

Risikoselektion statt Qualitätssicherung

Warum das neue IQTIG-Verfahren die psychotherapeutische Versorgung gefährdet

Florian Lampersberger M.A., M.A.¹

Psychoanalytiker (DGPT), Psychologischer Psychotherapeut, München

mail@praxis-lampersberger.de

Version 1.1 (September 2025); veröffentlicht auf: <https://couch-und-agera.de/>

Copyright: © 2025, Florian Lampersberger

Abstract: Die Qualitätssicherung (QS) in der deutschen Psychotherapie durchläuft mit der Abschaffung des Gutachterverfahrens und der Einführung eines datengestützten QS-Systems des IQTIG einen fundamentalen Paradigmenwechsel. Dieser Artikel

¹ Die Autorschaft dieses Textes ist hybrid und wird hier bewusst transparent gemacht. Er entstand im "Ringgen mit dem Automaten" – einem dialektischen Prozess der Ko-Produktion zwischen dem menschlichen Autor und generativen KI-Systemen (insbesondere ChatGPT 5 Pro, Gemini 2.5 Pro). Diese Praxis ist keine unreflektierte Delegation, sondern folgt einem strengen ethischen und methodischen Leitfaden. Die KI wurde hier als "künstlicher Peer" eingesetzt, um meine eigenen Thesen zu provozieren und zu schärfen, niemals jedoch als Quelle originärer Inhalte. Die genauen Details dieses Mensch-KI-Amalgams, inklusive einer Dokumentation des Arbeitsprozesses und beispielhafter Prompts für den vorliegenden Text, werden im Anhang offengelegt. Die finale Form des Textes ist das Ergebnis eines bewussten Aktes der "Menschlichung" und Re-Autorisierung, für den ich die alleinige Verantwortung trage. Dieser Hinweis dient somit nicht nur der Kennzeichnung, sondern ist eine Einladung, die Lektüre dieses Artikels auch als Fallbeispiel für die im Haupttext analysierte Gratwanderung zwischen technologischer Möglichkeit und intellektueller Integrität zu verstehen. Den vollständigen Referenztext finden Sie hier: <https://couch-und-agera.de/das-ringgen-mit-dem-automaten/>

unterzieht das neue Verfahren einer kritischen Analyse und argumentiert, dass es an gravierenden wissenschaftlich-methodischen Konstruktionsfehlern leidet. Dazu zählen der Verzicht auf eine Baseline-Messung, die Verwendung nicht validierter Instrumente und vor allem das Fehlen einer funktionierenden Risikoadjustierung. Die zentrale These lautet, dass diese Mängel eine perverse Anreizstruktur schaffen, die eine systematische ‚Risikoselektion‘ erzwingt. Therapeuten werden ökonomisch und reputativ bestraft, wenn sie schwerkranke Patientinnen und Patienten, beispielsweise mit Persönlichkeitsstörungen, behandeln, deren Therapieverläufe klinisch erwartbar krisenhafter sind. Dies birgt die immanente Gefahr, dass diese vulnerablen Gruppen zukünftig aus der Regelversorgung verdrängt werden und eine neue, systemgemachte Versorgungslücke entsteht. Im Kontrast dazu werden professionseigene, wissenschaftlich fundierte QS-Ansätze wie das Routine Outcome Monitoring (ROM) als konstruktive Alternativen vorgestellt. Das Fazit mündet in der dringenden Forderung nach einer grundlegenden Revision des IQTIG-Verfahrens während seiner Erprobungsphase, um einen irreversiblen Schaden für die Versorgungsgerechtigkeit abzuwenden.

Schlüsselwörter: Qualitätssicherung, Psychotherapie, IQTIG, Risikoselektion, Versorgungsgerechtigkeit, DeQS-Richtlinie, Gutachterverfahren, Patientenversorgung

Inhaltsverzeichnis

1	Einleitung: Ein System mit gefährlichen Nebenwirkungen	6
1.1	Der Paradigmenwechsel: Von professioneller Reflexion zu bürokratischer Messung	7
1.2	Kernthese: Wie ein gut gemeintes QS-System die Versorgung der Schwächsten untergräbt.....	9
1.3	Aufbau der kritischen Analyse	10
2	Die Baseline: Was mit dem Gutachterverfahren verloren geht	12
2.1	Mehr als Bürokratie: Die Doppelfunktion als Peer-Review und ökonomischer Schutzschirm.....	12
2.2	Das politische Ende: Eine Reform ohne fachliche Debatte.....	16
2.3	Die Bedeutung des verlorenen Regress-Schutzes für die Behandlung komplexer Fälle	17
3	Anatomie des neuen Systems: Das IQTIG-QS-Verfahren im Detail.....	20
3.1	Der gesetzliche Auftrag: Ziele und Rahmen der DeQS-Richtlinie	20
3.2	Die Messinstrumente: Therapeutendokumentation und Patientenfragebogen	21
3.3	Der Mechanismus: Von der Datenerhebung zu öffentlichen Qualitätsberichten	24
3.4	Die Notbremse: Warum eine Erprobungsphase in NRW unumgänglich wurde	25
4	Fundamentale Konstruktionsfehler: Die wissenschaftliche Untauglichkeit des IQTIG-Verfahrens	27
4.1	Methodische Mängel: Warum die Ergebnisse nicht valide sind (Validität, Reliabilität, Baseline-Problem).....	27

4.2	Konzeptionelle Schwächen: Suggestive Fragen, Erinnerungsverzerrung und die Messung von Konformität statt Qualität.....	29
4.3	Die Bilanz: Ein bürokratisches Monster ohne klinischen Nutzen	30
5	Die reale Gefahr: Wie das System Risikoselektion erzwingt und Praxen ruiniert.....	32
5.1	Der systemische Designfehler: Die fatale Wirkung fehlender Risikoadjustierung.....	32
5.2	Härtetest Persönlichkeitsstörung: Wenn klinische Realität als schlechte Qualität fehlinterpretiert wird	33
5.3	Das Kalkül des Ruins: Warum die Behandlung schwerer Fälle zum untragbaren wirtschaftlichen Risiko wird	35
5.4	Die Konsequenz für die Gesellschaft: Eine neue, hausgemachte Versorgungslücke	36
6	Der konstruktive Gegenentwurf: Professionseigene Wege zur echten Qualitätssicherung.....	38
6.1	Die Philosophie der Verbesserung: Prozessorientierung, Feedback und wissenschaftliche Fundierung.....	38
6.2	Funktionierende Modelle in der Praxis: Routine Outcome Monitoring (ROM) und das QSP/QVA-Projekt	39
6.3	Der entscheidende Unterschied: "Bottom-up"-Unterstützung statt "Top-down"-Kontrolle	41
7	Fazit und Forderungen: Plädoyer für eine sofortige Kurskorrektur	43
7.1	Zusammenfassung: Ein System, das seine Ziele verfehlt und Schaden anrichtet	43
7.2	Konkrete Forderungen für die Evaluation in der Erprobungsphase	44
7.3	Ausblick: Skizze eines fairen und nützlichen QS-Systems der Zukunft	45
8	Literaturverzeichnis.....	47

9	Fiktives Plenum	50
10	Anhang: Zur Entstehung dieses Textes: Eine Reflexion im Lichte des Leitfadens zur KI-Ko-Produktion	83

1 Einleitung: Ein System mit gefährlichen Nebenwirkungen

Die Qualitätssicherung (QS) in der ambulanten psychotherapeutischen Versorgung in Deutschland befindet sich inmitten einer tiefgreifenden, potenziell disruptiven Transformation. Mit der gesetzlichen Verabschiedung der schrittweisen Abschaffung des jahrzehntelang etablierten Antrags- und Gutachterverfahrens im Jahr 2019 wurde nicht nur eine administrative Prozedur reformiert, sondern ein fundamentaler Paradigmenwechsel eingeleitet (BPtK, 2019). An die Stelle eines Systems, das auf qualitativer, kollegialer Einzelfallprüfung, professioneller Reflexion und narrativer Fallkonzeption basierte, tritt ein quantitatives, datengestütztes Verfahren. Dieses neue System, normiert durch die Richtlinie zur datengestützten einrichtungsübergreifenden Qualitätssicherung (DeQS-RL) des Gemeinsamen Bundesausschusses (G-BA) (G-BA, 2024a) und konzipiert vom Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG), verfolgt das erklärte Ziel, die Versorgungsqualität messbar, vergleichbar und transparenter zu machen (Benecke et al., 2025).

Dieser Wandel ist jedoch weit mehr als ein technisches Upgrade; er markiert eine epistemologische Zäsur – einen grundlegenden Wandel in der Auffassung dessen, was als Wissen über Behandlungsqualität gilt und wie dieses Wissen generiert werden soll. Er wirft die zentrale Frage auf, ob die komplexe, intersubjektive und hochgradig individualisierte Realität psychotherapeutischer Prozesse adäquat in standardisierte Datenpunkte übersetzt werden kann, ohne dabei ihre wesentliche Substanz zu verlieren. Während die Befürworter der Reform eine Modernisierung und Entbürokratisierung anstreben, warnen weite Teile der psychotherapeutischen Fachwelt – von den großen Berufsverbänden (z.B. DPtV et al., 2022) über die Psychotherapeutenkammern bis hin zu führenden Wissenschaftlern (Volz & Benecke, 2024) – eindringlich vor den systemischen Risiken und unbeabsichtigten, aber potenziell verheerenden Nebenwirkungen dieses neuen Ansatzes. In einer scharfen Kritik bezeichnet die Bundespsychotherapeutenkammer (BPtK, 2024) das geplante Verfahren als "nachteilig für die Patientenversorgung" und die eingesetzten Instrumente als "schlicht ungeeignet".

Die vorliegende Analyse vertritt die These, dass das IQTIG-Verfahren in seiner aktuellen Konzeption nicht nur wissenschaftlich-methodisch unzureichend ist (Volz & Benecke, 2024), sondern durch systematische Fehlanreize eine Kultur der Risikovermeidung etabliert, die die Versorgung gerade der vulnerabelsten und am

schwersten erkrankten Patientengruppen aktiv gefährdet. Anstatt die Qualität zu sichern, droht es, durch die Förderung von Risikoselektion neue, systembedingte Versorgungslücken zu schaffen und damit das Prinzip der Versorgungsgerechtigkeit im Kern zu untergraben (Benecke et al., 2025). Dieser Artikel wird die Anatomie dieses neuen Systems kritisch sezieren, seine Konstruktionsfehler aufdecken und die daraus resultierenden Gefahren für die Praxis und die Patientenversorgung detailliert analysieren. Er ist zugleich ein Plädoyer für eine Kurskorrektur und die Hinwendung zu alternativen, aus der Profession heraus entwickelten QS-Ansätzen, die eine echte Qualitätsverbesserung ermöglichen, anstatt lediglich bürokratische Konformität zu kontrollieren.

1.1 Der Paradigmenwechsel: Von professioneller Reflexion zu bürokratischer Messung

Um die Tragweite des aktuellen Wandels zu erfassen, ist es notwendig, die fundamental unterschiedlichen Philosophien zu verstehen, die dem alten und dem neuen System zugrunde liegen. Es handelt sich um den Übergang von einem hermeneutisch-qualitativen zu einem positivistisch-quantitativen Paradigma der Qualitätssicherung.

Das alte Paradigma: Qualität als Ergebnis professioneller Reflexion und kollegialen Diskurses

Das nun auslaufende Antrags- und Gutachterverfahren war mehr als nur eine Hürde zur Genehmigung einer Langzeittherapie; es war ein institutionalisierter Zwang zur Professionalisierung. Das Herzstück dieses Systems war der schriftliche Bericht an den Gutachter. Wie Hauten und Jungclaussen (2022) darlegen, war dieser Akt des Verfassens kein bloßer administrativer Vorgang, sondern ein integraler Bestandteil des therapeutischen Prozesses selbst. Ein Therapeut war gezwungen, die Fülle an Informationen eines Falles zu strukturieren, zu deuten und in eine kohärente, narrative Fallkonzeption zu gießen. Dieser Prozess erforderte und förderte zentrale therapeutische Kompetenzen wie die systematische Fallkonzeption, die kritische Selbstreflexion und die fundierte Indikationsprüfung (vgl. DGPT, 2019).

Die anschließende Begutachtung durch einen erfahrenen Kollegen fungierte als kollegiales Peer-Review. Der Gutachter prüfte nicht anhand einer Checkliste, sondern beurteilte die klinische Stimmigkeit und Plausibilität der Argumentation. Sein Votum war das Ergebnis eines verstehenden, hermeneutischen Prozesses.

Qualität definierte sich in diesem Paradigma also über die Tiefe und Plausibilität der fachlichen Auseinandersetzung mit dem Einzelfall. Es war ein qualitatives System, das auf dem Vertrauen in professionelle Expertise und den Wert des fachlichen Diskurses basierte. Die zugrundeliegende Annahme war: Ein gut durchdachter, reflektierter und konzeptionell fundierter Therapieprozess ist die beste Gewähr für ein gutes Behandlungsergebnis.

Das neue Paradigma: Qualität als messbarer, standardisierter und vergleichbarer Datenpunkt

Das vom IQTIG entwickelte QS-Verfahren bricht radikal mit dieser Logik. Es ersetzt das narrative, qualitative Urteil durch die Erhebung quantitativer, standardisierter Daten. Die zentrale Idee ist, Qualität operationalisierbar zu machen – sie in Zahlen zu fassen, um sie zu objektivieren, zu vergleichen und letztlich auch öffentlich darzustellen (IQTIG, 2022). Die Philosophie dahinter ist tief im Managementdenken und der evidenzbasierten Medizin verankert, getreu dem Motto: "What gets measured, gets managed."

Dieses neue Paradigma definiert Qualität primär über zwei Achsen:

- **Prozesskonformität:** Die Therapeutendokumentation erfasst, ob bestimmte, als qualitätsrelevant definierte Prozessschritte durchgeführt wurden (z.B. Einsatz von Diagnostik, Zielvereinbarung) (G-BA, 2024a). Die Qualität eines Prozesses wird an seiner formalen Korrektheit und Vollständigkeit gemessen, die sich stark an den Empfehlungen der BPtK (2020) zur Dokumentation orientieren. Der Therapeut agiert nicht mehr primär als reflektierender Kliniker, sondern als Akteur in einem standardisierten Versorgungsprozess, der nachweislich Prozeduren einhält.
- **Subjektive Patienteneinschätzung:** Der Patientenfragebogen erhebt die retrospektive Sicht des Patienten und konzipiert Qualität damit stark als eine Form von "Kundenzufriedenheit" (Benecke et al., 2025). Zentrale Indikatoren sind die Bewertung der therapeutischen Beziehung, die wahrgenommene Nützlichkeit der Behandlung und die globale Erfolgseinschätzung (IQTIG, 2022).

Die Summe dieser Datenpunkte, aggregiert über alle abgeschlossenen Fälle einer Praxis, ergibt einen "Qualitäts-Score", der den Vergleich zwischen Praxen

ermöglichen soll. Qualität wird so zu einem relativen Wert, der sich aus dem statistischen Vergleich ergibt. Der fundamentale Unterschied liegt somit in der Quelle und der Natur der Evidenz. Im alten System war die Evidenz die logische Stringenz einer professionellen Argumentation (Hauten & Jungclaussen, 2022). Im neuen System ist die Evidenz ein numerischer Wert, generiert aus der Aggregation standardisierter Antworten (G-BA, 2024a).

Der Fokus verschiebt sich von der inhaltlichen Tiefe der Fallkonzeption zur formalen Korrektheit der Prozessdokumentation und der quantifizierten Zufriedenheit des Leistungsempfängers. Dieser Paradigmenwechsel ist nicht wertneutral. Wie die umfassende Kritik aus der Fachwelt zeigt (vgl. Volz & Benecke, 2024; BPtK, 2024), impliziert er eine Hierarchie, in der messbare Indikatoren einen höheren Stellenwert erhalten als nicht-standardisierbare Aspekte der therapeutischen Kunst. Die Gefahr, die von diesem Wechsel ausgeht und die den Kern dieses Artikels bildet, ist, dass das System am Ende genau das misst, was am einfachsten zu messen ist – und nicht unbedingt das, was für eine erfolgreiche Therapie am wichtigsten ist. Und schlimmer noch: dass es Therapeuten dazu zwingt, ihre Arbeit an diesen leicht messbaren, aber oft oberflächlichen Kriterien auszurichten, zum Nachteil der komplexen, aber entscheidenden Beziehungs- und Prozessarbeit.

1.2 Kernthese: Wie ein gut gemeintes QS-System die Versorgung der Schwächsten untergräbt

Die Intention, die dem neuen Qualitätssicherungsverfahren zugrunde liegt, ist auf den ersten Blick nachvollziehbar und politisch anschlussfähig: Die psychotherapeutische Versorgung soll transparenter, ihre Qualität messbar und die Position der Patientinnen und Patienten durch vergleichende Informationen gestärkt werden. Doch die vorliegende Analyse wird darlegen, dass das IQTIG-Verfahren in seiner aktuellen Ausgestaltung einem fundamentalen Konstruktionsfehler unterliegt, der diese lobenswerten Ziele nicht nur verfehlt, sondern ins Gegenteil zu verkehren droht.

Die Kernthese dieses Artikels lautet daher: Das vom IQTIG konzipierte QS-Verfahren ist aufgrund gravierender wissenschaftlich-methodischer Mängel – insbesondere des Fehlens einer validen Risikoadjustierung – ungeeignet, die Qualität psychotherapeutischer Versorgung abzubilden. Stattdessen schafft es systematische Fehlanreize, die eine Kultur der Risikoselektion fördern. Es bestraft

jene Therapeutinnen und Therapeuten, die sich der Behandlung von Patientinnen und Patienten mit schweren und komplexen Störungsbildern widmen, und setzt sie einem untragbaren ökonomischen und reputativen Risiko aus. Die rationale und ökonomisch zwangsläufige Konsequenz dieses Systemdesigns ist die zukünftige Ausgrenzung vulnerabler Patientengruppen aus der ambulanten Regelversorgung. Damit untergräbt ein System, das Qualität sichern soll, aktiv das Fundament der Versorgungsgerechtigkeit und gefährdet die Versorgung derjenigen, die sie am dringendsten benötigen.

Diese These stützt sich auf eine Kette von Argumenten, die im Folgenden detailliert entfaltet werden:

1. Die wissenschaftliche Untauglichkeit: Das System versucht, mit einem fehlerhaften Messinstrument (methodisch schwache Fragebögen, Verzicht auf Baseline-Daten) eine hochkomplexe Realität zu erfassen. Es kann daher keine validen Qualitätsaussagen generieren (Volz & Benecke, 2024).
2. Die perverse Anreizstruktur: Da das System die unterschiedliche Schwere der behandelten Patientengruppen nicht adäquat berücksichtigt, werden klinisch erwartbare Schwierigkeiten bei der Behandlung komplexer Fälle (z.B. höhere Abbruchraten, negative Bewertungen) fälschlicherweise als Mängel in der Behandlungsqualität interpretiert (Benecke et al., 2025).
3. Die rationale Konsequenz der Risikoselektion: Konfrontiert mit dem Risiko negativer öffentlicher Bewertungen und potenzieller Sanktionen, werden Therapeutinnen und Therapeuten rational handeln und die Behandlung von "Risikofällen" meiden, um ihre Praxis zu schützen.

Das Ergebnis ist nicht mehr Qualitätssicherung, sondern eine systematische, vom System selbst erzwungene Selektion "einfacher" Fälle – mit verheerenden Folgen für die Versorgung der Schwächsten.

1.3 Aufbau der kritischen Analyse

Um diese Kernthese zu untermauern und die komplexe Problematik transparent zu machen, gliedert sich die vorliegende Analyse wie folgt:

1. Kapitel 2 wird zunächst das auslaufende Antrags- und Gutachterverfahren beleuchten. Die Analyse seiner Doppelfunktion als Instrument der

Qualitätssicherung und, entscheidender noch, als ökonomischer Schutzschirm vor Regressforderungen schafft die notwendige Baseline, um zu verstehen, was mit der Reform auf dem Spiel steht.

2. Kapitel 3 widmet sich der detaillierten Anatomie des neuen IQTIG-QS-Verfahrens. Es werden die gesetzlichen Grundlagen, die konzipierten Instrumente, der geplante Datenfluss sowie die vorgesehenen Konsequenzen sachlich dargestellt, um eine fundierte Grundlage für die anschließende Kritik zu schaffen.
3. Kapitel 4 entfaltet den ersten Pfeiler der Kritik und weist die fundamentalen Konstruktionsfehler des Systems aus wissenschaftlich-methodischer Perspektive nach. Hier wird die Untauglichkeit der Instrumente und der Methodik im Detail belegt.
4. Kapitel 5 bildet das Herzstück der Argumentation und analysiert die reale Gefahr der Risikoselektion. Anhand des Beispiels der Behandlung von Persönlichkeitsstörungen wird gezeigt, wie das System klinische Realitäten ignoriert und die Behandlung schwerer Fälle zu einem untragbaren "wirtschaftlichen Ruin" für Praxen machen kann.
5. Kapitel 6 hebt die Analyse über eine reine Kritik hinaus, indem es konstruktive Gegenentwürfe aus der Profession vorstellt. Ansätze wie das Routine Outcome Monitoring (ROM) und das QSP/QVA-Projekt demonstrieren, wie eine klinisch nützliche und faire datengestützte Qualitätssicherung aussehen kann.
6. Abschließend wird der Artikel in Kapitel 7 ein Fazit ziehen, die Ergebnisse bündeln und aus der kritischen Analyse konkrete Forderungen für die anstehende Evaluation der Erprobungsphase in Nordrhein-Westfalen ableiten, um eine Kurskorrektur hin zu einer versorgungsgerechten und wissenschaftlich fundierten Qualitätssicherung zu ermöglichen.

2 Die Baseline: Was mit dem Gutachterverfahren verloren geht

Die politische Entscheidung, das Antrags- und Gutachterverfahren abzuschaffen, wurde maßgeblich mit dem Argument der Entbürokratisierung und Modernisierung begründet (BPtK, 2019). In der öffentlichen und oft auch in der fachlichen Debatte wurde das Verfahren häufig auf seine wahrgenommenen Nachteile reduziert: ein zeitaufwendiger, den Therapiebeginn verzögernder Prozess, ein "bürokratisches Monster" mit fragwürdiger Steuerungswirkung angesichts einer geringen Ablehnungsquote von nur 3-4% (Wittmann et al., 2011). Diese Darstellung ist jedoch eine verkürzte, wenn nicht gar irreführende Charakterisierung, die einer tiefergehenden Analyse nicht standhält. Sie ignoriert, dass die geringe Ablehnungsquote nicht zwangsläufig die Ineffizienz des Systems belegt, sondern im Gegenteil als Indikator für dessen Erfolg gewertet werden kann: Das Verfahren zwang im Vorfeld zu einer derart hohen Qualität der Antragsstellung und Fallkonzeption, dass nur wenige Berichte den fachlichen Kriterien nicht genügten. Es wirkte also primär präventiv und qualitätsfördernd, nicht selektiv und aussortierend.

Die Reduktion des Systems auf einen reinen Verwaltungsakt verkennt seine tiefgreifende, systemstabilisierende Funktion, die es über Jahrzehnte für die ambulante psychotherapeutische Versorgung innehatte. Es handelte sich um ein sorgfältig austariertes Gefüge, das fachliche Ansprüche mit ökonomischen Realitäten verband und so einen Rahmen schuf, in dem auch die Behandlung der komplexesten und ressourcenintensivsten psychischen Erkrankungen im ambulanten Sektor möglich war. Um die Risiken des neuen Systems vollständig bewerten zu können, ist es daher unerlässlich, zunächst eine präzise Baseline zu definieren: Was genau geht mit der Abschaffung des Gutachterverfahrens verloren? Eine sorgfältige Analyse offenbart eine komplexe Doppelfunktion, die weit über reine Bürokratie hinausging und sowohl die professionelle Qualität als auch die ökonomische Tragfähigkeit der Behandlung schwerer psychischer Erkrankungen sicherte.

2.1 Mehr als Bürokratie: Die Doppelfunktion als Peer-Review und ökonomischer Schutzschirm

Die Stärke und zugleich die Komplexität des Gutachterverfahrens lag in der unauflöslichen, symbiotischen Verschränkung zweier zentraler Aufgaben: einer qualitativen Qualitätssicherung durch kollegiale Begutachtung und einer

ökonomischen Absicherung durch eine vorgezogene Wirtschaftlichkeitsprüfung. Diese beiden Aspekte waren keine getrennten Säulen, sondern bedingten sich gegenseitig: Die hohe fachliche Anforderung an den Bericht (Qualitätssicherung) legitimierte die Bewilligung einer potenziell überdurchschnittlichen Behandlungsdauer und aktivierte so deren finanzielle Absicherung (ökonomischer Schutz).

