

Gesundheitsökonomie als Entscheidungshilfe:  
Kritische Diskussion der Grundlagen,  
-Anwendung auf die Behandlung von Schizophrenie und  
schwerer Sepsis-  
Empirischer Test.

Dissertation zur Erlangung des  
Doktorgrades (Dr. rer.soc.)  
des Fachbereichs Gesellschaftswissenschaften  
der Justus-Liebig-Universität Gießen

Vorgelegt von  
Johannes Clouth

aus Darmstadt, Deutschland  
Februar 2007

Erstgutachter: Prof. Dr. Adalbert Evers

Ich erkläre:

Ich habe die vorgelegte Dissertation selbstständig und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündliche Auskünfte beruhen, sind als solche kenntlich gemacht.

Johannes Clouth

# Inhaltsverzeichnis

	Seite
<b>1. Einführung.....</b>	<b>7</b>
<b>1.1 Zielsetzung .....</b>	<b>7</b>
<b>1.2 Begriffsbestimmungen .....</b>	<b>7</b>
<b>1.3 Stand der Forschung.....</b>	<b>10</b>
<b>1.4 Mögliche Beiträge der Gesundheitsökonomie .....</b>	<b>16</b>
<b>1.5 Anwendungshindernisse für gesundheitsökonomische Beiträge.....</b>	<b>18</b>
1.5.1 Gesundheitspolitische Rahmenbedingungen .....	18
1.5.2 Methodische Probleme.....	23
1.5.2.1 Probleme der Definition des Nutzens .....	24
1.5.2.1.1. <i>Psychologische Dimensionen des Nutzens.....</i>	<i>24</i>
1.5.2.1.2 <i>Nutzen in den Wirtschaftswissenschaften.....</i>	<i>25</i>
1.5.2.1.3 <i>Probleme der Definition des Nutzens .....</i>	<i>25</i>
1.5.2.1.4 <i>Probleme der Messung des Nutzens.....</i>	<i>26</i>
1.5.2.1.5 <i>Nutzen und Grenznutzen .....</i>	<i>27</i>
1.5.2.2 Analyse des QALY-Konzept .....	30
1.5.2.3 Analyse patientenrelevanter Endpunkte.....	33
<b>1.6 Zusammenfassung: Einführung und Analyse der Beiträge der eingereichten Artikel.....</b>	<b>44</b>
1.6.1 Zusammenfassung Einführung.....	44
1.6.2 Analyse der Beiträge der eingereichten Artikel .....	45
<b>2. Begründung der ausgewählten Krankheiten .....</b>	<b>48</b>
<b>2.1 Schizophrenie .....</b>	<b>48</b>
<b>2.2 Schwere Sepsis.....</b>	<b>49</b>
<b>3. Begründung der eingereichten Arbeiten .....</b>	<b>50</b>
<b>3.1 Wissenschaftstheoretische Überlegungen zur Werturteilsproblematik dargestellt am Beispiel der so genannten Vierten Hürde.....</b>	<b>50</b>

3.1.1 Kontext des Entstehens .....	50
3.1.2 Beiträge des Artikels zur Fragestellung der Arbeit.....	51
3.1.3 Limitierungen.....	51
3.1.4 Inhalt .....	52
<b>3.2 Kosten der Frühverrentung am Beispiel der Schizophrenie .....</b>	<b>53</b>
3.2.1 Kontext des Entstehens .....	53
3.2.2 Beiträge des Artikels zur Fragestellung der Arbeit.....	53
3.2.3 Limitierungen.....	54
3.2.4 Inhalt .....	54
<b>3.3 Kostenanalyse der Schizophrenie-Behandlung in Deutschland .....</b>	<b>55</b>
3.3.1 Kontext des Entstehens .....	55
3.3.2 Beiträge des Artikels zur Fragestellung der Arbeit.....	56
3.3.3 Limitierungen.....	56
3.3.4 Inhalt .....	56
<b>3.4 Constructing an Index for the Subjective Well-being Under Neuroleptics Scale (SWN), Short Form: Applying Structural Equation Modelling for testing Reliability and Validity of the Index...</b>	<b>57</b>
3.4.1 Kontext des Entstehens .....	57
3.4.2 Beiträge des Artikels zur Fragestellung der Arbeit .....	58
3.4.3 Limitierungen.....	58
3.4.4 Inhalt .....	59
<b>4. Eingereichte Arbeiten .....</b>	<b>60</b>
<b>4.1 Artikel:</b>	
<b>Wissenschaftstheoretische Überlegungen zur Werturteilsproble-</b> <b>matik der so genannten Vierten Hürde.....</b>	<b>60</b>
<b>4.2 Artikel:</b>	
<b>Kosten der Frühverrentung am Beispiel der Schizophrenie .....</b>	<b>73</b>
<b>4.3 Artikel:</b>	
<b>Kostenanalyse der Schizophrenie-Behandlung in Deutschland .....</b>	<b>90</b>

**4.4 Artikel:**

**Constructing an Index for the Subjective Well-being Under Neuroleptics Scale (SWN), Short Form: Applying Structural Equation Modeling for Testing Reliability and Validity of the Index ..... 104**

**5. Schlussfolgerungen..... 134**

**5.1 Mögliche Beiträge der Gesundheitsökonomie zur Überwindung der durch politische Rahmenbedingungen erzeugten Probleme ..... 134**

5.1.1 Theoretische Grundlage ..... 134

5.1.2 Festlegung von Grenzwerten ..... 135

5.1.3 Emissionsgebühr ..... 136

5.1.4 Emissionszertifikate ..... 137

5.1.5 Integrierte Versorgung ..... 138

5.1.6 Disease-Management-Programme ..... 139

5.1.7 Gesundheitskarte ..... 139

5.1.8 Modellierung ..... 141

5.1.9 Zentralregister ..... 142

**5.2 Mögliche Beiträge der Gesundheitsökonomie zur Definition und Messung des Begriffs „Nutzens“ ..... 142**

**5.3. Mögliche Beiträge der Gesundheitsökonomie zur Überwindung methodischer Probleme..... 149**

**5.4 Ausblick..... 153**

**6. Literatur..... 156**

# 1. Einführung

## 1.1 Zielsetzung

Gegenstand der vorliegenden Promotion ist die Analyse des Beitrags der Gesundheitsökonomie als Entscheidungshilfe im Allgemeinen und an ausgewählten Beispielen. Diese Zielsetzung untergliedert sich in drei Subziele.

- Zum einen wird untersucht, welche **Beiträge** die Gesundheitsökonomie für die Entscheidungshilfe leisten kann.
- Zum anderen werden aber auch die **Anwendungshindernisse** gesundheitsökonomischer Erkenntnisse analysiert.
- Auf dieser Grundlage aufbauend werden als dritte Zielsetzung mögliche Ansätze vorgestellt, die den Beitrag der Gesundheitsökonomie als Entscheidungshilfe **verbessern** können.

Die Terminologie in der Gesundheitsökonomie ist nicht immer einheitlich (Drummond, S. 143, 1997). Aus diesem Grund werden die Fachtermini vorab erklärt.

Die Analyse gesundheitsökonomischer Evaluationen beziehungsweise deren Begrenzung ist nur möglich, wenn der aktuelle Stand der Forschung mit neueren Ansätzen verglichen wird. Aus diesem Grund wird nach den Begriffsbestimmungen der aktuelle Stand der Forschung beschrieben.

## 1.2 Begriffsbestimmungen

Der erste Terminus Technicus, der beschrieben wird, ist die „**Gesundheitsökonomie**“.

Im Sozialgesetzbuch (SGB) V wurde 1988 die Beitragsstabilität für die gesetzliche Krankenversicherung festgeschrieben, so dass es nun darum geht die zur Verfügung stehenden knappen Mittel optimal einzusetzen. Dies ist das Betätigungsfeld der Wirtschaftswissenschaften. Die Zielsetzung der Gesundheitsökonomie liegt in der Optimierung der Allokation knapper Ressourcen im Gesundheitswesen (Drummond, S. 7, 1997). Diese Aufgabenstellung lässt sich an folgendem hypothetischen Beispiel verdeutlichen. Nehmen wir an, der Gesundheitsminister eines Landes verfüge über eine

begrenzte Menge Geld und müsse diesen Betrag zur Behandlung der Krankheiten Aids und Diabetes einsetzen, wobei weitere Krankheiten nicht existierten. Welche Daten müssen erhoben und welche Methoden eingesetzt werden, um dieses Problem lösen zu können? Genau diese Fragestellung ist Gegenstand der Gesundheitsökonomie und lässt sich exemplarisch wie folgt darstellen:

<b>Kosten life years gained</b>				
<b>Kosten life years gained</b>	<b>Kosten pro Patient</b>	<b>Anzahl Patienten</b>	<b>Gesamtkosten</b>	<b>Kum. Kosten</b>
Diabetes	3	20	60	60
Depression	5	10	50	110
Mukoviszidose	10	3	30	140
AIDS	20	1	20	160
Budget 160				

Abbildung 1

Die Kosten zur Heilung eines Patienten sind bei Diabetes am geringsten, allerdings ist die Anzahl der Patienten am größten. Die Kosten nehmen in unserem Beispiel ständig ab und die Prävalenz nimmt zu. Die Heilung eines Patienten mit Depression kostet weniger als bei Mukoviszidose und weniger als bei AIDS. Die vorgegebenen Mittel werden optimal eingesetzt. Jede andere Allokation würde zur Unterversorgung einer der genannten Gruppen führen.

Ein weiterer zentraler Begriff stellt der Terminus Technicus „**Kosten**“ dar. Dieser Begriff wird in direkte und indirekte Kosten unterteilt, wobei jeweils tangible oder intangible Kosten entstehen können (Schöffski, S.163, 2000). Unter Kosten versteht man den in Geld bewertbaren Ressourcenverbrauch, der mit der Ausführung einer Behandlung verbunden ist. Direkte tangible

Kosten messen den in Geld bewerteten Ressourcenverbrauch, der direkt durch die Behandlung verursacht ist. Hierzu zählen finanzielle Aufwendungen für die ärztliche Behandlung oder für die Medikation. Intangible Kosten sind z.B. Schmerzen bei der Behandlung. Indirekte tangible Kosten sind definiert durch die Verringerung der gesamtwirtschaftlichen Produktivität. Ein Beispiel für intangible indirekte Kosten wäre die Einbuße an Lebensqualität wegen Gefahr durch Ansteckung.

Das ökonomische Prinzip besagt, dass ein Ziel mit minimalem Einsatz oder dass mit gegebenem Einsatz möglichst viel erreicht werden soll. Das Ziel der Gesundheitsökonomie liegt in der Optimierung des „Nutzens“ (im obigen Beispiel wurde Heilung mit Nutzen gleichgesetzt, was nun differenzierter betrachtet werden soll). Damit wird der Nutzen zu einem zentralen Begriff in der Gesundheitsökonomie (Schöffski et al, 2000). Somit stellen Einsparungen bei den direkten tangiblen Kosten einen direkten tangiblen Nutzen dar. Ebenso wird der indirekte und intangible Nutzen definiert. Diese Definition ist kritisch zu sehen und deshalb wird der Terminus „Nutzen“ im Kapitel 1.5 „Anwendungshindernisse“ weiter analysiert und diskutiert.

Gesundheitsökonomische Studien über Maßnahmen sind unter Umständen nicht vergleichbar, weil die Kosten verschiedener Teilnehmer im Gesundheitssystem berechnet werden. Man spricht hier von der „Perspektive“. Es ist daher unabdingbar, die Perspektive offen zu legen. Man unterscheidet vier Perspektiven, wobei diese sich nicht ausschließen.

- Die gesellschaftliche Perspektive ist die umfassendste und berücksichtigt alle Kosten, egal wer sie trägt.
- Eine zweite Perspektive, die häufig gewählt wird, ist die Krankenkassensicht. Für die Krankenkassen sind nicht mehr alle Kosten wie bei der gesellschaftlichen Sichtweise relevant, da sie nicht alle in ihrer Budgetverantwortung liegen. Einsparungen z.B. bei indirekten Kosten wie Frühverrentung sind für die Krankenkassen nicht entscheidungsrelevant.
- Eine weitere Perspektive ist die Krankenhaussicht. Auch hier stehen die Verantwortlichen vor der Aufgabe ihr Budget zu optimieren. Fol-

gekosten, die außerhalb des Krankenhauses entstehen, sind zwar gesamtwirtschaftlich und für die Krankenkassen bedeutsam, aber nicht für das Krankenhaus.

- Die Perspektive der Patienten und Angehörigen nimmt insofern eine Sonderstellung ein, weil es hier um die Optimierung des Nutzens und nicht um die Optimierung eines Budgets geht.
- Letztendlich kann noch die Perspektive der niedergelassenen Ärzte eingenommen werden, die das ihnen zur Verfügung gestellte Budget optimieren müssen ohne Berücksichtigung möglicher Effekte auf die anderen Perspektiven.

### 1.3 Stand der Forschung

Hinter dem Terminus „gesundheitsökonomische Evaluation“ steht kein einheitliches Studiendesign. In einem ersten Schritt werden zuerst die Studienformen und anschließend der Ablauf gesundheitsökonomischer Evaluationen beschrieben.

Man unterscheidet verschiedene Studienformen, die Kosten- und Nutzenkomponenten unterschiedlich berücksichtigen. Die Wahl der Analyseart hängt von der Zielsetzung der Studie ab. Diese Studien kann man in die Kategorien vergleichende und nicht vergleichende Studien unterteilen. Obwohl für die Optimierung der Ressourcenallokation generell vergleichende Studien notwendig sind, haben für bestimmte Fragestellungen auch nicht-vergleichende Studien ihre Berechtigung.

Zu den nicht vergleichenden Studien zählen die Kosten-Analyse und die Krankheitskosten-Analyse, wobei bei Krankheitskosten-Analysen oft eine epidemiologische Studie eingeschlossen wird. Zu den vergleichenden Studien zählen die Kosten-Kosten-Analyse, die Kosten-Nutzen-Analyse, die Kosten-Wirksamkeitsanalyse (häufig auch Kosten-Effektivitätsanalyse genannt) und die Kosten-Nutzwert-Analyse.

Die einfachste Form ist die Kosten-Analyse. Hier werden ausschließlich die Kosten einer Maßnahme berechnet wie z.B. die Kosten für eine Stentimplantation. Allein aus dieser Kenntnis heraus lässt sich keine Allokations-

entscheidung treffen. Hierzu ist der Vergleich mit Alternativen notwendig. Dieser Vergleich wird möglich, wenn eine Alternative auch mit einer Kostenanalyse berechnet wird. In diesem Falle spricht man von einer Kosten-Kosten-Analyse.

Krankheitskostenstudien werden eingesetzt, um die gesamtwirtschaftliche Bedeutung von Krankheiten zu ermitteln. Dabei erfolgt keine Differenzierung nach einzelnen Maßnahmen, sondern vielmehr werden direkte und indirekte Kosten für eine Krankheit als Ganzes evaluiert. Bei Krankheitskostenstudien können zum einen die Kosten pro Patient ermittelt, zum anderen die Kosten aller Erkrankten ermittelt werden. Zur Berechnung der Kosten aller Erkrankten ist es notwendig, die Anzahl der Patienten möglichst aufgebrochen nach wichtigen Merkmalen wie Schweregrad oder Geschlecht zu kennen. Die Ermittlung dieser Zahlen erfolgt durch epidemiologische Studien.

Für die Allokation knapper Ressourcen ist es nicht nur wichtig zu wissen welche von mehreren die kostengünstigste darstellt, sondern welche Alternative den besten Nutzen generiert. Antwort auf diese Fragestellung versuchen die Kosten-Nutzen-Analyse, die Kosten-Wirksamkeitsanalyse und die Kosten-Nutzwert-Analyse zu geben.

Bei der Kosten-Nutzen-Analyse findet im Gegensatz zur Kosten-Kosten-Analyse nur eine Berechnung statt. Der Nutzen der zu bewertenden Maßnahme entspricht dabei den Kosten der alternativen Therapie, die durch deren Einsatz vermieden werden könnte. Dieser durch Geldeinheiten operationalisierte Nutzen wird von den Kosten der zu bewertenden Maßnahme subtrahiert und man erhält einen positiven oder negativen Saldo. Der Nutzen kann auch alternativ durch eine Division erfolgen. Bei der Kosten-Nutzen-Analyse werden also alle Kosten- und Nutzenkomponenten monetär bewertet. Dies gilt für Änderungen der Lebensqualität genauso wie für gewonnene Lebensjahre. Die Zuweisung eines bestimmten Geldbetrages für Lebensqualität oder für ein Lebensjahr wird in der Öffentlichkeit als ethisch fragwürdig angesehen. Die Reduzierung aller Effekte auf monetäre Größen hat sich nicht als gangbarer Weg zur Allokationsoptimierung im Gesund-

heitswesen erwiesen. Aus diesen Gründen wird allgemein von der Kosten-Nutzen-Analyse abgeraten.

Eine Möglichkeit Therapien zu evaluieren ohne die problematische Bewertung der Ergebnisse in Geldeinheiten, stellt die Kosten-Wirksamkeits-Analyse dar. Dabei werden die nicht in Geld bewertbaren Ergebnisse in natürlichen Einheiten gemessen, die meist von Medizinern festgelegt werden. Beispiele solcher so genannten Outcome-Parametern sind z.B. Senkung des Blutdrucks, Tumorgröße oder Lebensverlängerung in Jahren. Diesem messbaren Erfolg werden die Kosten der Behandlung gegenübergestellt. Mit diesem Vorgehen wird die Vergleichbarkeit zweier oder mehrerer Maßnahmen möglich. Das Ergebnis einer Kosten-Wirksamkeit können z.B. die Kosten für ein gewonnenes Lebensjahr sein. Kostet ein gewonnenes Lebensjahr bei Therapie A 10.000 € und bei Therapie B 15.000 €, dann ist Therapie A vorzuziehen.

Ein Kritikpunkt an der Kosten-Wirksamkeits-Analyse stellt die Tatsache dar, dass ein Patient zwar Lebensjahre gewinnt, aber u.U. nur unter großen Schmerzen existieren kann. Das Ziel aus Patientensicht sollte also nicht nur die Lebensverlängerung, sondern die Lebensverlängerung unter Berücksichtigung der Lebensqualität sein. Damit bleibt aber auch die Frage offen, welche Endpunkte herangezogen werden sollen (vgl. 1.5.2.3).

Zusätzlich soll eine Normierung der Ergebnisse für alle Krankheiten erfolgen, d.h. jede Therapie soll nach dem gleichen Muster bewertbar sein. Das hierfür am häufigsten eingesetzte Verfahren ist das QALY-Konzept (Quality adjusted life years), eine Kosten-Nutzwert-Analyse. Beim QALY-Konzept werden die beiden Dimensionen Lebensqualität und Lebenserwartung zu einem neuen Aggregat zusammengefasst, wobei es sich um ein eindimensionales Maß handelt. Im Gegensatz zur Kosten-Wirksamkeitsanalyse, bei der verschiedene Outcome-Parameter, wie gewonnene Lebensjahre, mit z.B. Lebensqualität verglichen werden müssen, findet die Kosten-Nutzwert-Analyse mit dem QALY einen gemeinsamen Nenner, der Allokationsentscheidungen ermöglicht.

Diese Studien werden in Deutschland in folgender Weise durchgeführt (Schöffski, S. 205ff, 2000):

- In einem ersten Schritt – und dies geschieht in Form einer epidemiologischen Studie – ermittelt man die Inzidenz und Prävalenz der behandelten Patienten, womit das gesamte Patientengut berechnet werden kann. Die methodische Vorgehensweise wird in einer Studie zur Ermittlung der Prävalenz der Schizophrenie in Deutschland anschaulich dargestellt (Besthorn, 1999). Zur Ermittlung der Anzahl der Schizophreniepatienten und ihrer Verteilung auf die Versorgungseinrichtungen wurde wie folgt vorgegangen. Zuerst wurden die betreuenden Institutionen ermittelt. Im Einzelnen waren dies niedergelassene Fachärzte, Institutsambulanzen, Krankenhäuser, Pflegeheime, betreutes Wohnen und Rehabilitation. Auf dieser Basis wurde eine repräsentative Stichprobe der betreuenden Institutionen gezogen. Anschließend wurde die Anzahl der Patienten in den betreuenden Institutionen gemessen. Letztendlich wurde die Gesamtzahl der Schizophreniepatienten in Deutschland hochgerechnet, sodass man nicht nur wusste, wie viele Patienten in Deutschland behandelt werden, sondern darüber hinaus, wie sich diese Patienten auf die betreuenden Institutionen verteilen.
- In einem zweiten Schritt, der Krankheitskostenstudie, erhebt man die Kosten pro Patient. Analog zu dem Design von Besthorn wurden die für die Behandlung relevanten Ressourcen in einer Untersuchung zu den Kosten der Schizophrenie in Deutschland erhoben (Schulenburg, 1998). Für die oben genannten Institutionen wurden die Ressourcen wie Medikamente, Krankenhaustage, Rehabilitation etc. gemessen und mit Preisen versehen und damit wurden sowohl die direkten als auch die indirekten Kosten erfasst. Die Studie ist insofern bemerkenswert, weil nicht nur die Gesamtkosten, sondern die Kosten der jeweils betreuenden Institutionen einzeln ausgewiesen wurden. Durch die Multiplikation der Kosten pro Patient mal Anzahl der Patienten erhält man den „Budget Impact“. Solche Analysen können, wie

die Studie des „Budget Impacts atypischer Neuroleptika“ (Eichmann, 2004) zeigt, sehr detailliert sein. Der Budget Impact wurde jeweils für Atypika und Typika für die Perspektiven der Krankenkasse, für die Perspektive der betreuenden Institutionen, der Perspektive der kassenärztlichen Vereinigungen, der gesellschaftlichen Perspektive ohne Frühverrentung und der gesellschaftlichen Perspektive mit Frühverrentung berechnet. Dabei wurde jeweils der Anteil der Atypika von 0% bis 100% berechnet. Mit dieser Analyse wissen nun die Entscheider im Gesundheitswesen, wie hoch die Kosten einer Behandlung je Perspektive sind.

Es stellt sich nun die Frage, ob sich diese Investition lohnt oder ob das Geld nicht in eine andere Therapie investiert werden soll. Eine Antwort auf diese Fragen geben die Kosteneffektivitätsstudie und die Kostennutzwertanalyse (Schöffski, S. 192ff, 2000).

- In einem dritten Schritt werden Kosteneffektivitätsstudien und Kostennutzwert-Analysen durchgeführt. Die Kosteneffektivitätsstudie gibt Antwort auf die Frage, wie viel Geld die Heilung eines Patienten kostet. Je geringer der Preis ist, desto eher lohnt es sich in diese Therapie zu investieren. Im Folgenden wird am Beispiel der schweren Sepsis gezeigt, wie die Wertigkeit eines Medikamentes berechnet werden kann (Neilson, 2003). Diese Krankheit eignet sich insofern gut für die Veranschaulichung einer Kosteneffektivitätsanalyse, weil die Lebensqualität nach Überleben einer schweren Sepsis nicht beeinflusst wird. Den Maßstab bilden also das Überleben bzw. die Kosten pro gewonnenem Lebensjahr. Zur Berechnung der Vorteilhaftigkeit einer Therapie geht man wie folgt vor: Mit Hilfe eines Entscheidungsbaums berechnet man die Kosten pro gewonnenem Lebensjahr sowohl für die Standardtherapie als auch für den Einsatz der innovativen Therapie.

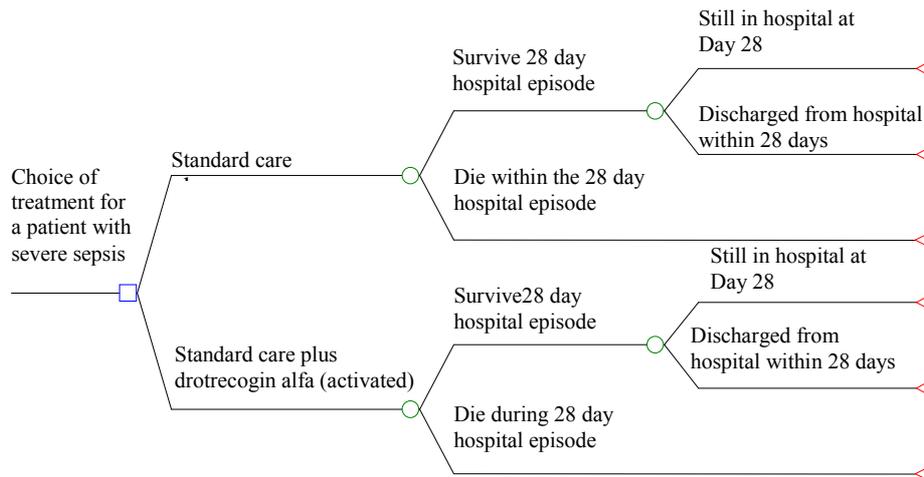


Abbildung 2: Beispiel eines Entscheidungsbaums für die Behandlung der schweren Sepsis.

Die Kosten der Standardtherapie sind aus Studien bekannt. Die Überlebensraten der zu vergleichenden Therapien und der zusätzliche Ressourcenverbrauch für die innovative Therapie, der sich hier auf die innovative Medikation beschränkt, stammen aus der doppel-blind randomisierten Zulassungsstudie. Die zusätzlichen Kosten bedingt durch die innovative Therapie werden durch die zusätzlich gewonnenen Lebensjahre dividiert und man erhält die Kosten pro gewonnenem Lebensjahr. Die Kosten pro gewonnenem Lebensjahr betragen für das Medikament „Xigris“ ca. 14.100 €. Ein Vergleich mit Therapien wie Herztransplantation oder Ovarian-Krebs zeigen, dass der Einsatz von „Xigris“ eine kostengünstige und damit kosteneffektive Alternative darstellt.

Die Kosten-Nutzwert-Analyse verfolgt eine anspruchsvollere Zielsetzung. Sie fragt nicht nur, ob ein Patient geheilt wird, sondern darüber hinaus, wie lange der Patient mit welcher Lebensqualität noch leben wird.

Bei dem QALY-Konzept wird davon ausgegangen, dass sich das menschliche Leben anhand der zwei Dimensionen „Restlebenserwartung“ und „Lebensqualität“ darstellen lässt (Schöffski, S. 367ff, 2000). Die Bewertung der Lebensqualität wird präferenzbasiert vorgenommen. Das am häufigsten eingesetzte Instrument zur Messung der präferenzbasierten Lebensqualität ist der EQ5D (EuroQoL). Dieses Instrument misst fünf Dimensionen mittels fünf Indikatoren, die jeweils drei Abstufungen enthalten. Damit ergeben

sich  $3^5$  bzw. 243 verschiedene Gesundheitszustände. Die Nutzwerte ausgewählter Gesundheitszustände wurden mit der Methode des Time Trade Off ermittelt. Damit konnten mit Hilfe eines regressionsanalytischen Ansatzes die Nutzenwerte für alle anderen Gesundheitszustände geschätzt werden. Die Werte vom bestmöglichen Zustand (=völlige Heilung) bis zum schlechtesten Zustand, hier der Tod, wurden zwischen 1 und 0 standardisiert. Diese Werte werden „Utilities“ genannt.

Durch Multiplikation der Utilities mit der Restlebenserwartung erhält man die QALIES.

Mit diesem Verfahren wurde erreicht, dass man medizinischen Maßnahmen eine einzige Kenngröße zuordnen kann. Durch diese Vorgehensweise werden Allokationsentscheidungen möglich. Stehen zwei Alternativen zur Auswahl, so ist jene mit dem geringeren QALY-Preis vorzuziehen. Eine Übersicht der Kosten pro QALY liefern die „League Tables“<sup>1</sup>. Die Frage, wie viel ein QALY überhaupt kosten darf, ist zur Zeit nicht beantwortet.

Zusammenfassend lässt sich der Stand der Forschung so beschreiben, dass mit den oben genannten Studienarten die Gesundheitsökonomie über ein differenziertes Instrumentarium verfügt, um Entscheidern auch im deutschen Gesundheitswesen Entscheidungsunterstützung bei Allokationsentscheidungen zu ermöglichen.

#### **1.4 Mögliche Beiträge der Gesundheitsökonomie**

Der Beitrag der Gesundheitsökonomie für die Entscheidungshilfe lässt sich an den vorab genannten Perspektiven zeigen.

Wie sieht die Perspektive für die **Krankenkassen** aus?

Für die Krankenkassen können gesundheitsökonomische Evaluationen wichtige Hinweise liefern. Studien, die zeigen, dass ein Behandlungsergebnis auf zwei Wegen mit unterschiedlichen Ausgaben erreicht wird, sollten bei Krankenkassen immer auf Interesse stoßen. Dies ist insbesondere dann

relevant, wenn die finanzielle Allokation verschiedener Sektoren wie Krankenhaus und niedergelassener Arzt betroffen sind. Ohne gesundheitsökonomische Studien ist die optimale Zuweisung der finanziellen Mittel an die entsprechenden Stellen nicht möglich.

Wie sieht die Perspektive der **niedergelassenen Ärzte** aus?

Das individuelle Einkommen der Ärzte wird durch den Vorwurf unwirtschaftlichen Verhaltens bedroht. Dabei konzentriert sich die Überprüfung der ärztlichen Wirtschaftlichkeit auf den Einsatzfaktor „Arzneimittel“. Überschreitet ein Arzt Grenzwerte, dann läuft er Gefahr, dass sich sein Einkommen um diesen Betrag vermindert. Es sei denn, er kann wirtschaftliches Handeln nachweisen. Hier können gesundheitsökonomische Studien aufzeigen, dass zwar der Ressourcenverbrauch und damit die Kosten im Bereich des niedergelassenen Arztes erhöht sind, dafür aber Kosten im Bereich des Krankenhauses eingespart werden.

Wie sieht die Perspektive des **Krankenhauses** aus?

Durch die Einführung der G-DRG (German Diagnosis Related Groups) zur Finanzierung der Krankenhäuser ist das Interesse an gesundheitsökonomischen Studien gewachsen. Ein Krankenhaus wird gemäß G-DRG für die Behandlung (Diagnose z.B. Blinddarmoperation) entlohnt. Dabei spielt es für die Krankenkasse keine Rolle wie lange der Patient im Krankenhaus liegt und welche Ressourcen er verbraucht. Für das Krankenhaus ist es deshalb lebensnotwendig zu wissen, ob das Entgelt die entstandenen Kosten deckt oder ob ein Verlust bzw. Gewinn pro Patient erwirtschaftet wurde.

Wie sieht die **gesellschaftliche Perspektive** aus?

Die gesellschaftliche Perspektive bezieht sich auf den Produktionsverlust und schließt damit den Arbeitsausfall für Arbeitgeber und die Rentenversicherung ein.

---

<sup>1</sup> Fußnote: Die Ergebnisse von Kostennutzwert-Analysen können zu einer Liste zusammengefasst werden. Man spricht dann von einer "League Table" (Schöffski, S. 376f, 2000)

Der Patient wird als gesonderte Perspektive gesehen. Der Einsatz von Atypika bei der Behandlung der Schizophrenie verringert die Frühverrentung um ca. 20% (Eichmann, 2005). Die vermeidbaren Rentenzahlungen und die entgangenen Beiträge zur Rentenversicherung interessieren jedoch die Rentenversicherer und die Beitragszahler. Insofern können hier die Rentenversicherer ein eigenes, aber auch ein gesellschaftliches Interesse artikulieren.

Neben Therapieeffekten auf die Lebenszeit ergeben sich auch Auswirkungen auf die Unterbrechung der Lebensarbeitszeit durch vorübergehende Arbeitsunfähigkeit. Bei in der gesetzlichen Krankenversicherung versicherten Patienten übernimmt der Arbeitgeber in der Regel die Lohnfortzahlung für die ersten sechs Wochen. Man kann nun seitens der Arbeitgeber ein Interesse an einer effizienten Gesundheitsversorgung unterstellen, weil eine Wechselwirkung zwischen Therapie und Dauer der vorübergehenden Arbeitslosigkeit besteht.

Wie sieht die Perspektive des **Patienten** aus?

Letztendlich gilt es noch die Perspektive des Patienten zu analysieren. Der suboptimale Einsatz knapper Ressourcen verursacht eine verkürzte Lebensdauer und/oder eine verminderte Lebensqualität. Dies hat neben der persönlichen Tragik auch gesellschaftliche Relevanz, da ein im Arbeitsleben stehender Patient durch vorzeitiges Ableben auch indirekte Kosten verursacht. Der sorgsame Umgang mit knappen Ressourcen sollte deshalb nicht nur ein ökonomisches, sondern auch ein ethisches Gebot sein.

## **1.5 Anwendungshindernisse für gesundheitsökonomische Beiträge**

### 1.5.1 Gesundheitspolitische Rahmenbedingungen

Das deutsche Gesundheitswesen ist sektoral aufgebaut. Einen Sektor bilden die niedergelassenen Ärzte, die mit den Kassenärztlichen Vereinigungen und diese wiederum kumuliert für alle Ärzte mit den Krankenkassen abrechnen. Aus diesem Grunde gehen patientenbezogene Daten, die eine medizinische und/oder ökonomische Vorteilhaftigkeit vorweisen können, verloren.

Ähnlich verhält es sich mit dem Krankenhaus. Die Krankenkassen verhandeln mit den Krankenhäusern die G-DRGs. Die separaten Verhandlungen mit den Kassenärztlichen Vereinigungen und mit den Krankenhäusern verhindern den Datenfluss eines Patienten über die Sektoren und damit die gesundheitsökonomische Evaluation einer Therapie.

Erschwerend kommt hinzu, dass der Planungshorizont der Krankenkassen sich auf ein Jahr beschränkt. Bei chronischen Krankheiten ist es deshalb nicht möglich, Therapien gesundheitsökonomisch zu vergleichen.

Die Leistungen bzw. Ergebnisse einer Therapie beschränken sich jedoch nicht nur auf die Sektoren der Kassenärztlichen Vereinigungen und der Krankenhäuser, sondern schließen auch die Rentenversicherung mit ein.

Die Finanzierungsstruktur der sozialen Sicherung in der Bundesrepublik Deutschland führt zu einer Gefangenendilemma-Situation (Schöffski, S. 479, 2000; Varian, S. 609ff, 2001). Die Krankenhäuser, die niedergelassenen Ärzte und die Krankenversicherer selbst können die Konsequenzen minderwertiger Therapien, sofern sie zu dauerhafter Arbeitsunfähigkeit und Frühverrentung führen, externalisieren. Ein Interesse an einer Therapie, die zwar in einem Sektor höhere Kosten verursacht, langfristig aber die Kosten senkt, besteht bei den Krankenversicherern zumindest nicht. Diese Anreizinkompatibilität bewirkt, dass sich zwar jeder Sektor für sich genommen rational verhält, das Gesamtsystem der sozialen Sicherung aber irrationales Verhalten aufweist. Das Verhalten der einzelnen Sektoren kann durch die Theorie der „Bounded Rationality“ von Herbert Simon erklärt werden. Danach wird ein Entscheidungsprozess abgebrochen, wenn ein befriedigendes Anspruchsniveau erreicht wird (Bendor, 2003). Ein Arzt oder ein Krankenhaus, das seine Budgets ausschöpft ohne sie zu überschreiten, hat ein entsprechendes Anspruchsniveau befriedigt und sieht keinen Anreiz, transsektorale Einsparmöglichkeiten zu verfolgen. Die Rationalität ist in diesem Falle auf den jeweiligen Sektor gebunden (bounded).

*Hinweis: Bei dem Gefangenen-Dilemma werden zwei Gefangene, die nicht miteinander kommunizieren dürfen, beschuldigt, gemeinschaftlich ein Verbrechen begangen zu haben. Wenn beide gestehen, werden sie zu je drei Monaten Haft verurteilt. Schweigen beide, werden sie nicht verurteilt. Gesteht nur ein Gefangener, wird der gestehende frei gelassen und der andere Gefangene zu sechs Monaten Haft verurteilt. Das Dilemma für die Gefangenen besteht darin, dass sie nicht wissen, welche Entscheidung ihr Gegenpart trifft. Folglich*

*weiß er nicht, wie er seinen Gewinn maximieren kann. In genau dieser Situation befinden sich die Entscheider im Gesundheitswesen. Für Krankenkassen, KVen und Krankenhäuser wäre Gewinnmaximierung nur möglich, wenn sie durch Kommunikation kooperieren könnten. Genau das können sie aber nicht, weil ihnen dazu Informationen aufgrund des sektoralen Gesundheitssystems fehlen (Varian, S. 132, 2000).*

Erschwerend kommt hinzu, dass die strikte Trennung von Informationen zwischen den Sektoren zur asymmetrischen Informationsvergabe führt. Mögliche Mechanismen wie Marktgeschehen oder Wettbewerb, die eine Allokation knapper Ressourcen verbessern könnten, werden so verhindert (Varian, S. 628, 2001). Eine Weiterentwicklung dieses Ansatzes, der sich m.E. auf das deutsche Gesundheitswesen übertragen lässt, wurde von Dörner entwickelt (Dörner, 1983).

Empirische Untersuchungen haben gezeigt, dass Menschen sich immer dann mit Problemlösungen schwer tun, wenn die Problemsituation durch Komplexität, Intransparenz, Vernetztheit und Dynamik gekennzeichnet ist. Diese Eigenschaften treffen auf das deutsche Gesundheitssystem mit seinen Sektoren, Substitutionseffekten, demografischen Schüben und dem starken Innovationsdruck zu.

Komplexität bedeutet, dass die zur Debatte stehende Situation, mit der man handelnd umgehen muss, zum einen aus einer Vielzahl von Elementen besteht, und dass zwischen diesen Elementen viele unterschiedliche Beziehungen herrschen. Die Komplexität besteht nicht nur aus einer meist unüberschaubaren Anzahl von Elementen, sondern erhöht sich durch die Beziehungen zwischen diesen Elementen. Damit impliziert die Komplexität noch zwei weitere Eigenschaften: Intransparenz und Vernetztheit.

Intransparenz bedeutet, dass auf den ersten Blick nicht alles sichtbar ist, was man sehen will. Dies bezieht sich sowohl auf die einzelnen Variablen des Systems als auch auf die Beziehungen zwischen den Variablen. Für den Entscheider bedeutet dies, dass er Entscheidungen hinsichtlich eines Systems, wie Festbeträge oder DRGs, zu fällen hat, dessen augenblickliche Merkmale er nur zum Teil, unklar oder gar nicht wahrnimmt.

Unter Vernetztheit versteht man das Ausmaß und die Art der Beziehungen, die zwischen den einzelnen Elementen bestehen. Ist ein Bereich stark vernetzt, so verändert sich automatisch sehr viel, wenn man nur an einem Ele-

ment verändernd eingreift. Senkt man z.B. die Preise durch Festbeträge, dann löst man automatisch Substitutionseffekte über die Menge und Alternativtherapien aus. Ob dann noch die erwünschten Effekte erzielt werden, ist zumindest fraglich.

Ein weiteres Merkmal ist die Dynamik, also die Tatsache, dass sich ohne eigenes Zutun das System ändert. Insbesondere Innovation und Demografie erzeugen im deutschen Gesundheitssystem Dynamik.

Interessant ist die Tatsache, dass mannigfaltige Problemlösefehler in intransparenten, komplexen, vernetzten und dynamischen Systemen wissenschaftlich untersucht wurden. Diese Erkenntnisse erlauben es, Ansatzpunkte zur Beseitigung dieser Fehler zu verfolgen.

Generell wird beim Problemlösen zweierlei benötigt. Das Wissen als solches (Epistemik) und Heuristiken. Unter Heuristiken versteht man die Wege zur Findung neuer Problemlösungen. Diese Wege beschreiben zwar die Art und Weise, sind aber inhaltlich selbst nicht festgelegt. Bei den Heuristiken unterscheidet man zwischen Suchraumeinengungs- und Suchraumerweiterungsheuristiken (Dörner, 1987). Bei einem Suchraumeinengungsheurismus wird ein definierter Suchraum auf Lösungen untersucht. Ein Beispiel ist das Schachspiel, das hoch komplex, aber begrenzt ist. Hier ist die Zwischenzielbildung ein passender Heurismus und zwar ein Suchraumeinengungsheurismus. Andere erfolgversprechende Heuristiken wären Zielbildung oder Modellbildung.

Bei neuartigen Problemen – zum Beispiel wie geht man mit Innovationen um – sind Suchraumerweiterungsheuristiken erfolgreich, wie Analogien oder ein systematischer Versuch-Irrtum-Ansatz.

Entscheidend ist es nun, für Probleme die adäquaten Heuristiken zu finden. Gerade dies ist aber in der Vergangenheit nicht gelungen. Das deutsche Gesundheitswesen ist durch das ständige Bemühen gekennzeichnet, finanzielle Zielsetzungen einzuhalten. Dieses Unterfangen ist bis heute nicht von Erfolg gekrönt. Seit Jahren folgt ein Kostendämpfungsversuch dem anderen. Festbeträge wurden eingeführt, Budgets, Richtgrößen, Arzneimittelrichtli-

nien oder Jumbogruppen für Festbeträge vorgegeben. Die Liste ließe sich fortführen.

Untersucht werden sollte die Frage, wie solche Denkfehler entstehen konnten.

Ein bei der Bewältigung von Komplexität häufig gemachter Fehler ist die Vernachlässigung von Zielen und Zwischenzielen bzw. eines logischen Zielsystems. So wird z.B. durch Regresse der Einsatz innovativer Medikamente eingeschränkt. Das Ziel der Budgeteinhaltung wird zwar damit erreicht, aber durch zusätzliche Kosten im Krankenhaus überkompensiert. Die Definition eines stringenten Zielsystems ist anstrengend. Der menschliche Denkapparat verfügt leider über zwei Eigenschaften, die ihn verleiten, anstrengende Denkprozesse zu meiden. Es handelt sich um die beschränkte mentale Kapazität und um das Anspruchsniveau. Man ist zufrieden mit dem, was man bisher erreicht hat und verfolgt diese Strategie weiter, zumal sich ja keine Katastrophen ereignet haben, das heißt, man „wurstelt sich durch“. Problemlöser „kapseln sich ein“, verschanzen sich dogmatisch hinter Paragraphen und man kommt der Problemlösung nicht näher.

Das begrenzte Anspruchsniveau ist sicherlich auch für die Akzeptanz der Intransparenz verantwortlich, wobei aber auch hier die mentalen Kapazitätsgrenzen eine Rolle spielen. So kapselt man sich ein, sucht seine Aufgaben nicht mehr nach der Wichtigkeit, sondern nach der Bewältigbarkeit. Man „wurstelt weiter“ z.B. bei den Festbeträgen, anstatt sich um transsektorale Probleme zu kümmern.

Besonders schwer fällt das Lösen von Problemen in dynamischen Systemen. In zahlreichen Untersuchungen wurde festgestellt, dass Vorhersagen meist linear vorgenommen werden. Bei nicht-linearen Verläufen kommt es dann automatisch zu fehlerhaften Prognosen. Dieses Fehlverhalten wird besonders dann problematisch, wenn positive Rückkopplungen die Entwicklung weiter beschleunigen. Gesellt sich zu dieser Konstellation auch noch Zeitdruck, dann versagen Problemlösungsversuche oft.

Aber nicht nur Loops (Rückkopplungsschleifen), auch Vernetztheit, erzeugen Denkfehler. Verwirrenderweise wirkt eine bessere medikamentöse The-

rapie beispielsweise erhöhend auf die direkten Kosten, verringert aber z.B. die Krankenhausliegedauer, was wiederum die direkten Kosten senkt, verbessert gleichzeitig die Wirkung der Psychotherapie, was aber dagegen die direkten Kosten erneut erhöht, senkt allerdings dadurch aber die Rückfallrate, was die direkten und indirekten Kosten senkt etc. Hier wären das Setzen von Zwischenzielen und/oder der Einsatz eines stringenten Zielsystems hilfreich. Beides ist oft nicht vorhanden.

Beim Problemlösen ist auch die Rolle von Emotionen nicht zu vernachlässigen. Thea Stäudel konnte in ihren Arbeiten aufzeigen, dass bei Frustrationen Problemlöser die Tendenz haben, sich eher zurückzuziehen und bei Erfolgen eher oberflächlich zu werden. Beides ist dem Problemlöseprozess abträglich (Stäudel, 1987). Schlussfolgernd lässt sich festhalten, dass Menschen und damit Entscheider im Gesundheitswesen Schwierigkeiten beim Problemlösen haben, wenn es sich um komplexe, intransparente, vernetzte und dynamische Systeme handelt, was auf das deutsche Gesundheitswesen zutrifft.

