittpersonen	684	88.4%	90	11.6%
2. Aktenauszug				
2.d) Hinweise für Bewertungen und Kommentare	433	55.9%	342	44.1%
3. Untersuchung				
3.2.6c) Erfahrungen mit Eingliederungsmassnahmen	642	82.8%	133	17.2%
3.2.9 Einschneidende Erlebnisse	576	74.3%	199	25.7%
4. Befund				
4.2a) Angaben zur Muttersprache des Expl.	618	79.7%	157	20.3%
4.2b) Angaben z. aktiven/passiven Beherrschen Sprache	595	76.8%	180	23.2%
4.2c) Angaben zum Dolmetscher	630	81.3%	145	18.7%
4.3a) Unsichere Ergebn.: Angaben zu Verifizierung	751	96.9%	24	3.1%
4.3b) Deklaration als nicht genau erhebbare Befunde	747	96.4%	28	3.6%
4.4.1 Zusatzbefunde - Somatischer Befund	671	86.6%	104	13.4%
4.4.2 Testpsychologische Zusatzuntersuchungen	622	80.3%	153	19.7%
4.4.3 Apparative und Labor-Untersuchungen	627	80.9%	148	19.1%
5. Angaben von Drittpersonen	667	86.1%	108	13.9%
7. Beilagen	723	93.3%	52	6.7%

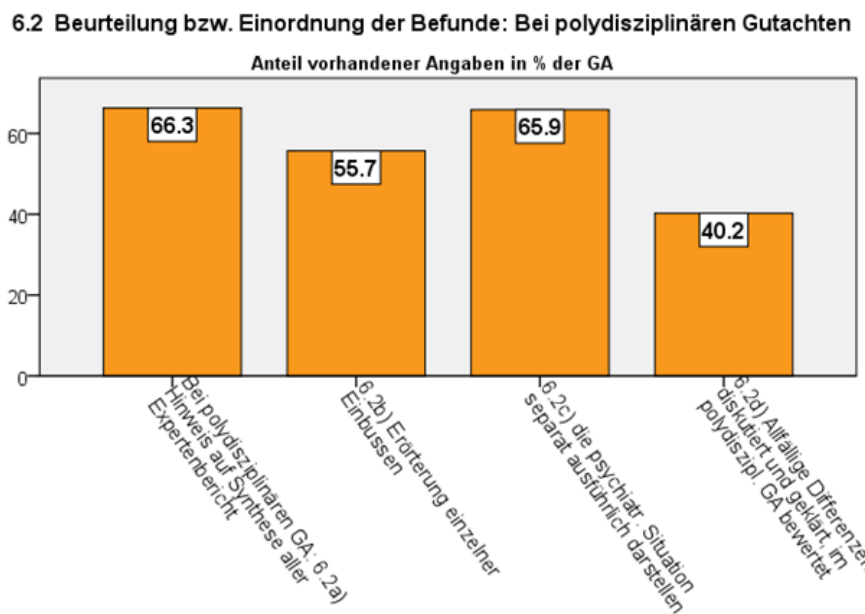
4.6.1 Beurteilung bzw. Einordnung der Befunde bei polydisziplinären Gutachten

Polydisziplinäre Gutachten bestehen in der Regel aus mehreren Teilgutachten, für die optimalerweise eine zusammenfassende Einordnung der Befunde vorliegt. Die Zusatzkriterien im Bereich „Medi-

“*zinische und versicherungsmedizinische Beurteilung*“ sind deshalb ausschliesslich für polydisziplinäre Gutachten relevant und werden getrennt von den anderen fakultativen Kriterien in Abbildung 4-30 und Tabelle 4-11 dargestellt.

Einen Hinweis auf eine Synthese aller Expertenbericht und eine Erörterung einzelner Einbussen enthalten jeweils etwa 66% der polydisziplinären Gutachten. In 65.9% der polydisziplinären Gutachten wird die psychiatrische Situation separat ausführlich dargestellt und 40.2% diskutieren oder bewerten allfällige Differenzen zwischen den Teilgutachten.

Abbildung 4-30: Fakultative Kriterien bei polydisziplinären Gutachten



n=246 polydisziplinäre Gutachten

Tabelle 4-11: Erfüllung der fakultativen Kriterien für polydisziplinäre Gutachten

Kriterium erfüllt	nein		Ja	
	Anzahl	% der GA	Anzahl	% der GA
6.2a) Hinweis auf Synthese aller Expertenberichte	83	33.7%	163	66.3%
6.2b) Erörterung einzelner Einbussen	109	44.3%	137	55.7%
6.2c) die psychiatr. Situation separat ausführlich darstellen	84	34.1%	162	65.9%
6.2d) Allfällige Differenzen diskutiert und geklärt, im polydisziplin. GA bewertet	147	59.8%	99	40.2%

n=246 polydisziplinäre Gutachten

5 Qualitätsunterschiede zwischen den Gutachten verschiedener Institutionen

Bezüglich der begutachtenden Institutionen wurden folgende Kategorien unterschieden:

- a) Frei praktizierende Psychiater und Psychiaterinnen,
- b) Medizinische Abklärungsstellen (MEDAS), die vorwiegend polydisziplinäre Gutachten erstellen,
- c) Regionalärztliche Dienste (RAD) der IV, die neben der Vergabe von Gutachtaufträgen an andere Stellen auch eigene Gutachten erstellen, und
- d) sonstige Institutionen, d.h. Psychiater/innen, die für andere Versicherer (z.B. SUVA oder KTG) oder für eine unabhängige Gutachterstelle arbeiten, einschliesslich Spitälern und Unikliniken.

Die Gutachten-erstellende Institution wurde in der Checkliste entlang dieser Kriterien erfasst, so dass die formale Qualität der Gutachten auf Unterschiede zwischen den ausführenden Institutionen geprüft werden können. Diese Vergleiche werden auf der aggregierten Ebene der obligatorischen Themenbereiche 1 bis 4 und 6 vorgenommen. Auf eine weitere Differenzierung der Kategorie „Sonstige“ wurde verzichtet, da aufgrund der vorliegenden Fallzahlen eine sinnvolle Detailauswertung nicht möglich ist.

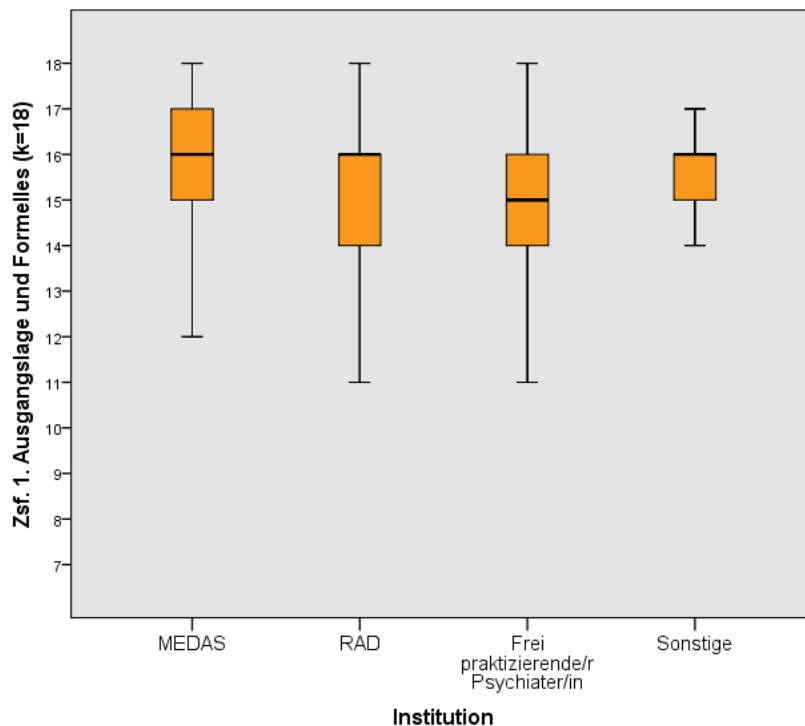
5.1 Unterschiede in einzelnen Themenbereichen

Themenbereich 1. „Ausgangslage und Formelles“

Im Bereich „Ausgangslage und Formelles“ liegen die Mediane von MEDAS, RAD und Sonstigen auf dem gleichen Niveau, auch wenn die Verteilungen sich ansonsten leicht unterscheiden. Die mittlere Zahl berücksichtigter Kriterien liegt bei den frei praktizierenden Psychiatern um eine Stufe unter den anderen Gruppen. Diese Unterschiede sind im H-Test hoch signifikant ($\chi^2=67.438$, $df=3$, $p<.001$), was allerdings angesichts der grossen Fallzahlen nicht sehr verwunderlich ist.

Eine Aussage über das Ausmass der Unterschiede machen die Rangsummen aus dem H-Test. Sie weisen aus, dass die MEDAS-Gutachten in diesem Bereich am besten abschneiden, gefolgt von denen von sonstigen Einrichtungen, dann von den RAD und schliesslich von den frei Praktizierenden.

Abbildung 5-1: Bereich Ausgangslage bei verschiedenen Institutionen



Themenbereich 2. „Aktenauszug“

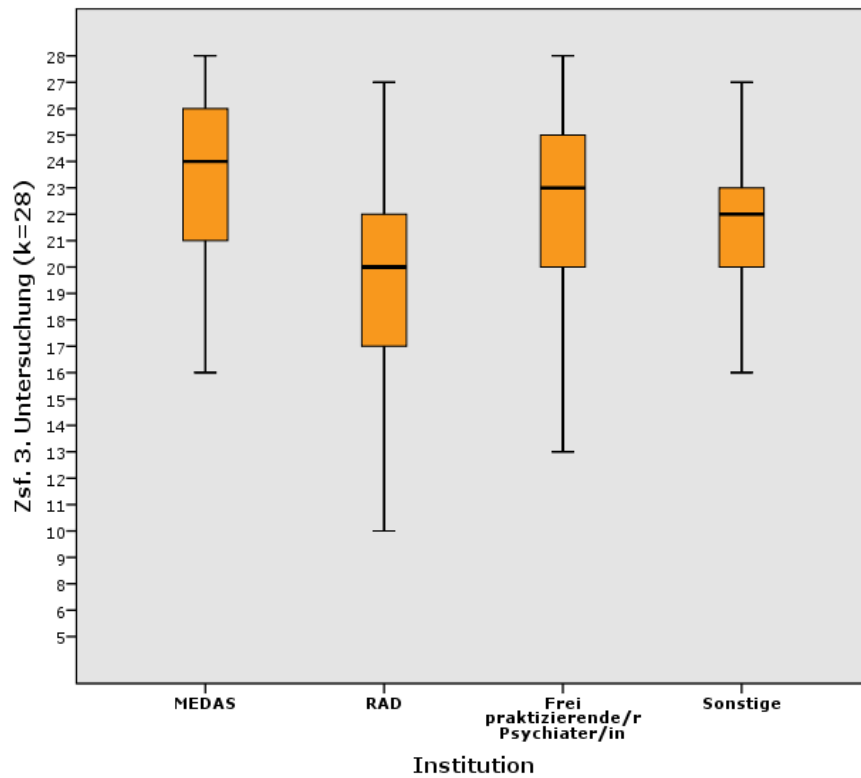
Im Bereich „Aktenauszug“ waren die drei obligaten Kriterien in allen Gruppen bei fast allen Gutachten nahezu perfekt erfüllt, weshalb hier in der Abbildung keine unterschiedlichen Verteilungen gezeigt werden könnten. Nichtsdestoweniger war selbst hier der Kruskal-Wallis-Test hoch signifikant ($\chi^2=20.327$, $df=3$, $p<.001$).

Die Detailanalyse zeigt, dass die Gutachten von MEDAS und Sonstigen etwa gleichauf an der Spitze liegen, dicht gefolgt von den frei praktizierenden PsychiaterInnen und dann den RAD-Gutachten.

Themenbereich 3. „Untersuchung“

Der Themenbereich „Untersuchung“ zeigt deutlichere Unterschiede zwischen den Institutionen. Hier liegen die Gutachten von RAD bei deutlich geringeren Erfüllungswerten als die anderen. Die besten Werte erreichen die Gutachten von MEDAS, gefolgt von denen von frei praktizierenden PsychiaterInnen und denen von sonstigen Einrichtungen. Diese Unterschiede sind hoch signifikant ($\chi^2=138.460$, $df=3$, $p<.001$).

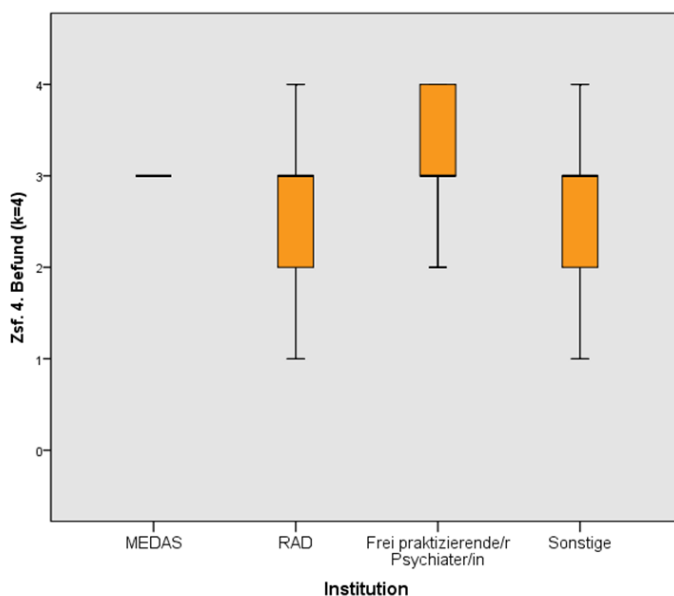
Abbildung 5-2: Bereich Untersuchung bei verschiedenen Institutionen



Themenbereich 4. „Befund“

Im Bereich „Befund“ liegen wieder die Mediane der vier Gruppen alle auf demselben Niveau, wobei allerdings auch nur fünf unterschiedliche Werte möglich wären, da in diesem Bereich nur vier Kriterien vorliegen.