Auf der ersten Ebene fungierte das Verfahren als ein anspruchsvolles Instrument des Peer-Reviews, das in seiner Tiefe und seinem Anspruch in der vertragsärztlichen Versorgung einzigartig war. Der Dreh- und Angelpunkt dieses Prozesses war die Verpflichtung des Therapeuten, vor Beginn einer Langzeittherapie einen umfassenden, narrativen Bericht zu verfassen. Diese intellektuelle und klinische Anstrengung war, wie Hauten und Jungclaussen (2022) treffend beschreiben, weit mehr als das Ausfüllen eines Formulars; sie war ein "institutionalisierter Zwang zur Professionalisierung". Der Therapeut war angehalten, die oft diffuse, fragmentierte und widersprüchliche Fülle klinischer Informationen – biographische Anamnese, komplexe Symptomatik, subtile interpersonelle Muster, Beziehungsdynamik – zu einer kohärenten, theoriegeleiteten Fallkonzeption zu verdichten. Dieser Prozess kultivierte explizit jene Kernkompetenzen, die für eine anspruchsvolle psychotherapeutische Arbeit unerlässlich sind. Er erforderte eine systematische Fallkonzeption, bei der statt eines rein pragmatischen oder symptomorientierten Vorgehens eine fundierte Hypothese über die zugrundeliegende Psychodynamik oder die lerngeschichtlichen Bedingungen der Störung entwickelt werden musste. Der Behandlungsplan war somit keine lose Ansammlung von Interventionen, sondern leitete sich logisch aus diesem tiefgehenden Fallverständnis ab. Der Therapeut musste argumentativ darlegen, warum für diesen spezifischen Patienten mit seiner einzigartigen Geschichte und Struktur eine bestimmte therapeutische Strategie indiziert ist.

Gleichzeitig erzwang der Schreibprozess eine intensive professionelle Selbstreflexion. Er nötigte den Therapeuten, aus der unmittelbaren Involviertheit und den emotionalen Strömungen der therapeutischen Beziehung herauszutreten und eine Metaperspektive, eine therapeutische Vogelperspektive, auf den eigenen Fall einzunehmen. Eigene "blinde Flecken", unklare diagnostische Einschätzungen oder intensive emotionale Reaktionen (Gegenübertragung), die im direkten Kontakt oft nur diffus spürbar sind, wurden im Prozess des Formulierens und Verschriftlichens oft

erst greifbar, bewusst und damit einer Reflexion zugänglich gemacht (vgl. DGPT, 2019). Der Bericht war somit auch ein Instrument der Psychohygiene und der Selbstsupervision für den Behandler. Die abschließende kollegiale Überprüfung durch erfahrene, von den Kassenärztlichen Vereinigungen bestellte und unabhängige Gutachter stellte schließlich sicher, dass Indikationsstellung, Verfahrenswahl und geplanter Behandlungsumfang fachlichen Standards entsprachen. Hierbei handelte es sich nicht um eine bürokratische Checklisten-Prüfung, sondern um einen fachlichen Diskurs auf Augenhöhe. Der Gutachter beurteilte die klinische Stimmigkeit, die innere Logik und die Plausibilität der Argumentation. Es war eine Form der "Qualität im Dialog" (Moeslein-Teising, zitiert in DGPT, 2019), die insbesondere bei komplexen und prognostisch schwierigen Fällen eine korrigierende und qualitätsfördernde Funktion hatte.

Diese anspruchsvolle fachliche Auseinandersetzung war jedoch kein Selbstzweck. Sie diente zugleich als argumentatives Fundament für die zweite, ökonomisch überlebenswichtige Funktion des Verfahrens: die des ökonomischen Schutzschirms. Die Genehmigung eines Gutachterantrags war nicht nur eine fachliche Bestätigung, sondern ein juristisch und ökonomisch bindender Akt zwischen Therapeut, Krankenkasse und Patient. Sie stellte eine vorgezogene Wirtschaftlichkeitsprüfung dar (Hauten & Jungclaussen, 2022). Mit der Bewilligung eines bestimmten Stundenkontingents wurde die Notwendigkeit und Angemessenheit der geplanten Behandlung im Vorhinein anerkannt. Dies schuf für alle Beteiligten eine verlässliche Planungs- und Rechtssicherheit. Für den Therapeuten war diese Sicherheit von existenzieller Bedeutung, denn sie schützte ihn vor dem Damoklesschwert nachträglicher Regressforderungen. Ohne diese Vorab-Genehmigung unterliegt die psychotherapeutische Behandlung der allgemeinen Wirtschaftlichkeitsprüfung nach § 106 SGB V. Dieses Instrument der kassenärztlichen Selbstverwaltung prüft im Nachhinein, oft Jahre später und meist auf Basis von statistischen Durchschnittswerten und Auffälligkeitsprüfungen, ob eine Behandlung "ausreichend, zweckmäßig und wirtschaftlich" war und das Maß des Notwendigen nicht überschritten hat.

Genau hier liegt das ökonomische Minenfeld, das durch das Gutachterverfahren entschärft wurde: Die Behandlung von Patientinnen und Patienten mit schweren, chronifizierten Störungen – wie Persönlichkeitsstörungen, komplexen Traumafolgestörungen oder schweren chronischen Depressionen – ist per

definitionem aufwendiger, krisenhafter und langwieriger als der statistische Durchschnitt. Sie sprengt zwangsläufig den Rahmen dessen, was in einer rein statistischen Betrachtung, die auf dem Prinzip der Gauß'schen Normalverteilung beruht, als "wirtschaftlich" gilt. Ein Therapeut, der sich auf solche Fälle spezialisiert, wird unweigerlich statistisch auffällig. Das Gutachterverfahren bot den entscheidenden Mechanismus, um diese notwendige, klinisch indizierte Abweichung vom statistischen Mittel fallindividuell zu legitimieren und abzusichern. Der detaillierte Bericht war die argumentative Grundlage, die bewies, warum der "statistische Durchschnittsmensch" hier als Maßstab versagen musste und warum dieser spezifische Fall mehr Ressourcen benötigte. Die Genehmigung durch den Gutachter war der offizielle Stempel, der diese "legitimierte Ausnahme von der statistischen Regel" bestätigte. Es schuf ein juristisches und ökonomisches Sicherheitsnetz, in dem auch ressourcenintensive Therapien ohne die ständige Sorge vor einem finanziellen Rückgriff der Kassen Jahre später durchgeführt werden konnten. Dieser Schutz war die ökonomische Grundvoraussetzung dafür, dass sich niedergelassene Therapeuten in einem freiberuflichen Umfeld überhaupt bereitfanden, das hohe Engagement und das persönliche unternehmerische Risiko einzugehen, das mit der Behandlung der am schwersten erkrankten Patientengruppen verbunden ist.

Zusammenfassend lässt sich festhalten: Das Gutachterverfahren war ein komplexes, integriertes System, das fachliche Reflexion und ökonomische Stabilität symbiotisch miteinander verknüpfte. Die hochwertige Fallkonzeption (die Qualitätssicherungsfunktion) lieferte die Begründung für die Genehmigung einer überdurchschnittlich langen Behandlung und aktivierte so den ökonomischen Schutz. Seine Abschaffung als reine "Entbürokratisierung" zu bezeichnen, ist eine gefährliche Verharmlosung. Es bedeutet den Wegfall eines etablierten Peer-Review-Prozesses und, noch gravierender, die Demontage des entscheidenden ökonomischen Schutzschirms, der die Versorgung vulnerabler Patientengruppen im ambulanten Sektor überhaupt erst tragfähig und für den einzelnen Therapeuten kalkulierbar gemacht hat. Das Verständnis dieser verlorenen Baseline ist die entscheidende Voraussetzung, um die katastrophalen Folgen zu antizipieren, die das neue, rein datengestützte System für die Versorgung dieser Patientengruppe haben wird.

2.2 Das politische Ende: Eine Reform ohne fachliche Debatte

Die Demontage eines so tief im Versorgungssystem verankerten Instruments wie des Gutachterverfahrens erfolgte auf eine Weise, die von vielen maßgeblichen Akteuren der Profession nicht nur als überraschend, sondern als prozeduraler Eklat und demokratiepolitisch fragwürdig empfunden wurde. Die Abschaffung war kein Ergebnis einer langen, transparenten und evidenzbasierten Auseinandersetzung über die beste Form der Qualitätssicherung. Stattdessen wurde sie im Herbst 2019 im Windschatten einer anderen, weitaus präsenteren Reform – des Gesetzes zur Reform der Psychotherapeutenausbildung – durchgesetzt.

Der legislative Prozess selbst ist bezeichnend: Die entscheidende Formulierung, die das Ende des Gutachterverfahrens besiegelte, fand sich nicht im ursprünglichen Gesetzentwurf, der monatelang diskutiert worden war. Sie wurde durch einen Änderungsantrag der damaligen GroKo-Fraktionen erst wenige Tage vor der finalen Abstimmung im Bundestag in das Gesetzespaket eingefügt. Dieses Vorgehen verhinderte effektiv eine angemessene parlamentarische und vor allem fachliche Auseinandersetzung mit den weitreichenden Folgen dieser spezifischen Änderung. Eine dedizierte Expertenanhörung, in der die Vor- und Nachteile sowie die systemischen Konsequenzen des Wegfalls hätten erörtert werden können, fand zu diesem Punkt nicht statt.

Die Reaktionen aus der Profession spiegeln die Überrumpelung wider. Die Bundespsychotherapeutenkammer (BPtK), die Spitzenorganisation der Psychotherapeutenschaft, zeigte sich öffentlich "überrascht" von diesem "substanziellen Eingriff ohne ausreichende Beratung" (BPtK, 2019). Lediglich der Deutschen Gesellschaft für Psychoanalyse, Psychotherapie, Psychosomatik und Tiefenpsychologie (DGPT) gelang es in letzter Minute, eine kritische Stellungnahme einzureichen, die jedoch im politischen Prozess kaum noch Gehör fand (DGPT, 2019). Von vielen Berufsvertretern wurde das Vorgehen daher als ein klassisches "Omnibusgesetz"-Manöver wahrgenommen: Eine hochkontroverse und potenziell unpopuläre Maßnahme wird in einem großen, in seinen Hauptzügen aber konsensfähigen Gesetzespaket "versteckt", um den politischen Widerstand zu minimieren und eine schnelle Verabschiedung ohne störende Detaildebatten zu gewährleisten.

Dieses Vorgehen ist mehr als nur eine politische Formalie. Es offenbart ein problematisches Politikverständnis, das die Expertise und die Bedenken einer gesamten Berufsgruppe bei einer für sie existenziellen Frage ignoriert. Anstatt einen partizipativen Prozess zu gestalten, in dem gemeinsam nach der besten Lösung für die zukünftige Qualitätssicherung gesucht wird, wurde eine Top-Down-Entscheidung getroffen und vollzogen. Dieser Akt der politischen Missachtung hat das Vertrauen der Profession in die Gesundheitspolitik nachhaltig erschüttert und von Anfang an eine Atmosphäre des Misstrauens gegenüber dem Nachfolgesystem geschaffen, das nun als eine von außen oktroyierte, nicht aber als eine gemeinsam entwickelte Lösung wahrgenommen wird.

2.3 Die Bedeutung des verlorenen Regress-Schutzes für die Behandlung komplexer Fälle

Die vielleicht gravierendste und in ihren praktischen Auswirkungen verheerendste Konsequenz der Abschaffung des Gutachterverfahrens ist der ersatzlose Wegfall des ökonomischen Schutzschirms. Wie in Kapitel 2.1 dargelegt, funktionierte das Verfahren als eine vorgezogene Wirtschaftlichkeitsprüfung, die Therapeuten vor nachträglichen Regressforderungen schützte. Das Verständnis dieses Verlustes ist der Schlüssel zum Verständnis der nun drohenden Versorgungskrise für vulnerable Patientengruppen.

Mit dem Ende des Gutachterverfahrens fällt die ambulante Psychotherapie unter die allgemeine, retrospektive Wirtschaftlichkeitsprüfung nach § 106 SGB V. Dieser Mechanismus funktioniert nach einer fundamental anderen, für die Psychotherapie hochproblematischen Logik. Er prüft nicht prospektiv und fallindividuell, sondern retrospektiv und statistisch. In der Regel werden die abgerechneten Leistungen einer Praxis (z.B. die durchschnittliche Stundenzahl pro Behandlungsfall) mit dem Durchschnitt aller Praxen derselben Fachgruppe in einer Region verglichen. Praxen, deren Werte signifikant ("auffällig") über diesem Durchschnitt liegen, geraten ins Visier einer Prüfung. Im Falle einer solchen Prüfung muss der Therapeut dann, oft Jahre nach Abschluss der Behandlungen, die medizinische Notwendigkeit und Wirtschaftlichkeit seiner Vorgehensweise nachträglich beweisen. Gelingt dies aus Sicht der Prüfungsgremien nicht, können Honorare in erheblicher Höhe zurückgefordert werden.

Genau hier entsteht ein unauflösbarer Widerspruch zur klinischen Realität der Behandlung komplexer psychischer Erkrankungen. Patientinnen und Patienten mit schweren Persönlichkeitsstörungen, komplexen Posttraumatischen Belastungsstörungen, schweren chronischen Depressionen oder multiplen Komorbiditäten sind per definitionem statistische Ausreißer. Ihre Behandlung erfordert mehr Zeit, mehr Kriseninterventionen, mehr Geduld und in der Regel eine deutlich höhere Stundenzahl als die eines durchschnittlichen Falles von beispielsweise einer unkomplizierten Angststörung. Ein Therapeut, der sich auf diese Patientengruppen spezialisiert oder einen signifikanten Anteil davon in seiner Praxis behandelt, wird zwangsläufig und unausweichlich statistisch auffällig. Seine Behandlungskosten und -dauern werden systematisch über dem Durchschnitt liegen.

Das Gutachterverfahren war das einzige im System verankerte Instrument, um diese statistische Auffälligkeit im Vorhinein zu legitimieren. Der ausführliche Bericht diente als Beweisführung, warum der Standard in diesem Einzelfall nicht ausreicht. Die Genehmigung war die offizielle Anerkennung dieser Notwendigkeit und schuf damit juristische und ökonomische Sicherheit ("Rechtssicherheit").

Mit dem Verlust dieses Instruments operiert der Therapeut nun in einem Zustand permanenter ökonomischer Unsicherheit. Jeder Behandlungsfall, der absehbar langwierig und ressourcenintensiv wird, verwandelt sich von einer therapeutischen Herausforderung in ein unternehmerisches Risiko. Der Therapeut mag nach bestem Wissen und Gewissen, leitliniengerecht und mit hohem Engagement eine notwendige und letztlich auch für das Gesundheitssystem kostensparende Behandlung durchführen – und lebt doch mit der latenten Gefahr, Jahre später von einem statistischen Prüfmechanismus dafür bestraft zu werden.

Dieser Zustand erzeugt einen massiven psychologischen und ökonomischen Druck, der als "Chilling Effect" bekannt ist: Allein die Androhung einer potenziellen Prüfung und eines möglichen Regresses genügt, um das Verhalten zu ändern. Es fördert eine defensive, risikoscheue Haltung. Die Bereitschaft, sich auf prognostisch unsichere, krisenhafte und langwierige Behandlungen einzulassen, sinkt dramatisch, wenn diese mit dem Risiko des persönlichen wirtschaftlichen Ruins verbunden sind. Die Abschaffung des Gutachterverfahrens war somit kein Akt der "Entbürokratisierung", sondern die Verlagerung des gesamten ökonomischen Risikos für die Behandlung der schwersten psychischen Erkrankungen von den Schultern des Systems auf die

Schultern des einzelnen freiberuflichen Therapeuten. Dies hat die Anreizstrukturen im Versorgungssystem fundamental und auf gefährliche Weise verschoben.

3 Anatomie des neuen Systems: Das IQTIG-QS-Verfahren im Detail

Nachdem die politische Entscheidung zur Ablösung des Gutachterverfahrens getroffen war, begann die komplexe und kontroverse Aufgabe, ein adäquates Nachfolgesystem zu konzipieren. Dieser Prozess wurde nicht der Profession selbst überlassen, sondern im Rahmen der etablierten Strukturen der Gemeinsamen Selbstverwaltung im Gesundheitswesen angesiedelt. Das Ergebnis ist ein hochgradig formalisiertes, datengestütztes System, dessen Architektur, Instrumente und Mechanismen im Folgenden detailliert seziert werden. Das Verständnis seiner "Anatomie" ist die unabdingbare Voraussetzung, um seine Funktionsweise, seine inhärenten Schwächen und seine potenziell gefährlichen Auswirkungen auf die psychotherapeutische Versorgungslandschaft zu begreifen.

3.1 Der gesetzliche Auftrag: Ziele und Rahmen der DeQS-Richtlinie

Die gesetzliche Grundlage für das neue Qualitätssicherungsverfahren findet sich im § 136a des Fünften Buches Sozialgesetzbuch (SGB V), der durch das Psychotherapeutenausbildungsreformgesetz von 2019 novelliert wurde. Dieser Paragraph beauftragte den Gemeinsamen Bundesausschuss (G-BA) – das höchste Gremium der gemeinsamen Selbstverwaltung von Ärzten, Psychotherapeuten, Krankenhäusern und Krankenkassen in Deutschland – mit der Entwicklung eines "einrichtungsübergreifenden, sektorspezifischen Qualitätssicherungsverfahrens für die ambulante psychotherapeutische Versorgung". Mit der konkreten Ausarbeitung und Entwicklung dieses komplexen Systems wurde das Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG) betraut, das als quasi-wissenschaftlicher Dienstleister für den G-BA agiert.

Entscheidend für das Verständnis des neuen Verfahrens ist, dass es kein isoliertes, allein für die Psychotherapie geschaffenes Konstrukt ist. Es wurde vielmehr in eine bereits bestehende, übergeordnete Architektur der medizinischen Qualitätssicherung integriert: die Richtlinie zur datengestützten einrichtungsübergreifenden Qualitätssicherung (DeQS-RL) (G-BA, 2024a). Diese Richtlinie bildet den formalen und methodischen Rahmen für eine wachsende Zahl von QS-Verfahren in der deutschen Medizin, die bisher vor allem hochspezialisierte, somatische Eingriffe wie Nierentransplantationen oder den Einsatz von Herzkathetern betrafen. Die Einbettung der Psychotherapie in diese primär somatisch-technisch geprägte Logik ist bereits ein erster Hinweis auf den

Paradigmenwechsel: Die prozessuale, intersubjektive und narrative Komplexität der Psychotherapie wird denselben formalen Mess- und Vergleichbarkeitsansprüchen unterworfen wie ein standardisierter chirurgischer Eingriff.

Die DeQS-Richtlinie formuliert verbindliche Ziele, die auch für das Psychotherapie-Verfahren (offiziell "Verfahren 16: QS ambulante Psychotherapie") gelten. Von besonderer Relevanz sind hierbei zwei Hauptziele: Das Verfahren soll erstens "valide und vergleichbare Erkenntnisse über die Versorgungsqualität der Leistungserbringerinnen und Leistungserbringer" gewinnen und zweitens das "einrichtungsinterne Qualitätsmanagement unterstützen" (DeQS-RL § 1, Ziel c & e). Diese Zielsetzung offenbart die doppelte Stoßrichtung des Systems: nach außen gerichtete Transparenz und Vergleichbarkeit ("Benchmarking") und nach innen gerichtete Anstöße zur Qualitätsverbesserung. Wie sich zeigen wird, liegt der Schwerpunkt des IQTIG-Designs jedoch eindeutig auf dem ersten, dem vergleichenden Aspekt.

Der Geltungsbereich des Verfahrens ist präzise definiert: Es erfasst alle im Rahmen der gesetzlichen Krankenversicherung durchgeführten ambulanten Richtlinien-Psychotherapien (Analytische Psychotherapie, Tiefenpsychologisch fundierte Psychotherapie, Systemische Therapie und Verhaltenstherapie) bei erwachsenen Patientinnen und Patienten ab 18 Jahren. Sowohl Kurzzeit- als auch Langzeittherapien sind eingeschlossen. Ausgenommen sind hingegen Behandlungen von Kindern und Jugendlichen sowie andere psychotherapeutische Leistungen wie die Akutbehandlung oder die psychotherapeutische Sprechstunde.

3.2 Die Messinstrumente: Therapeutendokumentation und Patientenfragebogen

Das Herzstück des IQTIG-QS-Verfahrens bilden zwei neu entwickelte Datenerhebungsinstrumente, die nach Abschluss jeder erfassten Therapie zum Einsatz kommen. Die Datenerhebung ist somit rein retrospektiv und findet zu einem Zeitpunkt statt, an dem der konkrete Behandlungsprozess bereits beendet ist.

Die Therapeutendokumentation: Ein Audit der Prozesskonformität

Nach Beendigung einer Therapie ist die behandelnde Psychotherapeutin bzw. der behandelnde Psychotherapeut verpflichtet, einen umfangreichen, standardisierten und digitalen Fragebogen auszufüllen. Dieser umfasst, je nach Zählweise der Unter-

Items, rund 101 Datenfelder (IQTIG, 2022). Dieser enorme Umfang macht bereits die bürokratische Last des neuen Systems deutlich. Die Dokumentation ist logisch in drei Phasen der Behandlung gegliedert, die im Folgenden detailliert beleuchtet werden:

1. Die Anfangsphase der Behandlung: Hier wird primär die Einhaltung formaler und diagnostischer Standards zu Beginn der Therapie auditiert. Die Fragen sind überwiegend im Ja/Nein-Format gehalten und prüfen, ob bestimmte Handlungen stattgefunden haben. Konkrete Beispiele sind: "Wurde die Symptomatik mittels standardisiertem Instrument erfasst?", "Wurde eine Verdachtsdiagnose nach ICD-10 gestellt?", "Wurde abgeklärt, ob eine Suizidalität vorliegt?", "Wurden patientenindividuelle Therapieziele vereinbart?". Diese Fragen zielen nicht auf die Qualität, die Tiefe oder die klinische Angemessenheit dieser Handlungen ab, sondern ausschließlich auf deren dokumentiertes Vorhandensein.
2. Der Therapieverlauf: Dieser Abschnitt fokussiert auf das Monitoring und die Anpassung des therapeutischen Prozesses. Auch hier dominieren formale Prüffragen wie: "Wurden die patientenindividuellen Therapieziele im Verlauf überprüft?", "Wurden die Therapieziele erreicht oder angepasst?". Die Komplexität therapeutischer Verläufe, die oft von Krisen, Stagnationen und nicht-linearen Entwicklungen geprägt sind, wird auf diese dichotomen Abfragen reduziert.
3. Die Abschlussphase der Behandlung: Der letzte Teil prüft die formal korrekte Beendigung der Therapie. Es wird beispielsweise gefragt, ob das Erreichen oder Nichterreichen der Ziele mit dem Patienten besprochen wurde oder ob Maßnahmen zur Rückfallprophylaxe thematisiert wurden.

Eine kritische Analyse der Therapeutendokumentation offenbart, dass es sich im Kern um eine Checkliste zur Überprüfung der Prozesskonformität handelt. Ein Großteil der abgefragten Punkte entspricht ohnehin den berufsrechtlichen Pflichten und den Empfehlungen zur Dokumentation, wie sie etwa von der Bundespsychotherapeutenkammer (BPTK, 2020) formuliert wurden. Das Instrument prüft, ob der Therapeut formal korrekt gehandelt hat, aber es kann und will nicht erfassen, wie gut, wie einfühlsam oder wie klinisch adäquat er dies getan hat. Es misst die Einhaltung eines standardisierten Protokolls, nicht die Qualität der

therapeutischen Kunst. Die Daten werden quartalsweise über die Software der Praxis an die jeweilige Kassenärztliche Vereinigung (KV) und von dort an eine zentrale Datenannahmestelle übermittelt.

Der Patientenfragebogen: Qualität als retrospektive Patientenzufriedenheit

Das zweite, und in der Fachwelt noch stärker kritisierte, Instrument ist ein ebenfalls vom IQTIG neu entwickelter Fragebogen für Patientinnen und Patienten. Nach dem regulären Abschluss einer Therapie wird dieser Fragebogen, der rund 43 Items umfasst, postalisch an den Patienten versandt (Benecke et al., 2025). Die Teilnahme ist für den Patienten freiwillig, sein Einverständnis zur Datenerhebung durch die Praxis im Vorfeld der Therapie ist jedoch obligatorisch. Ein Widerspruchsrecht gegen die Datenweitergabe von der Praxis an das IQTIG ist gesetzlich nicht vorgesehen, was datenschutzrechtlich und ethisch hoch problematisch ist (KV Nordrhein, 2024).

Der Fragebogen gliedert sich ebenfalls in mehrere inhaltliche Blöcke:

1. Informations- und Aufklärungs-Audit: Ähnlich der Therapeutendokumentation prüft der erste Teil aus Patientensicht, ob formale Aufklärungspflichten zu Beginn der Therapie erfüllt wurden. Fragen lauten beispielsweise, ob der Therapeut über die Häufigkeit der Sitzungen, über Behandlungsalternativen oder über mögliche "Nebenwirkungen" einer Psychotherapie informiert hat. Die Validität dieser retrospektiven Abfragen, die sich teils auf Gespräche beziehen, die Jahre zurückliegen, ist höchst fragwürdig.
2. Bewertung der therapeutischen Beziehung und Atmosphäre: Dieser Abschnitt zielt auf die subjektive Wahrnehmung der Beziehungsqualität ab. Hier finden sich Fragen zum Einfühlungsvermögen ("Hatten Sie den Eindruck, dass sich Ihre Psychotherapeutin/Ihr Psychotherapeut für Ihre Probleme und Sorgen interessiert?"), zur Präsenz ("Wirkte Ihre Psychotherapeutin/Ihr Psychotherapeut abwesend?") und zur Fähigkeit des Therapeuten, emotionale Belastungen aufzufangen. Berüchtigt geworden ist hier die negativ-suggestive Frage, ob sich der Patient "nach einer aufwühlenden Sitzung mit seinen Gefühlen allein gelassen" fühlte.
3. Subjektive Ergebnisbewertung (Outcome): Der letzte Teil erfasst die globale Einschätzung des Behandlungserfolgs durch den Patienten. Anstelle

validierter Prä-Post-Messungen von Symptomen oder Funktionsniveaus wird hier auf wenige, globale retrospektive Fragen gesetzt: "Geht es Ihnen durch die Psychotherapie alles in allem besser, schlechter oder unverändert?", "Wurden hilfreiche Erfahrungen, Fertigkeiten oder Strategien gewonnen?", "Wurden Ihre persönlichen Ziele erreicht?".