Die Fehler beim Problemlöseprozess sind auf bestimmte Denkfehler, wie die beschränkte Rationalität aufgrund des Anspruchsniveaus, des linearen Denkens, der beschränkten Kapazität oder des methodischen Verhaltens („Durchwursteln“, Einkapseln) zurückzuführen.

Auch Emotionen beeinflussen den Problemlöseprozess negativ.

All diese Faktoren wirken sich nachträglich auf die Umsetzung der Gesundheitsökonomie im deutschen Gesundheitswesen aus. Ansätze zur Überwindung dieser Probleme werden im Schlussteil aufgezeigt.

### 1.5.2 Methodische Probleme

Grundlage einer jeden wissenschaftlichen Analyse ist das Vorliegen eines Zieles (Dörner, S. 74ff., 1989). Die gesetzliche Grundlage zur Bewertung des Nutzens von Arzneimitteln findet sich im Sozialgesetzbuch (SGB V § 35b) (Porzsolt – in Review). Im deutschen Arzneimittelgesetz (AMG) wird der Begriff des Nutzens aber nicht verwendet. Es stellt sich damit die Frage, wie etwas evaluiert werden soll, wenn es nicht definiert ist. Dieser zentrale Punkt soll im Weiteren genauer untersucht werden.

Das Nutzendenken nimmt ausgehend vom Utilitarismus von J.S. Mill und J. Bentham nicht nur in der Sozialethik und in den Wirtschaftswissenschaften der Gegenwart eine zentrale Stellung ein, sondern durchdringt auch andere wissenschaftliche Disziplinen. Es basiert auf dem Gedanken, dass konsequentes und freies Streben der Gesellschaftsmitglieder nach Nutzenmaximierung nicht nur auf individueller Ebene, sondern auch für die Gesellschaft als Ganzes maximale Wohlfahrt garantiert. Deswegen soll man diesem Nutzenstreben freien Lauf lassen und die steuernden Aktivitäten des Staates auf ein notwendiges Minimum vermindern. Dieser Gedanke hat nicht zuletzt durch die fortschreitende Etablierung neoliberaler Wirtschaftskonzepte in den letzten beiden Dekaden an Bedeutung gewonnen (Krätke, 1999).

#### 1.5.2.1 Probleme der Definition des Nutzens

##### *1.5.2.1.1 Psychologische Dimensionen des Nutzens*

In der Psychologie wird Nutzen, der als subjektives Wohlbefinden definiert wird, in verschiedene Dimensionen aufgegliedert. Dabei werden angenehme Affekte, unangenehme Affekte und kognitive Lebenszufriedenheit unterschieden (Diener und Eunkook, 1999). Die diesem Optimierungsprozess zu Grunde liegenden Nutzenstrukturen unterliegen Veränderungen im Zeitablauf. Die Psychologie untersucht intensiv Einfluss und Wirkungsweise individueller Persönlichkeitsmerkmale sowie soziodemografische und kontextabhängige Faktoren. Durch Beiträge der Psychometrie wurden in letzter Zeit große Fortschritte im Hinblick auf die Messbarkeit von Nutzen und Zufriedenheit erzielt.

Es zeigt sich, dass der Begriff „Nutzen“ in der Psychologie wesentlich weiter gefasst ist als der eng auf Eigennutz ausgerichtete Nutzenbegriff des Utilitarismus. Nutzen im Sinne von Glück und Wohlbefinden hat eine komplexe Struktur, die weit über die Persönlichkeitsgrenze des Individuums hinaus in die soziale und materielle Umwelt hineinreicht. Die utilitaristische Reduktion menschlichen Handelns auf rein eigennützige Motive kann deshalb nur unter der Annahme der modelltheoretischen Perspektive der Wirt-

schaftswissenschaften akzeptiert werden, in welchen das zu großen Teilen irrationale Universum menschlichen Handelns auf den streng rationalen Homo Oeconomicus begrenzt wird, um zu gültigen Aussagen mit akzeptablem Gültigkeitsbereich zu gelangen.

#### *1.5.2.1.2 Nutzen in den Wirtschaftswissenschaften*

In Bezug auf klassische Konsumgüter wird jeder zustimmen, dass für die Beurteilung von deren Qualität der Nutzen von zentraler Bedeutung ist, weil kein rational handelndes Individuum Konsumgüter nachfragen würde, die nutzlos sind. Aus dem aggregierten Nutzenkalkül der Individuen resultiert die gesamtwirtschaftliche Nachfragestruktur. Diese Nachfrage ermöglicht es, den Produzenten auf den Märkten die entsprechende Güterstruktur anzubieten. Kein Produzent wird Güter produzieren, die niemand nachfragt.

Es gibt aber auch Güter, so genannte meritorische Güter, die zwar nachgefragt werden, aber aus denen sich aus verschiedenen Gründen noch keinen spezifischen Nutzen ableiten lassen. Diese Situation ist z.B. bei der Schulbildung und im Gesundheitssystem gegeben, wo traditionellerweise der Staat die Produktion steuert. Bei der Produktion meritorischer Güter gewinnt die Evaluation an Bedeutung, wenn sie es ermöglicht, die Eignung derartiger Güter zur Nutzengenerierung auf Nachfrageseite zu ermitteln und so einen effizienten Einsatz der gemeinschaftlichen Mittel zu gewährleisten.

#### *1.5.2.1.3 Probleme der Definition des Nutzens*

In der ökonomischen Theorie versteht man unter dem Nutzen das Maß für die Fähigkeit eines Gutes, das mehrere Eigenschaften (Nutzwerte) aufweisen kann, die Bedürfnisse eines wirtschaftlichen Akteurs (z.B. eines Konsumenten) zu befriedigen. Als ein Bedürfnis gilt das Streben, Unbehagen durch eine bestimmte Handlung zu beseitigen. Dieser Akt der Bedürfnisbeseitigung generiert Nutzen.

In einer übergreifenden Sichtweise wird Nutzen als Beitrag zur Bedürfnisbefriedigung durch Erreichung von Zielen von Individuen oder Gruppen definiert. Ziele müssen nicht zwingend vollständig realisierbar sein, auch

teilweise erreichte Ziele tragen zur Bedürfnisbefriedigung bei. Gleiche Ziele oder Zielerreichungsgrade haben für unterschiedliche Individuen oder Gruppen unterschiedlichen Nutzen. Ein Ziel ist hierbei die Beschreibung eines zukünftigen erstrebenswerten Ereignisses oder Zustandes, wie Gesundheit oder Schmerzfreiheit, der durch menschliches Handeln herbeigeführt werden soll.

#### *1.5.2.1.4 Probleme der Messung des Nutzens*

Die Klassiker der ökonomischen Theorie verwendeten einen objektiven, kardinalen Nutzenbegriff, der sich am anerkannten, meist eindimensionalen Wert eines Gutes orientiert, beispielsweise dem physikalischen Heizwert von Kohle oder der zugelassenen Höchstgeschwindigkeit eines Autos. Dieser Nutzen wurde als in objektiv anerkannten Einheiten messbar angenommen.

Die neoklassische Nutzentheorie geht demgegenüber von einem ordinalen, subjektiven Nutzenkonzept aus, in dem eine absolute Messbarkeit des Nutzens abgelehnt wird und vielmehr auf individuelle Präferenzhierarchien rekurriert wird. Das ordinale Nutzenkonzept ist hinreichend, um die Konsumententscheidungen zwischen Gütern zu erklären. Die jeweiligen Präferenzstufen der Güter werden empirisch aus der tatsächlich getroffenen Auswahl der Wirtschaftssubjekte abgeleitet und werden so zum Maß für den Nutzen ohne dessen weitere Differenzierung (Kruse, 1959).

Jüngere Tendenzen in der Ökonomie unterstellen hingegen wieder generell eine Messbarkeit des Nutzens und verwenden einen psychologischen Nutzenbegriff, der eher am individuellen Zustand der „Zufriedenheit“ oder des „Wohlbefindens“, denn an „Konsumpräferenzen“ ausgerichtet ist. Ein derart erweitertes Konzept wurde in der Ökonomie notwendig, weil bestimmte ökonomisch relevante Handlungsweisen mit dem neoklassischen Präferenzbegriff nicht dargestellt werden konnten. Soziales Verhalten beispielsweise lässt sich aus einer rein eigennützigem Präferenzstruktur, wie von den Neoklassikern unterstellt, kaum erklären. Überdies wurde in den Forschungsgebieten der Ökonomie und Psychologie in den letzten Jahren empirisch

fundiertes Wissen erarbeitet, das darauf hinweist, dass die Individuen durch rein eigennützig gesteuertes Handeln nicht zwingend auch ihr höchstes Wohlbefinden erzielen (Varian, S. 787ff, 2001).

Diese historische Veränderung des Nutzenbegriffs erforderte auch eine Änderung der Messmethode. Die ordinale Nutzenerfassung wurde durch die individuelle, psychologische Messung des Nutzens, die vor, nach oder während des Konsumvorganges erfolgen kann, substituiert. Mit Hilfe des Instruments der Befragung wird es möglich, Hinweise auf die Beurteilung des Wohlbefindens und der Lebenszufriedenheit von Individuen zu erhalten. Hinter der oftmals in Punkten auf einer Skala geäußerten Einschätzung steht ein kognitiver Prozess, bei dem Individuen ihre Lebenszufriedenheit insbesondere anhand der Lebensumstände und Aspirationen und im Vergleich zu anderen Personen, vergangenen Erfahrungen und Erwartungen für die Zukunft beurteilen und somit messbar machen (Stutzer, 1995). Die jeweils gewählte Skala, aus ca. 800 möglichen, ist eine theoretische Konvention zur wissenschaftlichen Beschreibung des Wohlbefindens. Mit diesen Instrumenten soll die individuell empfundene Lebensqualität beschrieben werden. Die Problematik dieser Aussage ist evident, da jedes Individuum den empfundenen Nutzen für sich nicht in diesen Skalen, sondern in jeweils anderen Entitäten misst. Es scheint m. E. sinnvoll davon auszugehen, dass Patienten in ihrer Therapiewahl analog ihr Nutzenmaximum suchen.

#### *1.5.2.1.5 Nutzen und Grenznutzen*

Die Haushalte der mikroökonomischen Konsumtheorie suchen in ihrer Konsumnachfrage das Nutzenmaximum. Sie bewerten die Qualität der Güter am Ausmaß des Nutzens. Bei den Klassikern um Adam Smith wurde noch zwischen Tausch- und Nutzwert unterschieden. Der Nutzwert entsprach dem objektiven Gebrauchswert des Gutes als dessen anerkannte Tauglichkeit zur Bedürfnisbefriedigung (z.B. dem physikalischen Heizwert der Kohle), während der Tauschwert mit dem Marktpreis des Gutes gleichgesetzt wurde und entweder durch dessen Nachfrage oder dessen Produktionskosten bestimmt war. Beide Wertgrößen konnten unterschiedliche Ausprägungen haben, wie

das Beispiel von Diamant und Wasser verdeutlicht. Für die Klassiker hatten Diamanten keinen Nutzwert, jedoch einen sehr hohen Tauschwert, dagegen war Wasser mit einem sehr hohen Nutzwert versehen, jedoch mit einem sehr geringen Tauschwert. Dieses so genannte Wertparadoxon war für die Klassiker nicht lösbar.

Erst die Einführung der Grenznutzenschule und des subjektiven Nutzenbegriffes im Sinne einer individuellen Bewertung eines Gutes als Tausch- und Nutzobjekt durch die bewertende Person führten zur Identität von Nutz- und Tauschwert: Der Wert eines Gutes wird zwar weiter objektiv als dessen anerkannte Brauchbarkeit zur Herbeiführung eines gewollten Erfolges bemessen, doch kommt ein Marktpreis als Tauschwert für ein Gut nur zustande, wenn mindestens ein Wirtschaftssubjekt diesen Erfolg wertvoller beurteilt als die Tauschgüter, die es dafür abgeben muss. Diese individuelle Beurteilung vollzieht sich auf Grundlage des empfangenen Grenznutzens. Dieser wird definiert als Veränderung des Gesamtnutzens, die eintritt, wenn bei gegebener Güterkombination der Konsum eines Gutes um eine minimal kleine Einheit erhöht wird. Dieses Kalkül unterliegt den beiden Gossenschen Gesetzen. Das erste Gesetz besagt, dass der Grenznutzen eines Gutes mit steigendem Verbrauch abnimmt (Gesetz der Bedürfnissättigung).

Das zweite Gossensche Gesetz besagt, dass der Grenznutzen in seinem Optimum für alle Verwendungsrichtungen der verfügbaren Güter gleich ist („Genussausgleichs-Gesetz“). Wenn im 2-Güter-Fall beispielsweise ein Gut mehr Nutzen bringt als das andere, wird das Wirtschaftssubjekt von diesem so lange mehr einsetzen (zu Lasten des anderen), bis der Grenznutzen der beiden Güter gleich ist.

Damit war das Wertparadoxon aufgelöst, denn es war nun darstellbar, dass Güter wie Wasser einen niedrigen Wert haben, weil sie reichlich vorhanden sind. Sein Gesamtnutzen ist zwar groß, doch der Genuszuwachs durch die letzte Einheit ist gering. Diamanten sind weniger wichtig als Wasser, ihr Gesamtnutzen ist relativ gering, doch erlangen sie durch ihre Seltenheit in der Regel einen hohen Grenznutzen, so dass ihr Tauschwert wegen der Seltenheit groß sein kann. Die Haushalte verhalten sich rational in dem Sinne,

dass der Gesamtnutzen ihrer Güter unter der Budgetbeschränkung des erzielten Einkommens so groß wie möglich wird.

Eine solche Grenznutzenanalyse kann aber nur bedingt angewendet werden, denn sie geht von kardinal messbaren Nutzenfunktionen aus, die in der Realität schwer durchführbar sind. Deshalb wird in der klassischen Mikroökonomie in der Regel auf die so genannte Indifferenzkurvenanalyse zurückgegriffen, die eine Nutzenanalyse mit Hilfe des ordinalen Messkonzepts (Ordnung: größer, gleich, kleiner) möglich macht.

Zusammenfassend ist das Nutzenverständnis in Psychologie, Wirtschaftswissenschaften und damit auch in der Medizin von den jeweiligen originären Problemstellungen dieser Disziplinen geprägt und dadurch keineswegs einheitlich. Die Wirtschaftswissenschaften reduzieren das Wirtschaftssubjekt, vereinfacht ausgedrückt, auf den Homo Oeconomicus, den allein auf Eigennutz bedachten, streng nach rationalen Kriterien handelnden Akteur. Der Homo Oeconomicus ist bestrebt, den Grenznutzen seiner verschiedenen Konsummöglichkeiten auszugleichen und so seinen Gesamtnutzen zu maximieren. Das Nutzenstreben aller Gesellschaftsmitglieder führt zur Konsumnachfrage, die über den Marktpreismechanismus die Angebotsstruktur bestimmt und so das so genannte Gütermarktgleichgewicht bewirkt.

Das Nutzenverständnis der Psychologie mit den Zielen „Glück“ und „Wohlbefinden“ kommt den Vorstellungen kranker Menschen näher als das wirtschaftswissenschaftliche Konzept. Eine Bewertung wird aber dadurch nicht erleichtert.

In der Medizin ist das Nutzenverständnis nicht standardisiert, was durch das Fehlen einer Theorie der Medizin zu erklären ist (Th. v. Uexküll, 2003). Weil sich die Medizin ohne theoretische Basis als Erfahrungswissenschaft entwickelt hat (die biomolekularen Grundlagen ersetzen keine Theorie der Medizin), fehlen einheitliche Definitionen von Gesundheit, Krankheit, Lebensbeginn und auch des Nutzens von Gesundheitsleistungen.

Es kann also festgestellt werden, dass es bisher keine einheitlich akzeptierte Definition des Nutzens von Gesundheitsleistungen gibt.

Bezieht man sich auf die oben genannten Definitionen, dann resultiert der Nutzen eines Medikamentes aus der Summe der Teilnutzen (unabhängig davon, wie sie gemessen werden) bzw. den durch ein Medikament erreichten Teilnutzen.

Daraus ergeben sich nun zwei Probleme. Zum einen stellt sich die Frage mit welcher Methode Nutzen bringende Endpunkte identifiziert und analysiert werden sollten. Diese Frage ist nicht trivial. So steht zur Entscheidung, ob ein Überleben von 10 Jahren mit Schmerzen einem Überleben von 8 Jahren ohne Schmerzen vorzuziehen sei.

Zum anderen wird die Nutzenbewertung im Gesetz vorgeschrieben. Wie bereits oben beschrieben, ist das Institut für Qualität und Wirtschaftlichkeit (IQWiG) durch das SGB V beauftragt, eine Nutzenbewertung durchzuführen, wobei der Terminus „Nutzen“ im Arzneimittelgesetz (AMG) nicht definiert ist.

Eine Analyse der Nutzenbewertung in 11 Ländern ergab, dass das QALY-Konzept eine in mehreren Ländern eingesetzte Methode der Nutzenbewertung ist (Zentner, 2005).

Das IQWiG zieht zur Analyse der Nutzenbewertung „patientenrelevante Endpunkte“ heran, ohne jedoch zu definieren, was Patientenrelevanz ausmacht.

Beide Ansätze, das QALY-Konzept und die Patientenrelevanz werden im Weiteren kritisch untersucht.

#### 1.5.2.2 Analyse des QALY-Konzept

Gemäß QALY-Konzept besteht das ultimative Ziel einer Therapie darin, dass ein Patient länger und besser lebt. Die Evaluation von Therapien kann durch dieses Konzept erfolgen (Schöffski, S. 367ff., 2000).

Das QALY-Konzept weist allerdings Schwachstellen auf, und zwar:

- es existieren mehrere Instrumente, die unterschiedliche Ergebnisse produzieren

- es tauchen Schwierigkeiten bei der Berechnung der präferenzbasierten Lebensqualität auf
- die Validität der QALY-Instrumente ist nicht gewährleistet
- die theoretische Fundierung der QALY-Instrumente ist fraglich

Zur Messung der präferenzbasierten Lebensqualität stehen **mehrere Instrumente** zur Verfügung, und zwar der HUI Marks 2, 15D, der QWB, der EQ-5D und der SF-6D. Alle diese Instrumente wurden zur Messung der präferenzbasierten Lebensqualität erstellt. Konsequenter Weise müssten alle Instrumente hoch miteinander korrelieren. Dies trifft leider nicht zu. Damit ist die externe Validität dieser Instrumente in Frage gestellt (Böhmer, S. 53ff, 2000).

Zur Messung der **präferenzbasierten Lebensqualität** müssen die einzelnen Lebensqualitätszustände berechnet werden. Dies ist aufgrund der vielen Gesundheitszustände der Patienten schwierig. Der EQ-5D unterscheidet 243 Gesundheitszustände, der SF-6D 9.000, der HUI Marks 2 24.000 und der 15D mehr als  $10^9$ . Mit den herkömmlichen Methoden wie Standard Tramble und Time Trade Off lässt sich diese Komplexität nicht bewältigen.

Trotzdem scheinen diese Probleme meines Erachtens nicht unüberwindbar. Die Bewältigung großer Zahlen von Gesundheitszuständen kann durch reduzierte Designs gelöst werden (Hujer, 1996). Ansätze sind sowohl für die Berechnung präferenzbasierter Lebensqualität unter Einbeziehung mit Conjoint Analysen (Monzini, 2003) als auch mit Discrete Choice Analysen (Backhouse, 2003) zu finden.

Ein weiteres Problem betrifft die **Validität der Instrumente** zur Messung der Lebensqualität. Wie oben erwähnt (Chassany, 2002) werden zur Prüfung der Validität der Cronbach alpha und unter Umständen explorative Faktorenanalysen eingesetzt. Strukturgleichungsmodelle ermöglichen die Durchführung konfirmatorischer Faktorenanalysen, womit die inhaltliche Konstruktvalidität exakt berechnet werden kann. Offensichtlich wird dieses

Instrument nicht häufig eingesetzt. Eine Literatursuche in EMBASE im Juni 2005 mit den Suchbegriffen EQ-5D, EuroQol, Euro-Qol, konfirmatorische Faktorenanalyse und confirmatory factor analysis ergab keinen Treffer. Es ist meines Erachtens aber trotzdem davon auszugehen, dass die konfirmatorische Faktorenanalyse in Zukunft einen wertvollen Beitrag zur Berechnung der inhaltlichen Konstruktvalidität leisten wird (Bühner, 2004). Der Vollständigkeit halber sei erwähnt, dass die externe Validität erfasst wird (Schumacher, 2003).

Mit Methoden wie Strukturgleichungsmodelle, konfirmatorische Faktorenanalyse, reduzierte Designs, Conjoint Analysen und Discrete Choice Analysen ergeben sich unter Umständen Ansätze, die Konstruktion eines umfassenden Instrumentes zur Messung der präferenzbasierten Lebensqualität in Angriff zu nehmen.

Es sollte erwähnt werden, dass ein wesentlicher Bestandteil des QALY-Konzeptes, nämlich die Nutzenmessung, bezüglich ihrer **theoretischen Fundierung** umstritten ist. Zum einen wird Rationalität, die in der Regel nicht vorliegt, unterstellt (Happich, S. 62ff, 2003). Zum anderen spielen die Gewichtung von Wahrscheinlichkeiten, die subjektiv in eine Entscheidung einfließen und die Existenz eines Referenzpunktes eine nicht zu unterschätzende Rolle. Ein solcher Referenzpunkt kann die normale Lage, der gegenwärtige Zustand, das individuelle Anspruchsniveau oder die Erwartung sein.

Die Schwierigkeiten des QALY-Konzepts beziehen sich aber nicht nur auf die empirischen Aussagen. Dies sei an zwei Beispielen verdeutlicht. Wendet man das QALY-Konzept bei der Therapie von Frühchen ein, so hat dies zur Folge, dass man aufgrund der zu erwartenden langen Lebenserwartung finanzielle Ressourcen vorrangig in diese Therapie alloziieren müsste. Genau gegenteilig sieht es in der Geriatrie aus. Aufgrund der geringen zu erwartenden Lebensdauer dürfte in diesen Bereich eigentlich nach dem QALY-Konzept kein Geld mehr fließen.

Von dieser Problemstellung ausgehend ergeben sich mehrere Fragen. Zum einen ist zu klären, welche Zielsetzung, Rettung der Frühchen oder Versor-

gung geriatrischer Patienten, verfolgt werden soll und zum anderen, mit welchen Methoden diese Fragestellung bearbeitet werden kann. Ein möglicher Ansatz wird im Folgenden beschrieben.

#### 1.5.2.3 Analyse patientenrelevanter Endpunkte

In der medizinischen Forschung und Versorgung entsteht ein Problem, nämlich dann, wenn zu entscheiden ist, anhand welcher Parameter die Wirksamkeit (output) oder der Nutzen (outcome) medizinischer Maßnahmen gemessen werden soll (Ruß, 2005). Bei Patienten mit z.B. Diabetes Mellitus entspricht die Senkung eines erhöhten Blutzuckerwertes einem „output“. Als „output“ bezeichnet man demnach die Ergebnisse aller Gesundheitsleistungen. Wenn das Ergebnis einer Gesundheitsleistung als gewonnene Lebensjahre oder gewonnene Verbesserung der Lebensqualität beschrieben werden kann, bezeichnet man dieses Ergebnis als „outcome“.

So stellt sich nun die Frage, ob neue Arzneimittel anhand der Verlängerung des Lebens (outcome), anhand der Remissionsrate (output) oder anhand beiden Parametern (das IQWiG spricht hier von Endpunkten) beurteilt werden sollen. Diese Fragestellung zeigt, dass die Bewertung von Gesundheitsleistungen ohne ein theoretisches Konzept kaum möglich sein wird. Hier nimmt man die Hilfe der Wissenschaftstheorie in Anspruch, die sich mit den betreffenden Fragen auseinandergesetzt hat. Wenn man diese Vorarbeiten anderer Wissenschaftsbereiche nutzen möchten, dient es dem Verständnis, sich zumindest mit einigen Grundbegriffen dieser Wissenschaft auseinanderzusetzen.

Die so genannten Brückenprinzipien sind Hilfsmittel, um exakt diese Fragen, die man Sollensfragen nennt, zu beantworten (Albert, 1985). Diese Hilfsmittel ermöglichen, empirisches Wissen für die Beurteilung von Sollensvorschlägen fruchtbar zu machen.

Statt von „empirischem Wissen“ wird im Folgenden auch von „empirischen Erkenntnissen“ gesprochen. Wird derartiges Wissen in Sprache gefasst, entstehen „empirische Aussagen“. Für den Begriff „Erfahrungswissenschaft“ wird auch der Terminus „empirische Wissenschaft“ gebraucht.

Für ein tieferes Verständnis der Idee der Brückenprinzipien ist es nützlich, einige wissenschaftstheoretische und (ursprünglich) moralphilosophische Überlegungen anzustellen. Es hat sich als problematisch herausgestellt, ein allgemein akzeptiertes und funktionierendes Verfahren zu entwickeln, um begründete Antworten auf Sollensfragen zu finden. Ein wesentlicher Grundsatz der Wissenschaftstheorie besagt, dass empirisches Wissen zwar nicht dazu dienen kann, Normen zu begründen – dazu bedarf es weiterer Faktoren – wohl aber dazu, vorgeschlagene Normen zu kritisieren. Normen, die dieser Kritik standhalten, werden beibehalten. Normen, die scheitern, werden verworfen.

Ob empirisches Wissen z.B. in der Gesundheitsökonomie in dieser Weise nutzbar gemacht werden kann, ist vom Charakter der empirischen Aussagen und von normativen Sätzen abhängig. Unser Beispiel, „Der HbA<sub>1C</sub> ist bei Patienten mit Diabetes Mellitus häufig erhöht“, beschreibt eine empirische Aussage, die deskriptiv ist. Die Aussage, „Der HbA<sub>1C</sub> soll bei Patienten mit Diabetes Mellitus regelmäßig bestimmt werden“, entspricht einem normativen Satz. Die Aussage „Wenn ein Patient behandelt wird, dann sinkt sein HbA<sub>1C</sub>“ stellt eine explikative Aussage dar. Empirische Aussagen und Aussagensysteme beschreiben oder erklären reale Zusammenhänge. Sie sind „deskriptiv“ oder „explikativ“. Das erste der oben verwendeten Beispiele ist deskriptiv. Das Beispiel „Wenn der HbA<sub>1C</sub> erhöht ist, dann sollte eine regelmäßige Bestimmung dieses Wertes durchgeführt werden“, entspricht einer gemischten Aussage. „Wenn der HbA<sub>1C</sub> erhöht ist“ stellt eine beschreibende Aussage dar, „dann sollte eine regelmäßige Kontrolle durchgeführt werden“ ist eine normative Aussage.

Deskriptive oder explikative Beschreibungen oder Erklärungen können grundsätzlich „wahr“ sein. Empirische Aussagen und Aussagensysteme, also Beobachtungsaussagen, Hypothesen, Theorien oder Erklärungen, sind prinzipiell wahrheitsfähig.

Normative Aussagen hingegen sind nicht wahrheitsfähig. Sie formulieren Handlungsanweisungen, bezeichnen also Handlungen als geboten, verboten oder erlaubt. Sie sind nicht beschreibenden, sondern vorschreibenden Cha-

rakters, sie sind „präskriptiv“. Als Handlungsanweisungen beanspruchen Normen „gültig“ zu sein.

Der Unterschied zwischen empirischen und normativen Aussagen wird an folgendem Beispiel verdeutlicht: Die Aussage der empirischen Wissenschaft „Durch Reduktion des Übergewichts können erhöhte Blutzuckerwerte bei Patienten mit Typ II Diabetes gesenkt werden“ ist empirisch und wahrheitsfähig und beschreibt einen realen Funktionszusammenhang

Die Aussage „Diabetiker sollen das Übergewicht reduzieren“ entspricht einer Norm. Sie ist empfehlend oder vorschreibend, d.h. präskriptiv und damit nicht wahrheitsfähig. Sie ist gültig oder ungültig. Darin unterscheiden sich empirische von normativen Aussagen. Kurzum, empirisches Wissen sagt etwas darüber aus, wie die Dinge sind, nicht wie sie sein sollen.

Für das Vorhaben, empirisches Wissen für die Beurteilung von Normen fruchtbar zu machen, kommt erschwerend hinzu, dass aus ausschließlich beschreibenden Sätzen auch keine rein normativen Sätze logisch gültig ableitbar sind (dies ist eine Formulierung des sog. ‚Humeschen Gesetzes‘). Der neuzeitliche Philosoph David Hume war auf dieses Problem im Rahmen seiner moralphilosophischen Untersuchungen aufmerksam geworden (Hume, 1975). Zunächst einmal könnten beschreibende oder erklärende Aussagen eine logisch gültige Ableitung von Sollensvorschriften erlauben. Aus empirischem Wissen allein folgen aber keine relevanten Handlungsnormen („relevant“ wird hier zur näheren Bestimmung gebraucht, weil Sätze mit normativer Komponente ableitbar sind, die aber in normativer Hinsicht irrelevant sind, weil sie gerade nicht besagen, was getan werden soll.). Aus der erfahrungswissenschaftlichen Erkenntnis etwa, dass eine Reduktion des Übergewichts dazu beiträgt, den Blutzuckerspiegel zu normalisieren, folgt nicht zwangsläufig, was getan werden soll. Betrachten wir dazu die beiden folgenden Fälle:

Man könnte sich einen Typ-II-Diabetiker vorstellen, der nicht daran interessiert ist, seinen Blutzuckerspiegel abzusenken. Aus seinem Wissen, dass er den Blutzuckerspiegel senken „kann“, wenn er sein Gewicht reduziert, folgt

noch nicht die Entscheidung, sein Gewicht zu reduzieren, weil er das Ziel nicht anstrebt, das er durch die Gewichtsreduktion erreichen könnte.

Ein anderer Typ-II-Diabetiker hat sich als Ziel gesetzt, seinen Blutzuckerspiegel ohne Arzneimitteltherapie, also durch Gewichtsreduktion, zu senken. In diesem Fall setzt er aus dem Wissen um die möglichen Effekte einer Gewichtsabnahme und aus dem Ziel, das er sich gesetzt hat, den Entschluss zur Gewichtsreduktion um.

Es zeigt sich also, dass unterschiedliche Handlungen als „gesollt“ bewertet werden können, obwohl in beiden Fällen die empirische Erkenntnis dieselbe ist. Welche Handlung als die richtige betrachtet wird, hängt wesentlich von den jeweiligen „Zielsetzungen“ ab. Dies impliziert, dass die Zielsetzungen selbst den Charakter von Normen haben. Im ersten der beiden Diabetiker Fälle handelt es sich um eine Erlaubnis- oder Verbotsnorm („der Zuckerspiegel darf abgesenkt werden, muss es aber nicht“ bzw. „der Zuckerspiegel soll nicht abgesenkt werden“). Im zweiten Fall handelt es sich um ein Gebot („der Blutzuckerspiegel soll abgesenkt werden“). Das empirische Wissen ist selbst nicht normativ und es lässt auch keine relevanten normativen Schlussfolgerungen zu. Es müssen Normen vorausgesetzt werden, um normative Schlussfolgerungen ziehen zu können. Wer dies nicht beachtet, begeht einen so genannten „Sein-Sollen-Fehlschluss“.

Dennoch ist empirisches Wissen relevant für die Beantwortung von Sollensfragen wie z. B. ob outputs oder outcomes verfolgt werden sollen. Es gibt einen logisch korrekten Weg, empirisches Wissen für die kritische Untersuchung von Normvorschlägen zu nutzen. Ziel und Zweck ist es, jene Normen zurückzuweisen, die der Kritik nicht standhalten. Dazu muss eine Brücke zwischen deskriptiven und normativen Sätzen geschlagen werden. Normen haben in aller Regel eine ganz bestimmte Funktion: Sie dienen dazu die menschliche Praxis anzuleiten. Es existieren nun verschiedene empirisch feststellbare Umstände unter welchen Normen sich diese Funktionen nicht oder nicht in der gewünschten Weise erfüllen können. Eine tragfähige Brücke zwischen deskriptiven und normativen Sätzen liegt vor, wenn argumen-

tativ gezeigt werden kann, dass eine Norm angesichts der empirischen Erkenntnis nicht als Regulativ der menschlichen Praxis in Betracht kommt. Sätze, die in dieser Weise empirische Aussagen mit Normen verknüpfen, werden als "Brückenprinzipien" bezeichnet.

Brückenprinzipien müssen begründet werden, sonst wären sie unbegründet und damit beliebig. Die Begründungen können an die eben genannte regulative Funktion von Normen anknüpfen. Um Sein-Sollen-Fehlschlüsse zu vermeiden, sind also begründete Brückenprinzipien notwendig. Nehmen wir etwa an, es würde eine Norm vorgeschlagen, die etwas Menschenunmögliches fordert. Dagegen kann eingewendet werden, dass eine derartige Forderung die menschliche Praxis nicht anzuleiten vermag, da die geforderte Handlungsweise nicht vollzogen werden kann. Entsprechend ist die betreffende Norm zurückzuweisen, sie ist damit ungültig. Was möglich ist und was nicht, ist aber eine empirische Frage. Somit können bestimmte empirische Erkenntnisse über die menschliche Natur zur Kritik von Normen herangezogen werden. Aber die kritische Zurückweisung kann nicht unmittelbar dem empirischen Wissen entnommen werden. Dieses ist deskriptiv und besagt nur, dass eine bestimmte Handlungsweise nicht vollzogen werden kann. Für die Normenkritik ist eine eigene Prämisse nötig, ein Brückenprinzip, welches durch die Einsicht in die Funktionsunfähigkeit der Norm begründet ist und eine Verbindung zwischen empirischem Wissen und der zu prüfenden Norm herstellt. Eine in einer bestimmten Norm geforderte Handlungsweise, die laut unserem empirischen Wissen nicht vollzogen werden kann, muss auch nicht vollzogen werden. Damit ist diese Norm hinfällig.

Diese Art der Normenkritik kann nun an verschiedenen, empirisch beurteilbaren Anforderungen an Normen ansetzen. Entsprechend resultieren daraus verschiedene Brückenprinzipien. Es ist nützlich, diese in einer ausführlichen konditionalen Weise anzugeben, weil dann die Art und Weise der Kritik klar zum Ausdruck kommt. Es geht also jeweils um die Beurteilung einer bestimmten Norm mittels empirischen Wissens. Sollensvorschriften, die derartigen Kritiken nicht standhalten, werden verworfen.

Ein Brückenprinzip wurde soeben genannt: das „Sollen-Können-Prinzip“, das auch als „Praktikabilitätsprinzip“ oder „Realisierbarkeitspostulat“ bezeichnet wird. Es kann nun wie folgt formuliert werden:

Wenn laut empirischem Wissen das nach der Norm gesollte Handeln nicht realisierbar ist, dann kommt diese Norm in empirischer Hinsicht nicht in Betracht.

Das oben genannte Beispiel erklärt, dass die Behandlung des Diabetes Mellitus gerechtfertigt werden kann, weil die erfolgreiche Reduktion schwerer Folgen des Diabetes nachgewiesen wird.

Die Formulierung, eine Norm komme „in empirischer Hinsicht nicht in Betracht“ weist darauf hin, dass Normen noch weitere Kriterien erfüllen müssen, um akzeptabel zu sein, etwa das Kriterium innerer logischer Widerspruchsfreiheit. Eine Begründung dieses Brückenprinzips wurde bereits angedeutet: Normen haben die Funktion, die menschliche Praxis anzuleiten. Wird ein Handeln gefordert, welches nicht realisierbar ist (so wird die Reduktion schwerwiegender Folgen des Diabetes Mellitus nicht in allen Fällen gemessen sein), kann die Norm nicht praktisch wirksam werden, also ihre Funktion nicht erfüllen. Ein Beispiel wäre ein Diabetes-Patient, der nicht in der Lage wäre, die dringend empfohlene körperliche Aktivität umzusetzen.

Drei Arten mangelnder Realisierbarkeit können unterschieden werden.

- Erstens eine nicht realisierbare Handlung wird vorgeschrieben.
- Zweitens, das Erreichen eines nicht realisierbaren Ziels wird vorgeschrieben.
- Drittens, wird eine Handlung vorgeschrieben, die ungeeignet ist, ein vorgegebenes Ziel zu erreichen.

Im zweiten Fall kann die fragliche Norm nicht praktisch wirksam werden, weil sie etwas zu realisieren vorschreibt, was nicht zu realisieren ist. Sie kann nicht den mit ihr gewünschten Effekt hervorrufen. Ähnliches gilt im dritten Fall. Dort kann auf die vorgeschriebene Weise der mit der Norm

intendierte Effekt nicht erreicht werden. Die Norm kommt nicht weiter in Betracht.

In allen drei Fällen sind die kritischen Instanzen empirischer Natur. Behauptungen, dass eine Handlung nicht realisierbar, ein Ziel nicht erreichbar oder ein Ziel nicht auf eine bestimmte Art und Weise erreichbar ist, sind deskriptiv und sagen für sich genommen noch nichts darüber aus, was getan werden soll. Erst zusammen mit den Brückenprinzipien können normative Schlüsse gezogen werden, nämlich Handlungsanweisungen akzeptiert oder zurückgewiesen werden.

Manchmal fließen in die Begründungen von Handlungsregeln Behauptungen ein, die einer empirischen Überprüfung nicht standhalten. Da die Begründungen dann fehlerhaft sind, sind die Normen unbegründet. Ihre Funktion der Regulierung der menschlichen Praxis wird damit problematisch, denn unbegründete Normen sind mit einem Befolgungsproblem konfrontiert. Das wird besonders dort deutlich, wo Normen bezweifelt werden oder Normenkonflikte auftreten. Wenn auf die Frage, warum gerade dieser Regel zu folgen sei oder warum diese Handlungsvorschrift anderen vorzuziehen sei, keine argumentativ haltbare Antwort gegeben werden kann, wird ihre Befolgung aus Einsicht unmöglich. Kliniker denken dabei an die Probleme bei der Umsetzung von Leitlinien. Entweder werden derlei Normen nicht beachtet, womit sie folglich nicht mehr praktisch wirksam wären oder ihre Einhaltung wird erzwungen. Es sind dann Sanktionen notwendig, die nur bedingt zur Befolgung von derlei Handlungsanweisungen führen. Im Fall der Leitlinien ist ihre praktische Umsetzung stets gefährdet. Unzureichende Begründungen unterhöheln die Bereitschaft zur Befolgung von Normen, sie vereiteln oder gefährden damit zumindest ihre praktische Wirksamkeit.

Normenbegründungen können angesichts empirischer Tatsachen fehlerhaft sein. Aus den genannten Gründen unterminiert dies die praxisregulierende Funktion. Eine Norm kommt dann nicht in Betracht. Das entsprechende

Brückenprinzip wird als „Kongruenzprinzip“ bezeichnet und kann folgendermaßen formuliert werden:

Wenn die Begründung einer bestimmten Norm im Lichte unseres empirischen Wissens nicht haltbar ist, dann kommt diese Norm aus empirischer Sicht nicht in Betracht.

Manche Handlungsnormierungen können, wenn sie praktisch wirksam werden, Folgen nach sich ziehen, die nicht akzeptabel sind. Angenommen es würde die Zulassung eines Medikaments gefordert, welches nachweislich in der Lage ist, bestimmte Beschwerden zu lindern. Jedoch seien mit diesem Medikament Nebenwirkungen verbunden, die als wesentlich gravierender bewertet werden als die ursprünglichen Beschwerden. Aufgrund inakzeptabler negativer Konsequenzen wäre die Forderung nach einer Zulassung zurückzuweisen.

In Fällen wie diesem können Normen nicht in der Weise wie es erwünscht ist praktisch wirksam werden, nämlich nicht ohne nicht hinnehmbare negative Folgeerscheinungen hervorzurufen. Dies ist die Begründung für das sogenannte „Verknüpfungsprinzip“.

Dieses Brückenprinzip lässt sich kurz auf die Formel bringen, dass nur hinsichtlich ihrer absehbaren Implikationen akzeptierte Normen angewandt werden sollen.

Wenn laut empirischem Wissen eine bestimmte Norm Folgen erwarten lässt, die nicht akzeptabel sind, kommt diese Norm aus empirischer Sicht nicht in Betracht.

Von Normen wird aber nicht nur erwartet, dass sie unser Handeln anleiten, sondern dass sie dazu auch „besser geeignet“ sind als mögliche Alternativnormen. Was bedeutet dies? Dieser Terminus bezieht sich auf verschiedene Aspekte der Normendiskussion. Verschiedene Handlungsregeln können etwa ein Ziel gleichermaßen zu realisieren erlauben, wobei manche jedoch besser geeignet sein werden. Ihre Umsetzung erfordert bspw. im Vergleich weniger finanzielle Mittel. In der Regel wird nun erwartet mit Ressourcen möglichst effektiv umzugehen. Dies gehört zu den erwünschten Verfahrensweisen, die kaum zu beanstanden sind. Einige Normen schneiden in

diesem Punkt besser als konkurrierende Vorschläge ab. Daneben sind Normen daraufhin überprüfbar, ob sie dem als richtig und gesollt Erkannten förderlicher sind als gegebene Alternativen, etwa weil sie auf geringere psychologische Widerstände stoßen. Weiterhin können sie danach beurteilt werden, ob sie im Konfliktfall Handlungen auf das unter dem als richtig und gesollt Erkannten hin orientieren, was zu Recht Priorität genießt. Es sind noch weitere Facetten einer ‚besseren Eignung‘ denkbar. Normen, die sich in solchen Fällen als schwächer erweisen, werden verworfen, weil sie die Praxis nicht in der erwünschten Weise anleiten können, nämlich nicht effizient genug bzw. nicht gemäß der allgemein gesetzten Prioritäten. Wie ressourcenschonend, förderlich und prioritätsgerecht Handlungsanweisungen sind, ist ein empirisches Problem. Aussagen wie “diese Handlungsanweisung ist effizienter als jene” sind deskriptiv. Um sie für die Normenkritik heranziehen zu können, ist ein Brückenprinzip nötig, welches als „Prinzip der komparativen Beurteilung“ bezeichnet werden kann.

Wenn eine bestimmte Norm unter empirischen Gesichtspunkten nicht die effektivste Alternative darstellt, kommt diese Norm aus empirischer Sicht nicht in Betracht.

Die verschiedenen Arten der Normbewertung fassen wir unter dem Begriff der Effektivität zusammen. Es geht also darum, Normen unter Effektivitätsgesichtspunkten zu beurteilen. Schwierige Abwägungsfragen können dort resultieren, wo Normierungen hinsichtlich gewisser Effektivitätsaspekte besser, bezüglich anderer aber schlechter abschneiden als konkurrierende Vorschläge. Was unter diesen Umständen als effektiv gelten kann, muss im Einzelfall entschieden werden.

Um dem Leser eine Anleitung zur praktischen Umsetzung der beschriebenen Prinzipien an die Hand zu geben, ist in *Tabelle 1* demonstriert, wie mit Hilfe der Brückenprinzipien geklärt werden kann, welche Endpunkte für die Bewertung von Gesundheitsleistungen (hier Therapieerfolge bei Diabetes Mellitus) „geeignet“, „unsicher“ oder „ungeeignet“ sind. Um die Aussagen quantifizieren zu können, haben wir dem Prädikat „geeignet“ drei Punkte, dem Prädikat „unsicher“ einen Punkt und dem Prädikat „ungeeignet“ null

Punkte zugeordnet. Für die Auswertung und Interpretation dieser Ergebnisse ist zu berücksichtigen, dass ein Endpunkt insgesamt als „ungeeignet“ einzuschätzen ist, wenn dieser Endpunkt bei Prüfung eines einzelnen Brückenkriteriums das Prädikat „ungeeignet“ erhält. Wenn die Prüfung aller Brückenkriterien die Prädikate „geeignet“ oder „unsicher“ ergeben, empfehlen wir die Score-Punkte zu einem Gesamt-Score zu addieren. Die Validität des Gesamt-Scores besitzt noch exemplarischen Charakter und ist noch zu bestätigen.