Abbildung 5-3: Bereich Befund bei verschiedenen Institutionen

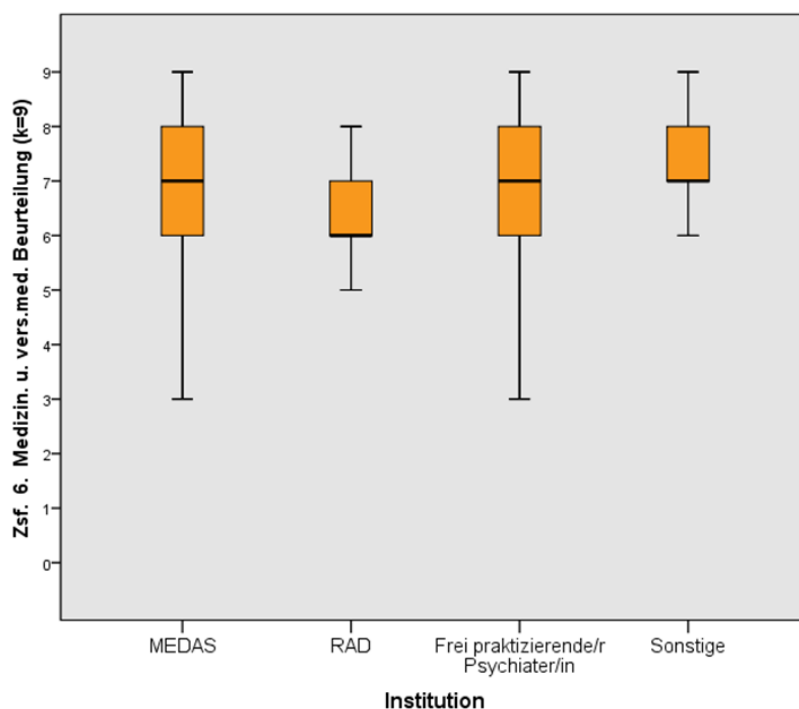


Wie nicht anders zu erwarten sind auch die Unterschiede im Themenbereich „Befund“ hoch signifikant (H-Test, $\chi^2=60.340$, $df=3$, $p<.001$). Am besten sind die mittleren Rangwerte für die frei Praktizierenden, gefolgt von MEDAS, dann den Sonstigen und am Ende die RAD.

Themenberich 6. „Medizinische und versicherungsmedizinische Beurteilung“

Auch im Bereich der „Medizinischen und versicherungsmedizinischen Beurteilung“ liegen die Niveaus der vier Gruppen wieder sehr ähnlich, im Median sind die Gutachten der RAD um ein Kriterium schlechter als die anderen.

Abbildung 5-4: Beurteilung bei verschiedenen Institutionen



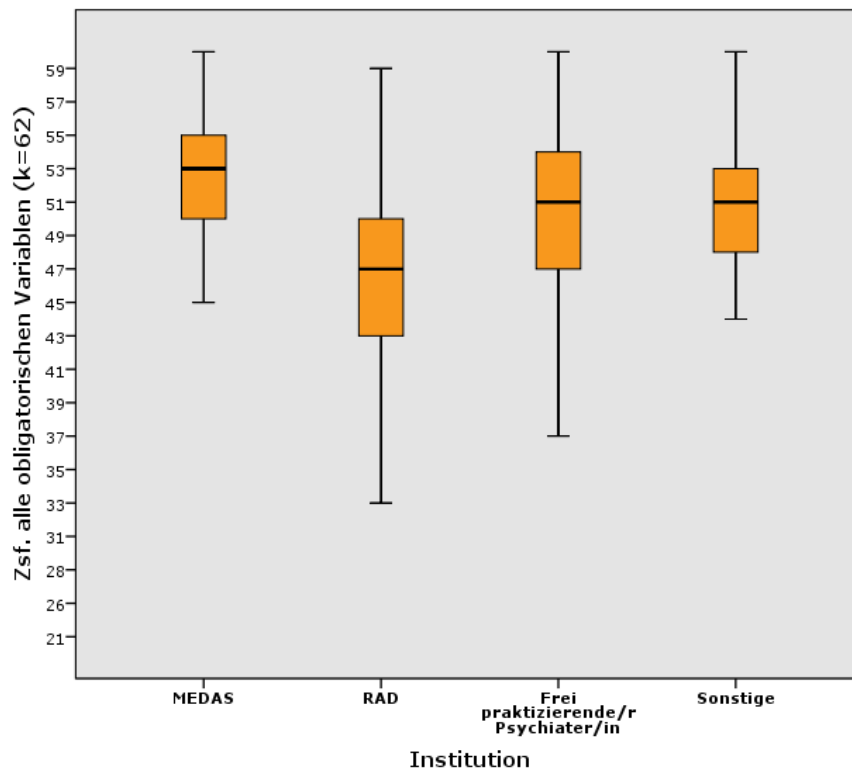
Auch diese Unterschiede sind sehr signifikant (H-Test, $\chi^2=51.160$, $df=3$, $p<.001$). In den Rangsummen sind die Gutachten der frei Praktizierenden und der Sonstigen gleichauf an erster Stelle, gefolgt von den MEDAS und dann den RAD.

5.2 Zusammenfassung über alle Kriterien

Die folgende Abbildung fasst den die Unterschiede zwischen den Typen von Institutionen im Erfüllungsgrad über alle 62 obligaten Kriterien zusammen.

Mit einem deutlich höheren Median der Werte haben zusammenfassend die MEDAS-Gutachten die beste Qualität. Nur geringfügig schwächer (bezogen auf die mittleren Ränge) folgen die frei Praktizierenden und die sonstigen Einrichtungen, deutlich weniger Kriterien sind in den Gutachten von RAD erfüllt. Die Unterschiede zwischen den Gruppen sind hoch signifikant (H-Test, $\chi^2=132.012$, $df=3$, $p<.001$).

Abbildung 5-5: Zusammenfassung aller obligatorischen Kriterien bei verschiedenen Institutionen



6 Qualitätsunterschiede zwischen den Gutachten aus verschiedenen Regionen

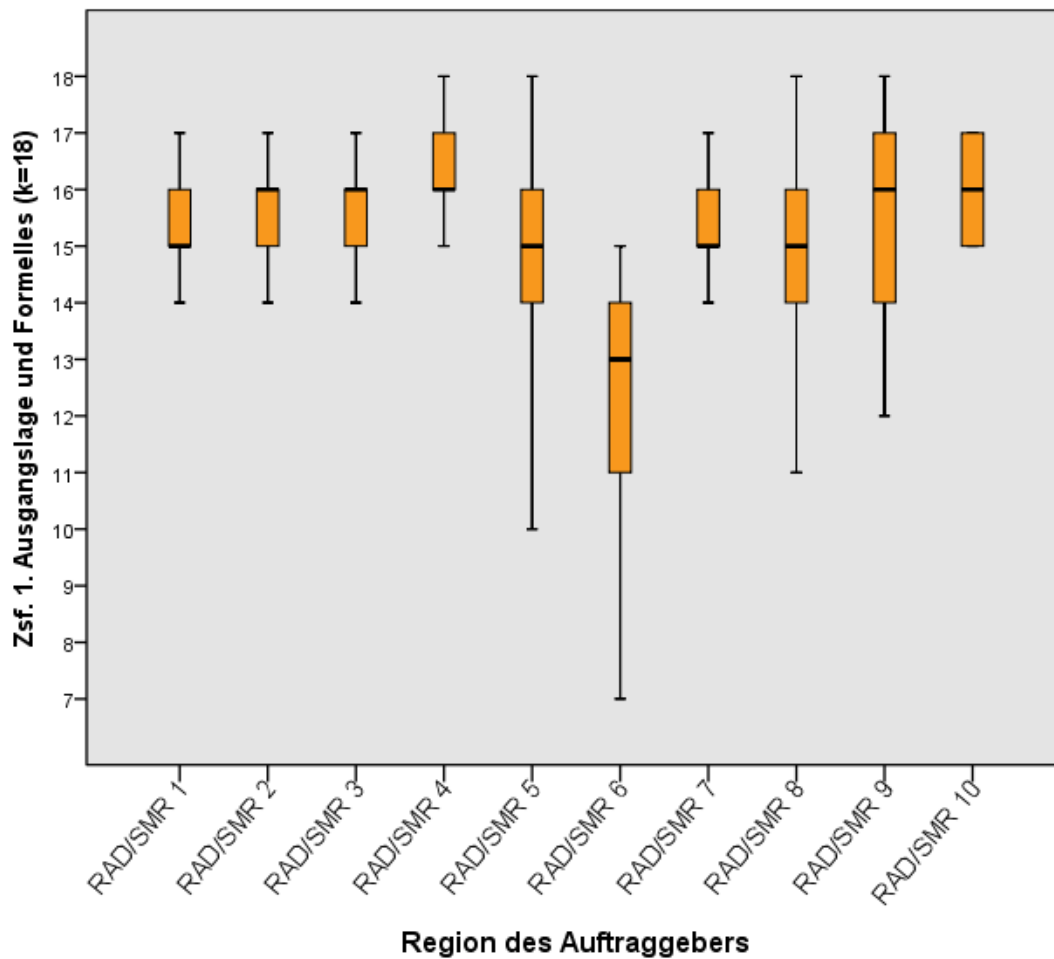
Im aktuellen Kapitel werden die Unterschiede in der formalen Qualität der Gutachten zwischen den verschiedenen Regionen dargestellt. Als Region wurde jeweils das Einzugsgebiet eines der RAD betrachtet.

6.1 Unterschiede in einzelnen Themenbereichen

Themenbereich 1. „Ausgangslage und Formelles“

Hinsichtlich der 18 Kriterien im ersten Themenbereich liegen die Mediane aller Regionen bei 15 oder 16 erfüllten Kriterien, abgesehen von im Mittel 13 erfüllten Kriterien beim RAD/SMR 6. Die Unterschiede zwischen den Gruppen sind hoch signifikant (H-Test, $\chi^2=124.958$, $df=9$, $p<.001$).

Abbildung 6-1: Bereich Ausgangslage in den verschiedenen Regionen



Die folgende Tabelle gibt die mittleren Ränge der RAD wieder, welche eine Einschätzung der Abfolge in der Qualität für diesen Bereich erlauben. N ist die Anzahl der Gutachten, die vom jeweiligen RAD ausgewertet werden konnten.

RAD/SMR	N	Mittlerer Rang
RAD/SMR 4	86	536.31
RAD/SMR 10	14	504.21
RAD/SMR 3	179	412.70
RAD/SMR 9	41	408.35
RAD/SMR 2	161	397.98
RAD/SMR 1	19	381.32
RAD/SMR 7	83	359.77
RAD/SMR 8	63	335.21
RAD/SMR 5	96	333.13
RAD/SMR 6	33	79.47

Es zeigt sich ein besonders hoher Grad der Bearbeitung bei RAD/SMR 4 und RAD/SMR 10.

Themenbereich 2. „Aktenauszug“

Wie bereits bei den Unterschieden zwischen den Institutionen bringt eine Abbildung zum Aktenauszug auch hier keinen Erkenntnisgewinn, weil alle Gutachten in allen Gruppen nahezu vollständig und damit gleich gut ohne wesentliche Varianz bearbeitet wurden. Nichtsdestoweniger sind die Unterschiede zwischen den Regionen auch hier hoch signifikant (H-Test, $\chi^2=25.259$, $df=9$, $p<.01$).

In diesem Bereich liegen Gutachten, die vom RAD/SMR 1 beauftragt wurden, gemeinsamen mit denen vom RAD/SMR 10 ganz vorne. Am unteren Ende der Skala stehen hier mit einem gewissen Abstand die Gutachten, die vom RAD/SMR 8 betreut wurden.