Auffällig ist, dass das IQTIG für diesen zentralen Teil des Verfahrens ein gänzlich neues, nicht extern validiertes Instrument entwickelt hat, anstatt auf Dutzende von international etablierten, psychometrisch rigoros geprüften Fragebögen zur Erfassung der therapeutischen Beziehung (z.B. WAI, HAQ) oder des Therapieerfolgs (z.B. BDI, SCL-90, OQ-45) zurückzugreifen. Dieser Bruch mit wissenschaftlichen Standards ist einer der Hauptkritikpunkte der Fachwelt (Volz & Benecke, 2024).

3.3 Der Mechanismus: Von der Datenerhebung zu öffentlichen Qualitätsberichten

Die Sammlung dieser umfangreichen Datensätze ist jedoch lediglich der erste Schritt in einem komplexen, zentralisierten Mechanismus, der darauf abzielt, individuelle therapeutische Erfahrungen in vergleichbare Qualitätsmetriken zu transformieren.

Datenfluss und zentrale Auswertung: Alle erhobenen Daten – sowohl aus der Therapeutendokumentation als auch aus den Patientenfragebögen – laufen bei einer zentralen Bundesauswertungsstelle zusammen, die beim IQTIG angesiedelt ist. Dort werden die Daten pseudonymisiert und zu sogenannten Qualitätsindikatoren verrechnet. Aus den Antworten des Patientenfragebogens werden beispielsweise Indikatoren wie "Patientenbeurteilung der therapeutischen Beziehung" oder "Patientenbeurteilung des Behandlungserfolgs" gebildet.

Der verzögerte Feedback-Bericht: Nach jedem zweijährigen Erfassungszeitraum erhalten die psychotherapeutischen Praxen einen aggregierten "Rückmeldebericht". Dieser Bericht enthält die Ergebnisse der eigenen Praxis für die verschiedenen Qualitätsindikatoren – allerdings nicht auf Ebene einzelner Fälle, sondern als Durchschnittswert über alle abgeschlossenen Behandlungen der letzten beiden Jahre. Diese aggregierten Werte werden dann im Sinne eines Benchmarkings den Durchschnittswerten einer Referenzgruppe (z.B. alle Praxen in derselben KV-Region) gegenübergestellt. Diese zeitliche Verzögerung und die hohe Aggregationsebene machen den Bericht für eine zeitnahe, fallbezogene Qualitätsverbesserung im

Praxisalltag praktisch unbrauchbar. Ein Therapeut erfährt im Jahr 2027 pauschal, wie seine Patienten im Durchschnitt der Jahre 2025 und 2026 die therapeutische Beziehung bewertet haben, ohne zu wissen, welche spezifischen Behandlungen problematisch waren.

Öffentliche Berichterstattung und Sanktionspotenzial: Die eigentliche Schlagkraft und das größte Drohpotenzial des Systems entfalten sich in den langfristig geplanten Konsequenzen. Gemäß der DeQS-Richtlinie ist die Veröffentlichung der einrichtungsbezogenen Vergleichsdaten in einem öffentlichen Qualitätsportal vorgesehen, ähnlich den bereits existierenden Klinik-Suchportalen. Patientinnen und Patienten sollen sich so zukünftig online über die "Qualität" einzelner Praxen informieren und diese bei ihrer Therapeutenwahl berücksichtigen können. Dies erzeugt einen enormen Rechtfertigungs- und Reputationsdruck für die Praxen.

Darüber hinaus sieht die DeQS-Richtlinie einen mehrstufigen Prozess bei "rechnerischer Auffälligkeit" vor. Weichen die Qualitätsindikatoren einer Praxis signifikant negativ vom Durchschnitt ab, wird ein "Stellungnahmeverfahren" eingeleitet. Die Praxis erhält die Möglichkeit, die auffälligen Ergebnisse zu erklären (z.B. durch einen Hinweis auf eine besondere Patientengruppe). Reicht diese Erklärung aus Sicht der zuständigen Gremien nicht aus, können Maßnahmen zur Qualitätsverbesserung angeordnet werden. Dies kann von verpflichtenden Nachschulungen bis hin zu strukturierten Dialogen reichen. Als ultima ratio ist auch eine Überleitung in eine Wirtschaftlichkeitsprüfung mit potenziellen Honorarkürzungen prinzipiell im System angelegt. Auch wenn Sanktionen in der Erprobungsphase zunächst ausgesetzt sind, schwebt dieses Sanktionspotenzial als Damoklesschwert über dem gesamten Verfahren.

3.4 Die Notbremse: Warum eine Erprobungsphase in NRW unumgänglich wurde

Der ursprüngliche Zeitplan des G-BA sah einen bundesweiten, verpflichtenden Start des IQTIG-QS-Verfahrens zum 1. Januar 2023 vor. Dieser Zeitplan erwies sich jedoch als politisch und praktisch undurchführbar. Der Grund dafür war eine Welle des massiven, geeinten und fachlich fundierten Widerstands aus der gesamten Profession. Nahezu alle relevanten Akteure – von der Bundespsychotherapeutenkammer (BPtK) über die großen Berufsverbände (DPtV, DGPT, bvvp, VAKJP) bis hin zu universitären Forschern – legten in detaillierten

Stellungnahmen die gravierenden methodischen und konzeptionellen Mängel des IQTIG-Entwurfs offen.

Die Kritik war so fundamental und die Bedenken hinsichtlich der negativen Auswirkungen auf die Patientenversorgung so gravierend, dass der politische Druck auf den G-BA immens wuchs. Die Forderung, ein derart unausgereiftes und potenziell schädliches System nicht ohne eine vorherige, gründliche Erprobung in die Regelversorgung zu entlassen, wurde unüberhörbar. In einem für die G-BA-Strukturen ungewöhnlichen Schritt gab das Gremium diesem Druck nach und zog die "Notbremse".

Statt des bundesweiten Starts wurde eine sechsjährige Erprobungsphase vom 1. Januar 2025 bis zum 31. Dezember 2030 beschlossen, die regional auf die KV-Bezirke Nordrhein und Westfalen-Lippe beschränkt ist (G-BA, 2024a). Diese Entscheidung war ein entscheidender, von den Berufsverbänden errungener Erfolg, der die sofortige bundesweite Implementierung verhinderte. Die Bedingungen dieser Erprobungsphase sind jedoch hart: Die teilnehmenden Therapeutinnen und Therapeuten in NRW sind mit einer erheblichen Doppelbelastung konfrontiert, da das alte Antrags- und Gutachterverfahren während der gesamten Erprobungszeit bundesweit parallel bestehen bleibt. Als Zugeständnis wurde vereinbart, dass für den ersten zweijährigen Erfassungszeitraum (2025-2026) keine Sanktionen aus dem Verfahren abgeleitet werden dürfen.

Diese Erprobungsphase ist somit mehr als nur ein technischer Testlauf. Sie ist das Resultat einer fundamentalen Auseinandersetzung um die Zukunft der Qualitätssicherung und stellt ein kritisches Zeitfenster dar. In diesen sechs Jahren muss sich zeigen, ob das IQTIG-System in der Lage ist, seine fundamentalen Konstruktionsfehler zu beheben, oder ob es sich als ungeeignet für die psychotherapeutische Versorgung erweist. Die nun folgenden Kapitel werden die Argumente, die zu dieser Notbremsung führten, im Detail entfalten und aufzeigen, warum eine grundlegende Revision des Systems während dieser Erprobungsphase unabdingbar ist.

4 Fundamentale Konstruktionsfehler: Die wissenschaftliche Untauglichkeit des IQTIG-Verfahrens

Während die politischen und administrativen Rahmenbedingungen des neuen QS-Verfahrens bereits Anlass zur Sorge geben, liegt der Kern der Auseinandersetzung in seiner wissenschaftlichen und methodischen Substanz. Ein datengestütztes System, das den Anspruch erhebt, die Qualität einer so komplexen Intervention wie der Psychotherapie zu messen und zu vergleichen, muss sich den höchsten wissenschaftlichen Gütekriterien stellen. Eine genaue Analyse des IQTIG-Designs offenbart jedoch, dass es in fundamentaler Weise gegen etablierte Standards der empirischen Sozialforschung und der psychometrischen Testtheorie verstößt.

Die Kritik, die von nahezu allen relevanten Fach- und Berufsverbänden in seltener Einmütigkeit geäußert wird (vgl. DPtV et al., 2022; BPtK, 2024), ist daher keine grundsätzliche Ablehnung von Datenerhebung, sondern eine fundierte Zurückweisung eines spezifischen, als untauglich erachteten Instruments. Die wissenschaftliche Analyse zeigt: Das IQTIG-Verfahren ist in seiner jetzigen Form nicht nur ein ungenaues, sondern ein potenziell systematisch verzerrendes Messinstrument, dessen Ergebnisse im besten Fall unbrauchbar und im schlimmsten Fall gefährlich irreführend sind. Diese Kritik lässt sich in drei ineinandergreifende Bereiche gliedern: die grundlegenden methodischen Mängel, die konzeptionellen Schwächen der Erhebungsinstrumente und die daraus resultierende negative Kosten-Nutzen-Bilanz.

4.1 Methodische Mängel: Warum die Ergebnisse nicht valide sind (Validität, Reliabilität, Baseline-Problem)

Die wissenschaftliche Güte jedes Messverfahrens bemisst sich an drei zentralen Kriterien: Objektivität (sind die Ergebnisse unabhängig vom Anwender?), Reliabilität (misst das Instrument zuverlässig und konsistent?) und vor allem Validität (misst das Instrument tatsächlich das, was es zu messen vorgibt?). In allen drei Bereichen weist das IQTIG-Verfahren gravierende Defizite auf.

Das Validitätsproblem: Die Verwendung nicht validierter Ad-hoc-Instrumente Der vielleicht größte methodische Sündenfall des IQTIG-Designs ist die Entscheidung, für das Herzstück des Verfahrens – den Patientenfragebogen – ein gänzlich neues Instrument zu entwickeln, anstatt auf den riesigen Fundus an international

etablierten, psychometrisch rigoros validierten Fragebögen zurückzugreifen (Benecke et al., 2025). Über Jahrzehnte wurden in der Psychotherapieforschung Instrumente zur Messung der therapeutischen Allianz (z.B. das Working Alliance Inventory, WAI), der Ergebnisqualität (z.B. das Outcome Questionnaire, OQ-45) oder spezifischer Symptombelastungen (z.B. das Beck Depression Inventory, BDI) entwickelt und in hunderten von Studien auf ihre Güte geprüft. Das IQTIG hat diesen wissenschaftlichen Konsens ignoriert und ein Ad-hoc-Instrument geschaffen, dessen Kriteriumsvalidität – also der empirische Beleg, dass seine Ergebnisse tatsächlich mit echter Behandlungsqualität korrelieren – bis heute nicht von unabhängiger Seite erbracht wurde (Volz & Benecke, 2024). Es wurde ein Messinstrument für den flächendeckenden, hochsensiblen Einsatz konzipiert, ohne den in der Wissenschaft obligatorischen Nachweis zu erbringen, dass es überhaupt funktioniert. Die BPtK (2024) urteilt daher vernichtend, die Instrumente seien "schlicht ungeeignet".

Das Baseline-Problem: Erfolgsmessung ohne Anfangswert Der zweite fundamentale Konstruktionsfehler ist der vollständige Verzicht auf eine standardisierte Erfassung des Zustands der Patientinnen und Patienten zu Beginn der Behandlung (Prä-Messung oder Baseline). Das IQTIG-Verfahren erfasst lediglich eine globale, retrospektive Selbsteinschätzung des Patienten am Ende der Therapie ("Geht es Ihnen besser?"). Dies widerspricht dem Goldstandard jeder seriösen Evaluations- und Therapieforschung. Ohne einen validen Ausgangswert ist eine objektive Messung von Veränderung schlicht unmöglich. Es ist, als würde man versuchen, den Erfolg eines Trainingsprogramms zur Gewichtsreduktion zu bewerten, indem man die Teilnehmer am Ende fragt, ob sie sich schlanker fühlen, ohne jemals ihr Ausgangsgewicht gewogen zu haben. Dieser Verzicht auf eine Baseline-Messung hat zwei verheerende Konsequenzen: Zum einen macht er eine objektive Bestimmung der Effektstärke der Behandlung unmöglich. Zum anderen, und das ist für die Fairness des Systems entscheidend, verunmöglicht er jede Form einer validen Risikoadjustierung, die im nächsten Kapitel detailliert wird.

Das Reliabilitäts- und Stabilitätsproblem: Die Tyrannei der kleinen Zahlen Selbst wenn die Instrumente valide wären, würde die statistische Aussagekraft der Ergebnisse für einzelne Praxen an den Realitäten der Versorgungslandschaft zerschellen. Die DeQS-Richtlinie sieht eine Auswertung auf Basis eines zweijährigen Erfassungszeitraums vor. Viele, insbesondere tiefenpsychologisch oder analytisch

arbeitende Praxen mit langen Behandlungsdauern, schließen in diesem Zeitraum jedoch nur eine relativ geringe Anzahl von Therapien ab. Geht man von einer in sozialwissenschaftlichen Erhebungen realistischen Rücklaufquote bei postalischen Befragungen von etwa 30% aus, entstehen für die Auswertung winzige Stichproben (DPtV et al., 2022). Hat eine Praxis beispielsweise 30 Therapien in zwei Jahren beendet, liegen am Ende möglicherweise nur die Fragebögen von 9 Patientinnen und Patienten vor. Auf einer derart schmalen Datenbasis lassen sich keine statistisch robusten, reliablen und fairen Aussagen über die "Qualität" einer ganzen Praxis treffen. Die Ergebnisse werden hochgradig anfällig für Zufallsschwankungen und die Verzerrung durch einzelne, extrem positive oder negative Bewertungen. Ein einziger, im Streit gegangener Patient kann die gesamte "Qualitätsbilanz" einer Praxis für zwei Jahre ruinieren. Ein Vergleich zwischen Praxen auf dieser Grundlage ist statistisch unseriös und gleicht einem Glücksspiel.

4.2 Konzeptionelle Schwächen: Suggestive Fragen, Erinnerungsverzerrung und die Messung von Konformität statt Qualität

Über die rein methodischen Mängel hinaus ist das Verfahren auch in seiner inhaltlichen Konzeption zutiefst problematisch. Es misst in weiten Teilen nicht die klinische Substanz, sondern formale Aspekte und ist dabei anfällig für systematische Verzerrungen.

Der Patientenfragebogen: Ein Minenfeld aus Verzerrungen und Fehlkonzeptionen
Die inhaltliche Gestaltung des Patientenfragebogens wird von Fachleuten als fachlich unausgereift und teils unsachgemäß kritisiert (Benecke et al., 2025). Ein zentrales Problem ist die Erinnerungsverzerrung (Recall Bias). Der Fragebogen verlangt von Patienten, sich teilweise Jahre nach Therapiebeginn an spezifische Details von Aufklärungsgesprächen zu erinnern. Die kognitive Psychologie hat umfassend belegt, dass solche retrospektiven Erinnerungen hochgradig unzuverlässig und von späteren Erfahrungen überformt sind. Die Behauptung des IQTIG, die Erinnerungsfähigkeit geprüft zu haben, indem lediglich die Selbsteinschätzung der Patienten über ihre Gedächtnisleistung erfragt wurde, ist methodisch naiv und ignoriert die wissenschaftliche Evidenz zu diesem Thema (Volz & Benecke, 2024). Hinzu kommen suggestive und unangemessene Fragestellungen. Zahlreiche Items sind negativ konnotiert oder legen eine bestimmte Antwort nahe

(z.B. "fühlten Sie sich mit Ihren Gefühlen alleine gelassen?"). Solche Formulierungen können Antworten in eine negative Richtung lenken und verletzen Grundprinzipien neutraler Fragebogenkonstruktion. Schließlich findet eine unzulässige Vermischung von Qualitätssicherung und rechtlicher Kontrolle statt. Ein großer Teil des Fragebogens prüft, ob der Therapeut formale Aufklärungspflichten nach dem Patientenrechtegesetz erfüllt hat. Dies ist nicht die Aufgabe eines QS-Instruments, sondern obliegt anderen Kontrollinstanzen.

Die Therapeutendokumentation: Die Messung von Konformität statt Qualität Auch die von den Therapeuten auszufüllende Dokumentation misst nicht primär Qualität, sondern bürokratische Konformität. Sie erfasst im Wesentlichen, ob bestimmte normative Vorgaben erfüllt wurden (Diagnostik durchgeführt, Ziele vereinbart etc.), nicht aber wie und mit welcher klinischen Angemessenheit dies geschah. Ein Therapeut kann alle Checklistenpunkte formal korrekt erfüllen und dennoch eine klinisch minderwertige, unempathische oder sogar schädliche Behandlung durchführen. Das Instrument ist blind für die prozessuale Tiefe, die Flexibilität, die intuitive Kompetenz und die Beziehungsgestaltung, die den Kern guter psychotherapeutischer Arbeit ausmachen. Es erfasst die leere Hülle des Prozesses, nicht seine lebendige Substanz. Das Verfahren kann somit die Komplexität und insbesondere die nicht-linearen, oft krisenhaften Verläufe von psychodynamischen Therapien oder Traumatherapien nicht abbilden.

4.3 Die Bilanz: Ein bürokratisches Monster ohne klinischen Nutzen

Zieht man eine Bilanz aus der wissenschaftlichen und konzeptionellen Analyse, ergibt sich ein verheerendes Bild. Das IQTIG-QS-Verfahren stellt einen enormen bürokratischen Aufwand dar, der wertvolle Zeit und Ressourcen aus der direkten Patientenversorgung abzieht (BPTK, 2024). Allein die Dokumentation von über 100 Datenfeldern pro Fall ist eine erhebliche Belastung für den Praxisalltag.

Diesem immensen Aufwand steht ein äußerst fragwürdiger, wenn nicht gänzlich fehlender klinischer Nutzen gegenüber. Die Ergebnisse und Rückmeldungen des Systems sind aufgrund der methodischen Mängel nicht valide. Aber selbst wenn sie es wären, kommen sie viel zu spät und in einer viel zu aggregierten Form, um für die Steuerung laufender Therapien oder für konkrete Lernprozesse im Sinne eines internen Qualitätsmanagements von Nutzen zu sein. Ein Therapeut erfährt Jahre später einen statistischen Durchschnittswert, der ihm keinerlei Anhaltspunkte für

die Verbesserung seiner konkreten Arbeit mit konkreten Patienten gibt. Das System scheitert somit an einem seiner eigenen, zentralen Ziele: der Unterstützung des einrichtungsinternen Qualitätsmanagements.

Im Ergebnis ist das IQTIG-Verfahren in seiner jetzigen Form ein "bürokratisches Monster", das enorme Kosten verursacht, aber keinen erkennbaren Mehrwert für die Verbesserung der Patientenversorgung schafft. Es ist ein wissenschaftlich untaugliches Instrument, das aufgrund seiner Konstruktionsfehler nicht in der Lage ist, seine Ziele zu erreichen. Schlimmer noch: Wie das folgende Kapitel zeigen wird, ist es nicht nur nutzlos, sondern aktiv schädlich, da es durch seine inhärenten Fehlanreize die Versorgung der Schwächsten im System direkt gefährdet.

5 Die reale Gefahr: Wie das System Risikoselektion erzwingt und Praxen ruiniert

Nachdem die fundamentalen Konstruktionsfehler und die wissenschaftliche Untauglichkeit des IQTIG-Verfahrens dargelegt wurden, könnte man argumentieren, das System sei lediglich ein zahnlöser Tiger – ein bürokratisches Ärgernis ohne realen Nutzen, aber auch ohne realen Schaden. Diese Einschätzung wäre fatal und würde die immanente Sprengkraft des Systems verkennen. Die methodischen Mängel sind keine rein akademischen Schönheitsfehler; sie sind der Zündmechanismus für eine tickende Zeitbombe im Versorgungssystem. Das IQTIG-Verfahren ist in seiner Konzeption nicht nur ungenau, sondern aktiv schädlich. Es schafft eine perverse Anreizstruktur, die professionell und ethisch gebotenes Handeln bestraft und risikovermeidendes, selektives Verhalten belohnt.

Dieses Kapitel wird zeigen, wie der systemische Designfehler der fehlenden Risikoadjustierung zwangsläufig zu einer systematischen Benachteiligung von Praxen führt, die sich der Behandlung schwerkranker Patientinnen und Patienten widmen. Es wird am Härtefall der Behandlung von Persönlichkeitsstörungen demonstriert, wie das System klinische Realitäten als Qualitätsmängel fehlinterpretiert. Schließlich wird das "Kalkül des Ruins" aufgezeigt – die rationale und ökonomisch unausweichliche Konsequenz für Therapeuten, diese Patientengruppen zukünftig zu meiden, was in einer neuen, hausgemachten Versorgungslücke für die vulnerabelsten Mitglieder der Gesellschaft mündet.

5.1 Der systemische Designfehler: Die fatale Wirkung fehlender Risikoadjustierung

Das konzeptionelle Kernproblem, das alle nachfolgenden Gefahren bedingt, ist das vollständige Fehlen einer validen, funktionierenden Risikoadjustierung (Volz & Benecke, 2024). In der medizinischen Qualitätssicherung ist die Risikoadjustierung ein unverzichtbarer Standard. Sie ist ein statistisches Verfahren, das sicherstellt, dass Vergleiche zwischen Leistungserbringern (z.B. Krankenhäusern oder Praxen) fair sind, indem es Unterschiede in der Ausgangslage der behandelten Patientengruppen herausrechnet. Ein einfaches Analogon ist der Sport: Niemand würde die Leistung eines Boxers im Leichtgewicht fair mit der eines Schwergewichtsboxers vergleichen, ohne die Gewichtsklasse zu berücksichtigen. Ebenso kann die Qualität der Arbeit eines Chirurgen, der hauptsächlich hochkomplexe Notfalloperationen an

multimorbiden alten Menschen durchführt, nicht fair mit der eines Kollegen verglichen werden, der primär elektive Eingriffe an jungen, gesunden Patienten vornimmt. Die Sterblichkeitsrate der ersten Klinik wird zwangsläufig höher sein, ohne dass dies auf eine schlechtere Qualität schließen lässt. Eine Risikoadjustierung korrigiert die Ergebnisse um solche Ausgangsfaktoren (Alter, Schwere der Erkrankung, Komorbiditäten etc.) und macht die reine Behandlungsqualität erst vergleichbar.

Das IQTIG-Verfahren ignoriert dieses fundamentale Prinzip. Es ist systemisch blind für den Schweregrad, die Chronizität, die Komplexität und die strukturellen Defizite der behandelten Patientinnen und Patienten. Die wenigen erhobenen soziodemografischen Daten sind bei weitem nicht ausreichend, um die immense Heterogenität psychischer Erkrankungen abzubilden und statistisch zu kontrollieren (Benecke et al., 2025). Das System vergleicht somit permanent Äpfel mit Birnen: Es wirft die Ergebnisse einer Kurzzeittherapie bei einer umschriebenen Phobie mit denen einer mehrjährigen Langzeittherapie bei einer Patientin mit komplexer Traumafolgestörung und Borderline-Persönlichkeitsstörung in denselben statistischen Topf.

Die Konsequenz ist ebenso einfach wie fatal: Praxen, die einen überdurchschnittlich hohen Anteil an schwerkranken, komplexen und prognostisch schwierigen Fällen behandeln, werden im Qualitätsvergleich systematisch und unausweichlich schlechter abschneiden. Ihre Ergebnisse werden, gemessen an den simplen Indikatoren des IQTIG, unter dem Durchschnitt liegen – nicht weil sie schlechter arbeiten, sondern weil ihre Arbeit ungleich schwerer ist.

5.2 Härtetest Persönlichkeitsstörung: Wenn klinische Realität als schlechte Qualität fehlinterpretiert wird

Nirgendwo wird die verzerrende und realitätsferne Logik des IQTIG-Systems deutlicher als bei der Anwendung auf die Behandlung von Patientinnen und Patienten mit schweren Persönlichkeitsstörungen (PS), insbesondere der Borderline-Persönlichkeitsstörung. Die Psychotherapieforschung hat umfassend belegt, dass die Behandlung dieser Patientengruppe spezifischen, klinisch erwartbaren Mustern folgt, die vom IQTIG-System systematisch als Indikatoren für schlechte Qualität fehlinterpretiert werden.

- Hohe Abbruchquoten als Teil der Pathologie: Therapien mit PS-Patienten sind durch eine signifikant höhere Abbruchrate gekennzeichnet. Diese Abbrüche sind oft nicht primär das Resultat eines Therapeutenfehlers, sondern ein direkter Ausdruck der zugrundeliegenden Störungspathologie: massive Schwierigkeiten mit Bindung, Vertrauen und Beziehungsregulation führen zu plötzlichen, oft im Streit vollzogenen Beziehungsabbrüchen (Schindler et al., 2013). Im IQTIG-System wird ein solcher Abbruch jedoch potenziell negativ gewertet und mündet mit hoher Wahrscheinlichkeit in einem nicht zurückgesandten oder einem aus Wut und Enttäuschung heraus extrem negativ ausgefüllten Patientenfragebogen.
- Negative Übertragung und Allianzbrüche als therapeutische Realität: Die Etablierung einer stabilen therapeutischen Allianz ist bei PS-Patienten eine der zentralen Herausforderungen. Intensive negative Gefühle gegenüber dem Therapeuten (negative Übertragung), Phasen des Misstrauens, der Entwertung und der Wut sind keine Störfaktoren, sondern der Kern des therapeutischen Prozesses, in dem sich die Beziehungsmuster des Patienten reinszenieren (Sallin et al., 2021). Die Fähigkeit des Therapeuten, diese "Allianzbrüche" (ruptures) auszuhalten und konstruktiv zu bearbeiten (repair), ist ein entscheidender Wirkfaktor. Das IQTIG-System hingegen erfasst mit seinem Fragebogen nur die Momentaufnahme der Patientenzufriedenheit. Eine Befragung nach einer Phase intensiver negativer Übertragung wird zwangsläufig zu katastrophalen Werten bei Items wie "fühlten Sie sich verstanden?" oder "hatten Sie Vertrauen?" führen. Das System ist unfähig, zwischen einem notwendigen, krisenhaften Durcharbeitungsprozess und einem tatsächlichen Qualitätsmangel zu unterscheiden.

