

Expertise zum Thema

Nutzen von Probandenbefragungen für die Qualitätssicherung in der (versicherungs-)medizinischen Begutachtung und speziell in Bezug auf die Fairness bzw. Zufriedenheit mit dem Begutachtungsablauf

im Auftrag der Eidg. Kommission für Qualitätssicherung in der medizinischen Begutachtung (EKQMB)

Autorenteam

Prof. Dr. Beate Muschalla¹

PD Dr. rer. nat. Felix Fischer²

Dipl.-Psych. Anne Meier-Credner¹

Prof. Dr. Michael Linden³

¹ Technische Universität Braunschweig, Psychotherapie und Diagnostik

² Charité Universitätsmedizin Berlin, Center for Patient-Centered Outcomes Research

³ Charité Universitätsmedizin Berlin, Forschungsgruppe Psychosomatische Rehabilitation

1. Auftrag

1.1. Auftraggeber

Am 26.10.22 hat Roman Schleifer, Leiter der Fachstelle der ausserparlamentarischen Eidgenössischen Kommission für Qualitätssicherung in der medizinischen Begutachtung (EKQMB) in der Schweiz, c/o Bundesamt für Sozialversicherungen (BSV), dem Autorenteam im Namen des EKQMB-Präsidenten Prof. Michael Liebrecht den Auftrag erteilt, eine Expertise zum Thema «Probandenbefragungen in sozialmedizinischen Begutachtungsprozessen» und speziell zum «*Nutzen von Probandenbefragungen in Bezug auf die Fairness bzw. Zufriedenheit mit dem Begutachtungsablauf*» zu erstellen.

1.2. Ausgangslage

Die ausserparlamentarische Eidgenössische Kommission für Qualitätssicherung in der medizinischen Begutachtung (EKQMB) hat den gesetzlichen Auftrag erhalten, Empfehlungen zu den Kriterien und Instrumenten für die Beurteilung der Qualität von Gutachten im Sozialversicherungsbereich zu erarbeiten (Art. 7p Abs. 1 Bst. d ATSV).

Die medizinische Abklärung im Rahmen des Verfahrens der Invalidenversicherung (IV) dient dazu, die notwendigen Informationen zur Beurteilung der medizinischen Voraussetzungen für den Anspruch auf IV-Leistungen zu ermitteln (KSVI, Stand 1.07.2022, Rz. 3048).

Gemäss nicht repräsentativer Auswertungen der Organisation Inclusion Handicap erlebten viele Probanden das Gesprächsklima als schlecht und die Begutachtenden als nicht

interessiert. Die Kritik an der Gesprächsführung führte bereits zu zahlreichen Vorstößen und Debatten im Schweizer Parlament und zum Beschluss der Einführung von Tonbandaufnahmen des gesamten Untersuchungsgesprächs (Art. 7k und 7l Abs. 2 ATSV).

Eine systematische Befragung der Versicherten zur Zufriedenheit mit dem Untersuchungsgespräch ist vom Gesetzgeber nicht vorgesehen, kann aber von der Kommission empfohlen werden, wenn der Nutzen einer solchen Befragung für die Beurteilung der Qualität von Gutachten begründet ist.

1.3. Zielsetzung und Forschungsfragen

Die Kommission hat deshalb beschlossen, zur Klärung der Frage des Nutzens der Feedbackerhebungen von Probandinnen und Probanden (Versicherten) eine literaturbasierte Studie in Auftrag zu geben. Die Studie sollte den wissenschaftlichen Stand einer solchen Methode aufzeigen und spezifizieren, unter welchen Bedingungen ein solches Instrument für die Qualitätsverbesserung in der medizinischen Begutachtung effektiv eingesetzt werden könnte.

Als Datenquellen sollten wissenschaftliche Publikationen im europäischen, angelsächsischen und insbesondere frankophonen Raum, inklusive Kanada, herangezogen werden, beispielsweise zu Stichworten wie «Erhebungsinstrument», «Patienten- bzw. Probandenzufriedenheit», «Nutzen», «Qualitätssicherung» und «medizinische Begutachtung».

Zusammenfassend beinhaltet der Auftrag eine literaturbasierte Studie zur Frage des Nutzens einer systematischen Feedbackerhebung von Probanden (Versicherten) für die Beurteilung der Qualität von Gutachten und des Gutachtenprozesses. In Rücksprache mit Dr. Schleifer wurden folgende Fragen eingegrenzt und im Rahmen eines Scoping-Reviews bearbeitet:

(a) Inwieweit liegen empirische Daten zur Art und Nutzung von Probandenselbstberichten während und nach Begutachtungen vor?

(b) Welchen Nutzen und welche Grenzen haben Probandenselbstauskünfte im Kontext sozialmedizinischer Begutachtungen?

(c) Unter welchen Bedingungen und mit welcher Zielsetzung sind Probandenselbstauskünfte für die Qualitätssicherung in der medizinischen Begutachtung hilfreich?

2. Literaturrecherche und Methode der Gutachtenerstellung

Das Autorenteam hat zur Literaturrecherche internationale Datenbanken (u. a. Google Scholar, PubMed, Medline) mittels der Begriffe «quality», «(patient) self-rating», «assessment», «instrument», «questionnaire», «PROM» bzw. «Patient-Reported Outcome Measure», «PRO» bzw. «Patient-Reported Outcome», «PREM» bzw. «Patient-Reported Experience Measure», «(patient) satisfaction, benefit», «social medicine», «expert opinion», «guidelines» bzw. der deutschsprachigen Termini «medizinische Begutachtung», «Qualität», «Patienten/Probandenselbstbericht/-selbstauskunft», «Patienten/Probandenzufriedenheit», «Nutzen» oder der französischen Begriffe «expertise médicale», «qualité», «rapport/auto rapport (du patient)», «satisfaction (du patient)», «utilité» in diversen Kombinationen durchsucht.

Zudem wurden die Literaturverzeichnisse relevanter identifizierter Publikationen auf weiterführende Quellen gesichtet. Relevante rein französischsprachige Arbeiten wurden nicht gefunden.

Zur Beantwortung der Fragestellung wurde auf allgemeingültige Grundsätze psychometrischer Qualitätsaspekte zur Validität von Untersuchungsinstrumenten und Dateninterpretation (z. B. Schmidt-Atzert & Amelang, 2018), sowie fachlicher Standards zur medizinischen Gutachtenerstellung (AWMF, 2019; SGVP, 2016) Bezug genommen.

Es wurden Qualitätssicherungsverfahren gesichtet und berücksichtigt, die in sozialmedizinischen, rehabilitativen und gutachterlichen Kontexten eingesetzt werden.

3. Theoretischer Hintergrund

3.1. Beurteilungsebenen der Gutachtenqualität

Um Aussagen darüber zu treffen, ob sich bestimmte Daten zur Bewertung von Gutachtenqualität eignen, soll zunächst definiert werden, was im konkreten Fall unter «Gutachtenqualität» zu verstehen ist (Schmidt-Atzert & Amelang, 2018).

3.1.1. Fachliche Korrektheit

Gutachtenqualität bezeichnet zunächst einmal die fachliche Korrektheit des Gutachtens. Medizinische Gutachten müssen medizinische und sozialrechtliche Fragestellungen (z. B. zur Arbeitsfähigkeit) mit passenden Methoden und einer Tatbestandserhebung und -beurteilung nach den Regeln des aktuellen Fachstandards (z. B. AWMF, 2019; SGVP, 2016) beantworten. Gutachten basieren auf Fakten, die im Rahmen von Probandenuntersuchungen (z. B. körperliche Untersuchung, Exploration, spezifische Testungen und Messungen mit fachlich anerkannten validierten Instrumenten, ggf. Selbstauskunftfragebögen) erhoben wurden, einschliesslich der Berücksichtigung von Vorbefunden und Vorgutachten. Die Darstellung der Befunde und Schlussfolgerungen im Gutachten muss konsistent und nachvollziehbar sein.

Eine Beurteilung, ob eine medizinische Begutachtung fachgerecht ist, ist nur im Abgleich mit Leitlinien zur medizinischen Begutachtung und Überprüfung der Fakten und Schlussfolgerungen möglich. Dies kann ausschliesslich im Rahmen einer «Zweit- oder Oberbegutachtung» durch andere Fachpersonen erfolgen, jedoch nicht durch die begutachtete Person.

