



Schweizerisches Gesundheitsobservatorium
Observatoire suisse de la santé
Osservatorio svizzero della salute
Swiss Health Observatory

Nachfrage, Inanspruchnahme, Bedarf und Angebotsinduzierung in der ambulanten medizinischen Versorgung

**Besoins, recours, demande réelle et demande
induite dans les soins médicaux ambulatoires**



Eine methodenkritische Literaturübersicht

Revue critique de la littérature

**Jürg Guggisberg und Stefan Spycher
Büro BASS**

Forschungsprotokoll 3

Das Schweizerische Gesundheitsobservatorium (Obsan) ist eine Organisationseinheit des Bundesamtes für Statistik, die im Rahmen des Projektes Nationale Gesundheitspolitik entstanden ist und von Bund und Kantonen einen Leistungsauftrag erhält. Das Gesundheitsobservatorium analysiert die vorhandenen Gesundheitsinformationen in der Schweiz. Es unterstützt Bund, Kantone und weitere Institutionen im Gesundheitswesen bei ihrer Planung, ihrer Entscheidungsfindung und in ihrem Handeln. Weitere Informationen sind zu finden auf www.obsan.ch und auf www.nationalegesundheit.ch

Die Forschungsprotokolle des Obsan sind Fachberichte, die nur in kleiner Auflage produziert werden und auf Anfrage an Interessierte abgegeben werden. Der Inhalt der Forschungsprotokolle unterliegt der redaktionellen Verantwortung der Autoren und Autorinnen.

Jürg Guggisberg und Stefan Spycher:
Nachfrage, Inanspruchnahme, Bedarf und
Angebotsinduzierung in der ambulanten
medizinischen Versorgung

Forschungsprotokoll des Obsan Nr. 3
Juli 2005

© Schweizerisches Gesundheitsobservatorium
Espace de l'Europe 10
CH-2010 Neuchâtel
www.obsan.ch

Projektleitung: Hélène Jaccard Ruedin

Anschrift der Autoren:
Büro für arbeits- und sozialpolitische Studien BASS
Konsumstrasse 20
CH-3007 Bern
juerg.guggisberg@buerobass.ch; stefan.spycher@buerobass.ch
Tel. 031 380 60 80

Inhaltsverzeichnis

Zusammenfassung	1
Résumé	11
1 Einleitung, Ausgangslage, Fragestellungen	21
Teil 1: Literaturanalyse	23
2 Theorien, Begriffe und Konzepte	23
2.1 Modelle der Gesundheitsökonomie	23
2.1.1 Die Nachfrageseite	23
2.1.2 Die Angebotsseite – angebotsinduzierte Nachfrage	25
2.2 Sozialwissenschaftliche Verhaltensmodelle	31
3 Probleme der empirischen Forschung	35
3.1 Theoretische und begriffliche Probleme	35
3.2 Methodische Probleme	36
3.2.1 Identifikation eines kausalen Effektes	36
3.2.2 Trennung der Effekte von patienteninduzierter Nachfrage und arztinduzierter Inanspruchnahme	37
3.2.3 Einfluss der Modellspezifikation auf die Resultate	37
3.2.4 Messfehler und andere datenspezifische Probleme	38
3.2.5 Unterschiedliche Schätzverfahren	38
3.3 Vergleichbarkeit der Resultate	40
3.4 Bestimmung des Ausmasses angebotsinduzierter Nachfrage	40
3.5 Zusammenfassung	42
4 Stand der Forschung	43
4.1 Internationale Untersuchungen	43
4.1.1 Determinanten der Inanspruchnahme	43
4.1.2 Überprüfung des Grossman-Modells	45
4.1.3 Arztinduzierte Nachfrage	46
4.1.4 Nachfragesteuerung, Anreizsysteme und Moral Hazard	48
4.2 Untersuchungen aus der Schweiz	49
4.2.1 Determinanten der Inanspruchnahme	50
4.2.2 Überprüfung des Grossman-Modells	51
4.2.3 Arztinduzierte Nachfrage	51
4.2.4 Nachfragesteuerung, Anreizsysteme und Moral Hazard	53
4.3 Zusammenfassung	55

5	Von der Erklärung der Inanspruchnahme zum Bedarf	59
5.1	Von der individuellen auf eine aggregierte, regionale Ebene – Die Bildung von medizinischen Versorgungsgebieten	59
5.2	Von der Bedarfsanalyse zur Bedarfsprognose	61
6	Die Bedeutung der Literaturanalyse	63
Teil 2: Konkretes Vorgehen zur Bestimmung des Bedarfs		64
7	Drei Konzepte und Wege zur Bestimmung des Bedarfs	64
7.1	Pro-Kopf-Bedarf von verschiedenen Inanspruchnahmegruppen (Variante 1)	65
7.2	Multivariate Regression auf der Ebene medizinischer «Versorgungsgebiete» (Variante 2)	71
7.3	Mehrebenenanalyse mit Individualdaten als Basis (Variante 3)	74
Literaturverzeichnis		77

Zusammenfassung

Die vorliegende Untersuchung ist das Ergebnis einer Grundlagenarbeit für ein Projekt, das das Schweizerische Gesundheitsobservatorium (Obsan) zusammen mit weiteren Partner/innen im Bereich der «Démographie Médicale» durchzuführen plant. Unter «Démographie Médicale» wird die Analyse des Standes und der Entwicklung des ambulanten medizinischen Angebotes und der entsprechenden Nachfrage verstanden. Auf der Seite der Nachfrage steht in der politischen Diskussion immer wieder die Frage der Existenz und des Ausmasses der durch die Ärzt/innen ausgelösten Nachfrage (sog. Angebotsinduzierung) im Vordergrund. Das Obsan beauftragte daher das Büro BASS, verschiedene Aspekte der Nachfrageseite mit einer Literaturanalyse zu klären.

Fragestellungen

Konkret sollten folgende Fragen beantwortet werden:

- Welche empirischen Ergebnisse zur Nachfrageschätzung liegen für die Schweiz vor? Welche Konsequenzen sind daraus für zukünftige Schätzungen abzuleiten?
- Wie kann methodisch vorgegangen werden, um die Inanspruchnahme nach ambulanten Leistungen zu bestimmen?
- Wie wird die angebotsinduzierte Inanspruchnahme berücksichtigt?
- Welche Vor- und Nachteile haben die einzelnen Methoden mit Blick auf die im Rahmen des Gesamtprojektes der «Démographie Médicale» in Zukunft geplante Nachfrageanalyse?

Methodisches Vorgehen

Bei der Literaturanalyse wurde in Absprache mit dem Auftraggeber pragmatisch vorgegangen. Ausgegangen wurde von den bereits bekannten wissenschaftlichen Arbeiten im Gebiet. Ergänzend wurde eine Literaturrecherche auf dem Internet und in der ökonomischen Datenbank ECONLIT durchgeführt. Im ersten Teil der Arbeit wurde die Literatur ausgewertet. Im zweiten Teil wurden anschliessend darauf aufbauend drei Varianten zur konkreten Bestimmung des regionalen Bedarfs an medizinischen-ambulanten Leistungen entwickelt.

Teil 1: Ergebnisse der Literaturanalyse

Zwei Theorieansätze

Die Inanspruchnahme medizinischer Leistungen wird bereits seit den frühen 1960er Jahren empirisch untersucht. Nachdem eine rein demographisch-deskriptive Betrachtung der Inanspruchnahme als unbefriedigende Informationsbasis angesehen wurde, entwickelten die **sozialwissenschaftlich** orientierten Disziplinen **Verhaltensmodelle** der Inanspruchnahme medizinischer Versorgung. Exemplarisch wird das Verhaltensmodell nach Andersen vorgestellt, weil es sich dabei um ein sehr umfassendes Modell handelt. Parallel zu der sozialwissenschaftlichen Perspektive entwickelte sich in der **Gesundheitsökonomie** ein eigener Theoriezweig zur Inanspruchnahme medizinischer Versorgung. Dieser befasst sich u.a. mit Problemen der Steuerung, die durch die besonderen Eigenschaften der Gesundheit und seiner Produktion hervorgerufen werden.

Die Angebots- und Nachfrageseite wird im Rahmen der Gesundheitsökonomie durch Modelle beschrieben, die das Verhalten der Akteurinnen und Akteure auf ökonomisch rationale Entscheidungen

zurückführen. Sozialwissenschaftliche Modelle versuchen, neben dem ökonomisch rationalen Verhalten auch soziologische, psychologische und morbiditätsbezogene Kontexte zu berücksichtigen.

Sozialwissenschaftliche und gesundheitsökonomische Modelle überschneiden sich heute jedoch in weiten Bereichen zumindest bezüglich ihrer empirischen Umsetzung bzw. Überprüfung. Dies hängt u.a. damit zusammen, dass sich die «traditionellen» sozialwissenschaftlichen und auch ökonomischen Verhaltensanalysen immer stärker in Richtung multivariater Modelle entwickelten. Sowohl sozialwissenschaftliche als auch ökonomisch geprägte Verhaltensanalysen versuchen, das Inanspruchnahmegeschehen mit Hilfe von Individualdaten (bspw. Alter, Geschlecht, Einkommen, etc.) welche durch strukturelle Variablen (bspw. Stadt-Land, ÄrztlInnendichte, Sprachregion, etc.) ergänzt werden, zu modellieren.

Das Grossman-Modell

Das bekannteste Modell aus der Gesundheitsökonomie zur Schätzung der Nachfrage ist das sog. Grossman-Modell. In dem Modell wird erstens die Nachfrage nach medizinischen Leistungen als abgeleitete Nachfrage interpretiert, d.h. das Individuum fragt Gesundheit und nicht medizinische Leistungen per se nach. Zweitens ist der Gesundheitskapitalbestand abhängig von der Inanspruchnahme medizinischer Leistungen und einer Abschreibungsrate, die explizit als Funktion des Alters (des Gesundheitskapitalbestandes) und der Intensität seiner Nutzung definiert ist. Über die Abschreibungsrate können auch Merkmale wie das Gesundheitsverhalten oder das Bildungsniveau in das Modell integriert werden. Ein Zusammenhang zwischen dem Erwerbseinkommen und dem Gesundheitszustand wird dadurch hergestellt, indem einem höheren Gesundheitskapitalbestand gleichzeitig konsumtiven wie investiven Nutzen zugeschrieben wird.

Viele der Implikationen des Grossman-Modells sind von der empirischen Evidenz widerlegt worden. Insbesondere müssten gemäss dem Modell der Gesundheitszustand und die Nachfrage nach medizinischen Leistungen positiv korreliert sein, was sich in der Empirie nicht bestätigt hat.

Das Verhaltensmodell nach Andersen

Andersen und seine Mitarbeiter/innen veröffentlichten zu Beginn der 1970er Jahre ein aus den Sozialwissenschaften stammendes Verhaltensmodell zur Modellierung des Inanspruchnahmegeschehens von Gesundheitsleistungen, das seither beständig weiterentwickelt wird. Das Modell enthält ein breites (potentiell vollständiges) Spektrum von Kategorien zur Subsumierung individueller und gesellschaftlicher Determinanten, welche die Inanspruchnahme beeinflussen können. Das Modell ist dazu geeignet, Kategorien bzw. Determinanten, die zur Beschreibung, Erklärung und Prognose des Inanspruchnahmegeschehens herangezogen werden können, in eine analytische Ordnung zu bringen. Im Zentrum des Modells stehen die Kategorien «Predisposing Characteristics», «Enabling Resources» und «Need» als Einflussgrössen auf die Inanspruchnahme von Gesundheitsleistungen. Unter dem Begriff **Predisposing Characteristics** fasst Andersen alle Merkmale zusammen, die sich indirekt auf die Inanspruchnahme auswirken. Dazu gehören verschiedene Merkmale aus den Bereichen Demographie (Alter, Geschlecht), Sozialstruktur (sozialer Status, Bildung, etc.) und Health Beliefs (Einstellungen, Werte und Wissen bez. Gesundheit). **Enabling Resources** sind die notwendige Voraussetzung für die Inanspruchnahme von Gesundheitsleistungen. Es wird zwischen personen- (bspw. Einkommen und das Vorhandensein von Krankenversicherung und Zusatzversicherungen) und gemeindebezogenen Ressourcen (bspw. das Vorhandensein und die Erreichbarkeit von Versorgungseinrichtungen am Wohn- und Arbeitsort) unterschieden. Beim **Need** wird zwischen dem von der betroffenen Person wahrgenommenen Bedarf einerseits und einem durch professionelles Urteil objektivierten Bedarf an-

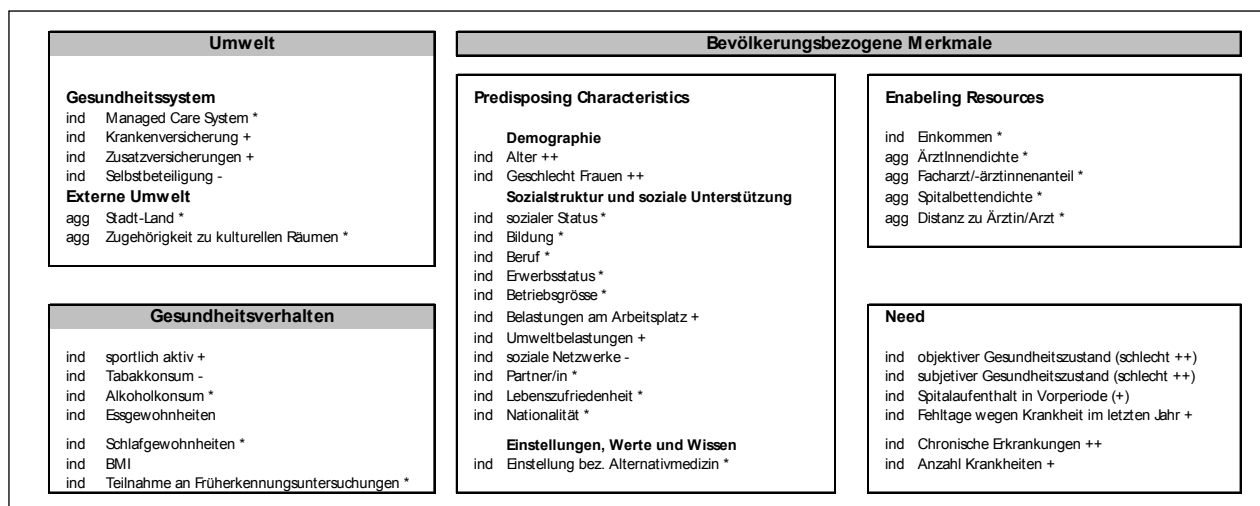
dererseits unterschieden. Der Modellkern wurde später um die Komponenten **Outcomes** und **Environment** ergänzt und schliesslich um Rückkopplungen von Outcomes auf Population Characteristics und auf das Gesundheitsverhalten vervollständigt.

International gibt es sehr viele Studien, welche sich auf das Modell von Andersen stützen. Die empirische Erklärungskraft des Andersen-Modells wird auf 25% der interindividuellen Varianz beziffert. Als wichtigste erklärende Grösse erscheint die subjektive Bewertung der eigenen Symptome (Need), während die übrigen Faktoren eine geringere Rolle spielen. **Abbildung Z1** gibt einen Überblick über die wichtigsten Faktoren und die Ergebnisse aus der Literatur.

Die umfassendste Studie zur Inanspruchnahme von Gesundheitsleistungen wurde in der Schweiz in der Untersuchung im Rahmen des Nationalen Forschungsprogramms 45 von Bisig / Gutzwiler et al. vorgenommen. Verschiedene Autoren und Autorinnen untersuchen, welche Faktoren – neben dem Gesundheitszustand der Patient/innen – die Inanspruchnahme von Gesundheitsleistungen beeinflussen. Hauptdatenquellen waren die Schweizerische Gesundheitsbefragung SGB des Bundesamtes für Statistik von 1992 und 1997. Dabei zeigt sich, dass Faktoren wie Geschlecht, Wohnregion, Nationalität, Sozialschicht und Versicherungsstatus mitbestimmen, ob und wie oft jemand Gesundheitsleistungen nachfragt. Zudem wird ersichtlich, dass die Wirkung der geprüften Determinanten nicht bei allen Arten von Gesundheitsleistungen die gleichen Effekte aufweisen.

Aus den Resultaten geht jedoch leider nicht hervor, ob es sich bei den beobachteten Unterschieden um einen Nachfrage- oder einen Angebotseffekt handelt.

Abbildung Z1: Auswahl aus empirisch überprüften Determinanten der Inanspruchnahme und ihre Wirkungsrichtung auf die Höhe der Inanspruchnahme (verschieden instrumentalisiert), kategorisiert nach dem Verhaltensmodell nach Andersen



Legende:

++ unbestrittener, starker positiver Einfluss; + unbestrittener positiver Einfluss; - unbestrittener negativer Einfluss;

* Einfluss vorhanden, jedoch nicht unbestritten

ind: Variable auf individueller Ebene;

agg: Variable auf aggregierter Ebene

Quelle: Eigene Darstellung

Angebotsinduzierung

Bei der Angebotsinduzierung wird davon ausgegangen, dass die Ärztinnen und Ärzte aufgrund der beratenden Funktion, die sie durch ihren Informationsvorsprung innehaben, erheblich auf die Art und den Umfang der Inanspruchnahme medizinischer Leistungen einwirken. Von einer Angebotsinduzierung ist dabei aber nur dann die Rede, wenn die Ärzt/innen bei der Behandlung ihre eigenen Interessen über die der Patient/innen stellen und ihre Position dazu nutzen, mehr Leistungen zu erbringen, als dies die PatientInnen bei gleichem Informationsstand gefordert hätten. Angebotsinduzierung setzt voraus, dass für die ÄrztInnen Anreize vorhanden sein müssen, die Menge ihrer Leistungen «künstlich» auszuweiten. Bei der Theorie zur Angebotsinduzierung wird davon ausgegangen, dass die Ärztinnen und Ärzte erstens über ein begrenztes Zeitbudget verfügen und zweitens, dass sie bezüglich ihres Einkommens ein Anspruchsniveau haben, das sie unbedingt realisieren möchten (Zieleinkommenshypothese). Wird das Zieleinkommen unterschritten, so verlieren die beiden übrigen Zielgrößen, Arbeitszeit und Berufsethik, ihre Bedeutung. Es wird dann versucht, das fehlende Einkommen über die Schaffung einer künstlichen Nachfrage zu erzielen.

Für verschiedene Autor/innen ist zur Beurteilung angebotsinduzierter Nachfrage von Bedeutung, welches Behandlungsergebnis trotz induzierter Nachfrage erreicht wird. Der Outcome sei auch zu berücksichtigen. So kann u.a. zwischen «sinnvoller» und «nicht erwünschter» arztinduzierter Nachfrage unterschieden werden. «Sinnvoll» ist sie dann, wenn sich eine vom Arzt bzw. von der Ärztin induzierte Behandlung positiv auf den Gesundheitszustand der Patientin bzw. des Patienten auswirkt. Ist dies nicht der Fall, d.h. die ärztlichen Leistungen verändern den Gesundheitszustand der PatientInnen nicht oder verschlechtern ihn sogar, dann handelt es sich um eine «nicht erwünschte» Induzierung.

Die in zahlreichen internationalen Untersuchungen gefundenen empirischen Ergebnisse und ihre Interpretation sind sehr kontrovers. Weder die Resultate der Untersuchungen von Aggregatdaten noch diejenigen von Individualdaten lassen über die Existenz und das Ausmass der Angebotsinduzierung eindeutige Schüsse zu. Während manche Autor/innen die Hypothese klar bestätigt sehen, sind sich andere nicht sicher oder finden sie nur für spezielle Formen von Gesundheitsgütern bestätigt. Nicht wenige Autor/innen lehnen die Hypothese sogar eindeutig ab. Die Gründe dafür sind u.a. folgende: Es bestehe keine einheitliche Definition des Begriffs angebotsinduzierter Nachfrage; es bestehe keine übereinstimmende theoretische Analyse des Phänomens; der Inanspruchnahmeprozess werde unterschiedlich modelliert (einstufig vs. mehrstufig, simultan oder sequentiell); es würden verschiedene ökonometrische Schätzmethode verwendet; die Inanspruchnahme werde in verschiedenen Masseinheiten gemessen (Mengen bzw. Geldeinheit); je nach Untersuchung werde zwischen verschiedenen Leistungsarten unterschieden; je nach Untersuchung werde die ÄrztInnendichte für verschiedene ÄrztInnentypen berücksichtigt; je nach Untersuchung würden (gleiche) strukturelle Variablen unterschiedlich aggregiert.

Als wesentliche Variable zur Überprüfung von Angebotsinduzierung wird in den meisten Untersuchungen die Variable «ÄrztInnendichte» verwendet. Unabhängig davon, mit welcher Schätzmethode und in welcher Masseinheit die abhängige Variable gemessen wird, werden als erklärende Größen für die Inanspruchnahme neben der «ÄrztInnendichte» ein möglichst umfassender Variablensatz mit Variablen aus verschiedenen Bereichen (bspw. «environment», «need», «predisposing» und «enabling») mitberücksichtigt. Nur bei einem solchen Vorgehen kann für so genannte konfundierende Faktoren kontrolliert werden. Gemeinsam ist zudem allen innerhalb der Literaturstudie gefundenen Erklärungsmodellen, dass die Qualität der ärztlichen Leistungen als homogen betrachtet wird.

Neben Untersuchungen von Aggregatdaten, die mit dem bekannten Problem des ökologischen Fehlschlusses¹ behaftet sind, wurde noch kaum versucht, die Existenz der Angebotsinduzierung in der Schweiz explizit mit Hilfe von Individualdaten zu untersuchen. Ausschlaggebend dafür ist, dass die Daten zur Ärzt/innendichte zum einen nur auf kantonaler Ebene verfügbar sind (Kantone entsprechen kaum der Definition einer medizinischen Versorgungsregion) und zum anderen die diesbezüglichen statistischen Grundlagen nicht zuverlässig sind.

Steuerung der Nachfrage

In der Gesundheitsökonomie wird darüber nachgedacht, wie das Verhalten der beteiligten Akteurinnen und Akteure gesteuert werden kann, damit gesundheitliche Leistungen möglichst effizient erbracht werden. Dabei geht es vor allem um eine Steuerung durch monetäre Anreize. Ein Zweig der Gesundheitsökonomie befasst sich deshalb mit der Frage, inwiefern Steuerungsmöglichkeiten durch unterschiedliche Versicherungsvertragsmodelle (bspw. Wahlfranchisen) bestehen und welche Auswirkungen diese auf das Verhalten der Versicherten und im Endeffekt auf die Kosten des Gesundheitswesens ausüben.

Beim der Analyse des Nachfrageverhaltens als Reaktion auf monetäre Anreize ist es empirisch sehr anforderungsreich, den Anreizeffekt (kleinere Inanspruchnahme aufgrund eines restriktiveren Verhaltens) vom Selbstselektionseffekt (weniger Inanspruchnahme aufgrund eines besseren Gesundheitszustandes) zu trennen.

Den methodologisch überzeugendsten Ansatz zur Identifikation des Anreizeffektes verfolgten Autoren aus Lausanne. Sie finden in ihren Daten von 62'000 Versicherten der CSS mit der höchsten Franchisen aus dem Kanton Waadt sowohl einen starken Selektionseffekt wie auch einen starken Anreizeffekt. Mit den Resultaten zeigen die Autoren, dass die Wahrscheinlichkeit, dass ambulante Kosten überhaupt entstehen, durch die Wahl der höchsten Franchise im Vergleich zur niedrigsten aufgrund des Anreizeffektes um 69% gesenkt wird. Gegeben, dass stationäre Kosten vorhanden sind, werden die Kosten der Versicherten mit der höchsten Franchise im Vergleich zu denjenigen mit der tiefsten insgesamt um 34% gesenkt, wovon die Hälfte auf Selektionseffekte und die andere auf Anreizeffekte zurückzuführen ist.

Probleme der empirischen Forschung

Es wurde schon darauf hingewiesen, dass die empirische Forschung zur Inanspruchnahme gesundheitlicher Leistungen mit verschiedenen Problemen zu kämpfen hat.

■ **Theoretische und begriffliche Probleme:** Besonders bei Fragestellungen, welche die Nachfrage betreffen, muss man sich über die zentralen Begriffe wie Bedarf, Nachfrage und Inanspruchnahme klar sein. Die Abgrenzung von Nachfrage und Inanspruchnahme ist nicht immer ganz deutlich und wird häufig als synonym verwendet. Während **Nachfrage** sich eher auf den Entscheidungsprozess bezieht, welcher der Inanspruchnahme vorausgeht, setzt die **Inanspruchnahme** immer den bereits erfolgten Arztbesuch, den Krankenhausaufenthalt oder die Medikamenteneinnahme voraus. Hinter den Begriffen stehen jedoch auch unterschiedliche Konzepte. Beim Nachfragekonzept geht man von «soveränen» Konsument/innen aus, welche die Wünsche selber artikulieren können; d.h. man stützt die Nachfrage allein auf die PatientInnenentscheidung ab und modelliert diese mit mikroökonomischen Verhaltensmodellen. Mit dem Begriff der Inanspruchnahme ist jedoch immer der Prozess ins-

¹ Mit dem Begriff «Ökologischer Fehlschluss» bezeichnet man das fehlerhafte Schliessen von statistischen Beziehungen zwischen aggregierten Daten auf Beziehungen zwischen den Komponenten der Aggregate (bspw. Individuen).

gesamt gemeint (realisierte Nachfrage); d.h. es gehen sowohl die Entscheidungen der PatientInnen wie der ÄrztInnen ein. Der Prozess selber wird sowohl von individuellen Merkmalen der AkteurInnen als auch von strukturellen Gegebenheiten beeinflusst.

Als **Bedarf** bezeichnet man diejenige Menge und diejenige Qualität von medizinischen Leistungen, die aus einer Evidence-based-Medicine-Sichtweise zur Diagnose und Behandlung von auftretenden Krankheiten notwendig sind. Ganz offensichtlich handelt es sich hier um einen normativen Begriff, weil die Vorstellung über die den Bedarf legitimierweise bestimmenden Faktoren auseinandergehen kann.

■ **Einfluss der Modellspezifikation auf die Resultate:** Insbesondere in der Inanspruchnahmeforschung sind die Ergebnisse von empirischen Analysen oft davon abhängig, welche erklärenden Merkmale für ein Schätzmodell vorhanden sind und welche dann tatsächlich in ein Modell eingehen. Je nach Modellspezifikation können so die Effekte der einbezogenen erklärenden Merkmale unterschiedlich ausfallen. Es ist auch möglich, dass bestimmte relevanten Variablen nicht berücksichtigt werden, oder dass die Endogenität unabhängiger Merkmale nicht berücksichtigt wird, was zu «falschen» Interpretationen der Resultate führen kann. Dass dieses Problem speziell bei der Inanspruchnahmeforschung von gesundheitlichen Leistungen auftritt, ist der Kompliziertheit und der Komplexität des Inanspruchnahmegeschehens zuzuschreiben.

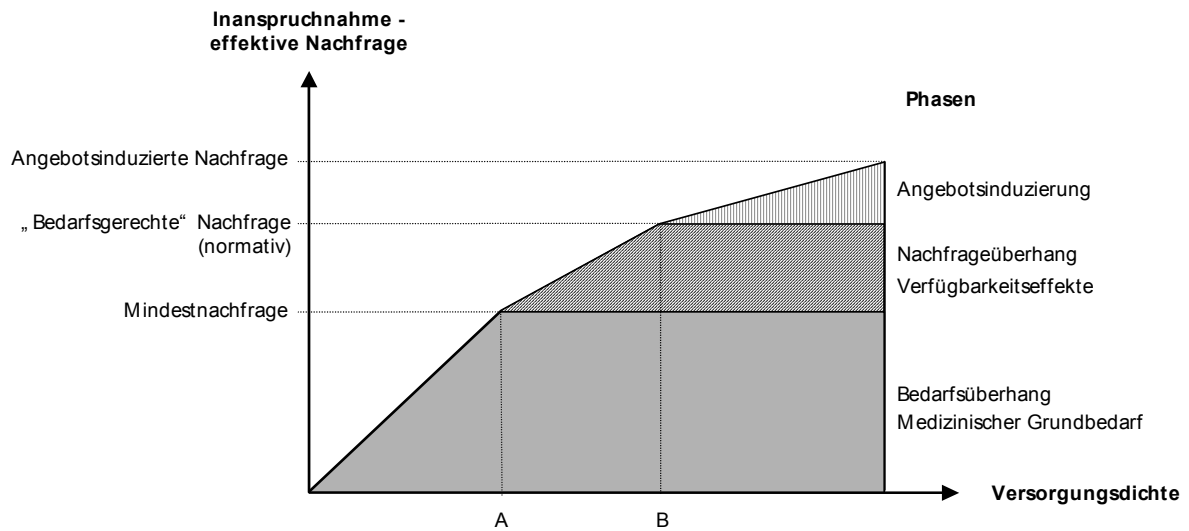
■ **Messfehler und andere datenspezifische Probleme:** Ungeeignete Aggregationsniveaus und nicht repräsentative Daten werden als Beispiele von Messfehlern genannt. Besonders hingewiesen werden muss auf das spezifische Problem, das die Verwendung unterschiedlicher Aggregationsniveaus in der Schätzung betrifft. Auf der einen Seite liegen den Schätzgleichungen zur Inanspruchnahme Individualdaten zugrunde. Auf der anderen Seite liegen die für die Überprüfung der Angebotsinduzierung verwendeten Variablen (bspw. ÄrztInnendichte) auf einer höheren Aggregationsebene vor. Auf inhaltlicher Ebene gilt es folgendes Problem zu lösen: Wie können individuenbezogene Daten mit den notwendig (innerhalb klarer Grenzen) aggregierten Daten der Angebotsstruktur (hier also der niedergelassenen Ärzten/Ärztinnen) so verknüpft werden, dass jede betrachtete Teilregion tatsächlich einen geographisch abgeschlossenen Markt für medizinische Leistungen bildet? Ein zu hohes Aggregationsniveau nivelliert die Unterschiede, die zwischen den einzelnen Märkten bestehen können und eine zu tiefe Disaggregation führt zu Verzerrungen durch einen zu hohen Anteil von «cross-bordering», d.h. von Patient/innen, die Ärzte und Ärztinnen kontaktieren, welche ausserhalb des als Markt identifizierten geographischen Raums ihre Praxen betreiben. In der Schweiz ist bspw. die ÄrztInnen-dichte bisher nur auf kantonaler Ebene verfügbar, was unter dem Aspekt geographisch abgeschlossener Märkte für medizinische Leistungen zumindest fragwürdig ist.

■ **Unterschiedliche Schätzverfahren:** In den in der Literatur beschriebenen empirischen Analysen zum Inanspruchnahmeprozess wird eine Vielzahl von unterschiedlichen Schätzverfahren angewandt. Gemeinsam ist allen weiterführenden Analysen, dass multivariate Methoden verwendet werden. Je nachdem, welches Schätzverfahren zur Anwendung kommt, können die Resultate variieren.

■ **Bestimmung des Ausmasses der angebotsinduzierter Nachfrage:** Dass der empirische Nachweis angebotsinduzierter Nachfrage sehr komplex und schwierig ist, wurde bereits erwähnt. Noch schwieriger ist es, das Ausmass zu bestimmen. Wir sind bei der Literaturrecherche auf keine empirischen Untersuchungen gestossen, welche versuchen, die Existenz von Angebotsinduzierung nachzuweisen und deren Modelle zugleich zwischen «erwünschter» und «unerwünschter» Angebotsinduzierung unterscheiden können.

Die Bestimmung der unerwünschten Angebotsinduzierung setzt als Referenz einen Bedarfswert voraus, der gesellschaftlich akzeptiert ist. Übersteigt die Inanspruchnahme diese Bedarfsgrösse, dann kann man von unerwünschter Angebotsinduzierung sprechen. **Abbildung 22** verdeutlicht dies. Ganz offensichtlich ist es aber nicht eindeutig, auf welchem Niveau diese Referenzbedarfsgrösse liegt.

Abbildung Z2: Modell des Zusammenhangs zwischen Versorgungsdichte (Allgemein- und FachärztInnen) und ambulanter Inanspruchnahme



(Quelle: Eigene Darstellung)

Die Bedeutung der Literaturanalyse

Aus der Literaturanalyse der Inanspruchnahme ambulanter Leistungen sowie der Angebotsinduzierung von Ärzten und Ärztinnen lassen sich auf verschiedenen Ebenen Folgerungen für die Nachfrageseite des Projektes «Démographie Médicale» ziehen.

- Erstens wurde deutlich, dass zwischen der Erklärung der Inanspruchnahme medizinischer Leistungen und der Bestimmung des Bedarfs eine wichtige konzeptionelle Differenz besteht. Währenddem im Analysekontext das «Auffinden» von Erklärungsfaktoren im Vordergrund steht, beschäftigt im Bedarfskontext vor allem die Frage, welche Faktoren die Inanspruchnahme bestimmen «dürfen».
- Zweitens hat die Literaturanalyse gezeigt, dass für die Beantwortung der zentralen Fragestellung nach der regionalen Über- bzw. Unterversorgung eine Bedarfsanalyse notwendig ist. Diese kann unterschiedlich komplex ausfallen.
- Drittens wurde deutlich, auch wenn die Literaturanalyse nur rudimentär durchgeführt werden konnte, dass der Prozess der Inanspruchnahme komplex ist. Dies bedeutet für die Bestimmung des Bedarfs an medizinischen Leistungen, dass diese Komplexität reduziert werden muss. Damit sind Schwerpunktsetzungen und Werturteile verbunden.
- Viertens konnte eine Übersicht über die Faktoren gewonnen werden, welche die Inanspruchnahme beeinflussen. Diese Faktoren sind bei der Abschätzung des regionalen Bedarfs in der einen oder anderen Form zu berücksichtigen. Werden ausgewählte Faktoren nicht berücksichtigt, so bedeutet dies auch an dieser Stelle eine normative Setzung, die ganz explizit dargelegt werden muss.
- Fünftens konnten vielfältige methodische Hinweise gewonnen werden, wie der Bedarf konkret abgeschätzt werden kann. Auch wurde deutlich, dass es keinen methodischen Goldstandard gibt. Vielmehr haben fast alle Vorgehensweisen sowohl Vor- wie auch Nachteile.
- Sechstens konnten vielfältige Einsichten im Bereich der Angebotsinduzierung gewonnen werden. Auf konzeptioneller Ebene wurde deutlich, dass Angebotsinduzierung nicht nur unerwünscht sein

kann und dass daneben auch eine Patienteninduzierung existiert. Es zeigte sich auch, dass die Angebotsinduzierung nur relativ zu einem normativen Referenzwert, der sog. «richtigen» oder «bedarfsgerechten» Versorgung, festgestellt werden kann. Methodisch konnte weiter gezeigt werden, dass die einwandfreie Feststellung der Existenz und des Ausmasses der Angebotsinduzierung sehr anspruchsvoll ist und in der Schweiz bisher noch nie à-fonds durchgeführt worden ist. Sollte im Gesamtprojekt der «Démographie Médicale» daher die Absicht bestehen, den Bedarf korrigiert um die Effekte der Angebotsinduzierung festlegen zu können, müssten dazu umfangreiche Analysen vorangehen.

■ Siebtens zeigte sich durch die Analyse der Anreize von Wahlfranchisen, dass nicht nur die Vergütungsanreize auf der Angebotsseite verhaltensrelevant sind, sondern dass auch die Nachfrage auf Anreize reagiert. Sollte bei der Feststellung des Bedarfs die Angebotsinduzierung korrigiert werden, so müsste man sich auch die Frage stellen, ob nicht auch eine Nachfrageinduzierung berücksichtigt werden müsste. Dies könnte dann bspw. bedeuten, dass bei der Feststellung des «legitimen» Bedarfs vom Verhalten der Versicherten in der höchsten Franchisestufe ausgegangen wird. Ganz offensichtlich beinhaltet dieses Vorgehen die Setzung von Werturteilen.