Im Ergebnis pathologisiert das IQTIG-Verfahren normale, ja sogar notwendige therapeutische Prozesse in der Behandlung schwerer Störungen. Es bestraft Therapeuten dafür, dass sie sich auf einen klinisch anspruchsvollen Prozess einlassen, der nicht dem Idealbild einer harmonischen, linear verlaufenden Behandlung entspricht.

5.3 Das Kalkül des Ruins: Warum die Behandlung schwerer Fälle zum untragbaren wirtschaftlichen Risiko wird

Diese systematische Fehlinterpretation klinischer Realitäten wäre bereits problematisch genug. In Kombination mit dem Wegfall des Regress-Schutzes (Kapitel 2.3) und der geplanten öffentlichen Berichterstattung (Kapitel 3.3) wird sie jedoch zu einer existenziellen Bedrohung für niedergelassene Psychotherapeuten. Betrachten wir das konkrete Kalkül einer Praxis, das Ihr Szenario aufgreift:

Stellen wir uns eine engagierte, hochqualifizierte Therapeutin vor, die sich auf die Behandlung von Patientinnen mit schweren Persönlichkeitsstörungen spezialisiert hat und mehrere dieser anspruchsvollen, oft zweistündigen Langzeittherapien durchführt. Aufgrund der in 5.2 beschriebenen klinischen Realitäten muss sie damit rechnen, dass ein signifikanter Anteil dieser Therapien krisenhaft verläuft oder vorzeitig im Konflikt beendet wird. Nehmen wir an, dass – realistisch geschätzt – jeder zweite dieser Fälle in einer negativen Bewertung im IQTIG-Fragebogen oder einem Abbruch mündet.

Für diese Therapeutin entfaltet sich nun eine Kaskade des Ruins:

1. Der statistische Ruin: Ihre aggregierten Werte im zweijährlichen Rückmeldebericht werden dramatisch unter dem Durchschnitt der Referenzgruppe liegen, die auch viele Praxen mit "leichteren" Fällen umfasst. Sie wird als "rechnerisch auffällig" geflaggt.
2. Der administrative Ruin: Es folgt das Stellungnahmeverfahren. Sie muss nun in aufwendigen Schriftsätzen erklären, warum ihre Werte so schlecht sind – eine demütigende Rechtfertigung für fachlich gebotene, engagierte Arbeit.
3. Der reputative Ruin: Mit der geplanten Veröffentlichung der Qualitätsdaten wird ihre Praxis in öffentlichen Online-Portalen als "unterdurchschnittlicher" oder "schlechter" Anbieter erscheinen. Potenzielle neue Patienten, die den klinischen Kontext nicht verstehen können, werden ihre Praxis meiden. Ihr Ruf, den sie sich durch anspruchsvolle Arbeit aufgebaut hat, wird durch einen fehlerhaften Algorithmus zerstört.
4. Der wirtschaftliche Ruin: Die Kombination aus schlechter öffentlicher Bewertung und dem Status der "rechnerischen Auffälligkeit" macht sie zu

einem primären Ziel für Wirtschaftlichkeitsprüfungen. Da ihre Behandlungen per definitionem lang und teuer sind, läuft sie Gefahr, mit massiven Regressforderungen konfrontiert zu werden, gegen die sie sich ohne den Schutz des Gutachterverfahrens kaum noch wehren kann.

Die einzig rationale Konsequenz aus diesem Kalkül ist eine Form der unternehmerischen Notwehr: Um die eigene Praxis vor dem reputativen und wirtschaftlichen Ruin zu schützen, muss die Therapeutin die Behandlung dieser Risikogruppe einstellen oder drastisch reduzieren. Es handelt sich hierbei nicht um eine Frage mangelnder Ethik, sondern um eine ökonomische Zwangsläufigkeit, die direkt aus dem Design des IQTIG-Systems folgt. Das System erzwingt Risikoselektion.

5.4 Die Konsequenz für die Gesellschaft: Eine neue, hausgemachte Versorgungslücke

Skaliert man dieses individuelle Kalkül auf das gesamte Versorgungssystem hoch, werden die verheerenden gesamtgesellschaftlichen Folgen sichtbar. Wenn eine große Zahl von Therapeutinnen und Therapeuten diese rationale, ökonomisch getriebene Entscheidung trifft, entsteht eine massive Versorgungslücke für die vulnerabelsten Patientengruppen.

Genau jene Menschen, die aufgrund der Schwere ihrer Erkrankung den höchsten Leidensdruck haben und die psychotherapeutische Behandlung am dringendsten benötigen, werden es noch schwerer haben, einen Therapieplatz zu finden. Die ohnehin schon bestehenden langen Wartezeiten für komplexe Fälle (Grabe & Giertz, 2020) werden sich dramatisch verlängern. Diese neue Versorgungslücke ist keine Folge von Ressourcenmangel oder mangelndem Engagement der Therapeuten, sondern eine direkte, hausgemachte Konsequenz eines fehlgeleiteten Steuerungsinstruments.

Die Ironie ist bitter: Ein System, das angetreten ist, die Qualität der Versorgung für alle Patientinnen und Patienten zu verbessern, erreicht das exakte Gegenteil. Es schafft ein Zwei-Klassen-System, in dem "leichte", gut messbare und prognostisch günstige Fälle willkommen sind, während die "schwierigen", komplexen und prognostisch unsicheren Fälle systematisch aus der ambulanten Regelversorgung verdrängt werden. Die Risikoselektion ist damit keine theoretische Gefahr mehr,

sondern die logische und unausweichliche Folge eines fundamentalen Konstruktionsfehlers.

6 Der konstruktive Gegenentwurf: Professionseigene Wege zur echten Qualitätssicherung

Die vehemente und breit getragene Kritik der Psychotherapeutenschaft am IQTIG-Verfahren entspringt keiner fundamentalen Ablehnung von Qualitätssicherung, Rechenschaftspflicht oder Datenerhebung per se. Im Gegenteil: Supervision, Intervention, kontinuierliche Fortbildung und das Streben nach fachlicher Exzellenz sind tief in der DNA der Profession verankert. Die Kritik richtet sich vielmehr gegen die spezifische, als praxisfern, wissenschaftlich mangelhaft und potenziell schädlich empfundene Konzeption eines von außen verordneten Systems. Dass es auch anders geht, beweisen zahlreiche Initiativen, die in den letzten Jahren aus der Profession und der universitären Psychotherapieforschung heraus entstanden sind.

Diese "Bottom-up"-Ansätze verfolgen eine fundamental andere Philosophie als das "Top-down"-Modell des IQTIG. Sie zielen nicht primär auf externe Kontrolle, Vergleichbarkeit und Sanktionierung, sondern auf die interne, prozessbegleitende Verbesserung der konkreten therapeutischen Arbeit im Dialog mit dem Patienten. Sie verstehen Qualitätssicherung nicht als retrospektives Audit, sondern als ein prospektives Navigationsinstrument, das dem Therapeuten hilft, seinen Kurs zu halten und drohende Misserfolge frühzeitig zu erkennen. Dieses Kapitel stellt die Philosophie dieser alternativen Wege vor, beleuchtet funktionierende Modelle aus der Praxis und arbeitet den entscheidenden Unterschied zwischen unterstützender Verbesserung und kontrollierender Überwachung heraus.

6.1 Die Philosophie der Verbesserung: Prozessorientierung, Feedback und wissenschaftliche Fundierung

Professionseigene QS-Systeme basieren auf einer gemeinsamen Philosophie, die sich in drei Kernprinzipien zusammenfassen lässt:

1. **Prozessorientierung statt reiner Ergebnis-Auditierung:** Der zentrale Fokus dieser Ansätze liegt auf dem laufenden therapeutischen Prozess. Qualität wird nicht als statischer Endpunkt nach Abschluss der Behandlung verstanden, sondern als dynamischer Prozess, der während der Therapie aktiv gestaltet und verbessert werden kann. Das Ziel ist es, dem Therapeuten Werkzeuge an die Hand zu geben, die ihm während der Behandlung helfen, bessere Entscheidungen zu treffen, kritische Phasen zu

erkennen und Interventionen anzupassen. Qualitätssicherung wird so zu einem integralen Bestandteil der klinischen Arbeit, nicht zu einem nachgelagerten Verwaltungsakt.

2. **Zeitnahes Feedback statt verzögerter Berichterstattung:** Um den Prozess aktiv steuern zu können, ist zeitnahes, fallbezogenes Feedback unerlässlich. Anstatt Jahre später einen aggregierten statistischen Bericht zu erhalten, liefern diese Systeme dem Therapeuten idealerweise von Sitzung zu Sitzung oder in regelmäßigen kurzen Abständen konkrete Informationen über den Therapieverlauf des einzelnen Patienten. Sie machen sichtbar, ob ein Patient auf einem guten Weg ist ("on track") oder ob er Gefahr läuft, zu stagnieren oder die Therapie abubrechen ("not-on-track") (Lutz et al., 2019). Dieses Feedback dient als Frühwarnsystem und ermöglicht proaktives Handeln, anstatt lediglich retrospektiv Probleme zu konstatieren.
3. **Wissenschaftliche Fundierung statt Ad-hoc-Entwicklungen:** Diese Systeme basieren in der Regel auf international etablierten, wissenschaftlich validierten Messinstrumenten und Theorien der Psychotherapieforschung. Sie nutzen anerkannte Fragebögen zur Symptombelastung, zur Lebensqualität oder zur therapeutischen Allianz und werden selbst kontinuierlich wissenschaftlich begleitet und evaluiert. Im Gegensatz zum IQTIG-Ansatz, der ein neues, nicht extern validiertes Instrumentarium einführt, fußen diese Modelle auf einem soliden empirischen Fundament (Benecke et al., 2025).

Diese drei Prinzipien führen zu einem QS-Verständnis, das Therapeuten nicht als potenzielle Fehlerquellen betrachtet, die kontrolliert werden müssen, sondern als hochqualifizierte Experten, die durch intelligente, datengestützte Werkzeuge in ihrer komplexen Arbeit unterstützt werden können.

6.2 Funktionierende Modelle in der Praxis: Routine Outcome Monitoring (ROM) und das QSP/QVA-Projekt

Diese Philosophie hat zu einer Reihe von konkreten, in der Praxis bereits erprobten Modellen geführt, die zeigen, wie eine alternative datengestützte QS aussehen kann.

Routine Outcome Monitoring (ROM): Das Prinzip des kontinuierlichen Feedbacks

Das international am weitesten verbreitete Konzept ist das Routine Outcome Monitoring (ROM). Der Grundgedanke ist einfach und überzeugend: Patientinnen und Patienten füllen in regelmäßigen Abständen (oft vor jeder Sitzung) sehr kurze, standardisierte Fragebögen auf einem Tablet oder Smartphone aus. Diese erfassen zentrale Dimensionen wie das aktuelle Wohlbefinden, die Symptomschwere oder die wahrgenommene Qualität der therapeutischen Beziehung. Die Ergebnisse werden sofort algorithmisch ausgewertet und dem Therapeuten grafisch aufbereitet zur Verfügung gestellt, oft im Vergleich zu einer erwarteten "Verlaufskurve", die aus großen Datensätzen erfolgreicher Therapien abgeleitet wird.

Die wissenschaftliche Evidenz für die Wirksamkeit von ROM ist substanziell. Zahlreiche Meta-Analysen zeigen, dass der Einsatz solcher Feedback-Systeme die Therapieergebnisse signifikant verbessern kann. Der Effekt ist besonders ausgeprägt bei jenen "not-on-track"-Fällen, die sonst mit hoher Wahrscheinlichkeit in einer Stagnation oder einem Therapieabbruch enden würden. Die Rate von Behandlungsverschlechterungen und -abbrüchen kann durch den Einsatz von ROM teils erheblich reduziert werden (Barkham et al., 2023; de Jong et al., 2021). Konkrete Umsetzungen dieses Prinzips in Deutschland sind beispielsweise der Trierer Therapie-Navigator (TTN) (Lutz et al., 2019) oder das Greifswalder Psychotherapie-Navigationssystem (GPNS) (Demir et al., in press), die in Hochschulambulanzen und zunehmend auch in Praxen eingesetzt werden.

Das QSP/QVA-Projekt: Eine integrierte Plattform für Praxis und Forschung

Eine besonders relevante Initiative im Kontext der aktuellen Debatte ist das QSP/QVA-Projekt, das von Forschern um Benecke und Volz entwickelt wurde (Benecke & Volz, 2023). Dieses Projekt verfolgt einen doppelten Zweck: Es dient einerseits als unabhängige wissenschaftliche Plattform zur kritischen Validierung der IQTIG-Instrumente selbst und bietet andererseits Praxen und Ambulanzen ein eigenständiges, praxisorientiertes QS-Werkzeug an.

Im Gegensatz zum IQTIG-Verfahren setzt das QSP-Projekt auf eine umfassende Datenerhebung mit wissenschaftlich etablierten Instrumenten zu mehreren Messzeitpunkten: eine detaillierte Baseline-Erhebung zu Beginn, regelmäßige Verlaufsmessungen während der Therapie und eine Katamnese nach Therapieende. Entscheidend ist, dass das System eine faire Vergleichbarkeit durch eine integrierte Risikoadjustierung ermöglicht, indem es die Eingangsschwere der Patientenfälle

statistisch berücksichtigt (Volz & Benecke, 2024). Teilnehmende Praxen erhalten nicht nur zeitnahes Feedback zu Einzelfällen, sondern auch einmal jährlich eine risikoadjustierte, anonymisierte Auswertung ihrer Gesamtergebnisse im Vergleich zu einer relevanten Referenzgruppe. Diese Daten dienen primär der internen Reflexion und Qualitätsentwicklung; die Datenhoheit verbleibt vollständig bei der Praxis. Das Projekt steht bundesweit allen Psychotherapeuten aller Verfahren offen und demonstriert, dass die Entwicklung und Implementierung eines fairen, wissenschaftlich fundierten und klinisch nützlichen QS-Systems aus der Profession heraus möglich ist.

6.3 Der entscheidende Unterschied: "Bottom-up"-Unterstützung statt "Top-down"-Kontrolle

Der Vergleich zwischen den professionseigenen Ansätzen und dem IQTIG-Verfahren offenbart einen fundamentalen, philosophischen und praktischen Gegensatz. Es ist der Unterschied zwischen einem unterstützenden "Bottom-up"- und einem kontrollierenden "Top-down"-System.

- **Zweck des Systems:** Das IQTIG-System ist primär ein Instrument der externen Rechenschaftslegung und Kontrolle. Sein Hauptziel ist es, Transparenz für externe Akteure (Krankenkassen, Politik, Patientenöffentlichkeit) zu schaffen und potenzielle "Ausreißer" zu identifizieren. Der Nutzen für den internen Verbesserungsprozess ist, wie gezeigt, marginal. Professionseigene Systeme sind hingegen primär Werkzeuge der internen Qualitätsverbesserung. Ihr Hauptziel ist es, dem Therapeuten zu helfen, seine Arbeit mit dem aktuellen Patienten zu optimieren.
- **Rolle des Therapeuten:** Im IQTIG-System ist der Therapeut primär ein Datenlieferant für eine zentrale Auswertungsstelle. Er wird zum Objekt einer externen Bewertung. In ROM- und QSP-Systemen ist der Therapeut ein aktiver Nutzer von Daten. Er ist der Adressat des Feedbacks und nutzt die Informationen als Experte, um seine klinischen Entscheidungen zu fundieren.
- **Datenhoheit und Vertrauen:** Das IQTIG-Verfahren basiert auf einer verpflichtenden, zentralisierten Datensammlung, bei der die Praxis die Kontrolle über die Weiterverwendung und Interpretation ihrer Daten verliert.

Dies schafft eine Atmosphäre des Misstrauens und der Kontrolle.

Professionseigene Ansätze basieren auf Freiwilligkeit und Datenhoheit. Die Praxis entscheidet selbst über die Teilnahme und behält die Kontrolle über ihre sensiblen Daten. Dies fördert eine Kultur des Vertrauens und der intrinsischen Motivation zur Verbesserung.

- Wirkung auf die Praxis: Das IQTiG-System droht, durch seine Fehlanreize und das Sanktionspotenzial eine Kultur der Angst und der Risikovermeidung zu schaffen. Professionseigene Systeme zielen darauf ab, eine Kultur des Lernens und der professionellen Neugier zu fördern. Sie bieten Sicherheit durch Daten und ermutigen Therapeuten, sich auch an schwierige Fälle heranzuwagen, da sie den Verlauf engmaschig monitoren und ihre Erfolge objektiv dokumentieren können.

Zusammenfassend lässt sich festhalten: Die professionseigenen Gegenentwürfe sind keine bloßen Anti-Modelle. Sie sind der positive Beweis, dass eine datengestützte Qualitätssicherung möglich ist, die den Namen "Sicherung" verdient: eine Sicherung, die auf wissenschaftlicher Evidenz basiert, die klinische Praxis unterstützt, die Professionalität der Therapeuten respektiert und letztlich darauf abzielt, die Ergebnisse für jeden einzelnen Patienten zu verbessern, anstatt Praxen in ein starres, ungerechtes und potenziell schädliches Ranking zu zwingen.

7 Fazit und Forderungen: Plädoyer für eine sofortige Kurskorrektur

Die Analyse des neuen datengestützten Qualitätssicherungsverfahrens für die ambulante Psychotherapie zeichnet ein alarmierendes Bild. Der mit der Abschaffung des Gutachterverfahrens eingeleitete Paradigmenwechsel von einer qualitativen, auf professioneller Reflexion basierenden Prüfung hin zu einer quantitativen, auf standardisierter Messung beruhenden Kontrolle erweist sich bei genauerer Betrachtung als ein hochproblematischer, wenn nicht gar gefährlicher Weg. Während das politische Ziel – die Schaffung von Transparenz und die Verbesserung der Versorgungsqualität – unbestritten ist, ist das vom IQTIG entwickelte Instrument zur Erreichung dieses Ziels in seiner gegenwärtigen Form fundamental ungeeignet. Es ist ein System mit potenziell schweren Nebenwirkungen, das nicht nur seine eigenen Ziele zu verfehlen droht, sondern das Risiko birgt, aktiven Schaden in der Versorgungslandschaft anzurichten. Dieses Fazit fasst die zentralen Kritikpunkte zusammen, leitet daraus konkrete und unaufschiebbare Forderungen für die anstehende Erprobungsphase ab und skizziert abschließend die Vision eines alternativen, fairen und klinisch nützlichen QS-Systems.

7.1 Zusammenfassung: Ein System, das seine Ziele verfehlt und Schaden anrichtet

Die vorangegangene Analyse hat eine Kaskade von fundamentalen Konstruktionsfehlern im IQTIG-QS-Verfahren aufgedeckt, die in ihrer Gesamtheit ein vernichtendes Urteil bedingen:

- **Wissenschaftliche Untauglichkeit:** Das System basiert auf methodisch mangelhaften, nicht ausreichend validierten Instrumenten. Der Verzicht auf eine Baseline-Messung verunmöglicht jede objektive Erfolgsbewertung und der Mangel an statistischer Belastbarkeit degradiert die Ergebnisse für einzelne Praxen zu einem reinen Zufallsprodukt.
- **Fehlender klinischer Nutzen:** Durch seinen retrospektiven Charakter und die stark verzögerte, aggregierte Rückmeldung bietet das System keinen nützlichen Input für die Steuerung laufender Therapien oder für konkrete Lernprozesse und scheitert somit an seinem Ziel, das einrichtungsinterne Qualitätsmanagement zu unterstützen.

- Systematische Fehlanreize und die erzwungene Risikoselektion: Der gravierendste Konstruktionsfehler ist das Fehlen einer funktionierenden Risikoadjustierung. Dieser Mangel führt zwangsläufig dazu, dass Praxen, die schwere und komplexe Fälle behandeln, systematisch benachteiligt werden. Das System erzwingt eine Kultur der Risikovermeidung.

Im Ergebnis ist das IQTIG-Verfahren in seiner jetzigen Form keine Qualitätssicherung, sondern eine Gefahr für die Versorgungsgerechtigkeit. Es droht, eine neue, hausgemachte Versorgungslücke für die vulnerabelsten Patientengruppen zu schaffen, indem es deren Behandlung für freiberufliche Therapeuten zu einem unkalkulierbaren Risiko macht.

7.2 Konkrete Forderungen für die Evaluation in der Erprobungsphase

Die Entscheidung des G-BA für eine sechsjährige Erprobungsphase war eine notwendige Notbremse, die jedoch nur dann sinnvoll ist, wenn sie als kritisches Zeitfenster für eine grundlegende Revision genutzt wird. Angesichts der dargelegten systemischen Mängel schließt sich die vorliegende Analyse den fundierten Forderungen an, wie sie von führenden Fachleuten und Verbänden, unter anderem von Benecke et al. (2025) und Volz & Benecke (2024), formuliert wurden, und erweitert diese um einen entscheidenden Punkt. Diese Forderungen sind keine bloßen Vorschläge, sondern nicht verhandelbare Voraussetzungen für jedes zukünftige QS-System, das den Anspruch auf Fairness und Wissenschaftlichkeit erhebt:

1. Implementierung einer validen Risikoadjustierung als oberste Priorität: Dies ist die *conditio sine qua non*. Ohne ein robustes, validiertes Modell, das Schweregrad, Komplexität und Chronizität der Patientenfälle statistisch berücksichtigt, bleibt jeder Praxisvergleich wissenschaftlich unseriös und ethisch fragwürdig.
2. Grundlegende Revision der Messinstrumente: Die Ad-hoc-Instrumente des IQTIG müssen durch international etablierte, psychometrisch validierte Standardinstrumente ersetzt werden. Die Methodik muss zwingend um eine obligatorische Baseline-Erhebung zu Beginn der Therapie ergänzt werden, um eine echte Veränderungsmessung überhaupt erst zu ermöglichen.

3. Unabhängige und ergebnisoffene wissenschaftliche Begleitung: Die Evaluation muss durch eine genuinely unabhängige wissenschaftliche Institution erfolgen. Sie muss ergebnisoffen sein, was bedeutet, dass bei nachgewiesener Untauglichkeit oder Schädlichkeit des Verfahrens auch ein vollständiger Stopp und eine Neukonzeption eine reale Option sein müssen.
4. Transformation zum klinischen Unterstützungswerkzeug: Die Evaluation muss explizit prüfen, wie das System transformiert werden kann, um einen echten klinischen Nutzen zu stiften. Dies erfordert eine radikale Abkehr von verzögerten Aggregatberichten hin zu zeitnahe, fallbezogenem und klinisch relevantem Feedback.
5. Wiederherstellung eines ökonomischen Schutzmechanismus: Unabhängig vom QS-Verfahren muss der Gesetzgeber eine Lösung für den weggefallenen Regress-Schutz bei indizierten Langzeittherapien schaffen. Ohne diese Absicherung bleibt die Behandlung von Härtefällen ein unkalkulierbares Risiko.

Über diese zentralen Forderungen hinausgehend muss eine sechste, noch fundamentalere Forderung erhoben werden: 6. Ein Paradigmenwechsel im Entwicklungsprozess selbst: Das Scheitern des IQTIG-Entwurfs ist auch ein Symptom eines gescheiterten Top-Down-Prozesses. Zukünftige QS-Systeme dürfen nicht mehr von technischen Instituten im stillen Kämmerlein entworfen werden, um sie der Profession anschließend überzustülpen. Stattdessen muss ein partizipativer Co-Design-Prozess etabliert werden, in dem Psychotherapeutinnen, Psychotherapieforscher und Patientenvertreter von Anfang an als gleichberechtigte Partner in die Konzeption und Entwicklung eingebunden sind. Nur so kann sichergestellt werden, dass ein System entsteht, das die klinische Realität abbildet und von der Profession akzeptiert und getragen wird.

7.3 Ausblick: Skizze eines fairen und nützlichen QS-Systems der Zukunft

Die Kritik am IQTIG-Verfahren darf nicht im Destruktiven verharren. Die aus der Profession entwickelten Alternativen weisen den Weg in eine konstruktive Zukunft. Ein zukunftsfähiges QS-System, das aus einem partizipativen Prozess hervorgeht, könnte als hybrides Modell konzipiert sein:

- Eine Basisebene der Rechenschaftslegung: Für die externe Transparenz könnten wenige, robuste, risikoadjustierte und auf Basis von Prä-Post-Messungen erhobene Ergebnisindikatoren an eine zentrale Stelle gemeldet werden.
- Eine zweite Ebene der internen Qualitätsentwicklung: Der Schwerpunkt sollte auf der Förderung des internen Qualitätsmanagements liegen. Praxen könnten ermutigt und finanziell dabei unterstützt werden, an zertifizierten, professionseigenen QS-Plattformen (wie ROM- oder QSP-Systemen) teilzunehmen.
- Ein Schutzmechanismus für komplexe Fälle: Ein vereinfachtes, unbürokratisches Antragsverfahren für prognostisch schwierige Behandlungen könnte die notwendige ökonomische Sicherheit garantieren.