3.1.2. Probandenzufriedenheit mit dem Gutachtenergebnis

Die Befragung von begutachteten Personen zur Zufriedenheit mit einem Gutachten kann zwei Aspekte umfassen:

- Zufriedenheit mit dem Gutachtenergebnis (Ergebniszufriedenheit)
- Zufriedenheit mit dem zugrundeliegenden Untersuchungsprozess (Prozesszufriedenheit).

Die Beurteilung des Gutachtenergebnisses durch eine begutachtete Person spiegelt zunächst einmal wider, ob das Ergebnis der Begutachtung mit den Erwartungen der begutachteten Person übereinstimmt. Es handelt sich also um eine subjektive Bewertung des Gutachtenergebnisses. Dies ist von Relevanz, wenn eine begutachtete Person Beschwerde gegen die Entscheidung der Sozialversicherung, die auf der Grundlage des erstellten Gutachtens getroffen wurden, einreichen will.

Dazu gehört jedoch auch die Beurteilung des Gutachtenprozesses durch die Probandin oder den Probanden, wenn sie oder er geltend machen will, dass die Gutachterin oder der Gutachter entscheidungsrelevante Fakten nicht erhoben oder nicht zur Kenntnis genommen hat (z. B. angeblich nicht zugehört hat).

Die Entscheidung, ob der Beschwerde der Probandin oder des Probanden entsprochen wird und ihre bzw. seine Einschätzung getragen wird oder nicht, obliegt Sozialversicherungen und Gerichten. Sie entscheiden damit primär über die fachliche Gutachtenqualität. Die Gutachtenqualität im Ergebnis wie im Erhebungsprozesse wird in diesem Zusammenhang regelmässig von der Rechtsvertretung der Probandinnen und Probanden kritisch kommentiert, mit richtigen oder falschen Schlussfolgerungen. Letztlich ist es Aufgabe der Gerichte zu entscheiden, ob sie diese Auffassungen teilen oder nicht.

3.1.3. Probandenzufriedenheit mit dem Gutachtenprozess

Die Probandenzufriedenheit mit dem Untersuchungsprozess, der Gesprächsführung und der Gesprächsatmosphäre spielt für die Beurteilung der Gutachtenqualität primär keine Rolle. Das Fairnesserleben oder die Zufriedenheit einer Probandin oder eines Probanden mit dem Untersuchungsablauf einerseits und die Durchführung und das Ergebnis der Begutachtung andererseits sind im Prinzip voneinander unabhängig, da es sich um verschiedene Gegenstände handelt.

Faktisch und in der Realität sind sie jedoch eng verwoben, da Probandinnen und Probanden sich dann ggf. über den Gutachtenprozess beschweren werden, wenn das Ergebnis nicht ihren Vorstellungen entspricht, während sie womöglich sogar eine belastende Gutachtenprozedur gut finden werden, wenn das dadurch gewonnen Ergebnis in ihrem Sinne gut war.

Die Probandenbewertung des Begutachtungsablaufs und der Gesprächsatmosphäre kann jedoch Relevanz für die Bewertung der Gutachtengüte bekommen, wenn die untersuchte Person geltend macht, dass wichtige Aspekte seiner Problemschilderung nicht zur Kenntnis genommen wurden. Das Argument ist dann, dass die Gutachtengüte angefochten wird, weil die Faktenerhebung unvollständig war wegen eines nicht adäquaten Begutachtungsprozesses. Solche Beschwerden sind seitens der Behörde als Anlass zu nehmen, das betreffende Gutachten zu prüfen mit der Frage, ob die von der Probandin oder vom Probanden geltend gemachten Faktenlücken tatsächlich gegeben sind.

Unabhängig von der Frage der Vollständigkeit der Faktenerhebung, ist die Probandenzufriedenheit mit dem Gutachtenprozess auch insofern von Relevanz, als ein subjektiv positives Erleben der Begutachtungssituation und insbesondere das subjektive Gefühl von Fairness für die Akzeptanz des Gutachtenergebnisses und damit den Rechtsfrieden von Bedeutung sein kann (Echterhoff et al., 2010; Rosburg, 2021).

Grundsätzlich gilt daher, dass ein Probandenzufriedenheitsrating kein notwendiger Bestandteil einer medizinischen Begutachtung ist. Fraglich ist jedoch, ob die Gutachterin oder der Gutachter im Gutachten Aussagen zur Interaktion zwischen sachverständiger und begutachteter Person machen sollte, und inwieweit Probandenrückmeldungen zur subjektiven Prozesszufriedenheit ein Kontrollinstrument für den Auftraggeber sein können, um Gutachten wie Sachverständige einer Überprüfung zu unterziehen.

Im Folgenden werden Kriterien, die an subjektive Bewertungen zu stellen sind, wenn sie sinnvolle und verlässliche Informationen liefern sollen, sowie die Einflussfaktoren auf Probandenbeurteilungen, die das Ergebnis beeinflussen, erläutert.

3.2. Reliabilität der Probandeneinschätzungen

Reliabilität bedeutet, dass zwei Beobachtungen dasselbe Ergebnis bringen. Die Reliabilität kann zwischen 0 (keinerlei Übereinstimmung) bis 1 (völlige Übereinstimmung) schwanken (Schmidt-Atzert & Amelang, 2018).

Die Anforderungen an die Höhe der Reliabilität hängen vom Nutzungszweck der Beobachtung ab. Zu unterscheiden ist dabei insbesondere zwischen Beobachtungen zum Zwecke der Berichterstattung einerseits (Public Reporting darüber, wie das Gutachtersystem im Allgemeinen nach Urteil bestimmter Personen oder unter Anlegung bestimmter Instrumente funktioniert) und Beobachtungen zur Bewertung im Einzelfall oder zur Prozesssteuerung andererseits (Beurteilung der fachlichen Qualität einer Gutachterin oder eines Gutachters oder Steuerung von Entscheidungsprozessen).

Bei gruppenstatistischen Felderhebungen oder orientierenden Probandenuntersuchungen werden Reliabilitätswerte zwischen 0,5 und 0,7 als hinreichend angesehen. Bei Beobachtungen zur Steuerung von Systemen und Prozessen müssen die Reliabilitätswerte nahe 1 sein.

Mit Blick auf die Reliabilitätsanforderungen sind also der Zweck und die Konsequenzen der Probandeneinschätzungen entscheidend und beispielsweise seitens der EKQMB zu erläutern und vorzugeben. Falls die EKQMB auf politischer Ebene zeigen will, dass die in Auftrag von Sozialversicherungen tätigen Sachverständigen ihre Arbeit mit Achtung vor den Probandinnen und Probanden ausüben, dann sind geringere Reliabilitätsanforderungen hinreichend. Falls die Befragung darüber entscheiden soll, ob eine Gutachterin oder ein Gutachter weiterhin Aufträge erhalten soll, oder ob das Gutachtenergebnis seitens des Auftraggebers und der EKQMB zu überprüfen ist, dann muss der Reliabilitätswert nahe 1 liegen. Das Autorenteam der Expertise verfügt über keine Informationen zu den konkreten Erhebungszielen.

3.3. Validität von Probandeneinschätzungen

Die Reliabilität allein ist kein hinreichender Faktor zur Bewertung einer Probandeneinschätzung. Mindestens ebenso wichtig ist die Validität (Schmidt-Atzert & Amelang, 2018), d. h. die Gültigkeit dessen, was gemessen wird (z. B. Probandenzufriedenheit mit dem Ergebnis oder dem Gutachtenprozess oder der eigenen Befindlichkeit).

Die Validität bedeutet, dass die Beobachtungskategorien tatsächlich das zu beobachtende Objekt erfassen (z. B. Zufriedenheit mit dem Gutachtenergebnis, subjektives Erleben der Interaktionsatmosphäre oder Abbildung des psychischen Zustands der Probandin oder des Probanden).