<b>Brückenprinzipien</b>	<b>Endpunkt: Tod</b>	<b>Endpunkt: HbA<sub>1c</sub></b>	<b>Endpunkt: gLQ</b>
<b>Sollen impliziert Können (Realisierbarkeitspostulat)</b> (wenn etwas gefordert wird, muss es erfüllbar sein)	Eine Studie zur Überprüfung des Endpunktes 'Tod' würde eine Beobachtung über mehr als 10 Jahre erfordern und erhebliche Ressourcen binden.	HbA <sub>1c</sub> ist einfach zu bestimmen	Die gLQ ist leicht und zu akzeptablen Kosten zu messen.
	<i>UNGEEIGNET</i>	<i>GEEIGNET</i>	<i>GEEIGNET</i>
<b>Kongruenzprinzip</b> (besteht ein gesicherter Zusammenhang zwischen Messung und Aussage)	Der Zusammenhang zwischen vorzeitigem Tod infolge eines Schlaganfalls oder Herzinfarkts mit einem unzureichend eingestellten Diabetes Mellitus ist zutreffend	Postprandiale Blutzucker Spitzenwerte scheinen einen höheren prädiktiven Wert haben als HbA <sub>1c</sub>	Der Zusammenhang zwischen gLQ und Spätschäden ist unbestritten.
	<i>GEEIGNET</i>	<i>UNSICHER</i>	<i>GEEIGNET</i>
<b>Verknüpfungsprinzip</b> (nur hinsichtlich ihrer erkannten Implikationen akzeptierte Normen sollen angewandt werden; Normen mit inakzeptablen Folgen sind zurückzuweisen)	Es fehlt die Bereitschaft, die für den Nachweis erforderlichen, langfristigen Studien zu finanzieren.	Es ist anerkannt, dass Kontrollen regelmäßig durchzuführen sind. Die monetären Kosten sind akzeptabel	Implikationen sind durch die Messung der gLQ nicht zu erwarten, da Daten der gLQ bisher kaum handlungsrelevant sind
	<i>UNGEEIGNET</i>	<i>GEEIGNET</i>	<i>UNSICHER</i>
<b>Komparative Beurteilung</b> (eignen sich andere Kriterien besser für die Beurteilung als dieses Kriterium)	Sind andere Endpunkte, z.B. gLQ, Erblindung, Amputation, Herzinfarkt, sinnvoller als der Endpunkt „Tod“?	Die Messung postprandialer Blutzuckerspitzenwerte könnte sinnvoller sein als die Messung von HbA <sub>1c</sub>	Im Vergleich zu anderen Endpunkten liefert die Messung der gLQ weniger handlungsrelevante Daten.
	<i>UNSICHER</i>	<i>UNSICHER</i>	<i>UNGEEIGNET</i>
<b>Gesamturteil</b>	<b>UNGEEIGNET</b>	<b>GEEIGNET (8 Punkte)</b>	<b>UNGEEIGNET</b>

Tabelle 1: Eignung verschiedener Endpunkte, „Tod“, „HbA<sub>1c</sub>“ oder „gesundheitsbezogene Lebensqualität (gLQ)“ zur Bewertung von Therapieerfolgen bei Patienten mit Diabetes Mellitus. Die Einschätzung der Eignung ist jeweils kursiv dargestellt. Wenn ein Endpunkt

*zur Beschreibung eines einzelnen Brückenprinzips ungeeignet ist, wird dieser Endpunkt insgesamt als „ungeeignet“ eingestuft. Andernfalls werden für die Ermittlung des Gesamt-Scores die Einzel-Scores addiert. Bei einem Gesamt-Score von 4 bis 7 Punkten wird ein Endpunkt als „unsicher“, bei einem Gesamt-Score von 8 bis 12 Punkten als „geeignet“ bezeichnet.*

Die Tabelle liest sich wie folgt: In der Spalte ganz links sind die Brückenprinzipien stichwortartig aufgeführt. Die darauf folgende Spalte zeigt den Endpunkt Tod. Dieser Endpunkt stellt eine Norm dar. Nach dem Brückenprinzip „Sollen impliziert Können“ stellt sich die Frage, ob der Tod eines an Diabetes erkrankten Menschen überhaupt gemessen werden kann. Spätfolgen bei Diabetes, Erkrankungen, die zum Tode führen können, treten erst nach zehn oder mehr Jahren ein. Dieser lange Zeitraum impliziert zwei Probleme. Zum einen ist die wissenschaftliche Beobachtung eines Patienten in einer doppel-blind randomisierten Studie extrem aufwendig und faktisch nicht finanzierbar. Zum anderen kann in diesen zehn Jahren eine neue Therapie dazu führen, dass das Leben diabetischer Patienten verlängert werden kann und eine solche Langzeitstudie unethisch wäre. Nach diesem Brückenprinzip wäre dieser Endpunkt abzulehnen. Gemäß des Kongruenzprinzips wäre dieser Endpunkt allerdings zu akzeptieren. Laut herrschender Lehrmeinung besteht ein Zusammenhang zwischen vorzeitigem Tod infolge eines Schlaganfalls oder eines Myokardinfarkts bei einem unzureichend eingestellten Diabetiker.

Nach dem Verknüpfungsprinzip sollten mögliche Folgen, wie zum Beispiel eine hohe finanzielle Belastung zur Finanzierung einer Studie oder das Vorhalten innovativer lebensrettender Therapien akzeptiert werden. Dies auch durchzuführen ist unwahrscheinlich.

Nach dem Prinzip der komparativen Beurteilung muss dieser Endpunkt im Lichte von Alternativen geprüft werden. Hier bieten sich weitere Endpunkte wie schwere Hypoklykämien, Sehschwäche, Erblindung, Neuropathie oder akute Koronarsyndrome an. Zusammenfassend kann festgestellt werden, dass der Tod kaum ein geeigneter Endpunkt sein wird.

Die nächste Spalte diskutiert den HbA<sub>1c</sub> als zu kritisierende Norm.

Der Blutzuckerspiegel des Diabetikers kann über den HbA<sub>1c</sub> bestimmt werden. Dieser Wert ist einfach zu messen. Es besteht allerdings kein Konsens darüber, ob der HbA<sub>1c</sub> ein sicherer Prädiktor für Spätschäden darstellt. Zur Zeit werden die postprandialen Spitzen als mögliche Ursache diskutiert. Die möglichen Folgen, also der Aufwand für das Messen für den Patienten und die finanzielle Belastbarkeit liegen im Bereich des Akzeptablen. Kritisch ist hingegen zu sehen, dass die Messung postprandialer Spitzen durchaus eine Alternative zum HbA<sub>1c</sub> darstellt.

Die Spalte ganz rechts betrachtet die gesundheitsbezogene Lebensqualität als zu kritisierende Norm.

Die Lebensqualität als Endpunkt ist leicht mit akzeptablen Kosten zu messen. Der Zusammenhang zwischen Spätfolgen wie Erblindung oder Amputation und einer beeinträchtigten Lebensqualität ist unbestritten.

Implikationen sind durch die Messungen der Lebensqualität nicht zu erwarten. Im Vergleich mit anderen Endpunkten liefert die Lebensqualität wenige handlungsrelevante Daten, weil die für die Lebensqualität relevanten Faktoren vor Jahren hätten beeinflusst werden müssen.

## **1.6. Zusammenfassung: Einführung und Analyse der Beiträge der eingereichten Artikel**

### 1.6.1 Zusammenfassung: Einführung

Die Zielsetzung der Gesundheitsökonomie besteht in der Optimierung der Allokation knapper Ressourcen. Der Stand der Forschung hat gezeigt, dass mit epidemiologischen- Krankheitskosten-, Kosten-Effektivitäts- und Kosten-Nutzwert-Studien Instrumente sowohl zur Erfassung der Kosten als auch zur Allokation auch aufgegliedert für die oben genannten einzelnen Perspektiven zur Verfügung stehen.

Anwendungshindernisse verringern die Beiträge der Gesundheitsökonomie. Solche sind zum einen das sektorale Gesundheitssystem mit seiner Anreizinkompatibilität, zum anderen methodische Probleme. Insbesondere die Messung und Analyse patientenrelevanter Endpunkte stellt ein bisher vernachlässigtes Feld dar.

Zusammenfassend kann festgestellt werden, dass für die Erfassung der Kosten und für die Allokation zwar Instrumente zur Verfügung stehen, diese in ihrer Wirkung aber beschränkt sind.

Es stellt sich nun die Frage, welche Schlussfolgerungen gezogen werden müssten, damit die Gesundheitsökonomie Beiträge für die Entscheidungshilfe leisten kann.

Diese Thematik wird im Kapitel „Schlussfolgerungen“ weiter untersucht werden.

### 1.6.2 Analyse der Beiträge der eingereichten Artikel

Die Analyse vorliegender Arbeit wird zeigen, dass Rahmenbedingungen, Probleme des QALY-Konzepts und die Analyse patientenrelevanter Endpunkte wie z.B. der Nutzen Hinderungsgründe für die Beiträge der Gesundheitsökonomie darstellen. Im Rahmen dieser Analyse stellt sich die Frage, welche Beiträge die eingereichten Artikel leisten.

Der Artikel „Wissenschaftstheoretische Überlegungen zur Werturteilsproblematik der so genannten Vierten Hürde“ leistet zweierlei.

Zum einen stellt er bzgl. der Rahmenbedingungen ein Erklärungsmodell bereit, das zeigt, warum sich Menschen beim Problemlösen in komplexen, vernetzten, intransparenten und dynamischen System schwer tun.

Zum anderen zeigt er für die Analyse des QALY-Konzepts und der patientenrelevanten Endpunkte mögliche Problemlösungen auf.

Bei einem QALY wird ein Preis festgelegt. Dieser Preis ist ein Werturteil. Bisher ist meines Wissens weder in der Gesundheitsökonomie noch in der Gesundheitspolitik ein methodisch stringenter Ansatz zu finden, um Werturteile zu diskutieren oder gar zu begründen. Genau diese Problematik behandelt obiger Artikel unter Zuhilfenahme der Brückenprinzipien.

Die patientenrelevanten Endpunkte sind durch die Festlegung des IQWiG in Diskussion geraten. Patientenrelevante Endpunkte sind Werturteile und werden zur Zeit weder geprüft noch ist eine Diskussion durch das IQWiG

erlaubt. Auch bei der Analyse patientenrelevanter Endpunkte können die Brückenprinzipien herangezogen werden.

Die Beiträge oben genannten Artikels bestehen darin, dass sowohl mit dem Problemlösungsansatz als auch mit den Brückenprinzipien Werkzeuge für die aktuellen Probleme zur Verfügung gestellt werden.

Der Artikel „Die Kosten der Frühverrentung am Beispiel der Schizophrenie“ fokussiert weniger auf das QALY-Konzept oder die patientenrelevanten Endpunkte als auf die Rahmenbedingungen.

Obiger Artikel beleuchtet die transsektoralen Probleme. Die strikte Trennung wird zwar häufig diskutiert, aber die Tragweite der Bedeutung des sektoralen Denkens wurde meines Wissens empirisch bisher nicht erfasst. Genau diese Problematik analysiert obiger Artikel. Er zeigt, dass Einsparungen durch mögliche transsektoraler Transfers nicht im einstelligen Prozentbereich liegen, sondern die Größenordnung des gesamten zur Diskussion stehenden Marktes erreichen. Der Wert dieses Artikels liegt darin, dass er die Bedeutung transsektoraler Sichtweisen verdeutlicht.

Der Artikel „Kostenanalyse der Schizophreniebehandlung in Deutschland“ ist insofern bemerkenswert, weil er alle oben genannten Hindernisse wie Rahmenbedingungen, QALY-Konzept und patientenrelevante Endpunkte beleuchtet. Mit Hilfe einer Markov-Modellierung zeigt er beispielhaft, wie Transfers über Sektoren analysiert und gestaltet werden können. Darüber hinaus analysiert dieses Forschungsprojekt QALIES und patientenrelevante Endpunkte. Insbesondere die Tatsache, dass die Ergebnisse aller Analysen in eine Richtung verlaufen, also nicht widersprüchlich sind, ist aus mehreren Gründen bemerkenswert. Zum einen ermöglicht die Korrelation der Ergebnisse eine externe Validierung und bietet damit zum anderen eine solide Entscheidungsgrundlage.

Der Beitrag dieses Artikels liegt in seiner komplexen Betrachtungsweise, die sowohl Rahmenbedingungen als auch Methodenprobleme berücksichtigt.

Die Kosten pro QALY werden von Institutionen wie NICE gefordert und dementsprechend von Anbietern von Gesundheitsleistungen für unterschiedliche Populationen wie geriatrische Patienten, Kinder oder Psychotiker berechnet. Dabei stellt sich eine grundsätzliche Frage: „Ist das von mir eingesetzte Instrument zur Berechnung der QALIES bei der von mir untersuchten Population valide?“

Mit genau dieser Fragestellung befasst sich der Artikel „Constructing an Index for the subjective Well-being under Neuroleptic Scale (SWN)“, indem er Forschern detailliert aufzeigt, wie Validität berechnet werden kann.

Ein anderer Beitrag ist mehr impliziter Natur. In jüngster Zeit drängt der Begriff „Nutzen“ mehr und mehr in die öffentliche Diskussion. „Nutzen“ lässt sich als latente Variable in Form einer konfirmatorischen Faktorenanalyse darstellen. Genau diese Konstruktion wird detailliert in oben genannter Arbeit beschrieben..

## 2. Begründung der ausgewählten Krankheiten

### 2.1 Schizophrenie

Die Schizophrenie gehört zu den endogenen Psychosen im Triadischen System der Psychiatrie (Clouth, 2003).

Allerdings sind Ursachen, Erscheinungsbilder und der Verlauf nicht vollständig erforscht. Es handelt sich um eine chronische Erkrankung, die meist im Erwachsenenalter ausbricht. Das Krankheitsbild ist durch das gleichzeitige Auftreten verschiedener Symptome gekennzeichnet. Dies sind die Positiv-Symptomatik (Halluzinationen, Angst etc.) und die Negativ-Symptomatik (Apathie, emotionale Zurückgezogenheit etc.). Eine repräsentative Umfrage bei niedergelassenen Psychiatern, Institutsambulanzen und stationären psychiatrischen Einrichtungen im Jahre 1997 ergab, dass rund 136.000 Patienten in Deutschland behandelt wurden. Diese Zahl ist sicherlich eine konservative Schätzung, weil bestimmte Ärztegruppen wie Allgemeinärzte in der Befragung nicht eingeschlossen waren und weil sich die Prävalenz aufgrund des demografischen Schubs wahrscheinlich zusätzlich erhöht hat. Die Kosten pro Patient beliefen sich in oben genannter Studie für die direkten Kosten auf rund 10.000 €, für die indirekten Kosten auf rund 20.000 €. Der Budget Impact lag für die direkten Kosten bei 1,5 Milliarden €, für die indirekten Kosten bei rund 5,6 Milliarden €.

Interessant ist, dass die bei der Behandlung der Schizophrenie in Anspruch genommenen Ressourcen aus mehreren Sektoren kommen. Aus der Perspektive der Kasse sind dies das Krankenhaus, der niedergelassene Bereich sowie die Rentenkasse.

Die Krankheit Schizophrenie bietet sich für das gestellte Thema an, weil es sich nicht nur um eine bzgl. Kosten und Prävalenz bedeutende Krankheit handelt, sondern weil sich alle Instrumente und alle in Teil 1.5 genannten Anwendungshindernisse exemplarisch darstellen lassen.

## **2.2 Schwere Sepsis**

Eine Sepsis ist eine systemische Reaktion auf eine schwere Infektion. Treten zu dieser Reaktion ein- oder mehrere Organversagen hinzu, so handelt es sich um eine schwere Sepsis (Schmid, 2002).

Die Mortalitätsrate der schweren Sepsis liegt bei 47% wobei in Deutschland jährlich 75.000 Menschen an einer schweren Sepsis erkranken (Brunkhorst, 2005), die Kosten belaufen sich in Deutschland auf rund 6,3 Milliarden €.

Zwei Gründe sprechen dafür, diese Erkrankung als Beispiel aufzuführen.

Zum einen gibt es nur ein Medikament, Xigris, das für die Behandlung der schweren Sepsis eine Zulassung erhalten hat. Die Berechnung der Kosteneffektivität ist damit extrem transparent darzustellen und das bei einer Krankheit mit sehr hoher Mortalität. Zum anderen ist insbesondere die Frage des Endpunktes äußerst transparent, weil nach dem Überleben der schweren Sepsis die Lebensqualität nicht tangiert wird. Insofern zeigt dieses Krankheitsbild im Gegensatz zur Schizophrenie, dass nicht nur hoch komplexe Problemfelder Einsatzgebiete der Gesundheitsökonomie sein können, sondern auch klar abgegrenzte Gebiete.

## 3. Begründung der eingereichten Arbeiten

### 3.1 Wissenschaftstheoretische Überlegungen zur Werturteilsproblematik dargestellt am Beispiel der so genannten „Vierten Hürde“

Bei Diskussionen über den Entscheidungsfindungsprozess im deutschen Gesundheitswesen schleichen sich krypto-normative Aussagen ein. Beispiele hierfür sind die patientenrelevanten Endpunkte, die Prof. Sawitcki vom Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) für die Bewertung von Arzneimitteln festsetzt, die Festlegung bestimmter Methoden zur Analyse des Patientennutzens oder die Forderung nach regulierten Preisen. Alle diese Beispiele unterliegen dem Sein-Sollen-Fehlschluss (Ruß, 2005). Patientenrelevante Endpunkte oder die Auswahl bestimmter Methoden sind Werturteile, die systematisch diskutiert werden müssten. Eine solche Wertediskussion findet in Deutschland jedoch nicht statt.

Das liegt meines Erachtens daran, dass die Instrumente zur Beurteilung von Werturteilen bei den Entscheidern im deutschen Gesundheitswesen, falls überhaupt, nur begrenzt Eingang gefunden haben. Der Artikel „Wissenschaftstheoretische Überlegungen zur Werturteilsproblematik dargestellt am Beispiel der so genannten „Vierten Hürde“ ist der erste Beitrag, der Instrumente zur Werturteilsbegründung für das deutsche Gesundheitswesen an aktuellen Beispielen darstellt und damit eine Basis für eine wissenschaftlich fundierte Wertediskussion bereitstellt. Aus diesem Grund findet er Berücksichtigung in vorliegender Arbeit.

#### 3.1.1 Kontext des Entstehens

Die Aussage der Gesundheitsministerin Ulla Schmidt, dass Medikamente, die 10% mehr leisten, nicht 300% mehr kosten dürfen (BMGS Reden), fachte eine Diskussion über die Wertigkeit von Medikamenten an. Insbesondere die Arbeit des National Institut of Clinical Excellence (NICE) ließ diese Frage für Deutschland dringlich erscheinen. Es stellte sich das Problem, welche medizinischen Endpunkte erhoben und auf welche Weise

sie ausgewertet werden müssen, um die Frage der Wertigkeit beantworten zu können. Im Kontext dieser Diskussion, die sich in den Jahren 2002 - 2003 entfachte, entstand dieser Artikel.

### 3.1.2 Beiträge des Artikels zur Fragestellung der Arbeit

Professor Busse analysierte weltweit alle Institute, die Nutzenbewertung durchführten (Zentner, 2005). Diese Analyse kam zu zweierlei Erkenntnissen:

1. Die Zielgrößen, die untersucht wurden, sind uneinheitlich. Für die einen stehen QALYs an oberster Stelle, für andere die Lebensqualität oder Mortalität.
2. Alle Institutionen entscheiden sich für Zielgrößen und fällen damit Werturteile, ohne diese zu begründen. Sie handeln also krypto-normativ.

Der Artikel "Wissenschaftstheoretische Überlegungen zur Werturteilsproblematik dargestellt am Beispiel der so genannten "Vierten Hürde" liefert zwei Beiträge zur Bearbeitung krypto-normativer Aussagen.

Zum einen zeigt er auf, dass bei der Auswahl von Endpunkten bzw. Zielen krypto-normative Aussagen einfließen und damit die Optimierung knapper Ressourcen auf nicht begründeten Zielparametern bzw. Endpunkten basiert. Damit wird konsequenterweise die Zielsetzung „Optimierung der Allokation“ tangiert.

Zum anderen stellt der Artikel Instrumente um krypto-normative Aussagen aufdecken, diskutieren und begründen zu können bereit.

### 3.1.3 Limitierungen

Offen bleibt das Problem, wer oder welche Institutionen eine solche auf Brückenprinzipien basierende Diskussion führen und Entscheidungen treffen könnte.

Dieses Problem wird im Schlussteil bearbeitet werden.

### 3.1.4 Inhalt

Der oben genannte Artikel behandelt wissenschaftstheoretische Grundlagen, die in besonderer Weise gesundheitsökonomische Fragestellungen in Deutschland betreffen. Kosten-Effektivitätsanalysen und Kosten-Nutzwert-Analysen sind unter Umständen in der Lage, Antworten auf die Frage zu geben, ob eine Therapie A einer Therapie B vorzuziehen sei. Grundlage einer solchen Analyse ist die Existenz eines Entscheidungskriteriums bzw. eines klinischen Endpunktes. Nun kann für den Therapieerfolg durchaus das Problem entstehen, dass mehrere Endpunkte zur Verfügung stehen und damit je nach Auswahl des klinischen Endpunktes eine Manipulation der Ergebnisse möglich ist. Die Auswahl von Endpunkten und oder Methoden fließen hier als krypto-normative Aussagen ein (Raffée, S.46, 1974). Ein möglicher Ansatz dieses Problem zu lösen, bilden die Brückenprinzipien, mit denen Zielparameter und damit verbunden Werturteile beurteilt werden können (Raffée et al., S. 218ff, 1979).

Im Einzelnen werden das Praktikabilitätsprinzip, das Kongruenz-Postulat, das Verknüpfungsprinzip sowie das Prinzip der komparativen Beurteilung beleuchtet.

Das Praktikabilitätsprinzip fordert, dass nur das gewünscht werden soll, was auch erreichbar ist. Für die Gesundheitsökonomie bedeutet dies, dass klinische Endpunkte, deren Messbarkeit zum Beispiel wie bei chronischen Krankheiten (wie Osteoporose oder Diabetes) erst nach zwanzig Jahren eintritt, aus Zeit- und Kostengründen nicht gefordert werden sollen.

Das Kongruenz-Postulat fordert, dass Normen dann kritisiert werden können, wenn sie auf Tatsachen beruhen, die sich als falsch erwiesen haben. Dieses Postulat betrifft die Gesundheitsökonomie in besonderem Maße, weil gerade der Einsatz von Surrogatparametern bei chronischen Krankheiten faktisch unverzichtbar ist.

Das Verknüpfungsprinzip postuliert, dass nur solche Normen akzeptiert werden sollen, deren Implikationen anerkannt werden. Kosteneffektivitätsanalysen und Kostennutzwertanalysen sagen zwar etwas über die Vorteilhaftigkeit einer Therapie aus, nicht aber über deren Budget Impact. Nun

kann es sein, dass eine Therapie zwar kosteneffektiv ist, aber eine beträchtliche finanzielle Belastung für das Gesundheitssystem darstellt. Diese Implikation muss ex ante explizit akzeptiert werden, wenn man sich für die kosteneffektivere Alternative entscheiden will.

Das Prinzip der komparativen Beurteilung beleuchtet Normen im Licht möglicher Alternativen. So stellt sich die Frage, ob es Sinn macht, die Überlebensdauer bei Krebspatienten als klinischen Endpunkt zu wählen oder ob alternativ nicht die Lebensqualität mit einfließen sollte. Nach diesem Postulat soll jenes gewählt werden, das den breitesten Konsens vereinen kann.

### **3.2 Kosten der Frühverrentung am Beispiel der Schizophrenie**

#### 3.2.1 Kontext des Entstehens

Die Studie über die Kosten der Schizophrenie „Untersuchungen zu den direkten und indirekten Kosten der Schizophrenie“ (Schulenburg, 1998) zeigte, dass die indirekten Kosten die direkten Kosten um ca. 100% übersteigen. Interessant war, dass die indirekten Kosten zu 75% durch Frühverrentung verursacht waren. Die Fragestellung transssektoraler Lösungsansätze, hier zwischen Krankenkassen und Rentenkasse, wurde Ende der 90er Jahre gestellt, aber nicht in Angriff genommen. Ein erster Ansatz, um diese Frage zu beantworten, wurde durch die oben genannte Studie insofern in Angriff genommen, indem sie die finanzielle Belastung der Rentenkasse durch die Frühverrentung der an Schizophrenie erkrankten Patienten ermittelte.

#### 3.2.2 Beiträge des Artikels zur Fragestellung der Arbeit

Die Zielsetzung der Gesundheitsökonomie liegt in der Optimierung der Allokation knapper Ressourcen und zwar für alle Perspektiven, die des Patienten, der Zahler, des Krankenhauses, der KV sowie der Gesellschaft. Durchforstet man die Literatur, so fällt auf, dass meist nur ausgewählte Perspektiven wie die der Krankenhäuser, der Krankenkasse oder der kassenärztlichen Vereinigung eingenommen werden. Teilweise werden auch die durch Krankheitstage bedingten Einkommensverluste als indirekte Kosten berücksichtigt. Diese Sichtweise ist insofern eng, weil mit der Frühverren-

tung der größte Kostenblock ausgeschlossen bleibt, wie oben genannter Artikel zeigt. Es scheint jedoch plausibel, dass jene Ansatzpunkte berücksichtigt werden sollten, die auch den stärksten Beitrag für die Allokation leisten können. Obiger Artikel zeigt, dass die Kosten der Frühverrentung den wichtigsten Kostenblock darstellen und transsektorale Perspektiven für gesundheitsökonomische Analysen sinnvoll sein können.

### 3.2.3 Limitierungen

Die Arbeit „Kosten der Frühverrentung am Beispiel der Schizophrenie“ lässt drei Fragen unbeantwortet. Erstens wird nicht analysiert, welche Faktoren eine Frühverrentung durch Schizophrenie verursachen. Dieser Frage ging eine gesonderte Hypothesen generierende Studie nach, deren Ergebnisse zur Zeit zur Veröffentlichung eingereicht sind. Mögliche Faktoren können z.B. der Schweregrad der Krankheit oder die Qualität der Behandlung sein. Die Kenntnisse solcher Faktoren sind notwendig, um mögliche Schritte transsektoraler Problemlösungsinstrumente anzudenken.

Darin liegt die zweite Limitierung dieser Arbeit. Mögliche Problemlösungen wie „Integrierte Versorgung“ oder andere Modellversuche werden nicht diskutiert.

Drittens ist anzumerken, dass die Werturteilsproblematik kurz angesprochen, aber nicht ausreichend verfolgt wird.

Diese noch offenen Fragen werden im Schlussteil diskutiert werden.

### 3.2.4 Inhalt

Die Arbeit „Kosten der Frühverrentung am Beispiel der Schizophrenie“ zeigt, dass die Gesundheitsökonomie weiter greifen kann und andere Bereiche unseres Sozialsystems, wie zum Beispiel die Perspektive der Rentenversicherung, mit einbeziehen kann.

Die Schizophrenie ist eine psychotische Krankheit, die früh ausbricht (Häfner, S. 130, 2000) und die Arbeitsfähigkeit so nachhaltig herabsetzt, dass es zu einer Erwerbsunfähigkeit und damit zu einer Verrentung kommt. Die Krankheit verursacht damit erhebliche Kosten, die sich insbesondere in

den Ausgaben der Rentenversicherung widerspiegeln. Die Studie wertete Statistiken der GRV zur Frühverrentung aus. In Deutschland werden jährlich weit über 16.000 Männer und Frauen mit der Diagnose Schizophrenie (295 nach ICD-9) neu verrentet. Das durchschnittliche Verrentungsalter mit dieser Diagnose liegt bei Männern bei 39 Jahren und bei Frauen bei 42 Jahren. Bei Männern erfolgen 14,7% aller Verrentungen unter 40 Jahren mit der Diagnose Schizophrenie. Entsprechend hoch sind die Ausgaben der GRV. Allein der Barwert der Rentenzahlungen für einen durchschnittlichen männlichen Rentenfall beträgt 215.000 €. Werden auch Einnahmeausfälle berücksichtigt, dann ergibt sich ein Barwert von 560.000 €. Eine Simulation der Bestandszahlen auf Basis der Neuzugänge ergab für das Jahr 2000 einen Rentenbestand von 215.000 Personen mit einem Alter von unter 65 Jahren. Die GRV wird damit jährlich mit Rentenzahlungen in Höhe von 1,3 Milliarden € belastet. Hinzu kommen Einnahmeausfälle der Sozialversicherung und des Fiskus in Höhe von nochmals zwei Milliarden €.

Diese Kosten müssen von allen berufstätigen Bürgern finanziert werden. Es ist offensichtlich, dass die hohen Kosten der Frühverrentung durch Schizophrenie einen Einfluss auf die Lohnnebenkosten und damit auf die Arbeitslosenquote ausüben.

### **3.3 Kostenanalyse der Schizophrenie-Behandlung in Deutschland**

#### **3.3.1. Kontext des Entstehens**

Die Einführung der so genannten Atypika der zweiten Generation mit Risperdal (1996) und Zyprexa (1997) führte zu einer starken Verringerung der Liegedauer im Krankenhaus, womit sich die direkten Kosten der Krankenkasse durch die gestiegenen Arzneimittelpreise zwar erhöhten, aber durch die verringerte Liegedauer aus Kassensicht senken ließen. Wie in Abschnitt 1.5 „Anwendungshindernisse“ gezeigt, verfügen die Krankenkassen nicht über die Daten, um solche, die Sektoren Krankenhaus und kassenärztliche Vereinigung überschreitenden Analysen durchzuführen. Aus diesem Kontext heraus, das heißt um die Wirtschaftlichkeit des Einsatzes der Atypika der zweiten Generation zu demonstrieren, entstand diese Arbeit.

### 3.3.2 Beiträge des Artikels „Kostenanalyse der Schizophreniebehandlung in Deutschland“ zur Arbeit

In dem Gliederungspunkt 1.3 „Stand der Forschung“ wurden die Instrumente der Gesundheitsökonomie dargestellt. Die oben genannte Arbeit ist insofern bemerkenswert, weil sie eine Vielzahl von Instrumenten einsetzt. Die Kostenanalyse der Schizophreniebehandlung in Deutschland beinhaltet eine Krankheitskosten-Analyse, eine Kosten-Effektivitäts-Analyse und eine Markov-Modellierung. Die Studie selbst enthält auch eine Kosten-Nutzwert-Analyse, die als Poster veröffentlicht wurde (Spannheimer, 1999). Dieser massive und vielfältige Einsatz gesundheitsökonomischer Instrumente in einem Artikel ermöglicht die Vermittlung wichtiger zur Zeit gängiger gesundheitsökonomischer Instrumente.

### 3.3.3 Limitierungen

Obwohl die Arbeit umfangreich ist, bleiben dennoch drei Fragen unbeantwortet.

Erstens wurde zwar die gesellschaftliche Perspektive in Form verlorener Arbeitstage erhoben, die Kosten der Frühverrentung wurden aber nicht berechnet.

Nicht angesprochen wurde außerdem die Werturteilsdiskussion. Gerade bei dieser Arbeit, in der Endpunkte, wie Rückfälle und QALIES (Spannheimer, 1999), berechnet wurden, hätte die Wertediskussion aber vertiefende Beiträge leisten können.

Letztendlich regen die Autoren eine Diskussion transsektoraler Probleme an, unterbreiten jedoch keine Vorschläge.

### 3.3.4 Inhalt

Ziel der Studie war eine pharmako-ökonomische Analyse der Behandlung von Schizophrenie-Patienten mit Olanzapin, Risperidon und Haloperidol in Deutschland. Die Behandlung mit Haloperidol entsprach hierbei der traditionellen Methode. Olanzapin und Risperidon sind Vertreter der Atypika der zweiten Generation.

Die Analyse wurde mit Hilfe eines klinischen Entscheidungsmodells durchgeführt, das von einem internationalen Expertenpanel entwickelt wurde. Das Modell einschließlich der klinischen Daten (u. a. aus multinationalen klinischen Studien) über den Verlauf der Erkrankung wurde übernommen. Daten zur Praxis bei der medizinischen Behandlung von Schizophrenie sowie Kostendaten für Deutschland wurden in einem nationalen Expertenpanel ermittelt und in das Modell integriert. Anschließend wurden eine Kosten-Analyse und eine Kosten-Effektivitäts-Analyse durchgeführt

Die Gesamtkosten der Therapie lagen bei allen 3 Therapien in der gleichen Größenordnung (Risperidon DM 123.004, Haloperidol DM 121.868, Olanzapin DM 119.572, jeweils pro Patient für 5 Jahre). Die Kosteneffektivität von Risperidon bzw. Olanzapin war um den Faktor 2,3 bzw. 2,5 besser als die von Haloperidol.

Bei Betrachtung der Gesamtkosten werden die hohen Tagestherapiekosten der atypischen Neuroleptika durch Einsparungen bei der Behandlung von Negativsymptomatik und Rückfällen ausgeglichen.

### **3.4 Constructing an Index for the Subjective Well-being Under Neuroleptics Scale (SWN), Short Form: Applying Structural Equation Modeling for testing Reliability and Validity of the Index**

#### 3.4.1 Kontext des Entstehens

Die Allokation knapper Ressourcen setzt Endpunkte bzw. Ziele voraus. Die gesundheitsbezogene Lebensqualität stellt einen solchen Endpunkt dar. Es muss allerdings erwähnt werden, dass die Validität von Lebensqualitätsanalysen kritisch gesehen werden kann. Grundlage dieser Kritik ist die Analyse von ca. 160 Arbeiten durch die Jury des Quality of Life Preises der Firma Lilly aus den Jahren 1996 bis 2004 ([www.lilly-pharma.de/Lilly\\_Quality\\_of\\_Life\\_Preis.587.0.html](http://www.lilly-pharma.de/Lilly_Quality_of_Life_Preis.587.0.html)).

Die Analyse zeigte, dass viele Arbeiten nicht valide waren. Zur Berechnung der inhaltlichen Konstruktvalidität wird die konfirmatorische Faktorenanalyse eingesetzt. Von den 160 eingereichten Arbeiten wurde nur bei drei Arbeiten mit Hilfe der konfirmatorischen Faktorenanalyse die inhaltliche

Konstruktvalidität berechnet. Ebenso wurde die Bildung von Summenscores teilweise willkürlich vorgenommen.

Der Artikel "Constructing an Index for the Subjective Well-being Under Neuroleptics Scale (SWN), Short Form: Applying Structural Equation Modelling for testing Reliability and Validity of the Index" zeigt am Beispiel des SWN auf welche Art und Weise mit Hilfe einer konfirmatorischen Faktorenanalyse die inhaltliche Konstruktvalidität berechnet werden kann. Darüber hinaus wird gezeigt, wie Kurzformen konstruiert und Indexe berechnet werden.

#### 3.4.2 Beiträge des Artikels zur Fragestellung der Arbeit

Meines Wissens ist dies der einzige Artikel, der Forschern ein Werkzeug an die Hand gibt, die Validität ihrer Ergebnisse zu belegen. Er ist insofern wichtig, weil Lebensqualitätsuntersuchungen häufig an Populationen durchgeführt werden, für die die Validität der eingesetzten Instrumente a priori nicht unbedingt unterstellt werden kann. Beispiele hierfür sind die Geriatrie oder die Pädiatrie.

Damit leistet der Artikel einen Beitrag zur Allokation.

#### 3.4.3 Limitierungen

Obwohl die Arbeit konkrete Anweisungen enthält, bleiben zwei Probleme unbeantwortet.

Zum einen wird eine mögliche Umsetzung der Ergebnisse nicht diskutiert. Fragen, wie z. B. wie man die Kosten pro gewonnenen Lebensqualitäts-Index-Punkten berechnen könnte oder wie man die Kurzform zur Berechnung von QALIES einsetzen könnte, werden nicht behandelt.

Zum anderen bleibt auch die Frage der Werturteilsproblematik offen. So kann die Frage, ob ein Gewinn an Lebensqualität für Diabetiker höher einzuschätzen ist als für einen an Schizophrenie erkrankten Menschen, auf Basis dieses Artikels nicht beantwortet werden.

#### 3.4.4 Inhalt

Strukturgleichungsmodelle (SEM) werden in der Psychologie und Soziologie häufig zur Testung der inhaltlichen Konstruktvalidität eingesetzt. Diese Technik spielt in der Lebensqualitätsforschung jedoch nur eine untergeordnete Rolle.

Die Zielsetzung dieser Arbeit besteht in der Aufgabe, das Potenzial von SEM für die Testung der Validität und der Konstruktion von Kurzformen aufzuzeigen und zwar am Beispiel des SWN.

Für die Arbeit wurden die Daten der GEO-Studie (**G**esundheitsökonomische **E**valuation von **O**lanzapin in Deutschland) genutzt. GEO ist eine prospektive, naturalistische, nicht intervenierende Beobachtungsstudie. 646 Patienten wurden in die Studie eingeschlossen, von denen 416 mit Olanzapin behandelt wurden, bei 230 wurde Haloperidol eingesetzt. Von diesen 646 Patienten gaben 360 Patienten ihre Einwilligung zur Befragung zur Lebensqualität mit Hilfe des Instrumentes SWN.

Dieses Instrument enthält 20 Items, die mit Hilfe einer konfirmatorischen Faktorenanalyse bzw. eines Strukturgleichungsmodells ohne Validitätsverlust zuerst auf zehn und letztendlich auf fünf Items gekürzt wurde.

## 4. Eingereichte Arbeiten

### 4.1 Artikel

#### **Wissenschaftstheoretische Überlegungen zur Werturteilsproblematik der so genannten „Vierten Hürde“**

Johannes Clouth - Katharina Schreeb

RPG 10.1 [2004], ©pmi Verlag AG, 2004

Johannes Clouth – Katharina Schreeb, Lilly Deutschland GmbH, Saalburgstraße 153,  
61350 Bad Homburg

Das bundesdeutsche Gesundheitswesen sieht sich mit stärker wachsenden Ausgaben bei gleichzeitiger Wachstumsschwäche der Einnahmen konfrontiert. Die Allokation knapper finanzieller Mittel wird damit umso dringlicher. Der Gesundheitsministerin Ulla Schmidt wird die Aussage zugesprochen, dass „Medikamente, die 10% mehr leisten, nicht 300% mehr kosten dürfen (BMGS Reden). In diesem Rahmen spielt die aktuelle Diskussion um eine so genannte „Vierte Hürde“ eine besondere Rolle. Die Vierte Hürde soll eine Antwort auf die Frage geben, welcher Erstattungsbeitrag für die Krankenkasse angemessen ist.

Dieser Entscheidungsprozess basiert nicht nur auf objektiven Daten, sondern beinhaltet auch werturteilsprägende Elemente. Zu betonen ist u.E., dass es sich bei dem Gesundheitsmarkt um einen Markt *sui generis* handelt. Der Konsument (Patient) trifft keine Entscheidung und er bezahlt auch nicht. Der Arzt trifft für ihn die Entscheidung und die Krankenversicherung bezahlt. Allerdings sind dem Arzt durch das Budget Grenzen gesetzt, so dass er ein Maximum an Therapieerfolg unter knappen Ressourcen erreichen muss. Häufig geschieht dies auch unter Einbeziehung des Patienten. In dieser Hinsicht fließen Marktelemente in den Prozess ein. Wenn basierend auf Werturteilsergebnissen marktpreispolitische Entscheidungen getroffen werden sollen, (z.B. Vierte Hürde, Kosten-Nutzen Bewertung von Arzneimitt-

teln<sup>2</sup>) ist es nützlich sich der wissenschaftstheoretischen Charakteristika urteilender Aussagen bewusst zu sein.

So einleuchtend dieser Versuch auf den ersten Blick scheint, wirft er doch Fragen auf. Wie beurteilt man eigentlich ein Werturteil und wie stellt sich in diesem Licht die Problematik der Vierten Hürde dar?

Ziel dieser Arbeit ist die Erarbeitung wissenschaftstheoretischer Kriterien und Methoden mit deren Hilfe obige Aussage analysiert und bewertet werden kann.

### ***Gang der Untersuchung***

In einem ersten Schritt wird das wissenschaftstheoretische Aussagesystem analysiert, um darauf aufbauend obige Aussage einordnen zu können. Anschließend werden Kriterien zur Beurteilung von Werturteilen identifiziert. Im Abgleich mit diesen Kriterien wird obige Aussage einer wissenschaftstheoretischen Analyse und Diskussion unterzogen. Zur Verdeutlichung dieser abstrakten Materie werden die einzelnen Kriterien jeweils an Beispielen unter anderem aus Politik, Wissenschaft und Gesundheitswesen erklärt. Anhand dieser Kriterien und Beispiele wird obige Aussage aus verschiedenen Gesichtspunkten beleuchtet und „wertend“ diskutiert.

### ***Wissenschaftstheoretisches Aussagesystem***

Zur Gliederung wissenschaftlicher Aussagen können zwei Dimensionen betrachtet werden. Zum einen ist zu prüfen, ob es sich um wahrheitsfähige oder um nicht wahrheitsfähige Aussagen handelt. Zum anderen, ob empirische oder nicht empirische Aussagen vorliegen (Raffée Grundprobleme S. 21 ff.)

Bei den wahrheitsfähigen Aussagen unterscheidet man zwischen Realdefinitionen, deskriptiven Aussagen und explikativen Aussagen.

---

<sup>2</sup> Die Pharmako-ökonomische Beurteilung bezieht sich auf die Bestimmung der Kosten-Effektivität eines Medikamentes. Sie wird als 4. Hürde beschrieben, sobald die Kosten-Effektivität in Verbindung mit der Sicherheit, Wirksamkeit sowie Qualität als explizites Kriterium für die Marktzulassung und weiterführend für die Kostenerstattungsfähigkeit eines neuen Medikamentes durch die Gesetzlichen Krankenkassen gebraucht wird (Lauterbach-Volmer).

Bei Definitionen handelt es sich um die Festlegung von Begriffsinhalt und Begriffsumfang (Seiffert S. 42).

Deskriptive Aussagen haben stets einen speziellen Raum-Zeit-Bezug und geben Antwort auf die Frage „was der Fall ist“. Beispiel: Die Bundesrepublik Deutschland hatte im Jahr 2002 80 Millionen Einwohner.

Die Prüfung des Wahrheitsgehaltes deskriptiver Aussagen kann z. B. durch statistische Tests erfolgen. Will man zum Beispiel wissen, ob die Patientengruppe A einen pathologisch erhöhten Laborwert gegenüber der Patientengruppe B aufweist, so lässt sich dies statistisch signifikant feststellen.

Explikative Aussagen lassen sich in Form von „Wenn-, Dann-Aussagen“ formulieren. Wenn einem Patienten bei einer schweren Sepsis das Medikament „X“ verabreicht wird, dann steigt seine Überlebenschance. Die Richtigkeit dieser Aussage kann durch eine kontrollierte Studie überprüft werden.

Realdefinitionen, deskriptive und explikative Aussagen sind wahrheitsfähig und empirisch. Rein logische Aussagen wie in der Mathematik sind zwar wahrheitsfähig, aber nicht empirisch.

Nicht wahrheitsfähige Aussagen sind metaphysischer oder normativer Art. Soll-Sätze stellen normative Aussagen dar und finden sich z.B. in Zielen oder Werturteilen wieder.

Werturteile liegen immer dann vor, wenn eine Aussage den jeweils anvisierten Sachverhalt in positiver oder negativer Weise für eine Stellungnahme oder für unser Handeln auszeichnet, dabei ein normatives Prinzip als gültig unterstellt, das ein entsprechendes Verhalten fordert und eine präskriptive Erwartung erfüllt (Albert S. 148).

Die eindeutige Einordnung der Aussage „dass Medikamente kosteneffektiv sein sollen“ in das oben genannte System ist zwar grundsätzlich durchführbar. Eine mögliche Ursache von bewusster oder unbewusster Beeinflussung liegt aber in versteckten Werturteilen, sogenannten krypto-normativen Aussagen. Die wissenschaftliche Beurteilung einer Aussage ist erst nach eindeutiger Einordnung möglich.