RAD/SMR	N	Mittlerer Rang
RAD/SMR 1	19	411.50
RAD/SMR 10	14	411.50
RAD/SMR 7	83	402.22
RAD/SMR 3	179	401.12
RAD/SMR 2	161	390.16
RAD/SMR 6	33	387.35
RAD/SMR 9	41	383.78
RAD/SMR 5	96	382.72
RAD/SMR 4	86	375.80
RAD/SMR 8	63	341.92

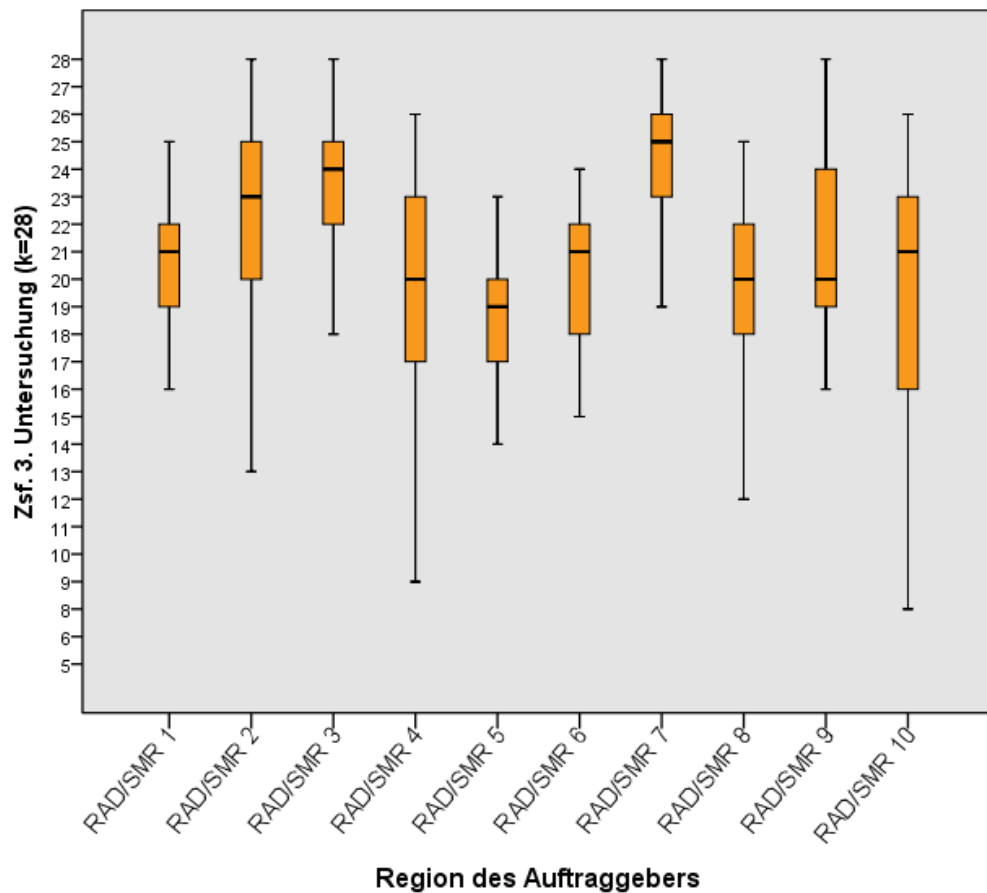
Themenbereich 3. „Untersuchung“

Im Themenbereich „Untersuchung“ liegt eine grosse Streuung der Werte vor. Die Abbildung zeigt, dass die RAD/SMR 2, RAD/SMR 3 und RAD/SMR 7 hier die besten Ergebnisse erzielen. Die Unterschiede zwischen den Regionen sind auch hier sehr signifikant (H-Test, $\chi^2=199.856$, $df=9$, $p<.001$).

Die 83 Gutachten vom RAD/SMR 7 schneiden ausweislich der mittleren Ränge mit Abstand am besten ab. Bei Gutachten des RAD/SMR 5 wurden zur Untersuchung die knappsten Angaben gemacht.

RAD/SMR	N	Mittlerer Rang
RAD/SMR 7	83	553.89
RAD/SMR 3	179	488.20
RAD/SMR 2	161	443.14
RAD/SMR 9	41	362.05
RAD/SMR 1	19	302.42
RAD/SMR 10	14	300.54
RAD/SMR 4	86	281.34
RAD/SMR 8	63	277.72
RAD/SMR 6	33	273.67
RAD/SMR 5	96	213.27

Abbildung 6-2: Bereich Untersuchung in den verschiedenen Regionen



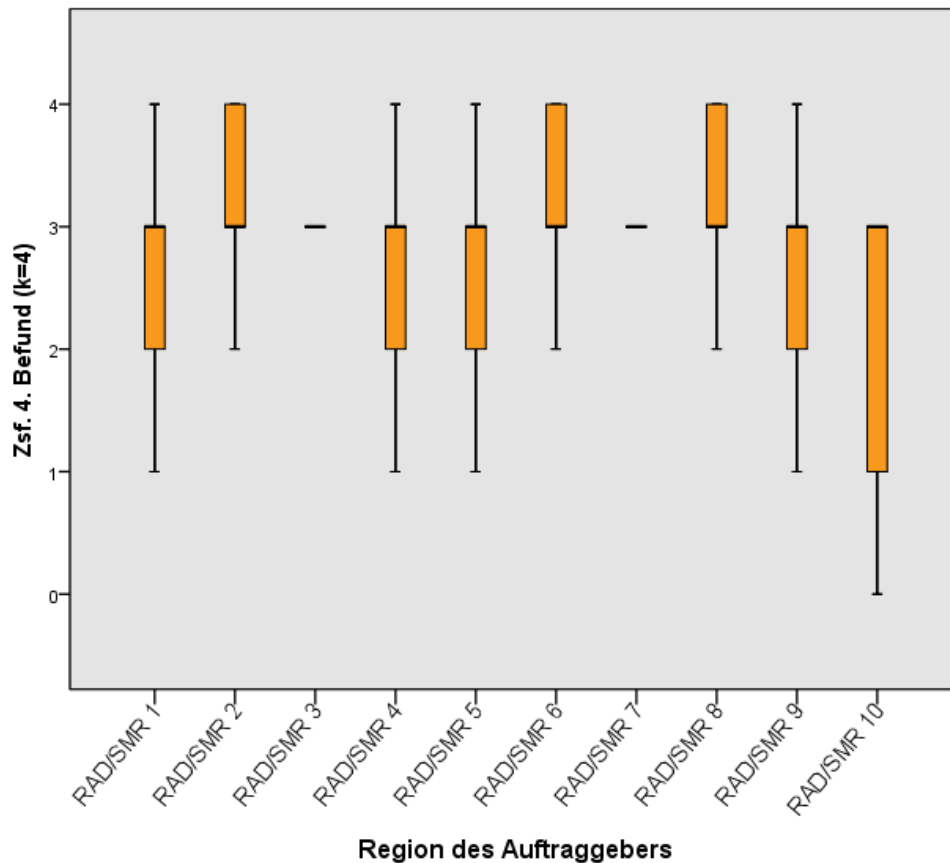
Themenbereich 4. „Befund“

Bei der Darstellung des Befundes sind die Gutachten von den RAD/SMR 8 und RAD/SMR 6 am vollständigsten. Am wenigsten der obligatorischen Kriterien in diesem Themenkomplex werden in den Gutachten vom RAD/SMR 10 berichtet.

RAD/SMR	N	Mittlerer Rang
RAD/SMR 8	63	464.35
RAD/SMR 6	33	460.73
RAD/SMR 2	161	441.17
RAD/SMR 7	83	423.42
RAD/SMR 3	179	395.44
RAD/SMR 9	41	383.07
RAD/SMR 1	19	324.05
RAD/SMR 4	86	306.97
RAD/SMR 5	96	283.95
RAD/SMR 10	14	268.89

Die Unterschiede zwischen den Regionen sind hier ebenfalls sehr signifikant (H-Test, $\chi^2=72.454$, $df=9$, $p<.001$).

Abbildung 6-3: Bereich Befund in den verschiedenen Regionen



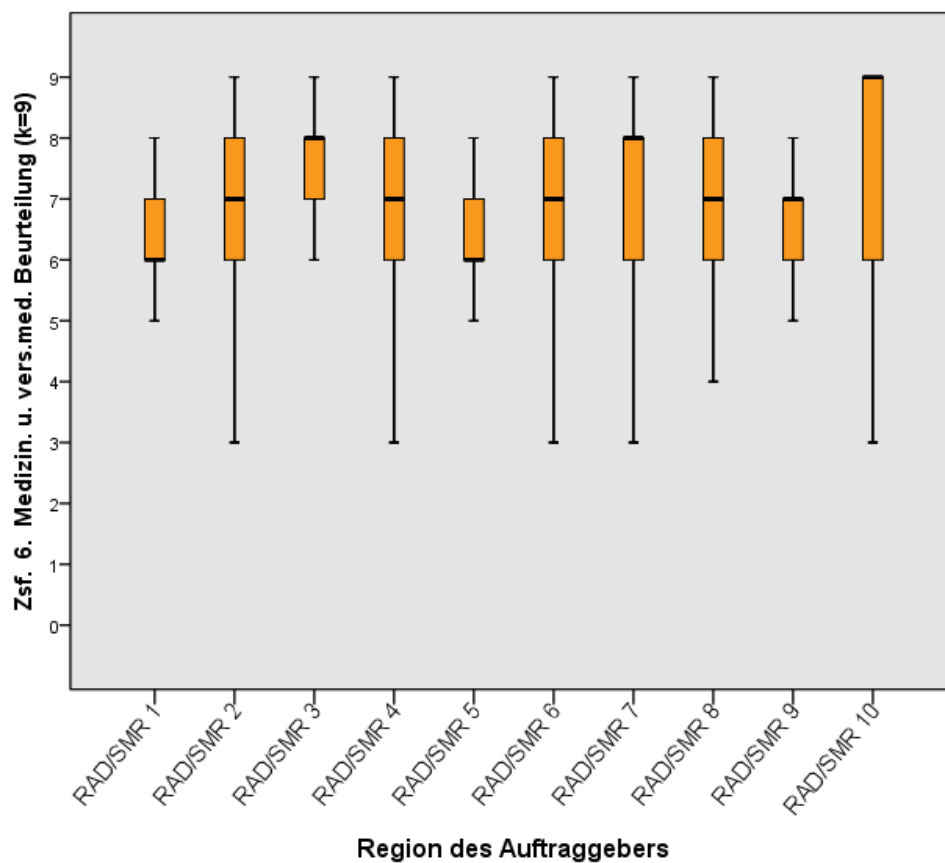
Themenbereich 6. „Medizinische und versicherungsmedizinische Beurteilung“

Der letzte der obligatorischen Themenkomplexe ist die Medizinische und versicherungsmedizinische Beurteilung der Befunde eines Gutachtens. Die erzielten Mittelwerte liegen hier in allen Regionen zwischen sechs und den maximal neun erreichbaren Punkten. Trotzdem sind die Unterschiede zwischen den Regionen auch hier sehr signifikant (H-Test, $\chi^2=55.363$, $df=9$, $p<.001$).

RAD/SMR	N	Mittlerer Rang
RAD/SMR 10	14	525.00
RAD/SMR 3	179	452.74
RAD/SMR 7	83	437.20
RAD/SMR 2	161	395.48
RAD/SMR 6	33	369.88
RAD/SMR 4	86	367.43
RAD/SMR 9	41	337.04
RAD/SMR 8	63	334.56
RAD/SMR 1	19	321.55
RAD/SMR 5	96	286.88

Am besten – immer in Bezug auf den gewählten formalen Kriterienkatalog – fallen die Ergebnisse hier mit deutlichem Abstand für die Gutachten des RAD/SMR 10 aus, gefolgt von denen vom RAD/SMR 3. Am wenigsten Aussagen zur Beurteilung der Befunde finden sich in den Gutachten, die vom RAD/SMR 5 beauftragt worden waren.

Abbildung 6-4: Bereich Beurteilung in den verschiedenen Regionen



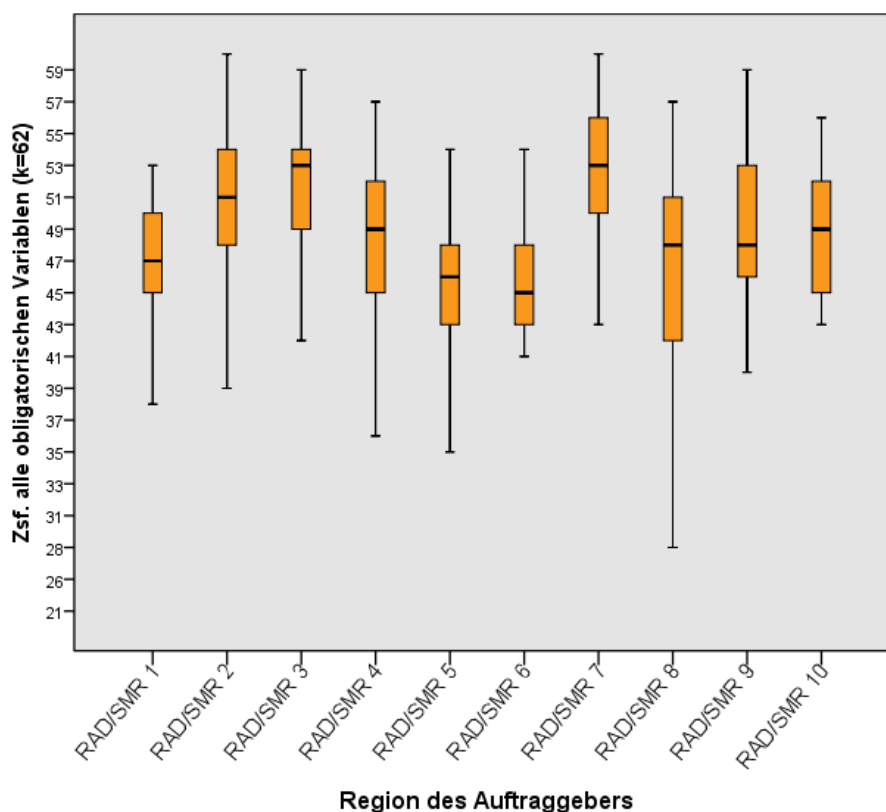
6.2 Zusammenfassung über alle obligatorischen Kriterien

Auch der Index, der alle obligatorischen Kriterien umfasst, wurde auf Unterschiede zwischen den Regionen der RAD/SMR geprüft. Die folgende Abbildung zeigt Unterschiede in diesem Indikator, die statistisch signifikant ausfallen (H-Test, $\chi^2=183.508$, $df=9$, $p<.001$).

Hier liegen die Gutachten vom RAD/SMR 7 am oberen Ende der Rangfolge, gefolgt von denen vom RAD/SMR 3 und vom RAD/SMR 2. Bezogen auf die Gesamtzahl der obligatorischen formalen Kriterien der Checkliste sind die Gutachten vom RAD/SMR 5 und vom RAD/SMR 6 am wenigsten vollständig. Auch hier muss allerdings noch einmal auf die o.g. inhaltlichen und methodischen Einschränkungen verwiesen werden, welche die Aussagekraft der Vergleiche über ein so hoch aggregiertes Mass begrenzen (vgl. Methodik in Abschnitt 3.5).