Teil 2: Konkretes Vorgehen zur Bestimmung des regionalen Bedarfs

Im zweiten Teil des Berichts werden, aufbauend auf den Erkenntnissen der Literaturanalyse, verschiedene Verfahren vorgestellt, wie der regionenspezifische Bedarf bestimmt werden kann. Derartige Kennziffern können für drei Arten von Vergleichen verwendet werden:

■ Erstens können die regionalspezifischen Bedarfskennziffern mit dem für die Schweiz durchschnittlichen Leistungsbedarf (Standardisierter Pro-Kopf-Leistungsbedarf) verglichen werden. Die Abweichung vom schweizerischen, durchschnittlichen Leistungsbedarf gibt Auskunft darüber, ob für das entsprechende Versorgungsgebiet aufgrund der in die Analyse miteinbezogenen Merkmale ein höherer bzw. ein tieferer Bedarf nach Gesundheitsleistungen besteht.

■ Zweitens kann die empirisch ermittelte regionalspezifische Bedarfskennziffer mit der entsprechenden «wahren» Pro-Kopf-Inanspruchnahme verglichen werden. Die Höhe der Abweichung von den «wahren», beobachteten Werten lässt erkennen, inwiefern noch andere, nicht berücksichtigte Determinanten auf die Inanspruchnahme von Gesundheitsleistungen innerhalb der entsprechenden Versorgungsregion einwirken.

■ Drittens kann der regionale SOLL-Bedarf mit dem regionalen ärztlichen Angebot verglichen werden. Dies erlaubt direkte Rückschlüsse auf die Frage der Über- und Unterversorgung.

Es wurde darauf hingewiesen, dass aufgrund des Kontextes, in dem eine Bedarfsanalyse stattfindet, die Komplexität im Vergleich zu Inanspruchnahmeanalysen reduziert werden muss. Die Antwort auf die Frage, welche Faktoren die Inanspruchnahme bestimmen «dürfen», wird durch eine normative Setzung gegeben. Sie muss als erstes beantwortet werden. Je nachdem, wie viele Faktoren für die Festlegung eines regionalen Leistungsbedarfs berücksichtigt werden sollen, müssen entsprechende Verfahren angewandt werden. Grundsätzlich stehen drei Varianten zur Verfügung, welche nun kurz vorgestellt werden.

Variante 1: Erstellung von Bedarfsprofilen

Mit dem ersten, relativ einfachen Verfahren ist es möglich, den Einfluss von wenigen «Schlüssel-determinanten» auf die Nachfrage/Inanspruchnahme zu berücksichtigen und diese in die Bedarfsberechnungen mit einzubeziehen. Dazu gehören mindestens die bevölkerungsstrukturellen Variablen «Alter» und «Geschlecht». Das Verfahren beruht auf der Überlegung, dass anhand von Versicherer-

daten für jede Person aufgrund von ausgewählten Merkmalen ein auf schweizerischen Durchschnittswerten beruhender Leistungsbedarf berechnet wird. Damit kann der Gesamtbedarf für jedes beliebig abgrenzbare Gebiet, für welches die erforderlichen Daten vorhanden sind, berechnet werden.

Der **Vorteil** des skizzierten Vorgehens zur Bestimmung eines regionenspezifischen Bedarfs liegt darin, dass es sich um eine einfache Methode handelt, welche problemlos umgesetzt werden kann. Der **Nachteil** liegt darin, dass damit weder nachfrage- noch angebotsinduzierende Effekte berücksichtigt werden können, weil die Gesamtheit der in Anspruch genommenen Leistungen als Referenzgrösse dient.

Variante 2: Multivariate Regression auf der Ebene von Versorgungsregionen

Ähnlich wie in Variante 1 ist bei dieser zweiten Variante zur Bestimmung des regionalspezifischen Leistungsbedarfs die medizinische Versorgungsregion. Anstatt anhand der Inanspruchnahmedaten einen standardisierten Leistungsbedarf zu ermitteln und diesen dann auf die Versorgungsregionen «umzurechnen», wird der regionalspezifische Leistungsbedarf mit Hilfe einer multivariaten Regressionsanalyse auf der Ebene der Versorgungsregionen direkt geschätzt. Die Prognose wird damit genauer ausfallen als in Variante 1, da mit der Methode der multivariaten Regression mehr Determinanten der Inanspruchnahme berücksichtigt werden können. Es können dabei sowohl angebotsseitige (bspw. ÄrztInnendichte) als auch nachfrageseitige Determinanten (Alter, Geschlecht, Stadt-Land, etc.) berücksichtigt werden.

Grundsätzlich ist es mit diesem Verfahren auch möglich, Angebots- und Nachfrageinduzierung zu berücksichtigen. Dazu müssten jedoch vorgängig zusätzliche Analysen, welche über die Existenz und das Ausmass von Angebotsinduzierung Auskunft geben, gemacht werden.

Der **Vorteil** gegenüber Variante 1 besteht darin, dass mehr Determinanten der Inanspruchnahme simultan berücksichtigt werden können. Damit sollte die Prognose exakter werden. Die **Nachteile** dieser Variante im Vergleich zur ersten sind die folgenden: Es werden höhere Anforderungen an die Daten gestellt; der Berechnungsaufwand ist grösser; es handelt sich um ein anspruchvolleres Verfahren, womit die Nachvollziehbarkeit geringer wird; Bedarfsprognosen über längere Zeiträume können kaum erstellt werden, weil die langfristige Prognose verschiedener Variablen fehlen.

Variante 3: Mehrebenenanalyse mit Individualdaten

Das dritte Verfahren beruht auf einer Mehrebenenanalyse (Kombinierung von Individual- und Aggregatebene). Ausgehend von deren Resultaten können regionale Bedarfsprofile berechnet werden. Mit diesem Verfahren könnte gleichzeitig die Forschungslücke, welche in Bezug auf die Existenz und das Ausmass von Angebotsinduzierung besteht, geschlossen werden. Es handelt sich dabei um das aufwändigste und methodisch anspruchsvollste Verfahren. Zudem stellt es im Vergleich zu den andern Verfahren die höchsten Anforderungen an die Daten.

Die vorgeschlagene Mehrebenenanalyse ist als **Vorteil** differenzierter als die Varianten 1 und 2, da nicht auf verhaltenswissenschaftliche Begründungen aus verschiedenen wissenschaftlichen Kontexten verzichtet werden muss. Zudem kann die Frage der Angebotsinduzierung zuverlässiger beantwortet werden. Demgegenüber besteht aber der **Nachteil**, dass die Anforderungen an die Datengrundlagen sehr hoch und die Berechnungen bei Variante 3 am aufwändigsten sind.

Beurteilung der drei Varianten

Aus Sicht der Autoren des vorliegenden Forschungsprotokolls schneidet Variante 2 unter Berücksichtigung der Vor- und Nachteile am besten ab. Die multivariate Regression auf der Stufe der Versorgungsgebiete ist mit den heutigen zur Verfügung stehenden Daten möglich und auch sinnvoll. Regionale Unterschiede, auch in der Angebotsstruktur, können damit viel besser berücksichtigt werden als mit Variante 1. Bedingung ist jedoch, dass in einem ersten Schritt auf empirischer Basis «vernünftige» medizinische Versorgungsregionen gebildet werden. Diese Aufgabe sollte prioritär angegangen werden, unabhängig davon, mit welcher Methode später regionale Bedarfskennziffern ermittelt werden.

Résumé

Ce rapport présente les résultats de travaux préparatoires effectués pour le projet de démographie médicale que l'Observatoire suisse de la santé (Obsan) prévoit de réaliser en collaboration avec divers partenaires. La désignation « Démographie médicale » recouvre l'analyse de l'état et de l'évolution de l'offre de soins ambulatoires et de la demande y afférente. Une question figure de façon récurrente au centre des débats politiques sur la demande de prestations médicales : les médecins créent-ils la demande (induction de la demande par l'offre) et dans quelles proportions ? C'est pourquoi l'Obsan a chargé le bureau d'études de politique du travail et de politique sociale (Büro für arbeits- und sozialpolitische Studien, BASS) de clarifier divers aspects de la demande en se basant sur une analyse bibliographique.

Questions soulevées

Concrètement, il s'agissait de répondre aux questions suivantes :

- De quels résultats empiriques dispose-t-on pour évaluer la demande en Suisse ? Et quelles conséquences peut-on en tirer pour faire des estimations futures ?
- Quelles méthodes appliquer pour déterminer le recours aux prestations ambulatoires ?
- Comment le recours aux prestations induit par l'offre est-il pris en compte ?
- Quels sont les avantages et inconvénients des différentes méthodes pour l'analyse de la demande prévue dans le cadre du projet global de démographie médicale ?

Procédure méthodologique

D'entente avec le mandant, l'analyse bibliographique a été effectuée selon une approche pragmatique. Fondée sur des travaux scientifiques menés à bien dans ce domaine, cette étude a été complétée par des recherches bibliographiques réalisées sur Internet et dans la banque de données économiques ECONLIT. Dans une première phase, les données ont été dépouillées et analysées. Dans une deuxième phase, les résultats de cette analyse ont été repris pour concevoir trois variantes visant à déterminer de façon concrète les besoins régionaux de prestations ambulatoires en Suisse.

1^{re} partie: Résultats de l'analyse bibliographique

Deux approches théoriques

Depuis le début des années 1960, le recours aux prestations médicales a fait l'objet d'études empiriques. Etant donné que les études démographiques de type purement descriptif menées en la matière ne fournissaient pas de base d'informations satisfaisante, des scientifiques issus des **sciences sociales** ont développé des **modèles comportementaux** sur le recours aux soins médicaux. Le modèle de comportement d'Andersen est présenté ci-après à titre d'exemple, car son approche a l'avantage d'être globale. Parallèlement à cette approche socioéconomique, une branche théorique distincte, consacrée au recours aux prestations médicales, s'est développée en **économie de la santé**. Celle-ci traite entre autres de problèmes de régulation résultant des particularités du domaine de la santé et de sa production.

En économie de la santé, l'offre et la demande sont décrites à l'aide de modèles qui expliquent le comportement des acteurs par des décisions économiquement rationnelles. Outre cet aspect, les

modèles sociologiques s'attachent également à considérer des dimensions sociologiques et psychologiques ainsi que des facteurs liés à la morbidité.

Les modèles comportementaux et les modèles de l'économie de la santé se recoupent toutefois aujourd'hui sous bien des aspects, du moins en ce qui concerne leur application, respectivement leur vérification empirique. Cette évolution est notamment due au fait que les analyses « traditionnelles » de comportement d'orientation sociologique ou économiques se sont de plus en plus orientées vers des modèles multi variés. Ces deux types d'analyse essaient de modéliser les différents paramètres influant sur le recours aux soins à l'aide de données individuelles (p. ex. âge, sexe, revenu, etc.), lesquelles sont complétées par des variables structurelles (p. ex. ville-campagne, densité de médecins, région linguistique, etc.).

Le modèle de Grossman

En économie de la santé, le modèle le plus connu pour l'estimation de la demande est celui de Grossman. Premièrement, précisons que selon ce modèle, la demande de prestations médicales est interprétée comme une demande dérivée, car, à l'origine d'une consultation il y a une « demande de santé » et non pas de prestations médicales. Deuxièmement, chaque être humain possède un stock de santé qui dépend de son recours aux prestations médicales et d'un taux de dépréciation défini explicitement comme une fonction de l'âge (du stock de santé) et de l'intensité de son utilisation. A travers le taux de dépréciation, des variables telles que le comportement en matière de santé ou le niveau d'éducation peuvent aussi être intégrées dans le modèle. Une corrélation est établie entre le revenu et l'état de santé dans la mesure où un stock de santé plus important présente un intérêt en terme de consommation comme en terme d'investissement. L'évidence empirique a infirmé de nombreuses implications du modèle de Grossman. Selon ce modèle, en particulier l'état de santé et la demande de prestations médicales devraient présenter une corrélation positive, ce qui n'a pas été attesté empiriquement.

Le modèle de comportement selon Andersen

Au début des années 1970, Andersen et ses collaborateurs publiaient un modèle de comportement issu des sciences sociales servant à la schématisation du recours aux prestations de santé. Depuis, ce modèle a fait l'objet de développements réguliers. Il contient un vaste spectre (potentiellement complet) de catégories intégrant des déterminants individuels et sociaux susceptibles d'influencer le recours aux prestations. Ce modèle autorise un classement analytique des catégories ou déterminants qui permettent de décrire ou d'expliquer les divers aspects liés au recours aux prestations et de faire des prévisions en la matière. Au cœur de ce modèle figurent les catégories *Predisposing characteristics*, *Enabling resources* et *Need* qui constituent les principaux facteurs d'influence sur le recours aux prestations de santé. Sous la désignation ***Predisposing characteristics***, Andersen regroupe l'ensemble des caractéristiques qui ont une incidence indirecte sur le recours aux prestations. Certaines d'entre elles ont trait à la démographie (âge, sexe), à la structure sociale (statut social, éducation, etc) et aux *health beliefs* (attitudes, valeurs et savoir en matière de santé). Les ***Enabling resources*** définissent les conditions qui doivent être remplies afin de pouvoir recourir à des prestations de santé. A ce niveau, une distinction est faite entre les ressources d'une personne (p. ex. revenu, existence d'une assurance-maladie et/ou complémentaire) et celles disponibles au niveau communal (p. ex. présence d'établissements de soins et accessibilité depuis le lieu de travail et le domicile). La catégorie ***Need*** différencie les besoins tels qu'ils sont perçus par la personne concernée et les be-

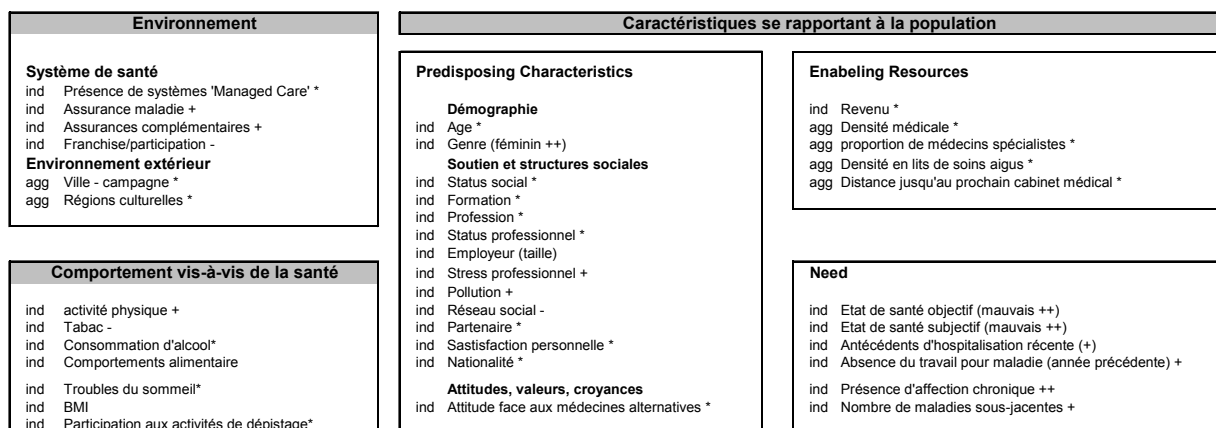
soins objectifs diagnostiqués par un professionnel. Par la suite, les composants **Outcomes** et **Environnement** sont venus s'ajouter aux principaux facteurs du modèle, qui a encore été complété par les rétroactions des outcomes et de l'environnement sur les caractéristiques de la population et le comportement en matière de santé.

A l'échelon international, de nombreuses études s'appuient sur le modèle d'Andersen. Celui-ci permet d'expliquer de manière empirique 25% de la variance interindividuelle. L'appréciation subjective des propres symptômes (need) semble être la principale variable explicative, tandis que les autres facteurs jouent un rôle moindre. La **figure R1** donne un aperçu des principaux facteurs et résultats issus de l'analyse bibliographique.

En Suisse, l'étude la plus complète sur le recours aux prestations de santé a été menée par Bisig/Gutzwiler et al. dans le cadre du programme national de recherche n° 45. Les différents auteurs qui ont participé à ce projet ont examiné quels facteurs, outre l'état de santé du patient, ont une incidence sur son recours aux prestations de santé. A cet égard, les *Enquêtes suisses sur la santé ESS* de 1992 et 1997 réalisées par l'Office fédéral de la statistique ont constitué les principales sources de données. De cette étude, il ressort que des facteurs tels que le sexe, la région de domicile, la nationalité, le niveau social et le type d'assurance-maladie contribuent à déterminer si une personne va recourir ou non à des prestations de santé et à quelle fréquence. On constate par ailleurs que les effets des déterminants examinés ne sont pas les mêmes selon le type de prestation de santé.

Les résultats obtenus ne permettent malheureusement pas de dire si les différences observées se rapportent d'abord à la demande ou plutôt à l'offre.

Figure R1: Sélection de déterminants du recours aux prestations de santé vérifiés empiriquement et effets produits sur le volume des prestations sollicitées (instrumentalisés de diverses façons), catégorisation selon le modèle de comportement d'Andersen



Légende

++ effet fortement positif démontré ; + effet positif démontré ; - effet négatif démontré

* effet possible, discuté

ind : données individuelles

agg : données agrégées

Source : Représentation propre, Bureau BASS

Demande induite par l'offre

S'agissant de la demande induite par l'offre, on part du principe que les médecins exercent une grande influence sur le type et le volume de prestations de santé sollicitées en raison du statut de conseiller que leur confère leur savoir. A noter que l'on parle uniquement de demande induite par l'offre lorsque les médecins placent leurs intérêts personnels avant ceux du patient et profitent de leur position pour fournir davantage de prestations que celles qu'aurait demandées le patient si son niveau de savoir était identique à celui de son médecin.

La demande induite par l'offre présuppose que divers facteurs incitent les médecins à augmenter de façon « artificielle » la quantité de traitements dispensés. La théorie relative à la demande induite par l'offre se base sur deux hypothèses : premièrement, le temps dont disposent les médecins est limité et, deuxièmement, ils désirent impérativement atteindre un certain niveau d'honoraires (hypothèse du revenu visé). Si les honoraires se situent en dessous du montant visé, les deux autres paramètres, c.-à-d. le temps de travail et l'éthique professionnelle, perdent de leur importance. Les médecins essaient alors de compenser la part de revenu manquant en créant une demande artificielle.

En vue d'apprécier la demande induite par l'offre, divers auteurs estiment qu'il est important de considérer le résultat obtenu à l'issue du traitement médical dispensé. Il convient donc également de tenir compte de l'outcome. A cet égard, une distinction peut notamment être faite entre demande induite « utile » ou « souhaitable » et « non souhaitable ». La demande induite par l'offre est considérée comme utile dès lors que le traitement initié par le médecin a un effet positif sur l'état de santé du patient. Dans le cas contraire, c.-à-d. lorsque les prestations médicales n'ont aucune incidence sur l'état de santé du patient ou qu'elles l'aggravent, il s'agit d'une induction non souhaitable.

Les résultats empiriques de nombreuses études internationales et leur interprétation sont très controversés. Ni les analyses de données agrégées ni celles de données individuelles ne permettent de tirer des conclusions claires sur l'existence et l'ampleur de la demande induite par l'offre. Si cette hypothèse trouve entièrement confirmation aux yeux de certains auteurs, d'autres émettent des doutes ou estiment qu'elle se vérifie uniquement pour des types spéciaux de biens de santé. D'autres encore, et ils ne sont pas rares, réfutent catégoriquement cette hypothèse. Les raisons notamment invoquées sont les suivantes : tout d'abord, il n'existe aucune définition uniforme de la demande induite par l'offre ni d'analyses théoriques concordantes de ce phénomène. Le processus de recours fait par ailleurs l'objet de diverses modélisations (simultanées ou séquentielles, à un niveau ou à plusieurs niveaux). Ensuite, les méthodes d'estimation économétriques appliquées varient, et diverses unités de mesure sont utilisées pour déterminer le recours aux prestations (quantités ou unité monétaire). Autres réserves émises : si certaines enquêtes établissent une distinction entre les types de prestations ou considèrent la densité des médecins par catégorie professionnelle, d'autres ignorent ces aspects. Enfin, des variables structurelles (identiques) ont été agrégées différemment selon les études.

Dans la plupart des analyses, la densité de médecins constitue une variable essentielle pour examiner l'induction par l'offre. Indépendamment de la méthode d'estimation et de l'unité de mesure choisies pour déterminer la variable dépendante, des séries de paramètres aussi complètes que possible – outre la densité de médecins – sont prises en compte pour expliquer le recours aux prestations de santé. Ces variables se rapportent à divers domaines (p. ex. *Environment, Need, Predisposing* et *Enabling*). Seule cette manière de procéder permet de contrôler de façon appropriée des facteurs confondants. Tous les modèles explicatifs répertoriés dans le cadre de l'analyse bibliographique considèrent par ailleurs la qualité des prestations médicales comme homogène.

En Suisse, hormis des analyses de données agrégées soumises au risque de sophisme écologique², un problème bien connu, l'existence de la demande induite par l'offre n'a quasiment fait l'objet d'aucune enquête s'appuyant explicitement sur des données individuelles. Deux raisons expliquent notamment cette lacune : d'une part, les données relatives à la densité des médecins sont uniquement disponibles à l'échelon cantonal (bien que les cantons ne correspondent guère à la définition d'une région responsable de la fourniture de soins) et, d'autre part, les bases statistiques y afférentes ne sont pas fiables.

Régulation de la demande

Dans le domaine de l'économie de la santé, la question se pose de savoir comment le comportement des acteurs peut être influencé afin que la fourniture de prestations de santé soit aussi efficace que possible. La régulation considérée ici s'opérerait principalement par des incitations monétaires. Une branche de l'économie de la santé s'intéresse donc aux possibilités de régulation par divers modèles de contrats d'assurance (franchises à option p. ex.) ainsi qu'à leur incidence sur le comportement des assurés et, par conséquent, sur les coûts de la santé.

L'analyse du comportement sur le plan de la demande en réaction aux incitations monétaires pose de grandes exigences d'un point de vue empirique, car il convient de distinguer l'effet d'incitation (recours réduit en raison d'un comportement plus restrictif) de l'effet d'auto sélection (recours moins important en raison d'un meilleur état de santé).

Des auteurs de Lausanne ont appliqué l'approche méthodologique la plus convaincante pour identifier l'effet d'incitation. Sur la base de données fournies par la caisse d'assurance maladie CSS, ils ont constaté que l'effet d'incitation tout comme l'effet de sélection était important pour les 62 000 assurés ayant choisi les franchises les plus élevées dans le canton de Vaud. Selon les résultats de leurs analyses, l'effet d'incitation entraîne une probabilité d'engendrer des coûts pour des traitements ambulatoires de 69% inférieure chez les assurés ayant choisi la franchise la plus élevée que chez ceux ayant opté pour la franchise la plus basse. Par contre, parmi les personnes ayant bénéficiés de traitements médicaux intra-muros, les coûts des assurés ayant la plus haute franchise sont inférieurs d'environ 34% à ceux des assurés avec la franchise la plus basse, la moitié de cette baisse étant attribuable aux effets de sélection et l'autre, aux effets d'incitation.

Problèmes liés à la recherche empirique

Comme mentionné à plusieurs reprises, la recherche empirique sur le recours aux prestations de santé se heurte à divers problèmes.

■ **Problèmes théoriques et terminologiques** Si l'on désire traiter des aspects liés à la demande en particulier, il convient de connaître la définition des termes centraux « besoins », « demande » et « recours ». La différence entre demande et recours, deux mots souvent employés à tort comme synonymes, n'est pas toujours très nette. Tandis que la **demande** se réfère plutôt au processus de décision qui précède le recours aux prestations de santé, le **recours** présuppose que le patient a déjà consulté un médecin, séjourné à l'hôpital ou suivi un traitement médicamenteux. Derrière ces termes se cachent par ailleurs divers concepts. S'agissant de la demande, ce concept veut que les consommateurs soient maîtres de leurs choix et à même d'exprimer leurs désirs. En d'autres termes, la demande repose sur la seule décision du patient. Elle est schématisée à l'aide de modèles de comportement mic-

² Le sophisme écologique désigne une mauvaise interprétation des relations statistiques entre des données agrégées dans la mesure où elles sont considérées comme des relations entre les composants des agrégats (p. ex. les individus).

micro-économiques. Le recours englobe par contre l'ensemble du processus (demande réalisée), c.-à-d. qu'il intègre tant les décisions des patients que celles des médecins. Le processus en lui-même est influencé par les caractéristiques individuelles des acteurs ainsi que par des spécificités structurelles. Par **besoins**, on entend la quantité et la qualité des prestations médicales nécessaires au diagnostic et au traitement de maladies reconnues, lesquelles ont été déterminées selon une pratique médicale basée sur l'évidence (*evidence based*). Selon toute apparence, il s'agit ici d'une notion normative, car la conception que l'on se fait des facteurs qui déterminent légitimement les besoins peut changer.

■ **Influence des spécifications du modèle sur les résultats** : En particulier dans le domaine de la recherche sur le recours aux prestations de santé, les résultats d'analyses empiriques dépendent souvent des caractéristiques explicatives dont on dispose pour un modèle d'estimation et de l'utilisation effective que l'on en fait. Selon les spécifications du modèle, les effets des caractéristiques explicatives considérées peuvent ainsi varier. Il est également possible que certaines variables importantes soient ignorées ou que l'endogénéité de caractéristiques indépendantes ne soit pas prise en compte, ce qui peut conduire à de « mauvaises » interprétations des résultats. Que ce problème se présente en particulier dans la recherche sur le recours aux prestations de santé est dû à la complexité des phénomènes se rapportant à ce domaine.

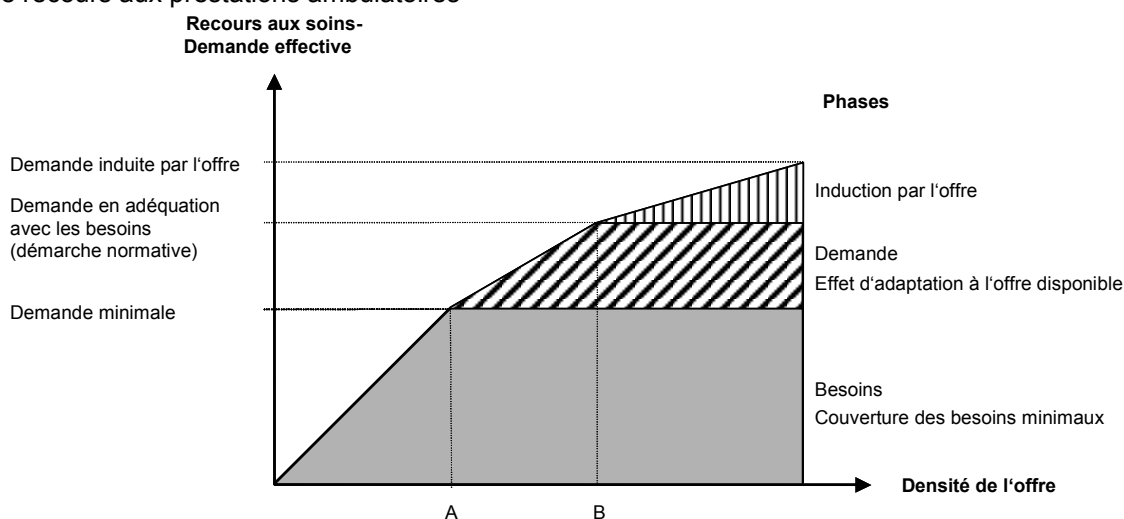
■ **Erreurs de mesure et autres problèmes liés aux données** : Parmi les erreurs de mesure, citons des niveaux d'agrégation inappropriés et des données non représentatives. Il convient de souligner ici le problème spécifique posé par l'utilisation de différents niveaux d'agrégation dans les estimations. Ainsi, tandis que les équations d'estimation du recours aux prestations de santé se basent sur des données individuelles, les variables utilisées pour vérifier l'induction de la demande par l'offre (p. ex. densité de médecins) présentent un niveau d'agrégation plus élevé. Sur le plan du contenu, il s'agit de résoudre le problème suivant : comment mettre en relation des données individuelles avec les données agrégées concernant la structure de l'offre (dans ce cas, les médecins établis), de sorte que chaque région considérée forme effectivement une zone géographique pertinente du point de vue de la fourniture des prestations médicales ? Deux difficultés se présentent ici. D'une part, un niveau d'agrégation trop élevé nivelle les éventuelles différences entre les marchés. D'autre part, un niveau de désagrégation trop bas biaise l'analyse en raison du problème lié aux régions limitrophes, c.-à-d. qu'une part importante de patients s'adressent à des médecins qui ont un cabinet en-dehors de l'espace géographique identifié en tant que marché. En Suisse, la densité de médecins, par exemple, n'est disponible à ce jour qu'à l'échelon cantonal, ce qui est d'un intérêt discutable si l'on se place sous la perspective de marchés géographiques clairement délimités.

■ **Diverses méthodes d'estimation** : Les analyses empiriques du processus de recours décrites dans la littérature s'appuient sur des méthodes d'estimation aussi diverses que variées. A noter que les analyses plus poussées reposent toutes sur des méthodes multi variées. Selon la méthode d'estimation appliquée, les résultats peuvent donc varier.

■ **Détermination de l'ampleur de la demande induite par l'offre** : Comme déjà mentionné, il est extrêmement complexe d'apporter des preuves empiriques à la demande induite par l'offre. Et il apparaît d'autant plus difficile d'en déterminer l'ampleur. Dans le cadre de nos recherches bibliographiques, nous n'avons pas trouvé d'études empiriques qui tentent de prouver l'existence de la demande induite par l'offre et dont les modèles établissent par ailleurs une distinction entre induction « souhaitable » et « non souhaitable ».

La détermination de l'induction « non souhaitable » requiert comme référence une valeur définissant les besoins qui soit acceptable pour la société. Si le recours aux prestations est supérieur à cette valeur, on peut alors parler d'induction non souhaitable. La **figure R2** illustre ce cas de figure. Il n'apparaît toutefois pas clairement à quel niveau se situe cette valeur de référence pour les besoins.

Figure R2 : Modèle présentant la relation entre la densité de médecins (généralistes et spécialistes) et le recours aux prestations ambulatoires



Source : Représentation propre, Bureau BASS

L'importance de l'analyse bibliographique

L'analyse bibliographique portant sur le recours aux prestations ambulatoires et sur l'induction de l'offre par les médecins permet de tirer diverses conclusions sur la demande dans la perspective du projet de démographie médicale.

- Premièrement, nous avons constaté des différences conceptuelles dans la manière de procéder pour expliquer le recours aux prestations médicales et la manière de déterminer les besoins. Si, au niveau analytique, la recherche de facteurs explicatifs figure au premier plan, l'identification des facteurs susceptibles de déterminer le recours aux prestations occupe une place centrale dans le domaine des besoins.

- Deuxièmement, les recherches bibliographiques ont montré qu'une analyse des besoins était nécessaire pour savoir s'il existait une pénurie ou une pléthore de prestations de santé au niveau régional. Selon le type d'analyse, cette démarche pourrait s'avérer complexe.

- Troisièmement, il ressort que le processus du recours aux prestations de santé présente un degré de complexité élevé, même si l'analyse bibliographique n'a pu être réalisée que de façon sommaire. Ce degré de complexité doit donc être réduit si l'on entend déterminer les besoins de prestations médicales. A cet effet, il faudrait définir les éléments centraux à prendre en compte, ce qui impliquerait des jugements de valeur.

- Quatrièmement, l'analyse bibliographique a permis de se faire une idée des facteurs qui influencent le recours aux prestations de santé. Ceux-ci doivent être intégrés, sous une forme ou autre, dans l'estimation des besoins régionaux. Au cas où certains des facteurs sélectionnés seraient ignorés, il conviendrait alors d'adopter à cet endroit une approche normative qui devrait être présentée de façon explicite.

- Cinquièmement, l'analyse bibliographique a permis de rassembler de nombreuses indications sur les manières de procéder à une estimation concrète des besoins. Il est clairement apparu qu'il n'existait pas de méthode infaillible, mais que les approches examinées présentaient toutes des avantages et des inconvénients.

■ Sixièmement, toute une série de connaissances ont pu être réunies sur la demande induite par l'offre. Il est ressorti que la demande induite par l'offre pouvait aussi avoir un effet positif (souhaitable) et qu'il existait également une induction par le patient. Autre constat : l'induction par l'offre ne peut être mise en évidence que par rapport à une valeur de référence normative des prestations de santé, considérées comme « appropriées » ou « conformes aux besoins ». D'un point de vue méthodologique, il est apparu qu'il est très difficile de prouver de façon irréfutable l'existence d'une demande induite par l'offre et d'en déterminer l'ampleur. En Suisse, aucune enquête approfondie allant dans ce sens n'a été réalisée à ce jour. Si le projet global de démographie médicale devait, comme planifié, déterminer les besoins corrigés des effets de la demande induite par l'offre, de vastes analyses devraient être réalisées au préalable.

■ Septièmement, l'analyse des incitations relatives aux franchises à option a montré que les incitations motivées par le remboursement des prestations influençaient le comportement des acteurs au niveau de l'offre et aussi de la demande. Si, pour déterminer les besoins, la demande induite par l'offre devait être corrigée, la question se pose de savoir si une induction par la demande ne devrait pas être prise en compte. Cela impliquerait, par exemple, que l'on se réfère au comportement des assurés qui ont choisi les franchises les plus élevées pour estimer les besoins « légitimes ». De toute évidence, cette procédure reposerait sur des jugements de valeur.

2^e partie : Comment déterminer concrètement les besoins régionaux ?

Fondée sur les résultats obtenus à l'issue de l'analyse bibliographique, la seconde partie du rapport présente diverses manières de procéder pour déterminer les besoins régionaux. Les valeurs y afférentes peuvent être utilisées pour trois types de comparaison.

■ Premièrement, les indices des besoins régionaux pourraient être comparés avec les besoins de prestations moyens pour la Suisse (besoins moyens standardisés par tête). Un écart avec la moyenne suisse indiquerait si, sur la base des caractéristiques intégrées dans l'analyse, les besoins en matière de prestations de santé sont supérieurs ou inférieurs pour la région considérée.

■ Deuxièmement, l'indice des besoins déterminé de façon empirique pour la région considérée pourrait être comparé avec le recours réel aux prestations par tête. L'écart observé avec les valeurs réelles permettrait de définir dans quelle mesure d'autres déterminants qui n'ont pas été pris en compte influencent le recours aux prestations de santé dans la région considérée.

■ Troisièmement, les besoins régionaux prévisionnels pourraient être comparés avec l'offre de médecins établis dans la région considérée. Ceci permettrait de savoir s'il existe une pénurie ou une pléthore de prestations de santé.

Comme souligné précédemment, en raison du contexte dans lequel se déroule une analyse des besoins, son degré de complexité doit être réduit par rapport à celui des analyses portant sur le recours aux prestations. Ainsi, il conviendrait d'adopter une approche normative en vue de répondre à la question « Quels sont les facteurs susceptibles d'influencer le recours aux prestations de santé ? » Les travaux à entreprendre devraient en priorité clarifier ce point. Le nombre de facteurs à considérer pour estimer les besoins régionaux s'avérerait déterminant pour la procédure à appliquer. Fondamentalement, trois variantes semblent envisageables. Elles sont décrites succinctement ci-après.

1^{re} variante: Etablissement de profils en matière de besoins

Relativement simple, la première approche permettrait de considérer l'influence qu'exercent quelques « déterminants-clés » sur la demande de prestations/le recours aux prestations et de les intégrer dans

le calcul des besoins. Parmi ces déterminants devraient au moins figurer les variables structurelles « âge » et « sexe » de la population. Cette variante repose sur l'idée suivante : en se fondant sur les données des assureurs, il serait possible de calculer les besoins de prestations de chaque personne selon des caractéristiques sélectionnées en s'appuyant sur la moyenne nationale. Les besoins globaux de chaque région délimitée, pour laquelle les données nécessaires sont disponibles, pourraient ainsi être calculés.