Betrachten zum Beispiel wir den Satz „Ein Arzneimittel, das 10 % mehr leistet, darf nicht 300 % mehr kosten“ näher. Die obige Aussage weist alle Merkmale eines Werturteils auf und ist damit als solches einzuordnen wie nachfolgend gezeigt wird.

Der Sachverhalt wird eindeutig in positiver bzw. negativer Weise anvisiert. Ein bestimmtes Arzneimittel darf „nicht“ mehr kosten. Es ist auch klar ersichtlich, dass vielfach Wertestandard unterstellt wird. Der hier diskutierte Wertestandard könnte „Angemessenheit“ sein. Allerdings könnte hier eine Wertediskussion in Gang gesetzt werden (was nicht Gegenstand dieser Arbeit ist). Man könnte als Norm auch die Akzeptierung des Marktmechanismus unterstellen. Im Produktverbund wären dann Rücklagen und weitere Mittel für Forschung und Innovation mit einzubeziehen. Dies wird hier offensichtlich nicht getan.

Erkennbar ist auch, dass man mit einer präskriptiven Erwartung, in diesem Falle von der Pharmaindustrie, rechnet und zwar in Form einer der Leistung des Medikamentes entsprechenden Preisfestlegung.

### ***Wie beurteilt man nun Werturteile?***

Zur Überbrückung der Kluft zwischen wahrheitsfähigen und nicht wahrheitsfähigen Aussagen benötigt man die sogenannten Brückenprinzipien (Raffée-Abel S.218 ff).

Das wohl bekannteste Brückenprinzip ist das „Praktikabilitätsprinzip“, das als „Sollen impliziert Können“ formuliert werden kann. Dieses Prinzip besagt, dass nur sinnvoll gefordert werden kann, was auch herstellbar ist. Zur Beantwortung dieser Frage können uns empirische, wahrheitsfähige Aussagen helfen.

Die Beantwortung der Frage, ob ein Medikament, das 100 % teurer ist auch 100 % mehr leistet, ist nicht so leicht zu beantworten. Ein geeignetes Instrumentarium dafür stellt die Kosten-Effektivitätsanalyse dar. Bei einer solchen Analyse wird das Verhältnis der in Geldeinheiten eingesetzten Mittel für eine Therapie dem Ergebnis gegenübergestellt. Dabei stellen sich unter anderem zwei Schwierigkeiten.

Zum einen ist häufig nicht klar, welche Referenztherapie und welcher damit verbundene Ressourcenverbrauch herangezogen werden soll. So herrscht etwa ein wissenschaftlicher Disput, ob allen Patienten ab einem gewissen Alter bzw. solchen mit leicht erhöhtem kardiovaskulärem Risiko obligat ein Thrombozytenaggregationshemmer verabreicht werden sollte. Solche Fragen müssen im Vorfeld geklärt sein, sonst schleichen sich krypto-normative Aussagen ein. Einen wertvollen Beitrag zu dieser Problematik können standardisierte Vorgehensweisen liefern, wie sie in der klinischen Forschung schon vorhanden sind. Damit wären Transparenz und Qualität der Analysen gewährleistet.

Angenommen das Problem der Standardtherapie sei geklärt. Dann müssen der Ressourcenverbrauch repräsentativ inklusive der klinische Endpunkte erfasst und ausgewertet werden. Die Durchführung eines solchen Projektes ist, und das zeigt die Einführung der DRGs in Deutschland, aufgrund der nicht vorhandenen Erhebungs- und Verarbeitungskapazitäten, ein kurzfristig nicht realisierbares Unternehmen.

Nehmen wir aber einmal an, die Frage der Standardtherapie und der benötigten Kapazitäten seien geklärt und man wäre in der Lage zu sagen, dass ein bestimmtes Medikament kosteneffektiv sei, dann stellt sich die Frage, ob diese Erkenntnisse in unserem Gesundheitssystem sinnvoll umgesetzt werden können.

Dies ist ernsthaft zu bezweifeln, wie im folgenden ein weiteres Beispiel für das Praktikabilitätsprinzip zeigt.

Zielsetzung der Berechnung der Kosteneffektivität ist die optimale Allokation knapper, finanzieller Ressourcen. Nach dem Prinzip „Sollen impliziert können“ müsste eine optimale Allokation durch Kosteneffektivität Erfolg zeigen. Folgendes Beispiel zeigt, dass dies in unserem sektoralen Gesundheitswesen nicht der Fall ist. Verschreibt nun ein Arzt ein innovatives, teures und kosteneffektives Medikament, dann läuft er Gefahr, wegen Budgetüberschreitung einem Regressverfahren unterzogen zu werden. Aufgrund eingesparter Krankenhaustage, bedingt durch die innovative Medikation, hat die Krankenkasse aber keinen Nachteil. Trotzdem wird der

behandelnde Arzt das Medikament restriktiv einsetzen. Der Patient trägt den Schaden. Aber nehmen wir einmal an, dieses Problem wäre gelöst, dann bliebe immer noch das Problem der Zeitdimension. Die gesetzliche Krankenversicherung ist durch Beiträge finanziert, die jährlich budgetiert werden. Rücklagen liegen nicht vor. Nun ist es bei Krankheiten wie Diabetes oder Osteoporose leider so, dass der finanzielle Nutzen einer Therapie erst Jahre oder Jahrzehnte später eintritt. Bei Diabetes handelt es sich um vermiedene Folgeerkrankungen wie Retinopathie, Nephropathie, Neuropathie, Schlaganfall oder Herz-Kreislaufferkrankungen. Bei der Osteoporose werden vertebrale Brüche und Brüche des Oberschenkelhalses vermieden. Der finanzielle Bedarf kosteneffektiver Medikamente kann in unserem Gesundheitssystem, das immer nur in einem Zeitrahmen von einem Jahr rechnet, aufgrund fehlender Rücklagen nicht gedeckt werden. Die Allokation, nicht nur im Querschnitt, sondern über die Zeit kann nicht geleistet werden. Die Kosteneffektivität kann zwar wertvolle Beiträge leisten, nur die Umsetzung ist bislang unmöglich.

Krypto-normative Aussagen können sich auch über die gewählte Perspektive einschleichen. Entscheidet man sich für die Perspektive der KV, des Krankenhauses, der Krankenkasse oder der Gesellschaft?

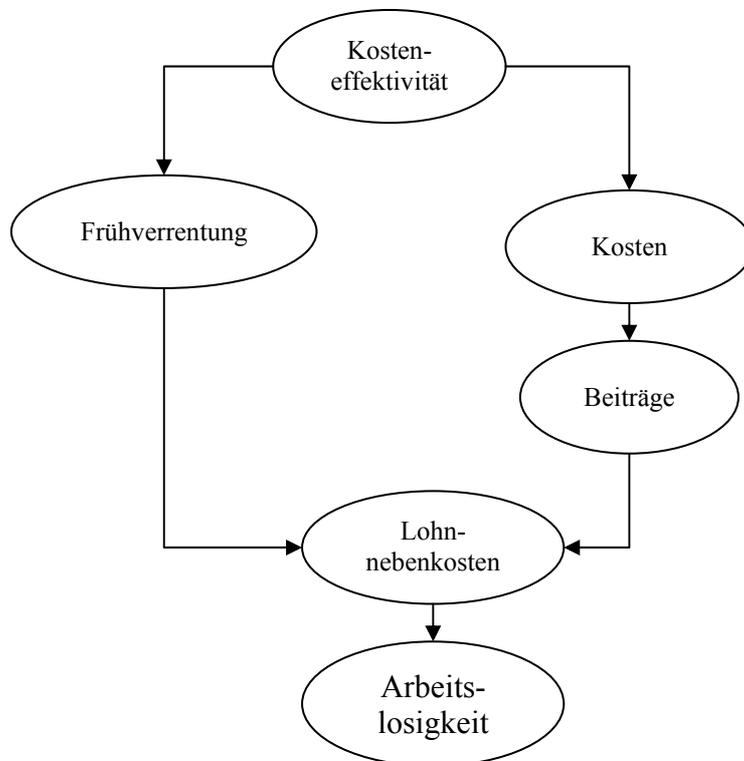
Das Beispiel der Schizophrenie zeigt, dass dieser Punkt nicht trivial ist. An Schizophrenie erkrankte Personen beanspruchen nicht nur Ressourcen der KV oder des Krankenhauses, sondern darüber hinaus des Sozialamts, des Arbeitsamts oder der Rentenversicherung. Die finanzielle Belastung bei der Schizophrenie ist im nichtärztlichen Komplementärangebot mit ca. 25.000 € für 5 Jahre extrem hoch (Deckert, S. 164). Die Frühverrentung mit 1.300.000.000 € pro Jahr, was mehr als das Fünffache des gesamten Marktes für atypische Neuroleptika ausmacht, übersteigt diesen Betrag noch bei weitem (Schnabel).

Transsektorale Betrachtungen sind bei der Berechnung der Kosteneffektivität vorab zu klären. Die Kosteneffektivitätsanalyse muss auf einem gesamtwirtschaftlichem Zielsystem aufbauen, das allerdings durch die sektorale Gliederung unseres Gesundheitssystems zur Zeit nicht möglich ist. Ein

Transfer über die Sektoren ist zur Zeit aber nicht möglich. Kosteneffektivität kann somit noch nicht einen Beitrag zu Allokation liefern.

Neben dem Praktikabilitätsprinzip kann zur Beurteilung von Werturteilen das Verknüpfungsprinzip herangezogen werden. Nach diesem Prinzip, sollen nur jene Normen befürwortet werden, deren Implikationen erkannt und akzeptiert werden.

Entscheider bewegen sich oft in komplexen, intransparenten, dynamischen und vernetzten Systemen, ohne das selbst zu bemerken (Dörner Logik S. 58 ff). Betrachten wir nun die Problemsituation der Kosteneffektivität. Ein Medikament, das 100 % mehr leistet und dementsprechend 100 % mehr kostet, verursacht einen zusätzlichen finanziellen Aufwand. Dies bedeutet mehr Geld in das System. Erfolgt dies über Beiträge, dann steigen die Lohnnebenkosten und damit möglicherweise die Arbeitslosigkeit und damit sinken wieder die Einnahmen für das Gesundheitssystem. Neben diesen direkten Effekten gibt es aber auch indirekte Effekte. So kann ein teures Medikament über die Vermeidung der Frühverrentung die Lohnnebenkosten senken.



Legende:

Die Pfeile deuten einen Zusammenhang an.  $A \rightarrow B$ .

A hat B zur Folge.

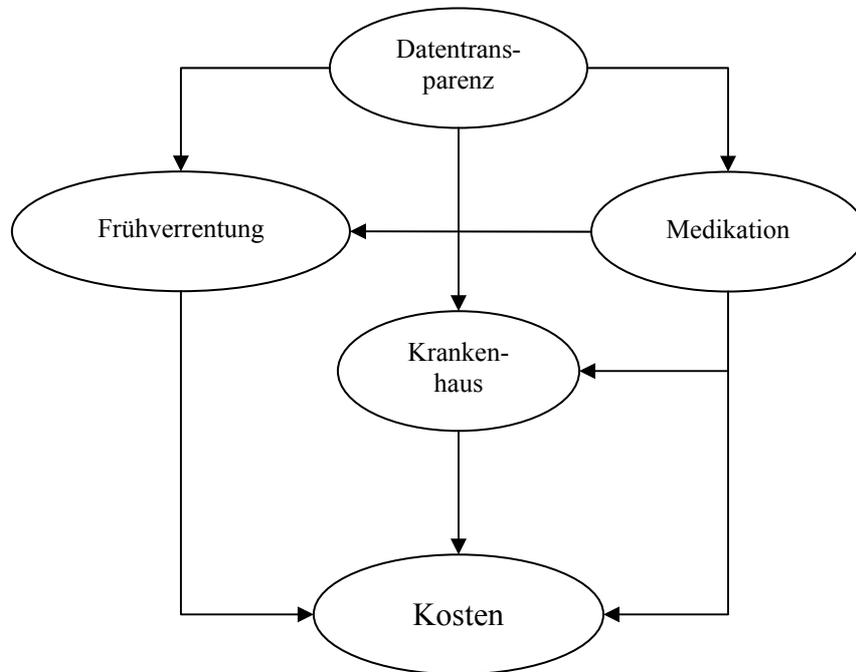
Menschen tun sich sehr schwer bei Entscheidungs- und Problemlösungsprozessen in komplexen, intransparenten, dynamischen und vernetzten Systemen (Dörner Lohhausen S. 24 ff). Die Frage, ob man Neben- und indirekte Effekte akzeptiert, setzt voraus, dass diese erfasst und quantifiziert werden. Markov-Modelle Strukturgleichungsmodelle, ökonometrische oder System Dynamics Modelle können hier Hilfestellung leisten (Deckert, Backhaus, Sterman). Gerade für das deutsche Gesundheitswesen liegen interessante Modellierungen vor (Neilson, Deckert).

Der Ansatz, Systemgrenzen für die Berechnung der Kosteneffizienz zu definieren und systemimmanente Beziehungen zu quantifizieren und zu diskutieren, ist bisher aber nicht zu erkennen. Damit wird eine Möglichkeit, Kosteneffizienz als Allokationsinstrument zu nutzen, nicht ausgeschöpft. Ansatzpunkte dazu bietet das nächste Brückenprinzip.

Das Prinzip der komparativen Beurteilung verlangt Normen und die durch sie ausgezeichneten Tatbestände im Lichte von Alternativen zu beurteilen. Anders ausgedrückt: „Wäre es nicht sinnvoller, weil wirkungsvoller, andere Wege einzuschlagen?“ Ein hypothetisches Beispiel soll dies verdeutlichen. Um Kosten einzusparen kann man sich die Medikamente, das Krankenhaus oder die Frühverrentung anschauen. Ein Ansatz kann darin bestehen, die Preise für Medikamente zu senken, die Verweildauer im Krankenhaus zu kürzen oder die Frühverrentung zurückzufahren. Nun kann es aber sein, dass Beziehungen zwischen diesen Variablen bestehen. Berücksichtigt man diese nicht, so kann es zu Fehlentscheidungen kommen.

Durch Datentransparenz, in Form eines Marktmechanismus oder durch Patientendaten könnten die Wechselwirkungen zwischen Medikation und Krankenhaus und Medikation und Frühverrentung erfasst und einem monokausalen Ansatz gegenübergestellt werden. Es ist durchaus denkbar, dass ein, die Wechselwirkungen umfassender Ansatz durch die eingesparten Krankenhaustage und Verminderung der Frühverrentung zu größeren Einsparungen kommt. In diesem Falle sollte man sich nicht auf die Einsparungen durch Medikamente beschränken. Selbst für den Fall, dass Kosteneffizienz zu Einsparungen im Arzneimittelmarkt führt, stellt sich

doch immer noch die Frage, ob die eigenen Ressourcen nicht für Alternativen wie z.B. Schaffung der Datentransparenz durch Studien oder ordnungspolitische Maßnahmen eingesetzt werden können, die einen höheren Wirkungsgrad beim Problemlösungsprozess versprechen.



Neben diesen drei Brückenprinzipien kann noch das sogenannte Kongruenz-Postulat formuliert werden. Mit diesem Prinzip können Normen dann kritisiert werden, wenn zu ihrer Befürwortung auf Erkenntnisse rekurriert werden muss, die sich als falsch erwiesen haben.

Durch die Etablierung einer „Vierten Hürde“ besteht nun die Gefahr einer Regulierung der Preise, was unweigerlich die Außer-Kraft-Setzung des Marktmechanismus bedeutet. Auch wenn es sich, wie oben erwähnt, um einen Markt „sui generis“ handelt, so sind wir doch der Meinung, dass Fehlallokationen durch Regulierungen Vorschub geleistet würde. Beispiele der Vergangenheit und auch der Gegenwart zeigen uns, dass damit Wohlfahrtsverluste einhergehen oder kontraintuitive Effekte erzielt wurden. Eindrucksvolle Beispiele sind der Milchsee und der Butterberg in der europäischen Gemeinschaft.

Am Beispiel der Parallelimporte wird deutlich, dass Regulierungen, sei es über den Preis oder über die Menge, zu kontraintuitiven Effekten führt. Die

Regierung schrieb eine feste Absatzmenge pro Apotheke vor. Dies bedeutete für die Reimporteure eine Absatzgarantie. Es war, da die Menge garantiert war, nur allzu verständlich, dass die Reimporteure die Preise fast auf Originalpräparateniveau erhöht hatten. Damit konnte der erwartete Einspar-effekt nicht realisiert werden.

Warum ist das so? Weil Eingriffe in das Markt-Preis-System die optimale Allokation zu verhindern drohten. Ein Instrument zur Optimierung der Allokation ist der Markt. Durch freien Tausch ist es möglich die Wohlfahrt zu optimieren. Man spricht auch von einem „pareto-effizienten“ Zustand (Varian S. 14). Die Theorie des „Reinen Tausches“ zeigt, dass durch freie Entscheidungen über Menge und Preis und nicht über Regulierungen die Wohlfahrt durch eine bestmögliche Allokation optimiert werden kann (Varian S. 491).

Preisfestsetzung erzeugt aber noch ein zusätzliches Problem, nämlich das des „Trittbrettfahrers“. Die Welt der forschenden Arzneimittelindustrie ist endlich und es gibt wenige Länder, die den forschenden Firmen die Möglichkeit geben, durch freie Marktwirtschaft ihre Gewinne zu optimieren. Die Konsequenz ist, dass andere Länder von der Innovation profitieren, aber selbst durch regulierte Preise keinen Beitrag dazu leisten. Dieses Verhalten kann aus Sicht eines Landes optimal sein. Es ist jedoch vom Standpunkt der Gesellschaft als Ganze nicht „pareto-effizient“ (Varian S. 611). Unter der Annahme, dass die Vierte Hürde mit Preiskontrollen verbunden wäre, ergäbe sich ein zusätzliches Problem. Betrachtet man die Zeit, die vergeht, bis ein Produkt nach Zulassung in einem Land eingeführt wird, dann fällt auf, dass Länder mit Preiskontrollen länger dafür benötigen. In Frankreich betrug die mittlere Zeit ca. 500 Tage wohingegen in den USA ca. 100 Tage verstrichen (Mullins, ISPOR 02). Was bedeutet dies aus dem Blickwinkel der Wohlfahrtsoptimierung? Die Amerikaner bezahlen einen hohen Preis, kommen dafür aber schnell in den Genuss guter Arzneimittel. Die Franzosen müssen erst das Preisniveau sondieren, was Zeit braucht, um dann i.d.R. niedrige Preise festsetzen zu können. Was bedeutet dies für Patienten? Betrachten wir einmal Arzneimittel zur Behandlung von an Schizophrenie er-

kranken Personen. Die Suicidrate bei an Schizophrenie erkrankten Menschen ist überdurchschnittlich hoch. Dies impliziert, dass in den USA zwar ein höherer Preis gezahlt wird, in Frankreich dafür aber mehr Menschen bei diesem Patientengut sterben.

Der Marktmechanismus kann wertvolle Beiträge zur Verbesserung der Allokation leisten. Allerdings stellt der Markt kein hinreichendes, sondern ein notwendiges Instrumentarium zur Allokationsoptimierung dar. Der Markt funktioniert aber nicht von selbst. Gründe für das Marktversagen liegen unter anderem im Anspruchsniveau (Simon) oder in der asymmetrischen Verteilung von Informationen (Stieglitz).

Eine asymmetrische Informationsverteilung läge vor, wenn z.B. nur die Anbieter das Preis-Leitungsverhältnis eines Medikamentes kennen würden.

Auch die Theorie der begrenzten Rationalität von Herbert Simon erklärt das Marktversagen. Danach wird ein Entscheidungsprozess abgebrochen, wenn ein befriedigendes Anspruchsniveau erreicht ist. Warum sollte ein Patient z.B. unzufrieden sein, wenn ein ihm verschriebenes Medikament seine Krankheit heilt und es ansonsten für ihn keine Effekte gibt? In der Summe führt ein solches Verhalten aber zu suboptimalen Entscheidungen.

Staatliche Eingriffe, die den Abbau einer Asymmetrie der Informationen und die Rationalität von Entscheidungsprozessen unterstützt, können den Marktmechanismus und damit verbunden die Allokation knapper Güter verbessern.

Es gibt Fälle, bei denen z.B. durch Asymmetrie der Informationsvergabe und durch ein gesättigtes Anspruchsniveau die Nachfrage eines Individuums nach einem Gut hinter dem gesellschaftlich gewünschten Ausmaßes zurückbleibt. Gesundheit ist ein solches Gut. Fördert der Staat dieses Gut z.B. durch Eingriffe wie Werbeverbot im Fernsehen für Rauchen, dann spricht man von demeritorischen Gütern (Musgrave).

Theater oder Weiterbildung können als meritorische Güter betrachtet werden.

Von besonderem Interesse sind meritorische Güter, die den Marktmechanismus unterstützen. Ein Beispiel stellt die Stiftung Warentest dar. Diese

Institution, die sich selbst aus eigenen Mitteln nicht tragen kann, stellt allen Marktbeteiligten das Preis-Leistungsverhältnis sowie einen Marktüberblick zur Verfügung.

Die Einrichtung einer sogenannten Vierten Hürde könnte, ähnlich wie die Stiftung Warentest, Informationsdefizite vermindern und so zu einer Stärkung des Marktmechanismus und damit der Allokation und letztendlich der Volksgesundheit beitragen.

Als Fazit lässt sich folgendes festhalten.

Bei der Einführung der Vierten Hürde würde es sich um ein Werturteil handeln.

Werturteile können anhand der Brückenprinzipien beurteilt werden.

Die Analyse brachte zutage, dass es nach dem Praktikabilitätsprinzip kaum möglich ist eine Vierte Hürde sinnvoll umzusetzen.

Das Verknüpfungsprinzip zeigt, dass die Effekte einer Vierten Hürde in ihrer Bandbreite noch nicht erkannt sind. Kosteneffektive Medikamente könnten bedeuten, dass mehr Geld für das System benötigt wird.

Das Prinzip der komparativen Beurteilung offenbarte, dass andere Aktivitäten, die zur Vorbereitung einer möglichen Einführung einer Vierten Hürde der Implementierung selbst vorzuziehen sind.

Das Kongruenz-Postulat fordert, dass nur auf solche Erkenntnisse zurückgegriffen werden darf, die sich als richtig erwiesen haben. Die Zerstörung des Marktmechanismus zieht aber Wohlfahrtsdefizite nach sich.

Die Verminderung der asymmetrischen Informationsverteilung und der begrenzten Rationalität durch eine Vierte Hürde könnte aber den Marktmechanismus stärken.

Die Einführung einer Vierten Hürde, würde sie als Preisregulierung eingesetzt, wiese darüber hinaus auf subtile Weise krypto-normative Elemente auf. Das Bild des mündigen Patienten würde ersetzt durch den unmündigen Menschen, dem man nichts zutraut und für den alles durch eine Staatsmedizin geregelt sein muss.

Schlussfolgernd kann festgestellt werden, dass die Optimierung der Allokation finanzieller Ressourcen im Gesundheitswesen durch kosteneffektive Medikamente grundsätzlich ein richtiger Ansatz darstellt.

Bevor dieses Ziel angestrebt werden kann, und das zeigen die Brückenprinzipien, müssen die dafür benötigten Instrumente und Bedingungen erfüllt werden.

Es liegt nahe, erst andere, wichtigere und grundlegendere Probleme zu lösen. Auf keinen Fall sollte die Vierte Hürde als Preiskontrolle installiert werden, vielmehr sollten Transparenz, Wettbewerb und Patienteneinbindung gefördert werden, wie dies in den Niederlanden angestrebt wird (Wever ISPOR 02).

Die Etablierung einer sogenannten Vierten Hürde ohne Preiskontrolle könnte allerdings Beiträge leisten, Asymmetrien der Informationen zu mindern und als meritorisches Gut, ähnlich wie Stiftung Warentest, den Marktmechanismus in diesem Markt „sui generis“ stärken.

## 4.2 Artikel

### **Kosten der Frühverrentung am Beispiel der Schizophrenie**

Johannes Clouth

Psychiat Prax 2004; 31, Supplement 2: S328-S245 © Georg Thieme Verlag KG Stuttgart,  
New York, DOI 10.1055/s-2004-828476, ISSN 1611-8332  
Johannes Clouth, Lilly Deutschland GmbH, Saalburgstraße 153, 61350 Bad Homburg

Die Schizophrenie ist eine schwere psychotische Krankheit, die relativ frühzeitig im Lebenszyklus auftritt und häufig die Arbeitsfähigkeit so nachhaltig herabsetzt, dass es zu einer Erwerbsunfähigkeit kommt. Die Krankheit verursacht somit nicht nur erhebliche direkte, sondern auch indirekte Kosten, die sich insbesondere in den Ausgaben der gesetzlichen Rentenversicherung (GRV) widerspiegeln. In dieser Studie werden Statistiken der GRV zur Frühverrentung ausgewertet. Jährlich werden in Deutschland weit über sechstausend Männer und Frauen mit der ersten Diagnose Schizophrenie (295 nach ICD-9) neu verrentet. Das durchschnittliche Verrentungsalter mit dieser Diagnose liegt für Männer bei 39 und für Frauen bei 42 Jahren. Bei den Männern erfolgen 14,7 Prozent aller Verrentungen der unter 40-Jährigen mit der Diagnose Schizophrenie. Entsprechend hoch sind die Ausgaben der GRV. Allein der Barwert der Rentenzahlungen für einen durchschnittlichen männlichen Rentenfall beträgt 215.000 €. Werden auch die Einnahmeausfälle der gesamten Sozialversicherung und der Einkommenssteuer berücksichtigt, ergibt sich ein Barwert von 560.000 €. Eine Simulation der Bestandszahlen auf Basis der Neuzugänge ergibt für das Jahr 2000 einen Rentenbestand von 125.000 Personen im Alter unter 65 Jahren in der GRV, die ursprünglich mit der ersten Diagnose Schizophrenie verrentet wurden. Die GRV wird hierdurch jährlich mit Rentenzahlungen in Höhe von etwa 1,3 Milliarden € belastet; hinzukommen Einnahmeausfälle der Sozialversicherung und des Fiskus von nochmals etwa 2 Milliarden €. Da in der GRV nur 2/3 der Personen im Erwerbsalter pflichtversichert sind, spiegeln diese Zahlen nicht die gesamten Kosten wieder.

### *Problemstellung, Zielsetzung und Methodik*

Die Problematik von rein kostenorientierten Sparmaßnahmen im Gesundheitswesen ist in den vergangenen Jahren zunehmend thematisiert worden. Die getrennte Budgetverantwortung der einzelnen Sektoren kann zu fehlerhaften Anreizen für die Allokation finanzieller Ressourcen führen. Eine rationale Gesundheitspolitik muss daher auch transsektorale Aspekte in Betracht ziehen (Greiner 2000, Über 2000). Dies gilt umso mehr, als heute in der wirtschaftspolitischen Diskussion nicht mehr die Beiträge zur GKV isoliert wahrgenommen werden, sondern die Summe aller Beiträge als Einheit gesehen wird.

Diese Arbeit rückt die Gesetzliche Rentenversicherung (GRV) in den Blickpunkt, die mit jährlichen Ausgaben von über 200 Mrd. € den weitaus größten Teilbereich der sozialen Sicherung darstellt. Ein erheblicher Teil dieser Ausgaben entsteht durch Renten wegen Erwerbsunfähigkeit. Am Beispiel der *Schizophrenie* wird demonstriert, welche Belastungen durch Frühverrentung im Bereich der GRV entstehen und welche Schäden durch eine falsche Weichenstellung ausgelöst werden können.

### *Krankheitsbild und Epidemiologie der Schizophrenie*

Die Schizophrenie ist eine schwere psychotische Krankheit. Generell kann man unter Schizophrenie eine Störung des Ich-Erlebnisse in der Form verstehen, dass Betroffene der Überzeugung sind, von anderen gelenkt oder beeinflusst zu werden. Es handelt sich um eine der schwerwiegendsten psychischen Erkrankungen, die zumeist zwischen der Pubertät und dem 30. Lebensjahr erstmals diagnostiziert wird. Die Symptome der schizophrenen Psychose sind vielschichtig und zeigen sich in inhaltlichem Wahn und formalen Denkstörungen, in Sinnestäuschungen, psychomotorischer Erregung, emotionaler Zurückgezogenheit, Antriebslosigkeit, Störungen des Affektes und motorischer Verlangsamung. Die Schizophrenie stellt nicht nur für die betroffenen Patienten, sondern auch für alle Familienangehörigen und Freunde, eine starke Belastung dar (Häfner 2000).

Die Schätzungen zu Inzidenz und Prävalenz der Schizophrenie sind mit großen Unsicherheiten behaftet. Nach Häfner (2000) tritt die Schizophrenie in allen Kulturkreisen und Bevölkerungsschichten mit ähnlich hoher Häufigkeit auf. Das Lebenszeitrisko an Schizophrenie zu erkranken beträgt nach mittelstrengen Konzepten etwa 1 Prozent der Bevölkerung ab dem fünfzehnten Lebensjahr (Lebenszeitprävalenz). Die bisher einzige repräsentative Studie zur Prävalenz der Schizophrenie in Deutschland von Besthorn et al. (1999) ergab, dass rund 136.000 Patienten mit der Diagnose Schizophrenie (ICD-10 F20) in Deutschland *fachärztlich* versorgt wurden.

Abbildung 1 zeigt, dass der weit überwiegende Teil der Patienten von einem niedergelassenen Psychiater oder in einer Institutsambulanz ambulant behandelt wird.

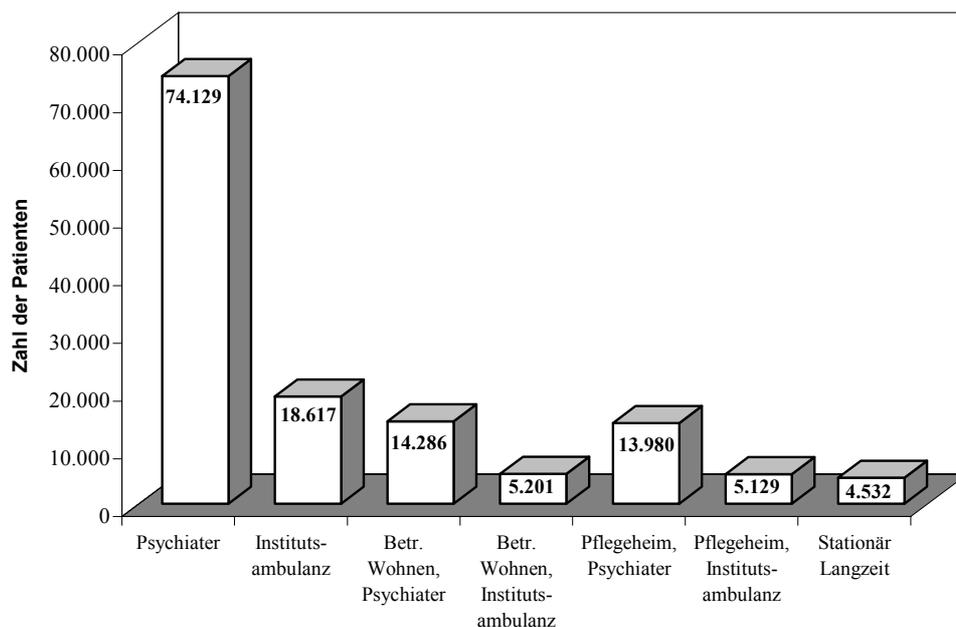


Abbildung 1: Verteilung der behandelten Patienten auf die betreuenden Institutionen. Quelle: Besthorn et al. 1999.

Nimmt die Krankheit einen Verlauf, der die Erwerbsfähigkeit erheblich und länger andauernd herabsetzt (vgl. VDR 2001, S.31-35), dann werden die Betroffenen in der Statistik der Rentenversicherung auftauchen, sofern sie die versicherungsrechtlichen Anspruchsvoraussetzungen erfüllen. Aus den

Statistiken der Rentenversicherung lässt sich demnach ebenfalls Aufschluss über die Bedeutung der Schizophrenie gewinnen.

#### *Direkte und indirekte Kosten der Schizophrenie*

In den Studien von v. d. Schulenburg et al. (1998) und Kissling et al. (1999) wurden die direkten und indirekten Kosten der Schizophrenie abgeschätzt. Hierfür wurden 180 Patienten in einer nach Versorgungsinstitutionen geschichteten Zufallsstichprobe selektiert und retrospektiv für einen Zeitraum von 12 Monaten in den Jahren 1995 und 1996 betrachtet. Eine Vergleichbarkeit der Kosten wurde durch eine Normierung auf einen einheitlichen Zeitraum von einem Jahr ermöglicht. Die geringsten direkten Kosten entstehen in der ambulanten Behandlung durch niedergelassene Psychiater und in der Institutsambulanz (2960 bzw. 3588 € pro Jahr und Patient). Im Vergleich hierzu sind die Kosten des betreuten Wohnens mit 21.735 €, der Langzeitpflege mit 22.725 €, des stationären Aufenthaltes mit 33.060 € und der Rehabilitation mit 40900 € deutlich höher (vgl. Abbildung 2).

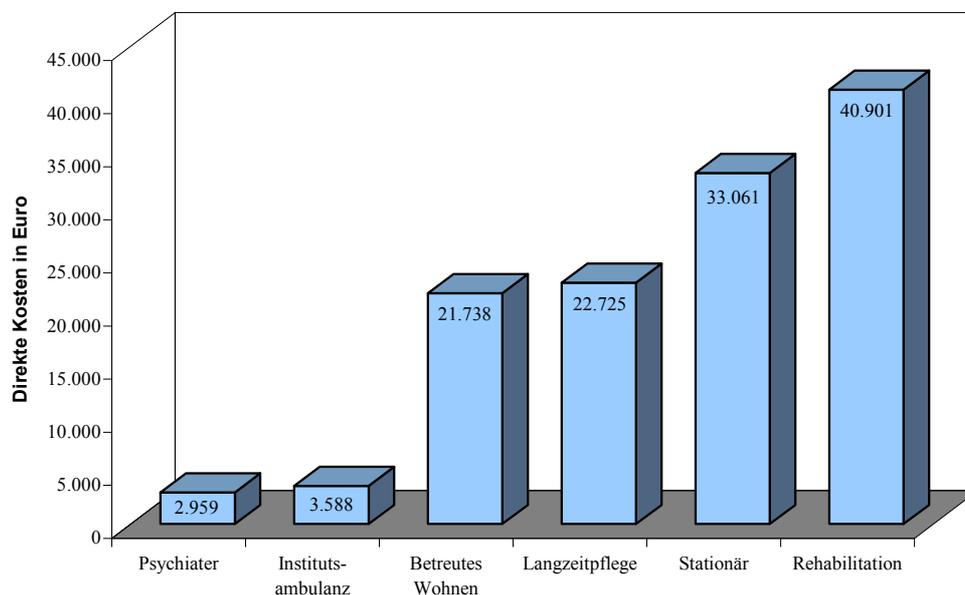


Abbildung 2: Direkte Kosten der Schizophrenie. Quelle: v.d. Schulenburg et al. 1998.

Die durchschnittlichen direkten Kosten pro Patienten lassen sich als *gewichtetes* Mittel der einzelnen Kosten bilden. Sie betragen nach v.d. Schulenburg

et al. (1998) 14.200 €, wobei allerdings das Gewicht der Rehabilitation mit 8 Prozent etwas hoch angesetzt wurde. Nach der Auswertung von Bestehorn et al. (1999) ist hingegen der relative Anteil der Rehabilitation wesentlich niedriger anzusetzen (Gewichtung mit der durchschnittlichen Verweildauer). Die durchschnittlichen jährlichen direkten Kosten von Patienten in fachärztlicher Versorgung beliefen sich damit auf 11.100 € im Jahr 1995.

Werden die Schätzungen zur Prävalenz der Schizophrenie von Bestehorn et al. (1999) zugrunde gelegt, dann ergeben sich für die 136.000 in fachärztlicher Behandlung befindlichen Patienten direkte jährliche Kosten in Höhe von rund 1,5 Mrd. € (ebenfalls für das Jahr 1995).

Nach Schätzung von Schulenburg et al. (1998, S. 84) betragen die auf ein Jahr bezogenen ungewichteten durchschnittlichen *indirekten* Kosten pro Patient knapp 22.000 €. Dabei entfallen 70% der indirekten Kosten auf die Folgen der Erwerbsunfähigkeit (entgangenes Arbeitsentgelt, soziale Ersatzleistungen und entgangene Abgaben).

## **Sektoren der Ressourcenbereitstellung**

### *Krankenkassen*

Die Krankenkassen tragen den größten Teil der *direkten* Kosten der Erkrankung, die durch Medikation, ambulante und stationäre Behandlung, sowie auch durch Ergotherapie und medizinische Rehabilitation entstehen. Darüber hinaus werden bis zu einem bestimmten Betrag die Kosten für häusliche Pflegedienste, Tagesstätten, Wohn- und Pflegeheime übernommen. Die Krankenkassen sind außerdem durch das Krankengeld, das an erwerbstätige Versicherte ab der 7. Woche einer Arbeitsunfähigkeit gezahlt wird, an den *indirekten* Kosten der Erkrankung beteiligt.<sup>3</sup> Dabei entfällt natürlich de facto auch die Zahlung der Versicherungsbeiträge. Die Krankenkassen müssen jedoch Lohnersatzleistungen nur für eine vorübergehende Arbeitsunfähigkeit aufbringen.

---

<sup>3</sup> In den ersten sechs Wochen einer Arbeitsunfähigkeit ist der Arbeitgeber zur Lohnfortzahlung einschließlich der Beiträge zur Sozialversicherung verpflichtet und trägt somit einen Teil der Kosten.

### *Rentenversicherung*

Stellt sich heraus, dass eine längerfristige Einschränkung der Arbeitsfähigkeit besteht, dann ist gegebenenfalls ein Rentenantrag zu stellen. Die Gesetzliche Rentenversicherung übernimmt, falls ein entsprechender Anspruch besteht, die Zahlung einer Erwerbsminderungsrente (Frührente) und/oder die Kosten von Maßnahmen zur Rehabilitation. Grundvoraussetzung für die Zahlung einer Rente ist die Erfüllung der allgemeinen Wartezeit (d.h. eine Mindestversicherungszeit von 60 Monaten). Die Zahlung einer Erwerbsminderungsrente setzt darüber hinaus eine Pflichtversicherung (als Beschäftigter oder Arbeitsloser) voraus; freiwillig Versicherte sind daher häufig vom Bezug einer Erwerbsminderungsrente ausgeschlossen. Die Rentenzahlungen der Gesetzlichen Rentenversicherung richten sich nach der Höhe der von dem Versicherten erworbenen Anwartschaften und unterstellen nach dem neuesten Reformgesetz eine Weiterarbeit bis zum 60. Lebensjahr (Zurechnungszeit). Die mittlere Rente wegen Erwerbsunfähigkeit betrug 816 € im Jahre 2000 (westdeutsche Männer). Beamte sind direkt durch ihren Dienstherrn abgesichert und erhalten eine in aller Regel großzügigere Pension, die sich nach ihrer Besoldungsstufe und ihrem Dienstalter bemisst.

### *Pflegeversicherung und Sozialhilfe*

Im Falle der Pflegebedürftigkeit übernimmt die *Pflegeversicherung* die Pflegekosten im Rahmen der Höchstsätze. Die Pflege kann ambulant (in privaten Haushalten oder Tagesstätten) oder stationär (in Wohn- und Pflegeheimen) erfolgen. Die Kosten der Unterbringung (die sogenannten „Hotelkosten“) sind von dem Pflegebedürftigen zu tragen. Bei stationärer Pflege übersteigen diese Kosten in der Regel die finanziellen Ressourcen des Kranken, so dass auch hier als letztes die *Sozialhilfe* einspringen muss („Hilfe in besonderen Lebenslagen“).

### *Allokation finanzieller Ressourcen*

Werden nun innovative Therapien eingesetzt, dann fallen finanzielle Belastungen in einem Teilbereich, beispielsweise den Medikamentenbudgets der

Krankenkassen, an und bewirken Einsparungen in anderen Bereichen. Diese Einsparungen können im Teilbereich der stationären Behandlung (vgl. Lichtenberg 2001); sie können aber auch in einem anderen Sektor entstehen, etwa in der Gesetzlichen Rentenversicherung durch eine Senkung der Frühverrentungsquoten. Dies führt zu einem klassischen Problem der Anreizverträglichkeit (incentive compatibility), denn die finanzielle Entlohnung folgt nicht der Leistung.<sup>4</sup> Allokationsprobleme ähnlicher Art werden in der Gesundheitsökonomik schon seit langem thematisiert. Hier tritt jedoch das Anreizproblem verstärkt auf, da dem Gesundheitswesen fremde Sektoren betroffen sind. Eine anreizkompatible Umschichtung von Mitteln über die Grenzen der Sektoren hinweg ist ungleich schwieriger als Modellversuche innerhalb des Gesundheitswesens.

#### *Frühverrentung in der Gesetzlichen Rentenversicherung*

In diesem Abschnitt wird die Frühverrentung mit der Diagnose „Schizophrenie“ in der Gesetzlichen Rentenversicherung dargestellt. Das Datenmaterial beruht in weiten Teilen auf einer Auswertung von Statistiken des Verbandes der Rentenversicherungsträger (VDR) zur Frühverrentung<sup>5</sup>. Es sind hierbei drei Anmerkungen zu beachten. Erstens werden durch die Frühverrentung nur solche Krankheitsbilder aktenkundig, die so schwerwiegend sind, dass sie zum Verlust der Erwerbsfähigkeit führen. Leichte oder erfolgreich behandelte Fälle werden damit nicht erfasst. Zweitens lassen sich aus dem verfügbaren statistischen Material nur die Zugangszahlen, nicht aber die Bestandszahlen ersehen. Drittens wird nicht die gesamte durch Schizophrenie ausgelöste Erwerbsunfähigkeit erfasst, da aus den oben dargelegten Gründen auch viele Fälle (Beamte, Selbständige, Hausfrauen und sonstige Nichterwerbstätige) unerfasst bleiben. Es handelt sich also um die Grundgesamtheit der Pflichtversicherten der GRV. Die Fallzahlen der

---

<sup>4</sup> Eine alternative Sichtweise auf dieses Problem ist die Theorie der externen Effekte. Das Gesundheitswesen produziert positive externe Effekte für das Budget der GRV, die aber nicht honoriert werden, was zu einer Unterversorgung führt. Die klassische Lösung dieses Problems besteht in der Internalisierung – beispielsweise durch eine Pigou-Steuer.

GRV lassen sich jedoch auf der Basis der bekannten Bevölkerungsanteile auf die Gesamtbevölkerung hochrechnen unter der Annahme, dass die Prävalenzen sich nicht unterscheiden. Dies ist eine vorsichtige Annahme, denn es muss davon ausgegangen werden, dass die Prävalenz in der Gruppe der Nichterwerbstätigen höher ist, weil diese Gruppe durch den Selektionsmechanismus eine positive Korrelation mit der Krankheit Schizophrenie aufweist: je früher die Krankheit ausbricht, desto unwahrscheinlicher ist es, dass der Kranke in der GRV abgesichert ist – und desto höher ist die Wahrscheinlichkeit, dass der Kranke auf Sozialhilfe angewiesen bleibt. Die Statistiken der Gesetzlichen Rentenversicherungen lassen sich damit als Ausgangspunkt verwenden, um eine untere Schranke für die Betroffenheit durch Schizophrenie abzuschätzen.