Wenn Beobachtungen dazu dienen sollen, Systeme zu bewerten und zu steuern, die Auswirkungen auf das Wohl von Menschen haben (z. B. medizinische Begutachtung der Arbeits- und Leistungsfähigkeit einer bestimmten Probandin oder eines bestimmten Probanden), dann sind auch an die Validität hohe Anforderungen zu stellen:

- Vorab ist festzulegen, wer was wissen will.
- Die Daten und Beurteilungen müssen eine inhaltliche Validität im Hinblick auf eine genau definierte Beobachtungsgrösse haben.
- Die Daten müssen zeitnah zur Verfügung stehen.
- Aus den Daten müssen konkrete Massnahmen ableitbar sein, und die ableitbaren Massnahmen müssen effektiv sein, d. h. zu einer Veränderung führen.

Wenn es um gruppenstatistische Auswertungen und globale Bewertungen geht (z. B. wie zufrieden die Probandinnen und Probanden im Allgemeinen mit ihrer Begutachtung sind, sei es inhaltlich oder auf den Prozess bezogen?), dann gilt:

- Vorab ist festzulegen, wer was wissen will.
- Die Daten und Beurteilungen müssen insofern eine inhaltliche Validität aufweisen, dass im Prinzip substantiierte Aussagen erlauben (wenn z. B. nach der Güte des Gutachtenprozesses gefragt wird, dann sollte die Selbstbewertung kein Bericht über den aktuellen Krankheitszustand der Probandin oder des Probanden sein)
- Die Daten müssen nicht zeitnah zur Verfügung stehen, sondern erst dann, wenn ein Bericht erstellt werden soll (z. B. einmal im Jahr).
- Auf der Basis der Daten kann nur eine Diskussion angestoßen und ggf. ein weiterer Prüfprozess eingeleitet werden, jedoch keine konkreten Änderungen begründet werden.

3.4. Motivationale und befindlichkeitsabhängige Wahrnehmungs- und Erinnerungsverfälschung

Probandeneinschätzungen geben das Erleben und die Wahrnehmung der begutachteten Personen wieder. Ein Problem sind Wahrnehmungs- und Erinnerungsverfälschungen. Es gibt eine Reihe von psychologischen Prozessen, die diesbezüglich von Bedeutung sind. Dazu gehören z. B. Fehlerinnerungen, Verdeutlichungstendenzen, Aggravation, Simulation, Dissimulation, Motivation, Suggestion, Wunschvorstellungen u. a (z. B. Kieser, 2017; Merten, 2020; Muschalla & Schönborn, 2021; Caussé-Versaveau et al. 2018). Diese Phänomene werden beispielsweise in der forensischen Psychologie unter dem Stichwort der Glaubwürdigkeitsbeurteilung bzw. in der Aussagepsychologie untersucht (z. B. Toolin et al., 2022; Pfundmair, 2020). Wenn beispielsweise jemand zur eigenen Arbeitsfähigkeit befragt wird, dann wird sie oder er die aus ihrer oder seiner Sicht wichtigen und dem eigenen Standpunkt und der eigenen Motivation förderlichen Aspekte aggravierend verdeutlichen und die widersprechenden Aspekte herabspielen.

Ein anderer wichtiger Aspekt im Kontext von sozialrechtlichen Begutachtungen ist das psychologische Prinzip des «state dependent memory» und «state dependent reasoning». Erinnerungen und Wahrnehmungen von Menschen wie auch Weltbewertungen und Argumentationslinien hängen unmittelbar vom aktuellen Gemütszustand einer Person ab (Blanchette & Richards 2010; Luo & Yu 2015). Wer in ausgeglichener Stimmung ist, wird denselben Sachverhalt anders beurteilen, als wenn er oder sie in schlechter Stimmungslage ist. Probandinnen und Probanden in Begutachtungssituationen dürften häufig in beeinträchtigtem psychischem Zustand sein, so dass auch das Erleben und die Beurteilung einer Begutachtung in vielen Fällen einem Negativbias unterliegen kann.

3.5. Die Besonderheit der Interaktion zwischen Gutachter und Proband

Probandeneinschätzungen zur Zufriedenheit mit dem Gutachtenprozess stellen eine Beurteilung der Interaktion zwischen begutachteter und sachverständiger Person dar. Diese Art der zwischenmenschlichen Beziehung hat eine Reihe von Besonderheiten, die naturgegeben auch zu Irritationen bei der begutachteten Person führen und die Probandeneinschätzung beeinflussen können.

Die Probandin oder der Proband kommt vorbelastet in die Gutachtensituation. Sie oder er leidet und hat aber insbesondere auch Erwartungen an das Ergebnis des Gutachtens, was zu Ängsten oder auch Enttäuschungen führen kann.

Die Sachverständigen sind keine Therapeutinnen oder Therapeuten und nicht «auf der Seite» der begutachteten Person, sondern werden im Auftrag von Sozialversicherungen, Gerichten oder Dritter tätig. Darüber sollte die begutachtete Person explizit informiert werden. Es

besteht keine Vertragsbeziehung zwischen sachverständiger und begutachteter Person. Die Gutachterin oder der Gutachter ist zur Neutralität verpflichtet und kann sich sogar dem Vorwurf mangelnder Distanz aussetzen, wenn sie oder er sich der begutachteten Person nach Therapeutenart sehr empathisch zuwendet und alles «versteht».

Sachverständige müssen Fragen beantworten, die der Auftraggeber stellt, und dürfen keine darüber hinausgehenden Problemschilderungen der Probandin oder des Probanden aufnehmen, oder gar kommentieren. Sachverständige haben alles, was gutachtenrelevant ist und sie in der Gutachtensituation erfahren, festzuhalten. Es gibt in diesem Sinne keine Vertraulichkeit oder Schweigepflicht. Sachverständige sind an gesetzliche Vorgaben gebunden, d. h. sie müssen beispielsweise die «Arbeitsunfähigkeit» beurteilen, was etwas anderes ist, als ein Urteil über die Beschwerden der Probandin oder des Probanden abzugeben.

Zudem sind die Sachverständigen gehalten, eine Konsistenzprüfung der Probandenaussagen vorzunehmen. Sie dürfen sich nicht ausschliesslich auf die subjektiven Angaben der Probandin oder des Probanden verlassen, sondern müssen diese mit anderen verfügbaren Informationen, einschliesslich der Verhaltensbeobachtung in der Gutachtensituation abgleichen. Die Sachverständigen sind aus Probandensicht sogar potenziell «gefährlich», da sie zu Feststellungen kommen können, die den Interessen der Probandinnen und Probanden widersprechen. Von daher stellt sich die Frage, was in einer Probandeneinschätzung des Gutachtenprozesses inhaltlich abgebildet wird oder abgebildet werden kann.

Für die Probandinnen und Probanden selbst sind solche juristischen und fachlichen Differenzierungen in der Regel nicht nachvollziehbar (z. B. «Wie kann man sagen, dass ich arbeitsfähig bin, wenn ich doch krank bin?»). Daher sind Missverständnisse bezüglich der Begutachtungssituation und des Gutachterverhaltens naheliegend.

3.6. Reaktivität und Nebenwirkungen

Bei der Diskussion um die Zweckmässigkeit von Probandeneinschätzungen müssen auch Negativwirkungen bedacht werden. Beurteilungen, die in systemischen Strukturen abgegeben werden, sind nicht inert, sondern haben Rückwirkungen sowohl auf die beurteilende als auch auf die beurteilte Person (Muschalla, 2018). Das wird als «Reaktivität» bezeichnet. Darunter versteht man den Grad der Veränderung, den eine Beobachtung am Beobachtungsobjekt bewirkt.

Wenn Probandinnen und Probanden Zufriedenheitsratings zum Gutachtenprozess abgeben, dann kann es sogar Zweck der Erhebung sein, dass sich das Verhalten der Gutachterin oder des Gutachters daraufhin ändert, dass sie oder er also z. B. zugewandter und aufmerksamer wird. Dies wäre eine positive Reaktivität. Es kann aber auch zu negativen Folgen kommen, wenn die Gutachterin oder der Gutachter beispielsweise meint, Zufriedenheitsurteile durch Gefälligkeitsgutachten erkaufen zu können.