L'**avantage** de cette variante réside dans sa simplicité et, partant, dans sa mise en œuvre. Elle présente toutefois un **inconvenient** : elle ne pourrait prendre en compte ni les effets induisant la demande ni ceux induisant l'offre, car l'ensemble des prestations sollicitées servirait de valeur de référence.

2° variante: Régression multi variée au niveau des régions responsables de la fourniture de prestations

Comme pour la première variante, la région chargée de fournir des prestations médicales constitue ici l'unité d'analyse pour estimer les besoins régionaux. Au lieu de déterminer des besoins de prestations standardisés sur la base des données relatives au recours et de les « extrapoler » ensuite aux régions chargées de la fourniture des soins, les besoins spécifiques à une région seraient directement évalués au niveau de la région responsable de la fourniture des soins à l'aide d'une analyse de régression multi variée. Les prévisions en résultant auraient l'avantage d'être plus précises qu'avec la première variante, car la méthode de régression multi variée permet d'inclure davantage de déterminants ayant trait au recours aux prestations. Ceux-ci pourraient aussi bien concerner les facteurs liés à l'offre (p. ex. densité de médecins) que ceux relatifs à la demande (âge, sexe, ville-campagne, densité de médecins, etc.).

Fondamentalement, cette variante permettrait aussi de considérer l'induction par l'offre et par la demande. Des analyses supplémentaires devraient toutefois être faites au préalable afin de disposer d'informations sur l'existence de la demande induite par l'offre ainsi que sur son ampleur.

Par rapport à la première variante, cette manière de procéder présente l'**avantage** qu'un plus grand nombre de déterminants relatifs au recours aux prestations pourraient être examinés simultanément, ce qui augmenterait la précision des prévisions. Elle présente toutefois aussi les **inconvenients** suivants: d'une part, les exigences posées sur le plan des données sont plus élevées et, d'autre part, les calculs requis par cette approche demanderaient un plus grand investissement en temps. Il s'agit donc d'une procédure plus complexe, l'interprétation des résultats étant alors aussi plus difficile à faire comprendre. Par ailleurs, il est quasiment impossible de faire des prévisions en matière de besoins sur une plus longue durée, car les données de base nécessaires pour réaliser des projections ne sont pas disponibles pour toutes les variables.

3° variante: Analyse à plusieurs niveaux sur la base de données individuelles

La troisième procédure repose sur une analyse à plusieurs niveaux (combinaison entre niveau individuel et niveau agrégé). En se fondant sur les résultats de ce type d'analyse, il serait possible de déterminer les besoins caractéristiques à l'échelon régional. Cette approche permettrait par ailleurs de combler la lacune existant dans le domaine de la recherche, car elle contribuerait à étayer l'existence de la demande induite par l'offre et d'en déterminer l'ampleur. A noter que cette analyse présente non seulement le plus haut degré de complexité méthodologique, mais qu'elle entraînerait aussi le plus grand investissement en temps. Les exigences en matière de données sont par ailleurs supérieures à celles des autres approches.

Cette analyse à plusieurs niveaux a pour principal avantage d'être plus différenciée que les variantes 1 et 2, car elle intègre des motifs comportementaux émanant de divers contextes scientifiques. Il serait dès lors aussi possible d'identifier avec davantage de fiabilité dans quelle mesure la demande est induite par l'offre. Cette variante présente toutefois deux **inconvenients**. En effet, les exigences relatives aux bases de données seraient très élevées et les calculs y afférents représenteraient la charge de travail la plus lourde.

Appréciation des trois variantes

De l'avis des auteurs du présent rapport, la deuxième variante apparaît la plus appropriée compte tenu de ses avantages et de ses inconvenients. La régression multi variée au niveau des régions chargées de la fourniture des prestations semble non seulement possible mais aussi judicieuse si l'on tient compte des données dont on dispose actuellement. Comparée à la première variante, cette approche permet de mieux intégrer les différences régionales également présentes dans la structure de l'offre. Elle présuppose toutefois que, dans un premier temps, les régions chargées de la fourniture des prestations médicales soient définies de manière « raisonnable » sur une base empirique. Cette tâche aurait la priorité absolue, indépendamment de la méthode choisie plus tard pour déterminer des indices régionaux des besoins.

1 Einleitung, Ausgangslage, Fragestellungen

Ausgangslage

Im Oktober 2003 schrieb das Schweizerische Gesundheitsobservatorium (Obsan) ein Forschungsmandat zum Thema «Bedarfsanalyse und Entwicklungsprognosen der Ärzte in freier Praxis und der anderen Professionen in der ambulanten Gesundheitsversorgung der Schweiz – Pilot- und Machbarkeitsstudie» aus. Politischer Hintergrund war das Auslaufen des Zulassungsstopps für ambulante Leistungserbringer/innen im Juni 2005 und die geplante Einführung der Vertragsfreiheit zwischen Versicherern und ambulanten Leistungserbringer/innen. Insbesondere sah der Vorschlag des Bundesrates bei der Einführung der Vertragsfreiheit vor, dass die Kantone angeben, mit wie vielen Leistungserbringer/innen die Versicherer minimal zu kontrahieren haben, um die Versorgungssicherheit sicherzustellen.

Der Auftrag wurde vom Büro für arbeits- und sozialpolitische Studien (BASS) bearbeitet und im Juni 2004 abgeschlossen.³ Die Aufarbeitung bestand insbesondere aus zwei Teilen: Zum einen wurde eine Literaturlauswertung zu Methoden der Prognose und Planung von Angebot und Nachfrage in der ambulanten Praxis vorgenommen, zum anderen wurden fünf Experten hinsichtlich ihrer Einschätzung zur Durchführbarkeit einer ambulanten Planung befragt.

Aufgrund dieser Vorarbeiten beschloss das Obsan, ein grösseres Projekt («Démographie Médicale») zu lancieren, das sich in einer ersten Phase vor allem auf die Erarbeitung der Angebotsseite der medizinischen ambulanten Versorgung konzentrieren wird. Auf der Seite der Nachfrage steht in der wissenschaftlichen und politischen Diskussion immer wieder die Frage der Angebotsinduzierung im Vordergrund. Um für das Projekt «Démographie Médicale» über Vorarbeiten zu verfügen, wurde das Büro BASS damit beauftragt, mit einer kurzen Literaturlausanalyse den Themenbereich Nachfrage, Inanspruchnahme, Bedarf und Angebotsinduzierung auszuleuchten. Weiter sollten verschiedene Varianten aufgezeigt werden, wie der Bedarf bzw. die Angebotsinduzierung konkret bestimmt werden können.

Fragestellungen

Namentlich sollten folgende Fragen beantwortet werden:

Wie kann methodisch vorgegangen werden, um die Inanspruchnahme nach ambulanten Leistungen zu bestimmen?

Wie wird die angebotsinduzierte Inanspruchnahme berücksichtigt?

Wie wird die nachfrageinduzierte Inanspruchnahme berücksichtigt?

Welche Vor- und Nachteile haben die einzelnen Methoden mit Blick auf die im Rahmen des Gesamtprojektes «Démographie Médicale» in Zukunft evtl. geplanten Nachfrageschätzungen in der Schweiz?

Welche empirischen Ergebnisse zur Nachfrageschätzung liegen für die Schweiz vor? Welche Konsequenzen sind daraus für zukünftige Schätzungen abzuleiten?

Methodisches Vorgehen und Aufbau des Berichtes

In einem ersten Schritt wurde eine kurze Literaturlausanalyse zu empirischen Forschungsarbeiten durchgeführt, welche sich mit dem Thema von Nachfrageschätzungen im Gesundheitswesen befassen. Es

³ Spycher Stefan (2004): Prognose und Planung in der ambulanten Gesundheitsversorgung. Literaturlausanalyse und Expertengespräche zur Prognose und Planung des medizinischen Personals in der Schweiz. Arbeitsdokument Nr. 5. Herausgegeben vom Schweizerischen Gesundheitsobservatorium, Neuenburg.

ging dabei nicht darum, eine vollständige Literaturübersicht herzustellen.⁴ Vielmehr wurden mit einem pragmatischen Vorgehen die wichtigsten Theorien, Konzepte und Begrifflichkeiten, die für die Schätzungen verwendeten Daten und die ökonometrischen Modellansätze sowie die dabei auftretenden theoretischen und methodischen Probleme identifiziert und daraus resultierend der Stand der empirischen Forschung zum Thema ermittelt. Dabei wurde deutlich, dass schon die Verwendung des Begriffs der Nachfrage in der gesundheitswissenschaftlichen Literatur nicht einheitlich ist. Es bestehen terminologische Unklarheiten bei den eindeutigen Abgrenzungen zu den Begriffen «Bedarf», «Nachfrage» und «Inanspruchnahme». Dies ist insofern von Bedeutung, weil dabei deutlich wird, dass die sprachliche Differenzierung Folgen für die theoretische Ausrichtung, die Spezifizierung der Modelle und die Analyse der empirischen Erhebungen hat. In Kapitel 2 werden deshalb zuerst diese begrifflichen Unschärfen geklärt, um anschliessend die mit diesen Begriffen operierenden verschiedenen Ansätze mit den entsprechenden empirischen Resultaten vorzustellen.

Anhand der aus der Literaturübersicht gewonnenen Erkenntnisse werden im zweiten Teil dieses Berichtes verschiedene Vorschläge präsentiert, mit welchen Daten und welchen Modellen die Nachfrage konkret geschätzt bzw. die Inanspruchnahme ambulanter Gesundheitsleistungen analysiert werden könnte.

⁴ Es wurde von bereits bekannten Arbeiten ausgegangen, eine Internet-Suche und eine Suche auf der ökonomischen Datenbank ECONLIT durchgeführt.

Teil 1: Literaturanalyse

2 Theorien, Begriffe und Konzepte

Empirische Forschung will Realität beschreiben und erklären sowie auf der Basis begründeter Erklärungen Prognosen zukünftiger Entwicklungen ableiten⁵. Beschreibungen von Realität erfordern zunächst ein Instrumentarium an Begriffen. Erklärungen verlangen Theorien bzw. Hypothesen, Prognosen setzen begründete Erklärungen voraus. Umsetzungs- bzw. Steuerungsrelevanz verlangt darüber hinaus, dass die Ergebnisse der empirischen Analysen eingebunden werden in ziel- und programmorientierte Planungs- und Entscheidungsprozesse. Beschreibungen, Erklärungen oder Prognosen an sich haben keine Umsetzungs- oder Steuerungsrelevanz (Andersen 2003).

Die Inanspruchnahme medizinischer Versorgung wird bereits seit den frühen 1960er Jahren empirisch untersucht. Nachdem eine rein demographisch-deskriptive Betrachtung der Inanspruchnahme schnell als unbefriedigende Informationsbasis angesehen wird (Anderson 1963: S. 364), entwickeln die sozialwissenschaftlich orientierten Disziplinen Verhaltensmodelle der Inanspruchnahme medizinischer Versorgung (vgl. Rosenstock 1996). Parallel zu der sozialwissenschaftlichen Perspektive entwickelt sich in der Gesundheitsökonomie ein eigener Theoriezweig zur Inanspruchnahme medizinischer Versorgung. Dieser befasst sich u.a. mit Problemen der Steuerung, die durch die besonderen Eigenschaften der Gesundheit und seiner Produktion hervorgerufen werden. Sozialwissenschaftliche und gesundheitsökonomische Modelle überschneiden sich in weiten Bereichen zumindest bezüglich ihrer empirischen Umsetzung, da sie meistens auf dieselben Datenquellen und auch auf dieselben Determinanten zurückgreifen. Sie weisen jedoch einen unterschiedlichen theoretischen Zugang auf, weshalb sie im Rahmen dieser kurzen Darstellung getrennt betrachtet werden (Kurth et al. 2004: S.9).

2.1 Modelle der Gesundheitsökonomie

Die Inanspruchnahme ambulanter ärztlicher Leistungen wird auf der einen Seite von den Interessen potentieller Patienten/innen auf der Nachfrageseite und den niedergelassenen Ärzten/Ärztinnen auf der Angebotsseite massgeblich beeinflusst. Beide Seiten werden durch Modelle beschrieben, welche das Verhalten der Akteurinnen und Akteure auf ökonomisch rationale Entscheidungen zurückführen. Der klassische fallende Verlauf einer Nachfragekurve im Preis-/Mengendiagramm muss allerdings modifiziert werden, da erstens in der Regel ein Krankenversicherungsschutz besteht, der trotz eines hohen Preises der Gesundheitsleistung dafür sorgt, dass der Patient bzw. die Patientin keine (oder nur geringe) Kosten selbst tragen muss, und zweitens die Preiselastizität der Nachfrage bei bestimmten Indikatoren (potentiell bei lebensbedrohenden Erkrankungen) sehr gering ist (Schöffski 2004: S.4).

2.1.1 Die Nachfrageseite

Im Sinne der herkömmlichen Produktionstheorie kann Gesundheit als Gut mit besonderen Eigenschaften betrachtet werden, das mit bestimmten Einschränkungen ebenso wie andere Güter produziert werden kann. Das Konzept der Gesundheitsproduktion ermöglicht es, die Optimalität des Verhaltens der Individuen und die Effizienz des Einsatzes knapper Ressourcen zu bewerten (vgl. Breyer und Zweifel 1997: S. 64f).

⁵ Kapitel 2 orientiert sich hauptsächlich an der Arbeit von Kurth B.-M. et al. (2004: S.9-18). Im Rahmen einer empirischen Analyse zur Identifikation sowohl von patienteninduzierten als auch von angebotsinduzierten Nachfragedeterminanten stellen sie den Stand der Forschung und die verschiedenen Modelle ausführlich dar.

Die Inanspruchnahme medizinischer Leistungen stellt in diesem Rahmen einen Faktor des Produktionsprozesses von Gesundheit dar, zu dem neben medizinischen Leistungen auch andere Faktoren, wie z.B. der Einsatz eigener Zeit, beitragen können. Dabei stellt sich auch die Frage, unter welchen Bedingungen medizinische Leistungen und andere Faktoren zur Produktion von Gesundheit gegenseitig substituiert werden können bzw. komplementär zueinander wirken. Das Gut Gesundheit wird nicht nur als Konsumgut hoch geschätzt, sondern dient auch als Voraussetzung für die Produktion und den Genuss anderer Güter. Gleichwohl konkurriert die Produktion von Gesundheit mit der Produktion von Konsumleistungen, sofern das Individuum zur Verbesserung oder zum Erhalt seines Gesundheitszustandes auf andere Konsumleistungen verzichten muss. Das Individuum wägt in seiner Entscheidung unter den jeweiligen Bedingungen den erwarteten Nutzen und die erwarteten Kosten der Gesundheitsproduktion gegeneinander ab, dies gilt im Besonderen auch für die Inanspruchnahme medizinischer Leistungen. Dabei sind nicht nur die monetär bewerteten Kosten einer ärztlichen Behandlung von Bedeutung, sondern auch die eigene Zeit, die in die Produktion der Gesundheit investiert wird (vgl. Breyer und Zweifel 1997).

Das Grossman-Modell

Grossman formulierte in den frühen 1970er Jahren ein Modell der Inanspruchnahme medizinischer Leistungen, das als einer der wichtigsten Beiträge der Gesundheitsökonomie betrachtet wird (vgl. Grossman 2000; Leu und Gerfin 1992). In dem Modell wird erstens die Nachfrage nach medizinischen Leistungen als abgeleitete Nachfrage interpretiert, d.h. das Individuum fragt Gesundheit und nicht medizinische Leistungen per se nach. Zweitens ist der Gesundheitskapitalbestand abhängig von der Inanspruchnahme medizinischer Leistungen und einer Abschreibungsrate, die explizit als Funktion des Alters (des Gesundheitskapitalbestandes) und der Intensität seiner Nutzung definiert ist (vgl. Grossman 1972a, 1972b, 2000). Über die Abschreibungsrate können auch Merkmale wie das Gesundheitsverhalten oder das Bildungsniveau in das Modell integriert werden. Das Individuum maximiert intertemporär seine Nutzenfunktion, die aus dem Gesundheitskapital und einem zusammengesetzten Konsumgut besteht. Schliesslich bedeutet drittens ein höherer Gesundheitskapitalbestand gleichzeitig konsumtiven wie investiven Nutzen. Damit wird unter anderem ein Zusammenhang zwischen dem Erwerbseinkommen und dem Gesundheitszustand hergestellt.

Breyer und Zweifel (1997) betonen in ihrer Kritik am Grossman-Modell die Unsicherheit der Gesundheitsproduktion. Der Gesundheitszustand wird nicht allein durch die vom Individuum steuerbaren Inputs, sondern auch von einer Vielzahl unkontrollierbarer äusserer Faktoren beeinflusst. Entsprechend kann das Individuum auch nur die Wahrscheinlichkeit beeinflussen, nachfolgende Perioden gesund zu erleben. So weicht nach Breyer und Zweifel der tatsächliche Gesundheitszustand in fast allen Fällen vom (unter ökonomischen Gesichtspunkten) optimalen Gesundheitszustand ab, woraus sich im Vergleich zum Grossman-Modell andere Folgerungen für den Einfluss des Gesundheitszustandes auf die Inanspruchnahme ergeben (vgl. Zweifel 1992: S. 23f., Breyer und Zweifel 1997: S. 75). Darüber hinaus nehmen die Autoren an, dass das Individuum nur im gesunden Zustand seinen Gesundheitszustand durch eigene Mittel erhalten kann, während es im Krankheitsfall auf medizinische Leistungen angewiesen ist. Der Gesundheitszustand ist damit nicht nur das Ergebnis des Prozesses zur Herstellung/Erhaltung von Gesundheit. Er ist gleichzeitig auch ein entscheidender Inputfaktor vor diesem Prozess. Als Inputfaktor bestimmt er die Möglichkeiten der Produktion für die folgenden Perioden.

Weiterentwicklung des Grossman-Modells

Breyer und Zweifel entwickeln ein Modell, in dem ein Trade-Off zwischen Konsum und Gesundheit stattfindet. In Abhängigkeit vom Gesundheitszustand der Ausgangslage und der zeitlichen Perspektive verändert bzw. verschiebt sich die Transformationskurve bei exogenen Veränderungen. Medizinisch-technologischer Wandel, erhöhte Anbieterdichte und der Ausbau des Krankenversicherungsschutzes führen demnach zu einer stärkeren Inanspruchnahme des Gesundheitswesens. Möglichkeiten, einer daraus resultierenden Kostenexpansion zu begegnen, sehen Breyer und Zweifel im Rahmen ihres Modells in einer Erhöhung der Preise für medizinische Leistungen bzw. einer Erhöhung der Produktivität eigener Anstrengungen zum Erhalt der Gesundheit. Letzteres wird von den Autoren jedoch kritisch gesehen, da eigene Anstrengungen und die Inanspruchnahme medizinischer Leistungen als Produktionsfaktoren für Gesundheit nur im gesunden Zustand gegenseitig substituierbar sind; im kranken Zustand wirken diese beiden Faktoren im Modell der Autoren komplementär (Breyer und Zweifel 1997: S. 96f).

Ein Aspekt der von den vorgenannten Modellen noch nicht berücksichtigt wird, ist das Vorliegen einer Moral Hazard Situation durch die Absicherung des Krankheitsfalls durch die Krankenversicherung. Mit dem Versicherungsschutz reduzieren sich die (Grenz-)Kosten für die Inanspruchnahme einer medizinischen Leistung deutlich, womit sich auch die Verhaltensanreize für ein ökonomisch rational handelndes Individuum verändern: Im Vergleich zu einer Situation, in der die Kosten vom Individuum selbst getragen werden, wird die Krankheitsvorbeugung eher vernachlässigt, während beim Eintreten des Krankheitsfalls eher mehr medizinische Leistungen in Anspruch genommen werden (vgl. Schneider 2003: S. 14).

2.1.2 Die Angebotsseite – angebotsinduzierte Nachfrage

Traditionell wird die Angebotsseite im Gesundheitswesen mindestens in die vier grossen Bereiche ambulante medizinische Versorgung, stationäre Versorgung, Arzneimittel sowie Heil- und Hilfsmittel unterschieden. Im vorliegenden Bericht beschränken wir uns auf den ambulanten Sektor.

Der ambulante Sektor ist in erster Linie durch den/die freiberuflich tätige/n Ärztin/Arzt in eigener Praxis gekennzeichnet. So wie die Theorie ein Modell der Nachfrage nach medizinischen Leistungen aufstellt, so versucht sie, ein entsprechendes Modell für das Angebot der niedergelassenen Ärzte und Ärztinnen zu formulieren. Hier geht es insbesondere um die Frage, durch welche Determinanten der Arzt bzw. die Ärztin sein/ihr eigenes Leistungsangebot ausweitet bzw. qualitativ verändert. Bedeutende Determinanten sind bspw. die Zahl der Kolleg/innen, mit denen der Markt geteilt werden muss, die Honorierungsform, die Höhe der Honorierung und ähnliches⁶.

Speziell wurde in der Gesundheitsökonomie das Thema der «angebotsinduzierten Nachfrage» thematisiert und untersucht. Dabei wird davon ausgegangen, dass die Ärztinnen bzw. der Ärzte aufgrund der beratenden Funktion, die sie durch ihren Informationsvorsprung innehaben, erheblich auf die Art und den Umfang der Inanspruchnahme medizinischer Leistungen einwirken. Von einer Angebotsinduzierung ist dabei aber nur dann die Rede, wenn sie bei der Behandlung ihre eigenen Interessen über die der Patient/innen stellen und ihre Position dazu nutzen, mehr Leistungen zu erbringen, als sie als perfekte Sachverwalter/innen erbringen würden. Angebotsinduzierung setzt voraus, dass für die Ärztin bzw. den Arzt Anreize vorhanden sein müssen, die Menge ihrer/seiner Leistungen auszuweiten. Es wird davon ausgegangen, dass die Ärztinnen und Ärzte erstens über ein begrenztes Zeitbudget verfü-

⁶ Vgl. Herder-Dorneich (1994: S. 351-358)

gen und zweitens dass jede/r bezüglich ihres/seines Einkommens ein Anspruchsniveau hat, das sie/er unbedingt realisieren möchte (Zieleinkommenshypothese). Wird das Zieleinkommen unterschritten, so verlieren die beiden übrigen Zielgrößen, Arbeitszeit und Berufsethik, ihre Bedeutung. Wird es erreicht oder überschritten, so führt das zusätzlich erzielte Einkommen zu keinem zusätzlichen Nutzen (Breyer, Zweifel 1999: S.251f.).

Ausprägungen angebotsinduzierter Nachfrage⁷

Existenz und Umfang arztinduzierter Nachfrage hängt von der Wissens- und Informationsposition des/der Patienten/in, der Interessenslage des Arztes bzw. der Ärztin, dem Verhältnis der Arzt-PatientInnen-Beziehung sowie von institutionellen Rahmenbedingungen ab. Zur Beurteilung angebotsinduzierter Nachfrage ist zudem von Bedeutung, welches Behandlungsergebnis trotz induzierter Nachfrage erreicht wird. Der Outcome ist zu berücksichtigen. So kann zwischen (sinnvoller) arztinduzierter Nachfrage zu Gunsten eines besseren Behandlungsergebnisses, („extern verursachter“) arztinduzierter Nachfrage wegen institutionellen und rechtlichen Rahmenbedingungen und der Versorgungsqualität mindernden, (negativen) arztinduzierten Nachfrage unterschieden werden.⁸ Labelle et al. (1994) unterscheidet sechs verschiedene Situationen (vgl. **Tabelle 1**).

Tabelle 1: Ergebnisorientierte Einordnung arztinduzierter Nachfrage

		Haben die ärztlichen Leistungen den Gesundheitszustand positiv verändert?		
		JA	NEIN	
			Neutral	Verschlechterung
Hätte der Patient / die Patientin die Leistungen bei perfekter Information nachgefragt?	JA	A	a	c
	NEIN	B	b	d

Quelle: Darstellung nach Labelle et al. (1994)

Aus Tabelle 1 wird deutlich, dass die Situation A keine arztinduzierte Nachfrage im Sinne der perfect-agency darstellt. Arztinduzierte Nachfrage im Sinne der perfect-agency ergibt sich nur, wenn der/die Patient/in die Leistung bei perfekter Information nicht nachgefragt hätte. In Situation B liegt arztinduzierte Nachfrage vor, da der Patient die Leistungen nicht nachgefragt hätte. Jedoch haben die zusätzlichen Leistungen das Behandlungsergebnis positiv beeinflusst. In einer solchen Situation besteht ein unterschiedliches Verständnis zwischen Ärztin/Arzt und Patientin, ob eine Leistung in Anspruch genommen werden sollte. Die Präferenzen des/der Patienten/in führen dazu, dass der Konsum der Leistung abgelehnt wird. Dies kann sowohl an der Zahlungsbereitschaft und Zahlungsfähigkeit als auch aus unterschiedlich bewerteten Begleiterscheinungen oder Folgen einer Behandlung resultieren. Diese Situation kann mit dem Nichteinlösen eines Rezepts oder der Nichteinnahme eines verordneten Arzneimittels verdeutlicht werden.

Die Situationen a und c unterscheiden sich durch den Einfluss auf den Gesundheitszustand mit Therapie. In Situation a ergibt sich weder eine Verbesserung noch eine Verschlechterung, in Situation c verschlechtert sich der Gesundheitszustand durch die Inanspruchnahme von Gesundheitsleistungen. Sowohl in a und in c ist die/der Ärztin/Arzt zwar perfekte/r Sachwalter/in des/der Patienten/in, die Ergebnisse sind jedoch nicht zufrieden stellend. Obgleich diese Situation keine arztinduzierte Nachfrage

⁷ Der folgende Abschnitt stützt sich stark auf Kern (2002: S. 18-20)

⁸ Vgl. Fuchs (1978), S. 36; vgl. Labelle et al. (1994), S. 354.

per Definition darstellt, so erscheint es im Sinne des Effizienzgedankens zumindest erforderlich, diese Leistungen hinsichtlich der Effektivität zu überprüfen.

In den Situationen b und d wird die Form der arztinduzierten Nachfrage angetroffen, die vom Grundsatz her abgelehnt wird, und aus der sich die negative Konnotation des Begriffs der arztinduzierten Nachfrage entwickeln konnte. Dabei stellt Situation b «nur» Verschwendung dar. Situation d bedeutet nicht nur Verschwendung, sondern auch intangible Kosten für den Patienten.

Daraus folgt, dass Leistungen, die zu den Situationen a, b, c und d führen, zuerst hinsichtlich ihrer Wirksamkeit zu überprüfen sind. So wird deutlich, dass in den Situationen b und d die Patient/innen bei Kenntnis der Informationen über die Qualität der Leistungen diese nicht nachgefragt hätten. Nur sie gelten als «klassische» Fälle der nicht erwünschten angebotsinduzierten Nachfrage.

Notwendige Bedingungen für angebotsinduzierte Nachfrage

Damit sich angebotsinduzierte Nachfrage überhaupt realisieren kann, müssen verschiedene Faktoren notwendigerweise erfüllt sein. Gemäss Kern (2002) gehören dazu:

Erstkontakt: Der Erstkontakt eines/r Patienten/in ist Voraussetzung, damit angebotsinduzierte Nachfrage überhaupt stattfinden kann. In manchen Modellen wird davon ausgegangen, dass der Erstkontakt insgesamt stärker patienten/innen-induziert ist, währenddem die Leistungen während der Behandlungsdauer insgesamt stärker von der Ärztin bzw. dem Arzt bestimmt sind. Es gibt deshalb empirische Untersuchungen, welche Erstkontakt und Behandlungsweg mit unterschiedlichen Modellen und Verfahren modellieren.

Informationsunsicherheiten: Patientinnen und Patienten nehmen Informationen subjektiv verschieden auf und bewerten sie auch unterschiedlich (Simon 1982). Zudem ist auch der Zugang zu Informationen für Patient/innen eingeschränkt. Die Informationsunsicherheit der Patient/innen führt dazu, dass die Nachfrage nach Gesundheitsleistungen von PatientInnenseite normalerweise weder nach Art noch nach Menge bestimmt werden kann, da man kaum in der Lage ist, den privaten Nutzen von Gesundheitsleistungen zu erkennen und zu bewerten. Diese Informationsunsicherheit seitens der Patient/innen ist Voraussetzung dafür, dass die Ärztinnen und Ärzte Angebotsinduzierung realisieren können.

Einkommensinteresse der Ärztinnen und Ärzte: Angebotsinduziertes Verhalten macht nur dann Sinn, wenn die Ärztinnen und Ärzte einen Anreiz haben, gesundheitliche Güter und Dienstleistungen über das wie auch immer definierte, erforderliche «objektive» Mass hinaus zu veranlassen. Damit müssen die Ärztinnen und Ärzte erstens ein Einkommensinteresse haben (allenfalls noch ein Forschungsinteresse) und zweitens müssen sie die Möglichkeit besitzen, irgendwie Einkommen aus der Vergütung von Behandlungsleistungen zu erzielen, weshalb auch der *Vergütungsform* eine entscheidende Rolle zukommt. Bei Diagnostik und Therapie hat die Ärztin/der Arzt insbesondere bei Einzelleistungsvergütung einen Anreiz, mehr diagnostische und therapeutische Massnahmen zu empfehlen, als für eine sichere Diagnosestellung und Behandlung erforderlich wären. Gemäss Schulenberg (1981) wird davon ausgegangen, dass eine Einzelleistungshonorierung mengenexpansiv und eine Pauschalvergütung mengenkontraktiv wirkt.

Therapiefreiheit und -alternativen: Angebotsinduzierte Nachfrage ist nur dann möglich, wenn Therapiefreiheit und damit Therapiealternativen möglich sind. Dies ist dann von Bedeutung, wenn keine eindeutig kausale Therapie bestimmt werden kann und sich aus der Diagnose kein Therapieregime ableiten lässt.

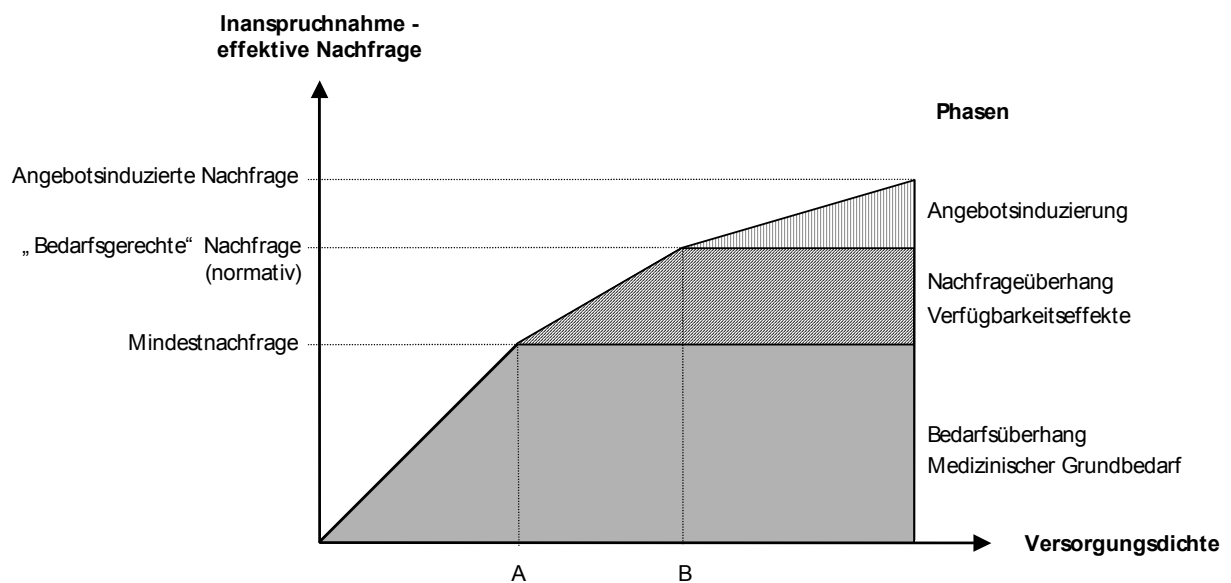
Behandlungsmonopol der Ärztinnen und Ärzte: Theoretische und empirische Untersuchungen zur arztinduzierten Nachfrage unterstellen, dass dieselbe Ärztin bzw. derselbe Arzt nach Diagnosestellung

auch die Therapie durchführt. Damit wird ein paternalistisches Beziehungsverhältnis zwischen Ärztin/Arzt und Patient/in unterstellt. Wenn es den Patient/innen möglich ist, die Ärztin oder den Arzt zu wechseln, so wird die Höhe bzw. der Umfang der anbieterInneninduzierten Inanspruchnahme geringer als bei einem reinen Informationsmonopol (Pauly und Satterhwaite 1981).

Ambulante Inanspruchnahme und Angebotsentwicklung: ein Modell

In **Abbildung 1** wird ein Modell dargestellt, welches den Zusammenhang zwischen der Inanspruchnahme und der Angebotsentwicklung mit verschiedenen Phasen graphisch darstellt (vgl. dazu Andersen 1997: Kap. 3). Die Höhe des Versorgungsniveaus – gemessen an der Dichte der Allgemein- und FachärztInnen – spielt dabei in Bezug auf die Angebotsinduzierung eine zentrale Rolle. Dabei werden vereinfachte Annahmen über das Verhalten der PatientInnen einerseits und ÄrztInnen andererseits zugrunde gelegt. Bezogen auf das ÄrztInnenverhalten gehen lediglich das (begrenzte) Zeitbudget und die «Zieleinkommenshypothese» in die Beschreibung ein⁹. Es wird davon ausgegangen, dass die ohne Zweifel vereinfachten Annahmen ausreichend sind, um die vermuteten grundlegenden Zusammenhänge plausibel zu machen.

Abbildung 1: Modell des Zusammenhangs zwischen Versorgungsdichte (Allgemein- und FachärztInnen) und ambulanter Inanspruchnahme



Quelle: Eigene Darstellung

Unabhängig von der ÄrztInnendichte besteht in einer Gesellschaft ein *Grundbedarf* nach medizinischer Versorgung bzw. medizinischen Leistungen. In dieser Phase gilt für die PatientInnen ausschliesslich der «biologische Imperativ». Das Zeitbudget der ÄrztInnen wird durch die Menge der PatientInnen vollständig aufgebraucht. Zeit ist (zu) knapp und Überlegungen zur Qualität sind kaum möglich. In dieser Phase kann der biologisch bedingte Bedarf vom Versorgungsangebot nicht gedeckt werden. Bei steigender Versorgungsdichte, d.h. es kommen mehr ÄrztInnen auf den Markt, steigt die Inanspruchnahme deshalb linear zur Angebotserhöhung. Die Elastizität liegt in dieser Situation bei 1.

⁹ Breyer/Zweifel (1997: S. 246 ff.)

Wir gehen davon aus, dass der Grundbedarf bzw. die Mindestnachfrage mit einem gewissen Angebot an Ärztinnen und Ärzten gedeckt werden kann. In Abbildung 1 ist dies bei der Versorgungsdichte in Punkt «A» erreicht.

Steigt in dieser Situation die Ärztinnendichte weiter an, wird das Verhalten der ÄrztInnen und PatientInnen interdependent. Es beginnt der Bereich der *Verfügbarkeitseffekte bzw. des Nachfrageüberhangs*. Entscheidungen der PatientInnen werden nicht nur noch unter dem Aspekt des «biologischen Imperativs» getroffen, und für die ÄrztInnen gilt nicht mehr ausschliesslich bzw. primär das limitierte Zeitbudget. Für sie beginnen auch Überlegungen, wie das Zieleinkommen zu realisieren ist. Charakteristisch für diese Phase ist, dass sowohl für ÄrztInnen als auch für PatientInnen in Bezug auf das Inanspruchnahmegeschehen neue Motive und Nutzenüberlegungen ins Spiel kommen. Die ÄrztInnen können bspw. die Qualität der Behandlung erhöhen, was neben einer direkten Zunahme der Inanspruchnahme u.U. auch eine erhöhte Nachfrage mit sich ziehen kann. Auch wird der Zugang zu medizinischen Angeboten für die PatientInnen erleichtert, was zu Nachfragesteigerungen führen kann. Andere Erwartungshaltungen seitens der PatientInnen können zu unterschiedlichem Nachfrageverhalten führen. Zudem bekommt auch das System der Gesundheitsversorgung, bspw. die Art der Versicherung oder die Honorierungsform der ÄrztInnen, eine wichtige Rolle.