#### *Rentenzugang im Jahr 1999 mit Diagnose Schizophrenie*

Im Jahr 1999 wurden in Deutschland im Bereich der Gesetzlichen Rentenversicherung 3045 Frauen und 3603 Männer mit der ersten Diagnose „schizophrene Psychose“ (ICD-9: 295) neu verrentet (VDR, 1999, S.30 ff.).<sup>6</sup> Dies entspricht bei den Frauen 3,77 Prozent und bei den Männern 2,62 Prozent aller *neuen* Verrentungen des Jahres 1999 wegen verminderter Erwerbsfähigkeit („Frührenten“). Schizophrenie hat bei Frauen damit einerseits einen etwas höheren Anteil an den Erwerbsminderungsrenten als bei Männern, doch werden andererseits Frauen etwas später betroffen. Ein besonderes Gewicht erhält die Schizophrenie im Vergleich zu anderen Diagnosen dadurch, dass die erheblichen Erwerbsminderungen in einem vergleichsweise frühen Stadium des Lebenszyklus erfolgen. Die

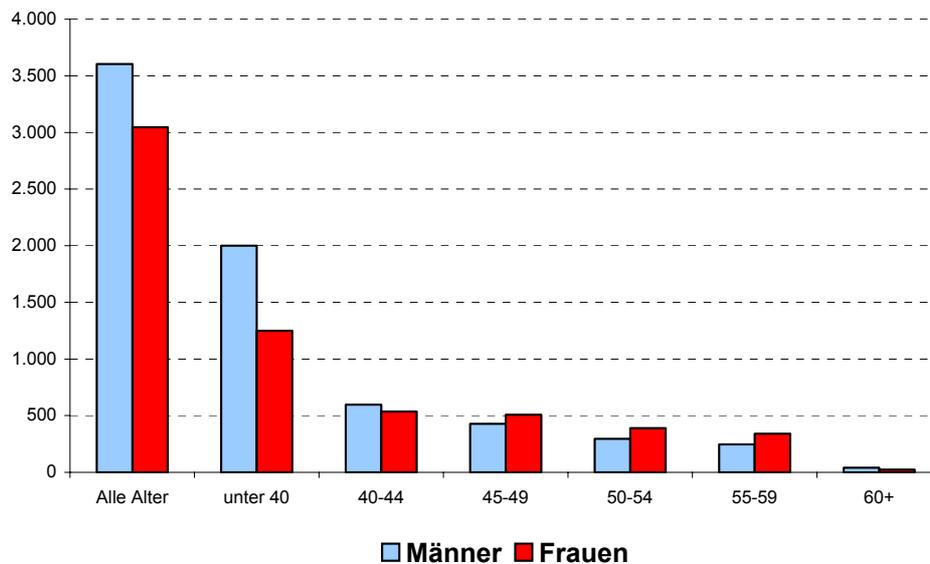
---

<sup>5</sup> Seit 1992 wird im Rentenrecht der Begriff „Renten wegen verminderter Erwerbsfähigkeit“ zur Abgrenzung von „Altersrenten“ verwendet. Der Begriff „Frührente“ wird in diesem Aufsatz der Einfachheit halber als Synonym weiterverwendet.

<sup>6</sup> Vgl. VDR-Statistik Rentenzugang, 1999. Bis zum Jahr 1999 einschließlich wurde die Klassifikation nach ICD-9 verwendet. Im Jahr 2000 erfolgte die Umstellung auf ICD-10, doch wurde leider ein großer Teil der Meldungen nach ICD-9 abgegeben. Daher werden im Folgenden nur Daten bis zum Jahr 1999 verwendet.

Altersverteilung der Frühverrentungen mit der Erst-Diagnose Schizophrenie wird in Abbildung 3 wiedergegeben.<sup>7</sup>

*Abbildung 3: Altersverteilung des Rentenzugangs mit der ersten Diagnose Schizophrenie – Männer und Frauen, Deutschland.*



*Quelle: VDR (Verband Deutscher Rentenversicherungsträger), VDR-Statistik Rentenzugang 1999, Rentenzugang wegen verminderter Erwerbsfähigkeit nach Diagnosen (1. Diagnose).*

Das Rentenzugangsalter von Männern mit Hauptdiagnose Schizophrenie lag bei nur 39,1 Jahren; dasjenige von Frauen liegt mit 42,3 Jahren im Schnitt um 3 Jahre höher (vgl. Abbildung 4). Schizophrenie ist damit sowohl bei Männern als auch bei Frauen *die* Erkrankung, die im Mittel die frühesten Verrentungen auslöst. Im Vergleich hierzu liegt das mittlere Zugangsalter für alle Erwerbsminderungsrenten bei 52,3 Jahren für Männer beziehungsweise bei 50,5 Jahren für Frauen. Die Regelaltersgrenze der Rentenversicherung liegt bei 65 Jahren.

<sup>7</sup> Es ist davon auszugehen, dass Schizophrenie aufgrund von Komorbidität an weiteren Fällen von Erwerbsunfähigkeitsrenten beteiligt ist. Über die Komorbidität gibt es jedoch bisher kein publiziertes Datenmaterial, da standardmäßig nur die „erste“ Diagnose statistisch ausgewertet wird. Außerdem können sich Abgrenzungsprobleme zu ähnlichen Diagnosen ergeben. Die Fallzahlen des Zugangs in eine Erwerbsminderungsrente in den Diagnosegruppen 295-299 betragen 6106 Männer und 6351 Frauen.

Abbildung 4: Durchschnittliches Rentenzugangsalter – Hauptdiagnose Schizophrenie und alle Diagnosen.

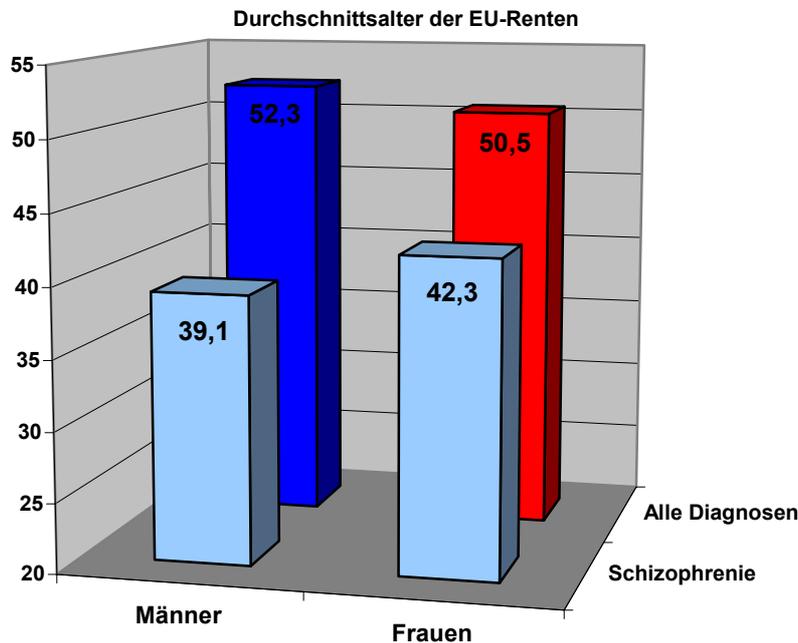


Abbildung 5: Anteil der Schizophrenie an der Frühverrentung nach Alter und Geschlecht. Quelle: VDR-Statistik Rentenzugang, verschiedene Jahrgänge.



### Schizophrenie – bedeutendste Ursache für Erwerbsunfähigkeit junger Menschen

Die Bedeutung der Schizophrenie für die Frühverrentung wird unterstrichen durch die Betrachtung der jüngeren Altersgruppen (vgl. Abbildung 4). Schi-

zophrenie ist in der Altersgruppe der unter 40-Jährigen die mit großem Abstand *wichtigste einzelne* Ursache für Erwerbsunfähigkeit, sowohl bei Männern (2000 neue Rentenfälle im Jahr 1999) als auch bei Frauen (1249 neue Rentenfälle im Jahr 1999). Damit sind bei den Männern der Altersgruppe unter 40 Jahren 14,7 Prozent aller Frühverrentungen auf Schizophrenie zurückzuführen. In der Altersgruppe der 40- bis 45-jährigen beträgt der Anteil der Schizophrenie an den Frühverrentungen immer noch 6,0 Prozent. In den oberen Altersgruppen geht die relative Bedeutung dieser Diagnose zurück. Bei den Frauen unter 40 Jahren beträgt der Anteil 12,1 Prozent. In der Altersgruppe der 40- bis 45-jährigen beträgt der Anteil an allen Frühverrentungen immer noch 7,23 Prozent.

#### *Budgetwirkungen und Kosten der Frühverrentung mit Diagnose Schizophrenie*

Ein Schizophreniekranker, der im durchschnittlichen Alter von 39,1 Jahren verrentet wird, bezieht vor dem Erreichen der Regelaltersgrenze für die Dauer von 25,9 Jahren eine Frührente (d.h. vor Vollendung des 65. Lebensjahres). Entsprechend hoch ist die Belastung des Budgets der Gesetzlichen Rentenversicherung. Der durchschnittliche Frührentner, der im Alter von 52,3 Jahren in Rente geht, bezieht hingegen „nur“ für die Dauer von 12,7 Jahren eine Frührente. Ein an Schizophrenie Erkrankter erzeugt damit bei gleichem monatlichem Rentenanspruch rechnerisch eine doppelt so hohe finanzielle Belastung bis zum Alter 65 wie der durchschnittliche Frührentner, vorausgesetzt, er erreicht das Alter von 65.

Weitere Einnahmeausfälle entstehen in den anderen Zweigen der Sozialversicherung und bei der Einkommensteuer. Die Einnahmeausfälle im Bereich der gesamten Sozialversicherung und der Einkommensteuer addieren sich auf etwa 17.100 € im Jahr und übersteigen damit die direkten Rentenzahlungen beträchtlich. Für die soziale Sicherung und die Steuereinnahmen addieren sich die Gesamtkosten aus Einnahmeverlusten und Rentenzahlungen auf 27.600 € pro Jahr. Bezogen auf eine Frührente vom Alter 39 bis

zum Alter 65 summieren sich mithin die Gesamtkosten auf über 560 Tausend € im Barwert. Diese Ausgaben müssen letztendlich durch die Versicherten, die Unternehmen und die Steuerzahler finanziert werden.<sup>8</sup>

#### *Ursachen der Frühverrentung im Zeitablauf 1983-1999*

Allgemein lässt sich feststellen, dass die Bedeutung von psychischen Erkrankungen als Ursache der Frühverrentung im Verlauf der letzten 15 Jahre dramatisch zugenommen hat. Die absoluten Fallzahlen haben sich hier nahezu verdoppelt, während gleichzeitig die absolute Fallzahl bei den übrigen (somatischen) Diagnosen zurückgegangen ist. Folglich ist die relative Bedeutung der psychischen Erkrankungen dramatisch angestiegen. Der Anteil psychischer Erkrankungen an allen neuen Erwerbsunfähigkeitsrenten hat sich bei den Männern von 8% auf 19% mehr als verdoppelt (vgl. Abbildung 5). Bei den Frauen (Abbildung 6) hat sich der Anteil sogar von 9% auf 32% mehr als verdreifacht, so dass auf der Ebene der Diagnose-Hauptgruppen bei den Frauen die psychischen Erkrankungen seit einigen Jahren den ersten Platz unter den Ursachen für Erwerbsunfähigkeit einnehmen – eine Entwicklung, die übrigens nicht auf die Bundesrepublik Deutschland beschränkt ist, sondern beispielsweise auch in den Vereinigten Staaten zu beobachten ist.

---

<sup>8</sup> Ein noch umfassenderes Maß für die volkswirtschaftlichen Kosten ist das entgangene Bruttoarbeitsentgelt einschließlich der Arbeitgeberbeiträge. Wird das durchschnittliche Arbeitsentgelt der sozialversicherungspflichtig Beschäftigten des Jahres 2000 zugrunde gelegt, dann beträgt der volkswirtschaftliche Verlust über 33.750 € pro Jahr.

Abbildung 6: Rentenzugang von Männer nach Diagnosehauptgruppen 1983-1999. Bundesrepublik Deutschland ohne Neue Bundesländer.

Quelle: VDR, Rentenversicherung in Zeitreihen, 2001.

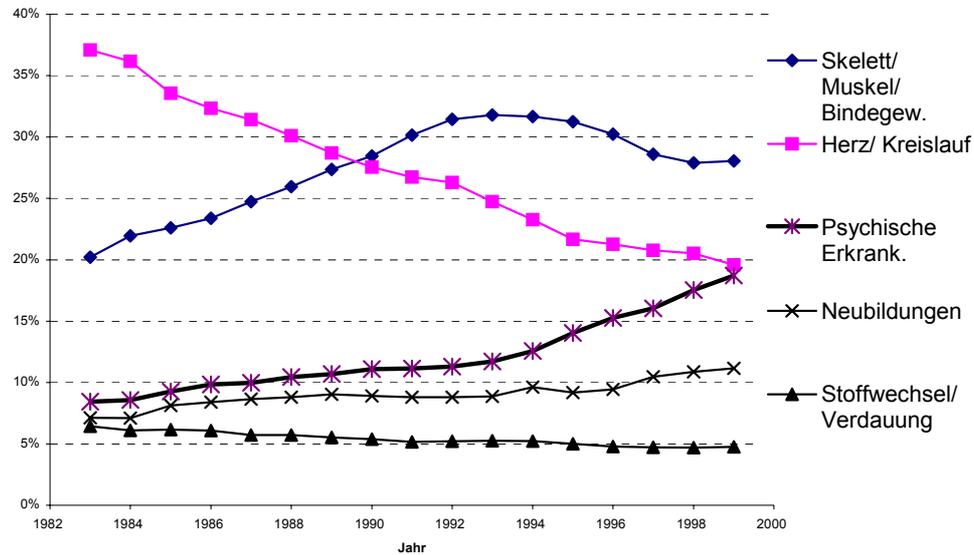
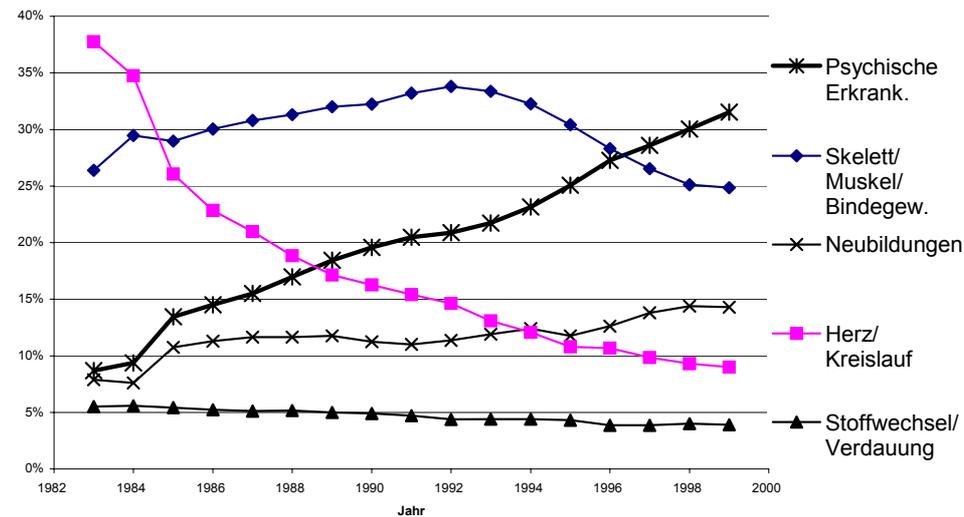


Abbildung 7: Rentenzugang von Frauen nach Diagnosehauptgruppen 1983-1999. Bundesrepublik Deutschland ohne Neue Bundesländer.

Quelle: VDR, Rentenversicherung in Zeitreihen, 2001.



Frühverrentung wegen Schizophrenie 1994-1999

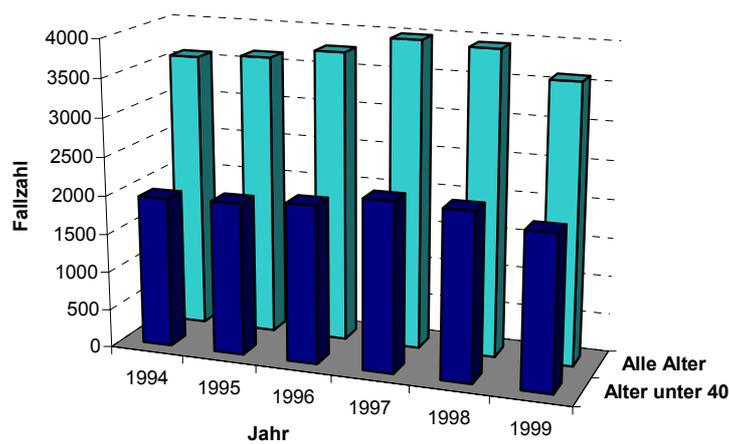
Die absolute Fallzahl der Frührenten mit der ersten Diagnose Schizophrenie (295 nach ICD-9) ist von 1994 bis zu den Jahren 1997/98 gestiegen, doch

sind in jüngster Zeit die Fallzahlen wieder rückläufig (vgl. Abbildungen 8 und 9). Der Rückgang bei den Frauen ist hierbei ausgeprägter als bei den Männern. Leider ist das Jahr wegen der Umstellungsprobleme auf ICD-10 (noch) nicht auswertbar.

Der *relative Anteil* der Schizophrenie an den Frühverrentungen ist jedoch von 1994 bis 1999 gestiegen, weil die Zahl der Frühverrentungen insgesamt deutlich abgenommen hat. Wurden im Jahr 1994 noch 12,4 Prozent aller Frühverrentungen von Männern der Altersgruppe „unter 40“ durch Schizophrenie bedingt, so waren es im Jahr 1999 bereits 14,7 Prozent. Für die Frauen betragen die entsprechenden Anteile 10,4 Prozent und 12,8 Prozent.

Abbildung 8: Frühverrentung mit erster Diagnose Schizophrenie – Männer.

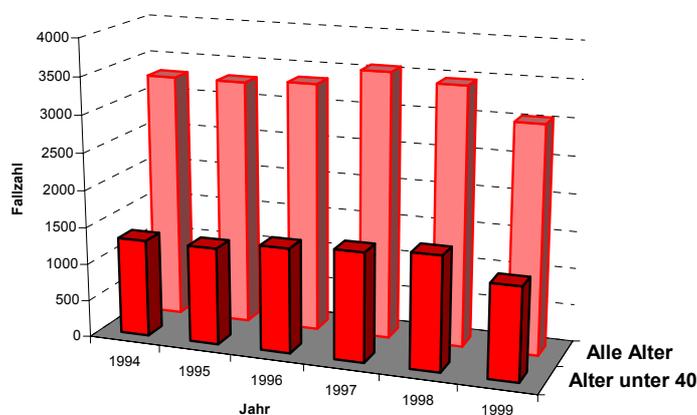
Quelle: VDR Statistik Rentenzugang, verschiedene Jahrgänge.



	1994	1995	1996	1997	1998	1999
■ Alter unter 40	1952	1981	2050	2200	2172	2000
■ Alle Alter	3591	3645	3773	3987	3939	3603

## Abbildung 9: Frühverrentung mit erster Diagnose Schizophrenie – Frauen.

Quelle: VDR Statistik Rentenzugang, verschiedene Jahrgänge.



	1994	1995	1996	1997	1998	1999
Alter unter 40	1310	1307	1413	1467	1534	1249
Alle Alter	3301	3299	3346	3570	3464	3045

*Simulation der Bestandszahlen auf Basis der Zugangszahlen*

Eine Abschätzung der jährlichen Ausgaben der Rentenversicherung erfordert eine Abschätzung der Zahl der *Bestandsrentner*, die wegen Schizophrenie in Rente gegangen sind. Da derzeit Angaben über den Rentenbestand aufgeschlüsselt nach Diagnosen nicht verfügbar sind, müssen die Zahlen aus dem *Rentenzugang* hochgerechnet werden. Aus den jährlichen Neuzugängen mit der Diagnose Schizophrenie wurden zunächst die altersspezifischen Hazardraten bestimmt als Anteil der Personen einer Altersgruppe, die mit dieser Diagnose verrentet wurden. Mit diesen Hazardraten und der Bevölkerungsstruktur der Jahre 1952 bis 1993 wurden sodann die Zugänge in den vergangenen Perioden geschätzt. Die Kumulation der Zugänge ergibt schließlich unter Berücksichtigung der Sterblichkeit eine Abschätzung des Bestands an Rentnern. Hierbei werden verschiedene Annahmen gemacht. Zum einen wird unterstellt, dass die Verrentungen irreversibel sind; zum zweiten wird eine zeitliche Stabilität der Hazardraten angenommen; zum dritten muss eine Annahme über die Sterblichkeit erfolgen.

Wird eine um das dreifache erhöhte Sterblichkeit gegenüber der allgemeinen Sterbetafel unterstellt, dann ergibt sich ein *Bestand* von 125 Tausend Personen im Alter *unter 65 Jahren*, die im Jahr 1999 eine GRV-Rente bezo-

gen haben und die ursprünglich mit der ersten Diagnose Schizophrenie frühverrentet worden waren.<sup>9</sup> Berücksichtigt man alle Altersgruppen, dann ergibt sich ein Bestand von 160.000 Rentnern. Diese Zahlen umfassen möglicherweise auch Personen, bei denen sich die Krankheit gebessert hat und die weiterhin eine Frührente bezogen haben. Daher darf die Zahl nicht als Jahresprävalenz der Krankheit Schizophrenie interpretiert werden. Sie gibt vielmehr die Zahl der Personen an, die zu einem (früheren) Zeitpunkt ihres Lebens so stark durch die Krankheit beeinträchtigt worden sind, dass sie verrentet wurden. Andererseits werden über die Frühverrentung weder die minder schweren Fälle noch die Fälle mit günstiger Entwicklung erfasst. Die durch Schizophrenie verursachten Frührenten führen hochgerechnet in der GRV zu Ausgaben in Höhe von etwa 1,3 Milliarden € pro Jahr.<sup>10</sup> Die gesamte Summe aus Frührenten sowie aus Einnahmenausfällen der Sozialversicherung und des Fiskus sind mit 3,5 Mrd. € pro Jahr deutlich höher anzusetzen. Diese Zahlen liegen auch deutlich höher als die Schätzungen von Schulenburg et al. (1998).

#### *Fazit und Ausblick*

Die Zahl der an Schizophrenie erkrankten Personen wird aus Gründen der demographischen Entwicklung weiterhin hoch bleiben, bis die „Babyboom-Kohorten“ über das Alter von 40 Jahren hinaus sind. So erreicht beispielsweise die Zahl der 40-jährigen im Jahr 2004 ein Maximum und wird erst im Jahr 2010 wieder auf das Niveau des Jahres 1998 zurückfallen. Frühverrentung wegen Schizophrenie wird demnach *ceteris paribus* aus rein demographischen Gründen weiterhin ein großes finanzielles Problem bleiben. Eine Verbesserung der Behandlung könnte auch außerhalb der (gesetzlichen) Krankenversicherung zu einer Entlastung der Budgets beitragen. Der Stand der Frühverrentung beim Einsatz atypischer Neuroleptika war bei einer naturalistischen Beobachtungsstudie Studie ca. 20% geringer. (Eich-

---

<sup>9</sup> Eine Sensitivitätsanalyse zeigt, dass selbst bei einer 10-fach erhöhten Sterblichkeit immer noch mit 99.000 Bestandsrentnern zu rechnen ist.

<sup>10</sup> Hinzu kommen noch pensionierte Beamte, die in der Regel deutlich höhere Renten beziehen.

mann 2004). Es sollte nach Wegen gesucht werden, diese transsektoralen Entlastungen in die Entscheidungsbildung einzubeziehen, um die gesamte Effizienz der Sozialen Sicherung zu steigern. In diesem Beitrag wurde insbesondere das Problem der Frühverrentung in der Gesetzlichen Rentenversicherung betrachtet. Angesichts relativ geringer Raten der Rehabilitationsmaßnahmen bei psychisch Kranken sind auch die Möglichkeiten verstärkter Rehabilitation zu prüfen. Doch auch in der residualen Sozialhilfe entstehen hohe Ausgaben durch Behandlung, Sicherung des Lebensunterhalts und „Hilfe in besonderen Lebenslagen“ (Pflege und Heimunterbringung), die in diesem Beitrag nicht betrachtet wurden. Auch in diesem Bereich könnten verbesserte Behandlung und Rehabilitation zu einer Senkung der Kosten führen.

### 4.3 Artikel

## **Kostenanalyse der Schizophrenie-Behandlung in Deutschland**

### **Vergleich von Olanzapin, Risperidon und Haloperidol mit Hilfe eines klinischen Entscheidungsmodells**

C. Deckert<sup>1</sup>, J. Höffler<sup>2</sup>, J. Kortmann<sup>3</sup>, M. Linden<sup>4</sup>, G.-D. Roth<sup>5</sup>, M. Struck<sup>6</sup>,  
J. Clouth<sup>7</sup>, J. Czekalla<sup>7</sup>, A. Spannheimer<sup>8</sup>, B. Wieseler<sup>8</sup>, T. Schlereth<sup>8</sup>

<sup>1</sup> Moorweidenstraße 12, 20148 Hamburg

<sup>2</sup> Westfälische Klinik Gütersloh

<sup>3</sup> Landeskrankenhaus Osnabrück

<sup>4</sup> Forschungsgruppe Ambulante Therapie an der Psychiatrischen Klinik und Poliklinik der  
FU, Berlin

<sup>5</sup> Hindenburgstraße 35, 73760 Ostfildern-Nellingen

<sup>6</sup> Westfälisches Zentrum für Psychiatrie, Paderborn

<sup>7</sup> Lilly Deutschland GmbH, Bad Homburg

<sup>8</sup> Kendle GmbH & Co. GMI KG, München

Gesundh ökon Qual manag 2001; 6: 161-166, © Georg Thieme Verlag Stuttgart – New  
York, ISSN 1432-2625

*Zielsetzung:* Ziel der Studie war eine pharmako-ökonomische Analyse der  
Behandlung von Schizophrenie-Patienten mit Olanzapin, Risperidon und  
Haloperidol in Deutschland.

*Methodik:* Die Analyse wurde mit Hilfe eines klinischen Entscheidungs-  
modells durchgeführt, das von einem internationalen Expertenpanel entwickelt  
wurde. Das Modell einschließlich der klinischen Daten (u.a. aus multinatio-  
nalen klinischen Studien) über den Verlauf der Erkrankung wurde über-  
nommen. Daten zur Praxis bei der medizinischen Behandlung von  
Schizophrenie sowie Kostendaten für Deutschland wurden in einem natio-  
nalen Expertenpanel ermittelt und in das Modell integriert. Anschließend  
wurden eine Kosten-Analyse und eine Kosten-Effektivitäts-Analyse durch-  
geführt.

*Ergebnisse:* Die Gesamtkosten der Therapie lagen bei allen 3 Therapien in der gleichen Größenordnung (Risperidon DM 123.004, Haloperidol DM 121.868, Olanzapin DM 119.572, jeweils pro Patient für 5 Jahre). Die Kosteneffektivität von Risperidon bzw. Olanzapin war um den Faktor 2,3 bzw. 2,5 besser als die von Haloperidol.

*Schlussfolgerungen:* Bei Betrachtung der Gesamtkosten werden die hohen Tagestherapiekosten der atypischen Neuroleptika durch Einsparungen bei der Behandlung von Negativsymptomatik und Rückfällen ausgeglichen.

### ***Einleitung***

Schizophrenie ist eine Erkrankung, die bis zu 1 % der Bevölkerung betrifft [16]. Die unbehandelte Erkrankung führt zu einer erheblichen Einschränkung der Lebensqualität der Patienten und ihres familiären Umfelds [11]. Die Behandlung der Erkrankung ist mit beträchtlichen Kosten verbunden. Da die Erkrankung in der Regel bei jungen Erwachsenen diagnostiziert wird, ist eine lang andauernde Behandlung notwendig. Die Patienten benötigen zumeist eine intensive medizinische Betreuung; es kommt bei Rückfällen in einen akuten Schub einer schizophrenen Psychose häufig zu Krankenhausaufenthalten. Darüber hinaus treten hohe indirekte Krankheitskosten, z.B. durch Arbeitsunfähigkeit, Arbeitslosigkeit oder Erwerbsunfähigkeit auf. Die genauen Kosten der Schizophrenie lassen sich nur schwer ermitteln, verschiedene Studien haben einen Anteil von 1,6% bis 2,5% der Gesamtausgaben für das Gesundheitswesen in Industrieländern ergeben [5]. Für die Behandlung schizophrener Psychosen stehen seit einigen Jahren neben den konventionellen Neuroleptika vom Haloperidol-Typ die sogenannten "atypischen Neuroleptika" (z.B. Olanzapin und Risperidon) zur Verfügung. "Atypische Neuroleptika" haben ein teilweise anderes Wirkprofil als die konventionellen Neuroleptika. Einige klinische Studien zeigten bei vergleichbarer oder besserer Effektivität eine geringere Rate von unerwünschten Ereignissen der atypischen Neuroleptika, insbesondere im Hinblick auf sogenannte extrapyramidal-motorische Nebenwirkungen [5]. Zudem zeigten atypische Neuroleptika einen günstigen Einfluß auf die Ne-

gativsymptomatik. Die Verminderung der extrapyramidal-motorischen Symptome führt darüber hinaus zu einer verbesserten Compliance der Patienten, die die Therapie unterstützt [18].

Durch die Bestrebungen zur Kostendämpfung im Gesundheitswesen werden die Kosten-Effektivitäts-Analysen von Therapiealternativen immer wichtiger. Betrachtet man im Fall der Therapie mit atypischen Neuroleptika bzw. Standardneuroleptika nur den Preis für die Medikation selbst, so ist die Therapie mit einem konventionellen Neuroleptikum sehr viel preiswerter als die Therapie mit einem atypischen Neuroleptikum. Die Gesamtkosten einer Therapie werden jedoch außer durch den Preis des Medikaments durch viele weitere Faktoren wie z.B. durch die Kosten für die Behandlung von Langzeitfolgen oder Nebenwirkungen beeinflusst. Um einen Vergleich der Wirtschaftlichkeit verschiedener Behandlungsalternativen durchführen zu können, sollten in einer pharmako-ökonomischen Evaluation möglichst alle relevanten und messbaren Kostenfaktoren einbezogen werden.

Eine Möglichkeit der pharmako-ökonomischen Analyse ist die Errechnung der Gesamtkosten von Therapiealternativen mit Hilfe von Krankheitsmodellen [14]. Diese Modelle bilden die möglichen Krankheitsverläufe ab. In das Modell werden die Wahrscheinlichkeiten, mit der jeder Krankheitsverlauf auftritt, sowie die Ressourcen, die jeder Verlauf der Erkrankung verbraucht, eingesetzt. Die Wahrscheinlichkeiten, die im Modell verwendet werden, basieren auf wissenschaftlichen Daten kontrollierter Studien zur Wirksamkeit und Verträglichkeit der Alternativtherapien. Daten zum Ressourcenverbrauch können durch ein Expertenvotum ermittelt werden. Mit Hilfe des Modells lassen sich dann die Gesamtkosten der untersuchten Therapien für hypothetische Patientenpopulationen über verschiedene Zeiträume errechnen. Die Verwendung von Krankheitsmodellen ist gerade bei der Betrachtung langer Zeiträume mit wiederkehrenden Krankheitszyklen, die häufig von prospektiven Studien nicht abgedeckt werden können, von Bedeutung. In der vorliegenden Evaluation wurde eine pharmako-ökonomische Analyse der Behandlung schizophrener Psychosen mit dem konventionellen Neuro-

leptikum Haloperidol und den atypischen Neuroleptika Olanzapin und Risperidon über einen Zeitraum von 5 Jahren vorgenommen.

### ***Methodik***

In dieser pharmako-ökonomischen Studie wurde ein Krankheitsmodell für schizophrene Psychosen verwendet, das durch ein internationales Expertenpanel entwickelt wurde und bereits in den USA und Großbritannien angewendet wurde [1, 4, 12, 13]. In dem Modell wurden klinische Daten (z.B. die Wahrscheinlichkeit für einen Rückfall oder das Auftreten von Positiv- und Negativsymptomatik) aus internationalen, randomisierten, doppelblinden klinischen Studien mit Olanzapin, Risperidon und Haloperidol verwendet [13, 15, 17]. Weitere klinische Daten (z.B. die Basis für die Hochrechnung der Rückfallrate für 5 Jahre [6]) wurden von dem internationalen Expertenpanel mit Hilfe der Literatur ermittelt. In das Modell wurden folgende Dosierung für die Neuroleptika eingesetzt: 10 mg/Tag für Olanzapin, 6 mg/Tag für Risperidon und 15 mg/Tag für Haloperidol. Für die hier vorliegende Studie wurde das Modell einschließlich der klinischen Daten übernommen. Die Daten zum Ressourcenverbrauch und zu den Kosten in Deutschland wurden dagegen mit Hilfe eines nationalen Expertenpanels erhoben und dann in das Modell eingesetzt. Damit ermöglicht das Krankheitsmodell die Abschätzung der Gesamtkosten der Behandlung der Schizophrenie mit Olanzapin, Risperidon und Haloperidol in Deutschland.

### ***Das Krankheitsmodell***

Das in dieser Studie verwendete Krankheitsmodell (Abbildung 1, erstellt von MEDTAP International) simuliert den Verlauf der Erkrankung für ein hypothetisches Patientenkollektiv bei Behandlung mit Olanzapin, Risperidon oder Haloperidol über einen Zeitraum von 5 Jahren. Der Zeitraum von 5 Jahren wurde gewählt, um Langzeitkosten, die z.B. durch Rückfälle verursacht werden, zu erfassen. Das Modell basiert auf einem Markov-Prozess, in dem Zyklen von jeweils 3 Monaten Dauer iterativ 20 mal durchlaufen werden. Im Initialzyklus wird die Entscheidung für die Behandlung mit einer

der 3 Therapiealternativen getroffen. Da bei Schizophrenie-Patienten die Wahrscheinlichkeit eines Suizids sehr hoch ist (jährliche Selbstmordrate: 8% der Patienten mit Schizophrenie [2]), wurde die Möglichkeit eines Suizids in das Modell mit aufgenommen. Patienten, die keinen vollendeten Suizid begehen, erreichen einen von 4 möglichen Symptomzuständen (Positiv-/ Negativsymptomatik, Negativsymptomatik, Positivsymptomatik oder ohne Positiv-/Negativsymptomatik). Aus jedem dieser Symptomzustände ist mit einer bestimmten Wahrscheinlichkeit der Rückfall in einen akuten Schub der Psychose möglich. Ein Rückfall kann wiederum zum Suizid führen. Im Fall, daß der Patient keinen Suizid begeht, kann sich wieder einer der 4 möglichen Symptomzustände einstellen. Auch für Patienten, die keinen Rückfall erleiden, ist von einem Zyklus zum anderen ein Übergang zwischen den Symptomzuständen möglich. Abbildung 1 zeigt den Entscheidungsbaum des Modells.

Im Gegensatz zu anderen Anwendungen des Modells [1, 13] wurde für die Berechnungen, die in der vorliegenden Arbeit präsentiert werden, davon ausgegangen, daß ein Patient über 5 Jahre das gleiche Medikament einnimmt. Ein Wechsel der Medikation oder ein Abbruch der Behandlung wurde also nicht angenommen.

### ***Datenquellen und Datenerhebung***

Die Wahrscheinlichkeiten, mit der die einzelnen Ereignisse des Entscheidungsbaums eintreten, waren Bestandteil des bereits vorliegenden Modells [1, 13].

Informationen zum Ressourcenverbrauch in Deutschland in den einzelnen Zyklen des Modells wurden von einem nationalen Expertenpanel zusammengestellt, das aus insgesamt 6 Psychiatern bzw. Psychologen aus der Klinik sowie aus dem niedergelassenen Bereich zusammensetzt war (4 Experten aus der Klinik und 2 Experten aus dem niedergelassenen Bereich; jeweils 1 Einzelinterview sowie 2 Plenumsdiskussionen aller Experten). Jedem Zyklus wurde ein Ressourcenverbrauch zugeordnet, dabei war für die verschiedenen Typen von Zyklen (z.B. Zyklus mit Negativsymptomatik) der

Ressourcenverbrauch unter den 3 Therapien im Wesentlichen gleich. Unterschiede ergaben sich nur durch die Medikation selbst sowie durch die Behandlung von extrapyramidalen Symptomen. Tabelle 1 zeigt exemplarisch den Ressourcenverbrauch für den Zyklus "Rückfall". Die Mengen der eingesetzten Ressourcen wurden anschließend mit Preisen bewertet. Die Kosten bzw. Quellen für die Kosten, mit denen die eingesetzten Ressourcen belegt wurden, sind in Tabelle 2 zusammengefasst.

### ***Pharmako-ökonomische Analyse***

Die Kostendaten wurden zusätzlich zu den klinischen Daten in das Modell integriert (Tree Age DATATM Software). Anschließend wurden eine Kosten-Analyse und eine Kosten-Effektivitäts-Analyse aus volkswirtschaftlicher Perspektive durchgeführt. Dabei wurde eine 5 %ige Diskontierung vorgenommen.

In der Kosten-Analyse werden nur die Faktoren berücksichtigt, die direkt in Kosten messbar sind, mit dem Ziel, die bezüglich der Gesamtkosten günstigste Alternative herauszufinden.

Die Kosten-Effektivitäts-Analyse berücksichtigt die Kosten in monetären Einheiten, die Effektivität der Therapie in nichtmonetären Einheiten. Ergebnis dieser Analyse sind nach Effektivität gewichtete Kosten. Als Effektivitätsparameter für diese Untersuchung wurde der Anteil der Patienten ohne Rückfall in eine akute Psychose ausgewählt. Die Kosten-Effektivitäts-Analyse errechnet also die Kosten, die innerhalb von 5 Jahren pro Patient ohne Rückfall auftreten. Entsprechend pharmako-ökonomischer Richtlinien wurde auch der Effektivitätsparameter (Anteil der Patienten ohne Rückfall) diskontiert (5%).

### ***Ergebnisse***

In einem ersten Schritt wurden die Gesamtkosten der drei Therapiealternativen für den Zeitraum von 5 Jahren errechnet (Kosten-Analyse). Die Gesamtkosten ergeben sich - außer aus den Kosten für Olanzapin, Risperidon und Haloperidol selbst - aus der Wahrscheinlichkeit für die verschiedenen

Verläufe [1,13] und deren Kosten pro Zyklus und Patient (Tabelle 3). Die Symptomatik, die im Verlauf der 5 Jahre unter den drei Behandlungen auftritt, bestimmt die Inanspruchnahme von Kostenfaktoren, so werden bei Rückfall oder Negativsymptomatik deutlich mehr Ressourcen gebraucht als bei Positivsymptomatik.

Die Ergebnisse der Kosten-Analyse werden in Tabelle 4 zusammengefasst. Die drei Therapiealternativen führten zu vergleichbaren Gesamtkosten (Risperidon DM 123.004, Haloperidol DM 121.868, Olanzapin DM 119.572, jeweils pro Patient für 5 Jahre). Tabelle 4 stellt auch die Beträge, die von den einzelnen Kostenfaktoren verursacht werden, dar. In diese Darstellung gehen die Wahrscheinlichkeiten für die Inanspruchnahme der Kostenfaktoren innerhalb von 5 Jahren sowie deren Kosten ein. So sind z.B. die Kosten für einen Tag Krankenhausaufenthalt unter allen drei Therapien gleich, weil es unter Haloperidol aber häufiger zu Hospitalisierungen kommt (höhere Wahrscheinlichkeit für Rückfall, [1,13]), wird für diesen Kostenfaktor bei Behandlung mit Haloperidol ein höherer Betrag aufgewendet als bei Behandlung mit Olanzapin oder Risperidon.

Wird die Medikation allein betrachtet, so ist Haloperidol um den Faktor 7 günstiger als Olanzapin und Risperidon. Dieser Kostenvorteil wird aber durch ungünstigere Symptomatik und höhere Rückfallquote unter Haloperidol [1, 13], die zu vermehrter Inanspruchnahme von Leistungen führt, mehr als ausgeglichen. Insbesondere verursacht Haloperidol höhere Kosten in den Bereichen stationäre Behandlung, Rehabilitation und nichtärztliches Komplementärangebot.

Zusätzlich zur Kosten-Analyse wurde eine Kosten-Effektivitäts-Analyse durchgeführt, die die Gesamtkosten in Relation zur Effektivität der Therapie setzt. Als Maß für die Effektivität der drei Therapiealternativen wurde der Anteil an Patienten ohne Rückfall in eine akute Psychose gewählt (siehe Tabelle 5). Die Effektivität der Behandlung mit Olanzapin war am höchsten,

mehr als ein Drittel der Patienten erlitten innerhalb von 5 Jahren keinen Rückfall. Ähnlich effektiv war Risperidon, Haloperidol dagegen zeigte eine wesentlich geringere Rate an Patienten ohne Rückfall als die beiden Vergleichspräparate. Die stark unterschiedliche Effektivität führt zu großen Differenzen in der Kosten-Effektivität, obwohl die Gesamtkosten der drei Therapiemöglichkeiten annähernd gleich sind. So muß bei einer fünfjährigen Behandlung mit Haloperidol im Vergleich zu Risperidon ein um den Faktor 2,3 höherer Betrag pro Patient ohne Rückfall aufgebracht werden. Im Vergleich zu Olanzapin hat Haloperidol sogar eine um den Faktor 2,5 schlechtere Kosten-Effektivität (Tabelle 5).

### ***Diskussion***

Die vorliegende pharmako-ökonomische Analyse wurde auf der Basis eines Krankheitsmodells durchgeführt [1, 4, 12]. Durch die Integration von länderspezifischen Behandlungsmustern, Ressourcenverbräuchen und Kostendaten kann das Modell an nationale Gegebenheiten angepaßt werden. Außer in den USA und in Deutschland wurde das Modell auch in Großbritannien, den Niederlanden und Spanien eingesetzt [1, 4, 12].

Die Analyse hat gezeigt, daß die Gesamtkosten der Therapie mit den atypischen Neuroleptika mit den Gesamtkosten der Behandlung mit konventionellen Neuroleptika vergleichbar sind. Dieses Ergebnis wurde auch in anderen pharmako-ökonomischen Studien erzielt [1, 3, 5, 10, 12]. In allen Untersuchungen wurden die höheren Tagestherapiekosten atypischer Neuroleptika durch Einsparungen ausgeglichen, die durch bessere Effektivität bzw. verringerte Nebenwirkungsraten möglich waren.

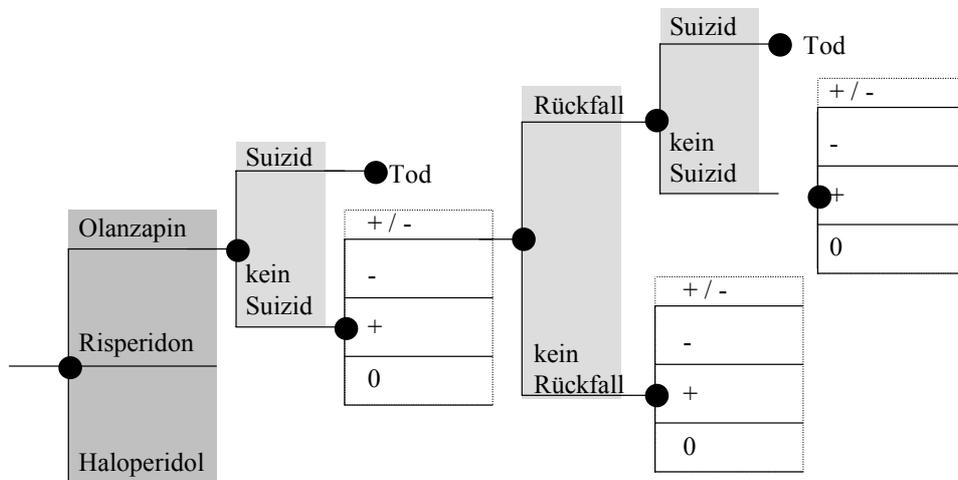
Bei der Behandlung schizophrener Psychosen führt besonders das Auftreten von Negativsymptomatik und Rückfällen zu hohen Kosten (Krankenhausaufenthalte, Rehabilitation) [7]. Die Wirksamkeit der atypischen Neuroleptika bei der Behandlung der Negativsymptomatik wirkt sich deshalb stark kostensenkend aus. Die bessere Effektivität der atypischen Neuroleptika

führte bei vergleichbaren Gesamtkosten zu einer besseren Kosten-Effektivität. Die Konsistenz der Ergebnisse in verschiedenen Studien unterstreicht die Glaubwürdigkeit der Resultate [1, 3, 5, 10, 12, 13].