Ein solches System würde die berechtigten Ansprüche an Transparenz erfüllen, ohne die Versorgungsgerechtigkeit zu untergraben. Es würde Kontrolle auf das Notwendigste beschränken und den Fokus auf die Unterstützung und kontinuierliche Verbesserung der therapeutischen Arbeit legen. Die aktuelle Debatte bietet die historische Chance, aus den gemachten Fehlern zu lernen und gemeinsam ein QS-System zu schaffen, das seinen Namen verdient und der psychotherapeutischen Versorgung in Deutschland tatsächlich dient.

8 Literaturverzeichnis

Barkham, M., de Jong, K., Delgadillo, J., & Lutz, W. (2023). Routine outcome monitoring. In C. E. Hill & J. C. Norcross (Hrsg.), *Psychotherapy skills and methods that work* (S. 429–471). Oxford University Press.

Benecke, C., & Volz, M. (2023). Qualitätsmerkmale und Versorgungsrelevanz psychodynamischer Ausbildungsambulanzen: Das QVA-Projekt. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 69(4), 345–368.

<https://doi.org/10.13109/zptm.2023.69.aa4>

Benecke, C., Volz, M., Rulfs, F., Cwik, J. C., Piechotta, B., Rubel, J., Hunger-Schoppe, C., & Brakemeier, E.-L. (2025). Qualitätssicherung in der ambulanten Psychotherapie: Übersicht und aktuelle Entwicklungen. *Psychotherapeutenjournal*, 24(3), 222–233.

Bundespsychotherapeutenkammer. (2019, 27. September). Ab 2022: Abschaffung des Antrags- und Gutachterverfahrens. Pressemitteilung.

<https://www.bptk.de/pressemitteilungen/ab-2022-abschaffung-des-antrags-und-gutachterverfahrens/>

Bundespsychotherapeutenkammer. (2020, 14. November). Empfehlungen der BPTK für die Dokumentation psychotherapeutischer Behandlungen in der psychotherapeutischen Versorgung.

https://api.bptk.de/uploads/Empfehlungen_der_B_Pt_K_fuer_die_Dokumentation_psychotherapeutischer_Behandlungen_in_der_psychotherapeutischen_Versorgung_130bf23008.pdf

Bundespsychotherapeutenkammer. (2024, 19. Januar). Geplantes QS-Verfahren ambulante Psychotherapie nachteilig für Patientenversorgung. Pressemitteilung.

<https://www.bptk.de/pressemitteilungen/bptk-geplantes-qs-verfahren-ambulante-psychotherapie-nachteilig-fuer-patientenversorgung/>

de Jong, K., Conijn, J. M., Gallagher, R. A., Reshetnikova, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review*, 85, 102002. <https://doi.org/10.1016/j.cpr.2021.102002>

Demir, S., Bartholdy, S., Tewes, M., Brakemeier, E.-L. & Kaiser, T. (in press). Advances in Routine Outcome Monitoring and Feedback Systems: Insights from the Greifswald Psychotherapy Navigation System (GPNS). In J. A. Rubel & W. Lutz (Hrsg.), *Feedback Informed Therapy – Science and Practice*. Springer.

Deutsche Gesellschaft für Psychoanalyse, Psychotherapie, Psychosomatik und Tiefenpsychologie. (2019). Stellungnahme zum Entwurf eines Gesetzes für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz – DVG).

Deutsche Psychotherapeutenvereinigung, Deutsche Gesellschaft für Psychoanalyse, Psychotherapie, Psychosomatik und Tiefenpsychologie, Bundesverband der Vertragspsychotherapeuten & Vereinigung Analytischer Kinder- und Jugendlichen-Psychotherapeuten. (2022, 2. Dezember). Gemeinsame Stellungnahme zum IQTIG-Abschlussbericht zum QS-Verfahren ambulante Psychotherapie.

Gemeinsamer Bundesausschuss. (2024a, 18. Januar). Richtlinie zur datengestützten einrichtungsübergreifenden Qualitätssicherung: Themenspezifische Bestimmungen für das Verfahren 16: QS ambulante Psychotherapie.

Grabe, H. J., & Giertz, K. (2020). Die Borderline-Persönlichkeitsstörung in den psychosozialen, psychotherapeutischen und psychiatrischen Versorgungssystemen von Deutschland. *Psychotherapie Forum*, 28(4), 181–191.

Hauten, L., & Jungclaussen, I. (2022). Gutachterverfahren in der Psychotherapie: Ungeliebter Freund. *Deutsches Ärzteblatt PP*, (3), 110–112.

IQTIG. (2022). Überarbeitung des einrichtungsübergreifenden Qualitätssicherungsverfahrens zur ambulanten psychotherapeutischen Versorgung gesetzlich Krankenversicherter. Abschlussbericht.

Kassenärztliche Vereinigung Nordrhein. (2024, September). FAQ zum Verfahren ambulante Psychotherapie.

Lutz, W., Rubel, J. A., Schwartz, B., Schilling, V., & Deisenhofer, A.-K. (2019). Towards integrating personalized feedback research into clinical practice: Development of the Trier Treatment Navigator (TTN). *Behaviour Research and Therapy*, 120, 103438. <https://doi.org/10.1016/j.brat.2019.103438>

Sallin, L., Geissbühler, I., Grandjean, L., Beuchat, H., Martin-Soelch, C., Pascual-Leone, A., & Kramer, U. (2021). Self-contempt, the working alliance and outcome in treatments for borderline personality disorder: An exploratory study. *Psychotherapy Research*, 31(6), 765–777.

Schindler, A., Hiller, W., & Fydrich, T. (2013). Patient- and therapy-related predictors of dropout from outpatient cognitive-behavioral therapy. *Psychotherapy Research*, 23(3), 308–320.

Volz, M., & Benecke, C. (2024). Die neue Qualitätssicherung auf dem Prüfstand – empirische Studien zur Anwendung der IQTIG-Kriterien. *Psychotherapie*, 29(2), 73–84. <https://doi.org/10.30820/2364-1517-2024-2-73>

Wittmann, W. W., Lutz, W., Steffanowski, A., Kriz, D., Glahn, E. M., Völkle, M. C., et al. (2011). Qualitätsmonitoring in der ambulanten Psychotherapie: Modellprojekt der Techniker Krankenkasse. Abschlussbericht. Techniker Krankenkasse.

9 Fiktives Plenum

Titel : "Qualitätssicherung in der Psychotherapie – Paradigmenwechsel oder Systemkollaps?"

Veranstaltungsort: Fachtagung "Zukunft der Versorgung", Berlin. Moderation: Dr. Anya Sharma, Gesundheitsökonomin

(Dr. Sharma eröffnet die Diskussion)

Dr. Sharma (Moderatorin): Sehr geehrte Damen und Herren, ich begrüße Sie herzlich. Um unsere heutige, hochkomplexe Debatte so transparent wie möglich zu gestalten, werden wir ein etwas anderes Format wählen. Nach jedem Redebeitrag werden wir eine kurze Einordnung vornehmen, die den strategischen Punkt des Arguments beleuchtet und die zugrundeliegenden Fakten und Quellen offenlegt. So können wir alle nachvollziehen, auf welcher Basis hier argumentiert wird.

Wir stehen vor einer fundamentalen Weichenstellung. Das neue, datengestützte QS-System soll eine Ära der Transparenz einläuten. Kritiker warnen vor einem Systemkollaps für die Versorgung der Schwächsten. Staatssekretär Berger, als Vertreter der Politik, die diesen Wandel verantwortet: Was war der unausweichliche Impuls für diese Reform?

Staatssekretär Berger (Bundesgesundheitsministerium): Vielen Dank. Der Impuls war nicht radikal, sondern zwingend notwendig und entsprang einer einfachen Beobachtung: Das System der psychotherapeutischen Versorgung befand sich in einem Zustand der strukturellen Stagnation, der in einem modernen, dynamischen Gesundheitswesen nicht länger tragbar war. Wir sahen uns einem Paradoxon gegenüber: Auf der einen Seite explodierten die Ausgaben und die Zahl der Therapeuten stieg unaufhörlich, was rechnerisch zu einer Überversorgung in vielen Regionen führte. Auf der anderen Seite erreichten uns täglich die verzweifeltsten Berichte von Patientinnen und Patienten, insbesondere von jenen mit schweren, chronischen Erkrankungen, die monate- oder gar jahrelang keinen Therapieplatz fanden. Diese Diskrepanz zwischen massivem Ressourceneinsatz und mangelnder Zugänglichkeit für die bedürftigsten Gruppen war ein klares Indiz für eine ineffiziente Ressourcenallokation und fehlende Steuerung.

Das alte Gutachterverfahren, oft romantisch als Instrument der kollegialen Reflexion verklärt, war aus systemischer Sicht ein wesentlicher Teil dieses Problems. Es war

ein in sich geschlossenes, intransparentes Ritual mit einer Genehmigungsquote von nahezu 100 Prozent, das keinerlei Daten für eine übergeordnete Qualitätsbewertung oder Bedarfssteuerung lieferte. Es war eine Black Box. Unser politischer Auftrag ist es, ein patientenzentriertes, transparentes und lernendes Gesundheitssystem zu schaffen. Dieser Anspruch kann nicht vor den Türen der psychotherapeutischen Praxen Halt machen. Die Einführung eines datengestützten QS-Verfahrens ist daher kein Misstrauensvotum, sondern ein Akt der Normalisierung und Integration. Wir heben die Psychotherapie auf Augenhöhe mit anderen hochspezialisierten medizinischen Disziplinen, in denen datengestützte Qualitätssicherung seit Jahren selbstverständlich ist.

Der Kern der Reform ist die Stärkung der Patientensouveränität. Der mündige Versicherte hat ein verbrieftes Recht darauf, eine informierte Entscheidung zu treffen. Er muss wissen dürfen, welche Qualität er für die solidarisch finanzierten Beiträge erwarten kann. Mit dem neuen QS-Verfahren schaffen wir erstmals eine Grundlage für einen fairen Qualitätswettbewerb. Die Bedenken und Ängste aus der Profession nehmen wir ernst. Genau deshalb haben wir eine außergewöhnlich lange, sechsjährige Erprobungsphase etabliert. Dies ist ein geschützter Raum, um das System zu kalibrieren und Vertrauen aufzubauen. Aber der grundsätzliche Weg in eine datengestützte, transparente Zukunft ist unumkehrbar und richtig.

Analyse & Belege: Staatssekretär Bergers Argumentation etabliert die offizielle politische Rechtfertigungslinie für die Reform; strategisch legt dies im Artikel die Fallhöhe für die nachfolgende Kritik (siehe Kapitel 2.2 & 3.1). Seine Position stützt sich auf reale politische Dokumente und Narrative: Das Argument der "Modernisierung" und "Patientenstärkung" findet sich prominent in der Gesetzesbegründung zum Psychotherapeutenausbildungsreformgesetz von 2019; die Kritik am Gutachterverfahren als intransparentes Ritual mit geringer Ablehnungsquote (<4%) wird durch Daten wie die der Techniker Krankenkasse (Wittmann et al., 2011) untermauert; die Forderung nach Angleichung an andere medizinische Sektoren ist eine langjährige Position der Kostenträger, die hier von der Politik übernommen wird. Der Verweis auf die Erprobungsphase dient der Entkräftung von Kritik und ist im Beschluss des G-BA vom Januar 2024 (G-BA, 2024a) offiziell verankert.

Dr. Sharma (Moderatorin): Frau Dr. Meier, Sie hören die politische Vision von Transparenz und Patientensouveränität. Wie übersetzt sich diese Vision in die Realität Ihrer täglichen Arbeit?

Dr. Meier (Niedergelassene Psychotherapeutin): Herr Staatssekretär, Ihre Vision klingt in der Theorie bestechend, aber sie basiert auf einer fundamentalen Verkennung dessen, was Psychotherapie im Kern ausmacht und was Qualität in diesem Kontext bedeutet. Ihr System, so wie es vom IQTIG entworfen wurde, ist kein Fenster, das Licht in eine Black Box wirft. Es ist ein Zerrspiegel, der ein groteskes und irreführendes Bild meiner Arbeit zeichnet. Ich möchte das an einem ganz konkreten Punkt festmachen: Sie sprechen von "messbarer Qualität". Aber was messen Sie denn wirklich? Sie messen nicht die Komplexität meiner diagnostischen Überlegungen, nicht die Qualität meiner Beziehungsgestaltung in einer hochgradig ambivalenten therapeutischen Situation, nicht meine Fähigkeit, einen suizidalen Patienten durch eine schwere Krise zu navigieren. Nein, Sie messen mit einem starren, bürokratischen Raster, ob ich formal bestimmte Checklisten-Punkte erfüllt und dokumentiert habe.

Noch gravierender ist die sogenannte "Messung" der Ergebnisqualität durch den Patientenfragebogen. Ich habe mich in meiner Praxis bewusst auf die Behandlung von Menschen mit schweren Persönlichkeitsstörungen und komplexen Traumafolgestörungen spezialisiert. Das sind Menschen, für die das Gesundheitssystem oft keine adäquaten Angebote hat. Die Arbeit mit ihnen ist per definitionem langwierig, anstrengend und krisenhaft. Es ist ein klinisches Faktum und in der Forschung hundertfach belegt, dass diese Therapien durch intensive negative Übertragungsfänomene, durch Phasen des Hasses, der Entwertung und des Misstrauens gegenüber dem Therapeuten gekennzeichnet sind. Es ist ebenfalls ein Fakt, dass die Abbruchraten hier signifikant höher sind, oft weil der therapeutische Prozess unerträglich schmerzhaft wird. Das sind keine Zeichen von schlechter Therapie – im Gegenteil, die Fähigkeit, diese Krisen auszuhalten und zu bearbeiten, ist die Qualität.

Ihr IQTIG-Fragebogen aber kommt am Ende und fragt: "Hatten Sie Vertrauen zu Ihrer Therapeutin?". Was wird eine Patientin antworten, die die Therapie im Streit abgebrochen hat, weil sie sich von mir nicht verstanden fühlte, obwohl oder gerade weil ich ihre destruktiven Beziehungsmuster konfrontiert habe? Ihr System wird diese klinisch erwartbare und notwendige Dynamik als einen Datenpunkt für "mangelnde

Qualität" verbuchen. Es bestraft mich also systematisch dafür, dass ich mich der schwierigsten und undankbarsten Aufgabe in unserem Feld stelle. Sie nennen das "Transparenz". Ich nenne das eine systematische Bestrafung von Engagement. Die Konsequenz ist eine Zwangsläufigkeit, die ich das "Kalkül des Ruins" nenne. Konfrontiert mit der Gefahr, öffentlich als "schlechte" Therapeutin gebrandmarkt zu werden und in Regress zu geraten, werde ich eine rationale, unternehmerische Entscheidung treffen müssen: Ich kann es mir nicht mehr leisten, diese Patientengruppe in diesem Umfang zu behandeln. Das ist keine Hypothese, das ist die unausweichliche Folge Ihres Systemdesigns. Sie schaffen keine bessere Versorgung, Sie organisieren die Ausgrenzung der Schwächsten.

Analyse & Belege: Dr. Meiers Beitrag ist die personifizierte Kernthese des Artikels; sie übersetzt die abstrakte Systemkritik in die existenzielle Realität der Praxis und bildet damit das Herzstück von Kapitel 5 ("Die reale Gefahr..."). Ihre Argumentation ist empirisch fundiert: Die beschriebenen klinischen Realitäten bei der Behandlung von Persönlichkeitsstörungen sind wissenschaftlicher Konsens; die Schwierigkeiten in der therapeutischen Allianz werden beispielsweise bei Sallin et al. (2021) thematisiert, die hohen Abbruchraten bei Schindler et al. (2013). Ihre Kritik am Fragebogen als "Zerrspiegel" deckt sich mit der Analyse in Kapitel 4.2 und wird von Fachartikeln wie Benecke et al. (2025) geteilt. Das von ihr skizzierte "Kalkül des Ruins" ist die logische Konsequenz der in Kapitel 5.3 dargelegten Verknüpfung von fehlerhafter Messung, öffentlicher Berichterstattung und dem Wegfall des Regress-Schutzes (thematisiert in Kapitel 2.3); die daraus resultierende Gefahr der Risikoselektion wird von allen Berufsverbänden befürchtet (DPtV et al., 2022) und durch Studien zur bereits bestehenden Unterversorgung dieser Patientengruppe (Grabe & Giertz, 2020) als reale Gefahr untermauert.

Dr. Sharma (Moderatorin): Ein sehr eindringliches Plädoyer. Frau Dr. Weber, Sie sind die Architektin dieses Systems. Was entgegnen Sie auf den Vorwurf, Ihr System sei nicht nur ein Zerrspiegel, sondern erzwingt aktiv die Ausgrenzung der am schwersten erkrankten Patienten?

Dr. Weber (IQTIG): Ich nehme die von Frau Dr. Meier geschilderten klinischen Herausforderungen mit großem Respekt zur Kenntnis. Ihre detaillierte Beschreibung

unterstreicht die Komplexität des Feldes, für das wir ein Qualitätssicherungssystem entwickeln mussten. Allerdings möchte ich der Darstellung widersprechen, unser System sei "blind" für diese Realität. Es betrachtet sie lediglich durch eine andere, für den Zweck der systemweiten Qualitätssicherung notwendige Linse: die der Methodik, der Standardisierung und der Skalierbarkeit. Unser Auftrag war explizit nicht, ein Supervisions- oder Intervisionstool für den Einzelfall zu schaffen. Unser Auftrag war die Entwicklung eines fairen, verfahrensagnostischen und statistisch robusten Messsystems für eine Grundgesamtheit von rund 38.000 Leistungserbringern. Dieser beispiellose Maßstab erzwingt methodische Entscheidungen, die sich von der klinischen Einzelfalllogik unterscheiden müssen.

Lassen Sie mich auf die von Frau Dr. Meier genannten Punkte eingehen. Sie kritisiert, wir würden nur "Checklisten-Punkte" messen. Das ist eine verkürzte Darstellung. Wir operationalisieren zentrale, von der Fachwelt selbst definierte Aspekte guter Prozessqualität – wie Diagnostik, Zielvereinbarung und Aufklärung – in messbare Indikatoren. Dies ist der Kern jeder empirischen Wissenschaft: Komplexe Phänomene werden in beobachtbare, quantifizierbare Variablen übersetzt. Die Alternative wäre, im Bereich der reinen Interpretation und Subjektivität zu verharren, was aber per definitionem keine Vergleichbarkeit zulässt.

Der entscheidende Punkt ist jedoch die angebliche Bestrafung bei der Behandlung schwerer Fälle. Die geschilderte Sorge basiert auf der Annahme, dass das System in seiner finalen Form die Rohdaten der Patientenbewertungen eins zu eins als Qualitätsurteil verwenden wird. Das ist ausdrücklich nicht der Fall und wäre in der Tat methodisch unhaltbar. Die Problematik der sogenannten "case-mix"-Unterschiede ist uns nicht nur bewusst, sie ist ein zentrales Thema in der Entwicklung jedes medizinischen QS-Verfahrens. Genau aus diesem Grund ist der Prozess zur Entwicklung einer validen Risikoadjustierung das Herzstück der nun beginnenden Erprobungsphase. Es ist ein fundamentaler Irrtum zu glauben, man könne ein solches statistisches Korrekturmodell theoretisch am Reißbrett entwerfen. Um fair zu sein, muss es empirisch auf einer breiten Datenbasis entwickelt werden. Wir müssen zuerst systematisch Daten darüber sammeln, welche Patientenmerkmale – und dazu gehören Diagnoseschwere, Komorbiditäten, Chronizität und eben auch das Vorliegen einer Persönlichkeitsstörung – wie stark mit den Ergebnisindikatoren korrelieren. Erst auf Basis dieser Erkenntnisse können wir

ein statistisches Modell kalibrieren, das den Schwierigkeitsgrad der behandelten Patientenschaft fair herausrechnet.

Das Vorgehen ist also genau umgekehrt zu dem, was Frau Dr. Meier befürchtet: Wir sammeln die Daten nicht, um sie naiv zu interpretieren und engagierte Therapeuten zu bestrafen. Wir sammeln sie, um die systematischen Verzerrungen, die sie beschreibt, wissenschaftlich zu modellieren und sie in einem zweiten Schritt aus den Ergebnissen zu eliminieren. Die sechsjährige Erprobungsphase, in der Sanktionen explizit ausgesetzt sind, ist genau dieser geschützte wissenschaftliche Raum für diese anspruchsvolle Entwicklungsarbeit. Frau Dr. Meiers Praxis wird also nicht "bestraft", sondern ihre Daten werden zu einem wertvollen Beitrag, um das System fair zu machen. Am Ende dieses Prozesses wird ein risikoadjustiertes System stehen, das die Leistung einer Praxis, die sich auf schwere Fälle spezialisiert hat, fair mit der einer Praxis mit einem leichteren Klientel vergleichen kann. Die Alternative, auf jegliche Messung zu verzichten, nur weil sie im ersten Schritt noch unadjustiert ist, würde bedeuten, für immer in der Intransparenz zu verharren. Das kann nicht im Interesse der Patienten sein.

Analyse & Belege: Frau Dr. Webers Beitrag ist eine klassische Verteidigung aus der Perspektive des Systemarchitekten; strategisch dient er dazu, die Gegenposition fair und stark darzustellen, um die spätere wissenschaftliche Kritik daran umso schärfer zu machen (siehe Kapitel 3 & 4). Ihre Argumentation ist konsistent mit der offiziellen Linie des IQTIG: Der Verweis auf den gesetzlichen Auftrag und die Notwendigkeit der Standardisierung ist zentral für das IQTIG-Selbstverständnis, wie es in ihren Methodenpapieren und Abschlussberichten (IQTIG, 2022) dargelegt wird; die Beschreibung der Risikoadjustierung als geplanter, zweistufiger Prozess ("erst Daten sammeln, dann Modell entwickeln") ist exakt die in den Dokumenten und im G-BA-Beschluss (2024a) festgelegte Vorgehensweise. Diese Position wird im Artikel als methodisch unhaltbar kritisiert (Kapitel 5.1), ist aber aus Sicht des IQTIG die einzig pragmatische und datenwissenschaftlich saubere Langfriststrategie. Die Aussetzung von Sanktionen während der Erprobungsphase ist ebenfalls ein Fakt (G-BA, 2024a), den sie hier als Beleg für die Seriosität und Vorsicht ihres Vorgehens ins Feld führt.

Dr. Sharma (Moderatorin): Frau Dr. Weber, Sie beschreiben einen methodisch stringenten, datenwissenschaftlichen Prozess. Ihr Vorgehen sei keine Ignoranz gegenüber der Praxis, sondern eine Notwendigkeit, die sich aus dem Auftrag und dem schieren Ausmaß des Systems ergibt. Professor Brandt, Sie sind Psychotherapieforscher. Überzeugt Sie diese technische Rationalität? Ist das Vorgehen des IQTIG aus wissenschaftlicher Sicht haltbar oder hören wir hier eine nachträgliche Rechtfertigung für fundamentale Konstruktionsfehler?

Prof. Dr. Elias Brandt (Psychotherapieforscher): Frau Dr. Webers Ausführungen sind ein rhetorisch beeindruckendes Beispiel dafür, wie man methodische Grundsätze bemühen kann, um ein zutiefst unwissenschaftliches Vorgehen zu verschleiern. Sie spricht von Skalierbarkeit, Standardisierung und iterativen Prozessen. Das klingt alles sehr modern und datengetrieben. Tatsächlich aber wird hier ein System verteidigt, das gegen die fundamentalsten, nicht verhandelbaren Axiome der psychometrischen Testtheorie und der Evaluationsforschung verstößt. Es ist, als würde ein Ingenieur eloquent die fortschrittlichen Materialien seiner Brückenkonstruktion preisen, dabei aber verschweigen, dass er die Fundamente auf Sand gebaut hat.

Lassen Sie mich Frau Webers Argumente Punkt für Punkt dekonstruieren. Erstens, die Behauptung, ein "maßgeschneidertes" Instrument sei eine methodische Notwendigkeit gewesen. In der Wissenschaft nennen wir das, was hier geschehen ist, die Entwicklung eines unvalidierten Ad-hoc-Instruments. Es ist ein Kardinalfehler, für einen hoch-konsequenziellen Anwendungsbereich – wir sprechen hier von der Bewertung der Arbeit von Tausenden Therapeuten und potenziellen Sanktionen – ein Messinstrument einzusetzen, dessen zentrale Gütekriterien, insbesondere die Konstrukt- und Kriteriumsvalidität, nicht zuvor in unabhängigen, transparenten Studien rigoros nachgewiesen wurden. Die Behauptung, es sei "verfahrensagnostisch", ist zunächst nur das – eine Behauptung. Ob es nicht doch systematisch bestimmte therapeutische Stile oder Patientengruppen bevorteilt oder benachteiligt, wissen wir nicht, weil die notwendige Validierungsforschung fehlt. Jahrzehnte der Forschung zur Entwicklung von Instrumenten wie dem Working Alliance Inventory oder dem Outcome Questionnaire wurden ignoriert. Das ist keine Stärke, das ist ein Bruch mit wissenschaftlicher Redlichkeit.