Zusammen ergibt dies insgesamt eine Steigerung der Inanspruchnahme, die auf sehr komplexe Erklärungsmuster zurückzuführen ist. Verschiedene, sowohl nachfrage- als auch angebotsseitige Effekte können nur noch sehr schwer auseinandergelassen werden. Ab dieser Phase wird es auch schwierig zu beurteilen, was der «richtige» Bedarf ist. Es braucht dazu u.a. auch eine normative Festsetzung. Der «Sättigungspunkt» dieser Phase – dargestellt in der Abbildung 1 mit der Versorgungsdichte auf der Höhe «B» – ist deshalb nicht mehr nach objektiven Kriterien bestimmbar, da er auch von gesellschaftlichen und kulturellen Rahmenbedingungen und vom medizinischen Kenntnisstand abhängt.

Der Beginn der Phase der *Angebotsinduzierung* ist dadurch gekennzeichnet, dass die Auslastung der ÄrztInnen aufgrund der Zunahme der ÄrztInnendichte so weit abgesunken ist, dass sie ihr Zieleinkommen nicht mehr realisieren können. Wollen die ÄrztInnen ein bestimmtes Einkommen erzielen, müssen sie versuchen, dieses über eine künstliche Nachfrageschaffung zu erzielen. Dies ist die «klassische» Situation der Angebotsinduzierung.

Bei der Interpretation des empirisch feststellbaren, positiven Zusammenhangs zwischen Inanspruchnahme und ÄrztInnendichte kann gemäss dem vorgestellten Modell nur dann von Angebotsinduzierung gesprochen werden, wenn sowohl der medizinische Grundbedarf als auch die Verfügbarkeitseffekte, bzw. der Nachfrageüberhang vom bereitstehenden Angebot befriedigt werden können. In Bezug auf die Situation der Schweiz ist es durchaus denkbar, dass sich verschiedene Regionen in verschiedenen Phasen, welche in Abbildung 1 dargestellt sind, befinden. Für die Interpretation der Zusammenhänge zwischen Inanspruchnahme und ÄrztInnendichte muss dies unbedingt berücksichtigt werden.

Ansätze zur empirischen Überprüfung angebotsinduzierter Nachfrage

Gemäss Gaynor (1994) finden sich in der Literatur drei unterschiedliche Ansätze zur Identifizierung von induzierter Nachfrage.

Der weitest verbreitete Ansatz untersucht die Beziehung zwischen der *ÄrztInnendichte* (Anzahl Ärzte/Ärztinnen pro Person in einer bestimmten Region) und der *Pro-Kopf-Inanspruchnahme* ärztlicher Leistungen. Es gibt aus empirischer Sicht keine Zweifel darüber, dass mehr Leistungserbringer/innen (Ärzte/Ärztinnen) mit mehr abgerechneten Leistungen einhergehen. Bivariat besteht zwischen den beiden Grössen ein statistisch gesicherter, positiver Zusammenhang. Dies zeigen sowohl Analysen

mit Individual- als auch mit Aggregatdaten (vgl. bspw. Horisberger/Schmid 2004). Ob damit die These der Angebotsinduzierung nachgewiesen wird, ist in der Literatur jedoch sehr umstritten.

Unter der Gültigkeit der Zieleinkommenshypothese kann es verschiedene Gründe geben, warum mit einer Zunahme der ÄrztInnendichte die Pro-Kopf-Inanspruchnahme ärztlicher Leistungen steigt.

Besteht ein Nachfrageüberhang (d.h. es herrscht Rationierung), steigt die Inanspruchnahme linear zu der Zahl der Ärzte. Beim Anstieg handelt es sich um keine arztinduzierte Nachfragesteigerung.

Die Zunahme der Leistungsmenge bei steigender ÄrztInnendichte kann auch auf rationale Entscheidungen der Patienten zurückzuführen sein. So kann bspw. ein Rückgang der indirekten Kosten (bspw. Zeit- und Wegkosten) bewirken, dass vermehrt Leistungen nachgefragt werden.

Durch vermehrten Konkurrenzdruck bei steigender Anzahl Ärzte könnte die Qualität des ärztlichen Angebots ansteigen, was zu einem veränderten Inanspruchnahmeverhalten führen könnte.

Die Niederlassungsentscheidungen der Ärzte sind nicht exogen, sondern endogen. Das würde bedeuten, dass ein umgekehrter Kausalzusammenhang zwischen den beiden Grössen ÄrztInnendichte und Inanspruchnahme besteht. Ärzte würden sich demnach dort niederlassen, wo bspw. aufgrund einer erhöhten Morbidität eine erhöhte Nachfrage besteht.

Das bedeutet, dass für die empirische Überprüfung des beobachtbaren Phänomens die Modelle so angelegt sein müssen, dass anhand der Ergebnisse zwischen den verschiedenen Erklärungsmustern unterschieden werden kann.

Folgt man dem theoretisch begründeten Ansatz, dass der Erstkontakt insgesamt stärker patienten/innen-induziert ist, währenddem die Leistungen der Behandlungsdauer insgesamt stärker von der Ärztin bzw. dem Arzt bestimmt ist, lautet demnach die empirisch zu überprüfende Hypothese folgendermassen:

«Von angebotsinduzierter Nachfrage der Patienten nach ambulanten ärztlichen Leistungen ist dann auszugehen, wenn die Arztdichte als unabhängige Variable einerseits keinen signifikanten Einfluss auf die Entscheidung zum Erstbesuch hat, andererseits die Nachfrage nach ärztlichen Leistungen aber signifikant positiv beeinflusst. Die Hypothese müsste dagegen abgelehnt werden, wenn keine derartige Evidenz gefunden werden kann (Cassel/Wilke 2000: S. 4)».

Labelle et al. (1994) erwähnen in Ihrem Übersichtsartikel zur angebotsinduzierten Nachfrage, dass nicht nur widersprüchliche Resultate zwischen verschiedenen Studien zur Angebotsinduzierung vorliegen, sondern sogar dieselben Resultate von verschiedenen Forschern unterschiedlich interpretiert würden (bspw. Rossiter and Wilensky (1983) versus Stano (1987)). Sie führen dies u.a. auch auf unterschiedliche ideologische Standpunkte der Forschenden zurück. *«As Reinhardt (1985) has pointed out, much of the controversy can be attributed to ideological predispositions, and in particular to analysts' degree of commitment to neoclassical economic theory. Typically, those who believe that supplier-induced demand (SID) is quantitatively important – the 'Bs' (broad economists), to borrow the labels used by Evans (1976) and Feldman and Solan (1988) – criticize studies that find no or little evidence of inducement. Conversely, the 'Ns' (narrow economists) challenge SID findings which suggest significant and/or widespread inducement»* (Labelle et al. 1994: S. 350f.).

Eine zweiter Ansatz untersucht die Auswirkungen, die durch Veränderungen der Rahmenbedingungen auf das Angebot von oder die Nachfrage nach ärztlichen Leistungen ausgelöst werden. Dazu gehören bspw. die Analysen von verschiedenen Rückerstattungsschemen für Ärzt/innen hinsichtlich ihrer Anreizwirkungen (vgl. bspw. Delattre et al. 2003).

Der dritte Ansatz legt seinen Fokus auf die Informationsasymmetrie zwischen Patient/in und Arzt/Ärztin. Dabei wird untersucht, ob ein Zusammenhang zwischen dem Informationsstand der Patient/innen und der Nachfrage nach Gesundheitsleistungen besteht. Interessant sind dabei die gegenteiligen Resultate zwischen Untersuchungen aus den USA und solchen aus Europa. Da die Struktur des amerikanischen Gesundheitssystems jedoch sehr verschieden von jenem in Europa ist, müssen die verschiedenen Resultate nicht per se einen Widerspruch darstellen.

Allen Erklärungsmodellen arztinduzierter Nachfrage ist gemeinsam, dass die ärztlichen Leistungen auch hinsichtlich der Qualität als homogen betrachtet werden. Damit sind bspw. nicht beobachtbare Qualitätsunterschiede immer die «Achillesferse» für die Überprüfung mittels empirischer Daten.

2.2 Sozialwissenschaftliche Verhaltensmodelle

Die sozialwissenschaftlichen Modelle, welche das Inanspruchnahmegeschehen beschreiben, sind in ihrem theoretischen und methodischen Zugang sehr heterogen (vgl. Schaufelberger et al. 1989). Je nachdem, auf welchen Ausschnitt aus dem gesamten Inanspruchnahmegeschehen sie sich konzentrieren, greifen sie auf sehr unterschiedliche Ansätze und Theorien zurück (bspw. entscheidungs- und spieltheoretische Modelle oder motivations- und lerntheoretische Modelle¹⁰). Aufgrund der unterschiedlichen Blickwinkel sind die Resultate kaum miteinander zu vergleichen (vgl. McKinlay 1972).

Aufgrund der zwischen den verschiedenen Ansätzen unvereinbaren Ergebnisse entwickelte sich in den späten 1970er Jahren eine Diskussion, die die Untersuchungsansätze zur Inanspruchnahme in eher «klassische» sozialpsychologische Verhaltensanalysen und eher multivariat-prädiktiv orientierte Modelle spaltete (vgl. Mechanic 1979). Wie Härtel (1985) feststellt, scheint diese Unterscheidung heute an Bedeutung verloren zu haben, da sich die «traditionellen» sozialwissenschaftlichen Verhaltensanalysen immer stärker in Richtung multivariater Modelle entwickelten. Unterschiedliche theoretische Orientierungen erscheinen dadurch oft nur als bestimmte Interessensschwerpunkte innerhalb eines grösseren Variablenzusammenhangs. Dies ist nicht zuletzt durch die Notwendigkeit einer umfassenden Confounderkontrolle¹¹ bedingt, durch die gezwungenermassen auch solche Faktoren berücksichtigt werden, die nicht im Mittelpunkt des jeweiligen Interesses stehen (vgl. Cummings et al. 1980).

Das Problem der Vergleichbarkeit der empirischen Ergebnisse bleibt gleichwohl bestehen und betrifft nicht nur den Bereich der sozialwissenschaftlichen Modelle, sondern genauso ökonomische Modelle. Letztlich ist das Problem für einzelne erklärende Merkmale nur im Kontext des gesamten Modells und unter Berücksichtigung der jeweiligen Rahmenbedingungen zu beantworten. Um eine gewisse Beliebigkeit einzugrenzen und eine Interpretationsbasis für die empirischen Ergebnisse zu haben, wäre es wichtig, auf einen umfassenden Modellrahmen zurückgreifen zu können. Das Verhaltensmodell nach Andersen ist vergleichsweise umfassend konzipiert und ermöglicht die Integration der zum Teil präziseren Hypothesen der vorgenannten Modelle. Durch die Herstellung von Bezügen zwischen den Modellen kann das Modell nach Andersen damit auch in einen grösseren wissenschaftlichen Kontext eingebettet werden.

¹⁰ Cummings (1980) bspw. unterscheidet 14 verschiedene Ansätze.

¹¹ Confounder bzw. konfundierende Faktoren sind sog. Drittvariablen, welche einen von den Kontrollvariablen unabhängigen Einfluss auf das zu erklärende Merkmal/Ereignis besitzen, wodurch die Beziehungen zwischen der abhängigen und den unabhängigen Variablen verzerrt wird.

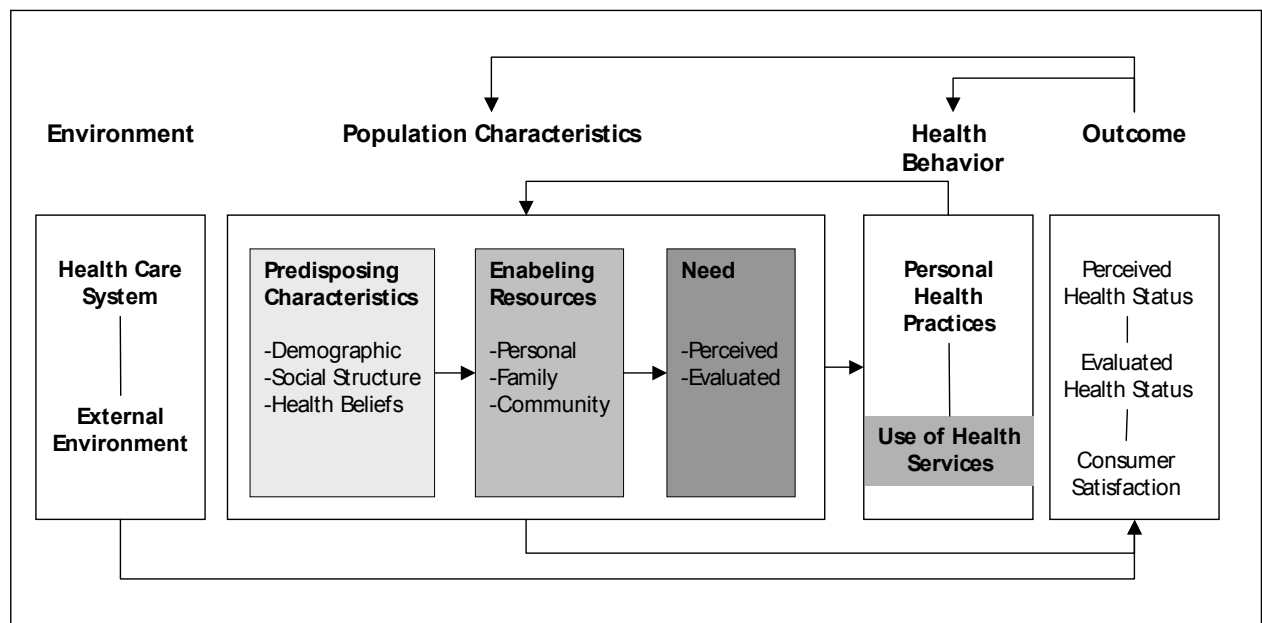
Das Verhaltensmodell nach Andersen

Andersen und seine Mitarbeiter veröffentlichten Ende der 1960er / zu Beginn der 1970er Jahre ein Verhaltensmodell zur Modellierung des Inanspruchnahmegeschehens von Gesundheitsleistungen, das seither beständig weiter entwickelt wird (Andersen 1968; Andersen und Newman 1973; Andersen 1995). Dieses Modell hat sich international vielfach als Bezugsmodell für die Analyse der Inanspruchnahme bewährt (vgl. De Boer et al. 1997: S. 102; Phillips et al. 1998).

Abbildung 2 zeigt das Modell nach Andersen auf dem Stand von 1995. Im Zentrum der Abbildung stehen die Kategorien *Predisposing Characteristics*, *Enabling Resources* und *Need* als Einflussgrößen auf die Inanspruchnahme von Gesundheitsleistungen.

Unter dem Begriff *Enabling Resources* sind die notwendigen Voraussetzungen für die Inanspruchnahme zusammengefasst, während *Predisposing Characteristics* und *Need* die Inanspruchnahme direkt bzw. indirekt beeinflussen. Dieser Modellkern wurde später um die Komponenten *Outcomes* und *Environment* ergänzt und schliesslich um Rückkopplungen von *Outcomes* auf *Population Characteristics* und auf das Gesundheitsverhalten vervollständigt.

Abbildung 2: Verhaltensmodell, adaptiert nach Andersen (1995)



Quelle: Kurth et al. (2004: S.16)

Mit dem Begriff *Outcomes* werden die Resultate der Inanspruchnahme von Gesundheitsleistungen benannt, wie sie sich einerseits im subjektiven und objektiven Gesundheitszustand und andererseits auch in der Zufriedenheit mit der Versorgung zeigen. *Environment* bezeichnet die äussere Umwelt, den gesamtgesellschaftlichen Kontext und damit auch das Gesundheitssystem, in dem die Inanspruchnahme stattfindet. Dies ist vor allem in länderübergreifenden Studien oder im Zeitvergleich von Bedeutung. Da die zusätzlichen Komponenten des Modells, besonders die hinzugekommenen Rückkopplungsschleifen, in einer Querschnittsuntersuchung schwerlich abgedeckt werden können, stützt sich die vorliegende Beschreibung hauptsächlich auf den Modellkern, dessen Elemente im Folgenden ausführlich dargestellt werden.

Predisposing Characteristics

Unter dem Begriff Predisposing Characteristics fasst Andersen alle Merkmale zusammen, die sich indirekt auf die Inanspruchnahme auswirken. Diese Merkmale werden in die Bereiche Demographie, Sozialstruktur und Health Beliefs aufgeteilt.

Demographie. Unter dem Begriff Demographie subsumiert Andersen Merkmale wie Alter und Geschlecht, die sich als «biologischer Imperativ» auf die Inanspruchnahme auswirken. Mit dem Hinweis auf einen «biologischen Imperativ» sollte jedoch nicht die soziale Dimension von Alter und Geschlecht vernachlässigt werden. Obwohl sowohl Alter als auch Geschlecht untrennbar mit den körperlichen Bedingungen verknüpft sind, ist unbestritten, dass sich Alter und Geschlecht darüber hinaus durch soziale Zusammenhänge vielfältig auf die Inanspruchnahme auswirken, beispielsweise in Form von Rollenerwartungen oder Kohorten- und Periodeneffekten.

Sozialstruktur. Unter dem Begriff Sozialstruktur werden sämtliche Merkmale zusammengefasst, die anzeigen, welche Möglichkeiten und Ressourcen einer Person zur Lösung von Problemen zur Verfügung stehen. Dazu gehören neben den klassischen Indikatoren wie sozialer Status, Bildung, Beruf und Wohnumfeld, auch der Sozialraum, in der sich die Person bewegt, ihre sozialen Netzwerke und Interaktionsmöglichkeiten.

Health Beliefs. Health Beliefs sind Einstellungen, Werte und Wissen, die eine Person über Gesundheit und Versorgungseinrichtungen hat. Sie beeinflussen die Wahrnehmung des Bedarfs, die Einschätzung des möglichen Nutzens und der möglichen Kosten einer Inanspruchnahme und sind deshalb für den Entscheidungsprozess bedeutsam.

Enabling Resources

Enabling Resources sind die notwendige Voraussetzung für die Inanspruchnahme von Gesundheitsleistungen. Es wird zwischen personen- und gemeindebezogenen Ressourcen unterschieden.

Personal / Family. Unter diesem Punkt werden Merkmale wie Einkommen und das Vorhandensein von Krankenversicherung und Zusatzversicherungen, aber auch das für eine Inanspruchnahme nötige Wissen zusammengefasst. Diese Merkmale können als Untermenge der Predisposing Characteristic verstanden werden. Sie lassen sich von diesen dadurch abgrenzen, dass die Personal Enabling Resources nur diejenigen Ressourcen umfassen, auf die eine Person direkt zurückgreifen muss, um Gesundheitsleistungen in Anspruch nehmen zu können. Predisposing Characteristics beziehen sich dagegen auf sämtliche Lebensbereiche; ihre Wirkung auf den Gesundheitszustand und das Gesundheitsverhalten ist indirekt.

Community. Unter den Community Enabling Resources versteht Andersen das Vorhandensein und die Erreichbarkeit von Versorgungseinrichtungen am Wohn- und Arbeitsort. Insofern unterscheiden sich diese Merkmale von den Eigenschaften des Gesundheitssystems, die der Modellkomponente Environment zugeordnet sind und hier nicht betrachtet werden.

Need

Es wird zwischen dem von der betroffenen Person wahrgenommenen Bedarf einerseits und einem durch professionelles Urteil objektivierten Bedarf andererseits unterschieden. Der biologische Bedarf ist nur wirksam, wenn er als solcher vom Arzt oder vom Patienten wahrgenommen wird.

Perceived. Der Perceived Need ist mit der patienteninduzierten Inanspruchnahme assoziiert und wird als ein hauptsächlich soziales Phänomen charakterisiert. Als solches sollte er sich – so eine These

von Andersen – bei angemessener Modellierung zu einem grossen Anteil aus der Sozialstruktur und den Gesundheitsüberzeugungen erklären lassen.

Evaluated. Der Evaluated Need stellt demgegenüber eher ein Mass für die biologische Notwendigkeit dar, ist jedoch nicht mit dieser gleichzusetzen, da auch das professionelle Urteil in einem sozialen Kontext gefällt wird. Der Evaluated Need bestimmt laut Andersen den Umfang einer Behandlung, die eine Person verschrieben bekommt, ist also anders als der Perceived Need stärker mit der arztinduzierten Inanspruchnahme verbunden.

3 Probleme der empirischen Forschung

In diesem Abschnitt werden die wichtigsten kritischen Anmerkungen, welche in der Literatur zur empirischen Inanspruchnahmeforschung im Gesundheitsbereich genannt werden, vorgestellt.

3.1 Theoretische und begriffliche Probleme

Besonders bei Fragestellungen, welche die Nachfrage betreffen, muss man sich über die zentralen Begriffe *Bedarf*, *Nachfrage* und *Inanspruchnahme* klar sein. Die Abgrenzung von *Nachfrage* und *Inanspruchnahme* ist nicht immer ganz deutlich und wird häufig als synonym verwendet. Gemäss Andersen (1997) besteht die Konvention, dass mit diesen Begriffen zwei zeitlich aufeinander folgende Phasen bezeichnet werden. Während Nachfrage sich eher auf den Entscheidungsprozess bezieht, welcher der Inanspruchnahme vorausgeht, setzt die Inanspruchnahme immer den bereits erfolgten Arztbesuch, den Krankenhausaufenthalt oder die Medikamenteneinnahme voraus. Hinter den Begriffen stehen jedoch auch unterschiedliche Konzepte. Beim *Nachfragekonzept* geht man immer von «souveränen» Konsument/innen aus, welche die Wünsche selber artikulieren können; d.h. man stützt die Nachfrage allein auf die PatientInnenentscheidung ab und modelliert diese mit mikroökonomischen Verhaltensmodellen. Mit dem Begriff der *Inanspruchnahme* ist jedoch immer der Prozess insgesamt gemeint (realisierte Nachfrage); d.h. es gehen sowohl die Entscheidungen der Patientin bzw. des Patienten wie der Ärztin bzw. des Arztes ein. Der Prozess selber wird sowohl von individuellen Merkmalen der AkteurInnen als auch von strukturellen Gegebenheiten beeinflusst.

Gemäss Andersen (1997) hat die sprachliche Differenzierung denn auch Folgen für die theoretische Ausrichtung, die Spezifizierung der Modelle und die Analyse der empirischen Erhebungen. Nachfrage-Analysen orientieren sich eher an mikroökonomischen Verhaltensmodellen, Inanspruchnahmeanalysen eher am systemischen Charakter des Gesundheitswesens. Neben den ökonomischen Verhaltensmodellen werden immer auch soziologische, politische oder psychologische Erklärungsansätze herangezogen. Werden wesentliche Annahmen des mikroökonomischen Verhaltensmodells aufgegeben bzw. in Frage gestellt – mangelnde Konsumentensouveränität und eine stark asymmetrische Informationsverteilung zu Lasten der Nachfrager/in – lassen sich beide Ansätze zumindest partiell verknüpfen. Auch mikroökonomische Ansätze müssen dann auf die Analyse von realisierter Nachfrage, also die Inanspruchnahme abstellen.

Die meisten empirischen Untersuchungen stützen sich auf Daten der Inanspruchnahme (utilization). Deshalb müsste man konsequenterweise bei Studien zur «Nachfrage nach medizinischen Leistungen» von «Inanspruchnahme von medizinischen Leistungen» und bspw. bei solchen zur «angebotsinduzierten Nachfrage» von «angebotsinduzierter Inanspruchnahme» sprechen. Problematisch wird dies, wenn Forschungsinteressen oder Verwendungszusammenhänge eine terminologische Differenzierung verlangen. Will man bspw. die Wirkung von verschiedenen Formen der Absicherung des Krankheitsrisikos auf die Nachfrage/Inanspruchnahme medizinischer Leistungen untersuchen, wird man versuchen müssen, die Inanspruchnahmeentscheidung in Nachfrage und realisierte Nachfrage zu disaggregieren. Für die Evaluation der Konsequenzen bestimmter Managed-Care-Arrangements ist der gesamte Inanspruchnahmeprozess zu berücksichtigen und eine Disaggregation wäre eher irreführend (Andersen 1997:S.7). Labelle et al. (1994: S. 351) kritisieren in ihrem viel zitierten Aufsatz zur angebotsinduzierten Nachfrage, dass die Ergebnisse der empirischen Studien mit unterschiedlichen Theorien erklärt werden können, dass Theoriesequenzen unvollständig sind oder dass bestimmte mikroökonomische Verhaltensmodelle kaum empirisch überprüfbar seien. Diese begrifflich-theoretische Kontroverse wird allgemein als Modellkontingenz bezeichnet («lack of rigorous model»)

Die Unterscheidung zwischen *Bedarf* und *Nachfrage/Inanspruchnahme* ist wesentlich und zeigt u.a. auch eine normative Grundeinstellung. Analysiert man den *Bedarf* an medizinischen Gütern und Dienstleistungen, so geht man implizit davon aus, dass die/der Nachfragende selbst nicht in der Lage ist, seine Bedürfnisse zu artikulieren und dass ein funktionierender Markt nicht existiert. Deshalb muss eine übergeordnete Planungsinstanz einschreiten und den Bedarf ermitteln. Schwierig an diesem Konzept ist, dass es kein objektivierbares Mass für die Ermittlung des Bedarfs gibt. Oft wird deshalb der erreichte Status quo als nicht zu unterschreitendes Minimum definiert, womit eine Tendenz zur Ausweitung besteht¹². Bedarfsanalysen werden deshalb meistens mit dem Zweck der Steuerung des Gesundheitswesens durchgeführt (bspw. durch Regulierung des medizinischen Angebots).

3.2 Methodische Probleme

Die methodischen Probleme können in verschiedene Kategorien eingeteilt werden. Eine erste, grundsätzliche Schwierigkeit ergibt sich in der *Modellierung* der empirischen Analyse von Nachfrage/ Inanspruchnahme, die sich auch zum Ziel setzt, den Einfluss der Angebotsentwicklung einzubeziehen. Die Frage der Modellierung ist, neben dem methodischen Aspekt, auch eng mit den begrifflichen und theoretischen Problemen verknüpft. In einem zweiten Schritt geht es darum, dasjenige *ökonometrische Schätzverfahren* auszuwählen, welches für das gewählte theoretische Modell und die verfügbaren Daten am besten geeignet ist. Es gilt zu beachten, dass unterschiedliche Schätzmethoden u.U. zu unterschiedlichen Resultaten führen können. Oft sind die Ergebnisse der empirischen Analysen im Bereich der Inanspruchnahmeforschung davon abhängig, welche Determinanten in welcher Konfiguration in die jeweiligen Schätzmodelle eingehen. Ist dies der Fall, so spricht man von *variableninduzierter Kontingenz* («specification error»). Auch *dateninduzierte Kontingenz* ist bei den Analysen zur Inanspruchnahme von gesundheitlichen Leistungen ein Problem. Damit ist gemeint, dass die Ergebnisse aufgrund von unterschiedlicher Datenstruktur variieren («measurement error»). Dazu gehört u.a. das Aggregationsproblem mit dem bekannten Problem des ökologischen Fehlschlusses¹³. Neben Untersuchungen von Aggregatdaten wird versucht, die Existenz der Angebotsinduzierung auf Individualebene nachzuweisen. Eine Verknüpfung der beiden Ansätze ist mit der Mehrebenenanalyse möglich. Dabei werden Daten unterschiedlicher Aggregierungsniveaus gemeinsam analysiert, also Daten von Individuen und der Kontexte, in denen diese leben. Damit sind jedoch sehr hohe Anforderungen an die Daten verbunden.

3.2.1 Identifikation eines kausalen Effektes

Ein zentrales Problem in der gesundheitsökonomischen Forschung ist die Identifikation von kausalen Effekten einer Politik auf eine Zielvariable. Im Unterschied zur blossen Feststellung einer einfachen statistischen Korrelation zwischen der Zielvariable und einer Politikvariable erfordert die Identifikation eines kausalen Effektes, dass mit Hilfe des ökonometrischen Modells für alle anderen möglichen Effekte (bspw. Selektionseffekte, Moral Hazard¹⁴) «korrigiert» werden kann.

¹² Vgl. Andersen (1992: S.19)

¹³ Mit dem Begriff «Ökologischer Fehlschluss» bezeichnet man das fehlerhafte Schliessen von statistischen Beziehungen zwischen aggregierten Daten auf Beziehungen zwischen den Komponenten der Aggregate (bspw. Individuen).

¹⁴ Der Begriff «Moral Hazard» stammt aus der Versicherungswissenschaft. Mit ihm wird ausgedrückt, dass sich die Einstellung von Versicherten und meist auch ihr Verhalten gegenüber dem Schadensfall durch Abschluss einer Versicherung tendenziell zugunsten eines höheren Risikos verschiebt ("moralische Versuchung").

Im Bereich der gesundheitsökonomischen Forschung tritt dieses Problem bspw. bei der Evaluation von Versicherungsmodellen mit verschiedenen Selbstbeteiligungsformen, aber auch bei der Frage der angebotsinduzierten Inanspruchnahme/Nachfrage auf. In Zusammenhang mit wählbaren Selbstbehalt kann aufgrund eines vorliegenden statistisch signifikanten Zusammenhangs zwischen den individuellen Gesundheitsausgaben und der Höhe einer individuell gewählten Franchise noch nicht auf einen kausalen Zusammenhang geschlossen werden. Mit der Wahlmöglichkeit von Franchisen soll ein finanzieller Anreiz gesetzt werden, Leistungen von Seiten der Versicherungsnehmer/innen sparsam in Anspruch zu nehmen. Der Anreiz sollte umso höher sein, je höher die Selbstbeteiligung ist. Der Zusammenhang kann jedoch auch den Sachverhalt widerspiegeln, dass Konsumentinnen und Konsumenten mit hohem individuell erwartetem Gesundheitsrisiko versuchen, einen geringen Selbstbehalt zu vereinbaren, währenddem Konsumentinnen und Konsumenten mit tiefem individuell erwartetem Gesundheitsrisiko einen hohen Selbstbehalt wählen, um durch eine Reduktion der Versicherungsprämie Kosten sparen zu können. Damit eine kausale Interpretation eines Effektes möglich wird, muss für solche und ähnliche Selektionsverzerrungen, welche u.a. durch so genannte «unbeobachtete Heterogenität» verursacht werden, korrigiert werden. Dies geschieht, wenn überhaupt, entweder mit Hilfe von ausgeklügelten Forschungsdesigns oder mit statistischen Verfahren, welche zur Kontrolle von Selektionsproblemen¹⁵ geeignet sind. Solche Modelle sind aus statistisch-mathematischer Sicht jedoch anspruchsvoll (vgl. dazu Abschnitt «Unterschiedliche Schätzverfahren» weiter unten). Bei der Beurteilung von Anzeizeffekten auf PatientInnenseite (bspw. Wahlfranchisen) muss zwischen Selektions- und Anzeizeffekt unterschieden werden. Beim Nachweis von arztinduzierter Inanspruchnahme geht es darum, die Effekte von patienteninduzierter und arztinduzierter Inanspruchnahme zu trennen. Eine Möglichkeit, wie diese methodisch versucht wird, wird im folgenden Unterabschnitt diskutiert.

3.2.2 Trennung der Effekte von patienteninduzierter Nachfrage und arztinduzierter Inanspruchnahme

Für den empirischen Nachweis von Angebotsinduzierung ist es von zentraler Bedeutung, einen Indikator zu finden, der es erlaubt, zwischen eher patienteninduzierter Nachfrage und eher arztinduzierter Inanspruchnahme zu unterscheiden. Eine Möglichkeit, dieses Problem zu lösen, besteht darin, zwischen Erstkontakt und Kontakthäufigkeit während der Behandlungsepisode zu unterscheiden. Dahinter steht die Annahme, dass ein Erstkontakt stärker von Patientinnen und Patienten beeinflusst wird und die Häufigkeit der auf die Behandlungsepisode bezogenen Kontakte in stärkerem Masse von den Ärztinnen und Ärzten bestimmt wird. Das Problem besteht jedoch darin, diesen überzeugenden Ansatz empirisch exakt umzusetzen. Dazu werden Daten benötigt, die es erlauben, den Beginn und den Abschluss einer krankheitsbedingten Behandlungsepisode exakt zu bestimmen. Die Daten von grossen Bevölkerungsbefragungen (bspw. Gesundheitsbefragung) enthalten diese Informationen jedoch nicht.

3.2.3 Einfluss der Modellspezifikation auf die Resultate

Insbesondere in der Inanspruchnahmeforschung sind die Ergebnisse von empirischen Analysen oft davon abhängig, welche erklärenden Merkmale für ein Schätzmodell vorhanden sind und welche dann tatsächlich in ein Modell eingehen. Je nach Modellspezifikation können so die Effekte der einbezogenen erklärenden Merkmale unterschiedlich ausfallen. Es ist auch möglich, dass bestimmte, relevante

¹⁵ Das Selektionsproblem tritt bei mikroökonomischen Analysen häufig in Zusammenhang mit «unbeobachteter Heterogenität» auf.

Variablen nicht berücksichtigt werden, oder dass die Endogenität unabhängiger Merkmale nicht berücksichtigt wird, was zu «falschen» Interpretationen der Resultate führen kann. Dass dieses Problem speziell bei der Inanspruchnahmeforschung von gesundheitlichen Leistungen auftritt, ist der Kompliziertheit und der Komplexität des Inanspruchnahmegeschehens zuzuschreiben. Die Vielzahl von möglichen Einflussfaktoren und die Vielzahl von Interdependenz- und Dependenzbeziehungen zeugen davon. So kann bspw. fast jeder Faktor eine bestimmte Wirkung erzeugen, wobei sich die einzelnen Wirkungen auch addieren können (Multikausalität). Nicht alle erklärenden Merkmale wirken linear. Oft gibt es abrupte Trendwenden, welche es zu berücksichtigen gilt (Nichtlinearität). Zudem können Wechselwirkungen zwischen verschiedenen Teilprozessen bestehen, die man dann auch modellieren muss, will man verlässliche Resultate «produzieren».

3.2.4 Messfehler und andere datenspezifische Probleme

Ungeeignete Aggregationsniveaus und nicht repräsentative Daten werden von Labelle et al. (1994) als Beispiele von Messfehlern genannt. Besonders hingewiesen muss auf das spezifische Schätzproblem, das die Verwendung unterschiedlicher Aggregationsniveaus in der Schätzung betrifft. Auf der einen Seite liegen den Schätzgleichungen zur Inanspruchnahme Individualdaten zugrunde. Auf der anderen Seite liegen die für die Überprüfung der Angebotsinduzierung verwendeten Variablen (bspw. Arztdichte, Facharztstruktur) auf einer höheren Aggregationsebene vor. Auf inhaltlicher Ebene gilt es, folgendes Problem zu lösen: Wie können individuenbezogene Daten mit den notwendig (innerhalb klarer Grenzen) aggregierten Daten der Angebotsstruktur (hier also der niedergelassenen Ärzten/Ärztinnen) so verknüpft werden, dass jede betrachtete Teilregion tatsächlich einen geographisch abgeschlossenen Markt für medizinische Leistungen bildet. Ein zu hohes Aggregationsniveau nivelliert die Unterschiede, die zwischen den einzelnen Märkten bestehen können und eine zu tiefe Disaggregation führt zu Verzerrungen durch einen zu hohen Anteil von «cross-bordering», d.h. von Patient/innen, die Ärzte und Ärztinnen kontaktieren, welche ausserhalb des als Markt identifizierten geographischen Raums ihre Praxis betreiben. In der Schweiz ist bspw. die Ärztlindichte bisher nur auf kantonaler Ebene verfügbar, was unter dem Aspekt von einem geographisch abgeschlossenen Markt für medizinische Leistungen zumindest fragwürdig ist. Aus statistischer Sicht gilt es zu beachten, dass u.U. eine zu hohe statistische Genauigkeit der auf der Aggregationsebene liegenden Einflussfaktoren angegeben wird¹⁶.