Die Ergebnisse der Studie zeigen, daß es aus volkswirtschaftlicher Perspektive keinen Grund gibt, den Patienten atypische Neuroleptika aus wirtschaftlichen Gründen vorzuenthalten. Eine Berücksichtigung dieser pharmako-ökonomischen Gesamtrechnung wird jedoch dadurch erschwert, daß die Ärzte, die die Therapie auswählen, aufgrund von Arzneimittelbudgets häufig zur Auswahl eines Medikamentes mit den niedrigeren Tagestherapiekosten tendieren. Eine Möglichkeit, dem Arzt die Entscheidung für die verträglichere und letztlich kostengünstigere Therapiealternative zu erleichtern, wären krankheitsbezogene kombinierte Budgets für ambulante und stationäre Behandlung.

In der vorliegenden Analyse waren die Kosten pro Zyklus mit einer bestimmten Symptomatik unter allen Therapien bis auf die Medikationskosten gleich. Damit spiegelt das Modell die Praxis nicht vollständig wider. Möglicherweise werden weitere Kostenfaktoren innerhalb einer Symptomatik durch die Auswahl der Therapie beeinflusst (z.B. Dauer des Krankenhausaufenthalts). In einer prospektiven Studie sollen deshalb Daten zum Ressourcenverbrauch unter Olanzapin, Risperidon und Haloperidol erhoben werden, um die Datenbasis für die Berechnungen zu verbessern.

Abbildung 1: Entscheidungsbaum des Modells



- + / - : Positiv- /Negativ-Symptomatik
- : Negativ-Symptomatik
- + : Positiv-Symptomatik
- 0 : Ohne Positiv-/Negativ-Symptomatik

Die 4 Symptomzustände wurden gemäß dem PANSS (Positive and Negative Symptom Scale) definiert [8]

**Tabelle 1:**  
**Exemplarischer Ressourcenverbrauch für den Zyklus "Rückfall"**

<b>Ressourcen</b>	<b>Anteil der Patienten</b>	<b>Menge</b>
Hospitalisierung (Psychiatrie)	60%	70 Tage
Teilstationäre Tagesklinik	5%	21 Tage
Institutsambulanz	35%	6 Visiten
Arztkonsultationen (ambulante Nachbehandlung im niedergelassenen Bereich)	60% 5% 35%	4,5 Visiten nach Hospitalisierung 15 Visiten nach teilstat. Behandlung 15 Visiten nach Institutsambulanz
Arbeitsunfähigkeit	9% 6%	10 Wochen 3 Wochen
Medikation	100% 3 Behandlungsgruppen	Olanzapin 91,25 Tage Risperidon 91,25 Tage Haloperidol 91,25 Tage
Labortests	100%	1 Blutbild 1 Lebertest
Behandlung der Extrapyramidalen Symptomatik	19% 20% 45%	Olanzapin-Gruppe Risperidon-Gruppe Haloperidol-Gruppe
Ergotherapie	10%	20 Behandlungen
Medizinische Rehabilitation	5%	91,25 Tage
Berufliche Rehabilitation	5%	91,25 Tage

**Tabelle 2:**  
**In der Analyse verwendete Kostendaten und Datenquellen**

<b>Ressource</b>	<b>Kostenquelle / Betrag</b>
Medikation	Rote Liste (1998)
Arztkonsultationen / Institutsambulanzen	EBM / GOÄ (1997)
Krankenhaus (psychiatrische Abteilung)	AOK (1997) Durchschnittlicher Pflegesatz: 384,57 DM/Tag
Teilstationäre Tagesklinik	Expertenpanel Durchschnittlicher Pflegesatz: 250 DM/Tag
Intensivstation	Deutsche Krankenhausgesellschaft Durchschnittlicher Pflegesatz: 2.088,62 DM/Tag
Labortests	EBM / GOÄ (1997)
Medizinische Rehabilitation	Expertenpanel Durchschnittlicher Tagessatz: 200 DM/Tag
Wohnheim	Expertenpanel Durchschnittlicher Tagessatz: 150 DM/Tag
Tagesstätten (5 Tage/Woche)	Expertenpanel Durchschnittlicher Tagessatz: 100 DM/Tag
Betreutes Wohnen	Expertenpanel Durchschnittlicher Tagessatz: 88 DM/Tag
Werkstattarbeitsplatz (5 Tage/Woche)	Expertenpanel Durchschnittlicher Tagessatz: 70 DM/Tag
Werkstatt Taschengeld	Expertenpanel Durchschnitt: 160 DM/Monat
Arbeitsunfähigkeit	Statistisches Bundesamt (1997) Durchschnittliches Gehalt: 137 DM/Tag, bis zur 7. Woche: 226 DM/Tag, ab 7. Woche: 110 DM/Tag
Berufliche Rehabilitation	Expertenpanel Durchschnittlicher Tagessatz: 160 DM/Tag
Notarztwagen	Bayerisches Rotes Kreuz Kosten pro Einsatz: 800 DM
Ergotherapie	Expertenpanel Kosten pro Behandlung: 33 DM
Ambulante Pflegedienste	Expertenpanel Kosten pro Besuch: 30 DM
Justiz/öffentliche Behörden	Expertenpanel, Vormundschaftsgericht München Einmaliges Betreuungsgutachten: 600 DM Inanspruchnahme eines Betreuers: 3500 DM/Jahr

**Tabelle 3: Kosten pro Zyklus pro Patient<sup>1</sup> bei Auftreten der verschiedenen Symptome**

	<b>Kosten pro Zyklus pro Patient [DM]</b>		
	<b>Olanzapin</b>	<b>Risperidon</b>	<b>Haloperidol</b>
Initialzyklus	17.372	17.437	16.819
Negativsymptomatik	10.251	10.359	9.284
Positiv-/Negativsymptomatik	10.251	10.359	9.284
Positivsymptomatik	1.456	1.570	432
Ohne	1.456	1.570	432
Positiv-/ Negativsymptomatik			
Rückfall	20.090	20.150	19.562
Suizid	46.383	46.383	46.383

<sup>1</sup>Die Kosten in Deutschland wurden durch ein nationales Expertenpanel ermittelt (siehe Tabellen 1 und 2)

**Tabelle 4: Kosten-Analyse  
(direkte und indirekte Kosten sowie Gesamtkosten für 5 Jahre)**

	<b>Durchschnittliche Kosten pro Patient für 5 Jahre [DM]</b>		
	<b>Olanzapin</b>	<b>Risperidon</b>	<b>Haloperidol</b>
Medikation	21.020	22.970	3.356
Arztkonsultationen	3.625	3.642	3.635
Behandlung extrapyramidaler Symptome	193	203	445
Laboruntersuchungen	549	550	534
Stationärer Bereich	32.236	32.871	37.797
Ergotherapie	601	610	733
Ambulanter Pflegedienst	721	729	881
Rehabilitation	9.642	9.778	11.752
Nichtärztliches Komplementärangebot	48.035	48.654	59.477
Arbeitsunfähigkeit	2.509	2.554	2.800
Justiz/öffentliche Behörden	295	295	295
Versuchter Suizid	145	148	163
Vollendeter Suizid	1	1	1
<b>Gesamtkosten</b>	<b>119.572</b>	<b>123.004</b>	<b>121.868</b>

**Tabelle 5: Kosten-Effektivitäts-Analyse**  
**(Effektivitätskriterium: Patienten ohne Rückfall in die akute Psychose)**

	<b>Olanzapin</b>	<b>Risperidon</b>	<b>Haloperidol</b>
Effektivität [% Patienten ohne Rückfall] <sup>1</sup>	36,2	33,0	14,5
Kosten-Effektivität [DM für 5 Jahre pro Patient ohne Rückfall]	330.309	372.739	840.469

<sup>1</sup>Die Daten wurden aus dem internationalen Modell entnommen [13]. In diesem Modell wurde die Rückfallrate für 5 Jahre hochgerechnet. Die Hochrechnung im Modell basierte auf Erfahrungswerten, die von Experten mit Hilfe von Literaturdaten [6] ermittelt wurden.

#### 4.4 Artikel

### **Constructing an Index for the Subjective Well-being Under Neuroleptics Scale (SWN), Short Form: Applying Structural Equation Modeling for Testing Reliability and Validity of the Index**

Peter Schmidt<sup>1</sup>, Johannes Clouth<sup>2</sup>, Lars Haggemüller<sup>1</sup>, Dieter Naber<sup>3</sup>,  
Ursula Reitberger<sup>4</sup>

<sup>1</sup>Justus-Liebig-Universität Gießen Karl-Glöckner Str. 21 E 35394 Gießen

<sup>2</sup>Lilly Deutschland GmbH, Saalburgstraße 153, 63150 Bad Homburg

<sup>3</sup>Universitäts-Krankenhaus Eppendorf (UKE), Martinistraße 52, 20246 Hamburg

<sup>4</sup>Kendle GmbH, München

Quality of Life Research (2006) 15: 1191-1202

#### Abstract

Structural equation modeling (SEM) has been widely used in psychology and sociology for testing validity of measurement instruments. However, this statistical technique has so far played minor role in quality-of-life research. The main objective of this paper is to demonstrate the potential of SEM for constructing and testing the validity of a Subjective Well-being under Neuroleptics (SWN) index for patients with schizophrenia. For these purposes, data from the GEO study (*Gesundheitsökonomische Evaluation von Olanzapin in Deutschland*; health economics study of olanzapine in the treatment of schizophrenia in Germany) were used. The GEO is a prospective, comparative, noninterventional, observational study. Six hundred and forty-six participants treated with either olanzapine (N=416) or haloperidol (N=230) were enrolled in the study; 360 patients were available for factor analyses. The short (20-item) form of the SWN scale was administered to assess patients' perspectives on their quality of life. The structural equation models (SEMs) were then applied to construct 5- and 10-item indexes based on SWN. The data indicate that the 5-item index is the most time-saving approach for evaluating perceptions of well-being (and thus, quality of life)

among patients with schizophrenia. The application of SEM showed no appreciable loss of validity of this index.

## Introduction

Structural equation modeling and quality of life (particularly, quality of life of patients with schizophrenia) are complex issues. Health economists need screening instruments and index instruments to analyze cost-effectiveness of different treatment options to improve quality of life of patients with schizophrenia.

Structural equation modeling is a statistical technique that plays a major role in psychology and sociology. This technique is commonly used for testing the factorial validity of a measurement instrument or for testing the validity of a hypothesized causal structure [1]. In quality-of-life research, however, structural equation modeling has so far played a minor role. Further options, such as using structural equation modeling for index construction, have not been widely investigated. Research of publications indexed in MEDLINE and EMBASE found that structural equation models (SEMs) have rarely been used for both validity testing and index construction [2-4].

Schizophrenia is a disease that places a high mental, social, and economic burden on those afflicted, their relatives, and society. The lifetime risk for acquiring schizophrenia is approximately 1% [5-8]. The yearly incidence of newly diagnosed schizophrenia is 150 per 100,000 inhabitants of Western industrialized nations [5]. Until recently, the causes, manifestations, and the course of schizophrenia have not been fully investigated. Frequently, the disease begins with marked changes in behavior; sudden appearance of psychotic symptoms such as hallucinations or delusions characterizes the acute phase of schizophrenia. Such acute phases of schizophrenia can occur only once in the patient's lifetime or can return chronically. In some patients, the length of time between the acute phases eventually becomes shorter and shorter. In addition to the 'positive' symptoms (changes in a person's perceptions, thoughts, feelings, and behavior), patients with schizophrenia frequently experience 'negative' symptoms such as social withdrawal,

depression, and lack of motivation. Schizophrenia is a chronic illness; antipsychotic drug therapy can fully control symptoms of schizophrenia in only ~20% of patients. In most of these cases, however, symptoms of schizophrenia can be treated by antipsychotics.

Symptom reduction alone was the most essential outcome parameter in treatment of schizophrenia for a long time. With the development of atypical antipsychotics, more ambitious success criteria (including patients' perspectives regarding their quality of life) are considered today. In the last 10 years, the attitude of practicing psychiatrists changed toward assessing the ability of patients with schizophrenia to validly express their emotions and evaluate their quality of life [9]. In 2001, the short (20-item) form of the Subjective Well-being under Neuroleptics (SWN) scale was developed by Naber and colleagues [10]. This self-report scale is a valid instrument for measuring subjective well-being in humans suffering from schizophrenia [11].

The objectives of this study were as follows:

1. To demonstrate the potential of structural equation modeling for validity testing and index construction in quality-of-life research.
2. To develop an abbreviated assessment instrument (5- and 10-item index) based on the SWN (short form) using structural equation models.
3. To test the internal construct validity of the new instrument (SWN index).

## Methods

To construct the index for the SWN scale and prove its validity using structural equation modeling, the baseline data from the GEO study (*Gesundheitsökonomische Evaluation von Olanzapin in Deutschland* [health economics study of olanzapine in the treatment of schizophrenia in Germany]) were used.

## Patient and Study Characteristics

GEO is a prospective, comparative, noninterventional, observational, multi-centric health economics study of patients with schizophrenia (ICD-10, F20). The GEO study is designed to last from May 2000 to June 2004. Six hundred and forty-six adult patients treated with either olanzapine (N=416) or haloperidol (N=230) were enrolled in the study after having signed the informed consent document. Patients in this study were not randomized to a certain treatment, but merely observed under the treatment that the investigators deemed suitable under everyday conditions. The study was conducted in accordance with the Helsinki Declaration. The “Ärztekammer Hamburg” (Medical Association of Hamburg) provided a general ethics approval for the study.

## Instruments

At enrollment in the study and every 6 months during the follow-up, patients are asked to complete a Visual Analogue Scale (VAS) and SF-12 questionnaire (the short form of SF-36 [Medical Outcomes Study 36-Item Short-Form Health Survey]) as generic instruments; the Munich Quality of Life Dimensions List (MLDL) as disease-specific instrument; and the Subjective Well-being under Neuroleptics (SWN) scale as treatment-specific instrument. The visual analogue scale (VAS) was used in addition to the SF-12 and is a separate instrument. It is a metric scale rated from 0 to 100 with 0 representing the worst possible health state and 100 representing the best possible health state.

### *SWN Scale (Short Form)*

In the GEO study, the short form of the SWN scale was applied. The objective of this self-report instrument is to measure the extent to which atypical antipsychotic treatment improves the patient’s subjective well-being, reflected by the following dimensions (factors): mental functioning, self-control, emotional regulation, physical functioning, and social integration. Each of these dimensions is measured by 4 items from the total of 20 items

for the whole instrument [10]. The items are summed for the five separate subscales [10].

The short form of the SWN scale was constructed following an item analysis based on data from 212 patients with schizophrenia [10] and showed sufficient internal consistency and external validity. However, a strict statistical test of the underlying measurement structure was not performed. This refers to the number of factors and the relations between items and factors. Using SEM allows to answer these questions. We get a formal test whether the postulated number of factors are sufficient to explain the observed variances and co-variances of the items and get a test whether the postulated measurement-model fits the data [12].

#### *SWN (Short Form) Item Phrasing and Rating*

The short form of the SWN scale is composed of 20 items (see the list below). Scores used to rate the positively worded SWN items (2, 3, 5, 7, 8, 12, 13, 15, 18, 19, and 20) range from 0 (“not at all”) to 5 (“very much”). (The rest of the scores for positively worded SWN items are as follows: 1 = “a little,” 2 = “somewhat,” 3 = “noticeable,” 4 = “much.”) The negatively worded items 1, 4, 6, 9, 10, 11, 12, 14, 16, and 17 were recoded. Hence, for these items, 0 means “very much,” 1 = “much,” 2 = “noticeable,” 3 = “somewhat,” 4 = “a little,” and 5 = “not at all.”

The SWN (short form) items are phrased as follows:

- 1) I feel powerless and not in control of myself.
- 2) I feel very comfortable with my body.
- 3) I find it easy to think.
- 4) I have no hope for the future.
- 5) My body feels familiar.
- 6) I am very shy about getting to know people.
- 7) I am imaginative and full of ideas.
- 8) My environment seems friendly and familiar to me.
- 9) I feel weak and exhausted.
- 10) My emotions and sensations are dull. Nothing matters to me.

- 11) My thinking is difficult and slow.
- 12) My feelings and behavior are inappropriate to situations. I get upset over small things, important ones hardly affect me.
- 13) I find it easy to keep in touch with people around me.
- 14) I perceive my environment as being changed, strange and threatening.
- 15) I find it easy to draw a line between myself and others.
- 16) My body is a burden to me.
- 17) My thoughts are flighty and undirected; I find it difficult to think clearly.
- 18) I am interested in what is happening around me, and it is important to me.
- 19) My feelings and behavior are appropriate in the particular situation.
- 20) I am full of confidence; everything will be all right.

Hereon, we will be referring to the SWN (short form) items as SWN variables (or simply variables).

## Statistical Analysis

### *Structural Equation Models*

Structural equation modeling (SEM) in its general form consists of two parts:

1. the measurement model and
2. the structural equation model [12].

The measurement model specifies how latent variables depend on or are indicated by the observed variables.

The structural equation model specifies the relationships between latent variables, describes effects, and assigns the explained and unexplained variance of the dependent (endogenous constructs). SEM integrates multivariate techniques such as linear regression analysis, analysis of covariance, and factor analysis [13].

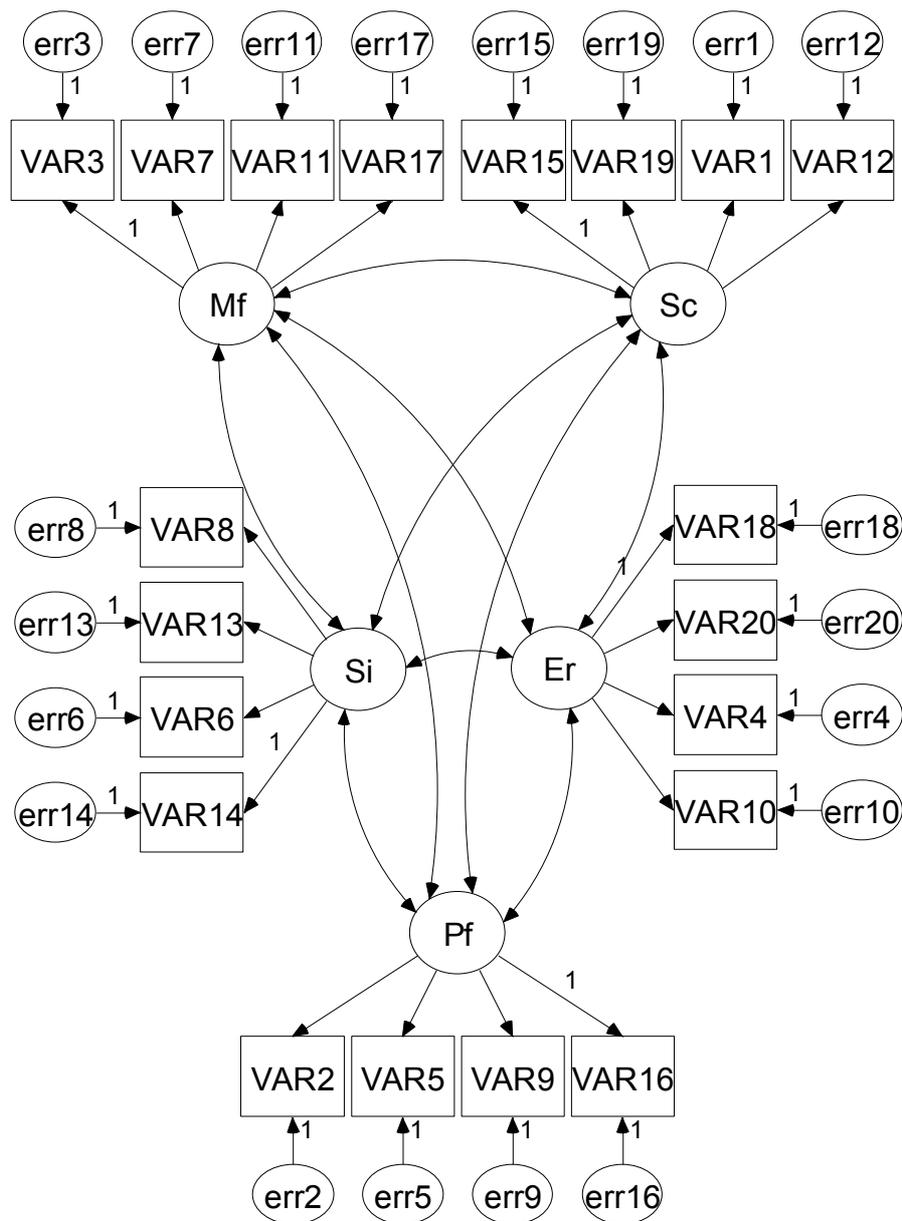
By comparing the covariance matrix derived by the model and the estimated parameters with the observed matrix, the validity of the model can be evaluated.

In this paper, graphical (rather than matrix equations) representation of SEM is used.

To start with, negatively worded SWN variables were recoded to conduct a consistent analysis (see the “SWN [Short Form] Item Phrasing and Rating” section). Initially, the measurement model originally formulated by Naber et al. [10] (the SWN scale, short form), which contains 20 items and five dimensions (factors), was reviewed using confirmatory factor analysis (CFA) [14].

#### *Baseline SEM and Confirmatory Factor Analysis*

The model postulated by Naber et al. [10] was tested as the baseline SEM for the explanation of covariances and correlations between the 20 selected items (variables). The correlations between errors reflect common sources of variation [14]. Figure 1 provides a graphic representation of the baseline SEM.



Naber et al. [10] postulated five SWN factors of first order – (for example, variables defining a latent variable by factor analysis such as “daily life” create the first-order factor “daily life”). Combining latent variables such as “daily life,” “physical functioning,” etc. using factor analysis creates a factor of second order (for example, a second-order latent variable “quality of life”). Higher-order factors are latent variables of a more abstract level (such as “general intelligence,” which influences first-order factors such as “verbal intelligence”).

In our study, mental functioning (Mf), self-control (Sc), emotional regulation (Er), physical functioning (Pf), and social integration (Sc) are considered first-order factors, whereas quality of life is a second-order factor. Four SWN (short form) variables were allocated to each of these five factors. Therefore, we were able to directly translate the model designed by Naber et al. [10] into a CFA and to empirically test the model using AMOS 4.0 computer program [15].

The new index's goodness of fit was measured using the Chi-square test. The probability value associated with the Chi-square test represents the likelihood of obtaining a Chi-square value that exceeds the Chi-square value when the null hypothesis ( $H_0$ ) is true. Thus, the higher the probability associated with the Chi-square value, the closer is the fit between the hypothesized model under  $H_0$  and the perfect fit [14]. The goodness-of-fit (GFI) value can fall between zero (0) and unity (1), where unity indicates a perfect fit. (See table 1 and Appendix for explanation of parameters used to test goodness of fit.)

As soon as one differentiates latent variables as endogenous and exogenous ones according to the substantial theory, a CFA is transformed into an SEM containing causal relations between latent variables [12].

An overview summarizes the key indices by which we will assess models and thresholds for acceptable fits are given in table 1 [16, 17].

**Table 1.** Key indices and fit measures

Descriptive Measures	Range	Criterion
GFI	0-1.0	$\geq .95$
AGFI	0-1.0	$\geq .95$
Chi <sup>2</sup> ÷ df	> 0	$3.0 \geq x \geq 1.0$
Statistical Inference		

Measurement of approximate fit:

RMSEA	0-1	$\leq .05$
PCLOSE	0-1	> .5

Information on theoretical measure:

AIC	not defined	the lower the better
CAIC	not defined	the lower the better
BIC	not defined	the lower the better

Abbreviations: GFI = goodness of fit index; AGFI = adjusted goodness of fit index; df = degrees of freedom; RMSEA = root mean squared error of approximation; AIC = Akaike information criterion; CAIC = Confirmatory Akaike information criterion; BIC = Bayes information criterion.

#### *Factor Scores and Factor Loadings*

Factor scores are necessary to construct weighted indexes. Based on the mode of their calculation, factor scores produce the strength of relation between items (variables) and target factors with prediction of the latter parameters [12]. Scores for both first-order factors and factors of higher order were calculated with AMOS 4.0. The factor scores were further processed with the program SPSS 9.0 (SPSS Inc.; Chicago, Illinois, USA).

Factor loadings are standardized or unstandardized partialized regression coefficients from the latent variables to the indicators. (The term “partialized” means that other independent variables [factors] are held constant [12].) Factor scores are the weight for predicting a latent variable from its observed indicators [18].

### *Modification Indexes*

We introduced 17 residual correlations between measurement errors of variables using modification indexes (MI).

MI allows one to examine all potential modifications in a single analysis. They can provide suggestions for modifications that will likely result in a better fit as shown by lower chi-square values [15].

## Results

The maximum of three hundred and sixty patients were available for factor analyses. The observations are all from cross-sectional data. Patients were medicated at the time of measurement. The missing value was due to patients lost to follow up. (Although we have chosen pairwise exclusion of missing values [not all patients completed the quality-of-life questionnaires] as an option, our sample size was reduced to a maximum of 360 because quality-of-life values were missing for a considerable number of patients.)

## Values for SWN (Short Form) Variables

Means and standard deviations for SWN (short form) variables are given in Table 2.

**Table 2.** Means and Standard Deviations for SWN (Short Form) Items<sup>a</sup>

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>	<b>N</b>
VAR1	4.12	1.23	358
VAR2	3.00	1.25	360
VAR3	3.02	1.32	352
VAR4	3.75	1.43	357
VAR5	3.41	1.26	356
VAR6	3.18	1.55	356
VAR7	2.69	1.35	354
VAR8	3.02	1.25	357
VAR9	3.20	1.47	355
VAR10	3.75	1.32	355
VAR11	3.53	1.36	355
VAR12	3.61	1.36	351
VAR13	3.00	1.37	355
VAR14	4.07	1.19	353
VAR15	2.83	1.26	353
VAR16	3.87	1.35	354
VAR17	3.66	1.38	357
VAR18	3.40	1.23	358
VAR19	3.23	1.09	351
VAR20	3.31	1.43	357

<sup>a</sup> In this context, the term “items” is synonymous with the term “variables.”

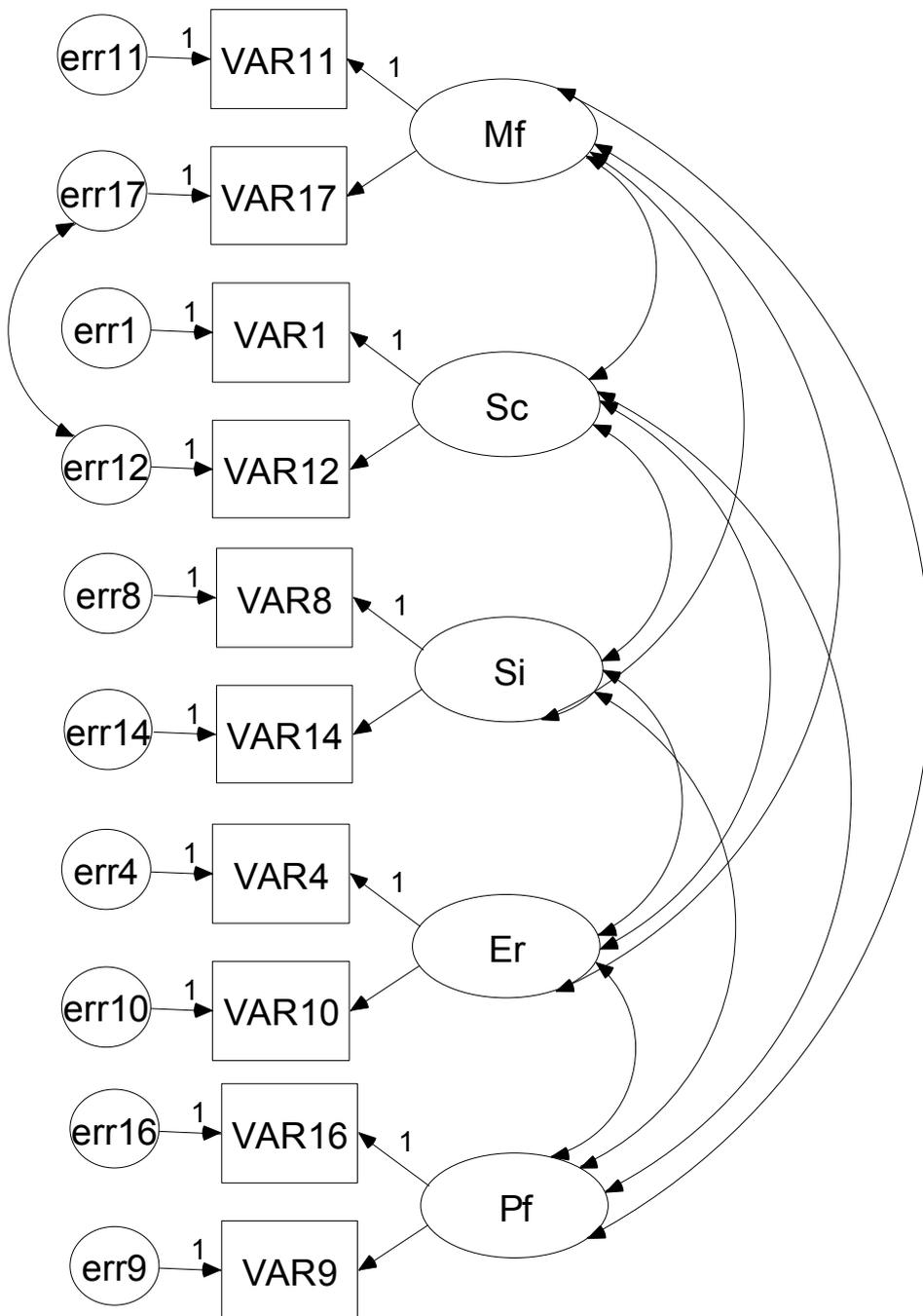
Abbreviation: VAR = variable.

The values for variables 1 and 15 are elucidated in the following: Variable 1 is negatively phrased and thus was recoded (see “SWN [Short Form] Item

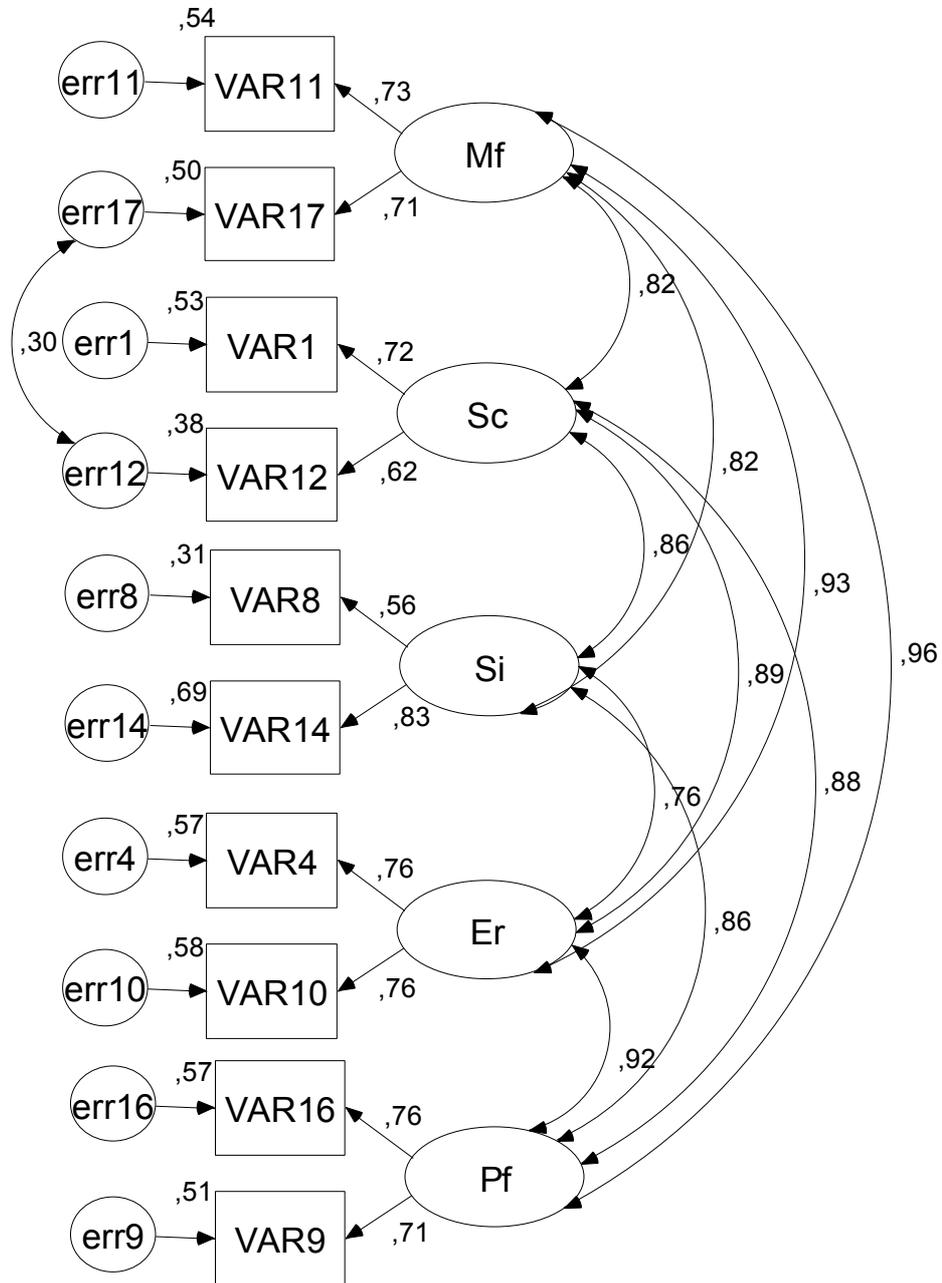
Phrasing and Rating”). For this variable, the mean of 4.12 is roughly equivalent to the answer category “a little.” (On average, patients in the studied sample replied to the question “I feel powerless and not in control of myself” with the answer category “a little.”) Variable 15 is positively phrased (“I find it easy to draw a line between myself and others”). For this variable, the mean of 2.83 is roughly equivalent to the answer category “noticeable.”

#### *Developing SEMs for SWN (Short Form) Indexes*

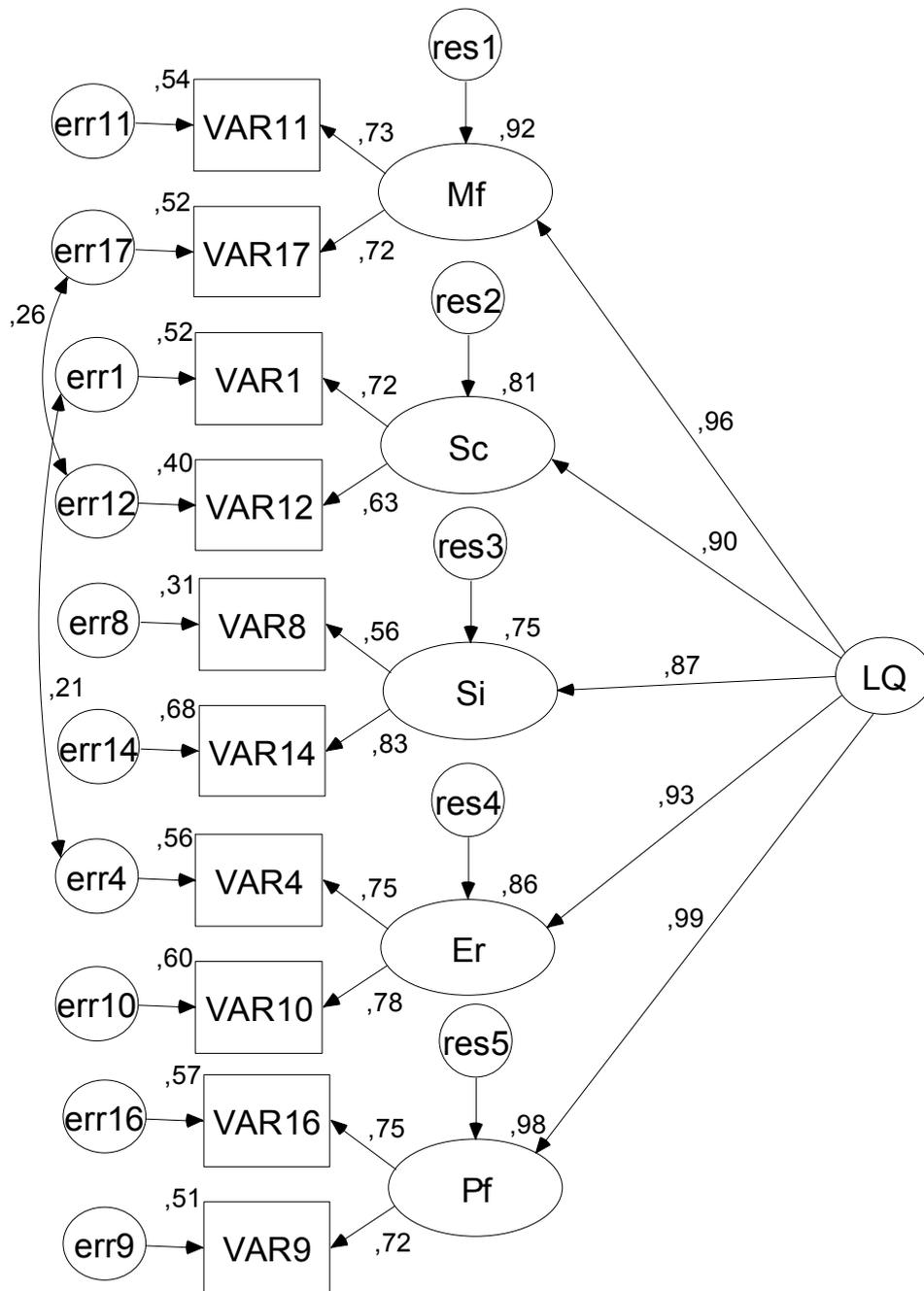
After applying a series of residual correlations based on respective modification indexes, a satisfactorily adapted final model was developed (model 1; Figure 1). Subsequently, in order to reduce the number of items for every latent variable to get a more economic instrument, we reduced the number of items by the following procedure. Based on the strength of the factor loadings of the items on their respective factors, we reduced the number of items from 20 to 10. This abridged scale (index) also contains five factors. Every factor is now measured with two items instead of four (see Figure 2). Estimation of a CFA model is known to be sensitive to multivariate kurtosis and skewness. However, as Boomsma and Hoagland [19] have shown, this is mainly influencing only the standard errors and the chi square test, if deviations are considerable. As the more complicated procedures like asymptotic distribution-free estimation methods need a much higher sample size, we have decided to use maximum likelihood estimation.



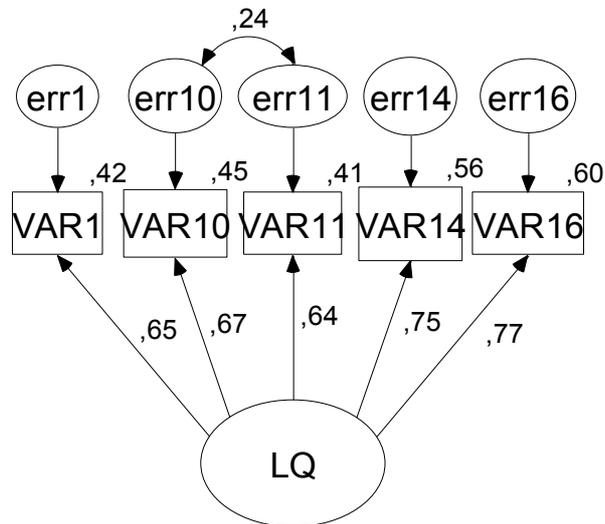
This second model (Figure 3) was again modified by introduction of an error correlation yielding good model fit.



To make allowances for the hypothesis of a factor of second order representing SWN, model 2 was then modified by the introduction of a factor of second order to form model 3 (Figure 4).



After the introduction of another error correlation, this model, too, exhibited good fit to the data. In a fourth and final step, the scale was again reduced by 5 items based on factor scores (model 4; Figure 5).



The resulting measurement instrument consisted of 5 variables illustrated by the only factor, “SWN.” Having specified an error correlation, this model fit the data very well.

*Model 1*

The application of modification indexes resulted in a more satisfactory fit of the baseline model to data (see Table 3).

**Table 3.** Global Goodness of Fit Parameters of Baseline and Final Model (Model 1)

<b>Parameter</b>	<b>Baseline Model</b>	<b>Final Model</b>
Chi <sup>2</sup> ÷ df	4.61	2.55
AGFI	0.74	0.86
RMSEA	0.10	0.07
PCLOSE	0.00	0.00
AIC (default to saturated)	837.46 to 420.00	499.12 to 420.00
BIC (default to saturated)	1181.14 to 1863.43	959.65 to 1863.43
CAIC (default to saturated)	1081.35 to 1444.32	825.93 to 1444.32

Abbreviations: df = degrees of freedom; AGFI = adjusted goodness of fit index; RMSEA = root mean squared error of approximation; AIC = Akaike information criterion; BIC = Bayes information criterion; CAIC = Confirmatory Akaike information criterion.

The standardized factor loadings and explained variances of variables ( $R^2$ ) of the final model are presented in Table 4.

**Table 4.** Standardized Factor Loadings and Explained Variances of the Variables of the Final Model (Model 1)

<b>Variable</b>	<b>Mf</b>	<b>Sc</b>	<b>Si</b>	<b>Er</b>	<b>Pf</b>	<b>R<sup>2</sup></b>
VAR3	0.60	—	—	—	—	0.36
VAR7	0.45	—	—	—	—	0.20
VAR11	0.73	—	—	—	—	0.53
VAR17	0.73	—	—	—	—	0.53
VAR15	—	0.37	—	—	—	0.14
VAR19	—	0.61	—	—	—	0.37
VAR1	—	0.70	—	—	—	0.50
VAR12	—	0.63	—	—	—	0.40
VAR14	—	—	0.75	—	—	0.57
VAR6	—	—	0.40	—	—	0.16
VAR13	—	—	0.46	—	—	0.21
VAR8	—	—	0.64	—	—	0.41
VAR18	—	—	—	0.50	—	0.25
VAR20	—	—	—	0.68	—	0.46
VAR4	—	—	—	0.75	—	0.56
VAR10	—	—	—	0.71	—	0.50
VAR16	—	—	—	—	0.75	0.56
VAR9	—	—	—	—	0.75	0.56
VAR5	—	—	—	—	0.65	0.42
VAR2	—	—	—	—	0.69	0.48

The em-dash (—) means “not applicable.”

Abbreviations: Mf = mental functioning; Sc =self-control; Si = social integration; Er = emotional regulation; Pf = physical functioning; VAR = variable.

The standardized factor loadings and explained variances showed that the relation between factors and variables was both valid and very reliable [20,

21]. Reliability of the items is defined as the explained variance of an item, whereas formal validity of an item is defined as the factor loading of the item on the latent variable [12].

Zero loadings in table 4 are the result of corresponding restrictions derived from theoretical predictions, which have been not falsified by the data.

Table 5 presents factor correlations for the SWN scale (short form). These factor correlations are, without exception, very high ranging (from .79 to .88). This indicates that the factors are all strongly related, which may be the effect of one second-order factor.

**Table 5.** SWN (Short Form) Factor Correlations (Model 1)

	Mf	Sc	Si	Er	Pf
Mf	1.00	0.79	0.83	0.88	0.83
Sc	—	1.00	0.89	0.85	0.84
Si	—	—	1.00	0.84	0.80
Er	—	—	—	1.00	0.86
Pf	—	—	—	—	1.00

The em-dash (—) means ‘not applicable.’

Abbreviations: SWN = Subjective Well-being under Neuroleptics; Mf = mental functioning; Sc = self-control; Si = social integration; Er = emotional regulation; Pf = physical functioning.

## Model 2

In the next step we chose those 10 items (variables) from the 20 SWN (short form) items with the highest loading on their respective factors (Figure 2). Again, the respective CFA model included five factors, but now only two variables per factor. This model was again tested with AMOS 4.0 and revealed satisfactory fit of the model to data (see Table 6). Since an error correlation of residuals was still high, we modified the model accordingly. The resulting model fits the data very well (see Table 6).