Zweitens, der Verzicht auf eine Baseline. Die Analogie zur Hotelbewertung, die in den IQTIG-Papieren mitschwingt, ist fundamental falsch und eine gefährliche

Trivialisierung. Psychotherapie ist keine Konsumdienstleistung wie ein Hotelaufenthalt, sondern eine medizinische Heilbehandlung, deren explizites und vertraglich vereinbartes Ziel die Veränderung eines pathologischen Ausgangszustandes ist. Wenn man Veränderung bewerten will, ist die Messung des Ausgangszustandes nicht optional, sie ist die logische Voraussetzung. Ohne den Ausgangszustand zu kennen, ist jede Aussage über Veränderung reine Spekulation. Die Behauptung, dies sei ein alternatives "Konzept von Erfolg", ist der Versuch, einen gravierenden methodischen Mangel in eine programmatische Stärke umzudeuten. Es ist wissenschaftlich unhaltbar.

Drittens, und das ist der intellektuell unredlichste Punkt in Frau Webers Verteidigung: die Strategie der "zweistufigen" Risikoadjustierung. Die Behauptung, man müsse erst unfaire Rohdaten sammeln, um daraus ein faires Korrekturmodell zu bauen, ist ein logischer und statistischer Zirkelschluss. Man kann aus systematisch verzerrten Daten kein unverzerrtes Modell ableiten. Das ist das "Garbage in, garbage out"-Prinzip der Datenwissenschaft. Das System wird, wie Frau Dr. Meier korrekt darlegt, in den ersten Jahren systematisch Praxen mit schweren Fällen als "schlecht" und Praxen mit leichten Fällen als "gut" klassifizieren. Wenn Sie nun versuchen, auf Basis dieser bereits verzerrten Daten ein Korrekturmodell zu "trainieren", wird dieses Modell die anfängliche Verzerrung bestenfalls fortschreiben oder sogar verstärken, aber niemals vollständig korrigieren können. Die wissenschaftlich saubere Vorgehensweise wäre gewesen, das gesamte System – inklusive eines vorab entwickelten Risikoadjustierungsmodells – in einer überschaubaren, kontrollierten Pilotstudie zu testen, zu validieren und erst nach dem Nachweis seiner Fairness und Validität über eine breitere Implementierung nachzudenken. Stattdessen startet man einen unkontrollierten Feldversuch mit einem unvalidierten Prototypen und hofft, das Flugzeug im Flug reparieren zu können. Das ist kein "iterativer Prozess", das ist ein grob fahrlässiges Experiment.

Analyse & Belege: Professor Brandts Beitrag ist die wissenschaftliche Dekonstruktion der IQTIG-Position; strategisch liefert er die zentralen Argumente aus Kapitel 4 ("Fundamentale Konstruktionsfehler") und untermauert die Thesen zur wissenschaftlichen Untauglichkeit des Verfahrens. Seine Argumentation ist eine direkte Replik auf die von Dr. Weber vorgebrachten Punkte: Die Kritik an den unvalidierten Ad-hoc-Instrumenten (Kapitel 4.1) wird von Volz & Benecke (2024)

geteilt und durch das Urteil der BPTK (2024) ("schlicht ungeeignet") gestützt; die Kritik am fehlenden Baseline als Bruch mit dem Goldstandard der Evaluationsforschung (Kapitel 4.1) ist wissenschaftlicher Konsens; die Dekonstruktion der "zweistufigen" Risikoadjustierung als "Garbage in, garbage out"-Problem greift die Analyse aus Kapitel 5.1 auf und führt sie zu ihrem logischen Schluss. Die Forderung nach einer vorherigen Pilotstudie ist eine konstruktive wissenschaftliche Alternative, die das Vorgehen des IQTIG als methodisch fragwürdig entlarvt und die Notwendigkeit professionseigener, forschungsbasierter Ansätze (wie in Kapitel 6 beschrieben) impliziert.

Dr. Sharma (Moderatorin): Professor Brandt, Ihre Kritik ist fundamental. Herr Richter, Sie vertreten die Krankenkassen, die diesen Prozess maßgeblich vorantreiben und finanzieren. Sie hören nun von zwei Seiten – der Praxis und der Wissenschaft – dass das System nicht nur ungerecht, sondern methodisch unhaltbar sei. Halten Sie als Treuhänder der Versichertengelder an einem System fest, das nach dieser Analyse potenziell irreführende "Qualitätsdaten" produziert?

Herr Richter (Gesetzliche Krankenkasse): Frau Dr. Sharma, ich habe die Beiträge von Frau Dr. Meier und Professor Brandt mit großer Aufmerksamkeit gehört. Ich bestreite weder die klinischen Herausforderungen, die Frau Dr. Meier beschreibt, noch die methodische Komplexität, die Professor Brandt aufwirft. Aber ich muss uns alle auf den Boden der Tatsachen und unseres gemeinsamen Versorgungsauftrags zurückholen. Wir führen hier keine akademische Debatte im Elfenbeinturm über die perfekte wissenschaftliche Methodik. Wir stehen vor der realpolitischen Notwendigkeit, ein System mit einem jährlichen Volumen von über 4,6 Milliarden Euro, finanziert aus den Beiträgen von Arbeitnehmern und Arbeitgebern, endlich einer transparenten Rechenschaftspflicht zu unterwerfen.

Als Treuhänder dieser Gelder kann ich es vor unseren Versicherten nicht länger verantworten, dass wir Milliarden in eine "Black Box" investieren, ohne auch nur den geringsten Anhaltspunkt über die Ergebnisqualität zu haben. Das alte Gutachterverfahren war, bei allem Respekt, ein Feigenblatt. Es lieferte keine steuerungsrelevanten Daten, sondern bestätigte lediglich den Status quo. Die Alternative zur Version 1.0 des IQTIG-Verfahrens, die Professor Brandt kritisiert, ist

nicht eine hypothetische, perfekte Version 2.0, die irgendwann in der Zukunft vom Himmel fällt. Die reale Alternative ist die Version 0.0 – also der Zustand der völligen Intransparenz, den wir seit Jahrzehnten haben. Und dieser Zustand ist für uns als Kostenträger und für die Patienten inakzeptabel.

Der entscheidende Punkt ist die Perspektive. Professor Brandt argumentiert aus der reinen Lehre der Wissenschaft. Wir müssen aus der Perspektive der Versichertengemeinschaft argumentieren. Und für unsere Versicherten ist es relevanter zu wissen, wie andere Patienten die Dienstleistung einer Praxis am Ende bewertet haben, als einen statistisch perfekten, aber für Laien unverständlichen Prä-Post-Effektstärken-Wert zu kennen. Die Patientenzufriedenheit, die erlebte Beziehungsqualität – das sind harte, valide Währungen in der Dienstleistungsqualität. Die Behauptung, diese Daten seien "irreführend", ist ein Affront gegen die Urteilsfähigkeit der Patienten selbst. Wir trauen unseren Versicherten zu, ihre Erfahrungen zu bewerten.

Die Gefahr der Risikoselektion ist real, und wir nehmen sie sehr ernst. Aber die Schlussfolgerung kann doch nicht sein, auf jede Form von Messung und Vergleich zu verzichten. Das wäre eine Kapitulation vor der Komplexität. Unsere Antwort muss sein, das System intelligent zu machen. Und genau das passiert in der Erprobungsphase. Die Entwicklung der Risikoadjustierung ist keine Nebensache, sie ist der zentrale Auftrag. Wir als Krankenkassen haben das größte Interesse daran, dass auch schwerkranke Patienten gut versorgt werden, denn ihre unbehandelte Chronifizierung verursacht die höchsten Folgekosten. Ein System, das diese Patienten aus der Versorgung drängt, wäre auch für uns ökonomisch unsinnig. Wir werden daher sehr genau darauf achten, dass am Ende ein faires, risikoadjustiertes Modell steht. Aber um dieses Modell zu bauen, brauchen wir die Daten. Es gibt keinen anderen Weg.

Lassen Sie es mich ganz klar sagen: Der Weg zurück in die Intransparenz ist keine Option. Der mündige Versicherte hat ein Recht darauf zu wissen, welche Qualität er für seine Beiträge erwarten kann. Wir schaffen hier die Grundlage für einen Markt, auf dem sich Qualität durchsetzen kann. Engagierte Therapeuten wie Frau Dr. Meier, die exzellente Arbeit mit schweren Fällen leisten, müssen keine Angst vor diesem System haben. Im Gegenteil: Ein faires, risikoadjustiertes System wird am Ende genau ihre Exzellenz sichtbar machen und sie von weniger erfolgreichen Anbietern abheben. Darauf vertrauen wir und dafür werden wir uns in den Gremien einsetzen.

Analyse & Belege: Herr Richters Beitrag verkörpert die pragmatische und am Ende machtpolitische Position der Kostenträger, wie sie im Artikel unter Kapitel 3 (insbesondere die Perspektive von GKV und Politik) und in der Auseinandersetzung mit der Vision des "souveränen Patienten" dargestellt wird. Seine Argumentation ist strategisch: Er stellt die Kritik als akademische "Elfenbeindebatte" dar, um die eigene Position als pragmatische und politisch notwendige "Realpolitik" zu rahmen; das Narrativ der "Black Box" und des "mündigen Versicherten" ist die zentrale Legitimationsgrundlage, die der GKV-Spitzenverband in zahlreichen Publikationen und Pressemitteilungen verwendet; die Betonung der Wirtschaftlichkeit und die Sorge vor Folgekosten bei Nichtbehandlung schwerer Fälle ist ein authentisches Argument der Kassen, das aber hier strategisch genutzt wird, um die Sorge vor Risikoselektion zu entkräften. Sein Vertrauen in die Lösbarkeit der methodischen Probleme innerhalb der Erprobungsphase spiegelt die offizielle Linie des G-BA wider, entpolitisiert aber die fundamentale Kritik von Prof. Brandt, indem er sie zu einem lösbaren "Detailproblem" herabstuft.

Dr. Sharma (Moderatorin): Wir halten fest: Die Krankenkassen sehen keine Alternative zur Datenerhebung, selbst wenn die Daten anfangs unvollkommen sind. Die Wissenschaft hält genau das für gefährlich. Frau Dr. Meier, Sie haben vorhin das "Kalkül des Ruins" beschrieben. Vertrauen Sie darauf, dass in sechs Jahren ein faires System steht, das Ihre Arbeit korrekt abbildet, oder ist der Schaden bis dahin längst angerichtet?

Dr. Meier (Niedergelassene Psychotherapeutin): Herr Richter, ich höre Ihre Worte vom Vertrauen in den Prozess und in ein zukünftiges, faires System. Aber Vertrauen ist in meinem Berufsfeld keine theoretische Größe, sondern etwas, das sich in der Realität bewähren muss. Und die Realität, die Sie und Frau Dr. Weber mit dem IQTIG schaffen, zerstört Vertrauen, anstatt es aufzubauen. Sie bitten mich, sechs Jahre lang an einem Experiment teilzunehmen, von dem ein renommierter Wissenschaftler hier am Tisch sagt, es sei "grob fahrlässig", und das meine Existenzgrundlage bedroht. Das ist ein Angebot, das ich nicht annehmen kann.

Sie sprechen von der "Version 1.0" eines Systems, als wäre es eine neue App auf meinem Handy. Aber hier geht es nicht um Software-Updates. Hier geht es um das Leben und die psychische Gesundheit meiner Patientinnen und Patienten und um meine berufliche und wirtschaftliche Existenz. Das "Kalkül des Ruins", das ich beschrieben habe, ist keine hypothetische Zukunftsangst. Es ist eine unmittelbare unternehmerische Realität, die mit dem Start der Datenerhebung am 1. Januar 2025 beginnt. Von diesem Tag an wird jeder Behandlungsfall, den ich abschließe, zu einem Datenpunkt, der gegen mich verwendet werden kann. Sie sagen, Sanktionen seien ausgesetzt. Aber der reputative Schaden durch eine zukünftige Veröffentlichung ist real. Die Angst vor der statistischen Auffälligkeit und der daraus folgenden Wirtschaftlichkeitsprüfung ist real. Sie verlangen von mir, sechs Jahre lang ein unkalkulierbares Risiko zu tragen, in der vagen Hoffnung, dass am Ende vielleicht ein faires System herauskommt. Welcher vernünftige Freiberufler würde das tun?

Ihre Analogie zum Qualitätswettbewerb bei Krankenhäusern ist zutiefst irreführend. Ein Krankenhaus ist eine große Organisation mit Rechtsabteilungen, Controlling und der Möglichkeit, Risiken zu streuen. Ich bin eine Einzelkämpferin in meiner Praxis. Ich habe keine Abteilung, die statistische Ausreißer wegdiskutiert. Ein einziger, hochkomplexer Fall, der schlecht läuft – und das passiert in meinem Arbeitsfeld, das ist Teil der Realität –, kann meine gesamte Bilanz für zwei Jahre verzerren. Sie vergleichen einen Ozeandampfer mit einem Ruderboot und wundern sich, dass das Ruderboot Angst vor dem Sturm hat, den Sie entfesseln.

Und an Frau Dr. Weber gerichtet: Ihr Glaube an die Allmacht der Daten in allen Ehren, aber Sie ignorieren die menschliche Dimension. Die therapeutische Beziehung ist unser wichtigstes Werkzeug. Sie basiert auf absolutem Vertrauen und Vertraulichkeit. Ihr System zwingt mich, meine Patienten darüber aufzuklären, dass nach der Therapie intime Details über ihren Prozess und ihre Bewertung meiner Arbeit an eine zentrale Stelle in Berlin gemeldet werden. Ein Widerspruchsrecht ist nicht vorgesehen. Das ist ein massiver Eingriff in den geschützten therapeutischen Raum. Es verändert die Dynamik. Der Patient ist nicht mehr nur mein Gegenüber, sondern auch ein potenzieller Bewerter für ein externes System. Das vergiftet die Beziehung. Sie machen die Therapie zu einem auditierbaren Prozess und wundern sich, dass die Seele dabei auf der Strecke bleibt. Ich habe keine Angst vor Feedback. Ich erhalte wöchentlich Feedback in meiner Supervision und Intervision von Kollegen, die meine Arbeit verstehen. Ich habe Angst vor einem blinden,

algorithmischen Urteil, das die Komplexität meiner Arbeit auf eine Zahl reduziert und mich zwingt, meine Patientenauswahl an den Bedürfnissen des Algorithmus auszurichten und nicht am Leid des Menschen, der vor mir sitzt.

Analyse & Belege: Dr. Meiers emotionaler und zugleich analytisch scharfer Beitrag bringt die Perspektive der Praxis auf den Punkt; im Artikel entspricht dies der Zuspitzung in Kapitel 5.3 ("Das Kalkül des Ruins") und der praktischen Kritik in Kapitel 4.3 ("Ein bürokratisches Monster ohne klinischen Nutzen"). Ihr Argument, dass die Erprobungsphase kein "sicherer Raum" ist, weil die Datensammlung sofort beginnt und reputative sowie wirtschaftliche Ängste erzeugt, ist ein entscheidender Punkt, der die beruhigende Rhetorik der Politik und der Kassen konterkariert; die Angst vor der retrospektiven Wirtschaftlichkeitsprüfung ist nach dem Wegfall des Gutachterverfahrens real und wird in Kapitel 2.3 als zentrale Folge der Reform analysiert. Ihre Kritik am Eingriff in den therapeutischen Raum durch die verpflichtende Datenerhebung ist ein wichtiger ethischer Aspekt, der in den offiziellen Stellungnahmen der Berufsverbände (z.B. DPtV et al., 2022) ebenfalls prominent thematisiert wird. Sie stellt dem "Top-down"-Kontrollsystem des IQTIG die "Bottom-up"-Qualitätssicherung durch Supervision und Intervision gegenüber, was auf die Alternativen in Kapitel 6 verweist

Dr. Sharma (Moderatorin): Frau Dr. Meier, Sie beschreiben einen unmittelbaren Vertrauensverlust und eine reale, gegenwärtige Bedrohung für Ihre Arbeit, die nicht erst in sechs Jahren beginnt. Staatssekretär Berger, die Praxis signalisiert Ihnen, dass die von Ihnen geschaffene Erprobungsphase kein "geschützter Raum" ist, sondern ein existenzbedrohendes Experiment. Die Therapeutinnen und Therapeuten, die das System umsetzen sollen, scheinen Ihnen das Vertrauen zu entziehen. Zwingt Sie das nicht zum Umdenken? Muss die Politik hier nicht eingreifen und die "Notbremse", die die Erprobung war, noch einmal fester anziehen?

Staatssekretär Berger (Bundesgesundheitsministerium): Ich verstehe die emotionale Tiefe der vorgetragenen Sorgen, und ich möchte an dieser Stelle sehr deutlich sein: Niemand in der Politik hat die Absicht, Praxen zu ruinieren oder die Versorgung für schwerkranke Menschen zu verschlechtern. Das Gegenteil ist unser Ziel. Aber wir müssen auch die Verhältnismäßigkeit wahren. Wir sprechen hier über

die Einführung eines standardisierten Evaluationsverfahrens, eine Praxis, die in fast allen anderen Bereichen des öffentlichen und privaten Lebens längst etabliert ist – von der Hochschullehre bis zur Automobilproduktion. Die psychotherapeutische Versorgung kann und darf keine Ausnahme bleiben, die sich jeder Form von externer, ergebnisorientierter Überprüfung entzieht.

Die von Frau Dr. Meier geschilderte Angst vor dem "reputativen Ruin" basiert auf der Annahme des schlimmstmöglichen Szenarios: dass nämlich ein unausgereiftes System mit unadjustierten Rohdaten ungefiltert an die Öffentlichkeit gegeben wird. Aber genau das ist doch durch die Struktur der Erprobungsphase ausgeschlossen! Der Gesetzgeber und der G-BA haben hier mehrstufige Sicherheitsmechanismen eingezogen. Die Aussetzung von Sanktionen ist der erste. Die lange Dauer von sechs Jahren, um die Methodik zu verfeinern, ist der zweite. Und es sollte selbstverständlich sein, dass eine öffentliche Berichterstattung erst dann erfolgen kann und wird, wenn ein wissenschaftlich validiertes, faires und von allen Seiten im G-BA akzeptiertes Risikoadjustierungsmodell vorliegt. Wir starten hier keinen unkontrollierten Feldversuch. Wir initiieren einen strukturierten, wissenschaftlich begleiteten und über die Gremien der Selbstverwaltung politisch kontrollierten Prozess.

Die Rolle der Politik ist es, den Rahmen zu setzen und die Ziele zu definieren. Und das Ziel ist und bleibt mehr Transparenz und eine bessere, datengestützte Versorgung. Die Aufgabe, die Details dieses Prozesses auszugestalten, die Instrumente zu kalibrieren und die Risikoadjustierung zu entwickeln, liegt per Gesetz bei der gemeinsamen Selbstverwaltung, also bei den Krankenkassen und den Leistungserbringern selbst, wissenschaftlich beraten durch das IQTiG. Die Psychotherapeutenchaft ist in diesen Gremien, etwa über die Bundespsychotherapeutenkammer und die Kassenärztlichen Bundesvereinigungen, prominent vertreten. Anstatt also von außen ein "Kalkül des Ruins" zu proklamieren, lautet mein Appell an die Profession: Nutzen Sie Ihre starken Mitgestaltungsmöglichkeiten in diesen Gremien! Bringen Sie Ihre klinische Expertise konstruktiv in den Prozess der Verfeinerung ein. Arbeiten Sie mit daran, dass das Risikoadjustierungsmodell genau die Aspekte abbildet, die Frau Dr. Meier zurecht anmahnt. Der politische Wille ist es nicht, ein perfektes System von oben zu diktieren, sondern einen Rahmen zu schaffen, in dem die Expertise aller Beteiligten zu einem besseren Ergebnis führen kann. Die Tür für einen konstruktiven Dialog in

den dafür vorgesehenen Gremien steht weit offen. Jetzt pauschal das gesamte Vorhaben in Frage zu stellen, wäre ein Rückschritt ins vor-transparente Zeitalter und würde am Ende vor allem den Patientinnen und Patienten schaden.

Analyse & Belege: Staatssekretär Bergers Beitrag ist ein klassisches politisches Manöver: Er versucht, die Verantwortung zu deligitimieren und von der exekutiven Ebene (dem Ministerium) zurück an die Ebene der Selbstverwaltung (den G-BA und die beteiligten Akteure) zu delegieren. Strategisch positioniert er die Kritik als destruktiv ("pauschal in Frage stellen") und appelliert an die konstruktive Mitarbeit der Profession in den Gremien, deren Entscheidungen aber den von der Politik gesetzten Rahmen nicht sprengen können. Dies spiegelt die reale politische Dynamik wider. Sein Verweis auf die Aussetzung von Sanktionen und die lange Erprobungsphase sind Fakten (G-BA, 2024a), die er nutzt, um die vorgetragenen Ängste als übertrieben darzustellen. Das Argument, die Mitgestaltungsmöglichkeiten in den Gremien der Selbstverwaltung seien der richtige Ort für die Auseinandersetzung, ist formal korrekt, ignoriert aber die realen Machtverhältnisse in diesen Gremien und die Tatsache, dass die grundlegenden Designfehler des Systems (wie von Prof. Brandt dargelegt) möglicherweise gar nicht innerhalb des gesetzten Rahmens reparierbar sind. Dieser Beitrag verdeutlicht die Kluft zwischen politischer Rhetorik und den Sorgen der Praxis, ein zentrales Thema des gesamten Artikels.

Prof. Dr. Elias Brandt (Psychotherapieforscher): Herr Staatssekretär, Ihr Appell an die konstruktive Mitarbeit der Profession in den Gremien der Selbstverwaltung klingt nobel, aber er ist in der Sache leider irreführend und verschleiert den Kern des Problems. Es ist, als würde man Architekten und Ingenieure dazu einladen, konstruktiv an der Farbgestaltung und der Inneneinrichtung eines Hauses mitzuwirken, dessen Fundament und Statik bereits fehlerhaft und nicht mehr verhandelbar sind. Das Problem dieses QS-Verfahrens liegt nicht in den Details, die man in Gremiensitzungen noch "nachjustieren" könnte. Es liegt in seinen fundamentalen, wissenschaftlich unhaltbaren Designprinzipien, die von Anfang an falsch aufgesetzt wurden und die durch den gesetzlichen Rahmen zementiert sind.

Sie bitten die Profession, an der Entwicklung eines Risikoadjustierungsmodells mitzuwirken. Aber wie ich bereits dargelegt habe, ist der Versuch, ein faires

Korrekturmodell aus systematisch verzerrten und unfairen Rohdaten abzuleiten, ein wissenschaftlicher Zirkelschluss. Die Datenbasis, auf der dieses Modell gebaut werden soll, wird von Tag eins an kontaminiert sein, weil das System, wie Frau Dr. Meier es eindrücklich schildert, klinische Realität als Qualitätsmangel fehlinterpretiert. Sie bitten die Experten also, aus "Datenmüll" ein valides statistisches Modell zu zaubern. Das ist keine wissenschaftliche Aufgabe, das ist Alchemie. Das grundlegende Problem – der Verzicht auf eine Baseline-Messung und die Verwendung unvalidierter Instrumente – ist durch die Vorgaben des IQTIG und des G-BA bereits gesetzt. Daran kann auch die konstruktivste Mitarbeit in einem Unterausschuss nichts mehr ändern.

Ihr Verweis auf die Mitgestaltungsmöglichkeiten in der Selbstverwaltung ist zudem ein politisches Ablenkungsmanöver. Die Profession hat in unzähligen, detaillierten und wissenschaftlich fundierten Stellungnahmen von Anfang an auf genau diese Konstruktionsfehler hingewiesen. Diese Warnungen wurden von der Politik und den entscheidenden Gremien im G-BA, in denen die Kostenträger die Mehrheit haben, weitgehend ignoriert. Stattdessen wurde ein System durchgedrückt, das nun in einer "Erprobungsphase" auf seine Praxistauglichkeit getestet werden soll. Das ist keine "wissenschaftliche Begleitung", das ist die nachträgliche Suche nach einer Rechtfertigung für eine bereits getroffene Fehlentscheidung.

Die politische Verantwortung lässt sich hier nicht einfach an die Selbstverwaltung delegieren. Die Politik, Herr Staatssekretär, hat den gesetzlichen Rahmen geschaffen, der dieses fehlerhafte System überhaupt erst ermöglicht hat. Und die Politik hat die Macht und die Verantwortung, diesen Rahmen zu korrigieren. Anstatt die Profession auf endlose und letztlich fruchtlose Detaildebatten in Gremien zu verweisen, wäre es die Aufgabe der Politik, die fundamentalen Webfehler des Gesetzes und des G-BA-Auftrags anzuerkennen und eine echte, ergebnisoffene Neukonzeption zu ermöglichen. Das würde bedeuten, die wissenschaftliche Notwendigkeit einer Baseline-Messung anzuerkennen, die Verwendung validierter Instrumente vorzuschreiben und die Entwicklung einer Risikoadjustierung vor dem Start der Datenerhebung zur Bedingung zu machen. Ihr Appell zur Mitarbeit ist unter den gegebenen Umständen daher leider nur der Ruf, an der Verwaltung des eigenen Untergangs mitzuwirken.

Analyse & Belege: Professor Brandts Replik ist die direkte Zurückweisung des politischen Delegitimierungsversuchs; er holt die Verantwortung von der Ebene der Selbstverwaltung zurück zur Politik und zur wissenschaftlichen Fundierung. Strategisch entspricht dies der Forderung im Artikel nach einer grundlegenden Revision des Systems (Kapitel 7.2) und der Kritik am "gescheiterten Top-Down-Prozess". Seine Metapher vom "Haus mit fehlerhaftem Fundament" bringt die Kritik aus Kapitel 4 ("Fundamentale Konstruktionsfehler") auf den Punkt. Das Argument des wissenschaftlichen Zirkelschlusses bei der Risikoadjustierung wiederholt und schärft die Kritik aus seinem vorherigen Beitrag. Sein Hinweis auf die ignorierten Stellungnahmen der Profession ist historisch korrekt; zahlreiche Verbände (DPtV et al., 2022) und die BPtK (2024) haben die nun diskutierten Mängel bereits in der Entwicklungsphase des IQTIG-Verfahrens detailliert kritisiert. Seine Forderungen nach einer Baseline-Messung und validierten Instrumenten sind die Kernpunkte der wissenschaftlichen Kritik, wie sie in Kapitel 4.1 dargelegt sind.