Ein Probandenurteil kann auch Rückwirkungen auf die Probandinnen und den Probanden selbst haben. Die Frage nach der aktuellen Befindlichkeit kann eine verstärkte Selbstbeobachtung auslösen, die Frage nach der Gutachtenqualität die Frage, ob man nicht Grund zur Beschwerde hat. Komplexe Sachverhalte können dichotomisiert werden, indem es nur gut oder schlecht gibt. Einmal eingenommene Positionen sind nur noch bedingt revidierbar. Es kann zu internen Dialogen kommen, in denen neue Argumente für das eigene Urteil gesucht und entwickelt werden. Dies geschieht insbesondere, wenn das eigene Urteil anschliessend von aussen hinterfragt wird. Es kann zu «pseudologischen» Argumentationsketten kommen, da einmal eingenommene Positionen mit immer neuen

Argumenten verteidigt werden, die nur bedingt die Realität wiedergeben. Dies könnte dann den Rechtsfrieden gefährden.

3.7. Emotionale versus verhaltensorientierte Befragungen

Messtheoretisch werden zwei Formen von Befragungen unterschieden, die emotionsorientierte und die verhaltens- und ereignisorientierte Befragung (Thierbach & Petschick 2019; Renecke 2022)

Die emotionsfokussierte qualitative Befragung nutzt Items, in denen Globalurteile zur Zufriedenheit und zum Erleben erfragt werden (z. B. «Der Gutachter hat mich ernst genommen»).

Verhaltensbeschreibende Items fragen nach konkreten Ereignissen und Abläufen (z. B. «Der Gutachter informierte mich»). Allerdings enthalten verhaltensbeschreibende Items stets auch emotionale Anteile und Bewertungen.

Diese Item-Formate sind für unterschiedliche Zwecke geeignet und je nach Ziel auszuwählen. Eine Liste an qualitativen Urteilen ist zumeist einfacher zu erstellen, als verhaltensbeschreibende Items.

Grundsätzlich ist anzumerken, dass die Ergebnisse von Befragungen durch die Auswahl und Formulierung der Fragen und die Vorgabe der Antwortalternativen beeinflusst werden können. Es ist möglich, Befragungen so zu konzipieren, dass hohe oder niedrige Zufriedenheitswerte das Ergebnis sind. Methodisch sind auf diesem Hintergrund verhaltensnahe und punktorientierte Items besser als qualitative und emotionale Globalratings.

4. Einsatzmöglichkeiten von Probandenbefragungen

4.1. Probandenbefragungen während der Begutachtung

4.1.1. Zustandserfassung

Während einer Begutachtung können zur Abbildung des eigenen Gesundheitszustands aus Sicht der Probandin oder des Probanden Selbstauskunftsinstrumente zum Einsatz kommen. Dies wird in vielen Gutachtenleitlinien empfohlen (Schmidt-Atzert & Amelang, 2018). Zumeist werden Selbstratingskalen zum eigenen Gesundheitszustand eingesetzt, wie bspw. Symptomskalen zur Erfassung des subjektiv von der Probandin oder vom Probanden erlebten Beschwerdeniveaus. Es können auch spezifische Leistungstests z. B. bei neuropsychologischen Fragestellungen verwendet werden. Schliesslich gibt es auch «Lügentests», um Aussageverfälschungen der Probandin oder des Probanden zu identifizieren.

Für alle Tests gelten immer die vorgenannten Anforderungen und Aussageeinschränkungen. Sie dürfen daher nicht als Sachtatbestandsfeststellung genommen werden (z. B. ein hoher Wert auf einer Depressionsskala beweist keine depressive Störung). Es bedarf stets einer Einordnung und Interpretation durch eine Person, die mit der Testtheorie und den Auswertungsvorschriften des konkreten Instruments und unter Bezugnahme auf die sonstigen Explorations- und Untersuchungsbefunde vertraut ist. Interpretationen und klinische Schlussfolgerungen sind nur auf der Basis einer sachgerechten Integration der Ratingergebnisse in die Gesamtsituation möglich und entsprechend zu begründen (AWMF, 2019; Schmidt-Atzert & Amelang, 2018).

4.1.2. Interaktionsbeurteilung

Beurteilungen des Interaktionsverhaltens in der Gutachtensituation oder unmittelbar im Anschluss unterliegen den gleichen methodischen Einschränkungen. Hinzu kommt das Problem der Vertraulichkeit. Wenn die Probandin oder der Proband davon ausgehen muss, dass die Gutachterin oder der Gutachter Einsicht in die Rückmeldung nehmen kann und dies möglicherweise dann auch Einfluss auf die gutachterliche Stellungnahme haben könnte, dann kann dies zu erheblichen Ratingverfälschungen im positiven wie negativen Sinne führen. Es ist auch nicht zu vermeiden, dass die Probandenäußerung Einfluss auf die gutachterliche Stellungnahme hat, da die Gutachterin oder der Gutachter zu diesem Zeitpunkt das Gutachten noch nicht abschliessend verfasst hat.

Eine messtechnische Methode, die insbesondere unter den Bedingungen von Ergebniserwartungen der Probandinnen und Probanden unbedenklich ist, ist dem Autorenteam nicht bekannt.

4.2. Probandenbefragungen nach der Begutachtung zu Händen Dritter

Probandenbefragungen können unmittelbar nach der Sitzung ohne Kenntnisnahme der Gutachterin oder des Gutachters wie auch Wochen und Monate später durchgeführt werden.

Dabei ist zu unterscheiden, ob diese Probandenankünfte anonym eingeholt werden oder mit Klarnamen. Hierbei stellt sich die Frage, wer ggf. tatsächlich oder in der Phantasie der Probandin oder des Probanden daraus welche Schlussfolgerungen ziehen könnte. Wenn die Probandin oder der Proband beispielsweise glaubt, ihre oder seine Aussage könne verdeutlichen, dass es ihr oder ihm schlechter gehe als der Gutachter gesagt habe, dann wird dies einen Einfluss auf die Probandenurteile haben.

Aus dieser Perspektive ist es auch wichtig zu unterscheiden, ob die Probandin oder der Proband zum Zeitpunkt der Befragung bereits Kenntnis vom Ergebnis der Begutachtung hatte oder nicht. Falls das Gutachtenergebnis bekannt ist und den Erwartungen der Probandin oder des Probanden entspricht, ist eine positivere und zufriedenerere Einschätzung des Gutachtenprozesses zu erwarten, als wenn die Probandin oder der Proband über das Gutachtenergebnis enttäuscht ist.

Diese Überlegungen zeigen, dass ein Probandenfeedback – auch zum Gutachtenprozess – mehrdimensional zu verstehen ist und keine eindeutige Tatsachenaussage darstellt.

5. Anwendungen von Feedbackbefragungen in der Qualitätssicherung

Es gibt einige Ansätze im Gutachtenkontext und in anderen medizinischen Bereichen, in denen Probandenankünfte oder -beurteilungen genutzt oder empfohlen werden.

5.1. PROs, PROMs und PREM s

Als Patient-Reported Outcomes (PROs) bezeichnet man nicht objektiv beobachtbare, sondern nur durch Befragung des Individuums erfassbare Konstrukte des Befindens oder der gesundheitsbezogenen Lebensqualität. Diese umfassen gemeinhin Dimensionen der körperlichen (z. B. körperliche Funktionsfähigkeit, Schmerzerleben), mentalen (Angst, Depressivität, Stresserleben) und sozialen (soziale Teilhabe) Gesundheit. PROs werden üblicherweise durch standardisierte Fragebögen, sogenannte Patient-Reported Outcome Measures (PROMs) erfasst. PROMs können krankheitsspezifisch oder generisch sein und

umfassen häufig mehrere relevante PROs, die durch mehrere standardisierten Fragen (Items) abgebildet werden.

Ein Ansatz, die Qualität von Instrumenten zur Messung von PROs zu sichern, ist seit dem Jahr 2004 die internationale Patient Reported Outcomes Measurement Information System (PROMIS)-Initiative (<https://promis-germany.de/>).

PROMs können genutzt werden, um Informationen über den aktuellen, selbst wahrgenommenen Gesundheitszustand zu erheben. Dabei können sowohl im Rahmen der Begutachtung selbst erhobene PROM-Daten genutzt werden, als auch solche, die im Rahmen früherer Behandlungen, zum Beispiel bei einem Krankenhausaufenthalt, gesammelt wurden, um den von der zu begutachtenden Person wahrgenommenen Gesundheitszustand zu beurteilen.