**Table 6.** Global Goodness of Fit Parameters of Baseline and Final Models (Model 2)

Parameters	Baseline Model	Final Model
Chi <sup>2</sup> ÷ df	2.70	1.90
AGFI	0.92	0.94
RMSEA	0.07	0.05
PCLOSE	0.05	0.46
AIC (default to saturated)	127.50 to 110.00	107.50 to 110.00
BIC (default to saturated)	312.91 to 449.92	299.08 to 449.92
CAIC (default to saturated)	273.83 to 378.28	258.70 to 378.28

Abbreviations: df = degrees of freedom; AGFI = adjusted goodness of fit index; RMSEA = root mean squared error of approximation; AIC = Akaike information criterion; BIC = Bayes information criterion; CAIC = Confirmatory Akaike information criterion.

The standardized coefficients and explained variances of indicators ( $R^2$ ) of the final model (model 2) are presented in Figure 3. The weight of factor loadings and explained variances showed that the relation between factors and variables was both valid and very reliable.

### Model 3

Thereafter, we modified this model and introduced SWN as a second-order factor. By this procedure, the first-order factors became endogenous (dependent) latent variables with an error component [12]. Modifications for this model were also satisfactory (see Table 7). Since, however, one MI still exhibited a higher value for one residual correlation, we used the respective residual correlation as a free parameter in the modified model. The fit measures of this model were excellent and are presented in Table 7.

**Table 7.** Global Goodness of Fit Parameters of Baseline and Final Model (Model 3)

Parameters	Baseline Model	Final Model
Chi <sup>2</sup> ÷ df	2.05	1.78
AGFI	0.94	0.95
RMSEA	0.05	0.048
PCLOSE	0.33	0.57
AIC (default to saturated)	111.49 to 110.00	103.80 to 110.00
BIC (default to saturated)	272.18 to 449.92	270.67 to 449.92
CAIC (default to saturated)	238.31 to 378.28	235.50 to 378.28

Abbreviations: df = degrees of freedom; AGFI = adjusted goodness of fit index; RMSEA = root mean squared error of approximation; AIC = Akaike information criterion; BIC = Bayes information criterion; CAIC = Confirmatory Akaike information criterion.

The standardized coefficients and explained variances of variables ( $R^2$ ) of the final model (model 3) are presented in Figure 4. The weight of factor loadings and explained variances showed that the relation between factors and variables was both valid and very reliable.

One can see that some of the regression coefficients of the second-order factor on the first-order factors are very high. This is also reflected in the high correlations between some of the first-order factors. However, as the items only load on the respective factors that they should measure, we can argue that the first-order factors really represent different constructs.

#### Model 4

As the last step, we specified and tested another abridged version of SWN, consisting of five items (variables). One variable with the highest factor weight was chosen per factor. Since only one variable per factor was used and factors were highly correlated, we postulate for this model that one factor is sufficient to explain the covariances between the five variables.

The overall goodness of fit for this model was relatively good (see Table 8). Yet, due to the extent of an MI, it was necessary to admit an error correlation between two residuals. The resulting final model fit data very well (see Table 8).

**Table 8.** Global Goodness of Fit Parameters of Baseline and Final Models (Model 4)

Parameters	Baseline Model	Final Model
Chi <sup>2</sup> ÷ df	3.75	1.45
AGFI	0.94	0.98
RMSEA	0.09	0.04
Pclose	0.06	0.58
AIC (default to saturated)	38.73 to 30.00	27.78 to 30.00
BIC (default to saturated)	93.52 to 112.18	88.05 to 112.18
CAIC (default to saturated)	87.43 to 103.04	81.34 to 103.04

Abbreviations: df = degrees of freedom; AGFI = Adjusted Goodness of Fit Index; RMSEA = root mean squared error of approximation; AIC = Akaike information criterion; BIC = Bayes information criterion; CAIC = Confirmatory Akaike information criterion.

The standardized coefficients and explained variances of variables ( $R^2$ ) of the final model are presented in Figure 5. The weight of factor loadings and explained variances showed that the relation between factors and items is both valid and very reliable.

## Constructing the SWN (Short Form) Index and Testing Its Validity

### *Factor Scores of Models 3 and 4*

For models 3 and 4 (final model each) we calculated the factor scores (see Tables 9 and 10) with regard to the “SWN” factor to obtain an index for assessing subjective well-being of patients with schizophrenia undergoing antipsychotic therapy. The factor scores were further processed with the SPSS 9.0 software. Finally, these new index variables were standardized using a Z value transformation to yield two standardized index variables.

**Table 9.** Factor Loadings (Model 3)

<b>Variable</b>	<b>Factor Loading</b>
VAR9	0.72
VAR16	0.75
VAR10	0.78
VAR4	0.75
VAR14	0.83
VAR8	0.56
VAR12	0.63
VAR1	0.72
VAR17	0.72
VAR11	0.73

Abbreviation: VAR = variable.

**Table 10.** Factor Loadings (Model 4)

<b>Variable</b>	<b>Factor Loading</b>
VAR16	0.77
VAR14	0.75
VAR11	0.64
VAR10	0.67
VAR1	0.65

Abbreviation: VAR = variable.

To calculate quantitatively the strength of the relation between the 10-item index and the 5-item index, the respective correlation was calculated using SPSS 9.0. This correlation yielded the extremely high score of 0.967, which confirms that the loss of information from 10 to 5 items is extraordinarily low. The mean of the newly formed SWN index of model 4 is 2.83, standard deviation is 0.73.

The following SWN (short form) items (variables) are included in the new SWN index for patients with schizophrenia:

- VAR1: I feel powerless and not in control of myself.
- VAR10: My emotions and sensations are dull. Nothing matters to me.
- VAR11: My thinking is difficult and slow.
- VAR14: I perceive my environment as being changed, strange and threatening.
- VAR16: My body is a burden to me.

## Discussion

Using confirmatory factor analysis and structural equation modeling, we have demonstrated in this paper that the SWN (short form) is a reliable and valid instrument for testing subjective well-being of patients with schizophrenia undergoing antipsychotic therapy. This method can be used for testing the validity of any quality-of-life (QoL) instrument with interval scales. In addition, we have succeeded in constructing an index based on SWN by means of SEM. The new index can be widely used in QoL research of patients with schizophrenia. Furthermore, it would be interesting to apply the instrument to cost-effectiveness studies, although for this purpose additional replications would be necessary.

We have shown that a reduction of the SWN (short form) to 10 and additionally to 5 items does not lead to a significant loss of the formal validity of the SWN. This is especially important for large multi-purpose studies (such

as health-outcomes or epidemiological surveys) where the instruments should not be too time-consuming and short scales are urgently necessary. Based on the results of our study, a 5-item version of the SWN can be easily used in future investigations as a disease-specific instrument for measuring subjective well-being (and thus, quality of life) of patients with schizophrenia undergoing antipsychotic therapy. An additional benefit of short instruments is the possibility for preference measurement (e.g., thorough conjoint analysis).

As mentioned earlier, SEM has not been widely used in QoL research for validation purposes. Keller et al. [2] provided detailed analysis of the SEM use for testing construct validity of the SF-36 Health Survey. Due to the partially different measures of fit, a direct comparison between the SF-36 and SWN is only possible in a limited way.

The introduction of higher-order factors resulted in better fits for both instruments. The underlying concept of using the “quality of life” latent variable as a higher-order factor may be too general for QoL research. Further research is urgently needed in this field to test the stability and validity of this concept in different settings and groups of patients.

## Recommendations

As a next step, it would be beneficial to test the external validity of the 5-item SWN index by comparing subjective well-being of patients with schizophrenia treated with atypical versus typical antipsychotics (for example, olanzapine or haloperidol) using additional external QoL criteria (such as ratings from physicians).

## Appendix

The following is the explanation of goodness of fit parameters per Arbuckle [15].

### ***CMIN/DF***

**CMIN/DF** is the minimum discrepancy,  $\hat{C}$ , divided by its degrees of freedom:

$$\frac{\hat{C}}{d} \cdot$$

### ***RMSEA***

Browne and Cudeck [22] suggested compensating for the effect of model complexity by dividing  $F_0$  by the number of degrees of freedom for testing the model. Taking the square root of the resulting ratio gives the population “root mean square error of approximation,” called **RMS** by Steiger and Lind, and **RMSEA** by Browne and Cudeck [22].

### ***PCLOSE***

**PCLOSE** =  $1 - \phi$  is a “p value” for testing the null hypothesis that the population **RMSEA** is no greater than .05. Browne and Cudeck [22] suggest that an **RMSEA** of .05 or less indicates a “close fit.” **PCLOSE** gives a test of close fit, whereas **P** gives a test of exact fit.

### ***AIC***

The Akaike information criterion [23, 24] is given by

$$AIC = \hat{C} + 2q.$$

### ***BIC***

In comparison with the **AIC**, **BCC** (Browne-Cudeck criterion), and **CAIC**, the **BIC** assigns a greater penalty to model complexity, and so has a greater

tendency to pick parsimonious models. (BCC is a measure developed specifically for analysis of moment structures. Browne and Cudeck [22] provided some empirical evidence suggesting that BCC may be superior to more generally applicable measures.)

### ***CAIC***

**CAIC** assigns a greater penalty to model complexity than either **AIC** or **BCC**, but not as great a penalty as does **BIC**.

### ***GFI***

The GFI (goodness of fit index) was devised by [25] for MI and ULS estimations, and generalized to other estimations criteria by [26].

### ***AGFI***

The AGFI (adjusted goodness of fit index) takes into account the degrees of freedom available for testing the model.

## ***Figure Legends***

### **Figure 1. Model 1**

CFA with 20 items (Naber's model). Baseline model and final model.

Abbreviations: err = error; VAR = variable; Mf = Mental functioning; Sc = Self-control; Er = Emotional regulation; Pf = Physical functioning; Si = Social integration; CFA = confirmatory factor analysis.

### **Figure 2. Model 2**

CFA of first order with 10 items.

Abbreviations: err = error; VAR = variable; Mf = Mental functioning; Sc = Self-control; Er = Emotional regulation; Pf = Physical functioning; Si = Social integration; CFA = confirmatory factor analysis.

**Figure 3. Standardized Output Path Diagram of the Final Model (Model 2)**

Abbreviations: err = error; VAR = variable; Mf = Mental functioning; Sc = Self-control; Er = Emotional regulation; Pf = Physical functioning; Si = Social integration.

**Figure 4. Standardized Output Path Diagram of Final Model (Model 3)**

Abbreviations: err = error; VAR = variable; res = residual; Mf = Mental functioning; Sc = Self-control; Er = Emotional regulation; Pf = Physical functioning; Si = Social integration; SWN = subjective well-being.

**Figure 5. Standardized Output Path Diagram of Final Model (Model 4)**

CFA of first order with five items and one factor (quality of life). Baseline model and final model.

Abbreviations: err = error; VAR = variable; SWN = subjective well-being; CFA = confirmatory factor analysis.

*Abbreviations*

<b>Abbreviation</b>	<b>Definition</b>
AGFI	Adjusted Goodness of Fit Index
AIC	Akaike information criterion
BCC	Brown-Cudeck criterion
BIC	Bayes information criterion
CAIC	Consistent Akaike information criterion
CFA	confirmatory factor analysis
<b>CMIN/DF</b>	Minimum value of the discrepancy function divided by degrees of freedom
<b>Er</b>	emotional regulation
err	error
EuroQol	European Quality of Life [questionnaire]
GEO	<i>Gesundheitsökonomische Evaluation von Olanzapin in Deutschland</i> (health economics study of olanzapine in the treatment of schizophrenia in Germany)
GFI	goodness of fit index

<b>Abbreviation</b>	<b>Definition</b>
ICD-10	<i>International Classification of Diseases, Tenth Edition</i>
Mf	mental functioning
MI	modification index/indices
MLDL	Munich Quality of Life Dimensions List
PCLOSE	“p value” for testing the null hypothesis of the close fit
Pf	physical functioning
“p value” QoL	P is the probability of getting as large a discrepancy as occurred with the present sample quality of life
res	residual
RMS	root mean square error of approximation (Steiger and Lind)
RMSEA	root mean squared error of approximation (Browne and Cudeck)
Sc	self-control
SEM	structural equation model(s)
SF-36	Medical Outcomes Study 36-Item Short-Form Health Survey
Si	social integration
SRM	Social Research Methodology [database]
SWN	Subjective Well-being under Neuroleptics
ULS	unweighted least squares
VAR	variable
VAS	Visual Analogue Scale

## 5. Schlussfolgerungen

### 5.1 Mögliche Beiträge der Gesundheitsökonomie zur Überwindung der durch politische Rahmenbedingungen erzeugten Probleme

#### 5.1.1 Theoretische Grundlage

Sowohl die Struktur des sozialen Sicherungssystems als auch des Subsystems Gesundheitswesen, das in unabhängige Sektoren aufgeteilt ist, führen wie erwähnt zu einer Gefangenendilemma-Situation.

Hierbei kommt es nicht notwendigerweise zu einer pareto-effizienten Lösung. Bei einem Gefangenen-Dilemma handelt es sich um eine Situation, bei der zwei Verbrecher, die gemeinsam ein Verbrechen begangen haben, in getrennten Räumen verhört werden, folglich keine Informationen austauschen können. Nehmen wir an, dass ein Gefangener die Möglichkeit hat, das Verbrechen zuzugeben oder zu leugnen. Wenn nur ein Gefangener gesteht, wird er entlassen und für den anderen Gefangenen würde eine Strafe von sechs Monaten verhängt werden. Gestehen beide, so erhalten sie jeweils drei Monate Gefängnisstrafe. Wenn beide Gefangenen leugnen, dann würden beide nur zu einem Monat verurteilt werden. Die Auszahlungsmatrix sieht wie folgt aus:

	Gestehen		Leugnen	
	Spieler A	Spieler B	Spieler A	Spieler B
Gestehen	-3	-3	0	-6
Leugnen	-6	0	-1	-1

Versetzen wir uns in die Lage des Spielers A. Wenn sich Spieler B entscheidet, das Verbrechen zu leugnen, dann ist A sicher besser gestellt, wenn er gesteht, da er freigelassen wird. Es ist für A ebenfalls günstiger zu gestehen, wenn B gesteht, weil A dann statt einer 6-Monatsstrafe nur eine 3-Monatsstrafe erhält. Umgekehrt gilt dasselbe für Spieler B.

Das für beide günstigste Ergebnis, nämlich ihre Freilassung, würden sie erzielen, wenn beide leugneten (Varian, S.482). Das Problem liegt darin, dass für beide keine Möglichkeit besteht, ihre Handlungen zu koordinieren. Genau dieses Problem, nicht koordinieren zu können, liegt im deutschen Gesundheitssystem vor. Krankenversicherungen können die Konsequenzen wie Frühverrentung oder Arbeitsunfähigkeit externalisieren. Unter einer Externalität versteht man, dass Aktivitäten, wie z.B. eine ärztliche Behandlung, zu Kosten führt, die nicht als Teil des Marktpreises erscheint, weil sie von der Krankenkasse bezahlt wird. Sie wird externalisiert, also außerhalb des Marktgeschehens abgewickelt (Pindyck, 2005). Ein weiteres wichtiges Beispiel ist die mögliche Vermeidung der Frühverrentung durch Schizophrenie (Schnabel et al., 2006; Eichmann, 2005). Grob geschätzt könnten durch den Einsatz moderner Atypika 260-300 Millionen € eingespart werden, was fast dem gesamten Arzneimittelmarkt der Atypika entspricht. Die Krankenkassen müssten also mehr in Atypika investieren und könnten damit die Rentenversicherung entlasten. Diese Zusatzbelastung „erspart“ sich aber die Krankenkasse und aufgrund mangelnder Möglichkeiten der Kooperation bleibt der Rentenversicherung nichts anderes übrig, als zu zahlen. Es stellt sich nun die Frage, wie dieses Gefangenendilemma im deutschen Gesundheitssystem gelöst werden könnte und welche Beiträge die Gesundheitsökonomie dazu leisten kann.

#### 5.1.2 Festlegung von Grenzwerten

Ein Grenzwert ist eine gesetzlich festgeschriebene Höchstmenge, die ein Unternehmen oder eine Krankenkasse nicht überschreiten darf. Eine Höchstmenge könnte z.B. die Anzahl der frühverrenteten Patienten sein, die nicht nach therapeutischen Richtlinien behandelt wurden. Eine Krankenkasse oder andere Anbieter von Dienstleistungen, die für die Einhaltung von Richtlinien verantwortlich sind, müssen bei der Überschreitung dieser Grenze mit Strafen rechnen. Die Idee von Grenzwerten wurde allerdings in der Politik bisher nicht einmal erwähnt. Trotzdem sollte man m. E. diese Idee

nicht vergessen, weil sie zumindest der Diskussion zur Korrektur von Externalitäten Auftrieb verschaffen kann.

### 5.1.3 Emissionsgebühr

Eine Emissionsgebühr ist eine Gebühr, die auf jede Emissionseinheit erhoben wird. Solch eine Emissionseinheit könnte eine vermeidbare Frühverrentung sein. Es wäre zu erwarten, dass die Anzahl der vermeidbaren Frühverrentungen rückläufig wäre. Der zugrunde liegende Mechanismus wird im folgenden erklärt.

Wie *Abbildung 3* zeigt, führt eine Emissionsgebühr von 3000 € zu einem effizienten Verhalten eines Dienstleistungsanbieters im Gesundheitswesen.

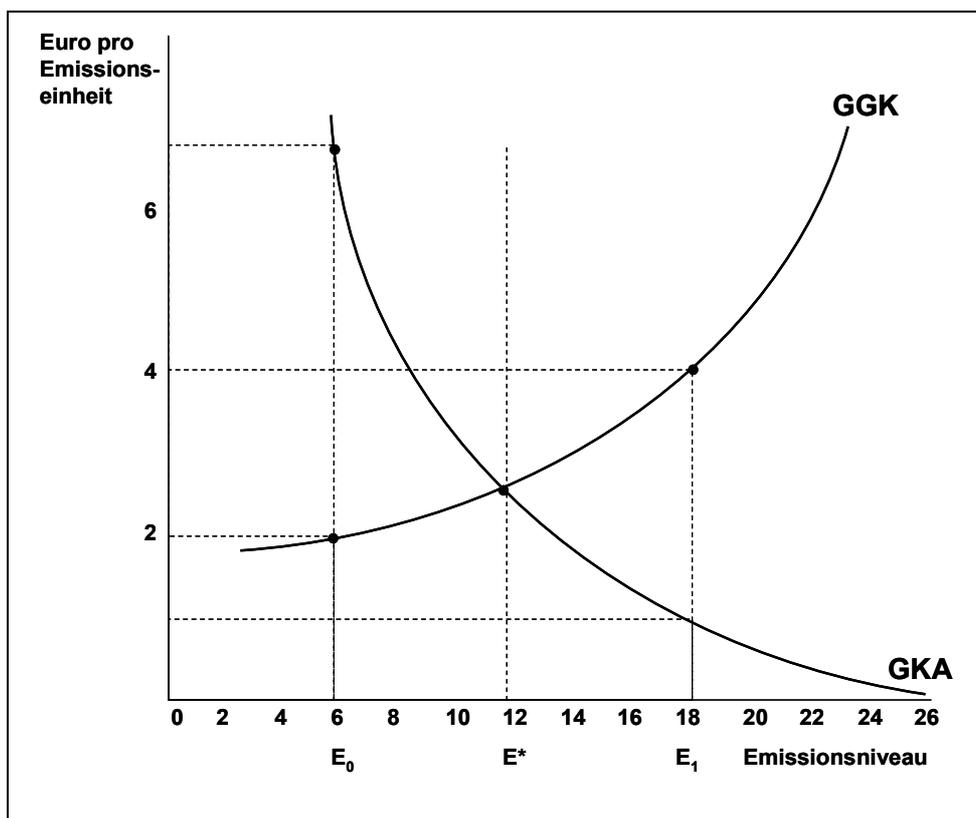


Abbildung 3

Mit diesem Betrag reduziert das Unternehmen, z.B. eine Krankenkasse, seine Kosten, indem es seine „Emissionen“ (vermeidbare Frühverrentungen) von 26 auf 12 Einheiten reduziert. Man erkennt, dass die erste Emissionseinheit eliminiert werden kann, ohne dass dies hohe Kosten verursacht. Das

Unternehmen muss also nur geringe Kosten aufbringen, um die 3000 € pro Emissionseinheit zu umgehen, weil die Grenzkosten der Emissionsvermeidung für alle Emissionsniveaus über 12 Einheiten unterhalb der Emissionsgebühr liegen. Das Unternehmen wird es hier also vorziehen, die Gebühr zu bezahlen. Unterhalb eines Niveaus von 12 Emissionseinheiten liegen die Grenzkosten der Emissionsvermeidung jedoch oberhalb dieser Gebühr und das Unternehmen wird es vorziehen, die Gebühr zu bezahlen, anstatt seine Emissionen zu reduzieren.

Der mögliche Einsatz von Emissionsgebühren zur Vermeidung von Externalität wird in Deutschland nicht genutzt, ja, nicht einmal diskutiert. Dabei ist das Einsparpotenzial durch Vermeidung von Externalität groß. Durch den Einsatz von modernen Atypika könnten z. B. zwischen 260 bis 300 Millionen € eingespart werden. Diesen Betrag könnte z. B. die Rentenversicherung der Krankenversicherung zahlen und dadurch Einsparungen in genau dieser Höhe erzielen. Damit wären sowohl Krankenkassen als auch das Gesundheitssystem finanziell entlastet. Bei Volkskrankheiten wie Osteoporose oder Diabetes kann man vermuten, dass die Einsparungen durch Vermeidung von Externalität noch wesentlich höher ausfallen würden.

#### 5.1.4 Emissionszertifikate

Unter der Voraussetzung, dass die Kosten einer Emissionsvermeidung bekannt wären, und die Kosten für alle Unternehmen gleich hoch wären, könnte man Grenzwerte festsetzen. Sind dagegen die Kosten der Unternehmen verschieden, dann kann weder ein Grenzwert noch eine Gebühr zu einem effizienten Ergebnis führen.

Ein effizientes Ergebnis könnte aber mit Hilfe von übertragbaren Emissionszertifikaten erreicht werden. Jeder Anbieter von Dienstleistungen im Gesundheitswesen muss die Erlaubnis haben, Frühverrentungen emittieren zu können. Jeder Anbieter, der Frühverrentungen emittiert, die nicht durch ein Zertifikat gedeckt sind, muss mit Strafen rechnen. Solche Zertifikate müssen frei handelbar sein.

Die Ansätze Grenzwerte, Gebühren und Zertifikate sind sicherlich Möglichkeiten, Externalitäten zu korrigieren. Allerdings sind weder der politische Wille noch die Datengrundlage zur Zeit vorhanden. Nichtsdestoweniger können diese Ansätze in der Zukunft Beiträge zur Korrektur von Externalitäten leisten.

Ein weiterer Ansatz, Externalitäten zu korrigieren, läge in der Aufhebung der Datenintransparenz. Zur Zeit liegen mit der integrierten Versorgung, den Disease-Management-Programmen und der Gesundheitskarte drei mögliche Problemlösungsansätze vor, um die durch die Sektoren verursachte Informationsintransparenz zu mindern bzw. zu beseitigen.

#### 5.1.5 Integrierte Versorgung

Nach §140a und §140b SGB V können Krankenkassen direkt mit niedergelassenen Ärzten und mit Krankenhäusern für ausgewählte Krankheiten wie Schizophrenie, Diabetes oder andere Indikationen oder aber für eine Gesamtversorgung innerhalb einer definierten Region Verträge abschließen. Im Rahmen dieser Verträge rechnen die Krankenkassen direkt mit den Ärzten und Krankenhäusern die Leistungen pro Patient ab. Der Informationsverlust, der dadurch entsteht, dass die Ärzte mit den Kassenärztlichen Vereinigungen und diese global mit den Kassen abrechnen, wird im Rahmen der Integrierten Versorgung umgangen. Nach § 130a können die Vertragspartner Krankenkassen, Ärzte und Krankenhäuser direkt mit anderen Anbietern von Gesundheitsleistungen, wie der Pharmaindustrie, Nachlässe verhandeln.

So vielversprechend die Integrierte Versorgung für die Minderung der Informationsintransparenz aussieht, weist sie doch Nachteile auf. Zum einen müssen sich, um zu einem von der Größe her befriedigenden Patientengut zu kommen, mehrere Krankenkassen an einem integrierten Versorgungsprojekt beteiligen. Dies bedeutet einen nicht unerheblichen Verhandlungsaufwand. Zum anderen sind diese Projekte regional begrenzt. Für Informationstransparenz müsste aber das gesamte Bundesgebiet im Rahmen eines integrierten Versorgungsprojektes gemanagt werden. Abgesehen da-

von, dass dann zwei bestehende Systeme konkurrierend existierten, wäre der Aufwand für eine flächendeckende Integrierte Versorgung m. E. nicht zu rechtfertigen.

Die Zahl der Verträge zur Integrierten Versorgung steigt zwar kontinuierlich an, von einer flächendeckenden Versorgung kann aber noch längst nicht die Rede sein. Zu diesem Ergebnis kommt eine Untersuchung, die am Lehrstuhl für Innovations- und Technologiemanagement der Technischen Universität Berlin von Prof. C. Schultz durchgeführt wurde (Schütze, 2006).

#### 5.1.6 Disease-Management-Programme

Der Gesetzgeber hat in §137f SGB V dem Gemeinsamen Bundesausschuss die Aufgabe der Disease-Management-Programme übertragen. Hierbei handelt es sich um strukturierte, sektorenübergreifende Behandlungsprogramme speziell für chronisch Kranke. Disease-Management-Programme gibt es bisher für Brustkrebs, Diabetes mellitus Typ 1 und 2, koronare Herzkrankheit und für Asthma/chronische obstruktive Lungenerkrankung (Häussler, S. 50ff., 2005). Die Abrechnung erfolgt analog der integrierten Versorgung. Im Gegensatz zur integrierten Versorgung haben die Disease-Management-Programme bundesweit Gültigkeit, sind aber auf vier Indikationen beschränkt. Nach Informationen des Gemeinsamen Bundesausschusses sind zur Zeit nur 1,6 Millionen Patienten in Disease-Management-Programmen erfasst ([http://www.g-ba.de/cms/front\\_content.php?idcat=212](http://www.g-ba.de/cms/front_content.php?idcat=212)). Von einer flächendeckenden Informationstransparenz kann also nicht die Rede sein.

#### 5.1.7 Gesundheitskarte

Der Ansatz über eine Gesundheitskarte Informationstransparenz herzustellen ist insofern interessant, weil faktisch alle Patienten der gesetzlichen und privaten Krankenversicherung, also fast 80 Millionen Menschen erfasst werden sollen. Diese neue Chipkarte kann selbst Daten speichern, dient aber auch als Schlüssel zu allen im Netz archivierten Informationen (Lütge, 2006). Die große Datenmenge der Patientenakte wird also nicht auf der Karte selbst, sondern zentral gespeichert. Die Gesundheitskarte wird mit einem

Zwei-Schlüssel-System gesichert. Nur der Patient kann es seinem Arzt erlauben, seine Gesundheitsdaten einzusehen. Der Zugang ist nur mit Gesundheitskarte und Heilberufausweis möglich.

220000 Ärzte/Zahnärzte werden künftig papierlos verordnen. 21000 Apotheken werden auf elektronischem Wege ihre Rezepte abwickeln. 2200 Krankenhäuser werden nach Einwilligung des Patienten Zugriff auf die elektronische Patientenakte haben. Ca. 300 Krankenversicherungen sind dafür zuständig, dass die Gesundheitskarte ausgegeben wird.

Die flächendeckende, transsektorale Erfassung von Patientendaten scheint nun in greifbarer Nähe.

Die Daten, die auf der Gesundheitskarte erfasst werden, teilen sich aber in Pflichtdaten und freiwillige Daten. Die Pflichtdaten umfassen die Versichertendaten und das elektronische Rezept. Die freiwilligen Daten beinhalten den Notfalldatensatz. Zu diesem gehören Diagnosen und hier insbesondere chronische Erkrankungen, Arzneimittelunverträglichkeiten, Allergien und operative Eingriffe. Zusätzlich zu dem Notfalldatensatz kommen die Arzneimitteldokumentation und die Patientenakte.

Die Pflichtdaten und die freiwilligen Daten würden einen Datensatz bilden, der eine transsektorale Transparenz ermöglichen würde, um das Gefangenendilemma der Krankenkassen aufzulösen. Mit einem umfassenden Datensatz könnte also Externalität vermieden (vgl. S. 141) und damit Einsparungen erzielt werden. Es stellt sich nun die Frage, wann mit einem solch umfassenden Datensatz durch die Gesundheitskarte gerechnet werden kann. Es ist meines Erachtens nicht damit zu rechnen, dass, zumindest kurzfristig, die freiwilligen Daten umfassend zentral gespeichert werden.

Die Beiträge der Gesundheitsökonomie zur Allokation knapper Ressourcen ist durch das Gefangenendilemma der Krankenkassen eingeschränkt. Ansätze zur Lösung dieses Problems wie Grenzwerte, Emissionsgebühren, Emissionszertifikate, integrierte Versorgung, Disease-Management Programme und die Gesundheitskarte lösen, wie obige Analyse zeigt, dieses Problem nur ansatzweise.

Es stellt sich nun die Frage, welche Beiträge unter diesen Bedingungen die Gesundheitsökonomie leisten kann und welche Beiträge dazu die erreichten Arbeiten leisten.

#### 5.1.8 Modellierung

Die Datentransparenz über Sektoren kann über das Instrument der Modellierung zumindest teilweise erreicht werden. Die Vorgehensweise gestaltet sich dabei wie folgt: In einem ersten Schritt wird ermittelt, wie sich kostenintensive Ressourcen unter verschiedenen Therapien entwickeln. So kann es z.B. sein, dass Therapie A die Einweisung in ein Krankenhaus und/oder die Liegedauer verkürzt. Methodisch kann dies mit einer randomisierten intervenierenden Studie ermittelt werden. Interessant an diesem Beispiel ist, dass die Therapie A z.B. im niedergelassenen Bereich erfolgt und Auswirkungen auf den Krankenhausbereich hat, die nun berechnet werden können. Die Datentransparenz ist insofern eingeschränkt, weil die Teststärke für randomisierte, intervenierende Studien auf Wirksamkeit berechnet wird und deshalb meist nicht repräsentativ ist. Außerdem hat eine Interventionsstudie nur für die Ein- und Ausschlusskriterien Gültigkeit. Dennoch nähert man sich mit dieser Methode der wahren Situation. Mit Hilfe einer solchen Modellierung kann man Kosteneffektivitäts-, Kostennutzwert- und Budget-Impact-Analysen erstellen. Mit Hilfe der Modellierung ist es also, wenn auch begrenzt, möglich, die Allokation knapper Ressourcen zu optimieren.

Die Arbeit von Deckert (Deckert, 2001) zeigt mit Hilfe einer Markov Modellierung wie für die Krankheit Schizophrenie eine solche Analyse durchgeführt und berechnet werden kann. Die Kosten der Frühverrentung wurden dabei nicht berücksichtigt.

Eine Arbeit von Schnabel geht insofern einen Schritt weiter, als dass ein transsektoraler Aspekt, und zwar der der Frühverrentung, beleuchtet wird (Schnabel, in review). Eine weitere Arbeit, die eine Modellierung berechnet, die Frühverrentung mit einbezieht, befindet sich zur Zeit in review.

Interessant ist bei dieser Untersuchung, dass es sich nicht um eine Markov-Modellierung, sondern um ein System Dynamic Modell handelt (Clouth,

2007). Wie schon erwähnt setzt NICE die Methode der Modellierung ein. In Deutschland ist diese Vorgehensweise, wie dem Methodenpapier des IQWiG ([www.IQWiG.de](http://www.IQWiG.de)) zu entnehmen ist, nicht akzeptiert.

Zur Validierung von Modellen können wiederum Krankheitskostenstudien oder epidemiologische Studien herangezogen werden. Die Arbeit über die Kosten der Frühverrentung ist insofern interessant, da sie zur Validierung des System Dynamic Modells herangezogen werden konnte.

#### 5.1.9 Zentralregister

Wie oben erwähnt könnten auf der Gesundheitskarte alle für eine Optimierung der Allokation wichtigen Daten erfasst werden.

Die für die Allokation relevanten Daten müssten Pflichtdaten werden. Dies beträfe also Informationen über Diagnose, Ressourcenverbrauch, Schweregrad der Krankheit etc. Diese Daten müssten zwar patientenbezogen sein, aber für die Berechnung einer Allokation von Ressourcen wäre der Rückschluss auf einen Patienten ohne jegliches Interesse. Die Daten könnten also anonym bleiben, womit die politische Durchsetzbarkeit eines solchen Registers erleichtert würde.

Zugriff zu diesem Register könnte jeder haben, der ein berechtigtes Interesse nachweisen kann wobei diese Definition noch gesetzlich geregelt werden müsste. Die Daten selbst könnten von einer unabhängigen Institution wie dem DIMDI oder dem IQWiG gespeichert werden. Der Datensatz eines Zentralregisters wäre dem der Gesundheitskarte ähnlich oder sogar mit ihm identisch. Mit Hilfe eines solchen Datensatzes könnten Allokationen knapper Ressourcen optimiert und entsprechende Einsparungen erzielt werden.

### **5.2 Mögliche Beiträge der Gesundheitsökonomie zur Definition und Messung des Begriffs „Nutzens“**

Obige Ausführungen haben gezeigt, dass der Begriff „Nutzen“ vielschichtig ist. So gilt es z.B. die Perspektive des Patienten zu berücksichtigen, die sich nicht nur auf rein medizinische Parameter wie den Blutzuckerspiegel oder die Knochendichte beschränkt, sondern auch subjektive Parameter wie die

krankheitsbezogene Lebensqualität mit einbezieht. Der Nutzen kann aber auch für den Arbeitgeber anfallen, indem die Arbeitsfähigkeit durch eine Therapie verbessert wird. Auch die Perspektive der Krankenkassen, die sich weiter in die Sektoren Krankenhaus und niedergelassener Bereich aufteilen lässt, gilt es zu berücksichtigen. Es ist zwar wünschenswert, dass die Ergebnisse einer Therapie verbessert werden, aber nutzlos, wenn sie unbezahlbar bleiben. Auch transsektoral kann Nutzen entstehen, wenn Frühverrentung vermieden wird.

Eine Definition des Nutzens wird kaum gelingen, wenn die Bewertung zum einen nur durch Experten und zum anderen von einem begrenzten Personenkreis, wie dem Gemeinsamen Bundesausschuss, durchgeführt wird, weil die Perspektive von Experten fokussiert ist und eine kleine Personengruppe die Perspektiven einer pluralen Gesellschaft kaum repräsentieren kann.

Einen Vorschlag zur Lösung dieses Problems unterbreitet eine Arbeitsgruppe von Franz Porzsolt (Porzsolt, S. 353 – 364, 2006).

Das Modell zur Operationalisierung des Konzepts sieht vor, ein Gremium aus 40 oder mehr Mitgliedern zu bilden, von welchen z.B.

- 30 Mitglieder die Nutznießer von Gesundheitsleistungen vertreten
  - o 12 Vertreter der Versicherten (inklusive Rentenversicherung)
  - o 12 Vertreter der Fachgesellschaften
  - o 6 Vertreter der Industrie
- 6 Mitglieder die Ersteller von Gesundheitsleistungen (GKV) sind und
- 4 unabhängige Wissenschaftler (z.B. das IQWiG und DIMDI) (siehe auch *Abbildung 4*).

Meines Erachtens ist kritisch anzumerken, dass auch die Betroffenen, also Patienten und deren Angehörige berücksichtigt werden sollten. Dieses Gremium sollte auf alle Fälle die Vielfalt der Perspektiven des „Nutzens“ abbilden.

Diese Mitglieder werden für begrenzte Zeit (z.B. 4 Jahre) berufen, wobei etwa 5 Mitglieder pro Halbjahr durch Nachfolger ersetzt werden sollen. Das Auswahlverfahren der Mitglieder innerhalb der einzelnen Gruppen soll an-

streben, das reale Spektrum der Versicherten, Leistungserbringer und Erstat-  
ter von Gesundheitsleistungen widerzuspiegeln.

Diesem Gremium werden die zur Bewertung anstehenden Therapien vorge-  
tragen. Die Mitglieder des Gremiums müssen in der Lage sein, wissen-  
schaftliche Daten zur Wirkung, zur Wirksamkeit und zur Validität der  
Daten zu verstehen und nach subjektiven Kriterien bewerten zu können.

Diese Daten werden den Mitgliedern des Gremiums von Wissenschaftlern  
vorgetragen. Die Mitglieder des Gremiums nehmen eine subjektive Bewer-  
tung der Validität der Daten und eine subjektive Bewertung der Leistungs-  
kategorie (z.B. Wert der Verhinderung eines Todes oder der Verhinderung  
einer zeitlich begrenzten Behinderung) vor.

### **Konsens zur Bewertung des intangiblen Nutzens von Gesundheitsleis- tungen**

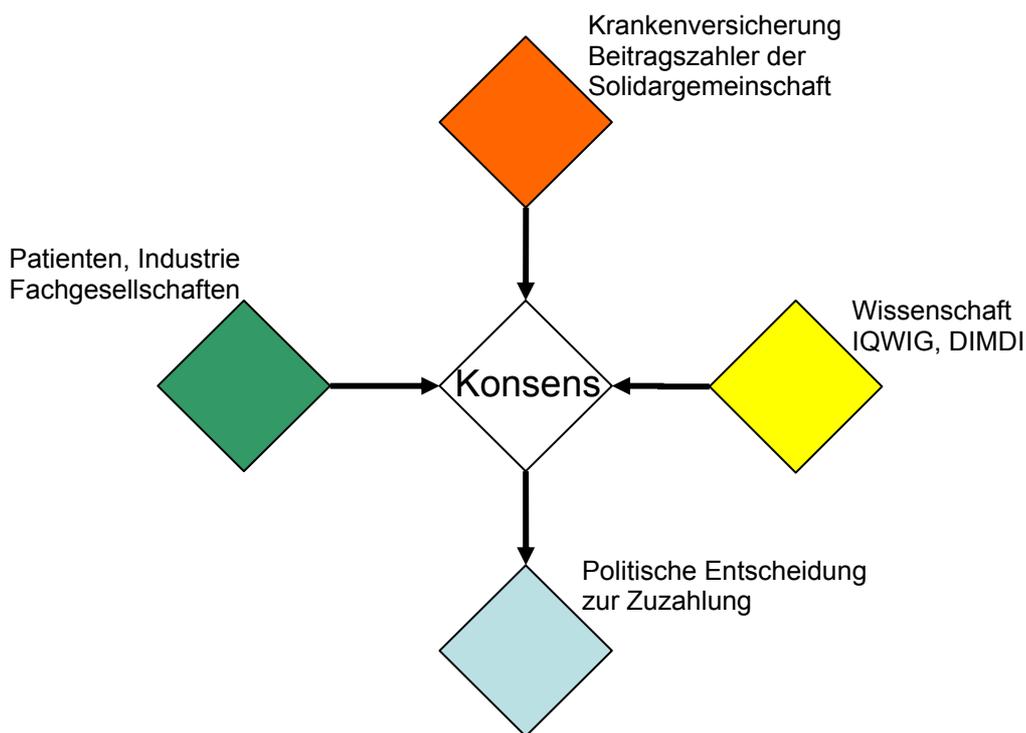


Abbildung 4: Konsensusgruppe zur Bewertung des intangiblen Nutzens von Gesundheitsleistungen.

Grün: Gruppe, die solidarisch finanzierte Gesundheitsleistungen nachfragt.

Rot: Gruppe, die auf Stabilität der Ausgaben achtet.

Gelb: Unabhängige Bewerter.

Blau: Politische Entscheidungsträger.

Das Gremium bewertet nur Gesundheitsleistungen, bei welchen von autorisierten Institutionen (BMGS oder g-BA) eine Diskussion der solidarischen Finanzierung vorgesehen ist.

Bei diesen Gesundheitsleistungen kann zur Ermittlung des erwarteten Mehrwerts folgendes Zwei-Schritt-Verfahren durchgeführt werden:

Der erste Schritt besteht darin, dass die zugänglichen wissenschaftlichen Daten, die der vergleichenden Bewertung mindestens zweier Verfahren zugrunde gelegt werden (Datenbasis), benannt werden.

Mit dem Gremium wird ein Termin zur Vorstellung der Frage und Präsentation der ersten Daten festgelegt. Die Datenbasis wird den Mitgliedern des Gremiums zur Verfügung gestellt.

Dazu stellen unabhängige Wissenschaftler Daten zum Wirkungsunterschied der Therapien und zur Validität der einbezogenen Studien vor.

Dem Gremium soll ausreichend Zeit eingeräumt werden, die präsentierten Ergebnisse anhand der vorliegenden Unterlagen nachzuvollziehen und ergänzende Fragen zu formulieren.

Der zweite Schritt wird innerhalb einer definierten Zeitspanne von z.B. 3 Wochen eingeleitet. Zu diesem Termin ermitteln die Mitglieder des Gremiums in einem schriftlichen Verfahren (die abstimmende Person ist nicht bekannt, aber ihre Zugehörigkeit zu einer der Gruppen) den erwarteten Gewinn an Wirksamkeit (EGW).

Dieser wird (siehe Beispiel in *Tabelle 2*) durch die Bewertung der Absoluten Risiko Reduktion (= Differenz [Delta] der Wirkungen) mit maximal 10 und minimal 5 Punkten ermittelt.

**Diskussionsvorschlag zur Berechnung des Nutzens**

Vergleich zweier Therapien zur Verhinderung der Reduktion der Lebensdauer oder Lebensqualität	<b>Ergebnisse beruhend auf Wiss Ergebnis, Konsensus</b>
Differenz der Wirkungen Orientierungswerte ARR > 40%: 10 Pkt; 39-25%: 9 Pkt; 24-15%: 8 Pkt, 14-10%: 7 Pkt; 9-5%: 6 Pkt; <5%: 5 Pkt	ARR=14%; NNT= 7
Validität (Multiplikator bei uneingeschränkter Validität = 1.00, bei sehr stark eingeschränkter Validität = 0.10)	Subjektiv festgelegter Validitätsfaktor: 0.90
Erwarteter Gewinn Wirksamkeit (EGW) (Produkt aus ARR-Punkten und Validitätsfaktor)	ARR-Pkt x Validität = EGW: 7.6 x 0.90 = 6.84 Punkte
Leistungskategorie (L) (Multiplikator bei Verhinderung von: Tod = 1.0; Lebensbedrohlichem Ereignis = 0.99 – 0.90; erheblicher Behinderung = 0.89 – 0.70; Behinderung = 0.69 – 0.40; Befindlichkeit = 0.39 – 0.10)	Subjektiv festgelegte Leistungskategorie: 0.92
Erwarteter Mehrwert (Nutzen) (Produkt aus EGW und Faktor der Leistungskategorie)	EGW x L = Nutzen 6,84 x 0.92 = 6,29 Punkte
<b>Aus einer ARR=14%, hoher Validität (0.90) und bedeutender Leistungskategorie (0.92) ergibt sich ein konsensueller Nutzen von 6.29 Punkten</b>	

Die beschriebene Validität der wissenschaftlichen Arbeiten wird mit einem subjektiven Validitätsfaktor zwischen 1,00 und 0,10 bewertet. Der Faktor 1,00 bedeutet, dass an der Validität der wissenschaftlichen Arbeit kein Zweifel besteht. Bei erheblichem Zweifel wird ein Wert nahe 0,10 gewählt. Es ist zu bedenken, dass mit diesem Validitätsfaktor ein virtueller Wert des wissenschaftlichen Ergebnisses festgelegt wird, der aus zwei Komponenten besteht.

Die erste Komponente weist darauf hin, ob das beschriebene Ergebnis tatsächlich das bewirkt, was es zu bewirken vorgibt. Die zweite Komponente versucht die Wahrscheinlichkeit abzuschätzen, mit der das Ergebnis, das

unter den idealen Bedingungen einer kontrollierten klinischen Studie erhoben wurde, im klinischen Alltag, also unter realen Bedingungen, ebenfalls beobachtet werden kann. Diese Abschätzung ist bedeutend, weil für die Finanzierung einer Leistung deren Wirkungen auf die Versorgungsqualität, also für alle Perspektiven und nicht deren Effekt in klinischen Studien bedeutend ist. Da der Effekt auf die Versorgungsqualität bisher direkt nur in Ausnahmefällen gemessen werden kann, bleibt vorerst keine andere Möglichkeit, als ihn annähernd zu schätzen.