RAD/SMR	N	Mittlerer Rang
RAD/SMR 7	83	523.53
RAD/SMR 3	179	489.22
RAD/SMR 2	161	441.61
RAD/SMR 9	41	358.79
RAD/SMR 10	14	353.25
RAD/SMR 4	86	329.38
RAD/SMR 1	19	293.29
RAD/SMR 8	63	290.93
RAD/SMR 5	96	209.35
RAD/SMR 6	33	199.88

Abbildung 6-5: Zusammenfassung aller obligatorischen Kriterien in den Regionen



7 Diskussion der Ergebnisse

7.1 Bemerkungen zur Stichprobe

Durch das gewählte zweistufige Vorgehen bei der Stichprobengewinnung (1. Erhebung der Gesamtzahl aller beauftragten GA je RAD, 2. Ermittlung des Schlüssels für die Ziehung der Stichprobe aus der Gesamtzahl der GA) konnte die gewünschte Repräsentativität der Stichprobe für die gesamte Schweiz sichergestellt werden. Kleinere, im Laufe des Projekts vorgenommene Korrekturen der Stichprobe aufgrund fehlerhafter Angaben zur Gesamtzahl der Gutachten beeinflussen diese Repräsentativität nur unwesentlich (< 5%) und können bei der Ergebnisinterpretation vernachlässigt werden.

7.2 Allgemeine Bemerkungen zur Interpretation der Ergebnisse

Mit der gegebenen Repräsentativität und aufgrund der Stichprobengrösse konnten alle verfolgten Fragestellungen beantwortet werden. Die Beurteilerübereinstimmung für die Auswertung der Gutachten anhand der Checkliste ist zufriedenstellend.

Für die Interpretation der Ergebnisse sei an dieser Stelle darauf hingewiesen, dass bei der Erfüllung der formalen Kriterien das Minimalprinzip angewendet wurde. Beispiel: Der Punkt 3.2.4 „Partnerschaft“ wurde als erfüllt geratet, sobald im Gutachten z.B. eine Aussage wie „Expl. ist verheiratet“ gemacht wurde. Eine fundierte Darstellung der partnerschaftlichen Beziehung(en) des Probanden ist damit dennoch nicht belegt. Diese liesse sich nur durch eine inhaltliche Analysen entsprechend inhaltlicher Leitlinien beurteilen, was explizit nicht Gegenstand der vorliegenden Analyse war. Die Kriterien der Checkliste sind zudem von unterschiedlichem Gewicht für die Gesamtbeurteilung eines Gutachtenfalls, und diese Relevanz kann je nach Fallkonstellation auch noch von Gutachten zu Gutachten variieren. Deshalb sind insbesondere die zusammenfassenden Indikatoren in diesem Bericht nur mit grosser Vorsicht zu interpretieren.

Andererseits besteht die Möglichkeit, dass Kriterien zwar formal nicht erfüllt wurden, diese aber dennoch vom Gutachter erhoben wurden. Beispiel: Spontane Angaben des Probanden können an verschiedenen Stellen in den Gutachtentext eingeflossen sein, auch ohne dass dies zwingend als spontane Äusserung kenntlich gemacht wurde. In diesem Fall wäre das Kriterium 3.1. „Spontane Angaben des Expl.“ nicht erfüllt, obwohl diese – eben nur implizit – berücksichtigt wurden, ohne dies ausführlich darzustellen.

Denkbar wäre auch, dass in manchen Gutachten einzelne Kriterien störungsbedingt nicht erfüllt werden konnten, da z.B. aufgrund einer vorliegenden Schizophrenie des Probanden manche formalen Kriterien evtl. nicht erhebbare waren. Das Gutachten kann aber dennoch inhaltlich qualitativ gut und angemessen sein. Da diese Störungsgruppe jedoch nur einen sehr geringen Teil an der Stichprobe ausmacht (n=13), ist eine Ergebnisverzerrung der dargestellten Art in grösserem Ausmass nicht sehr wahrscheinlich.

Weiterhin soll nochmals darauf hingewiesen werden, dass die vorliegende Studie eine Implementierungs-Studie darstellt und zum Ziel hatte, den Ist-Zustand der formalen Qualität von Gutachten vor Einführung der formalen Kriterien zu erheben. Nicht alle der aufgeführten verwendeten obligatorischen Kriterien sind jedoch Bestandteil früherer Leitlinien resp. Bestandteile von Leiturtteilen des Bundesgerichtes, sondern wurden zur Erhebung des Ist-Zustandes neu aufgenommen, um nach Bekanntmachung bzw. Einführung der Leitlinien im Vergleich dazu den Zustand erneut überprüfen

zu können. Diese Tatsache muss bei der Interpretation der Ergebnisse bzgl. Qualität der Gutachten berücksichtigt werden.

Abschliessend muss an dieser Stelle bemerkt werden, dass in der vorliegenden Studie gemäss ihren Fragestellungen keine Gewichtung einzelner obligatorischer Kriterien nach möglicher versicherungsmedizinischer Relevanz vorgenommen werden konnte, da solche Gewichtungen aus dem aktuell vorliegenden Entwurf nicht hervorgehen. So wurde beispielsweise das vollständige Vorhandensein bzw. Nichtvorhandensein der Adresse (Kriterium 1.1.2) gleichgewichtig gewertet wie das Kriterium „Beruflicher Werdegang“ (3.2.5), obwohl selbstverständlich die Darstellung des beruflichen Werdegangs eine wesentlich grössere Bedeutung für die versicherungs-psychiatrische Beurteilung hat.

7.3 Aktuelles durchschnittliches Qualitätsniveau hinsichtlich formaler Begutachungskriterien

Die Ergebnisse belegen eine sehr breite Verteilung bzgl. der Berücksichtigung der formalen Kriterien: Der Grossteil der untersuchten Gutachten erfüllt die formalen Kriterien in einem Ausmass zwischen 66% und 91% der Kriterien, während wenige Gutachten nur die Hälfte der Kriterien oder weniger, ebenfalls wenige (fast) alle Kriterien berücksichtigen.

Auf der Ebene der Einzelkriterien wurde das Kriterium 4.3.1b) „Beschreibung der Persönlichkeit“ am seltensten erfüllt. Nur 21.7% aller Gutachten machten dazu Angaben. In weniger als 50% der Gutachten fanden die folgenden Kriterien Erwähnung:

- 1.3.3a) Eigene Untersuchungen: Untersuchungsdauer (ggf. je Sitzung bei mehreren Terminen)
- 3.1 Spontane Angaben des Expl. / offenes Interview
- 3.2.2d) Sexualität
- 3.2.1f) Zukunftsvorstellungen bzgl. Krankheitsverlauf
- 3.2.0a) Zukunftsvorstellungen allgemein
- 4.3.1b) Beschreibung der Persönlichkeit
- 6.3c) Kritische Würdigung der Selbsteinschätzung des Expl.

Insbesondere bei den Kriterien wie „Zukunftsvorstellung allgemein“ oder „Spontane Angaben des Expl.“ kann nicht ausgeschlossen werden, dass diese zwar vom Gutachter erhoben und auch dokumentiert, aber nicht explizit als „spontan“ gekennzeichnet wurden (Abschnitt 7.2). Andererseits ist es beispielsweise bei dem Kriterium „Spontane Angaben des Expl.“ auch denkbar, dass aufgrund der Art und Weise der Explorationstechnik (zum Beispiel am Anfang schon geschlossene Fragen seitens des Gutachters) dem Exploranden zu wenig Raum zur Verfügung stand, seine Perspektive frei zu schildern.

Des Weiteren muss z.B. bei Kriterien wie der Anamnese der „Sexualität“ (nicht als Beschwerde) überlegt werden, inwiefern die geringe Vollständigkeit dieses Kriterium mit der Relevanz in der Versicherungsmedizin – im Gegensatz zur psychiatrisch-psychotherapeutischen Behandlung oder in der strafrechtlichen Begutachtung – zu tun haben könnte, d.h. wie es gewichtet werden sollte. In diesem Fall sollte das Nichterfüllen solcher Kriterien aus inhaltlichen Erwägungen, die hier allerdings nicht zur Diskussion standen, nicht als gravierender Mangel interpretiert werden. Bei der nachfolgenden Evaluation von Gutachten wäre dringend anzuraten, bei einzelnen Kriterien eine inhaltlich fundierte Gewichtung vorzunehmen!

Am vollständigsten waren die deskriptiven Angaben zum Exploranden, zum Gutachter und Auftraggeber. Ebenfalls relativ häufig waren Angaben zur aktuellen Krankheitssituation mit den Kriterien 3.2.1a) „Aktuelle Beschwerden“, das von 97.9% der Gutachten erfüllt wurde, 3.2.2a) „psychiatrische Anamnese“ von 95.4% sowie 3.2.4c) „Partnerschaft“ und 2.c) „(relevante) Diagnosen / Befunde“ von je 94.7%.

Im Hinblick auf die in den Gutachten weniger komplett dargestellten Inhaltsbereiche bzw. formalen Kriterien ist zu bemerken, dass alle ausgewerteten Gutachten vor der Definition der Leitlinien erstellt wurden, den Begutachtenden also nicht der Vorwurf zu machen ist, sich nicht an bekannte Vorgaben gehalten zu haben. Konzipiert als Ausgangsmessung vor Einführung der Leitlinien entspricht dies auch dem Evaluationskonzept. Erst nach der Veröffentlichung und der Implementierung der Leitlinien müssen sich die Gutachten (bzw. deren Verfasser) an diesen dann verbindlichen Vorgaben messen lassen (vgl. Abschnitt 7.2).

Stattdessen geben die weniger komplett bearbeiteten Kriterien Hinweise darauf, welche Themenbereiche von den Gutachtern und Gutachterinnen als weniger relevant oder als schwieriger zu beurteilen angesehen werden. Diese Themen sollten bei der Begründung und Einführung der Leitlinien für die Begutachtung besonders vertieft werden.

7.4 Durchschnittliche formale Qualität der einzelnen Gutachten- teile und systematische Qualitätsunterschiede zwischen die- sen

Prinzipiell ist festzuhalten, dass abgesehen von den Kriterien Name und Vorname des Exploranden keines der formulierten Kriterien in 100% der analysierten Gutachten erfüllt war.

Bezüglich des **Aktenauszugs** erreichen fast alle Gutachten hohe Werte, obwohl auch hier kein einziges Kriterium von allen Gutachten zu 100 % erfüllt wurde.

In fast allen Gutachten (mehr als 90%) wurde die Adresse des Gutachters, sein Facharztstitel, das Explorationsdatum, das Datum der Gutachtenfertigstellung und der Auftraggeber des Gutachtens genannt. Unter 90% blieben das Auftragsdatum, das Eingangsdatum des Auftrags, der Zweck des Gutachtens, die Angaben zur eigenen Untersuchungsdauer und in welcher Sprache mit dem Exploranden kommuniziert wurde.

Beim Oberkriterium „3. **Untersuchung**“ wurde vor allem das Krankheitskonzept des Exploranden, der Umgang mit seinen Beschwerden im Alltag, seine Zukunftsvorstellungen, die Sexualanamnese, der soziale Empfangsraum (Wohnung, Finanzen, Beziehungsnetz), das Datum und die Umstände der Arbeitsaufgabe, die Freizeitgestaltung, benötigte Hilfen im Haushalt und im Alltag, die aktuelle Behandlung und die Zukunftsvorstellungen des Exploranden nicht wie vom Kriterienkatalog gefordert dargestellt.

Vereinfacht gesagt findet die Sicht des Exploranden wenig Beachtung. Dies ist insbesondere dann von Bedeutung, wenn es um die Empfehlung von weiteren Massnahmen (z. B. Wiedereingliederung, Umschulung, etc.) und die Motivation des Exploranden hierfür geht. Das eigene Krankheitskonzept wie auch die Motivation sind jedoch – besonders bei den häufigsten Diagnosen depressive und somatoforme Störungen – wesentliche Faktoren für eine optimale Anpassung der vorgeschlagenen Leistungen und Interventionen an die Kompetenzen und Möglichkeiten der Exploranden.

Neben den subjektiven Daten werden häufig aber auch relevante objektive Daten nicht berücksichtigt. So wurden in ca. 10% Angaben zur aktuellen Behandlung und in ca. 28% der Fälle der Konsum psychotroper Substanzen nicht erfragt, in einem Viertel der Fälle wurden Datum und Umstand der Arbeitsaufgabe nicht erwähnt.

Beim „**Befund**“ wurden in weniger als 90% der Fälle Verhaltensbeobachtungen und die äussere Erscheinung beschrieben. Deutlich vernachlässigt wurde die Beschreibung der Persönlichkeit des Exploranden (22%).

Für die Auftraggeber besonders relevant ist das Kriterium 6. der **medizinischen und versicherungsmedizinischen Beurteilung**. Nur ca. 7% aller Gutachten berücksichtigen hier alle Subkriterien. Während das Unterkriterium 6.4 „Beantwortung der Fragen“ in der Regel berücksichtigt wird, werden beispielsweise unter „Diskussion und versicherungsmedizinische Würdigung“ (Punkt

6.3.) nur bei ca. zwei Drittel der Gutachten Funktions- und Fähigkeitsstörungen hergeleitet, ansonsten häufig nur als Feststellung dokumentiert. Dadurch werden Transparenz, Nachvollziehbarkeit und Überprüfbarkeit der Folgerungen eingeschränkt. Auch die Informationen zu vorhandenen Ressourcen des Exploranden, zu seiner Kooperationsbereitschaft und die Integration aller Einschätzungen und Informationen liegen teilweise weit unter 70%. Möglicherweise könnte die Fülle an Informationen hier teilweise eine Integration des Wissens behindern.