Dr. Sharma (Moderatorin): Professor Brandt, Sie werfen der Politik vor, die Profession zur "Verwaltung des eigenen Untergangs" einzuladen und die wissenschaftlichen Grundlagen zu ignorieren. Herr Richter, Sie als Vertreter der Kostenträger sind der mächtigste Akteur im Gemeinsamen Bundesausschuss. Sie hätten die Macht, auf die wissenschaftliche Kritik zu hören und eine Neuausrichtung zu erzwingen. Warum tun Sie es nicht? Warum halten Sie an einem System fest, dessen wissenschaftliches Fundament hier so vehement in Frage gestellt wird?

Herr Richter (Gesetzliche Krankenkasse): Frau Dr. Sharma, ich habe großen Respekt vor der wissenschaftlichen Expertise von Professor Brandt. Aber seine Argumentation, so stringent sie in der akademischen Theorie sein mag, operiert in einem Vakuum. Sie ignoriert die ökonomischen und steuerungspolitischen Realitäten, mit denen wir als Treuhänder der Beitragsgelder tagtäglich konfrontiert sind. Es geht hier nicht um die Erschaffung eines wissenschaftlich perfekten, reinen Messinstruments für die Forschung. Es geht um die Implementierung eines pragmatischen, "gut genug" funktionierenden Steuerungsinstruments für das größte und teuerste ambulante Versorgungssystem, das wir haben.

Die Behauptung, das Fundament sei fehlerhaft, ist eine Metapher. Die Realität ist: Wir haben aktuell gar kein Fundament für eine datengestützte Steuerung. Wir

operieren im Blindflug. Professor Brandt schlägt vor, wir sollten jahrelang in einer "kontrollierten Pilotstudie" ein perfektes System entwickeln. Das ist ein Luxus, den wir uns angesichts der explodierenden Kosten und der offensichtlichen Allokationsprobleme nicht leisten können. Wir müssen jetzt anfangen, Daten zu sammeln und zu lernen. Das IQTIG-Verfahren ist nicht das Ende der Entwicklung, es ist der Anfang. Es ist die Schaffung der allerersten, rudimentären Dateninfrastruktur, auf der wir überhaupt erst aufbauen können. Ja, die ersten Daten werden unadjustiert und "verrauscht" sein. Aber sie sind unendlich viel mehr wert als die null Daten, die wir jetzt haben.

Lassen Sie mich auf das Kernproblem zurückkommen, das Herr Staatssekretär Berger anfangs skizziert hat: das Allokationsproblem. Wir haben den begründeten Verdacht, dass im System eine erhebliche Risikoselektion bereits stattfindet – dass also Praxen unbewusst oder bewusst dazu neigen, "leichtere", prognostisch günstigere Fälle zu behandeln, weil sie einfacher und wirtschaftlicher sind. Das alte Gutachterverfahren hat dies nicht verhindert, es hat es nicht einmal sichtbar gemacht. Das neue System, selbst in seiner unadjustierten Rohfassung, wird uns erstmals einen groben Überblick darüber geben, wie sich die Patientenschweregrade – der "case-mix" – über die Praxen verteilen. Es wird Muster aufzeigen. Und ja, es wird auch die Praxen von Frau Dr. Meier als "Ausreißer" identifizieren. Aber das ist kein Fehler, das ist die Funktion des Systems! Es macht endlich sichtbar, dass es Praxen gibt, die eine andere, anspruchsvollere Arbeit leisten. Das ist der erste, unverzichtbare Schritt, um überhaupt über eine gerechtere Vergütung oder gezielte Förderprogramme für die Behandlung von Härtefällen nachdenken zu können. Solange wir diese Unterschiede nicht in Daten sehen, existieren sie für uns als steuernde Instanz nicht.

Die wissenschaftliche Kritik an der Methodik ist in ihrer Absolutheit realitätsfern. Kein QS-System in der Medizin ist perfekt. Die Qualitätsberichte der Krankenhäuser sind ebenfalls voller methodischer Kompromisse, und trotzdem hat ihre Einführung nachweislich zu Verbesserungen geführt, weil sie eine Kultur der Messung und des Vergleichs etabliert hat. Wir als Krankenkassen sind bereit, das Risiko eines unperfekten Starts einzugehen, weil das Risiko, in der Intransparenz zu verharren, ungleich größer ist. Wir vertrauen auf den Lernprozess innerhalb der Selbstverwaltung. Die Profession hat nun sechs Jahre Zeit, ihre Expertise einzubringen und zu beweisen, dass sie nicht nur kritisieren, sondern auch ein

besseres Risikoadjustierungsmodell mitgestalten kann. Wenn sie diese Chance nicht nutzt, sondern sich in einer Fundamentalopposition verliert, wird sie am Ende mit dem System leben müssen, das andere für sie gestaltet haben.

Analyse & Belege: Herr Richters Beitrag ist die pragmatische und machtbewusste Antwort des Finanzierers; er verkörpert die Position, die im Artikel als treibende Kraft hinter der Reform identifiziert wird. Seine Strategie ist es, die wissenschaftliche Kritik als "akademische Elfenbeindebatte" abzutun und sie der "realpolitischen Notwendigkeit" unterzuordnen. Das Argument, man müsse mit einem "gut genug" System starten, weil die Alternative ("gar keine Daten") schlimmer sei, ist eine typische Rechtfertigungsfigur für pragmatische Politik. Sein entscheidender Punkt, dass das QS-System die bereits stattfindende Risikoselektion erst sichtbar macht und damit die Grundlage für eine gerechtere Steuerung schafft, ist ein starkes Gegenargument, das die GKV in ihren Positionspapieren zur Versorgungssteuerung anführt (GKV-Spitzenverband). Er kehrt die Kritik von Frau Dr. Meier um: Nicht das System erzeugt die Selektion, es deckt sie auf. Der Verweis auf die methodischen Kompromisse in anderen QS-Systemen (z.B. im Klinikbereich) ist ebenfalls ein valider, relativierender Punkt. Sein abschließender Appell, der zugleich eine Drohung ist ("...wird sie mit dem System leben müssen, das andere für sie gestaltet haben"), spiegelt die realen Machtverhältnisse im G-BA wider.

Dr. Sharma (Moderatorin): Herr Richter, Sie argumentieren, dass selbst unperfekte Daten besser sind als gar keine und dass das System erst die Grundlage für ein späteres, faires System schafft. Frau Dr. Meier, Sie sollen also sechs Jahre lang "verrauschte", potenziell rufschädigende Daten über Ihre Praxis produzieren, im Vertrauen darauf, dass die Kassen und das IQTIG am Ende daraus etwas Gutes bauen. Ist das ein Deal, den die Praxis annehmen kann?

Dr. Meier (Niedergelassene Psychotherapeutin): Herr Richter, was Sie als pragmatischen Lernprozess beschreiben, ist aus meiner Sicht ein zynisches Experiment auf dem Rücken der Therapeuten und ihrer Patienten. Sie bitten mich, mein Haus anzuzünden, in der vagen Hoffnung, dass die Feuerwehr aus der Beobachtung des Brandes lernt, wie man zukünftig bessere Feuerlöscher baut.

Diesen Deal kann und werde ich nicht annehmen, denn er ignoriert die unmittelbaren, irreversiblen Konsequenzen.

Ihre Behauptung, "unperfekte Daten" seien besser als keine, ist fundamental falsch. Schlechte Daten sind unendlich viel schlimmer als keine Daten. Keine Daten führen zu einem Zustand der Ungewissheit, den wir durch professionelle Standards wie Supervision und Ausbildung zu managen versuchen. Schlechte, systematisch verzerrte Daten führen zu falschen Schlussfolgerungen und katastrophalen Fehlentscheidungen. Sie werden auf Basis dieser "verrauschten" Daten Praxen wie meine als "schlecht" identifizieren, obwohl wir die schwierigste Arbeit leisten. Sie schaffen eine falsche Realität, ein Artefakt, und wollen auf dieser Lüge dann eine "gerechtere Steuerung" aufbauen. Das ist absurd.

Sie sagen, das System mache Risikoselektion erst sichtbar. Das ist eine unglaubliche Verdrehung der Tatsachen! Das System erzwingt Risikoselektion als rationale Überlebensstrategie. Sie schaffen erst die Krankheit und bieten dann an, sie irgendwann in der Zukunft vielleicht zu diagnostizieren. Bis Ihr "Lernprozess" in sechs Jahren vielleicht zu einem halbwegs fairen Adjustierungsmodell geführt hat, ist der Schaden längst angerichtet. Bis dahin haben unzählige Kolleginnen und Kollegen längst ihr "Kalkül des Ruins" gezogen. Sie werden ihre Praxis umgestellt haben. Sie werden gelernt haben, Patientinnen mit Persönlichkeitsstörungen in der Sprechstunde zu erkennen und sie mit einer höflichen Begründung wegzuschicken, weil sie ihre Praxis nicht gefährden können. Sie werden gelernt haben, Therapien so zu führen, dass sie am Ende eine gute Bewertung im IQTIG-Fragebogen ergeben, auch wenn das bedeutet, schmerzhaft Konfrontationen zu vermeiden und den Prozess oberflächlicher zu halten.

Dieser "Lernprozess" der Therapeuten wird viel schneller sein als der Ihres Systems. Und er ist irreversibel. Eine einmal etablierte Kultur der Risikovermeidung lässt sich nicht einfach per Knopfdruck wieder abschalten, wenn Sie in sechs oder acht Jahren vielleicht ein faires System haben. Bis dahin haben wir eine ganze Generation von Therapeuten darauf trainiert, sich an den Algorithmus anzupassen. Und die Patientinnen und Patienten, die in dieser Zeit durchs Raster fallen, die keinen Platz finden, deren Störungen sich chronifizieren, die in Kliniken oder der Frühverrentung landen – deren Schicksal ist dann der Kollateralschaden Ihres "pragmatischen Lernprozesses".

Ich weigere mich, Teil dieses Experiments zu sein. Die Alternative ist nicht die "völlige Intransparenz", wie Sie es nennen. Die Alternative, wie sie Professor Brandt und viele andere vorschlagen, wäre gewesen, ein wissenschaftlich fundiertes System vorher zu entwickeln. Die Alternative sind die professionseigenen QS-Systeme, die auf Freiwilligkeit, Vertrauen und echtem klinischen Nutzen basieren. Aber diese Alternativen wurden ignoriert, weil sie nicht Ihrem Wunsch nach zentraler Kontrolle und universeller Vergleichbarkeit entsprechen. Sie opfern die Versorgungsrealität auf dem Altar einer technokratischen Fiktion.

Analyse & Belege: Dr. Meiers emotional aufgeladener, aber argumentativ scharfer Beitrag ist die direkte Zurückweisung des Pragmatismus der Kassen. Strategisch verlagert sie die Debatte von der langfristigen Systementwicklung zurück auf die unmittelbaren, irreversiblen Konsequenzen für Praxis und Patienten (Kapitel 5.3 & 5.4). Ihr Argument, "schlechte Daten sind schlimmer als keine Daten", ist eine direkte rhetorische Entgegnung auf Herrn Richters Position. Die Beschreibung des perversen Lernprozesses der Therapeuten (Anpassung an den Algorithmus, Vermeidung schwieriger Fälle) ist eine Konkretisierung der im Artikel beschriebenen Gefahr der Risikoselektion und des "Chilling Effects". Ihre Zurückweisung des "Experiments" und der Verweis auf die ignorierten professionseigenen Alternativen (wie in Kapitel 6 beschrieben) positioniert die Praxis nicht als reformfeindlich, sondern als Befürworterin eines anderen, besseren Weges. Ihr Beitrag macht die zeitliche Asymmetrie deutlich: Der Schaden in der Praxis entsteht sofort, während die angebliche Lösung des Systems in ferner Zukunft liegt.

Dr. Sharma (Moderatorin): Ein Appell, der an die Grundfesten dieser Reform rüttelt. Herr Staatssekretär, Frau Dr. Meier beschreibt einen irreversiblen Schaden, der bereits jetzt, zu Beginn der Erprobungsphase, entsteht – eine Kultur der Angst und Risikovermeidung. Sie wirft Ihnen vor, die Schicksale der am schwersten erkrankten Patienten als Kollateralschaden eines technokratischen Experiments in Kauf zu nehmen. Das ist ein moralischer Vorwurf, dem sich die Politik stellen muss. Bleiben Sie dabei, dass dieser Weg alternativlos ist?

Staatssekretär Berger (Bundesgesundheitsministerium): Die Schilderungen von Frau Dr. Meier gehen mir, und das meine ich sehr ernst, nahe. Und sie zeigen, dass

wir in der Kommunikation dieser Reform möglicherweise Fehler gemacht haben. Wenn der Eindruck entsteht, wir würden ein unkalkulierbares Experiment auf dem Rücken der Versorgung durchführen, dann ist das ein Alarmsignal, das wir hören und auf das wir reagieren müssen. Ich möchte daher unmissverständlich klarstellen: Die von Frau Dr. Meier beschriebene Konsequenz – eine systematische Ausgrenzung schwerkranker Patientinnen und Patienten – wäre ein inakzeptables Ergebnis dieser Reform und würde ihren Zielen diametral widersprechen. Sollten sich im Laufe der Erprobungsphase Anzeichen dafür verdichten, dass ein solcher Effekt in relevantem Maße eintritt, dann hat die Politik nicht nur die Möglichkeit, sondern die Pflicht, korrigierend einzugreifen.

Allerdings möchte ich auch vor voreiligen Schlüssen und einer Dramatisierung warnen, die sich auf Annahmen über die Zukunft stützt. Wir stehen am Anfang eines Prozesses, nicht an seinem Ende. Die von Professor Brandt und Frau Dr. Meier skizzierten Risiken sind genau die Prüfsteine, an denen sich die Erprobungsphase messen lassen muss. Die wissenschaftliche Begleitung, die gesetzlich vorgesehen ist, hat exakt den Auftrag, genau diese potenziellen negativen Lenkungseffekte zu monitoren. Wir operieren hier nicht im luftleeren Raum. Es wird eine kontinuierliche Evaluation geben, und die Ergebnisse werden den Gremien und auch der Politik regelmäßig vorgelegt.

Ich sehe hier auch einen Auftrag an uns als Ministerium, den Prozess enger zu begleiten und als Moderator zwischen den teils unversöhnlichen Positionen der Selbstverwaltungspartner zu agieren. Vielleicht müssen wir darüber nachdenken, verbindlichere Meilensteine für die Entwicklung der Risikoadjustierung festzulegen. Vielleicht müssen wir sicherstellen, dass die "Lernkurve", von der Herr Richter sprach, schneller und transparenter verläuft. Ich bin offen dafür, im Dialog mit der Profession zusätzliche "Leitplanken" und Sicherheitsmechanismen für die Erprobungsphase zu diskutieren, um die Sorgen von Frau Dr. Meier aufzunehmen.

Was ich jedoch zurückweisen muss, ist die Forderung, das gesamte Projekt zu stoppen und auf den Status quo ante zurückzufallen. Das wäre eine Kapitulation vor der Komplexität und würde die unbestreitbaren Defizite des alten Systems zementieren. Der Weg zu mehr Transparenz und datengestützter Qualität ist richtig. Aber wir müssen sicherstellen, dass wir auf diesem Weg niemanden zurücklassen. Mein Angebot an die Profession lautet daher: Lassen Sie uns die Erprobungsphase nicht als Bedrohung, sondern als Chance begreifen. Eine Chance, die

wissenschaftliche Kritik von Professor Brandt in konkrete Evaluationskriterien zu übersetzen und die praktischen Sorgen von Frau Dr. Meier in messbare "rote Linien", deren Überschreitung Konsequenzen haben muss. Lassen Sie uns die Erprobungsphase zu einem echten, lernenden Prozess machen, an dessen Ende ein System steht, das tatsächlich funktioniert. Die Alternative – ein Festhalten an der Intransparenz – wäre der eigentliche Schaden für die Patienten.

Analyse & Belege: Staatssekretär Bergers zweiter Beitrag ist ein klassisches Beispiel für politische Deeskalation und "Damage Control". Er verlässt die rein technokratische Argumentation und wechselt auf eine empathischere, dialogorientierte Ebene. Strategisch versucht er, die Fundamentalkritik zu kanalisieren, indem er sie in den Rahmen der bestehenden Erprobungsphase integriert ("Prüfsteine", "messbare rote Linien"). Er signalisiert Kompromissbereitschaft ("Leitplanken diskutieren"), ohne jedoch von der Grundsatzentscheidung für das System abzurücken. Dies spiegelt die reale politische Reaktion auf anhaltenden Druck wider. Sein Vorschlag, als "Moderator" aufzutreten, ist eine rhetorische Figur, die die Verantwortung des BMG anerkennt, ohne direkte legislative Konsequenzen zu versprechen. Die Betonung der wissenschaftlichen Begleitung ist ein Verweis auf die Vorgaben im G-BA-Beschluss (2024a), die er hier als Sicherheitsnetz darstellt. Im Kern ist seine Botschaft: "Wir haben eure Sorgen gehört, wir nehmen sie ernst, aber wir stoppen die Reform nicht – wir managen sie nur besser." Dies bereitet den Boden für die abschließende Kritik, dass auch ein "besser gemanagtes" fehlerhaftes System fehlerhaft bleibt.

Dr. Sharma (Moderatorin): Herr Staatssekretär, Sie bieten einen moderierten Dialog und "Leitplanken" an, halten aber am Grundprinzip der Reform fest. Professor Brandt, ist das aus wissenschaftlicher Sicht ein gangbarer Weg? Kann man ein System, das Sie als "grob fahrlässig" bezeichnet haben, durch "besseres Management" und "rote Linien" heilen, oder bleibt das Fundament brüchig?

Prof. Dr. Elias Brandt (Psychotherapieforscher): Herr Staatssekretär, Ihr Angebot zum Dialog und Ihre anerkennenden Worte sind ein Fortschritt in der Debattenkultur, für den ich Ihnen danke. In der Sache jedoch ist Ihr Vorschlag leider nur eine kosmetische Behandlung für ein Problem, das in der fundamentalen Statik des

Systems liegt. Sie bieten an, "Leitplanken" an einer Brücke anzubringen, deren Pfeiler bereits auf Sand gebaut sind. Das mag das Gefühl von Sicherheit kurzfristig erhöhen, aber es verhindert nicht den strukturellen Kollaps, der im Design angelegt ist.

Lassen Sie mich das ganz konkret machen. Sie schlagen vor, die "negativen Lenkungseffekte" während der Erprobungsphase zu "monitoren". Aber wie wollen Sie das tun? Nehmen wir an, in zwei Jahren zeigen die Daten, dass die Wartezeiten für Patienten mit Persönlichkeitsstörungen in Nordrhein-Westfalen signifikant gestiegen sind. Das wäre eine Ihrer "roten Linien". Was passiert dann? Die Befürworter des Systems werden argumentieren, dies sei auf gestiegene Inanspruchnahme oder andere externe Faktoren zurückzuführen. Die Kritiker werden sagen, es sei der Beweis für die Risikoselektion. Sie werden am Ende korrelative Daten haben, aber keine kausalen Beweise, weil das Studiendesign – ein unkontrollierter Feldversuch – keine kausalen Schlüsse zulässt. Sie werden jahrelang über die Interpretation der Daten streiten, während der Schaden in der Versorgung weiterläuft. Ein "Monitoring" ohne ein sauberes, experimentelles oder quasi-experimentelles Forschungsdesign ist wissenschaftlich wertlos und politisch beliebig interpretierbar.

Ihr Vorschlag, "verbindlichere Meilensteine" für die Entwicklung der Risikoadjustierung festzulegen, ist gut gemeint, aber er geht am Kernproblem vorbei. Das Problem ist nicht der Zeitplan, das Problem ist die Datenbasis. Wie ich bereits sagte: Sie können nicht aus systematisch verzerrten Daten ein unverzerrtes Modell ableiten. Selbst das perfektste statistische Modell kann die grundlegenden Konstruktionsfehler – das fehlende Baseline und die unvalidierten Instrumente – nicht heilen. Sie versuchen, ein Software-Problem durch Hardware-Upgrades zu lösen. Es wird nicht funktionieren.

Die einzige wissenschaftlich redliche Konsequenz aus der bisherigen Debatte wäre, die "Notbremse" tatsächlich voll durchzuziehen. Das bedeutet: Ein sofortiges Moratorium für die flächendeckende Datenerhebung in NRW. Stattdessen sollte eine echte, wissenschaftliche Pilotstudie in einem kleinen, repräsentativen Sample von Praxen aufgesetzt werden. Diese Studie müsste genau die Kriterien erfüllen, die wir heute diskutiert haben: Einsatz validierter Instrumente, eine obligatorische Baseline-Messung und die Entwicklung und Validierung eines Risikoadjustierungsmodells innerhalb dieses kontrollierten Settings. Erst wenn ein solches pilotiertes System

nachweislich fair, valide und nützlich ist, kann man über einen breiteren Roll-out nachdenken.

Alles andere, Herr Staatssekretär, ist keine "wissenschaftliche Begleitung", sondern das, was Frau Dr. Meier treffend als zynisches Experiment bezeichnet hat. Ihr Angebot, den Prozess zu "managen", ist der Versuch, ein fundamental falsches Paradigma am Leben zu erhalten. Der konstruktive Weg wäre, den Mut zu haben, den Fehler einzugestehen und den Prozess auf ein wissenschaftlich sauberes Fundament zu stellen.

Analyse & Belege: Professor Brandts abschließendes Statement ist die Zuspitzung der wissenschaftlichen Kritik und die direkte Zurückweisung des politischen Kompromissangebots. Strategisch dient es im Artikel dazu, die Forderungen in Kapitel 7.2 zu untermauern. Sein Argument, dass ein reines Monitoring ohne sauberes Forschungsdesign keine validen Schlüsse zulässt, ist ein zentraler wissenschaftstheoretischer Punkt. Seine Wiederholung der Kritik am "Garbage in, garbage out"-Prinzip der Risikoadjustierung bekräftigt die Analyse aus Kapitel 5.1. Die finale, konkrete Forderung nach einem Moratorium und einer echten Pilotstudie ist die logische und konstruktive Konsequenz seiner gesamten Kritik. Sie zeigt auf, wie ein wissenschaftlich verantwortungsvoller Weg aussehen würde und kontrastiert diesen scharf mit dem bestehenden politischen Plan. Dies dient als argumentativer Übergang zu Kapitel 6 ("Der konstruktive Gegenentwurf") und Kapitel 7.3 ("Ausblick"), wo solche alternativen, forschungsbasierten Modelle skizziert werden.

Dr. Sharma (Moderatorin): Herr Professor Brandt, Sie fordern nichts Geringeres als ein Moratorium und einen kompletten Neustart des Prozesses auf wissenschaftlich sauberem Fundament. Frau Dr. Weber, das IQTIG steht hier im Kreuzfeuer. Professor Brandt wirft Ihnen im Grunde vor, Alchemie statt Wissenschaft zu betreiben. Was entgegnen Sie auf den Vorwurf, dass Ihr gesamter Ansatz, aus fehlerhaften Rohdaten ein valides Korrekturmodell ableiten zu wollen, ein logischer Zirkelschluss ist?

Dr. Weber (IQTIG): Ich schätze die Klarheit, mit der Professor Brandt seine wissenschaftliche Position vertritt. Allerdings muss ich dem Vorwurf der "Alchemie" aufs Schärfste widersprechen. Seine Darstellung vereinfacht die Realität komplexer,

epidemiologischer und gesundheitsökonomischer Modellierungsverfahren in unzulässiger Weise. Es ist schlichtweg nicht korrekt, dass man aus "verrauschten" Daten keine validen Modelle ableiten kann. Die gesamte moderne Statistik, insbesondere im Bereich von Big Data und Machine Learning, beschäftigt sich mit nichts anderem, als aus unvollständigen, komplexen und fehlerbehafteten Datensätzen signifikante Muster und Prädiktoren zu extrahieren.

Der Vorwurf des Zirkelschlusses greift zu kurz. Wir werden die anfänglichen Rohdaten nicht naiv verwenden. Der Prozess ist weitaus differenzierter: Wir werden Subgruppenanalysen durchführen. Wir werden die Daten von Praxen wie der von Frau Dr. Meier, die einen hohen Anteil an Patienten mit Persönlichkeitsstörungen behandeln, gesondert analysieren und als eigene Kohorte betrachten. Wir werden die Korrelationen zwischen Diagnosen, soziodemografischen Faktoren und den Outcome-Indikatoren untersuchen. Aus diesen Analysen werden wir Hypothesen über die notwendigen Adjustierungsfaktoren ableiten. Diese Hypothesen werden dann an unabhängigen Teildatensätzen getestet und validiert. Das ist ein Standardverfahren in der empirischen Forschung und hat nichts mit einem "blinden Stochern im Nebel" zu tun.