Die Abkürzung PREM steht für Patient-Reported Experience Measures und bezieht sich ebenfalls auf standardisierte Fragebögen, die von Probandinnen und Probanden ausgefüllt werden, um ihre Erfahrungen und Einschätzungen im Zusammenhang mit der erlebten Gesundheitsversorgung zu beschreiben. PREMs bilden Strukturmerkmale (z. B. Gestaltung des Settings) und Prozessmerkmale (z. B. Pünktlichkeit, respektvolle Kommunikation) der Begutachtung aus der Sicht der Probandin oder des Probanden ab (IQTIG, 2017).

Die Probandenbeurteilung des Gutachtenprozesses wäre als PREM anzusehen. Beispiele von PROMs und PREMs finden sich bei der amerikanischen Agency for Health Care and Quality und der Consumer Assessment of Healthcare Providers and Systems (CAHPS, <https://www.ahrq.gov/cahps/index.html>). Die zahlreichen bislang eingesetzten Verfahren zur Erfassung der PROs, PROMs und PREMs sind von unterschiedlicher Inhaltsvalidität. Gute Inhaltsvalidität ergaben z. B. Items zu Angst, depressiven Symptomen, Schmerz einschränkungen, Erschöpfung, sowie zur Beziehung zu Gleichaltrigen bei Kindern mit Colitis Ulcerosa (Brenner et al., 2021).

Für erwachsene Rehabilitationspatientinnen und -patienten wiesen z. B. Kriz et al. (2008) zufriedenstellende Gütekriterien für den Fragebogen zur Erfassung der Zufriedenheit in der Therapie (ZUF-8, Schmidt et al., 1989) nach. In einer Studie von Black et al. (2014) an orthopädischen Patientinnen und Patienten in England zeigte sich, dass PROMs und PREMs schwach positiv zusammenhängen, wobei insbesondere die wahrgenommene Kommunikation und das Vertrauen in das Behandlungspersonal mit einer um 30 % geringeren Wahrscheinlichkeit für berichtete Komplikationen einherging. Zonjee et al. (2022) resümierten in einem systematischen Review zur Inhaltsvalidität von Messinstrumenten zur Erfassung selbstberichteter körperlicher Funktionsfähigkeit eine geringe bis mässige Qualität, wobei einschränkend methodische Qualitätsmängel der einbezogenen Studien erwähnt werden.

Verschiedentlich wurde darauf aufmerksam gemacht, dass eine Leistungsbewertung von Einrichtungen anhand der vorliegenden PROMs irreführend sein kann, da systematisch Daten bestimmter Probandengruppen fehlen (Gomes et al., 2016; Hutchings et al., 2014). Köhn et al. (2022) untersuchten Patienten der kardiologischen Rehabilitation in der Schweiz und fanden eine systematische Beantwortungsverzerrung, da etwa ein Drittel ihrer Stichprobe die PROMs nicht beantwortet hatten, d. h. insbesondere Frauen, Personen mit zunehmendem Lebensalter und solche, die nach der Rehabilitation wieder in die Akutversorgung entlassen wurden. Die Rückmeldefreudigkeit unterschied sich auch zwischen verschiedenen Reha-Einrichtungen, was bei Vergleichen von Einrichtungen berücksichtigt werden muss.

5.2. Beurteilung von Lehrveranstaltungen

Heute gehört es an Universitäten zum Standard, dass Studierende die Dozentinnen und Dozenten im Anschluss an eine Veranstaltung beurteilen. Gleiches wird auch von Ärzte- und Psychotherapeutenkammern verlangt, wenn sie eine Veranstaltung mit Fortbildungspunkten ausstatten sollen. Auch Dozentinnen und Dozenten können eine Beurteilung zu den Studierenden abgeben.

Dem Autorenteam ist nicht bekannt, wer welche Schlussfolgerungen aus den vielen Fragebögen zieht. Die Kammern entziehen nicht die Fortbildungspunkte, wenn Veranstaltungsteilnehmende unzufrieden waren. An Universitäten müssen die Lehrevaluationen der Dozentinnen oder Dozenten bei Evaluationsgesprächen (bspw. im Rahmen von Zielvereinbarungen für leistungsorientierte Mittelvergaben oder Weiterbeschäftigung bei Tenure-Track-Stellen) gelegentlich der Hochschulleitung vorgelegt werden. Es wird eine «hohe Lehrqualität» gefordert, wobei unklar ist, was damit genau gemeint ist. Auch hier stellt sich das Problem begrenzter Inhaltsvalidität, wenn die Beurteilung ausschliesslich auf Basis von Studierendenzufriedenheitsrückmeldungen erfolgt, ohne die Lehrziele, -anforderungen und -inhalte zu berücksichtigen. In Habilitationsverfahren (Qualifizierung zur Privatdozentin bzw. zum Privatdozenten) sind neben der Studierendenrückmeldungen auch Lehr- und/oder Fachvorträge vor Wissenschaftskolleginnen und -kollegen zu halten, um die Lehrbefähigung zu prüfen. Dies entspricht einem Peer-Review-Verfahren durch Fachvertreterinnen und -vertreter.

5.3. Versichertenbefragung durch die Rentenversicherung in Deutschland

Die Deutsche Rentenversicherung (DRV) ist sowohl Leistungsträger wie Leistungserbringer in der medizinischen Rehabilitation. Die DRV ist gesetzlich zu Qualitätssicherungsmassnahmen verpflichtet (§ 37 Sozialgesetzbuch IX). Für Rehabilitationskliniken gelten zum einen die vielfältigen Qualitätsanforderungen wie sie auch für sonstige medizinische Einrichtungen gelten. Darüber hinaus hat die DRV ein zusätzliches vielfach gestaffeltes Qualitätssicherungssystem aufgebaut. Dazu gehören detaillierte Beschreibungen der Strukturen und Prozesse in den von der DRV belegten Einrichtungen, Visitationen in den Kliniken, QM-Prozesse mit speziellen Visitationen, Aufzeichnung der klinisch-therapeutischen Leistungen (KTL), ein regelmässiges Peer-Review von Entlassungsberichten, ein differenziertes Beschwerdemanagement und schliesslich auch Patientenbefragungen einige Monate nach dem Aufenthalt in der Klinik.

Die DRV fasst diese Daten dann im Jahresabstand zusammen und erstellt Reports für die einzelnen Einrichtungen, in denen auch ein Benchmarking mit anderen Einrichtungen vorgenommen wird. Eine methodische Voraussetzung ist, dass die Einrichtung eine hinreichende Patientenzahl behandelt hat, so dass Verteilungswerte errechnet werden können.

Bei Auffälligkeiten geben solche Daten Anlass zu Rückfragen oder Visitationen in der jeweiligen Klinik. Die DRV benutzt diese Daten auch, um ein Qualitätsranking von Einrichtungen zu erstellen. Dies geschieht auf der Basis von Daten zur Patientenstruktur, zu Art und Umfang der klinisch-therapeutischen Leistungen, zum Rehabilitationserfolg und nicht zuletzt dem Ergebnis der Patientennachbefragung. Die DRV plant, die Belegung von Kliniken davon abhängig zu machen.

Aus der Rehabilitationsmedizin in Deutschland gibt es Befunde, wonach die subjektive Nutzenbewertungen der Patientinnen und Patienten mit der beruflichen Teilhabe nach der Rehabilitationsbehandlung zusammenhängt (Nübling et al., 2018). Solche Beobachtungsdaten lassen jedoch keine Kausalinterpretation dahingehend zu, dass subjektiv erlebter Misserfolg

durch einen schlechten Behandlungsprozess bewirkt wurde. Patientinnen und Patienten, deren Beitragszahlungen im Verlauf stärker absinken, können generell, d. h. auch bereits zu Beginn der Beobachtung, ausgeprägtere Teilhabeprobleme gehabt haben als Patientinnen und Patienten, deren Beitragszahlungen im Verlauf weniger stark rückläufig sind.

Zudem gibt es Hinweise, dass Patientenbeschwerden, die teilweise weitreichende Auflistungen vermeintlicher Mängel der Klinik und Therapiepersonen enthalten, dazu dienen, zu begründen, warum es zu einer Fehleinschätzung der Leistungsfähigkeit der Patientin oder des Patienten kam, wenn die sozialmedizinische Leistungsbeurteilung bei Entlassung nicht den Vorstellungen der Patientin oder des Patienten entsprochen hat.