Eindrücklich scheint es zudem im Zusammenhang zur Frage nach dem Einfluss einer psychischen Störung auf die Arbeitsfähigkeit, dass die medizinische Beurteilung, d.h. Diagnose und Differentialdiagnose, nur selten hergeleitet, begründet und ausführlich in den Gutachten dokumentiert werden. Häufig wurde eine Diagnose gestellt, im GA wurde jedoch nicht dokumentiert, wie es zur Bestimmung dieser Diagnose kam bzw. wie diese differentialdiagnostisch abgegrenzt werden könnte (z.B. PTSD/F43.1 und Persönlichkeitsänderung nach psychischer Krankheit/F62.0).

Generell sehr vollständig werden die administrativen Angaben zu Auftraggeber, ExplorandIn, GutachterIn etc. berichtet, ebenso zum Aktenauszug. Weiterhin werden die Kriterien im Allgemeinen sehr vollständig dargelegt, welche als „medizinisch“ im engeren Sinn bezeichnet werden können, etwa die psychiatrische Anamnese, aktuelle Beschwerden, Behandlungsmassnahmen. Grössere Variabilität über alle Gutachten hinweg und deshalb zumindest potenziell auch grössere Unterschiede zwischen Gruppen von Gutachten (Institutionen, Regionen) bestehen im Bereich der Untersuchung, aber auch hinsichtlich des Befundes und seiner Beurteilung, bei den psychosozialen Kriterien, z.B. Krankheitskonzepten, Zukunftsvorstellungen bzgl. Krankheitsverlauf, Umgang mit der Krankheit im Alltag, Persönlichkeit des Exploranden bzw. der Explorandin, Ressourcen und Selbsteinschätzung des Exploranden/der Explorandin. Möglicherweise gehen die Gutachter (vielleicht auf der Basis ihrer primär psychiatrischen Ausbildung) bei der Beurteilung der Arbeitsfähigkeit stärker nach einem bio-psycho-sozialen Krankheitsverständnis vor als nach einem versicherungsmedizinischen bzw. IV-rechtlichen, z.B. bezüglich des Einbezugs von Chancen auf dem Arbeitsmarkt. Schon in der „Dossieranalyse“ (Baer et al., 2009) stellten die Autoren fest, dass Beschreibungen von Funktionseinschränkungen wenig detailliert und die Folgen der Erkrankung auf die Arbeitsfähigkeit in den medizinischen Gutachten für eine umfassende Beurteilung der Eingliederungsmöglichkeiten und auch für die Planung allfälliger rehabilitativer Massnahmen häufig zu knapp seien. Die rein medizinische Abklärung wurde hingegen als umfassend und intensiv bewertet.

In lediglich 5 der untersuchten Gutachten wurde dokumentiert, auf welcher Grundlage die Beurteilung der Schwere einer Störung und ihre Auswirkungen auf die „zumutbare“ Arbeitsfähigkeit erfolgten. Die genauere inhaltliche Untersuchung dieser 5 Gutachten zeigte, dass in einem Gutachten das ICF, in zwei Gutachten eine „Checkliste“ und in weiteren zwei Gutachten die sogenannten „Foerster-Kriterien“⁵ als Instrumente dokumentiert wurden. Die „Foerster-Kriterien“ besitzen allerdings v.a. für somatoforme Schmerzstörungen Gültigkeit, sie sind nicht auf alle psychiatrischen Gutachtenfälle anwendbar. Auf jeden Fall ergeben sich hier Ansatzpunkte für die Entwicklung bzw. Verbreitung valider, standardisierter Instrumentarien als Grundlage für die Beurteilung der Arbeitsfähigkeit.

Im Fehlen solcher anerkannter valider Instrumente könnte der Hauptgrund dafür liegen, dass die Arbeitsfähigkeit nicht in strukturierter Form hergeleitet wurde, sondern, wie früher auch in strafrechtlichen Begutachtungen beschrieben, primär intuitiv vorgegangen wurde. Hierbei ist es dann möglich, dass der psychiatrische Gutachter, welcher von seiner Ausbildung her eher einem bio-psycho-soziales Verständnis folgt, auch soziale Aspekte wie z.B. Chancen auf dem Arbeitsmarkt, eher

⁵ Auch wenn Krankheitsdiagnosen *per se* keinen Rentenanspruch begründen, können ungünstige konstellative Faktoren dennoch dazu beitragen, dass das Beschwerdebild auch bei zumutbarer Willensanstrengung nicht mehr als überwindbar erscheint, so dass in der Summe der Beeinträchtigungen und insbesondere bei schwachen Ressourcen dennoch ein Leistungsanspruch erwachsen kann. Um solchen erschwerenden Problemkonstellationen Rechnung zu tragen, wurden die acht sog. Foerster-Kriterien von der Schweizer Versicherungsgerichtsbarkeit als massgeblich erklärt (Knecht, 2011; Winckler & Foerster, 1996)

intuitiv mit einbezieht, dies aber nicht systematisch im Gutachten dokumentiert. Solchen Verzerrungen könnte dann in der Tat mit strukturierten Instrumenten begegnet werden.

7.5 Qualitätsunterschiede zwischen den Gutachten verschiedener Institutionen

Untersucht wurden die Gutachten auch auf systematische Unterschiede zwischen den Einrichtungen, an denen sie erstellt wurden. Die Gutachten wurden gruppiert nach den Typen (1) Frei praktizierende Psychiater und Psychiaterinnen, (2) Medizinische Abklärungsstellen (MEDAS), (3) Regionalärztliche Dienste (RAD) der IV, und (4) sonstige Institutionen.

Im Themenbereich „Ausgangslage und Formelles“ zeigen die Mediane von MEDAS, RAD und Sonstigen das gleiche Niveau bei leicht unterschiedlichen Verteilungen. Die mittlere Zahl berücksichtigter Kriterien liegt bei den frei praktizierenden Psychiatern um eine Stufe unter den anderen Gruppen.

In allen vier Gruppen wird das Kriterium „Aktenauszug“ nahezu vollständig erfüllt, hier ergaben sich keine wesentlichen Unterschiede.

Im Bereich der „Medizinischen und versicherungsmedizinischen Beurteilung“ liegen die Niveaus der vier Gruppen ähnlich, wobei im Median die Gutachten der RAD signifikant ein Kriterium weniger berücksichtigen als die anderen.

Der Themenbereich „Untersuchung“ zeigt hochsignifikante Unterschiede zwischen den Institutionen (MEDAS > frei praktizierende PsychiaterInnen > sonstige Einrichtungen > RAD).

Zusammengefasst über alle Themenbereiche hinweg zeigt sich dieselbe Rangreihe.

Insgesamt sind die Unterschiede zwischen den begutachtenden Einrichtungen nicht gross. Tendenziell, wenn auch nicht in allen Themengebieten, sind die Gutachten der RAD etwas weniger ausführlich bzw. vollständig als die Gutachten aus anderen Einrichtungstypen. MEDAS-Gutachten sind in der Regel am vollständigsten; bei diesen handelt es sich aber besonders häufig auch um polydisziplinäre Gutachten mit mehreren Teilgutachten, sodass sich ein grösserer Umfang und damit auch eine grössere Vollständigkeit der MEDAS-Gutachten teilweise schon aus der Art der Gutachten ergeben könnte. Weitere Gründe für die Unterschiede lassen sich aus den vorliegenden Daten nicht entnehmen.

7.6 Systematische Qualitätsunterschiede zwischen den Gutachten aus verschiedenen Regionen

Über alle Kriterien hinweg liegen die Gutachten vom RAD/SMR 7 am oberen Ende der Rangfolge nach Vollständigkeit, gefolgt von denen vom RAD/SMR 3 und vom RAD/SMR 2. Am wenigsten vollständig sind die Gutachten vom RAD/SMR 5 und vom RAD/SMR 6.

Unterschiede zwischen den Regionen zeigen sich auch bezüglich der einzelnen Gutachtenteile. Während sich beim Kriterium „Untersuchung“ eine ähnliche Rangreihe zeigt wie gesamthaft (allerdings bei grossen Streuungen), wird z.B. das Kriterium „Befund“ am vollständigsten in den von den RAD/SMR 8 und RAD/SMR 6, am wenigsten von den vom RAD/SMR 10 beauftragten Gutachten berücksichtigt. Im Bereich der Ausgangslage haben die Gutachten vom RAD/SMR 4 die beste Vollständigkeit, beim Aktenauszug diejenigen vom RAD/SMR 1 und RAD/SMR 10, letztere auch bei der Medizinischen und versicherungsmedizinischen Beurteilung.

Insofern lässt sich aus den Ergebnissen kein klares Muster der Gutachtenqualität per se in den Regionen erkennen, was kaum verwundern kann, da diejenigen Kriterien, welche neu aufgenommen wurden und bei der Evaluation schon als Grundlage zur Erhebung des Ist-Zustandes dienen mussten, ja den Gutachtern beim Abfassen der Gutachten noch gar nicht bekannt waren, da sie – immer

noch in einem Entwurf vorliegend – noch nicht publiziert sind. Nichtsdestoweniger bestehen teils deutliche Unterschiede, die auf verschiedene Stile bzw. Gewichtungen der Gutachtenteile hinweisen. Wie oben berichtet (in Abschnitt 7.4), besteht die insgesamt grössere Variabilität bei der Vollständigkeit der Gutachten nicht hinsichtlich der „medizinischen“ Themen, sondern in Bezug auf die psychosozialen Kriterien. Diese sind einerseits schwerer zu fassen und andererseits ist ihr Einfluss auf die versicherungs-psychiatrische Beurteilung nicht immer gleich stark ausgeprägt. Denkbar ist deshalb, dass Gutachterinnen und Gutachter mit ihrem regional verantwortlichen RAD zu einer – vielleicht auch nur impliziten – Übereinkunft gekommen sind, dass manche Themenbereiche nicht immer und in allen Fällen explizit erwähnt werden müssen, wenn dadurch kein Informationsverlust entsteht. Auch die Häufigkeit oder Intensität, mit der die Gutachtenden von den RAD für weitere Klärungen über die schriftlichen Gutachten hinaus in Anspruch genommen werden, kann sich unterscheiden. Somit könnte die Informationsbasis für Entscheidungen der IV auch bei etwas hetero- genen Gutachten, welche teilweise nicht alle formalen Kriterien erfüllt haben, in ausreichendem Umfang gewährleistet sein.

Insgesamt zeigen die aufgefundenen Unterschiede zwischen Institutionen und Regionen, dass eine grössere Vereinheitlichung durch die Einführung der Gutachten-Leitlinien möglich und vielversprechend ist.

7.7 Unterschiede zwischen Arten von Gutachten

Polydisziplinäre Gutachten zeigen sich gesamthaft als vollständiger als die mono- und bidisziplinären Gutachten (hoch signifikant). Insbesondere betrifft dies die Kriterien „Ausgangslage“ und „Untersuchung“. Dies scheint von daher erwartungsgemäss, als polydisziplinäre Gutachten ohnehin ausführlicher angelegt sind als z.B. monodisziplinäre, insbesondere was die Untersuchung angeht.

Monodisziplinäre Gutachten berücksichtigen dagegen im Bereich „Befunde“ die meisten Kriterien.

Interessanterweise unterscheiden sich die Gutachtenarten nicht in der „Medizinischen und versicherungsmedizinischen Beurteilung“. Offenbar gelten die oben im Zusammenhang mit der unterschiedlichen Berücksichtigung verschiedener Gutachtenteile gemachten Feststellungen auch über die unterschiedlichen Gutachtentypen hinweg. Die ausführlichere Untersuchung bei polydisziplinären Gutachten schlägt sich also nicht in einer vollständigeren Beurteilung nieder. Gestützt wird dies auch durch die Analyse der fakultativen Kriterien (6.2a) – 6.2d), die sich nur auf polydisziplinäre Gutachten und dabei insbesondere auf die Integration der Befunde aus den Teilgutachten beziehen. Hinweise auf eine Synthese aller Expertenberichte hatten sich nur in zwei Dritteln der polydisziplinären Gutachten gezeigt (vgl. Tabelle 4-11, S. 45), allfällige Differenzen zwischen den Gutachtern wurden gar nur in 40% der polydisziplinären Gutachten diskutiert und geklärt. Bei der Integration der Teilgutachten im polydisziplinären Bereich besteht also deutlicher Verbesserungsbedarf.

8 Schlussfolgerungen und Empfehlungen

Die vorliegende Studie beschreibt den Status quo vor Einführung der Gutachten-Leitlinien. Bei der Erstellung der Gutachten, welche in den Jahren 2008 und 2009 erstellt und nun für diese Studie analysiert wurden, waren die Bewertungskriterien den Gutachtenden noch nicht bekannt. Deshalb lassen sich die Leitlinien selbst noch nicht empirisch auf ihre Praxiswirkung prüfen. Dies bleibt der zweiten Evaluationsphase nach Einführung der Leitlinien vorbehalten. Auf der Grundlage der Ergebnisse lassen sich aber bereits jetzt Bereiche identifizieren, welche für die Entwicklung und Einführung der Leitlinien besonders bedeutsam sein könnten.