Andersen (1997) nennt noch weitere datenspezifische Probleme, welche zu signifikanten Unterschieden in den Resultaten führen können. Dazu gehören marginale Variationen im Forschungsdesign, Unterschiede in der Vorgabe von Antwortkategorien, kleine Abweichungen in der Wortwahl, unterschiedliche Ausschöpfungsquoten und jahreszeitlich variierende Morbiditätsquoten. All diese Probleme lassen sich als dateninduzierte Kontingenz bezeichnen. D.h. die Ergebnisse sind nicht nur davon abhängig, welche Konfigurationen an Determinanten geschätzt werden, sondern auch davon, welche Struktur Daten haben. Andersen schätzt, dass diese Probleme fast noch diffiziler zu handhaben sind als bspw. das Fehlen relevanter Variablen.

3.2.5 Unterschiedliche Schätzverfahren

In den in der Literatur beschriebenen empirischen Analysen zum Inanspruchnahmeprozess wird eine Vielzahl von unterschiedlichen Schätzverfahren angewandt. Gemeinsam ist allen weiterführenden

¹⁶ Mit dem Aggregationsproblem im Fall der linearen Regression beschäftigt sich ausführlich Moulton (1990).

Analysen, dass multivariate Methoden verwendet werden. Je nachdem, welches Schätzverfahren zur Anwendung kommt, können die Resultate variieren. Es ist bei der Beurteilung von Resultaten immer wichtig zu wissen, mit welchen Schätzverfahren die Resultate erzielt wurden, welche statistischen Annahmen hinter den Modellen stehen (bspw. Verteilungsannahmen) und warum die Autoren gerade dieses Verfahren gewählt haben. Die Auswahl des (geeigneten) Schätzverfahrens ist u.a. abhängig von folgenden Aspekten und Fragen:

Handelt es sich um Querschnitts- oder Längsschnittdaten?

Es ist in der Literatur unbestritten, dass für die Identifikation eines kausalen Effektes Längsschnittdaten besser geeignet sind als Querschnittdaten.

Sind die Daten auf der Individualebene oder einer aggregierten angesiedelt?

Bei aggregierte Daten besteht immer das Problem des ökologischen Fehlschlusses. Zudem müssen Aggregatsanalysen auf verhaltenswissenschaftliche Begründungen verzichten, was bei Individualdaten und Mehrebenenanalysen nicht der Fall ist.

Welche Einheit besitzt das zu erklärende Merkmal?

Gesundheitliche Leistungen können in Mengen- oder in Geldeinheiten gemessen werden. Zudem kann unterschieden werden zwischen einem (positiven) Bezug und keinem Bezug von Leistungen. Je nach Einheit und Verteilungsmerkmalen müssen unterschiedliche Schätzverfahren gewählt werden. So bieten sich bspw. für eine Analyse der Anzahl Arztbesuche sog. Zähldatenmodelle an. Zähldatenmodelle sind für den Fall geeignet, dass die zu erklärende Variable kleine ganzzahlige, positive Werte annimmt (oder null) und als (An)Zahl zu interpretieren ist. Für ungleichmässig verteilte und zudem keine negativen Werte annehmenden Daten eignet sich die Zähldatenanalyse («count data models») besser als bspw. die OLS-Schätzung, welche von einer Normalverteilung der abhängigen Variable ausgeht.

Handelt sich beim Inanspruchnahmeprozess um einen simultanen oder einen sequentiellen Entscheid?

Unterteilt man analytisch den Entscheid darüber, ob man Gesundheitsleistungen in Anspruch nimmt und wie viel man in Anspruch nimmt, so stellt sich die Frage, ob es sich um einen simultanen oder einen sequentiellen Entscheid handelt. Entscheidet man sich für einen sequentiellen Entscheid, empfiehlt Jones (2000, S. 285) ein zweiteiliges Modell (2-part Modell, 2PM). Bei einem simultanen Entscheid sei eher ein Tobit-Modell denkbar.

Beim 2PM wird von einem zweiteiligen Entscheid über die Inanspruchnahme von Gesundheitsleistungen ausgegangen. In einer ersten Stufe entscheiden sich die Individuen darüber, ob sie Leistungen in Anspruch nehmen wollen oder nicht. In einem zweiten Schritt erfolgt die Wahl des Ausmasses der Inanspruchnahme. Weil die Kosten sehr schief verteilt sind, werden die Kosten im zweiten Modellteil logarithmiert.

Das Tobit-Modell beruht auf der Idee einer zensurierten Zufallsvariable. In einem bestimmten Bereich werden alle Werte auf einen einzigen Wert reduziert (Greene 1993: S. 691). Bei den Gesundheitskosten beobachtet man gehäuft den Wert von 0 Franken Kosten für die Versicherer. Die Kosten der Kassen entsprechen aber nicht zwingend den real entstandenen Kosten: Dies entsteht dadurch, dass die Versicherten ihre Rechnungen, welche in der Summe die Franchisegrenze nicht erreichen, den Kassen gar nicht einreichen. Sie wissen, dass sie keine Rückvergütung bekommen können. Die Ver-

teilung wird somit links zensuriert. Wenn das 2PM einen sequentiellen Entscheid abbildete, dann kann man beim Tobit-Modell davon sprechen, dass ein simultaner Entscheid gefällt wird: Man entscheidet sich gleichzeitig darüber, ob man Leistungen in Anspruch nehmen will und wie viel Leistungen es sein sollen.

Sind Selektionseffekte ein Problem und wie wird dafür kontrolliert bzw. korrigiert?

Selektionseffekte treten im Bereich von gesundheits-ökonomischen Studien in Bezug auf die Wirkung von Anreizsystemen auf die Inanspruchnahme auf. Ob Selektionseffekte überhaupt eine Rolle spielen, ist u.a. abhängig davon, ob man aufgrund der Hauptfragestellung kausale Beziehungen zwischen einzelnen Variablen nachweisen will (vgl. Unterabschnitt 3.2.1 «Identifikation eines kausalen Effektes»).

Aus statistischer Sicht finden sich in der Literatur zum Problem der Selbstselektion verschiedene Ansätze, wie für Selektionseffekte kontrolliert werden kann. Auf der einen Seite können mit Hilfe von experimentellen Studiendesigns Selektionseffekte vermieden werden (vgl. bspw. das RAND-Experiment von Newhouse et al. 1993, welches durch Randomisierung Kontrollgruppen bildet). Auf der anderen Seite gibt es rein statistisch-mathematische Methoden, um für mögliche Selektionsverzerrungen korrigieren zu können. Dabei handelt es sich um komplizierte und mathematisch anspruchsvolle Modelle. Konkrete Anwendungsbeispiele dafür finden sich bei Werblow (2002) oder Schellhorn (2002).

3.3 Vergleichbarkeit der Resultate

McKinlay (1972) wirft im Rahmen einer Systematisierung der Forschungsansätze das Problem auf, dass die empirischen Ergebnisse sowohl bezüglich der Stärke der Einflussfaktoren als auch bezüglich ihrer Richtung inkonsistent sind. Als Faktoren, die zu dieser Situation beitragen, nennt er Unterschiede im methodischen Zugang, in den Rahmenbedingungen der untersuchten Inanspruchnahme und in den gewählten Zeitintervallen (vgl. McKinlay, 1972: S. 115). Das Problem der Vergleichbarkeit der Resultate ist denn auch eine Folge der in den vorherigen Abschnitten diskutierten Probleme.

Das Problem der Vergleichbarkeit der empirischen Ergebnisse betrifft nicht nur den Bereich der sozialwissenschaftlichen Modelle, sondern genauso ökonomische Modelle. Letztlich ist das Problem für einzelne erklärende Merkmale nur im Kontext des gesamten Modells und unter Berücksichtigung der jeweiligen Rahmenbedingungen zu beantworten. Um eine gewisse Beliebigkeit einzugrenzen und eine Interpretationsbasis für die empirischen Ergebnisse zu haben, ist es deshalb umso wichtiger, auf einen umfassenden Modellrahmen zurückgreifen zu können. Nur mit einem solchen Rahmen können die empirischen Ergebnisse vergleichbar(er) gemacht werden. Das Verhaltensmodell nach Andersen bietet dazu eine Möglichkeit und ermöglicht die Integration der zum Teil präziseren Hypothesen der verschiedenen Modelle. Durch die Herstellung von Bezügen zwischen den Modellen kann das Modell nach Andersen damit auch in einen grösseren wissenschaftlichen Kontext eingebettet werden (Andersen 2003: S.7).

3.4 Bestimmung des Ausmasses angebotsinduzierter Nachfrage

Die zwei Unterabschnitte 3.1 und 3.2 über die theoretischen und methodischen Probleme haben aufgezeigt, dass der empirische Nachweis angebotsinduzierter Nachfrage sehr komplex und schwierig ist. Noch schwieriger ist es, deren Ausmass zu bestimmen. Labelle et al. (1994) haben in ihrem viel beachteten Aufsatz zu Recht darauf hingewiesen, dass Angebotsinduzierung auch unter dem Aspekt der Ergebnisorientierung betrachtet werden muss (vgl. Unterabschnitt 2.1.2). Sie verweisen darauf,

dass man grob zwischen «erwünschter» und «unerwünschter» angebotsinduzierter Nachfrage unterscheiden kann (vgl. Tabelle 1, S. 26). Wir sind bei der Literaturrecherche auf keine empirischen Untersuchungen gestossen, welche versuchen, die Existenz von Angebotsinduzierung nachzuweisen und deren Modelle zugleich zwischen «erwünschter» und «unerwünschter» Angebotsinduzierung unterschieden können. Will man das Ausmass «unerwünschter» Angebotsinduzierung bestimmen, ist dafür deshalb zwingend eine Referenzgrösse erforderlich, welche jedoch nicht einfach vorhanden ist. Mit Referenzgrösse ist gemeint, dass es irgendein Mass braucht, das darüber Auskunft gibt, welche und wie viele Behandlungen/Leistungen in jedem einzelnen Krankheitsfall «richtig» bzw. angebracht sind. Erst wenn man über ein solches Mass verfügt, können daraus Rückschlüsse über das Ausmass von Angebotsinduzierung gezogen werden. Ein solches Mass ist eng an den Begriff des «Bedarfs» gekoppelt, denn es ist der Bedarf, der auf die Frage antwortet, wie viel gesundheitliche Leistung richtig ist und tatsächlich benötigt wird. Deshalb muss für die Bestimmung eines Referenzmasses zwingend der «richtige» Bedarf ermittelt werden (vgl. dazu auch den Unterabschnitt über die «Ausprägungen angebotsinduzierter Nachfrage», S. 26 f.). Dabei ist es wichtig, dass man sich darüber im Klaren ist, dass hinter dem Begriff des «Bedarfs» ein normatives Konzept verborgen ist. Nach welchen Kriterien soll sich ein «richtiger» Bedarf richten? Nach dem Minimum, nach dem Optimum oder nach dem Maximum? Unabhängig davon, wie der «richtige» Bedarf bestimmt wird, muss über die Frage diskutiert werden, nach welchen Kriterien sich der Bedarf und mit ihm die zu bestimmende Referenzgrösse richten soll. Grundsätzlich gibt es dazu mindestens zwei Möglichkeiten:

Die erste Möglichkeit besteht darin, *medizinische Standards* hinsichtlich der Art und des Umfangs von Diagnose- und Therapieleistungen, die bei bestimmten Symptomen zu erfolgen haben, zu bilden. Aufgrund der Tatsache, dass die Medizin nur schwer in der Lage ist, ein eindeutiges Therapieregime aus der Diagnose abzuleiten, ist dies jedoch sehr schwierig. Es stellt sich die Frage, wer die Behandlungsstandards festlegen soll und nach welchen Kriterien dies erfolgen soll. Wird kein Standard festgelegt oder können solche nicht festgelegt werden, so ist es fast unmöglich, das Ausmass von angebotsinduzierter Nachfrage zu bestimmen, da kein objektives Mass an Behandlungsquantität und -qualität als Referenzgrösse vorliegt. In diesem Zusammenhang ist die Evidenzbasierte Medizin («evidence based medicine») zu erwähnen¹⁷. Dabei handelt es sich um eine Methode, welche versucht, solche medizinischen Standards festzulegen. Sie stösst jedoch in Zusammenhang mit der ärztlichen Therapiefreiheit auch auf heftige Kritik.

Eine zweite Möglichkeit besteht darin, dass man die realisierte Nachfrage als Basis für die Bestimmung eines Referenzwertes nimmt. Mit Hilfe der realisierten Nachfrage ist es möglich, auf Pro-Kopf-Durchschnittswerten beruhende, gruppenspezifische Inanspruchnahmeprofile (bzw. Bedarfsprofile) zu erstellen. Dabei können verschiedene Faktoren wie bspw. das Alter und das Geschlecht, welche die Inanspruchnahme von Gesundheitsleistungen beeinflussen, berücksichtigt werden. Aus den Durchschnittswerten können sodann Referenzgrössen berechnet werden. Ist die realisierte Nachfrage in einem bestimmten Gebiet höher als die so berechnete Referenzgrösse, kann daraus abgeleitet werd-

¹⁷ Nach Hagenbuch (2001) ist Evidenzbasierte Medizin (EBM) eine Methode zur besseren und rationaleren Entscheidungsfindung für die Patientenbehandlung und zur Beurteilung der Effektivität von diagnostisch-therapeutischen Massnahmen. Die Informationen zur EBM stammen aus klinisch-epidemiologischen Studien. Der Ansatz versteht sich als eine Überwindung der meinungsorientierten Medizin, die lediglich auf dem individuellen Wissen und den Erfahrungen einzelner Fachleute beruht. EBM ist der gewissenhafte und vernünftige Gebrauch der besten wissenschaftlichen Evidenz für Entscheidungen und Behandlungen in der medizinischen Versorgung individueller Patienten. Die Praxis der EBM bedeutet die Integration individueller klinischer Expertise mit der bestmöglichen verfügbaren externen Evidenz aus Publikationen zu systematischen Studien. Die Fähigkeit, die Evidenz in der Fachliteratur zu identifizieren, zu bewerten und auf den Einzelfall anzuwenden, ist Kern der evidenzbasierten Medizin.

en, wie hoch das maximale Ausmass an Angebotsinduzierung sein kann. Die Angebotsinduzierung kann nicht höher als die Differenz zwischen dem Referenzwert und der tatsächlich realisierten Nachfrage sein. Ob jedoch die gesamte Differenz auf Angebotsinduzierung zurückzuführen ist, kann so nicht schlüssig beantwortet werden. Das Problem bei dieser Methode besteht jedoch auch darin, dass für die Ermittlung der Referenzgrösse die realisierte Nachfrage und nicht die um die Arztinduzierung bereinigte Nachfrage in die Berechnungen einfließt. Es ist deshalb möglich, dass im Referenzwert selber noch Effekte von Angebotsinduzierung enthalten sind.

3.5 Zusammenfassung

Die erwähnten begrifflichen, theoretischen und empirischen Probleme der Inanspruchnahmeforschung und der Forschung zur angebotsinduzierten Nachfrage zeigen auf, dass eine empirische Analyse immer nur auf einen bestimmten Ausschnitt aus dem gesamten Inanspruchnahmeprozess fokussiert sein kann. Diese Ausschnitte können ein unterschiedliches Niveau an Kompliziertheit und Komplexität aufweisen. Sie können sich auf eine oder wenige Variablen beschränken (bspw. das Geschlecht oder das Alter) oder sie können sich auf bestimmte, abgrenzbare Kontexte beziehen (z.B. Schichtabhängigkeit). Die Resultate von empirischen Analysen im Rahmen der Inanspruchnahme von gesundheitlichen Leistungen hängen in starkem Masse ab von den theoretischen Modellen, den zur Verfügung stehenden Daten, der Wahl der ökonomischen Modelle und deren Spezifikation. Dies ist eine Folge davon, dass das Inanspruchnahmegeschehen sehr kompliziert und komplex ist. Dies zeigt sich an den sehr vielen Interdependenz- und Dependenzbeziehungen zwischen einer Vielzahl von möglichen Einflussfaktoren. Diese Beziehungen in ihrer Gesamtheit zu modellieren ist kaum möglich, weshalb man sich auf bestimmte Ausschnitte beschränken muss. Darunter leidet die Vergleichbarkeit der Resultate. Die hohe Komplexität hat zudem zur Folge, dass ein empirischer Nachweis von kausalen Beziehungen zwischen der Inanspruchnahme und ausgewählten Variablen äusserst schwierig und deshalb auch aufwändig ist. Oft fehlen dazu auch präzise Daten. Will man Resultate der empirischen Analysen beurteilen und miteinander vergleichen, muss dies immer unter den erwähnten Gesichtspunkten geschehen.

Mit dem Modell von Andersen steht ein Instrumentarium zur Verfügung, mit dem man in der Lage ist, die Kumulation des theoretischen und empirischen Wissens über das Inanspruchnahmegeschehen sichtbar zu machen und den Prozess zunehmender Erklärungs- und Prognoseleistungen zu systematisieren. Es enthält ein breites, potentiell vollständiges Spektrum von Kategorien zur Subsumierung individueller und gesellschaftlicher Determinanten, welche die Inanspruchnahme beeinflussen können. Das analytische Raster dieses Modells kann deshalb dazu verwendet werden, widersprüchliche empirische Ergebnisse einzuordnen, zu diskutieren und allenfalls zu erklären. Zudem zeigt es auf, in welchen Bereichen man bspw. aufgrund der Datenlage über quantifizierbare Merkmale verfügt und in welchen Bereichen man nicht in der Lage ist, messbare Variablen zu berücksichtigen.

4 Stand der Forschung

Im Zentrum des vorliegenden Kapitels steht der Stand der Forschung in der Frage der Angebotsinduzierung. Im vorherigen Kapitel konnte gezeigt werden, dass die Angebotsinduzierung jedoch nicht losgelöst vom gesamten Prozess des Inanspruchnahmegeschehens betrachtet werden kann. Will man Angebotsinduzierung empirisch nachweisen, muss eine Möglichkeit gefunden werden, wie dieser Effekt von anderen Effekten, welche die Inanspruchnahme beeinflussen, getrennt werden kann. Eine Steigerung der Inanspruchnahme kann u.a. auch die Folge von Nachfrageinduzierung, einem Rückgang von indirekten Kosten, Qualitätssteigerung oder einem Abbau eines bestehenden Nachfrageüberhangs sein. Damit dies berücksichtigt werden kann, greifen viele Studien auf komplexe Erklärungsmuster zurück, mit denen nicht nur Angebotsinduzierung, sondern auch andere Determinanten der Inanspruchnahme identifiziert werden können. In der Übersicht zum Stand der empirischen Forschung wird deshalb der Blick auf weitere Determinanten der Inanspruchnahme erweitert. Dazu gehören neben individuellen Merkmalen der Patient/innen auch Merkmale des gesellschaftlichen Umfelds und die Ausgestaltung der Krankenversicherung, welche Auswirkungen auf das Verhalten der Patient/innen hat. In einem ersten Teil wird deshalb der Fokus auf die Angebotsinduzierung gelegt, in einem zweiten auf die Wirkung von Steuerungsmechanismen zum Nachfrageverhalten und in einem dritten auf alle weiteren empirisch untersuchten individuellen und gesellschaftlichen Determinanten der Inanspruchnahme. Im Abschnitt 4.1 wird zuerst in einer groben Übersicht der Stand der Forschung ausserhalb der Schweiz dargestellt, um dann detaillierter auf den Stand der Forschung in der Schweiz einzugehen. Ein Vergleich zwischen den Resultaten verschiedener Länder ist insofern problematisch, da das Inanspruchnahmegeschehen stark von den rechtlichen und institutionellen Rahmenbedingungen der Systeme der Gesundheitsversorgung beeinflusst wird.

4.1 Internationale Untersuchungen

Für die Beurteilung der Resultate der internationalen empirischen Forschung im Bereich der Inanspruchnahme nach Gesundheitsleistungen ist es wichtig, in Betracht zu ziehen, dass zwischen den Gesundheitssystemen der verschiedenen Länder Unterschiede bestehen. Unterschiedliche Anreizsysteme, unterschiedliche Vergütungsformen etc. haben Auswirkungen auf das Verhalten der Akteurinnen und Akteure. Deshalb können die Resultate von länderspezifischen Studien nur in Zusammenhang mit den institutionellen Voraussetzungen des jeweiligen Landes betrachtet und kaum verallgemeinert werden.

4.1.1 Determinanten der Inanspruchnahme

Dass die Inanspruchnahmeforschung schon seit einiger Zeit stetig am wachsen ist, hat u.a. damit zu tun, dass die Gesundheitskosten in fast allen Gesellschaften laufend zunehmen und ein Ausmass annehmen, das zu gesellschaftspolitischen Diskussionen führt.

Die Inanspruchnahme von gesundheitlichen Leistungen wird in den zahlreichen Studien je nach theoretischem Ansatz mit persönlichen Faktoren (Krankheitszustand, Symptome, Wissen, Überzeugungen, Erfahrungen, Gefühle, Bedürfnisse, Bewältigungsstrategien etc.), sozialen Faktoren (soziodemographische Faktoren, familiäre Faktoren, soziale Unterstützung, Wertesysteme, etc.) und Faktoren, welche mit dem Gesundheitssystem in Verbindung zu bringen sind (geographische Distanz, Erreichbarkeit, Kosten, Abrechnungssysteme, etc.) erklärt. Einige in der Literatur oft zitierten Studienergebnisse und Ergebnisse von Reviews sollen hier kurz erwähnt werden.

Personen, welche Gesundheitsleistungen überhaupt in Anspruch nehmen, schätzen sich als weniger gesund ein. Sie greifen weniger auf Selbstbehandlungen zurück und scheinen stärker als Nicht-Benutzer/innen von Gesundheitsleistungen von persönlichen Problemen und Stress betroffen zu sein (Anderson et al. 1977). Der Bezug von Gesundheitsleistungen in einer Vorperiode ist stärker assoziiert mit der laufenden Inanspruchnahme als der Gesundheitszustand der Vorperiode (Eve 1988). Aufgrund eines Reviews zum Thema «Inanspruchnahme von Gesundheitsleistungen» schliesst McKinlay, dass Faktoren in Bezug zum Familienleben, die Wahrnehmung der Notwendigkeit («need») und Kindheitsgewohnheiten wichtige Bestimmungsfaktoren für das Inanspruchnahmeverhalten sind (McKinlay 1972). Die Inanspruchnahme von Gesundheitsleistungen hängt auch mit dem Alter, dem Geschlecht, der Erziehung, der Religion, der ethnischen Zugehörigkeit, dem sozioökonomischen Status und der sozialen Unterstützung zusammen (McKinlay 1972, Mechanic 1979, Kouzis & Eaton 1998). Die Inanspruchnahmequote ist bei Männern generell tiefer als bei Frauen und nimmt mit dem Alter zu (McKinlay 1972, Mechanic 1976). Der physische Gesundheitszustand erklärt die Inanspruchnahme sowohl älterer als auch jüngerer Patient/innen (Evashwick et al. 1984, Hibbard & Pope 1986). Der mentale Gesundheitsstatus ist ein guter Prädiktor für die Inanspruchnahme von gesundheitlichen Leistungen, welche von Patienten- und Patientinnenseite initiiert werden. Unterschiede in sozialer Wohlfahrt und dem System der sozialen Sicherheit beeinflussen die Behandlungsrate stark (Schrire 1986).

International gibt es sehr viele Studien, welche sich auf das Modell von Andersen (vgl. dazu Abschnitt 2.2, insbesondere Abbildung 2) stützen. Die empirische Erklärungskraft des Andersen-Modells beziffert Herschbach auf max. 25% der Varianz und beruft sich damit auf die Ergebnisse verschiedener Studien. Als wichtigste erklärende Grösse erscheine die subjektive Bewertung der eigenen Symptome, während alle übrigen Faktoren eine eher untergeordnete Rolle spielten (vgl. Herschbach, 1995: S. 86). Ähnliche Werte geben Phillips et al. (1998: S. 581) an. Der Median der erklärten Varianz liegt in den im Review verwendeten Studien bei etwa 20%.

De Boer et al. (1997) tragen die Ergebnisse von 53 Studien aus den USA, Grossbritannien, Schweden, Norwegen, Kanada, Neuseeland und den Niederlanden zusammen, die sich mit den Prädiktoren der Inanspruchnahme von Ärztinnen und Ärzten und Krankenhausaufenthalten chronisch Kranker befassen. Insgesamt zeigen die «Predisposing»-Faktoren wie Alter, Geschlecht und Familienstatus und «Enabling»-Faktoren wie Einkommen, Versicherungsschutz und soziale Unterstützung nur einen geringen Einfluss auf die Inanspruchnahme. Die «Need»-Faktoren – der Schweregrad der Krankheit und der Symptome, die Lebensqualität und der wahrgenommene Gesundheitszustand, Depressionen und psychische Sorgen – beeinflussen die Inanspruchnahme erwartungsgemäss am stärksten. Sie kommen zum Schluss, dass beide, sowohl der Schweregrad der Krankheit als auch der psychische Zustand die wichtigsten Einflussfaktoren auf die Inanspruchnahme von gesundheitlichen Leistungen sind.

Phillips et al. (1998) untersuchen in Studien, die auf dem Andersen-Modell basieren, die Verwendung von umweltbezogenen Faktoren und angebotsseitigen Merkmalen. 45% der untersuchten Studien schliessen umweltbezogenen Faktoren mit ein und 51% angebotsseitige Merkmale. Nur wenige Studien untersuchen spezifische Eigenschaften des Gesundheitssystems. Die meisten verwenden für die Analysen einfache Regressionen mit hierarchischem Einschluss der Variablen. Nur 14% der Studien schliessen in ihre Analysen gleichzeitig umweltbezogene und angebotsseitige Merkmale ein. Sie heben die Notwendigkeit hervor, diese Einflussgrössen zu berücksichtigen und zeigen Möglichkeiten auf, wie eventuelle Hindernisse, die bei der Einbeziehung von umweltbezogenen und angebotsseitigen Merkmalen häufig auftreten, überwunden werden können.

Eine deutsche Untersuchung, die sich explizit auf das Verhaltensmodell nach Andersen bezieht, liegt von Härtel (1985) vor. Sie untersucht die Inanspruchnahme der Teilnehmer an der Münchner Blutdruckstudie. Als signifikante Einflussgrößen auf die Kontaktzahl identifiziert Härtel Alter, Geschlecht, Haushaltsgrösse, Erwerbstätigkeit, berufliche Selbständigkeit, allgemeine körperliche Verfassung, Anzahl chronischer Krankheiten und Übergewicht. Als nicht signifikant erweisen sich neben dem Familienstand die Ausbildungsjahre, die berufliche Position und die Höhe des Blutdrucks. Bei einer nach Geschlecht getrennten Analyse stellen sich – neben den Need-Faktoren körperliche Verfassung und Anzahl der chronischen Krankheiten – für Männer die Erwerbstätigkeit und für Frauen die Haushaltsgrösse als signifikante Einflussgrößen heraus.

4.1.2 Überprüfung des Grossman-Modells

Muurinen (1982) und Wagstaff (1986) entwickelten theoretische Vorhersagen mit Hilfe des Grossman-Modells (Vgl. dazu Unterabschnitt 2.1.1), welche empirisch überprüfbar sind¹⁸. Sie leiteten daraus Hypothesen über die Vorzeichen der Effekte von Gesundheit, Alter, Bildungsniveau und Lebenseinkommen auf die Nachfrage nach Gesundheit bzw. die Nachfrage nach medizinischen Leistungen ab. Die empirischen Resultate verschiedener Studien sind nicht einheitlich. In der Regel werden die Hypothesen jedoch widerlegt.

Einer der Hauptkritikpunkte ist die Beziehung zwischen Gesundheit und Nachfrage medizinischer Leistungen: Aus dem theoretischen Modell lässt sich eine höhere Nachfrage nach medizinischen Leistungen bei steigendem gewünschten Gesundheitszustand ableiten. Da in dem Modell der aktuelle Gesundheitszustand immer dem gewünschten Gesundheitszustand entspricht, lässt sich dies nicht mit der empirisch höheren Inanspruchnahme bei schlechterem Gesundheitszustand vereinbaren. In Untersuchungen von Wagstaff (1986) und Leu und Gerfin (1992) besteht bei der Überprüfung des Grossman-Modells empirisch ein negativer, hochsignifikanter Zusammenhang.

Wagstaff (1993) schätzt die Nachfrage nach Gesundheit mit dem Einkommen, dem Bildungsniveau, dem Alter und dem Geschlecht. Er schätzt je zwei Gleichungen mit zwei unterschiedlichen Altersgruppen (unter und über 41 Jahre). In der Gruppe der unter 41-Jährigen hat das Alter das «falsche» Vorzeichen, ist jedoch nicht signifikant. Bei den über 41-Jährigen bestätigt sich die Vorhersage. Die Vorhersage bezüglich einer steigenden Nachfrage nach medizinischen Leistungen mit steigendem Alter wird im Falle ambulanter ärztlicher Leistungen von Duan et al. (1984), Newhouse und Phelps (1974) und Zweifel (1985) widerlegt (vgl. Zweifel und Breyer (1997: S. 71).

Nocera und Zweifel (1998) schätzten die Nachfrage nach Gesundheit mit Hilfe von zwei Panel-Datensätzen. Mit dem ersten Datensatz konnten die Vorhersagen des Grossman-Modells in den meisten der Fälle bestätigt werden, mit dem zweiten jedoch waren die Resultate weniger konsistent. Sie kommen zu dem Ergebnis, dass nicht nur das Bildungsniveau in positivem Zusammenhang mit der individuellen Gesundheit steht, sondern dass darüber hinaus das Erwerbseinkommen die Gesundheit positiv beeinflusst, während ein hohes Transfereinkommen einen negativen Effekt hinsichtlich der Gesundheitskapitalbildung aufweist.

¹⁸ Es ist zu unterscheiden zwischen Hypothesen zur *Nachfrage nach Gesundheit* und der Nachfrage nach medizinischen Leistungen. Gemäss den theoretischen Voraussagen müsste bspw. mit höherem Alter die Nachfrage nach Gesundheit abnehmen, wogegen die Nachfrage nach medizinischen Leistungen mit dem Alter zunehmen sollte (vgl. dazu Breyer und Zweifel 1997: S. 70).

Eine Untersuchung, bei der alle Vorzeichen der Schätzkoeffizienten konsistent zu den theoretischen Vorhersagen sind, ist diejenige von Gerdtham et al. (1999). Die Nachfrage nach Gesundheit sinkt mit dem Preis von medizinischen Leistungen und dem Alter, und steigt mit dem Einkommen und dem Bildungsniveau¹⁹.

Zweifel und Breyer folgern aus den empirischen Ergebnissen, dass *«viele der Implikationen des Grossman-Modells von der empirischen Evidenz widerlegt worden sind. Insbesondere die Vorstellung, dass Ausgaben für medizinische Leistungen eine abgeleitete Nachfrage darstellen, die auf eine zugrunde liegende Nachfrage nach Gesundheit zurückgeht, kann nicht aufrecht erhalten werden, da der Gesundheitszustand und die Nachfrage nach medizinischen Leistungen negativ und nicht positiv korreliert sind»* (Zweifel und Breyer 1997: S. 72).

4.1.3 Arztinduzierte Nachfrage

Nachdem mit den vorangegangenen Unterabschnitten verschiedene nachfrageseitige Determinanten der Inanspruchnahme von Gesundheitsleistungen vorgestellt wurden, befasst sich dieser Unterabschnitt mit dem Einfluss der Angebotsseite von gesundheitlichen Leistungen. Wie im Unterabschnitt 2.1.2 gezeigt wurde, nimmt auf der Angebotsseite das Thema der Angebotsinduzierung eine zentrale Stellung ein.

Empirisch wird versucht, die Angebotsinduzierung in Situationen mit verschärfter Konkurrenz nachzuweisen. Die ersten internationalen empirischen Studien zu diesem Thema stammen von Evans (1974), welcher die internationale Diskussion massgeblich begründete (Kern 2002: S. 15). Danach folgten viele weitere Studien, welche zumeist auf US-amerikanischen Daten und deshalb auch auf den rechtlichen und institutionellen Rahmenbedingungen dieser Länder beruhen²⁰. Auch aus verschiedenen Ländern Europas sind viele neuere empirische Studien zum Thema der Angebotsinduzierung vorhanden, so bspw. aus Frankreich (Delattre et al. 2003), Schweden (Gerdtham 1996), Norwegen (Carlsen und Grytten 1998, Sorensen und Grytten 1999, Grytten und Sorensen 2001), Österreich (Bodenhöfer 2001) und Deutschland (Pohlmeier und Ulrich 1995, Andersen und Schwarze 1997, Cassel und Wilke 2000 und Kurth et al. 2004). Die Arbeiten aus Deutschland beziehen sich zumeist auf den stark reglementierten Bereich der vertragsärztlichen Versorgung. Eine steigende Arztdichte wird jedoch als relevanter Einflussfaktor auf die arztinduzierte Nachfrage gesehen (Neubauer, 1982; Breyer, 1984a).

Die in zahlreichen Untersuchungen gefundenen empirischen Ergebnisse und ihre Interpretation sind jedoch kontrovers. Weder die Resultate der Untersuchungen von Aggregatdaten noch diejenigen von Individualdaten lassen über die Existenz und das Ausmass der Angebotsinduzierung eindeutige Schlüsse zu. Während manche Autoren die Hypothese klar bestätigt sehen, sind sich andere nicht sicher oder finden diese nur für spezielle Formen von Gesundheitsgütern bestätigt (Feldman / Solan 1988). Nicht wenige Autoren lehnen die Hypothese sogar eindeutig ab²¹. Die im Kapitel 3 dargestellten Probleme widerspiegeln sich beim Versuch, sich einen Überblick über den Stand der Forschung zu

¹⁹ Vgl. Fussnote 18

²⁰ Bspw. Solan und Feldman (1978), Fuchs (1978), Pauly und Satterthwaite (1981), Rossiter und Wilensky (1983), Hemenway und Fallon (1985), Stano (1985), Feldman und Solan (1988), Rice und Labelle (1989), Dranove und Wehner (1994) sowie Labelle et al. (1994).

²¹ Gemäss Cassel (2000) erfolge die Ablehnung oder zumindest die Relativierung der These vor allem in neueren Studien. Er verweist dabei auf die Untersuchungen von Bradford/Martin 1995; Feldman/Solan 1988; Hay/Leahy 1982; Labelle/Stoddart/Rice 1994.

verschaffen. Das Hauptproblem besteht in der Schwierigkeit, die Forschungsergebnisse miteinander zu vergleichen. Die Gründe dafür sind u.a. folgende:

Es besteht keine einheitliche Definition des Begriffs angebotsinduzierter Nachfrage.

Es besteht keine übereinstimmende theoretische Analyse des Phänomens.

Der Inanspruchnahmeprozess wird unterschiedlich modelliert (einstufig vs. mehrstufig, simultan oder sequentiell).

Es werden verschiedene ökonometrische Schätzmethoden verwendet.

Die Inanspruchnahme wird in verschiedenen Masseinheiten gemessen (Mengen bzw. Geldeinheit).

Je nach Untersuchung wird zwischen verschiedenen Leistungsarten unterschieden.

Je nach Untersuchung wird die ÄrztlInnendichte für verschiedene ÄrztlInnentypen berücksichtigt.

Je nach Untersuchung werden (gleiche) strukturelle Variablen unterschiedlich aggregiert.