In einer Publikation von Muschalla (2018) wurde dieses System analysiert und kritisch kommentiert. Es wurde angemerkt, dass die methodische Qualität der Instrumente bezüglich Reliabilität und Validität nicht hinreichend ist. Schwerwiegende Probleme der Nichtreaktivität sind nicht gelöst (z. B. da viele Diagnosen in das QM eingehen, wird in den Entlassungsberichten die maximal mögliche Zahl aufgeführt; da bestimmte KTL-Zahlen gefordert werden, werden Patientinnen und Patienten ohne Rücksicht auf die tatsächlichen Erfordernisse in Therapiemassnahmen einbezogen; da die Zufriedenheitsratings der Patientinnen und Patienten im QM berücksichtigt werden, werden Patientinnen und Patienten wunschgemäß arbeitsunfähig entlassen). Ebenso können keine kausalen Ableitungen getroffen werden (z. B. kranke Patientinnen und Patienten sind klagsam und klagen auch über den Aufenthalt). Wenn Rehabilitanden-Befragungsdaten nicht im Sinne einer Abbildung von Qualität medizinischer Rehabilitationsbehandlung (d. h. komplexer Diagnostik- und Behandlungsaktivitäten, vgl. § 42 SGB IX) valide sind, ist ihre Verwendung als Benchmarking- oder gar Steuerungsinstrument nicht möglich.

Sie können jedoch als Screening für besondere Auffälligkeiten genutzt werden. Beispielsweise kann man die Kliniken mit erhöhtem Unzufriedenheitsanteil prüfen und darauf aufbauend im Rahmen von Beobachtung und Exploration vor Ort (Visitationen, strukturierter Qualitätsdialog) herausfinden, was das Problem ist. Aussagen darüber, ob in Kliniken die «medizinische Rehabilitation» (§ 42 SGB IX) für den individuellen Patienten nach Fachstandard durchgeführt wurde, sind nur im Rahmen von Visitationen (durch Ärzte und Therapeuten der jeweiligen Indikation) und direkter Beobachtung durch Fachkräfte in der Klinikroutine möglich.

5.4. Basler Fairness Fragebogen BFF

Unter explizitem Fokus auf das Erleben des Begutachtungsprozesses wurde der Basler Fairness Fragebogen BFF (Rosburg et al., 2021; Lohss et al., 2020) entwickelt. Er ist als PREM anzusehen. In der aktuellen Form enthält er 25 sowohl emotionsorientierte («Ich fühlte mich ernst genommen»), wie verhaltensorientierte Items («Der Gutachter hat mich informiert»), die in den vier Subdimensionen Respekt, Interviewführung, Fallvertrautheit und Transparenz zusammengefasst werden können.

Es liegen erste Daten zu den Gütekriterien vor. Allerdings gibt es u. W. noch keine verlässlichen repräsentativen Daten, die Vergleichsdaten zur Einordnung von Einzelfällen liefern könnten. Die Autorenschaft des BFF weist auf die Interpretierbarkeitsbegrenzungen der Fragebogendaten hin und betont, dass man von einer Negativbeurteilung zu keinen Schlüssen über den Einzelfall gelangen kann. Befragungen mit dem BFF müssten sich zudem auf eine repräsentative Menge der Gutachten einer Gutachterin oder eines Gutachters beziehen, wobei unklar ist, wie viele Gutachten eine einzelne Gutachterin oder ein einzelner Gutachter erstellen müsste, um eine Aussage über die betreffende Gutachterin oder den betreffenden Gutachter als Person treffen zu können. Die Autorengruppe diskutiert auch, dass

evtl. weitere Dimensionen aufgenommen werden müssen, um den Gutachtenprozess vollumfänglich abzubilden.

Eine offene Frage ist auch, ob ein umfangreicher Fragebogen wie der BFF letztlich bessere Informationen liefert als ein Kurzinstrument von wenigen Items oder sogar nur ein Globalurteil. Die Autorenschaft des BFF spricht sich für ein detailliertes Feedback aus, da dies erlaube, unterschiedliche Aspekte der erlebten Fairness aus Sicht der gesuchstellenden Person abzubilden, was für gezielte Qualitätssicherungs-Massnahmen wichtig sei.

6. Zusammenfassung und Empfehlung

Seitens der EQMB ist festzulegen, wozu Probandenbewertungen des Gutachtenergebnisses und -prozesses dienen sollen: (a) zur Überprüfung der Vollständigkeit der Faktenerhebung und der Gültigkeit des Gutachtenergebnisses, (b) zur Generierung von Daten, die in der Fachöffentlichkeit und Politik einen Eindruck von den Abläufen beim Begutachtungsprozess geben können, (c) zur Veränderung des Gutachterverhalten und ggf. auch den Ausschluss einzelner Gutachterinnen und Gutachter, (d) als Sentinel zur Zweitsichtung und Überprüfung von Gutachten, die möglicherweise aus Verfahrensgründen inkorrekt sind, (e) zur präventiven Abwehr späterer Probandenvorwürfe.

6.1. Nutzung eines systematischen Probandenurteils zum Gutachtenergebnis

Eine Probandenbeurteilung zur Zufriedenheit mit dem Gutachtenergebnis kann erst nach Gutachtenfertigstellung und Übermittlung des schriftlichen Gutachtentextes an begutachtete Person erfolgen. Dies könnte sinnvoll sein, wenn der zuständige Versicherungsträger daran ein Review-Verfahren koppeln würde, um bei Negativurteilen intern vor einem Versicherungsentscheid das Gutachten auf seine Tragfähigkeit zu prüfen, als Massnahme zur Vermeidung von nachfolgenden Rechtsauseinandersetzungen und zur Förderung des Rechtsfriedens.

Methodisch sollte dies in Form eines Freitextes erfolgen, in dem die begutachtete Person oder deren Rechtsvertretung, die aus ihrer Sicht strittigen Punkte darlegt. Dies erfolgt in aller Regel bereits routinemässig bei Rekursen gegen einen auf dem Gutachten aufbauenden Versicherungsentscheid in Form von Spontanbeschwerden.

Ggf. könnte das im Ablauf systematisiert und vorgezogen werden mit einer routinemässigen Anfrage an den Probanden, ob es aus seiner Sicht eine Notwendigkeit zur Ergänzung oder Korrektur des Gutachtens gibt, bevor ein Entscheid erfolgt.

6.2. Nutzen von Probandenurteilen zur Gutachterbeurteilung und -auswahl

Die systematische Befragung von Begutachteten zu ihrer Zufriedenheit mit dem Begutachtungsprozess entspricht im engeren Sinne dem, was die Fragestellung der EKQMB ist («Nutzen einer systematischen Feedbackerhebung von Probandinnen und Probanden [Versicherten] für die Beurteilung der Qualität von Gutachten» und der «Zufriedenheit mit dem Untersuchungsgespräch»). Gemäss einer nicht repräsentativen Auswertung der Organisation Inclusion Handicap beurteilen viele Probandinnen und Probanden das Gesprächsklima während einer Begutachtung als schlecht und die Begutachtenden als nicht interessiert.

Nimmt man diese Beschwerden ernst und möchte man sich auf Sachverständige beschränken, die von den Probandinnen und Probanden positiv bewertet werden, dann müssen dies einerseits Rückmeldungen ohne Kenntnis des Gutachtenergebnisses sein.

Andererseits muss die Rückmeldung, wie oben begründet, mit höchst reliablen und vor allem validen Instrumenten erfolgen.

Schliesslich ist jedoch vor allem wichtig zu klären, was – unter Berücksichtigung des eigentlichen Gutachterauftrags, der oben beschriebenen Besonderheiten der Probanden-Gutachter-Beziehung, der möglichen Negativrückkopplungen auf die sachgerechte Gutachtenerstellung sowie der Auswirkungen des individuellen psychischen Zustands der begutachteten Person – unter einem «positiv erlebten Gutachterverhalten» zu verstehen ist.

Wie ausgeführt, ermöglicht ein Probandenurteil des Gesprächsklimas keine reliable, valide und objektive Aussage darüber, ob das Gutachterverhalten fachgerecht war. Dazu sind die vorstehenden Ausführungen zur Konfundierung von Befinden und Interaktionserleben, zur befindensabhängigen Rückerinnerung, zum Zusammenhang zwischen Gutachtenergebnis und Klage über die Gutachterin oder den Gutachter und zu den interaktionellen Besonderheiten der Gutachtensituation zu beachten.