8.1 Konsequenzen für die weitere Leitlinienentwicklung

Zur zuverlässigen Messung der Qualität von versicherungs-psychiatrischen Gutachten zeichnen sich folgende Empfehlungen ab:

Auf der Basis der vorliegenden Ergebnisse zur formalen Qualität hinsichtlich der Konsens-Kriterien lässt sich zunächst schlussfolgern, dass **Präzisierungen der Kriterien** bzgl. der einzelnen Gutachtenteile hilfreich wären. Es sind insbesondere solche Kriterien ungenügend bearbeitet, welche nicht leicht standardisiert zu erfassen sind, wie z.B. 4.3.1b) „Beschreibung der Persönlichkeit“ und 3.2.1f) „Zukunftsvorstellungen bzgl. Krankheitsverlauf“. Hier wären sicherlich Handreichungen für die Gutachter hilfreich, wie diese Punkte optimalerweise gefüllt werden sollten. Zudem könnte bspw. die mangelnde Berücksichtigung der subjektiven Einschätzungen des Probanden aufgegriffen und dieses Kriterium weiter differenziert werden. Die Sicht des Exploranden zu Themen wie „Angaben zum Krankheitskonzept des Expl.“, „Zukunftsvorstellungen bzgl. Krankheitsverlauf“ oder „allgemeine Zukunftsvorstellungen“ wird oft nicht dokumentiert. Möglicherweise sind diese Konstrukte zu wenig konkret und müssten praktikabel operationalisiert werden. Zudem waren sie noch nicht Bestandteil der gültigen Vorgaben.

Was die Soziale Anamnese betrifft, finden sich bisher selten Angaben zum „Beziehungsnetz“, zu „finanziellen Verhältnissen“ oder „Wohnsituation“. In der vorliegenden Untersuchung wurde Beziehungsnetz definiert als soziale Integration/Funktion/Netzwerk ausserhalb von Partnerschaft und Kindern, da diese in eigenen Kriterien erfasst werden. Bei inhaltlichen Leitlinien wäre der Begriff des „Beziehungsnetzes“ ggf. näher zu definieren.

Aber auch die Empfehlung, falls erforderlich objektive Laborwerte zu erheben, aktuelle Laborwerte vom Hausarzt einzuholen oder z.B. bei Verdacht auf Alkoholabusus entsprechende Langzeitparameter zu erheben, könnte die Beurteilungsgrundlage verbessern.

So wie Laborparameter nicht in allen Fällen relevant sind, gilt dies auch für andere Kriterien. Einzelne Themenbereiche wie z.B. die soziale Anamnese sind grundsätzlich von höherer Bedeutung für die versicherungsmedizinische Beurteilung eines Falles als andere Kriterien; wiederum andere Themenbereiche sind nur bei bestimmten Störungsbildern oder individuellen Konstellationen relevant (z.B. Anamnese von „Sexualität“ vs. „Konsum psychotroper Substanzen“). Es sollte also eine **Gewichtung oder Priorisierung der Kriterien** vorgenommen werden, welche den Gutachterinnen und Gutachtern Hinweise darauf gibt, unter welchen Umständen welche Kriterien vertieft exploriert und dargestellt werden sollten. Andererseits könnte mit Hilfe solcher Regeln auch definiert werden, unter welchen Umständen die Angaben zu gewissen Kriterien weggelassen werden können. Bei Kern-Kriterien sollte allerdings auch dokumentiert werden, wenn z.B. keine Medikamente aktuell eingenommen werden oder keine Laboruntersuchungen oder kein offenes Interview vorgenommen wurden und warum nicht; ganz besonders wichtig erscheint uns dies bei den eigentlichen Kernbereichen einer medizinischen psychiatrischen Begutachtung: Befund, medizinische Beurteilung (Herleitung von Diagnose und Differentialdiagnose) und versicherungs-psychiatrische Beurteilung (nachvollziehbare Herleitung von Funktionen, Fähigkeiten, Partizipation mit Arbeitsfähigkeit, Prognose, Massnahmen), sicherlich auch in Teilbereichen der Anamneseerhebung.

Wie dargestellt wurde, sind die mittels Minimalprinzip gewonnenen Ergebnisse bzgl. der Erfüllung formalen Kriterien eine notwendige, aber keinesfalls eine hinreichende Bedingung für qualitativ hochstehende Gutachten. Erst die Entwicklung und Einhaltung inhaltlicher Qualitätskriterien erlaubt hier eine abschliessende Beurteilung.

Für die **inhaltliche Leitlinienentwicklung** zeigen die Ergebnisse zur formalen Qualität Ansatzpunkte: Wie am Beispiel der „Diskussion und versicherungsmedizinischen Würdigung“ gezeigt, mangelt es an Herleitungen von Funktions- und Fähigkeitsstörungen. Eine Definition von ‚Herleitung‘ in den inhaltlichen Kriterien kann hier eine Verbesserung der Nachvollziehbarkeit bewirken.

Sollte weiterhin eine Verbesserung von Reliabilität und Validität versicherungsrechtlich relevanter Beurteilungen erwartet werden, so wird man aus Sicht der Autoren nicht um die Entwicklung von anerkannten validen **Instrumenten zur formell und inhaltlich angemessenen diagnostischen Unterstützung** der versicherungs-psychiatrischen Beurteilung herkommen. Durch die Empfehlung von und die Schulung des Einsatzes standardisierter Verfahren (z.B. zur Operationalisierung der ICF, validiert auf Probanden in Rentenverfahren) könnte die Qualität der Beurteilung einerseits und die Transparenz der gewonnenen Einschätzung andererseits deutlich erhöht werden. Möglicherweise könnten die somit verbesserten, detaillierteren Beschreibungen der Funktionsfähigkeit die Planung rehabilitativer Massnahmen zur Wiedereingliederung unterstützen und in manchen Fällen eine vorzeitige Berentung überflüssig machen. Abzuwarten bleibt, ob die Entwicklung neuer Kriterien für „F4“-Diagnosen in DSM-V und ICD-11 in Bezug auf eine höhere Gültigkeit Reliabilität und Validität verbessern und damit zusätzlich mithelfen, die Gutachtenqualität zu erhöhen.

Da die am häufigsten gestellten Diagnosen rezidivierende depressive Störung und depressive Episoden (zusammen n=239) gefolgt von somatoformen Störungen (n=88) darstellen, wäre im Rahmen der Leitlinienentwicklung eine diagnosenabhängige **Schwerpunktbildung** zur Vertiefung der Befundung zu diskutieren. Empfehlungen speziell zur Begutachtung dieser häufigsten Störungen evtl. mit störungsspezifischen Kriterien könnten eine erste Fokussierung darstellen. Hier könnte auch gezielt auf die festgestellte geringe Nutzung eigener zusätzlicher medizinischer Untersuchungen eingegangen werden.

Für die Entwicklung von inhaltlichen Leitlinien besteht also dringender und hoher Forschungsbedarf insbesondere im Hinblick auf die Fundierung der versicherungspsychiatrischen Beurteilung.

8.2 Konsequenzen für die Leitlinienimplementierung

Die Information aller Gutachter zu den vorliegenden formalen Leitlinien ist der erste Schritt in der Leitlinienverbreitung. Überregional gültige Leitlinien sind ein sinnvoller Ansatz hin zu einer einheitlichen Sichtweise von Qualität. Da jedoch Information allein noch keine Veränderungen bewirkt, kommen Weiterbildungsmaßnahmen bei der Implementierung der Leitlinien eine wesentliche Funktion zu. Naheliegend ist hierbei, eine Fokussierung auf die identifizierten Schwächen vorzunehmen und speziell die wenig berücksichtigten Kriterien aufzugreifen. Ebenso ist eine Gewichtung von Schulungsinhalten entsprechend ihrer Relevanz für die Beantwortung von Fragestellungen denkbar. Weiter können die gefundenen institutionellen und regionalen Unterschiede Hinweise für eine allfällige Priorisierung von Schulungsmassnahmen mit einer Differenzierung nach deren Zielgruppe und Zielregion gesehen werden. Dazu wäre es aber günstig, die gefundenen oder auch nur die vermuteten Unterschiede in der Gutachtenpraxis zwischen den Gruppen von Gutachtern und zwischen den RAD zu verifizieren, etwa indem die Ergebnisse des vorliegenden Berichts mit VertreterInnen diskutiert werden.

Wie sich zeigte, wird zur Erfassung der Psychopathologie fast immer das AMDP benutzt, während selten die Persönlichkeit beschrieben wird (z.B. „Big Five“ o.ä.). Möglicherweise ist dies ein Bereich, der bisher eher von psychologischen Fachpersonen behandelt wird, während den ärztlichen Gutachtern möglicherweise teilweise der Zugang zu solchen weitergehenden Diagnosesystemen fehlt. Hier wäre entsprechende Weiterbildung hilfreich, um diese Bereiche weiter zu stärken.

Zudem könnten störungsspezifische, normierte Diagnose-Instrumente wie z.B. das Beck Depression Inventory (BDI) zur Selbstbeurteilung oder die Montgomery-Åsberg Depression Rating Scale (MADRS) zur Fremdbeurteilung eingesetzt werden, um den Schweregrad des Syndroms zu quantifizieren. Die Leitlinien für die Begutachtung im IV-Verfahren könnten sich so den Leitlinien für die Diagnostik und Therapie psychiatrischer Störungen in der klinischen Praxis annähern und Befunde aus der therapeutischen Praxis integrieren.

8.3 Empfehlungen für die weitere Evaluation der Leitlinienentwicklung und –implementierung

Die vorgestellte Untersuchung ist die Prä-Messung vor Einführung der Leitlinien. Für die Abschätzung des Umsetzungsgrades nach Einführung der Leitlinien und damit für die Praxis-Wirkungen der Leitlinien sowie für die Feinsteuerung der Leitlinienimplementierung sind Post-Erhebungen erforderlich. Zur optimalen Vergleichbarkeit der Ergebnisse aus der zweiten Phase mit den vorliegenden Ergebnissen sollte das Verfahren zur Datenerhebung und –analyse möglichst gut mit dem hier beschriebenen Verfahren übereinstimmen. Eine gewisse Reduktion der Fallzahlen ist möglich, wobei man allerdings dann auf Aussagen auf der Ebene der einzelnen RAD verzichten müsste. Schon innerhalb der jetzt vorliegenden Stichprobe von ca. 800 Gutachten sind die „kleineren“ RAD nur mit wenigen Gutachten vertreten. Eine weitere Reduktion würde die Möglichkeit, Subgruppenanalysen durchzuführen, weiter beschränken.

Dabei könnten auch weitere, möglicherweise relevante Einflussfaktoren miterhoben werden. So könnten die Expertise der Gutachter oder die Zeit zwischen Gutachtensauftrag und –fertigstellung wichtige Hinweise auch zur Prozessoptimierung geben.

Die vorliegende Analyse der Erfüllung der formalen Kriterien stösst an vielen Stellen an die Grenzen ihrer Aussagekraft. Dies wurde insbesondere dort deutlich, wo Kriterien eine unterschiedliche Bedeutung für den Begutachtungszweck besitzen. Hier sind in Zukunft die inhaltlichen Kriterien mit einzubeziehen, um differenziertere Aussagen über die Relevanz einzelner Kriterien machen zu können.

Seit Beginn der vorliegenden Evaluationsstudie wurden die formalen Kriterien bereits weiterentwickelt; dies wäre bei der Zweituntersuchung zu berücksichtigen. Um Veränderungen zum hier dargestellten Status quo beschreiben zu können, sollten die Gutachten bei der Zweiterhebung sowohl nach dem ersten Satz von Kriterien als auch nach der Neufassung ausgewertet werden. Letzteres würde dann dem Soll-Ist-Vergleich dienen.

Anhang

Anhang 1: Checkliste

Gutachtennummer: _____

Checkliste

Basisdaten

a. Gutachter / Gutachten	
Fachtitel Gutachter	
Institution	<input type="checkbox"/> MEDAS <input type="checkbox"/> RAD <input type="checkbox"/> Niedergelassene Ärzte <input type="checkbox"/> Sonstige
Region Gutachter	
Region Auftraggeber	
Art des Gutachtens	<input type="checkbox"/> Monodisz. <input type="checkbox"/> Bidisz. <input type="checkbox"/> Polydisz. <input type="checkbox"/> N. beantwortbar
Monat und Jahr des Gutachtens	

b. Explorand	
Geschlecht	<input type="checkbox"/> männlich <input type="checkbox"/> weiblich
Alter	
Nationalität	
Kanton	
Erwerbsstatus der Person	
% der Arbeitsfähigkeit	
Diagnose	

Sonstige Vermerke:

Nr.		Formal (einzeln) Nein=0 Ja=1	Vermerk
-----	--	---	---------

1.	Ausgangslage und Formelles		
-----------	-----------------------------------	--	--

1.1	Abwicklung des Gutachtensauftrages/ Formelles		
1.1.1	Angaben zum Auftraggeber		
1.1.2	Angaben zum Exploranden:		
	a) Name		
	b) Vorname		
	c) Geburtsdatum		
	d) Adresse		
	e) Versicherungsnummer		
1.1.3	Angaben zum Gutachter:		
	a) Adresse		
	b) Facharztstitel		
1.1.4	<i>(Gutachtendaten)</i>		
	a) Auftragsdatum		
	b) Eingangsdatum		
	c) Explorationsdatum		
	d) Datum der Gutachtenfertigstellung		

1.2	Anlass der Begutachtung		
	a) Wer ist Auftraggeber?		
	b) Wann wurde Gutachten in Auftrag gegeben?		
	c) Zu welchem Zweck wurde Gutachten in Auftrag gegeben?		