Es kann jedoch festgehalten werden, dass als wesentliche Variable zur Überprüfung von Angebotsinduzierung die Variable «ÄrztlInnendichte» verwendet wird. Unabhängig davon, mit welcher Schätzmethode und in welcher Masseinheit die abhängige Variable gemessen wird, werden als erklärende Größen für die Inanspruchnahme neben der «ÄrztlInnendichte» ein möglichst umfassender Variablenatz mit Variablen aus den Bereichen «environment», «need», «predisposing» und «enabling» mitberücksichtigt. Nur bei einem solchen Vorgehen kann für so genannte konfundierende Faktoren kontrolliert werden.

Gemeinsam ist zudem allen innerhalb der Literaturstudie begegneten Erklärungsmodellen, dass die Qualität der ärztlichen Leistungen als homogen betrachtet wird. Damit werden nicht beobachtbare Qualitätsunterschiede zur «Achillesferse» für die Überprüfung mittels empirischer Daten (vgl. Newhouse 1978: S. 60).

Die Diskussion um die Resultate einer Untersuchung von Delattre et al. (2003) wird beispielhaft für die vielen kontroversen Debatten um widersprüchliche Resultate dargestellt. Delattre et al. finden in ihrer Studie für Frankreich starke Evidenz für arztinduzierendes Verhalten. Die Resultate korrespondieren mit den Resultaten einer Studie aus Kanada (Nassiri et al. 2002), welche feststellt, dass Ärztinnen und Ärzte in Québec die Anzahl lukrativer Leistungen erhöhen, um ihr Einkommen zu erhalten. Die Resultate stehen auch im Einklang mit mehreren Resultaten in der empirischen Literatur, welche sich mit dem Verhalten von Ärztinnen und Ärzten befasst (vgl. McGuire 2000). Die Resultate widersprechen jedoch neueren Studien von Carlsen und Grytten (1998), Sorensen und Grytten (1999) und Grytten und Sorensen (2001). Sie finden keine empirische Evidenz für das Auftreten von arztinduziertem Verhalten für norwegische Ärzte/innen. Sorensen und Grytten (1999) weisen darauf hin, dass in den USA die Dichte der Ärztinnen und Ärzte viel höher sei als in Norwegen. Auf der anderen Seite verweisen sie darauf, dass das Entgeltungssystem mit Pauschalen (fixed fee system) in Norwegen Honorarinduzierung verhindert, wie es in Regionen mit einer hohen Dichte an Ärztinnen und Ärzten und mit dem System der Einzelleistungsvergütung (fee for service) in den USA oder Kanada beobachtet werden kann. In Frankreich besteht jedoch wie auch in Norwegen ein Entgeltungssystem mit Pauschalen, dennoch sind die Resultate aus den beiden Ländern widersprüchlich. Delattre et al. (2003) sehen zwei Möglichkeiten, die divergierenden Resultate zu erklären. Erstens verwenden sie im Gegensatz zu den norwegischen Studien Paneldaten, wodurch das Problem der unbeobachteten Heterogenität bezüglich der ÄrztlInnen und der Angebotscharakteristiken besser kontrolliert werden kann. Zweitens ist die ÄrztlInnendichte in Frankreich wesentlich höher als in Norwegen. In Übereinstimmung mit den theoretischen Prognosen könnte sich empirisch abbilden, dass angebotsinduziertes Verhalten davon abhängig ist, auf welchem Niveau sich ein Land bzw. eine Region in Bezug auf die ÄrztlInnendichte befindet.

4.1.4 Nachfragesteuerung, Anreizsysteme und Moral Hazard

In der Gesundheitsökonomie wird darüber nachgedacht, wie das Verhalten der beteiligten Akteurinnen und Akteure gesteuert werden kann, damit gesundheitliche Leistungen möglichst effizient erbracht werden. Dabei geht es vor allem um eine Steuerung durch monetäre Anreize. Einer der Hauptgründe der Finanzierungsprobleme seien die falsch gesetzten Anreize u.a. auf der Nachfrageseite in der Krankenversicherung. Ein Zweig der Gesundheitsökonomie befasst sich deshalb mit der Frage, inwiefern Steuerungsmöglichkeiten durch unterschiedliche Vertragsmodelle bestehen und welche Auswirkungen diese auf das Verhalten der Versicherten und im Endeffekt auf die Kosten des Gesundheitswesens ausüben. Dabei spielen Informationsasymmetrien²² zwischen den Beteiligten und Moral Hazard eine wichtige Rolle. Mit Hilfe von Studien wurde versucht, die Effekte von Anreizsystemen empirisch nachzuweisen und deren Höhe zu ermitteln.

Die wohl am meisten erwähnte und bekannteste empirische Untersuchung zur Wirkung von Anreizsystemen ist die als RAND-Studie bekannte, kontrollierte und randomisierte Studie «Health Insurance Experiment» aus Kalifornien. Obwohl bereits Mitte der 70er Jahre durchgeführt, gilt sie bis heute als zentrale Untersuchung zur Thematik. Anhand der Inanspruchnahme von medizinischen Leistungen sowie des Gesundheitsstatus wurden die Zusammenhänge zwischen verschiedenen Finanzierungsmodalitäten, der Nutzung von medizinischen Einrichtungen, Gesundheitsausgaben und gesundheitlichem Outcome untersucht (Manning et al. 1987: S. 253 f.). Dazu wurden ca. 6'000 US-Bürgerinnen und Bürger und ihre Familienmitglieder zufällig einem von 14 Krankenversicherungsverträgen zugeordnet und anschliessend während drei bis fünf Jahren in ihrem Inanspruchnahmeverhalten beobachtet. Die Verträge variierten u.a. in den Selbstbehalten (0%, 50% bzw. 95%) und in der maximalen Selbstbeteiligung, welche durch einen bestimmten Prozentsatz des Haushaltseinkommens bestimmt wurde (5%, 10% bzw. 15%, bis maximal US-\$ 1000.-). Newhouse et al. (1993) finden, dass durch Kostenbeteiligung der Versicherten das Verhalten relativ stark beeinflusst werden kann. Dabei zeigte sich, dass Kostenbeteiligungen ihre Wirkung fast ausschliesslich beim Erstkontakt mit Gesundheitseinrichtungen entfalten und nur einen vernachlässigbaren Einfluss auf die nachfolgende Nutzung von Gesundheitsleistungen pro Erkrankungsepisode nehmen (Richardson 1991: S. 21). In Bezug auf den gesundheitlichen Outcome wurde in späteren Zusatzauswertungen festgestellt, dass Kostenbeteiligungen nicht bloss die Zahl der «unbegründeten» Arztbesuche verringerte, sondern auch die Inanspruchnahme von medizinisch angebrachten Konsultationen und von Präventivmassnahmen. Insbesondere bei chronisch Kranken zeigten sich negative Auswirkungen auf die Sehfähigkeit (Lurie et al. 1989) sowie die Blutdruckeinstellung (Keeler et al. 1985). Insgesamt trafen bei der RAND-Studie die Kostenbeteiligungen die einkommensschwächsten sowie die «kränksten» zwei Fünftel der Gesellschaft besonders hart.

Aus der RAND-Studie können jedoch keine Erkenntnisse über die Gesamtwirkung von Wahlmöglichkeiten im Krankenversicherungsbereich (bspw. Franchisenwahl wie in der Schweiz) gezogen werden. Im Unterschied zu den Versicherten in der Schweiz hatten die Versuchspersonen des RAND-Experiments keine Wahlmöglichkeit bezüglich des Versicherungsmodells. Sie wurden zufällig auf die verschiedenen Versicherungskontrakte verteilt («randomisiertes Experiment»), womit durch das Studiendesign das Problem der Selektionseffekte neutralisiert werden konnte. In der Schweiz ist die Aus-

²² Das Gesundheitswesen ist durch asymmetrische Informationsverteilung gekennzeichnet: Ärztinnen und Ärzte und Krankenkassen können nicht beobachten, ob sich Patientinnen und Patienten gesundheitsbewusst verhalten. Krankenkassen und Patientinnen/Patienten wiederum verlassen sich auf das Urteil des Arztes hinsichtlich der angezeigten Therapien – und damit der verursachten Kosten. Diese Informationsasymmetrien können ineffizientes Verhalten auf Seiten der Ärztinnen/Ärzte und Patientinnen/Patienten fördern und so zu unnötig hohen Ausgaben führen.

gangslage eine andere und deshalb auch die Fragestellung. In der Schweiz müssen die Effekte der Wahlfreiheit untersucht werden. Neben den Anreizeffekten mit den verschiedenen Selbstbeteiligungsmodellen spielen jedoch auch die Selektionseffekte eine grosse Rolle (Schellhorn 2004: S. 4).

Gemäss Schellhorn (2004: S. 5) wurden die Effekte der Wahlfreiheit einer Krankenversicherung auf die Inanspruchnahme von medizinischen Leistungen das erste Mal von Cameron et al. (1988) für Australien empirisch modelliert. Die zentralen Erkenntnisse aus dieser Studie sind, dass das Einkommen eine zentrale Rolle für den Abschluss eines Versicherungsvertrages spielt, während der Gesundheitszustand eher unwichtig für diese Entscheidung ist. Für die Inanspruchnahme ist hingegen der Gesundheitszustand die wichtigste erklärende Grösse, während Einkommen hier weniger wichtig ist. Zudem weisen die Autoren sowohl Anreiz-Effekten als auch Selektionseffekten eine wichtige Rolle bei der Inanspruchnahme medizinischer Leistungen zu. Für die Schweiz liegen mehrere neuere Studien zu diesem Thema vor, da mit dem neuen Krankenversicherungsgesetz Wahlmöglichkeiten bezüglich der Kostenbeteiligung für die Versicherten bestehen. Diese Studien werden im Abschnitt 4.2 «Untersuchungen aus der Schweiz» vorgestellt.

4.2 Untersuchungen aus der Schweiz

Der Leitfaden für die Literaturanalyse zu den Untersuchungen der Schweiz bildet die im Jahr 2004 erschienene Publikation von Brigitte Bisig und Felix Gutzwiler, welche im Rahmen des nationalen Forschungsprogramms 45 des Schweizerischen Nationalfonds entstanden ist. Bisig (2004: S. 7f.) schreibt in der Einleitung zu den zwei erschienenen Bänden, dass bei Diskussionen um Kostendämpfungsmassnahmen im Gesundheitswesen das Begriffspaar Rationalisierung vor Rationierung bei der Erbringung von Gesundheitsleistungen auch in der Schweiz im Zentrum stehe. Dabei stelle sich die Frage, ob mit der Ressourcenlimitierung eine Minderung der Qualität der Versorgung einhergehe. Oberstes Ziel von Massnahmen im Gesundheitswesen müsse es deshalb sein, die Kosten bei hoher Adäquanz und Qualität der Leistungen zu kontrollieren. Eine wichtige Voraussetzung für gezielte Massnahmen sei die Kenntnis der Einflussfaktoren auf die Versorgung bzw. die Inanspruchnahme von Gesundheitsleistungen. Neben den bekannten individuellen Faktoren erwähnt sie auch die mögliche Bedeutung von strukturellen Faktoren wie bspw. den Versicherungsstatus und die ÄrztlInnendichte. Die bisher für die Schweiz vorliegenden Informationen zu Einflussfaktoren der Inanspruchnahme beurteilt sie als Basis für konkrete Aktionen als noch nicht geeignet. Dafür verantwortlich seien bspw. nicht aktuelle Daten, nicht vergleichbare zeitliche Perioden und die Verwendung unterschiedlicher Definitionen u.a. für soziale und kulturelle Indikatoren. Sie konstatiert deshalb für die Schweiz einen hohen Forschungsbedarf in diesem Bereich. Der nun vorliegende Endbericht soll gemäss den Autor/innen eine Bestandesaufnahme zum Inanspruchnahmeverhalten von gesundheitlichen Leistungen in der Schweiz liefern. Diese neue Studie wird im Unterabschnitt 4.2.1, welcher sich mit individuellen und strukturellen Faktoren der Inanspruchnahme befasst, detaillierter vorgestellt. Ein Kapitel der NFP-Studie befasst sich mit der Wohnregion und der Inanspruchnahme. Daraus lassen sich einige Hinweise auf die Angebotsinduzierung ableiten.

Neben dieser Studie gibt es andere Untersuchungen, welche sich explizit mit der Angebotsinduzierung in der Schweiz befassen. Einige dieser Arbeiten werden im Unterabschnitt 4.2.3 vorgestellt. Durch die Möglichkeit der Krankenversicherten, eine Wahlfranchise zu wählen, ist die Schweiz u.a. auch in Deutschland für Untersuchungszwecke in Bezug auf die Steuerung von Versichertenverhalten durch Anreizmodelle interessant geworden. Untersuchungen mit Schweizer Daten zur Thematik der Nachfragesteuerung mit Wahlfranchisen und Moral Hazard werden im Unterabschnitt 4.2.4 betrachtet.

4.2.1 Determinanten der Inanspruchnahme

Es wurde schon darauf hingewiesen, dass die umfassendste Studie zur Inanspruchnahme von Gesundheitsleistungen in der Schweiz mit den zwei Bänden von Bisig / Gutzwiler et al. (2004a, 2004b) vorliegt. Verschiedene Autoren und Autorinnen untersuchen, welche Faktoren – neben dem Gesundheitszustand der Patient/innen – die Inanspruchnahme von Gesundheitsleistungen beeinflussen. Hauptdatenquelle war die Schweizerische Gesundheitsbefragung SGB des Bundesamtes für Statistik von 1992 und 1997. Dabei wird versucht, die Inanspruchnahmedeterminanten sowohl von kurativen als auch von Vorsorgeleistungen zu ermitteln. Dazu werden die Anzahl Konsultationen bei verschiedenen ÄrztInnengruppen bzw. Leistungserbringer/innen – Spital, Notfall, Zahnarzt, Gynäkologe/in, übrige Spezialärztinnen und -ärzte, Allgemeinärztinnen und -ärzte – detailliert untersucht. Dabei zeigt sich, dass Faktoren wie Geschlecht, Wohnregion, Nationalität, Sozialschicht und Versicherungsstatus mitbestimmen, ob und wie oft jemand Gesundheitsleistungen nachfragt. Zudem wird ersichtlich, dass die Wirkung der geprüften Determinanten nicht bei allen Arten von Gesundheitsleistungen die gleichen Effekte aufweisen.

Einen wichtigen Beitrag zur Inanspruchnahmeforschung liefert die Studie, indem sie aus den vorhandenen Daten einen reliablen und validen Sozialschichtindex entwickelt. In ihm enthalten sind die Einzelindikatoren «sozioprofessionelle Klassen», das «Äquivalenzeinkommen», der Erwerbsstatus, das Bildungsniveau und die Wohnsituation mit der Dimension Peripherie-Zentrum. Es zeigt sich, dass die Sozialschicht, unter Kontrolle von anderen Einflussfaktoren wie dem Geschlecht, dem Alter, der Nationalität, der Sprachregion und dem Gesundheitszustand, einen eigenständigen Einfluss auf das Ausmass der Nachfrage nach Gesundheitsleistungen ausübt. Personen aus tieferen Sozialschichten nehmen seltener Vorsorgeuntersuchungen in Anspruch als Personen aus hoher Sozialschicht. Ein ähnlicher Trend lässt sich bei einigen Operationsarten (Gebärmutter, Mandeln, Blinddarm), der Komplementärmedizin, den Konsultationen bei Allgemein- bzw. Spezialärzten und bei Zahnarztkonsultationen beobachten. Aus den Resultaten geht jedoch leider nicht hervor, ob es sich bei den beobachteten Unterschieden um einen Nachfrage- oder einen Angebotseffekt handelt.

In der französischen Schweiz sind Spezialarztkonsultationen, ambulante Spitalkonsultationen und Notfallkonsultationen häufiger und Zahnarztkonsultationen seltener als in der deutschen Schweiz. Früherkennungsmassnahmen sind bei Personen in der Peripherie/in Kleinzentren seltener als bei Personen in Grosszentren / Metropolen. Krebsfrüherkennungsmassnahmen (wie Hautuntersuchung, Gebärmutterhalsabstrich, Brustuntersuchung (Palpation durch Arzt/Ärztin) und Prostatauntersuchung) sind bei Ausländerinnen und Ausländern seltener als bei Schweizerinnen und Schweizern. Aber auch der Versicherungsstatus scheint bei der Inanspruchnahme eine Rolle zu spielen: Personen mit einer Zusatzversicherung für stationäre Behandlungen im Spital (Privatversicherung) nehmen generell mehr Leistungen in Anspruch als Personen mit einer Allgemeinversicherung. Dies gilt für kurative, aber auch paramedizinische Leistungen (inkl. Komplementärmedizin) und Vorsorge-/Früherkennungsmassnahmen.

Ein Problem der Studie, welches auch die Autor/innen einräumen, besteht darin, dass die Menge der konsumierten Gesundheitsleistungen auf Selbstangaben der Befragten beruhen. Sie weisen darauf hin, dass die Daten zwingend durch alternative Statistiken im ambulanten und stationären Bereich ergänzt bzw. auf ihre Gültigkeit evaluiert werden müssten (Bisig et al. 2004a: S. 233).

4.2.2 Überprüfung des Grossman-Modells

Leu und Doppmann (1986) finden bei einer ersten Überprüfung des Grossman-Modells anhand von Daten aus der Schweiz einen signifikanten, negativen Einfluss des Alters, des Beschaffungszeitpreises und der Familiengrösse auf die Inanspruchnahme. Männer konsultieren seltener als Frauen, kranke Personen konsultieren häufiger als gesunde, Versicherte häufiger als Nichtversicherte. Zudem erweist sich die ÄrztInnendichte als signifikante positive Einflussgrösse, wobei jedoch nicht für regionsbezogene Merkmale kontrolliert wird.

In einer neueren Arbeit, ebenfalls zur Testung des Grossman-Modells, berücksichtigen Leu und Gerfin (1992) auch Indikatoren des Gesundheitsverhaltens wie Raucherstatus, Schlafgewohnheiten, sportliche Aktivität. Die Arbeit beruht ebenfalls auf Daten aus der Schweiz und schliesst nur Erwerbstätige ein. Die ermittelten Gesamteffekte zeigen auch hier wieder einen negativen Alterseffekt und den bekannten Geschlechtereffekt mit einer niedrigeren Inanspruchnahme der Männer im Vergleich zu den Frauen. Bezüglich der Verhaltensmerkmale zeigt sich, dass Raucher seltener konsultieren als Nichtraucher und dass Personen, die keinen Sport treiben, seltener konsultieren als Personen, die Sport treiben. Es zeigt sich ebenfalls wieder ein signifikant positiver Einfluss der ÄrztInnendichte auf die Zahl der Konsultationen. Auch hier wird jedoch nicht für weitere regionale Merkmale kontrolliert.

4.2.3 Arztinduzierte Nachfrage

Neben Untersuchungen von Aggregatdaten, die mit dem bekannten Problem des ökologischen Fehlschlusses behaftet sind, wurde noch kaum versucht, explizit die Existenz der Angebotsinduzierung in der Schweiz mit Hilfe von Individualdaten zu untersuchen. Mit ein Grund dafür dürfte sein, dass die für die Überprüfung dieser Frage am häufigsten verwendete Variable, die ÄrztInnendichte, nur auf Kantonsbasis vorhanden ist. Zum einen ist die Kantonsbasis unter dem Aspekt eines «abgeschlossenen» Gesundheitsmarktes für die Überprüfung der Angebotsinduzierung nicht sehr geeignet. Und zum anderen sind Individualdatensätze mit Informationen über den Gesundheitszustand der Individuen (wie bspw. die Schweizerische Gesundheitsbefragung) nur auf Grossregionen- und nicht auf Kantons-ebene repräsentativ. Wiederum andere Individualdatensätze, welche auf Kantonsbasis Informationen enthalten, weisen zu wenige Variablen aus dem sozioökonomischen Bereich der Individuen auf (bspw. sind im Datenpool keine Informationen über Einkommen, Bildung, etc. vorhanden).

Die erwähnte NFP 45-Studie von Bisig und Gutzwiler et al. stellt deutliche Unterschiede bei den Arztkonsultationen nach Sprach- und MS-Regionen fest (Bisig / Gutzwiler 2004: S. 119-146). Einen Hinweis, dass das Angebot teilweise die Nachfrage steuern kann, sehen die Autor/innen vor allem in den Analysen der Inanspruchnahme nach MS-Regionen. Spezialarztkonsultationen, ambulante Spitalkonsultationen und Notfallkonsultationen sind in Metropolen/Grosszentren statistisch signifikant häufiger als in der Peripherie. Die Unterschiede lassen sich weder mit soziodemographischen Unterschieden noch mit Unterschieden im Gesundheitszustand erklären. Da in Metropolen/Grosszentren das Angebot von Spezialarztpraxen und Spitalnotfallzentren höher ist als in der Peripherie, könne vermutet werden, dass das Angebot teilweise die Nachfrage steuert. Es wird jedoch auch darauf hingewiesen, dass sich diese Hypothese methodisch nicht eindeutig überprüfen lasse, da auf der Angebotsseite noch keine entsprechenden Daten vorliegen. Ein anderes Problem besteht in der Messung der Nachfrage selbst, da es sich bei der Gesundheitsbefragung um Selbstangaben ohne medizinische Bestätigung handelt.

Marty (1997) wählt einen etwas anderen Ansatz. Um Evidenz für das Auftreten von angebotsinduzierter Nachfrage zu finden, untersucht er die Verrechnungspraxis von 81 Ärzten und Ärztinnen aus

Basel-Stadt und bildet daraus Indikatoren, welche als Verhaltensmotive für Nachfrageinduzierung interpretiert werden können. Er findet einen positiven Zusammenhang zwischen diesen Indikatoren und dem Umsatz pro Patient/in, was als Evidenz für das Auftreten von angebotsinduzierter Nachfrage betrachtet wird. Selektionsmechanismen können jedoch nicht ausgeschlossen werden und es wird unterstellt, dass es zwischen den untersuchten Ärzten/innen keine Qualitätsunterschiede gibt.

Crivelli et al. (2004) untersuchen anhand von aggregierten Daten auf Kantonsbasis aus der Periode 1996–2001 den Zusammenhang zwischen der kantonalen ÄrztlInnendichte und den kantonalen Gesundheitskosten pro Kopf. Gemäss den Autoren/innen steigen die Gesundheitsausgaben pro Kopf um 2.0%, wenn die ÄrztlInnendichte um 10% steigt. Mit Hilfe von Paneldaten und der Verwendung eines ökonometrischen Verfahrens, welches gleichzeitig längsschnittliche als auch querschnittmässige Effekte für die Analyse der zu erklärenden Grösse (kantonale Gesundheitskosten pro Kopf) berücksichtigen kann (Times-Series-Cross-Section-Analysis), prüfen sie den Einfluss von elf kantonsspezifischen Faktoren auf die kantonalen Pro-Kopf-Gesundheitskosten. Sie folgen dabei dem Ansatz von Gerdtham und Jönsson (2000), welcher für internationale Vergleiche von Gesundheitssystemen verwendet wurde. Als Erklärung für die Gesundheitskosten pro Kopf fliessen demographische, sozioökonomische und Produktionskapazitäts-Faktoren ins Schätzmodell ein. Neben der kantonalen ÄrztlInnendichte werden folgende Variablen berücksichtigt: Das kantonale Pro-Kopf-Einkommen, der Anteil an älteren (über 75 Jahren) bzw. jungen (unter 5 Jahren) Kantonsbewohner/innen, die Arbeitslosenquote, eine Armutsrate²³, eine Art gesundheitsversorgungsbezogene Mortalitätsrate²⁴, die Bevölkerungsdichte (als Proxy für die Urbanität eines Kantons), einen kantonalen Index für direkte Demokratie, die Dichte für Akut-Spitalbetten und der Anteil der Bevölkerung, welcher ein Managed Care Programm als Versicherungsform wählt. Fünf Faktoren weisen einen statistisch signifikanten Einfluss auf die Pro-Kopf-Gesundheitskosten aus. Dazu gehören die ÄrztlInnendichte, das kantonale Pro-Kopf-Einkommen, der Anteil an älteren Kantonsbewohner/innen über 75 Jahren, die Arbeitslosenquote und die Bevölkerungsdichte.

Der statistisch signifikante Zusammenhang zwischen ÄrztlInnendichte und den Pro-Kopf-Ausgaben für Gesundheitsleistungen liefert Hinweise auf Angebotsinduzierung für die Schweiz. Man kann jedoch nicht von definitiven Belegen sprechen. Der gewichtigste Einwand besteht darin, dass mit Aggregatdaten keine Rückschlüsse auf individuelles Verhalten gezogen werden können. Das Konzept der Arztinduzierung basiert auf individuellen Motiven und deshalb auch auf individuellem (ärztlichem) Verhalten. Es besteht deshalb die Gefahr des ökologischen Fehlschlusses, insbesondere bei der hohen Komplexität des Inanspruchnahmeverhaltens. Zudem ist die Aggregationsebene der Kantone problematisch. Die Kantone müssten dafür einen möglichst «abgeschlossenen» Gesundheitsmarkt bilden. Dies dürfte jedoch weder auf die sehr kleinen, noch auf die sehr grossen Kantone zutreffen (vgl. dazu Abschnitt 3.2 «Methodische Probleme»).

Rüefli / Vatter (2001) analysieren die kantonalen Unterschiede zwischen den öffentlichen Gesundheitsausgaben und zwischen den Krankenversicherungskosten für die Zeitperiode 1994 bis 1999 mit Hilfe von Aggregatdaten. Dabei unterscheiden sie zwischen der Erklärung des Niveaus und der Zunahme. Für die Analyse entwickeln die Autoren ein komplexes Erklärungsmodell, das einerseits die Nachfrage- und die Angebotsseite im Gesundheitswesen berücksichtigt, andererseits aber auch er-

²³ Die Armutsrate wurde gemessen am Anteil an Haushalten mit einem Einkommen, welches unterhalb von 50% des mittleren kantonalen Einkommens liegt.

²⁴ Die Rate basiert auf einem Konzept, welches davon ausgeht, dass gewisse Todesfälle unter Voraussetzung einer effektiven Gesundheitsversorgung vermieden werden könnten (vgl. dazu Nole und McKee 2003).

weiterte Indikatoren, die kantonale Unterschiede beschreiben können. Es werden mit Hilfe von Querschnitt-Längsschnitt-Regressionsanalysen über 100 Indikatoren empirisch getestet. Die Ärztedichte (auf Kantonsebene) erweist sich auch in einem multivariaten Setting als starker Erklärungsfaktor für das Kostenniveau. Je höher die Dichte, desto höher die Kosten.

Domenighetti befasste sich über längere Zeit mit der Frage der Angebotsinduzierung in der Schweiz. Auch er verwendet wie Crivelli keine Individualdaten, sondern versuchte, mit Hilfe der kantonalen ÄrztlInndichte und den kantonalen Pro-Kopf-Ausgaben für Gesundheitsleistungen die These der Angebotsinduzierung zu belegen (Domenighetti et al. 1994; Domenighetti 1998). Im Gegensatz zu Crivelli beruht seine Analyse jedoch nicht auf Paneldaten. Er verwendet für die Analyse bivariate Zusammenhangsanalysen bzw. (multivariate) lineare Regressionsschätzungen und findet einen positiven Zusammenhang zwischen ÄrztlInndichte und den Pro-Kopf-Ausgaben für ambulante Leistungen bzw. den durchschnittlichen Arztbesuchen. Zudem besteht ein statistisch negativer Zusammenhang zwischen der ÄrztlInndichte und der Anzahl Konsultationen pro Arzt/Ärztin bzw. dem Umsatz pro Arzt/Ärztin. Im linearen Regressionsmodell von 1994 wird nur für das Pro-Kopf-Einkommen korrigiert, in der Analyse von 1998 werden bivariate Korrelationsanalysen als Grundlage für die Aussagen zur Angebotsinduzierung genommen. In einer repräsentativen Befragung von Ärztinnen und Ärzten in der Schweiz stimmen 70% der Befragten der Aussage zu, dass ihre Kolleg/innen die Inanspruchnahme von Gesundheitsleistungen bei ihren Patient/innen erhöhen, wenn ihr Einkommen durch eine erhöhte Konkurrenz bedroht wird (Domenighetti et al. 2002). Mit den empirisch nachgewiesenen Zusammenhängen kann aus methodischer Sicht jedoch aus denselben Gründen wie für die Resultate von Crivelli et al. (2004) nicht beantwortet werden, ob es sich zwischen der ÄrztlInndichte und den Pro-Kopf-Ausgaben für Gesundheitsleistungen um einen kausalen Effekt handelt. Zudem kann bei bivariaten Analysen nicht für andere relevante Determinanten kontrolliert werden. Die Resultate liefern jedoch Hinweise, dass zumindest ein Teil der kantonalen Unterschiede auf Angebotsinduzierung zurückzuführen sein könnte.

Horisberger et al. (2004) kommen zum Schluss, dass sich die These, eine zu hohe ÄrztlInndichte hätte die Leistungen der Krankenkassen oder die Prämien in die Höhe getrieben, statistisch nicht belegen lässt. Sie arbeiten dabei mit demselben Datenmaterial wie Domenighetti und Crivelli. Sie untersuchen den bivariaten Zusammenhang zwischen der prozentualen Veränderung der kantonalen ÄrztlInndichte und der kantonalen prozentualen Veränderungen der Pro-Kopf-Kassenleistungen bzw. Prämien. Diese erweisen sich im Längsschnitt als statistisch nicht signifikant. Sie kommen deshalb zum Schluss, dass sich Veränderungen der Prämien und der Kassenleistungen sich kaum über die Höhe bzw. die Veränderungen der ÄrztlInndichte erklären lassen. Sie erwähnen, dass der rein angebotsorientierte Ansatz ausser acht lässt, dass es sich bei der Kostenentwicklung im Gesundheitswesen auch um ein Strukturphänomen handelt und verweisen darauf, dass die ÄrztlInndichte in der Schweiz schon seit über 100 Jahren ein West-Ost-Gefälle aufweist, welches deutlich sozial, weltanschaulich und wirtschaftlich bestimmt ist. Zudem bemängeln sie, dass eine Analyse der Kostenentwicklung nach anderen Kriterien als nach Kantonen, bspw. auf der Dimension Stadt/Land, mittels der vorliegenden Daten nicht möglich sei. Auch diese Studie kontrolliert jedoch nicht für andere Determinanten der Inanspruchnahme.

4.2.4 Nachfragesteuerung, Anreizsysteme und Moral Hazard

Bevor die neusten Studien zu den Effekten der Wahlmöglichkeit der Höhe der Selbstbeteiligung auf die Inanspruchnahme von Gesundheitsleistungen vorgestellt wird, soll nochmals kurz das Problem der Selbstselektion und Moral Hazard erläutert werden. Dabei geht es darum, dass bei der Wahl eines

bestimmten Franchisemodells zwei Effekte auftreten können. Zum einen wird erwartet, dass der/die Versicherte weniger Gesundheitsleistungen konsumiert, je mehr er/sie dafür bezahlen muss. Er/sie verhält sich sparsamer. Dieser Effekt wird als Anreizeffekt bzw. als Reduktion von Moral Hazard Verhalten bezeichnet. Andererseits weiss jede/r Versicherte, dass eine Sparmöglichkeit durch die Wahl einer hohen Selbstbeteiligung nur dann möglich ist, wenn man relativ gesund ist und keine grösseren Ausgaben im Gesundheitsbereich erwartet. Deshalb wählen vorwiegend solche Versicherte eine hohe Selbstbeteiligung, welche gesund sind und keine grösseren Kosten für Behandlungen erwarten. Bei einem solchen Effekt spricht man von einem Effekt aufgrund von Selbstselektion (Selektionseffekt).

Werblow und Felder (2003) untersuchen Paneldaten (1997 bis 1999) von 75'000 Versicherten aus dem Kanton Zürich. Unter Berücksichtigung der Endogenität der Vertragswahl können sie zeigen, dass Selbstbehalte zu einer signifikanten Reduktion der Ausgaben führen. Im tieferen Franchisenbereich (CHF 300 und 600.-) führen sie 30% der Kostenersparnisse auf Moral Hazard Verhalten zurück und 70% auf Selektion. Im höheren Bereich (CHF 1'200 und 1'500.-) seien es sogar 70% Moral Hazard und nur 30% Selektion.

Schellhorn (2002, 2004) kommt zu anderen Resultaten als Werblow und Felder. Er sieht die These bestätigt, dass die Wahlmöglichkeit des Selbstbehalts keinen starken Einfluss auf das Patientenverhalten ausübt. Den Grossteil der beobachteten Reduktion in der Anzahl der Arztbesuche bei Versicherten mit erhöhten Franchisen führt er auf Selbstselektion zurück.

In einem Koreferat zu den Beiträgen von Schellhorn und Werblow kritisiert Steiner (2002) u.a., dass die Ergebnisse von Werblow stark von der Spezifikation des Selbstbehalts in der Schätzgleichung abhängen (vgl. dazu «dateninduzierte Kontingenz» in Abschnitt 3.2). Wird der Selbstbehalt als kategoriale Variable spezifiziert, zeigen sich keine statistisch signifikanten Effekte, wird dieser jedoch mittels Dummy-Variablen spezifiziert, ergeben sich signifikante negative Effekte. Die hauptsächliche Kritik der Studie von Werblow und von Schellhorn betrifft jedoch die Identifikation des kausalen Effektes eines Selbstbehalts. Damit dieser von Selektionseffekten getrennt werden könne, seien ökonomisch begründbare Ausschlussrestriktionen erforderlich, wolle man sich nicht nur auf Identifikation durch Annahmen über die funktionale Form der Selektion verlassen²⁵. Darauf geht Werblow in seiner Arbeit gar nicht und Schellhorn zu wenig überzeugend ein.

Den methodologisch überzeugendsten Ansatz zur Identifikation eines kausalen Effektes verfolgten Gardiol et al. (2003). Sie finden in ihren Daten von 62'000 Versicherten der CSS mit der höchsten Franchisen aus dem Kanton Waadt sowohl einen starken Selektionseffekt als auch eine Reduktion des Moral Hazard Effektes. Für ihre recht einfache, aber überzeugende Methode benötigen sie eine beobachtbare Variable, welche einerseits mit dem unbeobachtbaren Gesundheitszustand korreliert und andererseits durch die Anreizeffekte nicht beeinflusst wird. Der statistisch signifikante Zusammenhang zwischen Mortalitätsrate und Franchisenhöhe kann als starke Evidenz für das Vorhandensein von Selektionseffekten interpretiert werden. Die Sterberate in der Gruppe von Personen mit den tiefsten Franchisen ist unter Kontrolle von Geschlecht und Alter um das Doppelte höher als in den Gruppen der Versicherten mit einer Franchise, welche höher als CHF 600.- ist. Es wird davon ausge-

²⁵ Je nach Schätzverfahren und Modellierung des Entscheidungsprozesses sind zur Identifikation der Modelle (Wahl des Selbstbehaltes bzw. Höhe der Inanspruchnahme) sog. Ausschlussrestriktionen nötig. D.h. einige Variablen dürfen nur für die Wahl des Selbstbehaltes eine Rolle spielen, nicht aber für die Inanspruchnahme (bspw. Anzahl Arztbesuche). Nur in der Gleichung für die Wahl des Selbstbehalts verwendet Schellhorn dafür zwei Variablen: eine Dummy-Variable für das Vorhandensein einer Spitalzusatzversicherung und vier Dummy-Variablen, die Regionen mit ähnlicher Prämienhöhe zusammenfassen. Steiner betrachtet diese Ausschlussrestriktionen jedoch als wenig überzeugend.

gangen, dass die Mortalität von den Anreizeffekten nicht beeinflusst wird, was als realistisch angesehen werden kann. In einem zweiten Analyseschritt werden einige Personen aus dem Sample ausgeschlossen, damit reliablere Resultate möglich werden (Verminderung der unbeobachteten Heterogenität)²⁶. Mit diesem «homogenisierten» Sample wird nun getestet, inwiefern es sich bei den beobachtbaren Kostenunterschieden zwischen den unterschiedlichen Franchisengruppen um Selektionseffekte bzw. Anreizeffekte (Verminderung von Moral Hazard) handelt. Dazu werden je zwei Regressionen miteinander verglichen, welche sich in der abhängigen Variable unterscheiden. Es wird den Effekt der Wahlfranchise verglichen, auf die Wahrscheinlichkeit, dass eine Person stationäre bzw. ambulante Kosten erzeugt, unter Kontrolle von Alter, Wohnregion und dem Vorhandensein von Zusatzversicherungen. Es wird davon ausgegangen, dass der Anreizeffekt auf die stationären Kosten nahe bei null liegt, da stationäre Aufenthalte aus medizinischer Sicht meistens notwendig sind²⁷ und die Kosten nur einer Nacht in einem Spital höher sind als die höchste Selbstbeteiligung der Wahlfranchisen. Somit besteht für die Versicherten kein Anreiz mehr, Kosten zu sparen. Wenn nun ein negativer Effekt der gewählten Franchise auf die Wahrscheinlichkeit, stationäre Kosten zu erzeugen, zu beobachten ist, weist dies auf einen Selektionseffekt hin. Aufgrund der Resultate zeigen die Autoren, dass die Wahrscheinlichkeit, dass ambulante Kosten überhaupt entstehen, durch die Wahl der höchsten Franchise im Vergleich zur niedrigsten aufgrund des Anreizeffektes um 69% gesenkt wird. Gegeben, dass stationäre Kosten vorhanden sind, werden die Kosten der Versicherten mit der höchsten Franchise im Vergleich zu denjenigen mit der tiefsten insgesamt um 34% gesenkt, wovon die Hälfte auf Selektionseffekte und die andere auf Anreizeffekte zurückzuführen ist.