6.3. Darstellung der Güte von Gutachtenprozessen in der interessierten Öffentlichkeit

Wenn das Ziel ist, in der politischen und Fachdiskussion (z. B. mit Betroffenenverbänden, der Politik, der Öffentlichkeit) datenbasierte Aussagen zur Qualität der Begutachtungsprozesse im Allgemeinen zu machen, dann könnte dies mithilfe einer systematischen Befragung der Versicherten erfolgen.

Gemäss den vorstehenden Erläuterungen unter Berücksichtigung der jeweiligen Vor- und Nachteile zu entscheiden, ob dies unmittelbar nach der Begutachtung (eher sinnvoll) oder mit zeitlichem Abstand (eher weniger sinnvoll) erfolgen soll, ohne Kenntnis des Gutachtenergebnisses (eher sinnvoll) oder nach Gutachtenveröffentlichung (eher weniger sinnvoll), mit Namensnennung an den Auftraggeber (eher sinnvoll) oder anonym (eher weniger sinnvoll).

Des Weiteren wäre zu entscheiden, ob man ein differenziertes Instrument wie den BFF einsetzt, was auch differenziertere Diskussion ermöglicht oder provoziert. Die Alternative können Kurzinstrumente sein, die beispielsweise nach Art einer visuellen Analogskala auf einem Zehn-Punkte-Kontinuum abfragen, ob die Probandin oder der Proband mehr oder weniger zufrieden war, z. B. (a) mit der Freundlichkeit und Gesprächsführung der Gutachterin oder des Gutachters und (b) mit der Bereitschaft, der Probandin oder dem Probanden zuzuhören.

6.4. Sentinel-Rückmeldungen zur Gutachtengüte

Systematische Befragungen zur Zufriedenheit mit der Gutachterin oder dem Gutachter und dem Gutachtenprozess können auch sinnvoll sein, im Sinne einer Sentinel-Erhebung. Damit ist gemeint, dass die begutachtete Person unmittelbar nach der Begutachtung die Gelegenheit erhält zu sagen, ob aus ihrer Sicht beim Gutachtenprozess alle relevanten Fakten zur Sprache kamen und ob die Interaktion mit der Gutachterin oder dem Gutachter zufriedenstellend war.

Wird dies von den begutachteten Personen verneint, dann wirft dies die Frage auf, ob das Gutachten möglicherweise ergänzungs- oder korrekturbedürftig ist. Dies könnte dann Anlass sein, das fertige Gutachten einer Überprüfung zu unterziehen, bevor daraus rentenrelevante

Entscheide abgeleitet werden. Die kritischen Anmerkungen der begutachteten Person könnten auch der Gutachterin oder dem Gutachter zur Kenntnis gegeben werden, mit der Bitte um Stellungnahme.

Auch hier stellt sich die Frage, ob ein differenziertes Messinstrument wie der BFF eingesetzt werden sollte oder eher eine fokussierte Frage nach konkreten Problemen in der Gutachtensituation mit der zusätzlichen Möglichkeit zu präzisierendem und ergänzendem Freitext.

7. Zusammenfassung

7.1. Inwieweit liegen empirische Daten zur Art und Nutzung von Probandenselbstberichten während und nach Begutachtungen vor und welchen Nutzen und welche Grenzen haben Probandenselbstauskünfte im Kontext sozialmedizinischer Begutachtungen?

- Aufgrund der Heterogenität dessen, was „Qualität medizinischer Gutachten“ in verschiedenen medizinischen Fachbereichen ausmacht (unterschiedliche spezifische Anforderungen in unterschiedlichen Untersuchungssettings, Indikationen, Probandenklientel, Gutachtenfragestellungen) kann es keine Global-Operationalisierung einer „guten Begutachtung“ geben.
- Bei Gutachterprozessen ist die Probandenzufriedenheit dem Ziel einer Korrektheit der Begutachtung untergeordnet.
- Es sind die methodischen Grenzen von Probandenbefragungen zu berücksichtigen.
- Bei Bewertungs- und Steuerungsprozessen sind höchste Anforderungen an die Reliabilität und Validität und Nicht-Reaktivität der Instrumente zu stellen.
- Bei Public Reporting und Gruppenerhebungen genügen geringer, d.h. übliche Methodenstandards.
- Grundsätzlich erlauben Probandenaussagen nur sehr begrenzt valide Rückschlüsse.
- Probandenbefragungen können als Anlass dienen, im Einzelfall ein Gutachten einer internen Prüfung zu unterziehen, bevor leistungsrelevante Entscheide getroffen werden.
- Es sind auch die Nebenwirkungen von Probandenbefragungen zu bedenken, wie unerwünschte Änderungen im Gutachterverhalten.
- Es gibt vielfältige Anwendungen von PROs, PROMs und PREMs im Allgemeinen. Ein gutachtenspezifisches Instrument zur Beurteilung des Begutachtungsprozesses ist der BFF.

7.2. Unter welchen Bedingungen und mit welcher Zielsetzung sind Probandenselbstauskünfte für die Qualitätssicherung in der medizinischen Begutachtung hilfreich?

- Probandinnen und Probanden werden am Ende der Begutachtung aufgefordert, eine kurze Rückmeldung zum Begutachtungsprozess zu geben.
- Es könnte der BFF oder ein Kurzverfahren verwendet werden, in dem explizit danach gefragt wird, ob die Gutachterin oder der Gutachter freundlich war und ob sie oder er alle Probandeneinlassungen angemessen zur Kenntnis genommen hat. In jedem Fall sollte noch die Möglichkeit gegeben werden, im Freitext zu sagen, was gefehlt hat.
- Die Probandin oder der Proband schickt diese Rückmeldung im verschlossenen Umschlag an den Versicherungsträger oder an eine unabhängige Stelle wie die EKQMB.

- Bei kritischen Rückmeldungen wird das Gutachten versicherungsintern nochmals einer fachlichen Prüfung unterzogen, bevor ein Versicherungsentscheid gefällt wird.
- Die Probandin oder der Proband wird neutral, ohne bewertende oder entschuldigende Stellungnahme seitens der Behörde über diesen Review informiert, damit sie oder er weiss, dass ihre oder seine Rückmeldungen zur Kenntnis genommen werden. Dieses Vorgehen kann den Rechtsfrieden fördern.
- Falls zu einem späteren Zeitpunkt durch die Probandinnen und Probanden oder ihre Rechtsvertretung gegen den Gutachtenprozess geklagt wird, kann die initiale Probandeneinschätzung zum Vergleich und in manchen Fällen auch zur Richtigstellung herangezogen werden.
- Die EKQMB oder das BSV haben die Möglichkeit, über diese Rückmeldungen in gebotenen Abständen Statistiken zu erstellen und zu veröffentlichen.
- Diese Statistiken können die politische Diskussion versachlichen.
- Im Sinne einer Qualitätsschulung und Qualitätssicherung können diese Statistiken bei geeigneter Gelegenheit mit den Sachverständigen besprochen werden.

Das skizzierte Vorgehen könnte einen Beitrag zur Verbesserung der inhaltlich-fachlichen Gutachtenqualität, zur Probandenzufriedenheit, zum Rechtsfrieden, zur Versachlichung der öffentlichen Diskussion und zur Qualitätssicherung und -verbesserung leisten.