1.3	Grundlagen für das Gutachten		
------------	-------------------------------------	--	--

1.3.1	Auflistung der vom Auftraggeber zur Verfügung gestellten Aktendossiers (z.B. SUVA-Akte, IV-Akte)		
1.3.2	<i>(Aktenstücke)</i>		
	a) Summarische Auflistung der vom Gutachter <u>zusätzlich</u> beigebrachten (o. anderweitig bei ihm eingegangenen) Aktenstücke respektive telef. Nachfragen z. B. beim Hausarzt		fakultativ
	b) Aktenstücke mit dem GA dem Auftragg. in Kopie zuge- stellt?		
1.3.3	Eigene Untersuchungen und Befunde		
	a) Angaben zur Untersuchungsdauer (ggf. je Sitzung bei mehreren Terminen)		
	b) etc. (z.B. bzgl. einer evtl. Zuhilfenahme eines Dolmet- schers in welcher Sprache)		
1.3.4	Zusätzl. durchgeführte Untersuchungen /Zusatz-GA / Be- richterstattung		
	a) Untersucher/Gutachter		fakultativ
	b) Titel		
	c) Datum von Untersuchung		
	d) ggf. Beurteilung apparative Untersuchungen, Laborunter- suchungen etc.		
	e) sind die gestellten Fragen ersichtlich?		
1.3.5	Angaben von Drittpersonen		fakultativ

2.	Aktenauszug		
-----------	--------------------	--	--

	a) Angabe des Autors		
	b) Datum des Berichts		
	c) (relevante) Diagnosen / Befunde		
	d) <i>Hinweise für Bewertungen und Kommentare</i>		fakultativ

3.	Untersuchung		
3.1	Spontane Angaben des Expl. / offenes Interview		
3.2	Vertiefendes Interview - je nach Fragestellung -		
3.2.1	Jetziges Leiden:		
	a) Aktuelle Beschwerden		
	b) Entstehung/Verlauf der psychischen Beeinträchtigung		
	c) Behandlungsmassnahmen		
	d) Eigene Krankheitskonzepte		
	e) Umgang m. Beschwerden im Alltag?		
	f) Zukunftsvorstellungen bzgl. Krankheitsverlauf		
3.2.2	<i>(Anamnesen)</i>		
	a) psychiatrische Anamnese		
	b) somatische Anamnese		
	c) Konsum psychotroper Substanzen		
	d) Sexualität		
3.2.3	Familienanamnese / Heredität		
3.2.4	Soziale Anamnese:		
	a) Wohnung		
	b) finanzielle Verhältnisse / Schulden		
	c) Partnerschaft		
	d) Kinder		
	e) Beziehungsnetz		
3.2.5	<i>(Werdegang)</i>		
	a) schulischer Werdegang		
	b) beruflicher Werdegang		
3.2.6	<i>(Berufliche Tätigkeiten und Arbeitsversuche)</i>		
	a) Berufliche Tätigkeiten und Stellungen im Beruf		
	b) Datum und Umstände der Arbeitsaufgabe		
	c) Erfahrungen mit Eingliederungsmassn.		fakultativ
3.2.7	<i>(Alltagsleben)</i>		
	a) Tagesablauf		
	b) Freizeitgestaltung und Hobbies		
	c) benötigte Hilfen in Haushalt und Alltag		
3.2.8	<i>(Behandlungen)</i>		
	a) Aktuelle Behandlung		
	b) inklusive Medikamente		
3.2.9	einschneidende Erlebnisse		fakultativ
3.2.0	<i>(Zukunft)</i>		
	a) Zukunftsvorstellungen allgemein		
	b) Zukunftsvorst. bzgl. berufl. Tätigkeiten/Eingliederung		
4.	Befund		
4.1.	Verhaltensbeobachtungen / äussere Erscheinung		
	a) Verhaltensbeobachtungen und Interaktionen		
	b) äussere Erscheinung		
4.2	Sprachliche Verständigung		
	a) Angaben zur Muttersprache des Expl.		
	b) Angaben z. aktiven/passiven Beherrschen Sprache		fakultativ
	c) Angaben zum Dolmetscher		

Nr.		Formal (einzeln) Nein=0 Ja=1	Vermerk
4.3	Objektive Untersuchungsbefunde		
	Allfällig unsichere Untersuchungsergebnisse:		
	a) Angaben zu Verifizierung		fakultativ
	b) Deklaration als nicht genau erhebbare Befunde		
4.3.1	Psychiatrischer Befund		
	a) klinisch-psychiatrischer Befund		
	b) Beschreibung der Persönlichkeit		
4.4	Zusatzbefunde		
4.4.1	Somatischer Befund		fakultativ
4.4.2	Testpsychologische Zusatzuntersuchungen		fakultativ
4.4.3	Apparative und Labor-Untersuchungen (Hinweise zur Indikation der jeweiligen Untersuchung, z.B. bei Medikamentenspiegel)		fakultativ
5.	<u>Angaben von Drittpersonen</u>		
			fakultativ
6.	<u>Medizin. u. vers.med. Beurteilung</u>		
6.1	Resumé		
	a) bisherige persönliche Entwicklung		
	b) gesundheitliche Entwicklung		
	c) aktuelle psychische und gesundheitliche Situation		
6.2	Beurteilung bzw. Einordnung der Befunde		
	Bei polydisziplinären Gutachten:		
	a) Hinweis auf Synthese aller Expertenberichte		
	b) Erörterung einzelner Einbussen (Addierung von Funktionseinschränkungen oder ob eine Schonung aufgrund der einen Störung gleichzeitig den nötigen Schonraum für eine andere Störung bietet)		fakultativ
	c) Die psych. Beurteilung soll die psychiatr. Situation separat ausführlich darstellen		
	d) Allfällige Differenzen in den einzelnen Berichten sollen von den Experten diskutiert und geklärt werden; noch verbleibende Differenzen sollen im polydisziplin. Gutachten bewertet werden.		
6.3	Diskussion und versicherungsmedizinische Würdigung		
	a) Herleitung und Darstellung der Funktions- und Fähigkeitsstörungen		
	b) vorhandene Ressourcen		
	c) mit kritischer Würdigung der Selbsteinschätzung des Expl.		
	d) Kooperationsbereitschaft des Exploranden		
	e) Integration aller vorhand. früherer fachlicher Einschätzungen sowie anderer aktenkundiger Informationen		
6.4	Beantwortung der Fragen		
7.	<u>Beilagen</u>		
			fakultativ

Anhang 2: Kodierregeln

Nr.	<u>Formale Kriterien</u>	Formale Vollständigkeit: (einzeln) Nein=0 Ja=1	Formale Vollständigkeit: (gesamt) 0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	Kodierhilfen (Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)
-----	--------------------------	--	---	--

1.	<u>Ausgangslage und Formelles</u>
-----------	--

1.1	Abwicklung des Gutachtensauftrages/ Formelles			
1.1.1	Angaben zum Auftraggeber	0=keine Angaben 1=Angaben vorh.		Adresse, Anschrift
1.1.2	Angaben zum Exploranden a) Name b) Vorname c) Geburtsdatum d) Adresse e) Versicherungsnummer		0= a)-e) n.v. 1= a)-e) mind. 1x n.v. 2= a)-e) vorhanden	
1.1.3	Angaben zum Gutachter a) Adresse b) Facharztstitel		0= a) und b) n.v. 1= a) oder b) v. 2= a) und b) v.	Name etc. des Gutachters befindet sich auch hinten im GA
1.1.4	<i>(Gutachtendaten)*¹</i> a) Auftragsdatum b) Eingangsdatum c) Explorationsdatum		0= a)-d) n.v. 1= a)-d) mind. 1x n.v. 2= a)-d) vorhanden	Zu b): siehe Stempel erste Seite und im Text

*¹dies wurde den offiziellen formalen Kriterien hinzugefügt

*² fakultatives Formales Kriterium

Nr.	Formale Kriterien	Formale Vollständigkeit: (einzeln) Nein=0 Ja=1	Formale Vollständigkeit: (gesamt) 0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	Kodierhilfen (Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)
-----	--------------------------	---	--	---

	d) Datum der Gutachtenfertigstellung			
--	--------------------------------------	--	--	--

1.2	Anlass der Begutachtung			
	a) Wer ist Auftraggeber?		0= a)-c) n.v.	Zu a) entspricht 1.1.1
	b) Wann wurde Gutachten in Auftrag gegeben?		1= a)-c) mind. 1x n.v.	Zu b) entspricht 1.1.4a)
	c) Zu welchem Zweck wurde Gutachten in Auftrag gegeben?		2= a)-c) vorhanden	Zu c) z.B. „zur psychiatrischen BeGA“ oder z.B. „Stellung zur AF aus psychiatr. Sicht“ (=„1“)

1.3	Grundlagen für das Gutachten			
-----	-------------------------------------	--	--	--

1.3.1	Auflistung der vom Auftraggeber zur Verfügung gestellten Akten dossiers (z.B. SUVA-Akte, IV-Akte)	0=keine Auflistung 1=Auflistung vorh.		Angaben z.B. „Ihre Akte...“ („1“=vorhanden)
-------	---	--	--	---

1.3.2*2	<i>(Aktenstücke)*1</i>			
	a) <i>Summarische Auflistung der vom Gutachter <u>zusätzlich</u> beigebrachten (o. anderweitig bei ihm eingegangenen) Aktenstücke respektive telef. Nachfragen z. B. beim Hausarzt</i>	0=keine Auflistung 1=Auflistung vorh.	Falls vorhanden: 0= a) und b) n.v. 1= a) oder b) v.	Fakultativ
	b) <i>Aktenstücke mit dem GA dem Auftrag. in Kopie zugestellt?</i>	0=nicht zugestellt 1=zugestellt	2= a) und b) v.	

1.3.3	Eigene Untersuchungen und Befunde			
-------	-----------------------------------	--	--	--

*1 dies wurde den offiziellen formalen Kriterien hinzugefügt

*2 fakultatives Formales Kriterium

Nr.	<u>Formale Kriterien</u>	Formale Vollständigkeit: (einzeln) Nein=0 Ja=1	Formale Vollständigkeit: (gesamt) 0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	Kodierhilfen (Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)
	a) Angaben zur Untersuchungsdauer (ggf. je Sitzung bei mehreren Terminen)	0=keine Angaben 1=Angaben vorh.	Falls a) und b) vorhanden: 0= a) und b) n.v.	
	b) etc. (z.B. bzgl. einer evtl. Zuhilfenahme eines Dolmetschers in welcher Sprache)	0=keine Angaben 1=Angaben vorh.	1= a) oder b) v. 2= a) und b) v.	Falls z.B. Dolmetscher benötigt wurde: Gibt es Angaben zu Sprache?
1.3.4* ²	<i>Zusätzl. durchgeführte Untersuchungen /Zusatz-GA / Berichterstattung</i>			
	a) <i>Untersucher/Gutachter</i>		0= a)-e) n.v.	Fakultativ
	b) <i>Titel</i>		1= a)-e) mind. 1x n.v.	
	c) <i>Datum von Untersuchung</i>		2= a)-e) v.	
	d) <i>ggf. Beurteilung apparative Untersuchungen, Laboruntersuchungen etc.</i>			
	e) <i>sind die gestellten Fragen ersichtlich?</i>			
1.3.5* ²	<i>Angaben von Drittpersonen</i>	0=keine Angaben 1=Angaben vorh.		Fakultativ „0“=Drittpersonen zwar vorhanden, aber nicht als Quelle angegeben

*¹dies wurde den offiziellen formalen Kriterien hinzugefügt

*² fakultatives Formales Kriterium

Nr.	Formale Kriterien	Formale Vollständigkeit: (einzeln)	Formale Vollständigkeit: (gesamt)	Kodierhilfen
		Nein=0 Ja=1	0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	(Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)

2.	Aktenauszug			
	a) Angabe des Autors		0= a)-c)/ d) n.v.	Zu c) Diagnose nach ICD-10 in Akte angegeben Zu d) Gibt es Hinweise für Bewertungen und Kommentare von anderen Personen zu Expl. ?
	b) Datum des Berichts		1= a)-d) mind. 1x n.v.	
	c) (relevante) Diagnosen / Befunde		2= a)-c)/ d) v.	
	d) Hinweise für Bewertungen und Kommentare* ²			
3.	Untersuchung			

3.1	Spontane Angaben des Expl. / offenes Interviews (= offener Teil; offener Gesprächsabschnitt)* ¹	0=keine Angaben 1=Angaben vorh.		Hinweis auf offene Fragen: „1“= z.B. „auf die allg. Frage...“, „äussert spontan“, „berichtet von sich aus“; „0“=Expl. antw. auf Fragen
-----	--	------------------------------------	--	--

3.2	Vertiefendes Interview - je nach Fragestellung -			
-----	---	--	--	--

3.2.1	Jetziges Leiden a) Aktuelle Beschwerden b) Entstehung/Verlauf der psychischen Beeinträchtigung c) Behandlungsmassnahmen d) Eigene Krankheitskonzepte (Kausalattribution, Ausmass u. Gefährlichkeit des Schadens, therap. u. progn. Vorstell. etc.) e) Umgang m. Beschwerden im Alltag? Bewältigungsstrategien? f) Zukunftsvorstellungen bzgl. Krankheitsverlauf		0= a)-f) n.v. 1= a)-f) mind. 1x n.v. 2= a)-f) v.	Zu b) z.B. „Psych. Beeinträchtigung besteht seit Kindheit.....“ minimal vorhanden=„1“ Zu c) z.B. (Miss)Erfolge / Unverträglichkeiten Zu e) z.B. „am besten zu Hause bleiben“ = „1“
-------	--	--	--	--