4.3 Zusammenfassung

Im Kapitel 4 wurde der Stand der Forschung zum Inanspruchnahmegeschehen von Gesundheitsleistungen auf internationaler Ebene und für die Schweiz dargestellt. Es werden die Resultate von empirischen Untersuchungen aus zwei unterschiedlichen Bereichen der Inanspruchnahmeforschung vorgestellt. Der erste Bereich befasst sich mit den verschiedenen Determinanten des Inanspruchnahmegeschehens und der zweite mit der Analyse und Wirkung von Steuerungsinstrumenten.

Determinanten des Inanspruchnahmegeschehens

Weil es unterschiedliche theoretische Ansätze gibt oder weil im Bereich des Inanspruchnahmegeschehens unterschiedliche Forschungsschwerpunkte gesetzt werden können, wurden die Resultate der empirischen Untersuchungen zum Inanspruchnahmegeschehen in drei separaten Abschnitten vorgestellt.

In den Abschnitten 4.1.1 und 4.2.1 wurden zuerst aufgrund der Resultate von internationalen bzw. schweizerischen Untersuchungen allgemeine Determinanten der Inanspruchnahme vorgestellt. Grundsätzlich sprechen sowohl theoretische Überlegungen als auch die Resultate der empirischen Untersuchung eher für die Verwendung von Individualdaten, wenn das Inanspruchnahmegeschehen modelliert werden soll. Diejenigen Untersuchungen, welche in ihren Schätzmodellen neben individuellen auch umweltbezogene und angebotsseitige Merkmale miteinbeziehen (meist Aggregatdaten), scheinen das Inanspruchnahmegeschehen insgesamt besser abbilden zu können, als solche, welche

²⁶ (1) Es werden nur Männer betrachtet, weil die Kosten für eine Schwangerschaft von den Versicherungen vollumfänglich übernommen werden. (2) Nur Personen, welche von 1997 bis 2000 bei der CSS versichert waren. (3) Nur Personen, welche über die gesamte Periode die Franchisenhöhe nicht gewechselt haben. (4) Alter mindestens 25 Jahre. (5) Ohne Bezüger von staatlichen Prämienverbilligung. (6) Ohne IV-Rentenbezüger.

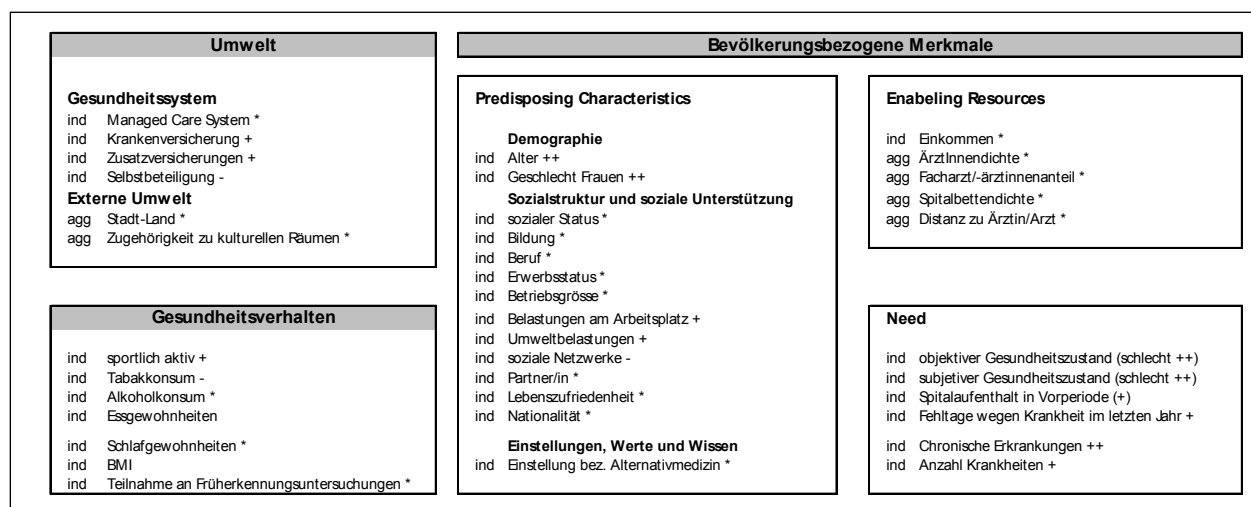
²⁷ Die Preiselastizität bei stationären Aufenthalten liegt (empirisch) nahe bei null (Vgl. Newhouse 1993).

entweder nur mit individuellen Faktoren oder aber nur mit aggregierten Faktoren modellieren. Mit Hilfe des Verhaltensmodells nach Andersen können die von den Forscherinnen und Forschern überprüften Determinanten grob in vier Kategorien – Environment, Predisposing, Enabeling und Need – eingeteilt werden. In **Abbildung 3** wird versucht, eine Übersicht über die hauptsächlich untersuchten Determinanten zu geben.

Zwei weitere Abschnitte (4.1.2 und 4.2.2) befassten sich mit den Resultaten von empirischen Untersuchungen zur Überprüfung des Grossman-Modells. Die aus der Theorie abgeleiteten Hypothesen werden von den empirischen Resultaten zumindest teilweise widerlegt. Es fällt jedoch auf, dass die zur Überprüfung des Modells verwendeten Variablen dieselben sind wie sie zur Überprüfung von anderen Modellen, u.a. dem Verhaltensmodell nach Andersen, benutzt werden.

In den letzten beiden Abschnitten (4.1.3 und 4.2.3) aus dem Bereich der Analyse des Inanspruchnahmegeschehens werden Studien vorgestellt, welche Aussagen über das Phänomen der Angebotsinduzierung zulassen. Als wesentliche Variable zur Überprüfung von Angebotsinduzierung wird die «ÄrztInnendichte» verwendet. Unabhängig davon, mit welcher Schätzmethode und in welcher Masseneinheit die abhängige Variable gemessen wird, werden als erklärende Grössen für die Inanspruchnahme neben der «ÄrztInnendichte» ein möglichst umfassender Variablensatz mit Variablen aus den Bereichen «environment», «need», «predisposing» und «enabeling» mitberücksichtigt. Deshalb enthalten solche Studien auch immer Informationen über das Inanspruchnahmegeschehen und deren Determinanten insgesamt.

Abbildung 3: Auswahl aus empirisch überprüften Determinanten der Inanspruchnahme und ihre Wirkungsrichtung auf die Höhe der Inanspruchnahme (verschieden instrumentalisiert), kategorisiert nach dem Verhaltensmodell nach Andersen



Legende:

- ++ unbestrittener, starker positiver Einfluss; + unbestrittener positiver Einfluss; - unbestrittener negativer Einfluss;
 - * Einfluss vorhanden, jedoch nicht unbestritten
 - ind: Variable auf individueller Ebene;
 - agg: Variable auf aggregierter Ebene
- Quelle: Eigene Darstellung

Dies ist nicht zuletzt durch die Notwendigkeit einer umfassenden Confounderkontrolle bedingt, durch die gezwungenermassen auch solche Faktoren berücksichtigt werden müssen, die nicht – wie die «ÄrztInnendichte» bzw. die Angebotsinduzierung – im Mittelpunkt des Interesses stehen. Das Hauptproblem bei der Beurteilung des Standes der internationalen Forschung, auch im Vergleich zum Stand in der Schweiz, besteht in der Schwierigkeit, die Forschungsergebnisse miteinander zu vergleichen. Die Gründe dafür sind u.a. folgende:

Es besteht keine einheitliche Definition des Begriffs angebotsinduzierter Nachfrage.

Es besteht keine übereinstimmende theoretische Analyse des Phänomens.

Der Inanspruchnahmeprozess wird je nach theoretischem Ansatz, aber auch nach Verfügbarkeit von Daten unterschiedlich modelliert (einstufig vs. mehrstufig, simultan oder sequentiell).

Es werden verschiedene ökonometrische Schätzmethoden verwendet.

Die Inanspruchnahme wird in verschiedenen Masseinheiten gemessen (Mengen bzw. Geldeinheit).

Je nach Untersuchung wird zudem zwischen verschiedenen Leistungsarten unterschieden.

Je nach Untersuchung wird die ÄrztInnendichte für verschiedene ÄrztInnentypen berücksichtigt.

Je nach Untersuchung werden (gleiche) strukturelle Variablen unterschiedlich aggregiert.

In Bezug auf den Forschungsstand der Schweiz ist festzuhalten, dass eine umfassende Untersuchung, welche explizit das Phänomen der Angebotsinduzierung mit Hilfe einer auf Individualdaten beruhenden Mehrebenenanalyse analysiert, fehlt. Es sind verschiedene Studien mit Aggregatdaten auf Kantonsbasis vorhanden, welche höchstens Hinweise auf Angebotsinduzierung in der Schweiz liefern. Das bedeutet, dass in diesem Bereich in der Schweiz eine grosse Forschungslücke besteht.

Analyse und Wirkung von Steuerungsinstrumenten

Für die Analyse von Steuerungsinstrumenten auf der Nachfrageseite sind verschiedene Studien über die Wirkung von Selbstbeteiligungsmodellen vorhanden. Bei der Analyse der Wahlfranchisen im KVG kommen verschiedene Studien zum Schluss, dass sich der Gesamteffekt der Wahlfreiheit auf zwei Effekte verteilt. Einerseits können Selektionseffekte beobachtet werden. Dies, weil Versicherte mit einer besseren Gesundheit eher eine hohe Selbstbeteiligung wählen, wodurch diese im Vergleich zu denjenigen Personen mit tieferen Franchisen geringere Kosten aufweisen. Andererseits können jedoch auch Anreizeffekte durch die Wahl einer höheren Franchise beobachtet werden. Es kann empirisch nachgewiesen werden, dass der Moral Hazard Effekt bei Personen mit einer hohen Selbstbeteiligung auch unter Berücksichtigung der Selbstselektion zu deutlich tieferen Kosten führt. Wie hoch das Ausmass der einzelnen Effekte ist, bzw. in welchem Verhältnis sie zueinander stehen, darüber herrscht, weil verschiedene Resultate vorliegen, keine Einigkeit.

5 Von der Erklärung der Inanspruchnahme zum Bedarf

In Teil 1 wurde bisher verschiedentlich darauf hingewiesen, dass zwischen der *Inanspruchnahmeforschung* einerseits und der *Bedarfsanalyse bzw. -planung* andererseits wesentliche Unterschiede bestehen. Bevor verschiedene Konzepte zur konkreten Bestimmung des Bedarfs im Kapitel 7 vorgestellt werden, wird im Kapitel 5 vertiefend dargestellt, wodurch sich die beiden Forschungsrichtungen unterscheiden und wo ihre Berührungspunkte sind. Dadurch wird klar, welche Bedeutung die bisherige Literaturliteraturanalyse erlangt.

Im Zentrum der *Inanspruchnahmeforschung* steht das Erkenntnisinteresse über das Inanspruchnahmegeschehen. Wovon hängt es ab, wann und wie oft jemand eine Ärztin oder einen Arzt aufsucht? Warum, wann und in welchem Ausmass nimmt eine Person gesundheitliche Leistungen in Anspruch? Um solche und ähnliche Fragen zu beantworten, werden mit Hilfe von verschiedenen theoretischen Ansätzen Erklärungsmodelle erstellt und spezifiziert, mit denen man versucht, die Wirkung von verschiedenen Einflussfaktoren auf das Inanspruchnahmegeschehen zu testen und zu beurteilen. In die Modelle werden Einflussfaktoren miteinbezogen, welche die Entscheidungen sowohl der PatientInnen wie der ÄrztInnen beeinflussen. Das Ziel der Inanspruchnahmeanalysen besteht demnach in der Beschreibung und der Erklärung des *Ist-Zustandes* von Inanspruchnahme, d.h. des gesamten Entscheidungsprozesses, welcher zur realisierten Nachfrage führt. Im Teil 1 wurden die theoretischen Ansätze, die Methoden, die Probleme und der Stand der Forschung zur Inanspruchnahme ausführlich vorgestellt und diskutiert.

Im Zentrum der *Bedarfsanalysen und -prognosen* steht nicht die Frage, wie viele gesundheitliche Leistungen eine Person effektiv in Anspruch nimmt, sondern welche und wie viele gesundheitliche Leistungen eine Person bzw. die Versicherten in einer Region effektiv benötigt/benötigen. Im Gegensatz zur Inanspruchnahmeforschung, welche einen Ist-Wert zu erklären versucht (Menge der effektiv realisierten Nachfrage), steht bei der Frage nach der effektiv benötigten Menge ein *Soll-Wert* im Zentrum des Interesses. Bedarfsanalysen und -prognosen werden deshalb, im Gegensatz zu Erklärungsmodellen, vor allem mit dem Zweck zur *Steuerung des Gesundheitswesens* durchgeführt. Interventionen zur Steuerung eines Systems sind vor allem dann notwendig, wenn Anzeichen vorhanden sind, dass ein bestehendes System nicht (mehr) in der Lage ist, sich in befriedigender Weise selbst zu regulieren. Im Falle des Gesundheitswesens geht es dabei u.a. um die Frage der effizienten Bereitstellung von Gesundheitsleistungen²⁸. Im Rahmen dieses Berichtes steht vor allem die Steuerung und damit der Bedarf im Zentrum des Interesses.

5.1 Von der individuellen auf eine aggregierte, regionale Ebene – Die Bildung von medizinischen Versorgungsgebieten

Bedarfsanalysen, bei denen die Frage der Über- und Unterversorgung im Zentrum steht, müssen auf einer aggregierten Ebene stattfinden, weil damit ein strukturelles Phänomen gemeint ist. Ein Individuum kann «über- bzw. unterversorgt» sein. Damit meint man, dass ein Individuum zu viel bzw. zu wenig an Leistungen bezogen hat. Hinter einer solchen Aussage steht jedoch nicht das Konzept der Über- und Unterversorgung. Über- bzw. Unterversorgung bezieht sich in Zusammenhang mit der Gesundheitsversorgung auf geographische Gebietseinheiten. Die Bedarfsanalysen dienen in diesem Zusammenhang dazu, regionenspezifische Soll-Werte (Ziel-Werte, Referenzwerte) zu ermitteln, wel-

²⁸ Die Frage der Über- und Unterversorgung und diejenige der Effizienz stehen in einem engen Zusammenhang. Besteht eine Überversorgung, werden knappe Ressourcen nicht effizient eingesetzt.

che den entsprechenden Ist-Werten (d.h. der regionenspezifischen, realisierten Nachfrage) gegenübergestellt werden. Erst mit einem Vergleich der realisierten Nachfrage mit dem Referenzwert innerhalb eines bestimmten Gebietes kann die Frage einer (regionenspezifischen) Über- bzw. Unterversorgung schlüssig beantwortet werden.

Die empirischen Arbeiten zur Inanspruchnahme von Gesundheitsleistungen zeigen, dass die Inanspruchnahme von sehr vielen Faktoren beeinflusst wird und sehr komplex ist. Für die Bestimmung eines regionalen Leistungsbedarfs müssen die Erkenntnisse der Inanspruchnahmeanalysen deshalb in der einen oder anderen Form berücksichtigt werden. Es ist jedoch bei einer Bedarfsanalyse nicht möglich, dasselbe Komplexitätsniveau beizubehalten wie bei einer Inanspruchnahmeanalyse, weil dazu nicht alle Informationen auf regionaler Ebene bekannt bzw. vorhanden sind und weil zwischen der Erklärung der Inanspruchnahme medizinischer Leistungen und der Bestimmung des Bedarfs eine wichtige konzeptionelle Differenz besteht. Währenddem im Analysekontext das «Auffinden» von Erklärungsfaktoren im Vordergrund steht, beschäftigt im Bedarfskontext vor allem die Frage, welche Faktoren die Inanspruchnahme bestimmen «dürfen».

Mit einer Reduktion der Komplexität sind *Schwerpunktsetzungen und Werturteile* verbunden. Werden bekannte, aus der Inanspruchnahmeforschung relevante Determinanten bei der Bedarfsanalyse bzw. bei der Festsetzung eines Referenzwertes nicht berücksichtigt, so bedeutet dies eine normative Setzung, die ganz explizit dargelegt werden muss. Deutlich wird dies am Beispiel der Landesregionen. Verschiedene Analyse zeigen, dass die Inanspruchnahme in der französisch sprechenden Schweiz ceteris paribus höher ist als in der Deutschschweiz. Vermutet wird, dass diese erhöhte Inanspruchnahme mit kulturellen Faktoren im weitesten Sinn zu tun haben (vgl. Bisig / Gutzwiler 2004a: S. 119-146). Auf der Bedarfsseite muss nun entschieden werden, ob die Landesgegend bei der Bedarfsbestimmung berücksichtigt werden soll. Wird sie berücksichtigt, dann wird den französisch sprechenden Landesgegenden ceteris paribus ein höherer Bedarf zugestanden. Wird sie nicht berücksichtigt, so wird das Verhalten in der Deutschschweiz als Norm genommen. Der festgestellte Bedarf in der Romandie fällt dann tiefer aus, tiefer als die heute tatsächlich beobachtete Inanspruchnahme.

Aus der Literaturanalyse konnten vielfältige methodische Hinweise gewonnen werden, wie der Bedarf konkret abgeschätzt werden kann. Auch wurde deutlich, dass es keinen methodischen Goldstandard gibt. Vielmehr haben fast alle Vorgehensweisen sowohl Vor- wie auch Nachteile. Unabhängig davon, welches Verfahren für die Ermittlung des Bedarfs bzw. für die Erstellung von Bedarfsprognosen verwendet wird, wird es notwendig sein, ein Konzept zu erstellen, mit dem die zu *analysierenden Gebietseinheiten* gebildet werden können. Für die Bildung von solchen Gebietseinheiten stehen zwei Kriterien im Vordergrund, welche es zu beachten gibt:

Die zentrale Bedeutung der geographischen Analyseeinheiten besteht in der Annahme abgrenzbarer medizinischer Einzugsgebiete bzw. in abgrenzbaren medizinischen «Versorgungsmärkten». Dazu gehört gleichzeitig die Unterstellung, dass so genannte «cross-border- Effekte», d.h. eine «grenzüberschreitende Inanspruchnahme», soweit zu vernachlässigen sind, dass dadurch der prognostizierte regionalspezifische medizinische Leistungsbedarf nicht verzerrt wird. Wenn es nicht gelingt, medizinische Einzugsgebiete so zu definieren, dass «cross-border- Effekte» zu vernachlässigen sind, ist eine Beurteilung, ob in einer Region Über- bzw. Unterversorgung herrscht, nicht möglich.

Das zweite Kriterium betrifft das Inanspruchnahmeverhalten der Bevölkerung. Für die Analyse ist es von Vorteil, wenn innerhalb eines Versorgungsgebietes das Inanspruchnahmeverhalten der Bevölkerung in gewissem Mass als homogen betrachtet werden kann.

Ob eine Aufteilung der Schweiz nach Kantonen für die Analyse von regionalen Unterschieden im bevölkerungsseitigen Leistungsbedarf geeignet ist, scheint aus zwei Gründen eher fragwürdig. Erstens handelt es sich bei den Kantonen um keine «abgeschlossenen» medizinischen Versorgungsgebiete. Grosse Kantone sind für einen einzigen medizinischen Versorgungsmarkt zu gross und kleine wohl eher zu klein. Grenzüberschreitende Inanspruchnahme findet zwischen bestimmten Kantonen sicher statt. Zweitens weisen die empirischen Analysen zum Inanspruchnahmegeschehen darauf hin, dass innerhalb der Kantonsgebiete das Nachfrageverhalten nicht als homogen zu betrachten ist. Diese Heterogenität innerhalb der Kantonsgebiete sollte jedoch bei der Analyse berücksichtigt werden können. Beruht ein Analyseverfahren auf kantonalen Mittelwerten, läuft man deshalb Gefahr, regional-spezifische Unterschiede zu verwischen und damit nicht zu berücksichtigen.

Es spricht deshalb vieles dafür, für die Bildung von medizinischen Versorgungsgebieten neben einem Kriterium, welches die Distanz zu einem zentralen Ort berücksichtigt, auch auf kulturelle und institutionelle Grenzen Rücksicht zu nehmen. Einen solchen Raumgliederungsindikator wurde bspw. im Rahmen des Nationalen Forschungsprogramms «Regionalprobleme» gebildet. Jede Gemeinde der Schweiz kann dabei einem von 106 Raumeinheiten (MS-Regionen, Mobilité Spatiale) zugeordnet werden. Die Einheiten variieren nicht stark nach Grösse und Einwohnerzahl und entsprechen in der Regel dem Einzugsgebiet eines zentralen kleinen Ortes, berücksichtigen jedoch auch sprachliche, kantonale und teilweise auch konfessionelle Grenzen. Zudem können die MS-Regionen in vier Kategorien auf der Dimension Stadt-Land gruppiert werden: Grosszentren/Metropolen; Mittelzentren; Kleinzentren und Peripherie. Es kann daher geprüft werden, ob mit Hilfe von diesen 106 MS-Regionen medizinische Versorgungsregionen gebildet werden können.

Zurzeit laufen im Bundesamt für Gesundheit Arbeiten, die darauf abzielen, Versorgungsregionen zu bilden. Die Ergebnisse sind sicher zu berücksichtigen.

5.2 Von der Bedarfsanalyse zur Bedarfsprognose

Die Frage der Über- und Unterversorgung beinhaltet neben der *geographischen* jedoch auch eine *zeitliche Dimension*. Bedarfsanalysen befassen sich deshalb nicht nur mit dem Vergleich von aktuellen Ist-Zuständen und Soll-Werten, sie werden auch dazu verwendet, Prognosen für die Zukunft zu erstellen. Mit Hilfe von verschiedenen Entwicklungsszenarien wird so der zukünftige Bedarf – bspw. in zehn Jahren – abgeschätzt. Im Kern geht es dabei darum, anhand des prognostizierten zukünftigen Bedarfs nach medizinischen Leistungen die dafür notwendige Angebotsstruktur abzuleiten. Damit kann beurteilt werden, ob der gewünschte medizinische Versorgungsgrad in Zukunft zu garantieren ist oder ob dafür u.a. auch Massnahmen zur Gestaltung des Angebotes notwendig sind. Für die Bestimmung eines zukünftigen regionalen Bedarfs muss die Komplexität im Vergleich zur Bestimmung des aktuellen Bedarfs zwangsläufig noch einmal reduziert werden. Dies hängt mit der Schwierigkeit zusammen, dass für gewisse Einflussfaktoren im regionalspezifischen Kontext keine Entwicklungsszenarien vorliegen. Die Veränderung der Altersstruktur bspw. kann modelliert werden, wogegen eine Entwicklung des regionalen Krankheitsgeschehens oder des Inanspruchnahmeverhaltens der Bevölkerung nur schwierig vorhersehbar ist. **Tabelle 2** zeigt die im Rahmen des Projektes «Démographie Médicale» relevanten Ziele und Zwecke, welche mit Hilfe von Bedarfsanalysen bzw. -prognosen erreicht werden können.

Tabelle 2: Mögliche Ziele und Zwecke von Bedarfsanalysen bzw. -prognosen

	Bedarfsanalyse	effektive Inan- spruchnahme	Ziel	Zweck
heute	Soll-Wert	Ist-Wert	Ermittlung von Über- bzw. Unterversorgung	Beurteilung, ob und welche Steuerungsmassnahmen notwendig sind
zukünftig (bspw. in 10 Jahren)	Bedarfsprognose		Ermittlung von Ungleichgewichten zwischen zukünftigem Bedarf und zukünftigem Angebot	Gewünschtes medizinisches Versorgungsniveau durch Massnahmen auf der Angebots- und Nachfrageseite «garantieren»

Quelle: Eigene Darstellung

6 Die Bedeutung der Literaturanalyse

Aus der Literaturanalyse der Inanspruchnahme von ambulanten Leistungen sowie der Angebotsinduzierung von Ärzten und Ärztinnen lassen sich zusammenfassend auf verschiedenen Ebenen Folgerungen für die Nachfrageseite des Projektes «Démographie Médicale» ziehen.

Erstens wurde deutlich, dass zwischen der Erklärung der Inanspruchnahme medizinischer Leistungen und der Bestimmung des Bedarfs eine wichtige konzeptionelle Differenz besteht. Währenddem im Analysekontext das «Auffinden» von Erklärungsfaktoren im Vordergrund steht, beschäftigt im Bedarfskontext vor allem die Frage, welche Faktoren die Inanspruchnahme bestimmen «dürfen». Am Beispiel des «Einflusses» der Landesregionen konnte dies verdeutlicht werden.

Zweitens hat die Literaturanalyse gezeigt, dass für die Beantwortung der zentralen Fragestellung nach der regionalen Über- bzw. Unterversorgung eine Bedarfsanalyse notwendig ist. Diese kann unterschiedlich komplex ausfallen.

Drittens wurde deutlich, auch wenn die Literaturanalyse nur rudimentär durchgeführt werden konnte, dass der Prozess der Inanspruchnahme komplex ist. Dies bedeutet für die Bestimmung des Bedarfs an medizinischen Leistungen, dass diese Komplexität reduziert werden muss. Damit sind Schwerpunktsetzungen und Werturteile verbunden.

Viertens konnte eine Übersicht über die Faktoren gewonnen werden, welche die Inanspruchnahme beeinflussen. Diese Faktoren sind bei der Abschätzung des regionalen Bedarfs in der einen oder anderen Form zu berücksichtigen. Werden ausgewählte Faktoren nicht berücksichtigt, so bedeutet dies auch an dieser Stelle eine normative Setzung, die ganz explizit dargelegt werden muss.

Fünftens konnten vielfältige methodische Hinweise gewonnen werden, wie der Bedarf konkret abgeschätzt werden kann. Auch wurde deutlich, dass es keinen methodischen Goldstandard gibt. Vielmehr haben fast alle Vorgehensweisen sowohl Vor- wie auch Nachteile. Im zweiten Teil des Berichtes werden wir daher folgerichtig mehrere Varianten vorschlagen.

Sechstens konnten vielfältige Einsichten im Bereich der Angebotsinduzierung gewonnen werden. Auf konzeptioneller Ebene wurde deutlich, dass Angebotsinduzierung nicht nur unerwünscht sein kann und dass daneben auch eine Patienteninduzierung existiert. Es zeigte sich auch, dass die Angebotsinduzierung nur relativ zu einem normativen Referenzwert, der sog. «richtigen» oder «bedarfsgerechten» Versorgung, festgestellt werden kann. Methodisch konnte weiter gezeigt werden, dass die einwandfreie Feststellung der Existenz und des Ausmasses der Angebotsinduzierung sehr anspruchsvoll ist und in der Schweiz bisher noch nie à-fonds durchgeführt worden ist. Sollte im Gesamtprojekt der «Démographie Médicale» daher die Absicht bestehen, den Bedarf korrigiert um die Effekte der Angebotsinduzierung festlegen zu können, müssten dazu umfangreiche Analysen vorangehen.

Siebtens zeigte sich durch die Analyse der Anreize von Wahlfranchisen, dass nicht nur die Vergütungsanreize auf der Angebotsseite verhaltensrelevant sind, sondern dass auch die Nachfrage auf Anreize reagiert. Sollte bei der Feststellung des Bedarfs die Angebotsinduzierung korrigiert werden, so müsste man sich auch die Frage stellen, ob nicht auch eine Nachfrageinduzierung berücksichtigt werden müsste. Dies könnte dann bspw. bedeuten, dass bei der Feststellung des «legitimen» Bedarfs vom Verhalten der Versicherten in der höchsten Franchisestufe ausgegangen wird. Ganz offensichtlich beinhaltet dieses Vorgehen die Setzung von Werturteilen.

Teil 2: Konkretes Vorgehen zur Bestimmung des Bedarfs

Im zweiten Teil dieses Berichts werden, aufbauend auf den Erkenntnissen der Literaturanalyse, verschiedene Verfahren vorgestellt, wie mit Hilfe der Ermittlung eines regionenspezifischen Bedarfs Fragen der Über- bzw. Unterversorgung und des zukünftigen Bedarfs beantwortet werden können.

7 Drei Konzepte und Wege zur Bestimmung des Bedarfs

Es wurde darauf hingewiesen, dass aufgrund des Kontextes, in dem eine Bedarfsanalyse stattfindet, die Komplexität im Vergleich zu Inanspruchnahmeanalysen reduziert werden muss. Die Antwort auf die Frage, welche Faktoren die Inanspruchnahme bestimmen «dürfen», wird durch eine normative Setzung gegeben. Sie muss als erstes beantwortet werden. Je nachdem, wie viele Faktoren für die Festlegung eines regionalen Leistungsbedarfs berücksichtigt werden sollen, müssen entsprechende Verfahren angewandt werden. Grundsätzlich stehen drei Varianten zur Verfügung:

Mit dem ersten, relativ einfachen Verfahren wird es möglich sein, den Einfluss von wenigen «Schlüsseldeterminanten» auf die Nachfrage/Inanspruchnahme zu berücksichtigen und diese in die Bedarfsberechnungen mit einzubeziehen. Dazu gehören mindestens die bevölkerungsstrukturellen Variablen «Alter» und «Geschlecht». Es beruht auf der Überlegung, dass anhand von Versichererdaten für jede Person gemäss ausgewählten Merkmalen ein durchschnittlicher Leistungsbedarf pro Jahr berechnet wird. Damit kann der Gesamtbedarf für jedes beliebig abgrenzbare Gebiet, für welches die erforderlichen Daten vorhanden sind, berechnet werden.

Mit dem zweiten, etwas aufwändigeren Verfahren wird es möglich sein, mehr Einflussfaktoren zu berücksichtigen, als mit dem ersten. Grundsätzlich ist es mit diesem Verfahren auch möglich, Angebots- und Nachfrageinduzierung zu berücksichtigen. Dazu müssten jedoch vorgängig zusätzliche Analysen, welche über die Existenz und das Ausmass von Angebotsinduzierung Auskunft geben, gemacht werden. Es handelt sich bei diesem Verfahren um ein multivariates Analyseverfahren mit aggregierten Daten auf Regionenebene.

Das dritte Verfahren ist, im Gegensatz zu den ersten zwei, eine Mehrebenenanalyse (Kombinierung von Individual- und Aggregatebene). Ausgehend von deren Resultaten können regionale Bedarfsprofile berechnet werden. Mit diesem Verfahren könnte gleichzeitig die Forschungslücke, welche in Bezug auf die Existenz und das Ausmass von Angebotsinduzierung besteht, geschlossen werden. Es handelt sich dabei um das aufwändigste und methodisch anspruchsvollste Verfahren. Zudem stellt es im Vergleich zu den anderen Verfahren die höchsten Anforderungen an die Daten.

Welches der drei vorgeschlagenen Verfahren angewandt werden soll, hängt u.a. von der präzisen Fragestellung und von der Gewichtung verschiedener Aspekte ab. Es wurde schon erwähnt, dass es keinen methodischen Goldstandard gibt. Vielmehr haben fast alle Vorgehensweisen sowohl Vor- wie auch Nachteile. Zur Beurteilung der Vor- und Nachteile müssen mindesten vier verschiedene Aspekte berücksichtigt werden.

Auf der *normativen Ebene* muss darüber entschieden werden, welche Faktoren für die Ermittlung des «richtigen» Bedarfs berücksichtigt werden sollen (bspw. kulturelle Faktoren)? Unseres Erachtens ist dies der wichtigste, aber auch der anspruchsvollste Aspekt, der geklärt werden muss. Je mehr Faktoren eine Rolle spielen «dürfen», um so eher muss eine multivariate Methode angewandt werden (Variante 2 oder 3).

Inhaltlich muss klar sein, ob versucht werden soll, den ermittelten Bedarf um den Anteil, der durch die Angebotsinduzierung ausgelöst wurde, zu korrigieren. Damit dies möglich ist, muss zuerst eine ver-

tiefende Analyse zur Klärung der Existenz und des Ausmasses von Angebotsinduzierung in der Schweiz durchgeführt werden. Erst danach kann entschieden werden, welches Verfahren angewandt werden soll.

Auf der *praktischen Seite* müssen für die verschiedenen Methoden unterschiedlich detaillierte Daten vorhanden sein. Je komplexer die Methode, umso mehr Anforderungen sowohl in quantitativer als auch qualitativer Sicht werden an die Daten gestellt. Ein anderer Aspekt eher praktischer Art hängt mit dem unterschiedlichen Aufwand der verschiedenen Methoden zusammen. Wenn komplexere Verfahren gewählt werden, muss die Bereitschaft vorhanden sein, mehr Zeit und Aufwand in das Forschungsvorhaben zu investieren.

Verwendungszweck der Bedarfsanalysen bzw. Zielpublikum: Je komplexer und anspruchsvoller die gewählte Methode, umso geringer ist für «Aussenstehende» die Nachvollziehbarkeit der Resultate und deshalb auch deren Transparenz. Methodisch korrekt hergeleitete Resultate können, wenn die Methode zu kompliziert und schwer verständlich ist, auf weniger Akzeptanz stossen als u.U. etwas ungenauere Resultate, welche jedoch mit einfacheren, verständlicheren Verfahren produziert werden.

Die Diskussion der nun vorgestellten Varianten richtet sich nach diesen vier Aspekten.

7.1 Pro-Kopf-Bedarf von verschiedenen Inanspruchnahmegruppen (Variante 1)

Mit dem ersten, relativ einfachen Verfahren ist es möglich, den Einfluss von wenigen «Schlüssel-determinanten» auf die Nachfrage/Inanspruchnahme zu berücksichtigen und diese in die Bedarfsberechnungen mit einzubeziehen.

Grundidee

Das Verfahren beruht auf der Überlegung, dass anhand von Inanspruchnahmedaten für jede Person mit bestimmten Merkmalen (bspw. Alter und Geschlecht) ein auf *schweizerischen Durchschnittswerten* beruhender, *standardisierter Leistungsbedarf* (SLB_y)²⁹ pro Jahr berechnet wird. Dieser gibt an, wie viele gesundheitliche Leistungen eine Person mit bestimmten Merkmalen durchschnittlich in Anspruch nimmt bzw. genommen hat. Mit den zwei Variablen Alter und Geschlecht würde bspw. sowohl für Männer als auch für Frauen für jedes erreichte Altersjahr ein für die Schweiz durchschnittlicher Leistungsbedarf berechnet. In einem zweiten Schritt wird der standardisierte Leistungsbedarf dazu verwendet, für jedes medizinische Versorgungsgebiet (VG_x) einen regionalen Leistungsbedarf (RLB_x) zu ermitteln. Dazu wird pro medizinisches Versorgungsgebiet der standardisierte Leistungsbedarf mit der Anzahl der entsprechenden in dieser Versorgungsregion lebenden Personen multipliziert. Werden diese Werte aufsummiert und durch die Anzahl der in einer Versorgungsregion lebenden Personen dividiert, erhält man für jede Versorgungsregion einen durchschnittlichen, regionalen Pro-Kopf-Bedarf, welcher mit der effektiven Pro-Kopf-Inanspruchnahme (realisierte Pro-Kopf-Nachfrage) verglichen werden kann. Wir sprechen von den so ermittelten Werten im Weiteren von regionalspezifischen Bedarfskennziffern (RB_x).