8. Literatur

- AWMF. (2019). *AWMF-Leitlinie zur Begutachtung psychischer und psychosomatischer Störungen. Teil II: Begutachtung der beruflichen Leistungsfähigkeit*. Berlin: AWMF Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Online abgerufen am 20.11.2022 <https://register.awmf.org/de/leitlinien/detail/051-029>
- Black, N., Varaganum, M., Hutchings, A. (2014). Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery. *BMJ Qual Safety*, 23, 534–542. <http://dx.doi.org/10.1136/bmjqs-2014-002993>
- Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion*, 24(4), 561-595.
- BPtK. (2018). *Muster-Berufsordnung für die Psychologischen Psychotherapeutinnen und Psychotherapeuten und Kinder- und Jugendlichenpsychotherapeutinnen und Kinder- und Jugendlichenpsychotherapeuten*. Dortmund: Deutscher Psychotherapeutentag.
- Brenner, E.J., Long, M.D., Mann, C.M., Lin, M.S., Chen, W., Reyes, C., Bahnson, K.M., Reeve, B.B. (2021). Validity and responsiveness of the Patient-Reported Outcomes Measurement Information System in Children with Ulcerative Colitis. *Pediatr Gastroenterol Nutrition*, 73(1), 67–72. doi:10.1097/MPG.0000000000003101.
- Caussé-Versaveau F, Assié P, Jammet P, Lacambre M, Courtet P. (2018). *L'expert psychiatre peut-il diagnostiquer la simulation en expertise de dommage corporel ?* L'Information psychiatrique. 94 (4): 271-4 doi:10.1684/ipe.2018.1789
- Deutscher Ärztetag. (2018). Muster-Berufsordnung für die in Deutschland tätigen Ärztinnen und Ärzte. *Deutsches Ärzteblatt*, 05/2018, A1–A9.
- DRV. (2017). *Medizinische Rehabilitation: Wie sie Ihnen hilft*. Berlin: Deutsche Rentenversicherung.
- Echterhoff, W., Heinecke, M., & Brückner, P.F. (2010). *Informationsbroschüre: Das rundum gute Gutachten. Qualität von Sachverständigen-Gutachten und ihr Beitrag zum Rechtsfrieden*. Bergische Universität Wuppertal und Sozialgericht Düsseldorf.
- Farin, E., & Jäckel, W.H. (2011). Qualitätssicherung und Qualitätsmanagement in der medizinischen Rehabilitation. *Bundesgesundheitsblatt*, 54, 176–184.
- Gomes, M., Gutacker, M., Gutacker, N., Bojke, C., Street, A. (2016). Addressing Missing Data in Patients-Reported Outcome Measures (PROMS): Implications for the Use of PROMS for Comparing Provider Performance. *Health Economics*, 25, 515–528. doi: 10.1002/hec.3173
- Hutchings, A., Neuburger, J., von der Meulen, J., Black, N. (2014). Estimating recruitment rates for routine use of patient reported outcome measures and the impact on provider comparisons. *BMC Health Services Research*, 14, 66. <http://www.biomedcentral.com/1472-6963/14/66>
- IQTIG. (2017). *Entwicklung und Durchführung von Qualitätssicherungsverfahren im Gesundheitswesen*. Institut für Qualitätssicherung und Transparenz im Gesundheitswesen.

- Kieser, U. (2017). *Verdeutlichung, Aggravation und Simulation im Sozialversicherungsrecht*. In Kieser, Ueli & Lendfers, Miriam (Hrsg.). *Jahrbuch zum Sozialversicherungsrecht 2017*. Zürich/St.Gallen: Dike Verlag, 2017, S. 185-194.
- Klosterhuis, H. (2010). Reha-Qualitätssicherung der Rentenversicherung – eine kritische Bestandsaufnahme. *RV aktuell*, 08/2010, 260–268.
- Köhn, S., Schlumbohm, A., Marquardt, M., Scheel-Sailer, A., Tobler, S., Vontobel, J., Menzi, L. (2022). Predicting non-response in patient-reported outcome measures - results from the Swiss quality assurance program in cardiac inpatient rehabilitation. *International Society for Quality in Health Care*, accepted for publication.
- Kriz, D., Nübling, R., Steffanowski, A., Wittmann, W.W., Schmidt, J. (2008). Patientenzufriedenheit in der stationären Rehabilitation: Psychometrische Reanalyse des ZUF-8 auf der Basis multizentrischer Stichproben verschiedener Indikation. *Zeitschrift für Medizinische Psychologie*, 17, 67 – 79.
- Linden, M. (2011). *Möglichkeiten und Grenzen der Beobachtungsforschung in der psychosomatischen Rehabilitation*. Vortrag im Rahmen des Forschungskolloquiums der Forschungsgruppe Psychosomatische Rehabilitation. Rehasentrum Seehof, Teltow.
- Linden, M., Muschalla, B., Noack, N., Heintze, C., & Döpfner, S. (2018). Treatment Changes in General Practice Patients with Chronic Mental Disorders following a Psychiatric-Psychosomatic Consultation. *Health Services Research and Managerial Epidemiology*, 5, 1-6.
- Lohss, R., Rosburg, T., Bachmann, M., Meyer, B.W., de Boer, W., Fischer, K., Kunz, R. (2020). Perceived fairness of claimants undergoing a work disability evaluation: Development and validation of the Basel Fairness Questionnaire. *PLoS One*, 15(9), e0238930.
- Luo, J., & Yu, R. (2015). Follow the heart or the head? The interactive influence model of emotion and cognition. *Frontiers in Psychology*, 6, 573.
- Merten, T. (2020). Beschwerdvalidierung in der psychologischen und medizinischen Begutachtung. *Praxis der Rechtspsychologie*, 30(2), 59-67.
- Muschalla, B. (2018). What can and what cannot be concluded from patient-reported outcomes as a quality indicator in medical rehabilitation? *Praxis Klinische Verhaltensmedizin und Rehabilitation*, 31, 227-238.
- Muschalla, B., & Schönborn, F. (2021). Induction of False Beliefs and False Memories in Laboratory Studies – A Systematic Review. *Clinical Psychology and Psychotherapy*, xx, accepted for publication. doi: 10.1002/cpp.2567
- Pfundmair, M. (2020). Aussagepsychologie. In: *Psychologie bei Gericht. Die Wirtschaftspsychologie*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-61796-0_1
- Reinecke, J. (2022). Grundlagen der standardisierten Befragung. In *Handbuch Methoden der empirischen Sozialforschung* (pp. 949-967). Wiesbaden: Springer Fachmedien Wiesbaden.
- Rogge, A. A., Fischer, F., Otto, L., & Rose, M. (2022). Empirische Erfassung patient*innenberichteter Merkmale: PROMs und PREMs. *AINS - Anästhesiologie · Intensivmedizin · Notfallmedizin · Schmerztherapie*, 57(02), 150–155. <https://doi.org/10.1055/a-1452-2788>

- Rosburg, T., Lohss, R., Bachmann, M.S., Walter Meyer, B., de Boer, W., Fischer, K., & Kunz, R. (2021). *Basler Fairness Fragebogen (BFF): Erlebte Fairness der Begutachtung*. SUVA.
- Schmidt-Atzert, L., & Amelang, M. (2018). *Psychologische Diagnostik*. Berlin: Springer.
- SGVP. (2016). *Qualitätsleitlinien für versicherungspsychiatrische Gutachten*. Schweizerische Gesellschaft für Versicherungspsychiatrie. *Qualitätsleitlinien für versicherungspsychiatrische Gutachten (17.10.2016)*. Online abgerufen am 20.11.2022: <https://www.sgvp.ch/leitlinien/>
- Thierbach, C., & Petschick, G. (2019). Beobachtung. *Handbuch Methoden der empirischen Sozialforschung*, 1165-1181.
- Toolin, K., van Langeraad, A., Hoi, V., Scott, A.J., & Gabbert, F. (2022). Psychological contributions to cold case investigations: A systematic review. *Forensic Sci Int Synerg.*, 19;5:100294. doi: 10.1016/j.fsisyn.2022.100294.
- WHO. (2001). *Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit ICF*. Genf: Weltgesundheitsorganisation.
- Widera, T. (2010). Aktuelles aus der Reha-Qualitätssicherung- Neue Ergebnisse der Rehabilitandenbefragung. *RV aktuell*, 4, 2010, 153–159.
- Widera, T., Beckmann, U., Zander, J., & Radoschewski, M. (2011). Wer hat Recht - Rehabilitand oder KTL-Dokumentation? *RV aktuell* 10, 2011, 291–298.
- Zonjee, V.J., Abma, I.L., de Mooij, M.J., van Schaik, S.M., Van den Berg-Vos, R.M., Roorda, L.D., Terwee, C.B. (2022). The patient-reported outcomes measurement information systems (PROMIS®) physical function and its derivative measures in adults: a systematic review of content validity. *Quality of Life Research*, 31, 3317–3330. <https://doi.org/10.1007/s11136-022-03151-w>