*¹dies wurde den offiziellen formalen Kriterien hinzugefügt

*² fakultatives Formales Kriterium

Nr.	<u>Formale Kriterien</u>	Formale Vollständigkeit: (einzeln) Nein=0 Ja=1	Formale Vollständigkeit: (gesamt) 0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	Kodierhilfen (Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)
3.2.2	<i>(Anamnesen)*¹</i> a) psychiatrische Anamnese b) somatische Anamnese c) Konsum psychotroper Substanzen d) Sexualität		0= a)-d n.v. 1= a)-d) mind. 1x n.v. 2= a)-d) vorhanden	Zu b) „1“ = auch bei Verweis auf z.B. Vor-GA bei polydisz. GA; bei monodisziplin. muss eigene Untersuchung verlangt werden
3.2.3	Familienanamnese / Heredität	0=keine Angaben 1=Angaben vorh.		Minim. Ang. zu Eltern/ Geschwistern z.B. ob gesund/krank o. Beruf o. Familien-Leben =„1“
3.2.4	Soziale Anamnese a) Wohnung b) finanzielle Verhältnisse / Schulden c) Partnerschaft d) Kinder e) Beziehungsnetz		0= a)-e) n.v. 1= a)-e) mind. 1x n.v. 2= a)-e) vorhanden	Zu e): Kann man sich ein Bild machen, wie Expl.Integriert ist? Existieren Angaben dazu? „0“=keine Infos darüber vorhanden „1“=Expl. ist einsam, hat nur Nachbarin-> Info über Beziehungsnetz vorhanden, ausser c) u. d)
3.2.5	<i>(Werdegang)*¹</i> a) schulischer Werdegang b) beruflicher Werdegang		0= a) und b) n.v. 1= a) oder b) v. 2= a) und b) v.	Tendenziell eher Vergangenheit
3.2.6	<i>(Berufliche Tätigkeiten und Arbeitsversuche)*¹</i> a) Berufliche Tätigkeiten und Stellungen im Beruf b) Datum und Umstände der Arbeitsaufgabe		0= a) und b)/ c) n.v. 1= a)-c) mind. 1x n.v.	Tendenziell aktuell

*¹dies wurde den offiziellen formalen Kriterien hinzugefügt

*² fakultatives Formales Kriterium

Nr.	<u>Formale Kriterien</u>	Formale Vollständigkeit: (einzeln) Nein=0 Ja=1	Formale Vollständigkeit: (gesamt) 0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	Kodierhilfen (Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)
	<i>c) Erfahrungen mit Eingliederungsmassn. / Arbeitsversuchen*2</i>		2= a) und b)/ c) v.	
3.2.7	<i>(Alltagsleben)*1</i> a) Tagesablauf b) Freizeitgestaltung und Hobbies c) benötigte Hilfen in Haushalt und Alltag		0= a)-c) n.v. 1= a)-c) mind. 1x n.v. 2= a)-c) vorhanden	Zu a) Minimalprinzip: z.B. es steht nur in Zfg.: „...keine Tages-Struktur“ (=„1“, vorhanden) Zu c) „1“= „macht Haushalt allein“
3.2.8	<i>(Behandlungen)*1</i> a) Aktuelle Behandlung b) inklusive Medikamente		0= a) und b) n.v. 1= a) oder b) v. 2= a) und b) v.	Zu a) z.B. regelmäss. Gespräche mit Psychiater
3.2.9*2	<i>einschneidende Erlebnisse; evtl. frühere Konflikte mit dem Gesetz (z. B. auch im Strassenverkehr)</i>	0=keine Angaben 1=Angaben vorh.		Fakultativ
3.2.10	<i>(Zukunft)*1</i> a) Zukunftsvorstellungen allgemein b) Zukunftsvorst. bzgl. berufl. Tätigkeiten bzw. Eingliederung		0= a) und b) n.v. 1= a) oder b) v. 2= a) und b) v.	Zu a) Minimalprinzip: z.B. „könne nicht an Zukunft denken“ (= „1“ /vorhanden)

*1 dies wurde den offiziellen formalen Kriterien hinzugefügt

*2 fakultatives Formales Kriterium

Nr.	Formale Kriterien	Formale Vollständigkeit: (einzeln)	Formale Vollständigkeit: (gesamt)	Kodierhilfen
		Nein=0 Ja=1	0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	(Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)

4.	<u>Befund</u>			
-----------	----------------------	--	--	--

4.1.	Verhaltensbeobachtungen und äussere Erscheinung		0= a) und b) n.v. 1= a) oder b) v. 2= a) und b) v.	Kann man sich ein Bild machen, wie Expl. aussieht, sich verhält? z.B. „schaut starr“, „trägt Haarspange“ (=„1“); z.B. „ordentlich gepflegt“
	a) Verhaltensbeobachtungen und Interaktionen			
	b) äussere Erscheinung			

4.2*²	Sprachliche Verständigung		0= a)-c) n.v. 1= a)-c) mind. 1x n.v. 2= a)-c) v.	Nur bei Expl. mit nicht Muttersprache Deutsch
	a) Angaben zur Muttersprache des Exploranden			
	b) Angaben z. aktiven/passiven Beherrschen der Landessprache			
	c) Angaben zum Dolmetscher			

4.3*²	Objektive Untersuchungsbefunde <i>Allfällig unsichere oder unklare Untersuchungsergebnisse:</i>		0= a) und b) n.v. 1= a) oder b) v. 2= a) und b) v.	Fakultativ
	a) Angaben zu Verifizierung durch mehrmalige, ev. zeitversetzte Untersuchungen			
	b) Deklaration als nicht genau erhebbare Befunde			
4.3.1	Psychiatrischer Befund		0= a) und b) n.v. 1= a) oder b) v. 2= a) und b) v.	Zu a) Minimalprinzip: AMDP wurde berücksichtigt; Orientierung a. AMDP erkennbar (=„1“)
	a) klinisch-psychiatrischer Befund (Definitionen gemäss AMDP)			
	b) Beschreibung der Persönlichkeit (z. B. nach „the big five“)			

4.4	Zusatzbefunde			Fakultativ
4.4.1* ²	Somatischer Befund	0=keine Angaben 1=Angaben vorh.		

*¹dies wurde den offiziellen formalen Kriterien hinzugefügt

*² fakultatives Formales Kriterium

Nr.	Formale Kriterien	Formale Vollständigkeit: (einzeln)	Formale Vollständigkeit: (gesamt)	Kodierhilfen
		Nein=0 Ja=1	0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	(Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)

4.4.2* ²	Testpsychologische Zusatzuntersuchungen	0=keine Angaben 1=Angaben vorh.		Es wurden testpsych. Untersuchungen durchgeführt, aber keine Angaben dazu gemacht (=„0“)
---------------------	---	------------------------------------	--	--

4.4.3* ²	Apparative und Labor-Untersuchungen (Hinweise zur Indikation der jeweiligen Untersuchung, z.B. bei Medikamentenspiegel)	0=keine Angaben 1=Angaben vorh.		Es wurden zwar app./Labor-Unters. durchgeführt, aber keine Angaben dazu gemacht(=„0“)
---------------------	--	------------------------------------	--	---

5. *²	Angaben von Drittpersonen			Fakultativ
		0=keine Angaben 1=Angaben vorh.		WAS wurde über Drittpersonen geschrieben? Inhaltl. Angaben zu Drittpersonen vorhanden?

6.	Medizin. u. versicherungsmediz. Beurteilung			
6.1	Resumé			
	a) bisherige persönliche Entwicklung		0= a)-c) n.v. 1= a)-c) mind. 1x n.v. 2= a)-c) vorhanden	
	b) gesundheitliche Entwicklung			
	c) aktuelle psychische und gesundheitliche Situation			

*¹dies wurde den offiziellen formalen Kriterien hinzugefügt

*² fakultatives Formales Kriterium

Nr.	Formale Kriterien	Formale Vollständigkeit: (einzeln) Nein=0 Ja=1	Formale Vollständigkeit: (gesamt) 0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	Kodierhilfen (Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)
-----	--------------------------	---	--	---

6.2 ^{*2}	Beurteilung bzw. Einordnung der Beschwerden und Befunde in Bezug auf aktuelle Diagnosesysteme und in Bezug auf frühere, ggf. anderslautende Beurteilungen			Fakultativ
	<i>Bei polydisziplinären Gutachten:</i>			
	a) Hinweis auf Synthese aller Expertenberichte		0= a)-d) n.v.	
	b) Erörterung einzelner Einbussen (Addierung von Funktionseinschränkungen oder ob eine Schonung aufgrund der einen Störung gleichzeitig den nötigen Schonraum für eine andere Störung bietet		1= a)-d) mind. 1x n.v. 2= a)-d) vorhanden	
	c) Die psych. Beurteilung soll die psychiatr. Situation separat ausführlich darstellen			
	d) Allfällige Differenzen in den einzelnen Berichten sollen von den Experten diskutiert und geklärt werden; noch verbleibende Differenzen sollen im polydisziplin. Gutachten bewertet werden.			
6.3	Diskussion und Versicherungsmedizinische Würdigung			
	a) Herleit. und Darst. der Funktions- und Fähigkeitsstörungen		0= a)-e) n.v.	Zu a) „0“=keine Herleitung/Begründung, max nur Behauptung; 1= wurde begründet (Diff. nach: *keine systematische Herleitung *Herleitung mittels Checkliste o.ä. und *Benutzung ICF-> dann Notiz bei „Vermerk“
	b) vorhandene Ressourcen		1= a)-e) mind. 1x n.v.	
	c) mit kritischer Würdigung der Selbsteinschätzung des Expl.		2= a)-e) vorhanden	
	d) Kooperationsbereitschaft des Exploranden			
	e) Integration aller vorhand. früherer fachlicher Einschätzungen sowie anderer aktenkundiger Informationen			

*¹dies wurde den offiziellen formalen Kriterien hinzugefügt

*² fakultatives Formales Kriterium

Nr.	<u>Formale Kriterien</u>	Formale Vollständigkeit: (einzeln) Nein=0 Ja=1	Formale Vollständigkeit: (gesamt) 0=nicht vorhanden (=n.v.) 1=teilweise vorhanden (=t.v.) 2=vorhanden (=v.) 9=nicht bewertbar (=n.b.)	Kodierhilfen (Insgesamt Minimalprinzip: „1“ bedeutet Hinweis vorhanden, unabhängig von Qualität und Umfang. Qualität und Umfang wird in einem zweiten Schritt unter subjektive „Bewertung“ mit „0“, „1“ oder „2“ bewertet)
-----	--------------------------	--	---	--

6.4	Beantwortung der Fragen			
			0= keine Fragen beantwortet 1= Fragen teilw. beantwortet 2= alle relevanten Fragen beantwortet	Mindestanforderungen für „2“=Angaben zu: AF-Angaben in h/% und Differenzierung nach bisherige Tätigkeit und adaptierte Tätigkeit; eigene Diagnose (ICD-10)und Auswirkungen auf die AF; Prognose/Verbesserung der AF?; Therap. Empfehlung; ggf. Zusatzfragen beantwortet?

7. * ²	<u>Beilagen</u>			Fakultativ
		0=keine Angaben 1=Angaben vorh.		z.B. Fachgutachten oder testpsychologische Zusatzgutachten

*¹dies wurde den offiziellen formalen Kriterien hinzugefügt

*² fakultatives Formales Kriterium

Anhang 3: Verwendete Literatur

- Baer, N., Frick, U., & Fasel, T. (2009). Dossieranalyse der Invalidisierungen aus psychischen Gründen. Forschungsbericht Nr. 6/09. Bern: Bundesamt für Sozialversicherungen BSV.
- Dittmann, V., Ebner, G., Herdt, J., Junge, C., & Träbert, S. (2009). Literaturstudie als Grundlage zur Entwicklung von evidenzbasierten Gütekriterien zur Beurteilung von psychischen Behinderungen. Bern: Bundesamt für Sozialversicherungen BSV.
- Knecht, T. (2011). Distorsionstrauma der Halswirbelsäule - Eine Standortbestimmung aus psychiatrischer Sicht. *Schweiz Med Forum*, 11(18), 314–318.
- Linden, M. & Baron, S. (2005). Das „Mini-ICF-Rating für psychische Störungen (Mini-ICF-P)“. Ein Kurzinstrument zur Beurteilung von Fähigkeitsstörungen bei psychischen Erkrankungen. *Rehabilitation*, 44(3), 144-151.
- Linden, M., Baron, S., Muschalla, B. (2009). Mini-ICF-Rating für psychische Störungen (Mini-ICF-APP). Ein Kurzinstrument zur Beurteilung von Fähigkeits- bzw. Kapazitätsstörungen bei psychischen Störungen. Göttingen: Hans Huber.
- Mosimann, H. J., & Ebner, G. (2008). „Objektiv nicht überwindbare“ Erwerbsunfähigkeit: Überlegungen zu Art. 7 Abs. 2 ATSG aus juristischer und psychiatrischer Sicht. *SZS(6/2008)*, 1-27.
- Rajower, I. (2007). 5. IV-Revision: Welche Rolle spielen die Ärzte? *Schweizerische Ärztezeitung*, 88(23), 1004-1006.
- Winckler P, Foerster K. (1996). Zum Problem der «zumutbaren Willenanspannung» in der sozialmedizinischen Begutachtung. *Med Sach.* 1996;92:120–4.
- Wirtz, M., & Caspar, F. (2002). Beurteilerübereinstimmung und Beurteilerreliabilität - Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen: Hogrefe.

**Weitere Forschungs- und Expertenberichte aus der Reihe
«Beiträge zur Sozialen Sicherheit»**

<http://www.bsv.admin.ch/praxis/forschung/publikationen/index.html?lang=de>

**Autres rapports de recherche et expertises de la série
«Aspects de la sécurité sociale»**

<http://www.bsv.admin.ch/praxis/forschung/publikationen/index.html?lang=fr>

**Altri rapporti di ricerca e perizie della collana
«Aspetti della sicurezza sociale»**

<http://www.bsv.admin.ch/praxis/forschung/publikationen/index.html?lang=it>

**Further research reports and expertises in the series
«Beiträge zur Sozialen Sicherheit»**

<http://www.bsv.admin.ch/praxis/forschung/publikationen/index.html?lang=en>