Durch Multiplikation des Punktwerts der Absoluten Risiko Reduktion mit dem Validitätsfaktor errechnet sich der Erwartete Gewinn an Wirksamkeit (EGW).

Anschließend ist dieser einer Leistungskategorie zuzuordnen, weil es für den Erwarteten Mehrwert nicht unerheblich ist, ob durch die eingeleitete Maßnahme ein Leben gerettet, die Arbeitsfähigkeit erhalten oder eine Befindlichkeitsstörung gelindert werden kann. Die Leistungskategorie wird analog zur Validität mit einem Faktor zwischen 1,00 und 0,10 ausgedrückt, wobei ein gerettetes Leben mit einem Faktor nahe 1,00 und die Linderung einer Befindlichkeitsstörung mit einem Faktor nahe 0,10 bewertet werden.

Aus der Multiplikation des Erwarteten Gewinns an Wirksamkeit (EGW) mit dem Faktor der Leistungskategorie ergibt sich rechnerisch der Erwartete Mehrwert, der als Nutzen definiert werden kann, wobei alle Perspektiven berücksichtigt werden.

Der Nutzen berechnet sich nach Porzsolt also aus (Punktwert absolute Risiko Reduktion) x (Validitätsfaktor) x (Faktor d. Leistungskategorie).

Dieser resultierende Nutzen dient den politischen Entscheidungsträgern als transparentes Maß (das Zustandekommen des Wertes ist nachvollziehbar) für den „solidarisch festgesetzten Wert“ einer Gesundheitsleistung. Anhand dieses Wertes können weitere Entscheidungen, wie z.B. die Finanzierung oder die Erstattungsfähigkeit getroffen werden.

Es stellt sich nun die Frage, ob ein solches Modell in der Zukunft realisierbar wäre.

Hervorgehoben werden sollte, dass alle errechneten Werte zunächst auf subjektiven Definitionen beruhen und validiert werden müssen. Man könnte auch weitere Kriterien berücksichtigen, allerdings bestünde die Gefahr, dass diese nicht die gleiche allgemeine Gültigkeit wie die drei Ausgangskriterien besitzen. Außerdem könnten sie die Verständlichkeit, die einfache Handhabung und Dokumentation sowie die geringe Störanfälligkeit des Systems negativ beeinträchtigen.

Ohne Dokumentation von subjektiven Entscheidungen lässt sich jedoch diese Validierung nicht durchführen. Dazu ist zunächst ein willkürlicher Entscheidungsmodus vorzugeben, der anhand der konkret getroffenen Entscheidung zu beurteilen und gegebenenfalls zu korrigieren ist.

Wenn diese Phase der Prüfung und Korrektur überwunden ist, könnte ein validiertes System zur Bewertung des Nutzens von Gesundheitsleistungen erwartet werden, das bisher in keinem anderen Gesundheitssystem zur Verfügung steht.

Der wesentliche Vorteil gegenüber bisherigen Bewertungssystemen, z.B. den publizierten League Tables (Schöffski, S. 367, 2000), besteht in dem Modus zur Ermittlung des Nutzens.

- Während bei den League Tables Experten den Nutzen subjektiv einschätzen und vorgeben, erfolgt der Entscheidungsprozess hier durch Gruppen, die als Vertreter partikularer Interessen benannt sind. Die Einbeziehung aller großen Interessensgruppen in den Entscheidungsprozess unterstützt den demokratischen Aspekt der Entscheidung, legitimiert die nachfolgenden Entscheidungen und ist deshalb als Vorteil zu werten.
- Die einzelnen Schritte des Entscheidungsprozesses werden transparent gemacht. Damit können Längs- und Querschnittsvergleiche durchgeführt werden.

- Falls erwünscht, lassen sich Ranglisten erstellen.
- Die Varianz der Entscheidungen wird messbar.
- Durch die Quantifizierung der Bewertung wird erstmals eine differenzierte Betrachtung und Diskussion des Nutzens von Gesundheitsleistungen möglich.

Die Arbeitsgruppe von Porzsolt unterbreitet zwar einen interessanten Ansatz, wie der Nutzen definiert werden könnte, dieser Ansatz unterstellt aber, dass zur Zeit Instrumente zur Verfügung stehen, die zumindest Teilaspekte des Nutzens valide messen können. Dies ist jedoch ansatzweise in Frage zu stellen. Insbesondere für die präferenzbasierte Lebensqualität lassen sich Verbesserungsvorschläge unterbreiten, was im Folgenden gezeigt wird.

### **5.3 Mögliche Beiträge der Gesundheitsökonomie zur Überwindung methodischer Probleme**

In Abschnitt 1.5.2.2 „Analyse des QALY-Konzepts“ wurden die methodischen Probleme der präferenzbasierten Messung der Lebensqualität dargestellt (Schöffski, S. 353ff, 2000). Im Folgenden werden mögliche Ansätze zur Überwindung dieser Hindernisse erörtert. Gegenstand der Analyse ist der EQ-5D (ehemals EuroQol), weil dieses Instrument das am häufigsten eingesetzte Instrument zur Messung der präferenzbasierten Lebensqualität darstellt.

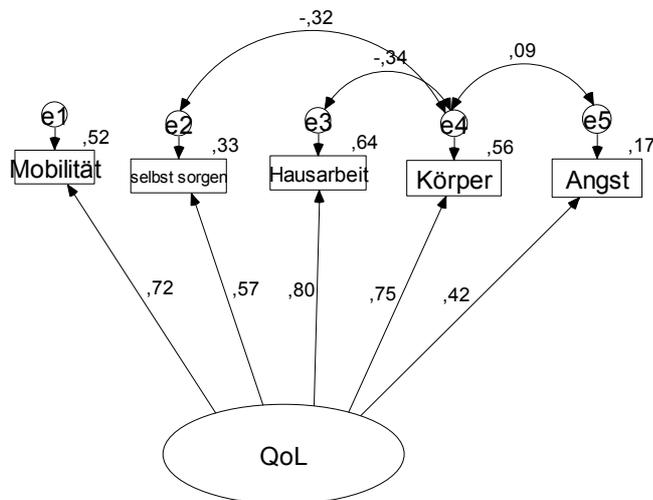
Der EQ-5D ist ein generisches Instrument zur Messung der Lebensqualität. Er stellt eine einfache Klassifizierung von Gesundheitszuständen dar und umfasst fünf Dimensionen mit je drei Antwortalternativen. Diese können m.E. (nach einer unveröffentlichten Analyse von Kohlmann) als intervallskaliert betrachtet werden. Allerdings wird jede Dimension nur durch ein einziges Item gemessen. Die Auswahl des deskriptiven Inhaltes basiert auf einem Review bereits existierender Instrumente. Der EQ-5D wurde bzgl. seiner Konstruktvalidität also nicht psychometrisch getestet.

Die Dimensionen des EQ-5D sind: Beweglichkeit/Mobilität, Für-sich-selbst-sorgen, allgemeine Tätigkeiten, Schmerzen/körperliche Beschwerden und Angst/Niedergeschlagenheit.

Mit drei Antwortalternativen bei drei Ausprägungen pro Item ergeben sich  $3^5$  Möglichkeiten also 243 Gesundheitszustände. Mit Hilfe des Time Trade Off Verfahrens wurden für ausgewählte Gesundheitszustände Präferenzen berechnet, auf deren Grundlage sich alle Präferenzwerte für alle Gesundheitszustände mittels einer Regression berechnen lassen.

Es stellen sich nun mehrere Fragen: „Ist die Konstruktvalidität des EQ-5D gegeben? Wie könnte man die Validität verbessern? Ist die Regressionsgleichung zur Schätzung der Präferenzwerte valide? Wie könnte man die Validität der Regressionsgleichung verbessern?“

Zur Analyse der Konstruktvalidität wurde eine konfirmatorische Faktorenanalyse gerechnet (eigene Berechnung). Der Datensatz hierbei stammt von einem Forschungsprojekt der Universität Leipzig (Hinz, 2006).



gfi=,998 agfi=,984  
df=2 p-value=,004 rmr=,001 chi-square=10,886 pclose=,528

Abbildung 5

Betrachtet man die Fitmaße des Modells, so kann man von einer mixed evidence sprechen. Sowohl GFI als auch AGFI entsprechen der Norm. Insbesondere der p-close weist einen guten Wert auf. Ebenso liegt der rmr im akzeptablen Bereich. Nur das Verhältnis chi-square zu df ist mit dem Wert

5 von dem Sollwert 3 nicht im akzeptablen Bereich. Im Großen und Ganzen kann das Modell aber mit dem vorhandenen Datensatz als valide angesehen werden (Reinicke, 2005).

Die Leistungsfähigkeit des EQ-5D ergibt sich aber auch aus dem Ausmaß der Aufklärung der einzelnen Dimensionen und der Stärke der Faktorladungen auf diese. Insbesondere die Dimension Angst/Niedergeschlagenheit weist mit einer aufgeklärten Varianz von nur 0,17 einen Wert auf, der als nicht akzeptabel zu bezeichnen ist. Ebenso liegt die Faktorladung von 0,42 außerhalb der Norm.

Der Grund dafür liegt m. E. darin, dass in einem Item zwei Fragen gestellt wurden. Damit verstießen die Konstrukteure des EQ-5D gegen elementare Regeln der Fragekonstruktion (Bühner, 2004). Es stellt sich die Frage, ob die Gesundheitszustände, die zur Messung der Präferenzen herangezogen wurden überhaupt valide sind und ob damit die Messergebnisse der Präferenzen selbst noch als valide bezeichnet werden können.

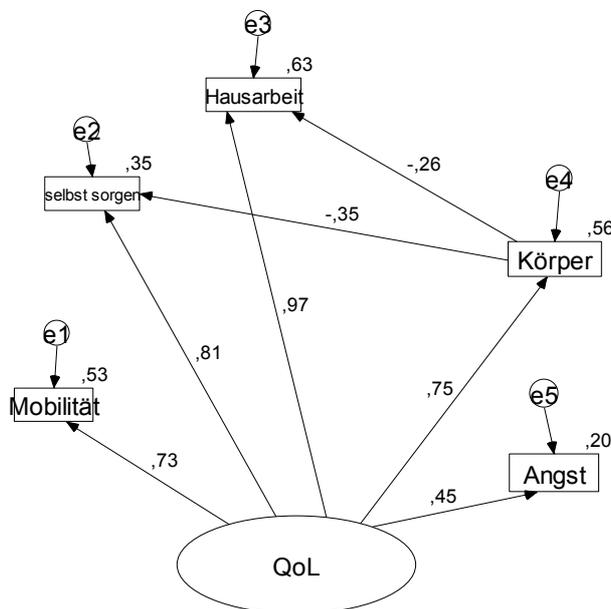
Die geringe Validität könnte auch erklären, warum die aufgeklärte Varianz der Regressionsgleichung mit 0,51 als suboptimal bezeichnet werden kann.

Ansatzpunkte zur Verbesserung des EQ-5D scheinen mir möglich. Aus test-theoretischen Überlegungen wäre es angebracht, die Frage zur Messung der Dimension „Angst/Niedergeschlagenheit“ neu zu formulieren. Ebenso wäre es dem messtheoretischen Standard entsprechend, die Dimensionen nicht nur mit einem Item, sondern mit zwei oder drei Items zu messen. Den damit verbundenen Zuwachs an Komplexität (es wären nicht mehr 35 sondern 315 Zustände der präferenzbasierten Lebensqualität zu messen) könnte man durch Methoden wie Conjoint oder Discrete Choice lösen. Im Conjoint Ansatz legt man zunächst die Eigenschaften und deren Ausprägungen fest. Aus der Menge der möglichen Gesundheitszustände extrahiert man mit Methoden wie dem Addelman-Plan oder dem lateinischen Quadrat Untermengen, die von Befragten in eine Rangordnung gebracht werden. Es gibt auch methodisch die Möglichkeit, direkt Nutzenurteile abzugeben (Hujer, 1996). Mit Hilfe einer Regressionsanalyse können dann die Präferenzen für alle Gesundheitszustände berechnet werden. In Punkt 1.5.2 wurde ausgeführt, dass

der Nutzen nicht durch Präferenzen, sondern durch Wahlentscheidungen definiert ist. Ist das Studiendesign einer Conjoint-Analyse entwickelt, so können Ja/Nein Entscheidungen zu Gesundheitszuständen mit Hilfe von Discrete Choice Modellen eingesetzt werden. Mit diesem Ansatz würde der Anforderung entsprochen, den Nutzen durch Wahlentscheidungen zu messen.

Trotzdem würde auch bei diesem Ansatz ein Schwachpunkt bestehen bleiben. Der regressionsanalytische Ansatz geht von der Unabhängigkeit der eingehenden Variablen bzw. Dimensionen der Lebensqualität aus. Dies ist aber weder empirisch noch von der Plausibilität her aufrechtzuerhalten. Es ist schwer vorstellbar, dass die Dimension „Schmerzen“ keinen Einfluss auf die Dimension „Beweglichkeit“ haben soll.

Eine Alternative wäre, nicht eine Regressionsgleichung einzusetzen, sondern ein Strukturgleichungsmodell, das die Multikollinearität berücksichtigt. Genau dies wurde mit dem vorliegenden Datensatz gerechnet.



gfi=,995 agfi=,974  
df=3 p-value=,000 rmr=,002 chi-square=26,017 pclose=,173

Abbildung 6

Sowohl GFI als auch AGFI weisen ebenso wie der *pclose* und der *rmsa* akzeptable Werte auf. Das Verhältnis *chi-square* zu *df* ist mit knapp 7 vom Grenzwert 3 entfernt. Man kann von einer *mixed evidence* sprechen. Sicherlich ließen sich die Fitmaße für ein, wie oben erwähnt, neues Modell noch verbessern. Mit der Berücksichtigung der Multikollinearität kann eine verbesserte aufgeklärte Varianz bei der Berechnung der Präferenzen von Gesundheitszuständen und bei Einsatz einer *Discrete Choice Analyse* der Berechnung der Nutzwerte erwartet werden.

Schlussfolgernd kann bestätigt werden, dass Ansatzpunkte zur Überwindung methodischer Probleme vorliegen.

In einem ersten Schritt müssten die Items des EQ-5D neu formuliert und mittels einer konfirmatorischen Faktorenanalyse validiert werden.

Die Präferenzen sollten mit Hilfe einer *Conjoint Analyse* und oder mit einer *Discrete-Choice Analyse* neu erhoben werden.

Die damit ermittelten Präferenz- und/oder Nutzenwerte sollten mit Hilfe eines Strukturgleichungsmodells neu geschätzt werden.

Sowohl die theoretische Fundierung als auch die empirischen Ergebnisse haben gezeigt, dass mit dieser Vorgehensweise validierte Ergebnisse der präferenzbasierten Lebensqualität zu erwarten sind. Weitere Forschung in diesem Bereich ist m. E. sinnvoll.

#### **5.4 Ausblick**

Vor dem Hintergrund der oben dargestellten Anwendungshindernisse und dem Blick auf die dargestellten Ansätze, diese Hindernisse zu überwinden sind für die Erschließung des Nutzenpotenzials gesundheitsökonomischer Studien und der damit verbundenen Optimierung der Allokation knapper Ressourcen für diejenigen, die für die politischen Rahmenbedingungen verantwortlich sind, m.E. folgende Aufgaben noch zu bewältigen:

- Entscheidungsinstanzen

Als Entscheidungsinstanz fungiert in Deutschland hauptsächlich der gemeinsame Bundesausschuss. Dieser ist sowohl was seine Legitimität als

auch seine Kompetenz angeht kritisch zu sehen. Es wäre wünschenswert eine Entscheidungsinstanz zu haben, deren Legitimität und Kompetenz außer Frage steht.

– Nutzendefinition

Der Terminus Technicus "Nutzen" ist weder im SGB V noch im AMG definiert. Ohne den Zielparameter "Nutzen" kann aber keine Optimierung der Allokation erfolgen.

– Datentransparenz

Die für eine Optimierung der Allokation notwendige Datentransparenz existiert nicht. Ansätze wie die Gesundheitskarte liegen vor und sollten umgesetzt werden.

– Validität der Allokationsinstrumente

Die Validität der QALY-Instrumente ist nicht gewährleistet und sollte deshalb in die Wege geleitet werden.

Entscheidungen im deutschen Gesundheitswesen werden in einzelnen Sektoren gefällt, die aber die gesamte Bevölkerung betreffen. Dies ist der Fall, weil einzelne Sektoren wie z.B. die Rentenversicherung aus gesundheitspolitischen Entscheidungen faktisch ausgeschlossen sind, oder ausgeschlossen werden, wie z.B. die Pharmaindustrie oder aber weil sie kein Stimmrecht haben, wie bspw. die Patientenvertreter. Daraus lässt sich die These ableiten:

1. Es muss ein von allen Seiten als legitim empfundenes und fachlich anerkanntes Entscheidungsgremium für das deutsche Gesundheitssystem geschaffen werden.

Wie oben ausführlich dargestellt ist der Zielparameter einer Optimierung einer Allokation im Gesetz zwar gefordert, aber nicht definiert. Daraus lässt sich eine zweite These ableiten.

2. Der Begriff „Nutzen“ muss durch das in These 1 geforderte Gremium definiert werden.

Die Allokation eines definierten Nutzens kann aber nur erfolgen, wenn die dafür benötigten Daten zur Verfügung stehen. Daraus lässt sich eine dritte These formulieren.

3. Die Transparenz und Verfügbarkeit der Daten, die für die Allokation benötigt werden, müssen gewährleistet sein.

Ein definierter Nutzenbegriff und verfügbare Daten allein reichen aber noch nicht aus, wenn nicht valide Instrumente zur Verfügung stehen. Daraus lässt sich eine vierte These aufstellen.

4. Für die Allokation knapper Ressourcen im Gesundheitswesen müssen valide Instrumente erstellt werden.

Würden diese Thesen umgesetzt werden, so käme es, und dafür sprechen die oben gezeigten Ergebnisse als auch die zugrunde liegende Theorie, zu einer Effizienzverbesserung. Schöffski drückte die mögliche Perspektive einer validen Optimierung der Allokation knapper Ressourcen so aus: „Gesundheitsökonomie mag zwar die Vertreibung aus dem Paradies der unbegrenzten Ressourcen erkennbar gemacht haben, ihre Anwendung brächte uns jedoch dem Paradies durch die effiziente Nutzung knapper Ressourcen wieder ein Stück näher“ (Schöffski, S. 495, 2000).

## 6. Literatur

### 1. Einführung

Drummond M. F.; O'Brien B.; Stoddart G. L.; Torrance G. W.  
Methods for the Evaluation of Health Care Programmes  
Oxford Medical Publications, Second edition 1997

Schöffski O.; Graf von der Schulenburg J.-M. (Hrsg)  
Gesundheitsökonomische Evaluationen  
Springer Verlag, zweite, vollständig neu überarbeitete Auflage 2000

Bestehorn et al.  
Repräsentative Studie zur Verteilung schizophrener Patienten auf medizini-  
sche Versorgungseinrichtungen in Deutschland  
Fortschritte der Neurologie Psychiatrie. 67 (1999)

Graf von der Schulenburg J.-M. et al.  
Untersuchungen zu den direkten und indirekten Kosten der Schizophrenie.  
Eine empirische Analyse  
Gesundheitsökonomie und Qualitätsmanagement 3 (1998) 81-87

Eichmann F.; Clouth J.  
„Budget Impact“ der atypischen Neuroleptika in Deutschland  
PharmacoEconomics, German Research Articles 2004; Bd. 2 Nr.1 S. 3-16

Neilson A. R.; Burchardi H., Chin C.; Clouth J.; Schneider H.; Angus D.  
Cost-Effectiveness Of Drotrecogin Alfa (Activated) for the Treatment of  
Severe Sepsis in Germany  
Journal of Critical Care, Vol. 18, No 4 (December) 2003: pp 217-227

Eichmann F, Reitberger U, Clouth J, Czekalla J  
GEO observational study: 24 months characteristics of socioeconomic and clinical status in schizophrenia patients treated with Olanzapine and Haloperidol in Germany Poster presentation ISPOR Florence November 2005.

Bendor J.

Titel: Herbert A. Simon: Political Scientist Annu. Rev. Polit. Sci. 2003.  
6:433-71

Varian, H. R.

Grundzüge der Mikroökonomik  
Verlag R. Oldenbourg, 5. Auflage 2001

Dörner D.; Kreuzig H. W.; Reither F.; Stäudel T. (Hrsg)  
Lohhausen: Vom Umgang mit Unbestimmtheit und Komplexität  
Verlag Hans Huber 1983

Dörner D.

Funkkolleg „Psycho-Biologie“  
Handeln – Problemlösen – Entscheiden  
Studienbegleitbrief 7, S. 97, 1987

Stäudel T.

Problemlösen, Emotionen und Kompetenz.  
1987 Regensburg

Dörner D.

Die Logik des Misslingens  
1989, Rowolth

Porzsolt F.; Ackermann M.; Amelung V.

Konsens zur Bewertung des intangiblen Nutzens von Gesundheitsleistungen

Gesundh ökon Qual manag 2006; 11: 353 – 364.

Thieme Verlag, 2006, Stuttgart

Krätke M.R.

Neoklassik als Weltreligion? Loccumer Initiative kritischer Wissenschaftle-

rinnen und Wissenschaftler (Hrsg): Die Illusion der neuen Freiheit Reali-

tätsverleugnung durch Wissenschaft (Kritische Interventionen 3). Offizin

Verlag 1999, Hannover

Diener E., Eunkook M. S.

National Differences in Subjective Well-Being.

In: Daniel Kahnemann, Ed Diener und Norbert Schwarz (Hrsg.). Well-

Being: The Foundations of Hedonic Psychology. New York, 1999, Russel

Sage Foundation: 434-450

Kruse A.

Geschichte der volkswirtschaftlichen Theorien. Duncker & Humblot,

Berlin 1959; S. 225

Stutzer A.

Eine ökonomische Analyse menschlichen Wohlbefindens.

Zürich, 2001; 20-26

v.Uexküll T.; Adler R. H.; Hermann J. M.

Psychosomatische Medizin.

6.Aufl. 2003, Urban & Fischer

Zentner A., Velasco-Garrido M., Busse R.

Methoden zur vergleichenden Bewertung pharmazeutischer Produkte.

In: Deutsche Agentur für Health Technology Assessment des Deutschen Instituts für Medizinische Dokumentation und Information (Hrsg): Informationssystem Health Technology Assessment (HTA) in der Bundesrepublik Deutschland. 1. Auflage 2005 DIMDI, Köln

Ruß H.; Clouth J.; Porzsolt F.

Zur Theorie der Brückenprinzipien – Gesellschaftspolitische Kommentare

Okt. 2005, 46. Jg., Nr. 10

Hume D.

A Treatise of Human Nature, Book III [1740],

ed. L.A. Selby-Bigge, Oxford 1975

Böhmer S.; Kohlmann T.

Verfahren zur Bewertung von Gesundheitszuständen und Lebensqualität

In: Ravens-Sieberer U.; Cieza A. (Hrsg)

Lebensqualität und Gesundheitsökonomie in der Medizin

Ecomed 2000

Hujer R.; Grammig J.; Fryns H.; Herterich R.

Preisfindung und optimale Marketingstrategien für neue pharmazeutische Produkte

zfbz Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung

3/1996 S. 219-232

Monzini et al.

Application of conjoint analysis to elicit preferences in hemophilia care: the

CONAN study group

ISPOR Poster presentation 2003 Barcelona

Backhouse M. et al.

The use of discrete choice analysis to assist with the interpretation of quality of life scores

ISPOR Poster presentation 2003 Barcelona

Chassany O. et al.

Patient-reported outcomes: the example of health-related quality of life – a european guidance document for the improved integration of health-related quality of life assessment in the drug regulatory process

Drug Information Journal 36: 209 - 238; 2002

Bühner M.

Einführung in die Test- und Fragebogenkonstruktion

Pearson Studium 2004

Schumacher J.; Klaiberg A.; Brähler E.

Diagnostische Verfahren zu Lebensqualität und Wohlbefinden

Hogrefe 2003

Happich M.

Die Bewertung von Gesundheit

Peter Lang Europäischer Verlag der Wissenschaften 2003

Ruß H. G.

Empirisches Wissen und Moralkonstruktion Eine Untersuchung zur Möglichkeit und Reichweite von Brückenprinzipien in der Natur- und Bioethik

2002 Dr. Hänsel Hohenhause AG

Albert H.

Treatise on Critical Reason

Princeton 1985

## 2. Begründung der ausgewählten Krankheiten

Clouth J.; Uber A.; Graf von der Schulenburg J. M.

Kosten der Schizophrenie Untertitel: Strukturgleichungsmodelle als Instrument der Entscheidungshilfe bei gesundheitsökonomischen Fragestellungen  
Journal: Gesund ökon Qual manag 2003:8:120-124

Schmid A.; Burchardi H.; Clouth J.; Schneider H.

Burden of Illness imposed by severe Sepsis in Germany  
Eur. J Health Econom 2002: 3:77-82 2002

Brunkhorst F. M. et al.

For the German Competence Network Sepsis:

Epidemiology of severe sepsis and septic shock in Germany – results from the German "Prevalence" Study;

Abstract 25th Congress of Intensive Care and Emergency Medicine, Critical Care, 9, Suppl. 1, March 2005; P196, 2005 (DOI 10.1186/cc3259).

### 3. Begründung der eingereichten Arbeiten

Ruß H.; Clouth J.; Porzsolt F.

Zur Theorie der Brückenprinzipien – Gesellschaftspolitische Kommentare

Okt. 2005, 46. Jg., Nr. 10

Zentner A.; Velasco-Garrido M.; Busse R.

Methoden zur vergleichenden Bewertung pharmazeutischer Produkte.

In: Deutsche Agentur für Health Technology Assessment des Deutschen Instituts für Medizinische Dokumentation und Information (Hrsg): Informationssystem Health Technology Assessment (HTA) in der Bundesrepublik Deutschland. 1. Auflage 2005 DIMDI, Köln

Raffée H.

Grundprobleme der Betriebswirtschaftslehre

UTB Vandenhoeck 1974

Raffée H.; Abel B. (hrsg)

Wissenschaftstheoretische Grundfragen der Wirtschaftswissenschaften

Verlag Vahlen 1979

Graf von der Schulenburg J.-M. et al.

Untersuchungen zu den direkten und indirekten Kosten der Schizophrenie.

Eine empirische Analyse

Gesundheitsökonomie und Qualitätsmanagement 3 (1998) 81-87

Häfner H.

Das Rätsel Schizophrenie: Eine Krankheit wird entschlüsselt,

Verlag Beck, 2000

Spannheimer A.; Clouth J.

Pharmacoeconomic Evaluation of the Treatment of Schizophrenia in Germany: A Comparison of Olanzapine and Haloperidol

Poster Presentation APA Washington 1999

[www.lilly-pharma.de/Lilly\\_Quality\\_of\\_Life\\_Preis.587.0.html](http://www.lilly-pharma.de/Lilly_Quality_of_Life_Preis.587.0.html)

## 4. Literatur zu den Arbeiten

### *4.1 Wissenschaftstheoretische Überlegungen zur Werturteilsproblematik der so genannten Vierten Hürde*

Raffée, Hans, Grundprobleme der Betriebswirtschaftslehre, UTB Vandenhoeck, Göttingen, 1974

Albert, Hans, Konstruktion und Kritik, Aufsätze zur Philosophie des kritischen Rationalismus, Hamburg, 1972

Seiffert, Helmut, Einführung in die Wissenschaftstheorie 1, C.H. Beck, München, 1975 Raffée Hans; Abel Bodo, Wissenschaftstheoretische Grundfragen der Wirtschaftswissenschaften, Verlag Vahlen, München, 1997

Ulla Schmidt bei der Veranstaltung der IG Metall in Dortmund, BMGS Reden 05.09.2002

Deckert C., Höffler J., Kortmann J., Linden M., Roth G.-D., Struck M., Clouth J., Czekalla J., Spannheimer A., Wieseler B., Schlereth T.  
Kostenanalyse der Schizophrenie-Behandlung in Deutschland  
Vergleich von Olanzapin, Risperidon und Haloperidol mit Hilfe eines klinischen Entscheidungsmodells  
Gesundh ökon Qual manag 2001; 6: 161-166, Georg Thieme Verlag Stuttgart – New York, ISSN 1432-2625

Foster, Rachel H et al., Olanzapine. A Pharmacoeconomic Review of its use in Schizophrenia, Pharmacoeconomics 1999 Jun;15 (6): 611-640, 1999

Schnabel, Clouth, Kosten der Frühverrentung am Beispiel der Schizophrenie, Gesundheitsökonomie und Qualitätsmanagement, Heft 7, S. 381-388, 2002

Dörner, Dietrich, Logik des Misslingens, Rowohlt, Reinbek, 1989

Dörner, Dietrich, Lohhausen, Verlag Hans  
Huber, Bern Stuttgart Wien, 1983

Neilson, Aileen et al., Poster Presentation; ESICM Barcelona Oct. 2002

Backhaus, Klaus et al., Multivariate Analysemethoden, Springer, Berlin,  
Heidelberg, New York, 2000

Sterman, John D., Business Dynamics, Irwin McGraw-Hill, Boston, 2000

Varian, Hal R., Grundzüge der Mikroökonomik, Oldenburg, 2000, 5. Auflage

Mullins, Unveröffentlichter Vortrag auf ISPOR Nov. 2002 in Rotterdam

Wever, Leon, Unveröffentlichter Vortrag auf ISPOR Nov. 2002 in Rotterdam

#### 4.2 Kosten der Frühverrentung am Beispiel der Schizophrenie

- Bestehorn, M., B. Tischer, P. Glaser, O. Mast, D. Schmidt (1999), „Repräsentative Studie zur Verteilung schizophrener Patienten auf medizinische Versorgungseinrichtungen in Deutschland“, Fortschritte Neurologisch Psychiatrie, No. 67 (1999) S487-492.
- Eichmann F, Reitberger U, Spannheimer A, et al. Patient and disease characteristics in schizophrenia outpatients and inpatients treated with olanzapine and haloperidol in Germany: Results from the GEO observational study (poster presentation). ISPOR 5th Annual European Congress; 2002 Nov 3-5; Rotterdam
- Greiner, W. (2000), Die Berechnung von Kosten und Nutzen im Gesundheitswesen, in: O. Schöffski und J.M. Graf von der Schulenburg (Hrsg.), Gesundheitsökonomische Evaluation, Springer Verlag.
- Häfner, Heinz (2000), „Das Rätsel Schizophrenie: Eine Krankheit wird entschlüsselt“, Verlag: Beck.
- Lichtenberg, Frank (2001), „The Benefits and Costs of Newer Drugs: Evidence from the 1996 Medical Expenditure Panel Survey“, NBER Working Paper No. 8147.
- Kissling, W. et al. (1999), „Die direkten und indirekten Kosten der Schizophrenie“, Fortschritte Neurologische Psychiatrie, No. 67 (1999), S. 29-36.
- Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen (2001), „Bedarfsgerechtigkeit und Wirtschaftlichkeit Bd.III – Über-, Unter- und Fehlversorgung“, Berlin.
- v.d. Schulenburg, J.-M. Graf v.d., A. Uber, K. Höffler, U. Trenckmann, W. Kissling, U. Seemann, P. Müller und E. Rüther (1998), „Untersuchungen zu den direkten und indirekten Kosten der Schizophrenie“, Gesundheitsökonomie und Qualitätsmanagement, 1998, 3, S.81-87.
- Statistisches Bundesamt (Hrsg.) (1998), „Gesundheitsbericht für Deutschland“, Stuttgart: Verlag Metzler-Poeschel.

- Über, A. (2000), „Die Kosten der Schizophrenie. Eine empirische Analyse aus versicherungsökonomischer und gesamtwirtschaftlicher Perspektive. Versicherungswissenschaft in Hannover, Hannoveraner Reihe, Band 11, Verlag Versicherungswirtschaft Karlsruhe.
- VDR – Verband Deutscher Rentenversicherungsträger (2001), „Empfehlungen für die sozialmedizinische Beurteilung psychischer Störungen“, DRV-Schriften, Band 30, Oktober 2001, Frankfurt am Main.
- VDR – Verband Deutscher Rentenversicherungsträger (versch. Jahrgänge), „VDR-Statistik Rentenzugang“, Frankfurt am Main.
- VDR – Verband Deutscher Rentenversicherungsträger (2001), „Rentenversicherung in Zeitreihen“, Frankfurt am Main.

### 4.3 Kostenanalyse der Schizophrenie-Behandlung in Deutschland

- 1 Almond, S., O. O'Donnell: Cost analysis of the treatment of schizophrenia in the UK: A comparison of Olanzapine and Haloperidol. *Pharmacoeconomics* 13 (1998), 575-588.
- 2 Cohen LJ, Test MA, Brown RL. Suicide and schizophrenia: data from a prospective community treatment study. *AM J Psychiatry* 1990; 147:5 .
- 3 Davies A., P.C. Langley, N.A. Keks, S.V. Catts, T. Lambert, Schweitzer I.: Risperidone versus haloperidol: II. Cost-effectiveness. *Clin. Ther.* 20 (1998), 196-213.
- 4 Edgell E.T., C.S. Palmer, D.A. Revicki et al.: A cost effectiveness clinical decision analysis for schizophrenia: results across four countries. Poster presentation at the 11th European College of Neuro-Psychopharmacology (ECNP) Congress, Paris, 1998.
- 5 Foster, R.H., K.L. Goa: Risperidone: a pharmacoeconomic review of its use in schizophrenia. *Pharmacoeconomics* 14 (1998), 97-133.
- 6 Gilbert P.L., M.J. Harris, L.A. McAdams, D.V. Jeste: Neuroleptic withdrawal in schizophrenic patients. *Arch. Gen. Psychiatry* (1995), 52.
- 7 Hamilton S.H., L.A. Genduso, D.A. Revicki: Medical resource use and work and social outcomes for olanzapine compared with haloperidol in the treatment of schizophrenia and other psychotic disorders. *Schizophrenia reserach* 29 (1998), 148-149.
- 8 Kay, S.R., L.A. Opler, A. Fiszbein: Positive and Negative Syndrome Scale Manual. North Tonawanda, NY: Multi-Health Systems, 1986.
- 9 Lehman A.F., J.W. Thompson, L.B. Dixon, F.E. Scott: Schizophrenia: treatment outcomes research. *Schiz. Bull.* 21 (1995) 561-566.
- 10 Mahmoud R.A., L.M. Engelhart, G. Oster et al.: Risperidone vs. Conventional antipsychotics: a prospective randomized naturalistic effectiveness trial of outcomes in chronic schizophrenia [poster]. 36th Annual Meeting of the American College of Neuropsychopharmacology: 1997 Dec 8-12, Kameula, Hawaii.

- 11 Naber D.: A self-rating to measure subjective effects of neuroleptic drugs, relationships to objective psychopathology, quality of life, compliance and other clinical variables. *Int. Clin. Psychopharmacol.* 10 Suppl. 3 (1995), 133-138.
- 12 Nigthengale B.S., L. Garrett, S. Waugh et al.: Economic outcomes associated with the use of risperidone in a naturalistic group practice setting. *Am. J. Manage. Care* 4 (1998), 360-366.
- 13 Palmer CS, D.A. Revicki, L.A. Genduso, S.H. Hamilton, R.E. Brown: A cost effectiveness clinical decision analysis model for schizophrenia. *The American Journal of Managed Care* 43 (1998), 345-355.
- 14 von der Schulenburg J.M., A. Uber, J. Höffler, U. Trenckmann, W. Kissling, U. Seemann, P. Müller, E. Rütger: Untersuchungen zu den direkten und indirekten Kosten der Schizophrenie. Eine empirische Analyse. *Gesundh. ökonom. Qual. manag.* 3 (1998), 81-87.
- 15 Tollefson G.D.; C.M. Beasley; P.V. Tran, J.S. Street; J.A. Krueger; R.N. Tamura, K.A. Graffeo, M.E. Thieme: Olanzapine versus haloperidol in the treatment of schizophrenia and schizoaffective and schizophreniform disorders: results of an international collaborative trial. *Am. J. Psychiatry* 154(4) (1997), 457-65.
- 16 Torrey E.F.: Prevalence studies in schizophrenia. *Br. J. Psychiatry* 150 (1987), 598-608.
- 17 Tran P.V.; S.H. Hamilton, A.J. Kuntz, J.H. Potvin, S.W. Andersen, C. Beasley; G.D. Tollefson: Double-blind comparison of olanzapine versus risperidone in the treatment of schizophrenia and other psychotic disorders. *J. Clin. Psychopharmacol* 17(5) (1997), 407-418.
- 18 Tran P.V.; M.A. Dellva; G.D. Tollefson; C.M. Beasley; J.H. Potvin; G.M. Kiesler: Extrapyramidal symptoms and tolerability of olanzapine versus haloperidol in the acute treatment of schizophrenia. *J. Clin. Psychiatry*, 58 (1997), 205-211.

*4.4 Constructing an Index for the Subjective Well-being under Neuroleptics Scale (SWN), Short Form: Applying Structural Equation Modeling for Testing Reliability and Validity of the Index*

*References*

1. Social Research Methodology (SRM) Database.  
<http://www.niwi.knaw.nl/srmonline/index.htm>. Last accessed May 15, 2003, 5:30 pm.
2. Keller SD, Ware JE Jr, Bentler PM, et al. Use of structural equation modeling to test the construct validity of the SF-36 Health Survey in ten countries: Results from the IQOLA Project. International Quality of Life Assessment. *J Clin Epidemiol* 1998; 51(11): 1179–1188.
3. Carlsson M, Hamrin E. Evaluation of the life satisfaction questionnaire (LSQ) using structural equation modeling (SEM). *Qual Life Res* 2002; 11(5): 415–425.
4. Saris WE, Scherpenzeel A, Veenhoven R. *A comparative study of satisfaction with life in Europe*. Budapest: Eötvös University Press, 1996.
5. Payck TR. *Checkliste Psychiatrie*. (In German) Stuttgart: Thieme, 1988.
6. NIMH (National Institute of Mental Health).  
<http://www.nimh.nih.gov/publicat/schizoph.htm>. Last accessed February 12, 2003, 4:34 pm.
7. Foster RH, Goa KL. Olanzapine: a pharmacoeconomic review of its use in schizophrenia. *Pharmacoeconomics* 1999; 15(6): 611–640.
8. Goldner EM, Hsu L, Waraich P, Somers JM. Prevalence and incidence studies of schizophrenic disorders: a systematic review of the literature. *Can J Psychiatry*. 2002; 47(9): 833-843.
9. Karow A, Naber D. Subjective well-being and quality of life under atypical antipsychotic treatment. *Psychopharmacology (Berl)* 2002; 162(1): 3–10.
10. Naber D, Moritz S, Lambert M, et al. Improvement of schizophrenic patients subjective well-being under atypical antipsychotic drugs. *Schizophr Res* 2001; 50(1–2): 79–88.

11. De Haan L, Weisfelt M, Dingemans PM, et al. Psychometrics properties of the Subjective Well-Being Under Neuroleptics scale and the Subjective Deficit Syndrome Scale. *Psychopharmacology (Berl)* 2002; 162(1): 24–28.
12. Bollen KA. *Structural equations with latent variables*. New York: Wiley-Interscience, 1989.
13. Jöreskog KG. Structural analysis of covariance and correlation matrices; *Psychometrika* 1978; 43: 443–472.
14. Byrne BM. *Structural Equation Modeling with AMOS. Basic Concepts, Applications and Programming (Multivariate Applications Series)*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 2001.
15. Arbuckle J (ed.) *Amos 4.0 User's Guide*. Chicago: SPPS. 1999.
16. Hu, L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* 1999; 6: 1-55.
17. Marsh HW, Hau KT, Wen Z (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralising Hu & Bentler's (1999) findings. *Structural Equation Modeling* 2004; 11: 320-341.
18. Harman, HH. *Modern Factor Analysis, 3rd edition*. Chicago: University of Chicago Press, 1976.
19. Hoogland JJ, Boomsma A. Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research* 1998; 26: 329–367.
20. Cronbach LJ, Meehl PE. *Construct Validity in Psychological Tests*. *Psychol Bull* 1955; 52: 281–302
21. Raykov T. *Estimation of composite reliability for congeneric measures*. *Applied Psychological Measurement* 1997; 21: 173–184.
22. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS (Eds.). *Testing Structural Equation Models*. Newbury Park, California: Sage, 1993: 136–162.

23. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F (Eds.). *Proceedings of the 2nd International Symposium on Information Theory*. Budapest: Akademiai Kiado, 1973: 267–281.
24. Akaike H. Factor analysis and AIC. *Psychometrika* 1987; 52: 317-332.
25. Jöreskog KG, Sörbom D. *LISREL 8 User's Reference Guide, 2nd edition*. Lincolnwood, IL: Scientific Software, 2001.
26. Tanaka JS, Huba GJ. A fit index for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology* 1985; 38, 197–201.

## 5. Schlussfolgerungen

Varian H. R.

Grundzüge der Mikroökonomik,  
Oldenburg, 2000, 5. Auflage

Pindyck R. S.; Rubinfeld D. L.

Mikroökonomie

6. Auflage 2005 Pearson Studium S409.

Schnabel R.; Friedel H.; Erfurth A.; Angermayer M.; Clouth J.;

Eichmann F.

Predisposing Factors for Early Retirement in Patients with Schizophrenia in  
Germany

European Journal of Health Economics

in review

Clouth J.; Knoll S.; Maier F.;

The costs of Schizophrenia in Germany - a System Dynamics modelling  
approach

Zeitschrift in review

Eichmann F.; Reitberger U.; Spannheimer A.; et al.

Patient and disease characteristics in schizophrenia outpatients and  
inpatients treated with olanzapine and haloperidol in Germany:

Results from the GEO observational study (poster presentation).

ISPOR 5th Annual European Congress; 2002 Nov 3-5; Rotterdam

SGB V

Leitfadenverlag Sudholt; Auflage: 11, März 2005

Schütze Brief Nr. 65/2006, Seite 6

Determinanten der Akzeptanz der Integrierten Versorgung – Empirische  
Ergebnisse

Häussler B.; Berger U.

Bedingungen für effektive Disease-Management-Programme  
Nomos Verlagsgesellschaft 2005 S. 50ff.

[http://www.g-ba.de/cms/front\\_content.php?idcat=212](http://www.g-ba.de/cms/front_content.php?idcat=212)

Lütge G.

Die Zeit Nr. 16, 2006, S. 25

Deckert et al.

Kostenanalyse der Schizophreniebehandlung in Deutschland – Vergleich  
von Olanzapin, Risperidon und Haloperidon mit Hilfe eines klinischen  
Entscheidungsmodells

Gesundh ökon Qual manag 2001; 6: 161-166

Georg Thieme Verlag Stuttgart – New York

Clouth et al.

The Costs of Schizophrenia in Germany – a systems dynamics modelling  
approach

in review

[www.IQWIG.de](http://www.IQWIG.de)

Porzsolt F.; Ackermann M.; Amelung V.

Konsens zur Bewertung des intangiblen Nutzens von Gesundheitsleistungen  
Gesundh ökon Qual manag 2006; 11: 353 – 364.

Thieme Verlag, 2006, Stuttgart

Schöffski O.; Graf von der Schulenburg J. M. (Hrsg)

Gesundheitsökonomische Evaluationen

Springer Verlag, zweite, vollständig neu überarbeitete Auflage 2000

Hinz A.; Klaiberg A.; Brähler E.; König H.

Der Lebensqualitätsfragebogen EQ-5D: Modelle und Normwerte für die  
Allgemeinbevölkerung

Psychotherapie, Psychosomatik, Medizinische Psychologie.

S. 56(2):42-8, 2006 Feb.

Reinecke J.

Strukturgleichungsmodelle in den Sozialwissenschaften

2005, Oldenburg Verlag

Bühner M.

Einführung in die Test- und Fragebogenkonstruktion

Pearson Studium 2004

Hujer R.; Grammig J.; Fryns H.; Herterich R.

Preisfindung und optimale Marketingstrategien für neue pharmazeutische  
Produkte

zfbz Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung

3/1996 S. 219-232