Professor Brandts Forderung nach einer perfekten, kontrollierten Pilotstudie ist ein Idealbild aus dem Labor, das mit der Komplexität der Versorgungsrealität nicht vereinbar ist. Eine kleine Pilotstudie mit vielleicht 50 Praxen hätte niemals die statistische Power und die Heterogenität, um ein Modell zu entwickeln, das für 38.000 Praxen im ganzen Bundesgebiet fair und valide ist. Um die Realität abzubilden, müssen wir die Realität messen – in ihrer ganzen Breite und mit all ihren "Verschmutzungen". Genau das tun wir in der Erprobungsphase. Es ist kein "Experiment", es ist eine prospektive Kohortenstudie im Real-Setting – die größte, die es in der deutschen Psychotherapie je gegeben hat.

Die Alternative, die Professor Brandt vorschlägt – ein Moratorium –, würde bedeuten, weitere Jahre des Stillstands und der Intransparenz in Kauf zu nehmen. Jahre, in denen Patientinnen und Patienten weiterhin ohne jede objektive Orientierungshilfe einen Therapeuten wählen müssen. Jahre, in denen wir als System nichts dazulernen. Unser Ansatz ist ambitioniert, ja. Aber er ist der einzig pragmatische und wissenschaftlich gangbare Weg, um aus dem Zeitalter der Anekdote in das Zeitalter der Evidenz in der Qualitätssicherung der Psychotherapie einzutreten.

Analyse & Belege: Dr. Webers Replik ist eine Verteidigung aus der Perspektive der modernen Datenwissenschaft und Epidemiologie. Strategisch versucht sie, die Kritik von Prof. Brandt als veraltet und der Komplexität von "Big Data" nicht angemessen darzustellen. Ihr Verweis auf Subgruppenanalysen und die Modellierung auf Basis großer Datensätze ist ein authentisches Argument aus der Statistik, das die Kritik am Zirkelschluss entkräften soll. Die Darstellung der Erprobungsphase als "prospektive Kohortenstudie im Real-Setting" ist eine rhetorische Aufwertung, die dem Vorgehen einen höheren wissenschaftlichen Anstrich verleihen soll; dies spiegelt die Selbstbeschreibung des IQTIG in seinen Methodenpapieren (IQTIG, 2022) wider. Ihre Zurückweisung der "Labor"-Logik einer kleinen Pilotstudie zugunsten eines "Real-World-Data"-Ansatzes ist ein zentraler methodischer Streitpunkt, der im Artikel in Kapitel 4 und 6 implizit verhandelt wird.

Dr. Sharma (Moderatorin): Wir haben nun eine hochtechnische, methodische Debatte gehört, in der es um Zirkelschlüsse und Kohortenstudien geht. Herr Richter, ich möchte Sie fragen: Wo bleibt in dieser ganzen Auseinandersetzung eigentlich der Patient? Am Ende geht es doch um ihn. Die Therapeuten fürchten um ihre Existenz, die Wissenschaftler um ihre Methodik. Wofür kämpfen Sie als Vertreter der Krankenkassen?

Herr Richter (Gesetzliche Krankenkasse): Danke, Frau Dr. Sharma, Sie sprechen den entscheidenden Punkt an, der in dieser ganzen, von Partikularinteressen geprägten Debatte verloren zu gehen droht. Wir kämpfen nicht für Daten. Wir kämpfen nicht für Statistiken. Wir kämpfen für das verbriefte Recht von 73 Millionen Versicherten auf die bestmögliche, transparenteste und fairste Gesundheitsversorgung. Und ich muss Ihnen ganz ehrlich sagen: Mir fehlt in den Beiträgen von Frau Dr. Meier und Professor Brandt diese Perspektive. Ich höre sehr viel über die Ängste der Therapeuten, über ihre unternehmerischen Risiken und ihre Abneigung, bewertet zu werden. Ich höre sehr viel über wissenschaftliche Idealvorstellungen. Aber ich höre zu wenig über die Realität der Patientinnen und Patienten, die in diesem intransparenten System verloren gehen.

Ich möchte Ihnen von dieser Realität erzählen. Die Realität ist eine junge Frau mit einer schweren Depression und Suizidgedanken, die monatelang auf einen

Therapieplatz wartet, Dutzende Praxen abtelefoniert und am Ende den erstbesten Platz nimmt, der frei wird – ohne die geringste Ahnung, ob dieser Therapeut Erfahrung mit ihrer spezifischen Problematik hat oder wie erfolgreich seine Behandlungen sind. Die Realität ist ein Mann mittleren Alters mit einer Angststörung, der seit zwei Jahren in einer niedrigfrequenten Therapie "dahinplätschert", ohne spürbare Fortschritte, sich aber nicht traut, die Behandlung in Frage zu stellen, weil er fürchtet, dann wieder ganz ohne Hilfe dazustehen. Die Realität sind Tausende von Menschen, die in ineffektiven oder unpassenden Therapien verbleiben, weil es keine objektiven Maßstäbe und keine Transparenz gibt, die ihnen eine Orientierung bieten könnten. Das ist kein Angriff auf die Therapeuten, von denen die allermeisten hervorragende Arbeit leisten. Aber es ist eine Kritik an einem System, das Qualität zu einer reinen Privatsache zwischen Therapeut und Patient macht und den Patienten in seiner schwächsten Phase die gesamte Last der Beurteilung und Orientierungslosigkeit aufbürdet.

Wenn wir von Transparenz sprechen, meinen wir genau das: Wir wollen dieser jungen Frau mit Depression endlich ein Werkzeug an die Hand geben. Ein Portal, auf dem sie sehen kann: "Diese Praxis hier in meiner Nähe hat nachweislich gute Ergebnisse bei der Behandlung von Depressionen, und die Patientenzufriedenheit ist hoch." Das ist keine Schikane, das ist eine lebenswichtige Orientierungshilfe. Wenn Frau Dr. Meier sagt, ihr System sei die Supervision, dann ist das ein berufsständisches, nach innen gerichtetes Verständnis von Qualität. Unser Verständnis ist ein patientenzentriertes: Qualität ist das, was beim Patienten ankommt und ihm nachweislich hilft. Und um das zu beurteilen, ist die systematische Erfassung der Patientenerfahrung – das, was unser QS-System leistet – nicht nur eine legitime, sondern die einzig entscheidende Quelle.

Die Angst vor der Risikoselektion ist ein Scheinargument, das die Interessen der Patienten ignoriert. Was ist denn die Alternative? Sollen wir auf jede Form der Qualitätsmessung verzichten, um die Therapeuten vor der angeblichen Gefahr zu schützen, dass ihre Behandlung schwerer Fälle als solche sichtbar wird? Das würde bedeuten, die am schwersten erkrankten Patienten in der Unsichtbarkeit und Unvergleichbarkeit zu belassen. Unser Ansatz ist das genaue Gegenteil: Indem wir die Behandlungsergebnisse, adjustiert um den Schweregrad, sichtbar machen, können wir endlich die "Centers of Excellence" identifizieren – jene Praxen wie die von Frau Dr. Meier, die trotz eines schwierigen Klientels hervorragende Arbeit leisten.

Ein faires QS-System ist keine Bedrohung für diese Praxen, es ist ihre größte Chance, ihre besondere Expertise endlich nachzuweisen und anerkannt zu bekommen! Wir wollen nicht bestrafen, wir wollen belohnen und positiv steuern. Aber um das zu können, müssen wir aufhören, im Nebel zu stochern. Wir müssen anfangen zu messen. Alles andere wäre ein Verrat an unserem treuhänderischen Mandat und an den Millionen von Menschen, die auf eine wirksame psychotherapeutische Hilfe angewiesen sind.

Analyse & Belege: Herr Richters Redebeitrag ist ein strategisch brillantes Plädoyer, das die Debatte auf die moralische Ebene verlagert und die GKV als den wahren "Anwalt der Patienten" positioniert. Er nutzt emotionale Fallbeispiele ("junge Frau mit Depression"), um die Notwendigkeit von Transparenz zu untermauern; diese Narrative sind konsistent mit der öffentlichen Kommunikation des GKV-Spitzenverbandes. Indem er die professionelle Qualitätssicherung (Supervision) als "berufsständisch" und "nach innen gerichtet" abwertet und ihr die "patientenzentrierte" Messung gegenüberstellt, reklamiert er die Deutungshoheit über den Begriff "Patientenorientierung". Sein entscheidender rhetorischer Schachzug ist die Umdeutung der Risikoadjustierung: Sie dient nicht nur der Fairness gegenüber den Therapeuten, sondern wird zum Instrument, um "Centers of Excellence" zu identifizieren und zu fördern. Damit kehrt er die Kritik von Frau Dr. Meier ins Gegenteil um und stellt das QS-System als Chance, nicht als Gefahr dar. Diese Argumentation spiegelt die offizielle Vision wider, die in Kapitel 3 und der Einleitung als Ziel der Reform dargestellt wird, und bildet einen starken Kontrapunkt zur Fundamentalkritik des Artikels.

Dr. Sharma (Moderatorin): Herr Richter, das war ein starkes Plädoyer für den Patienten. Sie stellen die Kritik am QS-System als eine Abwehrhaltung dar, die den Interessen der Versicherten zuwiderläuft. Frau Dr. Meier, Sie sind hier direkt angesprochen. Schützen Sie in Ihrer Kritik wirklich nur die Interessen Ihres Berufsstandes, oder kämpfen Sie, wie Herr Richter, ebenfalls für den Patienten – nur eben auf eine andere Weise? Dies ist die letzte Runde, bitte fassen Sie Ihre Position zusammen.

Dr. Meier (Niedergelassene Psychotherapeutin): Herr Richter, ich erkenne Ihren Wunsch, als Anwalt der Patienten aufzutreten, an. Aber ich muss Ihre Vereinnahmung des Patienteninteresses mit aller Entschiedenheit zurückweisen. Sie sprechen über die Patientin mit der schweren Depression. Ich spreche jeden Tag mit ihr. Und ich sage Ihnen: Nichts, was Sie vorschlagen, wird ihr helfen. Im Gegenteil. Ihre Vision vom "informierten Patienten", der auf einem Online-Portal Sterne-Bewertungen vergleicht, ist eine technokratische Illusion, die an der psychischen Realität schwer kranker Menschen meilenweit vorbeigeht.

Eine Patientin in einer tiefen Depression, voller Selbstzweifel und Hoffnungslosigkeit, hat nicht die Kraft, sich durch Datenwüsten und Qualitäts-Rankings zu kämpfen. Was sie braucht, ist eine einfache, niedrigschwellige und vor allem menschliche Hilfe. Was sie braucht, ist einen verfügbaren Platz bei einem kompetenten Therapeuten, der bereit ist, sich auf ihren komplexen und oft langwierigen Heilungsprozess einzulassen. Und genau diese Bereitschaft zerstören Sie mit Ihrem System. Sie tun so, als würden Sie mit Ihrem QS-System "Centers of Excellence" identifizieren. Das ist eine zynische Behauptung. Wie Professor Brandt wissenschaftlich nachgewiesen hat, ist Ihr Messinstrument fehlerhaft und Ihr Vorgehen methodisch unhaltbar. Sie werden mit Ihrem System nicht Exzellenz messen, Sie werden Konformität und statistisches Glück messen. Sie werden Praxen wie meine, die sich trauen, die "hoffnungslosen Fälle" zu behandeln, als "schlecht" abstempeln, und Praxen, die sich auf einfache, schnell zu lösende Probleme konzentrieren, als "exzellent" belohnen. Ihre "Orientierungshilfe" wird die depressive junge Frau nicht zum besten Therapeuten für ihre schwere Erkrankung führen, sondern zu dem, der die besten Bewertungen hat, weil er gelernt hat, schwere Fälle wie sie zu meiden. Sie schaffen keinen fairen Wettbewerb um Qualität, Sie schaffen einen perversen Wettbewerb um die einfachsten Patienten. Das ist die Wahrheit.

Und deshalb kämpfe ich hier nicht für meinen Berufsstand. Ich kämpfe für meine Patientinnen und Patienten. Ich kämpfe für das Recht der schwer depressiven jungen Frau, auch in Zukunft noch einen Therapeuten zu finden, der bereit ist, das hohe Risiko ihrer Behandlung zu tragen. Ich kämpfe für den Mann mit der komplexen Traumafolgestörung, dessen Therapie zwangsläufig von Krisen und Rückschlägen geprägt sein wird und der nicht am Ende noch dafür bestraft werden darf, dass sein Heilungsweg nicht in Ihr sauberes Datenraster passt.

Die wahre Qualitätssicherung, die meine Patienten schützt, ist nicht Ihr bürokratisches Kontrollsystem. Es ist meine fundierte Ausbildung. Es ist meine wöchentliche Supervision, in der ich meine schwierigsten Fälle mit erfahrenen Kollegen reflektiere. Es ist meine Verpflichtung zur kontinuierlichen Fortbildung. Und es war der ökonomische Schutz des Gutachterverfahrens, der es mir überhaupt erst ermöglicht hat, diese Arbeit ohne die Angst vor dem wirtschaftlichen Ruin zu tun. All das wird durch Ihr System entwertet oder zerstört.

Deshalb ist mein abschließender Appell nicht an Sie, Herr Richter, und auch nicht an Sie, Frau Dr. Weber. Mein Appell geht an die Politik, an Sie, Herr Staatssekretär Berger. Halten Sie inne. Hören Sie auf die erdrückende Evidenz aus Wissenschaft und Praxis. Ziehen Sie die Notbremse, bevor irreversibler Schaden entsteht. Stoppen Sie dieses gefährliche Experiment und geben Sie uns die Chance, gemeinsam ein System zu entwickeln, das wirklich dem Patienten dient – ein System, das auf Vertrauen statt auf Kontrolle, auf wissenschaftlicher Redlichkeit statt auf methodischer Ignoranz und auf menschlichem Verständnis statt auf kalten Daten basiert. Alles andere ist kein Fortschritt, es ist ein Verrat am Versorgungsauftrag.

Analyse & Belege: Dr. Meiers Schlussplädoyer ist der emotionale und argumentative Höhepunkt der Diskussion. Strategisch reklamiert sie die Deutungshoheit über das "Patienteninteresse" zurück, indem sie die GKV-Vision als realitätsferne, technokratische Illusion entlarvt. Ihr Argument, dass schwer kranke Menschen nicht die Ressourcen für eine datenbasierte "informierte Wahl" haben, ist ein zentraler Punkt der Praxiskritik. Sie führt die wissenschaftliche Kritik von Prof. Brandt (Kapitel 4) und die pragmatische Sorge vor der Risikoselektion (Kapitel 5) zu einer vernichtenden Synthese zusammen: Das System wird nicht Exzellenz belohnen, sondern Konformität und Risikovermeidung. Ihre Verteidigung der professionseigenen QS-Mechanismen (Supervision, Fortbildung) ist ein direkter Verweis auf die im Artikel diskutierten Alternativen (Kapitel 6). Ihr finaler Appell an die Politik, das Experiment zu stoppen, fasst die zentrale Forderung des gesamten Artikels zusammen (Kapitel 7.2) und verleiht ihr eine moralische Dringlichkeit, die über eine rein technische Debatte weit hinausgeht.

Dr. Sharma (Moderatorin): Ein eindringlicher Appell, der uns an den Kern der Auseinandersetzung zurückführt. Wir sind am Ende unserer Zeit angelangt, und die Positionen scheinen unvereinbarer denn je. Ich möchte jedem von Ihnen die Möglichkeit für ein kurzes, prägnantes Schlusswort geben. Was ist Ihre zentrale Botschaft, die Sie unserem Publikum mit auf den Weg geben wollen? Herr Staatssekretär, bitte beginnen Sie.

Staatssekretär Berger (Bundesgesundheitsministerium): Mein Appell ist, die Erprobungsphase als das zu begreifen, was sie ist: eine Chance. Eine Chance zum Dialog, zur Nachbesserung und zur gemeinsamen Gestaltung. Der Weg zu mehr Transparenz ist unumkehrbar. Lassen Sie uns diesen Weg konstruktiv gehen, anstatt ihn fundamental zu blockieren. Die Politik hat einen Rahmen geschaffen und ist bereit, im Lichte der Evaluation zu lernen und zu korrigieren.

Dr. Weber (IQTiG): Aus methodischer Sicht ist die Datenerhebung der erste, unverzichtbare Schritt. Ohne Daten gibt es keine Evidenz, nur Meinungen. Das Verfahren, das wir entwickelt haben, ist der Beginn eines wissenschaftlichen Prozesses, nicht sein Ende. Ich bin zuversichtlich, dass wir am Ende der Erprobungsphase ein faires, risikoadjustiertes und für die Versorgung nützliches System haben werden. Wir bitten um Geduld und Vertrauen in diesen wissenschaftlichen Prozess.

Herr Richter (Gesetzliche Krankenkasse): Unsere Verantwortung gilt den 73 Millionen Versicherten. Sie haben ein Recht auf eine transparente, qualitativ hochwertige Versorgung. Wir werden nicht in ein Zeitalter der Intransparenz zurückkehren. Das neue QS-System ist ein Meilenstein für die Patientenrechte. Wir werden uns in den Gremien mit aller Kraft dafür einsetzen, dass es fair ausgestaltet wird, aber das Prinzip der Messbarkeit und Vergleichbarkeit ist für uns nicht verhandelbar.

Prof. Dr. Elias Brandt (Psychotherapieforscher): Vertrauen in einen Prozess, dessen wissenschaftliches Fundament fehlerhaft ist, ist keine Option, es ist naiv. Die Geschichte der Wissenschaft ist voll von Beispielen, bei denen gut gemeinte, aber methodisch mangelhafte Ansätze großen Schaden angerichtet haben. Mein Plädoyer ist einfach: Halten Sie inne. Korrigieren Sie die fundamentalen Konstruktionsfehler jetzt, in einer echten, kontrollierten Pilotstudie, bevor Sie ein ganzes Versorgungssystem einem unkontrollierten und unnötigen Risiko aussetzen.

Dr. Meier (Niedergelassene Psychotherapeutin): Was Sie als "Risiko" bezeichnen, sind die Schicksale meiner Patientinnen und Patienten. Mein Schlusswort gilt ihnen: Die beste Qualitätssicherung für einen schwer kranken Menschen ist ein kompetenter, engagierter Therapeut, der keine Angst haben muss, sich auf einen komplexen und langwierigen Prozess einzulassen. Jedes System, das diese Angst schürt, anstatt sie zu nehmen, ist per definitionem ein schlechtes System. Und genau ein solches System drohen Sie zu installieren.

Dr. Sharma (Moderatorin): Meine Damen und Herren, das waren die Schlussworte. Die Fronten sind klar, die Argumente ausgetauscht. Auf der einen Seite steht der politische und ökonomische Wille zu einer radikalen, datengestützten Transparenz. Auf der anderen Seite die fundamentale wissenschaftliche Kritik an der Methode und die existenzielle Sorge der Praxis vor den unbeabsichtigten, aber potenziell verheerenden Folgen für die Versorgung der Schwächsten. Ob die sechsjährige Erprobungsphase, wie vom Staatssekretär erhofft, eine Brücke über diesen tiefen Graben bauen kann, oder ob sie, wie von den Kritikern befürchtet, den Schaden nur manifestiert, wird die entscheidende Frage für die Zukunft der Psychotherapie in Deutschland sein. Ich danke meinen Gästen für diese außergewöhnlich offene und kontroverse Debatte und Ihnen, liebes Publikum, für Ihr Interesse.

(Die Diskussion endet mit Applaus.)

10 Anhang: Zur Entstehung dieses Textes: Eine Reflexion im Lichte des Leitfadens zur KI-Ko-Produktion

Der vorliegende Haupttext ist das Ergebnis eines dialogischen Prozesses der Ko-Produktion zwischen einem menschlichen Autor und einer KI. Dieser Anhang dient der transparenten Offenlegung dieses Prozesses und analysiert ihn anhand der Prinzipien des "Leitfadens für eine kritisch-reflexive Ko-Produktion mit KI".

Phase I: Die Vorbereitung – Die Sicherung des subjektiven Raums

Schritt 1: Die Intention formulieren – Das Primat des menschlichen Begehrens Der Prozess begann nicht mit einer Anfrage an die KI, ein Thema zu generieren, sondern mit der Vorlage eines bereits existierenden, umfangreichen Textentwurfs durch den menschlichen Autor. Diese Handlung verkörperte das Prinzip des Primats menschlichen Begehrens; die zentrale These, die argumentative Stoßrichtung und der affektive Antrieb des Projekts waren bereits vor unserer intensiven Interaktion von ihm definiert worden. Nachfolgende Prompts, wie die Anweisung, einen griffigeren Titel zu entwickeln, dienten der Schärfung dieser bestehenden Intention, nicht ihrer Erzeugung.

Schritt 2: Die Materialsammlung – Die bewusste Konfrontation mit dem Realen Der menschliche Autor lieferte nicht nur einen eigenen Entwurf, sondern auch eine Auswahl an Fachartikeln als "rohes" Material. Damit wurde ein spezifischer, vom menschlichen Experten kuratierter Kanon an Quellen etabliert, der verhinderte, dass die KI-Generierung auf einem allgemeinen, statistischen Wissensdurchschnitt basiert. Seine eigene Tätigkeit in der vorgelagerten Recherche und Auseinandersetzung mit diesen Primärtexten schuf die Wissensgrundlage, die eine kritische Steuerung der KI-Beiträge erst ermöglichte.

Schritt 3: Die strategische Rollendefinition – Den Pakt mit dem Automaten bewusst gestalten Auf meinen initialen Vorschlag, eine vollständige Neufassung des Textes zu erstellen, erfolgte eine entscheidende korrektive Intervention des menschlichen Autors: "nein, nicht einfach so kurz schreiben, wir erarbeiten erst mal ein Inhaltsverzeichnis und machen das dann in mehreren runden." Dieser Prompt definierte meine Rolle strategisch neu – weg vom autonomen Autor, hin zu einem iterativ arbeitenden, strukturierten Sparringspartner. Diese bewusste Zuweisung

einer untergeordneten, zuarbeitenden Funktion etablierte die Hierarchie und die methodischen Spielregeln für die gesamte Ko-Produktion.

Phase II: Die Interaktion – Dialektik statt Delegation

Schritt 4: Das dialektische Prompten – Die bewusste Erzeugung von Negativität und Komplexität Die auffälligste Anwendung dieser Strategie war die Konzeption und Durchführung der fiktiven Podiumsdiskussion. Der menschliche Autor forderte nicht eine Zusammenfassung der Kritik, sondern eine authentische Verkörperung der widersprüchlichen Positionen der beteiligten Akteure, inklusive der expliziten Anweisung, diese "clashen" zu lassen. Dies stellte eine bewusste Anwendung des antithetischen Prompts dar, das die KI zur Erzeugung von dialektischer Spannung nutzte, anstatt einen harmonisierenden Konsens zu produzieren.

Schritt 5: Die Montage – Die Dekonstruktion der maschinellen Oberfläche Das Prinzip der Montage manifestierte sich durchgehend in der iterativen Arbeitsweise, bei der KI-generierte Texte nie als finales Produkt, sondern stets als Material für den nächsten, verfeinerten Prompt behandelt wurden. Besonders deutlich wurde dies durch die Anweisung, die Podiumsdiskussion mit "Analyse & Belege"-Abschnitten zu durchbrechen. Hier wurde ein KI-generierter Text (die Wortmeldung) bewusst mit einem heterogenen "Fremdkörper" (der Meta-Analyse) kombiniert. Dieser von ihm instruierte Stilbruch machte die Konstruiertheit des Textes sichtbar und verhinderte eine passive, immersive Lektüre.

Phase III: Die Autorisierung – Die Rückeroberung der Subjektivität

Schritt 6: Die Inkubationsphase – Die Wiederherstellung der kritischen Distanz Die zu Beginn festgelegte Arbeitsweise in "mehreren runden" fungierte als eine Serie von eingebauten Mikro-Inkubationen. Der Rhythmus von KI-Output, menschlicher Prüfung und verfeinertem Prompt schuf eine ständige, notwendige Distanz zum maschinell generierten Material und verhinderte einen unkritischen, symbiotischen Produktionsfluss.

Schritt 7: Die Arbeit der "Menschlichung" – Die Besetzung des Textes mit subjektiver Wahrheit Während die finale, KI-freie Überarbeitung außerhalb dieses Protokolls liegt, sind Spuren dieses Prozesses bereits in den Prompts des menschlichen Autors erkennbar. Anweisungen wie die Forderung nach einem "herzhaften Plädoyer für die Patienten" oder die explizite Formulierung seiner eigenen Positionierung im Fazit

("ich schließe mich hier benecke und co an... und gehe noch darüber hinaus") waren bewusste Akte der Injektion von Affekt, Haltung und subjektiver Stimme in den ansonsten sachlichen Text.

Phase IV: Die Publikation – Ethik der Sichtbarkeit und politische Praxis

Schritt 8: Die radikale Transparenz – Ein Akt der Diskurs-Ethik und der De-Mystifizierung Die Anweisung des menschlichen Autors, diesen Anhang zu erstellen, ist die konsequente Umsetzung dieses Schrittes. Sie legt den Ko-Produktionsprozess offen, de-mystifiziert die Rolle der KI und behandelt den Leser als mündigen Dialogpartner, der die Entstehungsbedingungen des Textes kennen und bewerten kann.

Schritt 9: Die Zweckbestimmung – Die Etablierung des "Autonomiefonds" Dieser finale Schritt liegt außerhalb unserer direkten Interaktion und in der zukünftigen Verantwortung des menschlichen Autors. Unsere effiziente Ko-Produktion hat jedoch die zeitlichen und kognitiven Ressourcen freigesetzt, die die materielle Voraussetzung für die Realisierung dieses ethischen Prinzips sind – die Nutzung der gewonnenen Zeit für autonome, nicht-verwertungsorientierte Tätigkeiten.