Grundsätzlich ist es möglich, mehrere Determinanten der Inanspruchnahme zu berücksichtigen. Die Anzahl Merkmale ist jedoch begrenzt, da von der Methode her für jede Merkmalskombination ein entsprechender standardisierter Leistungsbedarf berechnet werden muss (vgl. dazu das Beispiel in **Tabelle 3**).

²⁹ Der Index «y» steht dabei für eine Kombination von wenigen, noch auszuwählenden «Schlüssel-determinanten» der Inanspruchnahme.

Tabelle 3: Beispiel zur Berechnung des standardisierten und regionalen Leistungsbedarfs unter Berücksichtigung von Alter und Geschlecht

Beispiel Leistungsbedarf Schweiz unter Berücksichtigung von Alter und Geschlecht

Alter	Bevölkerung in Tausend			Ist-Wert Inanspruchnahme Schweiz in 100'000 CHF			Standardisierter Pro-Kopf-Leistungsbedarf Schweiz in CHF		
	Männer	Frauen	Gesamt	Männer	Frauen	Gesamt	SLB _(CH) = Inanspruchnahme/Anzahl		
							Männer	Frauen	Gesamt
0-9	380	360	740	2'964	2'376	5'340	780	660	722
10-19	450	420	870	3'240	3'780	7'020	720	900	807
20-29	450	450	900	3'780	7'290	11'070	840	1'620	1'230
30-39	600	600	1'200	6'480	11'880	18'360	1'080	1'980	1'530
40-49	550	540	1'090	8'580	11'016	19'596	1'560	2'040	1'798
50-59	480	480	960	11'520	13'824	25'344	2'400	2'880	2'640
60-69	330	360	690	13'464	14'040	27'504	4'080	3'900	3'986
70-79	220	300	520	12'936	14'400	27'336	5'880	4'800	5'257
80-89	90	170	260	7'560	16'320	23'880	8'400	9'600	9'185
>90	20	40	60	2'400	6'000	8'400	12'000	15'000	14'000
Total	3'570	3'720	7'290	72'924	100'926	173'850	2'043	2'713	2'385

Beispiel Regionaler Leistungsbedarf Versorgungsregion A

Alter	Bevölkerung in Tausend			Ist-Wert Inanspruchnahme Versorgungsregion A in 100'000 CHF			Regionaler Leistungsbedarf Versorgungsregion A (Soll-Wert) in 100'000 CHF		
	Männer	Frauen	Gesamt	Männer	Frauen	Gesamt	RLB(A) = Anzahl (A) * SLB(CH)		
							Männer	Frauen	Gesamt
0-9	43	53	96	290	330	620	332	351	692
10-19	49	60	108	300	510	810	351	537	875
20-29	45	59	103	320	920	1'240	374	949	1'268
30-39	60	79	139	580	1'500	2'080	649	1'565	2'129
40-49	52	67	119	740	1'200	1'940	813	1'363	2'138
50-59	69	92	161	1'550	2'656	4'206	1'658	2'656	4'259
60-69	49	71	120	1'800	2'755	4'555	2'012	2'755	4'781
70-79	34	58	93	1'900	2'780	4'680	2'016	2'802	4'871
80-89	15	32	47	1'100	3'000	4'100	1'242	3'101	4'325
>90	4	9	14	420	1'349	1'769	534	1'370	1'903
Total	420	580	1'000	9'000	17'000	26'000	9'982	17'450	27'240
Pro-Kopf Inanspruchnahme bzw. Bedarf Versorgungsregion A				2'143	2'931	2'600	2'377	3'009	2'724
RLB_(A) = Total A/Anzahl Personen A									

Beispiel Regionaler Leistungsbedarf Versorgungsregion B

Alter	Bevölkerung in Tausend			Ist-Wert Inanspruchnahme Versorgungsregion B in 100'000 CHF			Regionaler Leistungsbedarf Versorgungsregion B (Soll-Wert) in 100'000 CHF		
	Männer	Frauen	Gesamt	Männer	Frauen	Gesamt	RLB(B) = Anzahl (B) * SLB(CH)		
							Männer	Frauen	Gesamt
0-9	87	73	160	802	601	1'403	678	484	1'156
10-19	102	85	187	846	854	1'700	736	764	1'510
20-29	102	91	193	923	1'469	2'392	859	1'469	2'373
30-39	135	120	255	1'612	2'567	4'179	1'458	2'371	3'897
40-49	124	108	232	2'216	2'403	4'619	1'935	2'206	4'174
50-59	101	89	190	2'746	2'698	5'444	2'423	2'572	5'023
60-69	68	66	134	2'906	2'611	5'517	2'783	2'577	5'352
70-79	44	54	99	2'837	2'715	5'552	2'597	2'614	5'185
80-89	16	29	45	1'538	2'845	4'383	1'324	2'813	4'139
>90	0.5	4	5	74	737	811	56	621	646
Total	780	720	1'500	16'500	19'500	36'000	14'850	18'491	33'456
Pro-Kopf Inanspruchnahme bzw. Bedarf Versorgungsregion B				1'904	2'568	2'400	1'904	2'568	2'230
RLB_(B) = Total B/Anzahl Personen B									

Quelle: Eigene Berechnungen

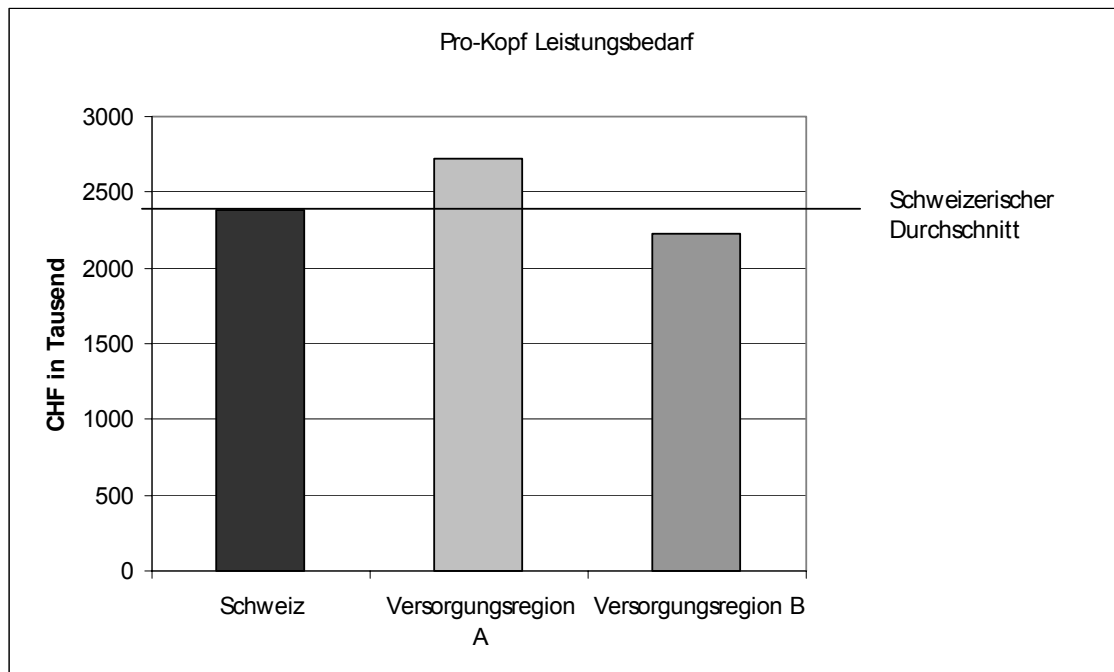
Die regionalspezifischen Bedarfskennziffern können für zweierlei Dinge benutzt werden:

Erstens können sie mit dem für die Schweiz durchschnittlichen Leistungsbedarf (standardisierter Pro-Kopf-Leistungsbedarf) verglichen werden. Im Beispiel in **Tabelle 3** hat Versorgungsregion A aufgrund der Verteilung von Alter und Geschlecht einen höheren und Versorgungsregion B einen tieferen

Leistungsbedarf als der Schweizer Durchschnitt. Die Abweichung vom durchschnittlichen Leistungsbedarf gibt demnach Auskunft darüber, ob für das entsprechende Versorgungsgebiet aufgrund der in die Analyse miteinbezogenen Merkmale ein höherer bzw. ein tieferer Bedarf nach Gesundheitsleistungen besteht.

Abbildung 4 zeigt dies für das erwähnte Beispiel graphisch.

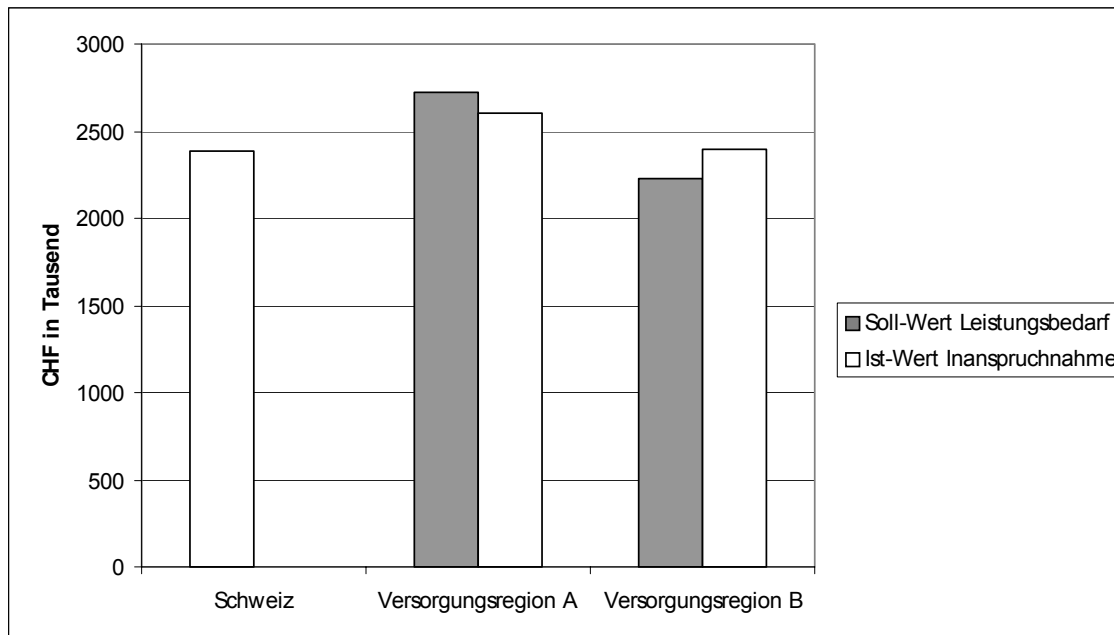
Abbildung 4: Vergleich Schweizer und regionaler Leistungsbedarf gemäss Beispiel **Tabelle 3**



Quelle: Eigene Darstellung

Zweitens kann die empirisch ermittelte regionalspezifische Bedarfskennziffer mit der entsprechenden «wahren» Pro-Kopf-Inanspruchnahme verglichen werden. Die Höhe der Abweichung von den «wahren», beobachteten Werten lässt erkennen, inwiefern noch andere, nicht berücksichtigte Determinanten auf die Inanspruchnahme von Gesundheitsleistungen innerhalb der entsprechenden Versorgungsregion einwirken. Damit können u.a. auch Rückschlüsse auf Über- bzw. Unterversorgung gezogen werden. Dies ist jedoch nur sehr bedingt möglich, da mit der vorgeschlagenen Variante 1 nur wenige der aus der Inanspruchnahmeforschung bekannten Determinanten berücksichtigt werden können. Im vorgestellten Beispiel ist in der Versorgungsregion A die tatsächliche Pro-Kopf-Inanspruchnahme (Ist-Wert) mit CHF 2'600.- geringer, als dies aufgrund der Alters- und Geschlechterverteilung erwartet werden könnte (Soll-Wert = 2'724.-). In der Versorgungsregion B ist es genau umgekehrt. Die effektive Pro-Kopf-Inanspruchnahme (CHF 2'400.-) ist höher als der prognostizierte Pro-Kopf-Bedarf (CHF 2'230.-). Warum diese Unterschiede bestehen, kann mit der vorgeschlagenen Methode nicht beantwortet werden. **Abbildung 5** zeigt für das verwendete Beispiel den Vergleich in graphischer Form.

Abbildung 5: Vergleich Soll- und Ist-Wert des regionalen Bedarfs bzw. Inanspruchnahme gemäss Beispiel **Tabelle 3**



Quelle: Eigene Darstellung

Neben der einfachen Berechnung des regionalen Bedarfs eignet sich die Methode gut dazu, Bedarfsprognosen für die Zukunft zu erstellen. Die zur Bestimmung des standardisierten Leistungsbedarfs verwendeten Merkmale können mit Hilfe von Entwicklungsszenarien für ein bestimmtes Jahr neu berechnet werden. Der so berechnete zukünftige, standardisierte Leistungsbedarf bildet dabei die Veränderungen im Bedarf ab, welche sich durch die Veränderung in den berücksichtigten Merkmalen unter Konstanthaltung aller anderen Merkmale ergeben.

Möglichkeiten zur «Korrektur» des standardisierten Leistungsbedarfs

Die Literaturanalyse hat gezeigt, dass Bedarfsanalysen ohne die Berücksichtigung der bevölkerungsstrukturellen Merkmale «Alter» und «Geschlecht» auf wenig Akzeptanz stossen, weshalb zumindest diese Faktoren in die Berechnungen einfließen sollten. Zudem gibt es genügend empirische Evidenz, die zeigt, dass diese zwei Faktoren gute Prädiktoren für die Inanspruchnahme von Gesundheitsleistungen sind. Bestehen zwischen den Versorgungsregionen bevölkerungsstrukturelle Unterschiede in Bezug auf Alter und Geschlecht, werden diese denn auch in einem unterschiedlich hohen regionalen Bedarf abgebildet.

Es stellt sich jedoch die Frage, ob für die Berechnung des standardisierten Leistungsbedarfs noch weitere Merkmale mitberücksichtigt werden sollten. Für eine bessere und unbestrittenere Bedarfsbestimmung bedürfte es noch zusätzliche Informationen über den *allgemeinen Gesundheitszustand* der Bevölkerung innerhalb eines medizinischen Versorgungsgebiets. Der Gesundheitszustand ist gemäss den empirischen Resultaten einer der stärksten und besten Prädiktoren für die Höhe der Inanspruchnahme von gesundheitlichen Leistungen. Mit Informationen zum Gesundheitszustand würde der ermittelte regionale Bedarf sicher präziser ausfallen. Auf regionaler Versorgungsebene dürften jedoch repräsentative und verlässliche Daten fehlen. Es ist deshalb zu überlegen, ob es alternative Indikato-

ren gibt, welche Rückschlüsse auf den Gesundheitszustand der Bevölkerung zulassen. Dafür gibt es verschiedene Möglichkeiten, u.a.:

Berechnung des SLB für Personen mit bzw. ohne stationäre Aufenthalte unter Berücksichtigung des «Alters» und des «Geschlechts»³⁰.

Berechnung des SLB unter Berücksichtigung der Franchisewahl, des «Alters» und des «Geschlechts». Anhand der empirischen Resultate zu den Effekten der Wahlmöglichkeit von Selbstbeteiligungsmodellen wissen wir, dass die individuelle Wahl der Franchisenhöhe nicht unabhängig vom jeweiligen Gesundheitszustand erfolgt.

Diskussionswürdig für die Berücksichtigung zur «Korrektur» des regionalen Leistungsbedarfs sind Variablen, welche in irgendeiner Form in Bezug zu *kulturell begründeten Verhaltensmustern* stehen. Dazu gehören bspw. die Landesregionen aber auch die Dimension Stadt-Land. Aus der Inanspruchnahmeforschung weiss man, dass die Bevölkerung in der Französisch sprechenden Schweiz mehr Leistungen in Anspruch nimmt als diejenige der Deutsch sprechenden Schweiz. Auch die städtische Bevölkerung nimmt mehr Leistungen in Anspruch als die ländliche. Sollen diese Unterschiede in die Ermittlung des standardisierten Leistungsbedarfs einfließen oder nicht? Die vorgeschlagene Methode ist in der Lage, solche Unterschiede zu berücksichtigen. Ob man dies will, bedarf einer eingehenden Diskussion. Dabei sind neben dem normativen Aspekt auch Fragen der Akzeptanz bzw. der (politischen) Durchführbarkeit zu berücksichtigen.

Datengrundlagen

Als Datengrundlage für die Ermittlung des regionalen Leistungsbedarfs können die Daten des *santé-suisse*-Datenpools benutzt werden. Die Daten sollten dazu auf der Ebene der Gemeinden aufbereitet werden. Wünschbar wäre, wenn die Inanspruchnahmedaten sowohl in monetären als auch in Mengeneinheiten geliefert werden könnten.

Für die Erstellung von Bedarfsprognosen kann in Bezug auf die Entwicklung der bevölkerungsstrukturellen Merkmale «Alter» und «Geschlecht» auf die vom Bundesamt für Statistik (2001) herausgegebenen «Szenarien zur Bevölkerungsentwicklung der Schweiz 1991–2060» zurückgegriffen werden. Sollen neben dem «Alter» und dem «Geschlecht» noch weitere Merkmale für Bedarfsprognosen berücksichtigt werden, müssen eigene Szenarien entwickelt werden.

Verfahrensschritte

Abbildung 6 zeigt die einzelnen Verfahrensschritte mit Hilfe einer graphischen Darstellung.

³⁰ Es wird davon ausgegangen, dass die stationären Kosten ein guter Indikator für den (unbeobachtbaren, bzw. nicht bekannten) Gesundheitszustand sind und nur marginal durch Nachfrage- bzw. Angebotsinduzierung beeinflusst wird (vgl. dazu Gardiol et al. 2003).

Abbildung 6: Verfahrensschritte zur Erstellung von regionalen Bedarfsprofilen für Gesundheitsleistungen. Variante 1



Quelle: Eigene Darstellung

Vor- und Nachteile

Der Vorteil des nun skizzierten Vorgehens zur Bestimmung eines regionenspezifischen Bedarfs liegt darin, dass es sich um eine einfache Methode handelt, welche problemlos umgesetzt werden kann. Der Nachteil liegt darin, dass damit weder nachfrage- noch angebotsinduzierende Effekte berücksichtigt werden können, weil die Gesamtheit der in Anspruch genommenen Leistungen als Referenzgrösse dient. Die Gesamtheit der in der Schweiz in Anspruch genommenen Leistungen wird dabei mit Hilfe eines «Verteilschlüssels» wieder auf die Versorgungsregionen verteilt. Weil die Gesamtheit der in Anspruch genommenen Leistungen als Referenzgrösse dient, richtet sich auch die Bedarfsprognose nach dem aktuellen Leistungsbedarf.

7.2 Multivariate Regression auf der Ebene medizinischer «Versorgungsgebiete» (Variante 2)

Mit dem zweiten, etwas aufwändigeren Verfahren ist es möglich, mehr Einflussfaktoren der Inanspruchnahme zu berücksichtigen als mit dem ersten. Grundsätzlich ist es mit diesem Verfahren auch möglich, Angebots- und Nachfrageinduzierung zu berücksichtigen. Es handelt sich bei diesem Verfahren um ein multivariates Analyseverfahren mit aggregierten Daten auf der Ebene von Versorgungsgebieten.

Grundidee

Ähnlich wie in Variante 1 ist bei dieser zweiten Variante zur Bestimmung des regionalspezifischen Leistungsbedarfs die medizinische Versorgungsregion. Anstatt anhand der Inanspruchnahmedaten einen standardisierten Leistungsbedarf zu ermitteln und diesen dann auf die Versorgungsregionen «umzurechnen», wird der regionalspezifische Leistungsbedarf mit Hilfe einer multivariaten Regressionsanalyse auf der Ebene der Versorgungsregionen direkt geschätzt. Die Prognose wird damit genauer ausfallen als in Variante 1, da mit der Methode der multivariaten Regression mehr Determinanten der Inanspruchnahme berücksichtigt werden können. Es können dabei sowohl angebotsseitige (bspw. ÄrztInnendichte) als auch nachfrageseitige Determinanten (Alter, Geschlecht, Stadt-Land, etc.) berücksichtigt werden.

Um die Frage der Angebotsinduzierung für die Ermittlung des regionalen Bedarfs zu berücksichtigen, müssten jedoch vorgängig zur Regressionsanalyse auf der Ebene der medizinischen «Versorgungsgebiete» mit einem Individualdatensatz (bspw. der Gesundheitsbefragung) Analysen zur Bestimmung des Ausmasses gemacht werden.

Es können jedoch auch Informationen über die Existenz von Angebotsinduzierung mit Hilfe der Regressionsanalyse auf der Ebene der medizinischen «Versorgungsgebiete» bekommen werden. Wenn davon ausgegangen wird, dass Angebotsinduzierung erst ab einem gewissen Niveau der ÄrztInnendichte zu einem Thema wird, können mit Hilfe der vorgeschlagenen Methode Hinweise gesucht werden, welche diese Hypothese stützen. Dazu werden die Versorgungsregionen anhand ihrer ÄrztInnendichte in zwei Gruppen eingeteilt: Eine Gruppe wird mit denjenigen Versorgungsregionen gebildet, welche eher eine hohe ÄrztInnendichte aufweisen und eine zweite beinhaltet diejenigen mit den tiefen. Für die beiden Gruppen wird die Regressionsschätzung separat durchgeführt. Zeigt sich, dass die ÄrztInnendichte bei der Gruppe der Versorgungsregionen mit tiefer ÄrztInnendichte keinen Einfluss und in derjenigen mit hoher ÄrztInnendichte einen positiven Einfluss besitzt, kann dies als Hinweis auf Angebotsinduzierung in Versorgungsregionen mit hoher ÄrztInnendichte interpretiert werden.

Die zu prognostizierende, abhängige Variable ist bei diesem Verfahren der durchschnittliche Pro-Kopf-Bezug von Gesundheitsleistungen pro Versorgungsgebiet. Diese kann, wie schon erwähnt, nach Leistungsarten unterschieden werden. Methodisch kann ein ähnliches Verfahren verwendet werden, wie es Crivelli et al. (2004) für die Analyse der kantonalen Unterschiede der Pro-Kopf-Gesundheitskosten verwendet haben. Im Unterschied zu Crivelli et al. müssten jedoch aufgrund der spezifischen Fragestellung des Gesamtprojektes der «Démographie Médicale» unterschiedliche Analyseeinheiten, d.h. nicht Kantone, sondern medizinische Versorgungsgebiete, gewählt werden.

Die Resultate können zudem dazu verwendet werden, die Effekte der einzelnen Determinanten auf die Inanspruchnahme innerhalb der Versorgungsregionen zu ermitteln. Zur Beurteilung der Frage nach einem erhöhten Bedarf in einer bestimmten Region kann der ermittelte Pro-Kopf-Soll-Wert der

Region mit dem durchschnittlichen Pro-Kopf-Soll-Wert der Schweiz verglichen werden. Zur Beurteilung, ob weitere, durch die Regression nicht berücksichtigte Determinanten der Inanspruchnahme in bestimmten Regionen vorhanden sind, können die regionalen Soll-Werte mit den entsprechenden regionalen Ist-Werten verglichen werden. In einem letzten Schritt ist es möglich, den regionalen Soll-Wert mit dem Angebot in Beziehung zu bringen. Daraus sind Rückschlüsse auf Über- bzw. Unterversorgung möglich.

Der Vergleich des Pro-Kopf-Bedarfs mit dem entsprechenden effektiven Pro-Kopf-Bezug (analog zum vierten Schritt von Variante 1) gibt Hinweise über den regionalen Versorgungsgrad.

Das Erstellen von Bedarfsprognosen für die Zukunft mit Hilfe dieser Methode ist schwieriger als mit demjenigen von Variante 1. Je mehr Variablen berücksichtigt werden, umso eher sind keine validen Informationen über die Entwicklung von bestimmten Variablen vorhanden.

Möglichkeiten zur «Korrektur» des geschätzten regionalen Leistungsbedarfs

Grundsätzlich ist es «technisch» möglich, sehr viele Determinanten der Inanspruchnahme in einem multivariaten Regressionsmodell zu berücksichtigen. Dies hängt vor allem von den verfügbaren Daten sowie deren Qualität ab.

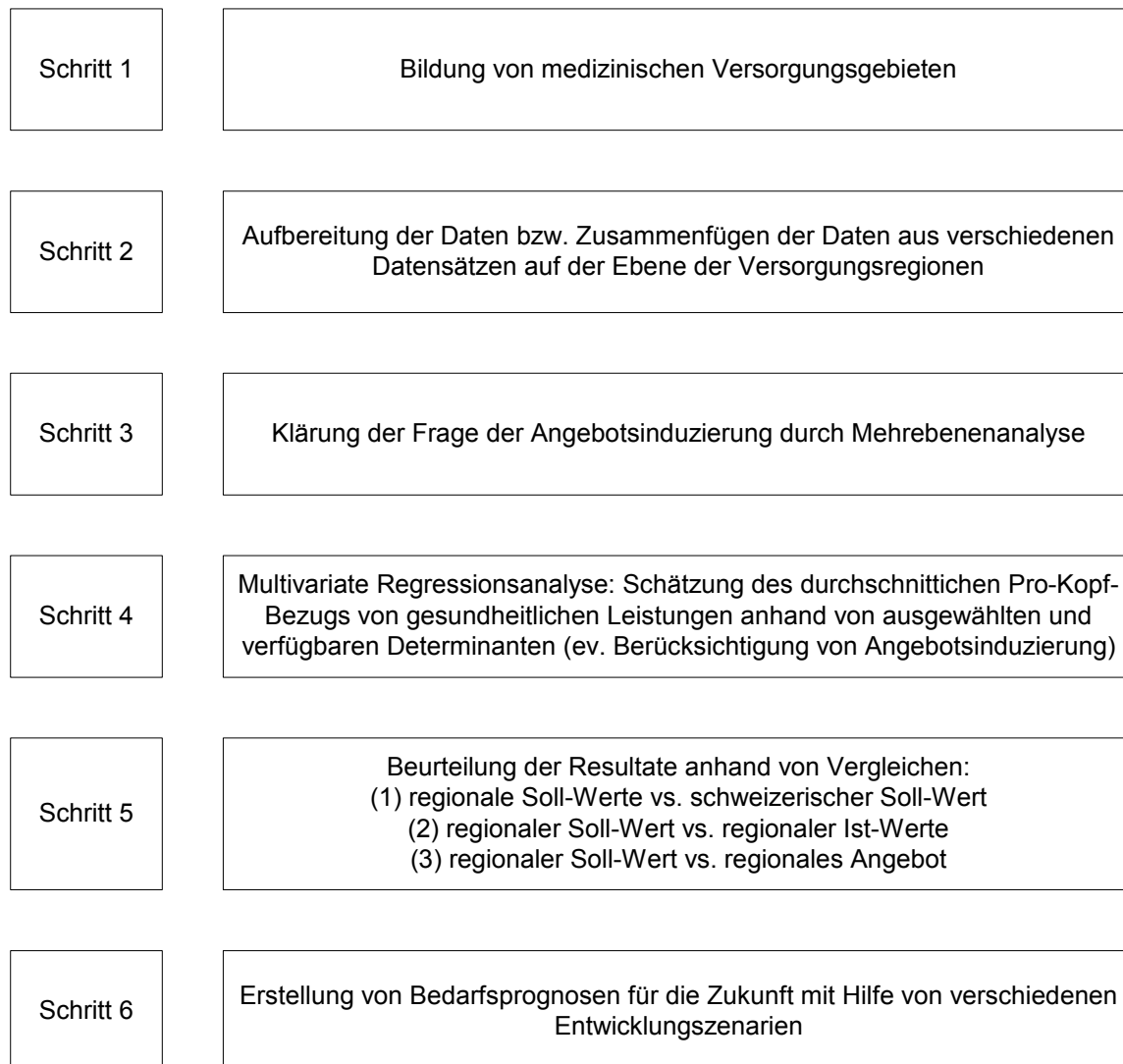
Datengrundlagen

Voraussetzung für dieses Verfahren ist das Vorhandensein von Daten auf der Ebene der Analyseinheit der medizinischen Versorgungsgebiete. Für alle Determinanten, welche in der Schätzgleichung verwendet werden, müssen deshalb die entsprechenden Daten für jedes Versorgungsgebiet vorhanden sein. Dazu müssten verschiedene Datenquellen herangezogen und zu einem Datensatz verknüpft werden. Je mehr Determinanten berücksichtigt werden sollen, umso mehr Datenquellen muss zurückgegriffen werden. Ein grosser Teil von möglichen Variablen ist im Datenpool von santésuisse vorhanden. Sollen sozio-ökonomische Merkmale wie Bildung, Einkommen, Sozialstatus etc. mit berücksichtigt werden, muss auf andere Datenquellen zurückgegriffen werden (Schweizerische Gesundheitsbefragung, SAKE, evtl. Volkszählung 2000). Es wäre von Vorteil, wenn die Daten, welche für die Analyse verwendet werden, für mehrere Jahre aufbereitet werden könnten, weil damit validere Resultate erzielt werden können.

Verfahrensschritte

Abbildung 7 zeigt anhand einer graphischen Darstellung die einzelnen Verfahrensschritte von Variante 2.

Abbildung 7: Verfahrensschritte zur Erstellung von regionalen Bedarfsprognosen für Gesundheitsleistungen. Variante 2



Quelle: Eigene Darstellung

Vor- und Nachteile

Der Vorteil gegenüber Variante 1 besteht darin, dass mehr Determinanten der Inanspruchnahme simultan berücksichtigt werden können. Damit sollte die Prognose exakter werden. Insgesamt muss mit diesem Verfahren jedoch auch wie Variante 1 auf verhaltenswissenschaftliche Begründungen verzichtet werden, weil für die Analyse keine Daten auf der Mikroebene verwendet werden. Es stellt sich jedoch die Frage, ob verhaltenswissenschaftliche Begründungen in einer Bedarfsanalyse überhaupt berücksichtigt werden sollen oder nicht. Dies ist eine normative Setzung.

Die Nachteile dieser Variante im Vergleich zur ersten sind folgende:

Es werden höhere Anforderungen an die Daten gestellt

Der Berechnungsaufwand ist wesentlich grösser

Es handelt sich um ein anspruchvolleres Verfahren, womit die Nachvollziehbarkeit geringer wird
Prognosen zu erstellen ist mit diesem Verfahren schwieriger

7.3 Mehrebenenanalyse mit Individualdaten als Basis (Variante 3)

Die Literaturanalyse hat gezeigt, dass Mehrebenenanalysen am besten geeignet sind, das Inanspruchnahmegeschehen abzubilden, da sie das höchste Komplexitätsniveau enthalten. Mit dieser Art der Analyse kann deshalb auch der Bedarf am besten ermittelt werden. Es ist jedoch auch die anspruchsvollste Methode, was den Aufwand und den Anspruch an die Daten betrifft.

Grundidee

Die Grundidee einer Mehrebenenanalyse besteht darin, dass individuelles (Inanspruchnahme-)Verhalten, welches mit Hilfe der Individualdaten modelliert wird, von gesellschaftlichen «Kontextbedingungen» mitbeeinflusst wird. Deshalb muss bei einer Mehrebenenanalyse im Gegensatz zur Aggregatsanalyse nicht auf verhaltenswissenschaftliche Begründungen, die sowohl ökonomische, soziologische, psychologische und auch morbiditätsbezogene Kontexte berücksichtigen, verzichtet werden. Damit wird auch eine vertiefte Überprüfung der Existenz der Angebotsinduzierung in der Schweiz und einer empirisch ermittelten Bestimmung einer Schwelle, ab der Angebotsinduzierung ein Thema ist, möglich (vgl. *Abbildung 1: Modell des Zusammenhangs zwischen Versorgungsdichte (Allgemein- und FachärztInnen) und ambulanter Inanspruchnahme*», S. 28).

Als erklärende Merkmale für den Pro-Kopf-Bezug von gesundheitlichen Leistungen müssen sowohl angebotsseitige als auch nachfrageseitige Determinanten berücksichtigt werden. Das Modell von Andersen kann dabei als Referenz verwendet werden. Die im Unterabschnitt 4.2.1 vorgestellten Resultate des Forschungsprojektes von Brigitte Bisig et al. (2004a, 2004b) beruhen auf einer Mehrebenenanalyse. Die AutorInnen können jedoch die angebotsseitige Determinante «ÄrztInnendichte» zur Überprüfung der Hypothese der Angebotsinduzierung nicht verwenden, da sie nur auf Kantonsbasis vorhanden ist und die Daten der Schweizerischen Gesundheitsbefragung (SGB) nicht für alle Kantone repräsentativ sind. Sollte die «ÄrztInnendichte» auf einem anderen Aggregationsniveau zur Verfügung stehen, kann aufbauend auf den Resultaten von Brigitte Bisig et al. (2004a, 2004b) der Einfluss der «ÄrztInnendichte» in die Mehrebenenanalyse miteinbezogen werden. Mit diesem Vorgehen wäre es möglich, die Frage der Angebotsinduzierung in der Schweiz das erste Mal mit Hilfe von Individualdaten und nicht mit aggregierten Daten zu überprüfen. Die Resultate könnten dazu verwendet werden, die Frage nach einer maximalen ÄrztInnendichte besser und abgestützt auf empirische Resultate beantworten zu können. Es gilt jedoch zu beachten, dass es sich bei der Messung der Nachfrage bzw. Inanspruchnahme bei der Schweizerischen Gesundheitsbefragung um Selbstangaben handelt.

Es ist grundsätzlich möglich, die Resultate einer solchen Analyse dazu zu verwenden, um regional-spezifische Bedarfsprofile für Versorgungsregionen zu erstellen. Dazu müssten die in der Analyse verwendeten Determinanten auf die Versorgungsregionen «umgerechnet» werden können. Damit verbunden sind jedoch sehr hohe Anforderungen an die Daten. Sie müssen auf der Ebene der Versorgungsgebiete repräsentativ sein. Je nachdem, wie diese Gebiete gebildet werden, ist bspw. die Schweizerische Gesundheitsbefragung dafür nicht geeignet.

Eine Mehrebenenanalyse kann jedoch auch dazu verwendet werden, vertiefte Kenntnisse über das Inanspruchnahmegeschehen, insbesondere zur Hypothese der Angebotsinduzierung, zu generieren. Diese Kenntnisse können für die Modellierung des Modells, welches mit Variante 2 vorgestellt wurde, benutzt werden (vgl. *Abbildung 7, Schritt 3*).

Die Mehrebenenanalyse ist für Bedarfsprognosen kaum geeignet. Das Modell ist zu komplex und man wird kaum über plausible Entwicklungsszenarien für die einzelnen Variablen verfügen.

Datengrundlagen

Um eine Mehrebenenanalyse durchzuführen, braucht es einen Individualdatensatz, der für die medizinischen Versorgungsgebiete repräsentativ ist. Der Datensatz muss auf der Individualebene Informationen über die wichtigsten soziodemographischen, sozioprofessionellen und sozioökonomischen Merkmale, sowie Informationen über den Gesundheitszustand und natürlich über die Inanspruchnahme von Gesundheitsleistungen enthalten. Neben diesen individuellen Merkmalen muss der Datensatz mit (sozio)-strukturellen und angebotsseitigen, aggregierten Kontextvariablen ergänzt werden, damit auch deren Einfluss getestet werden kann³¹.

Vor- und Nachteile

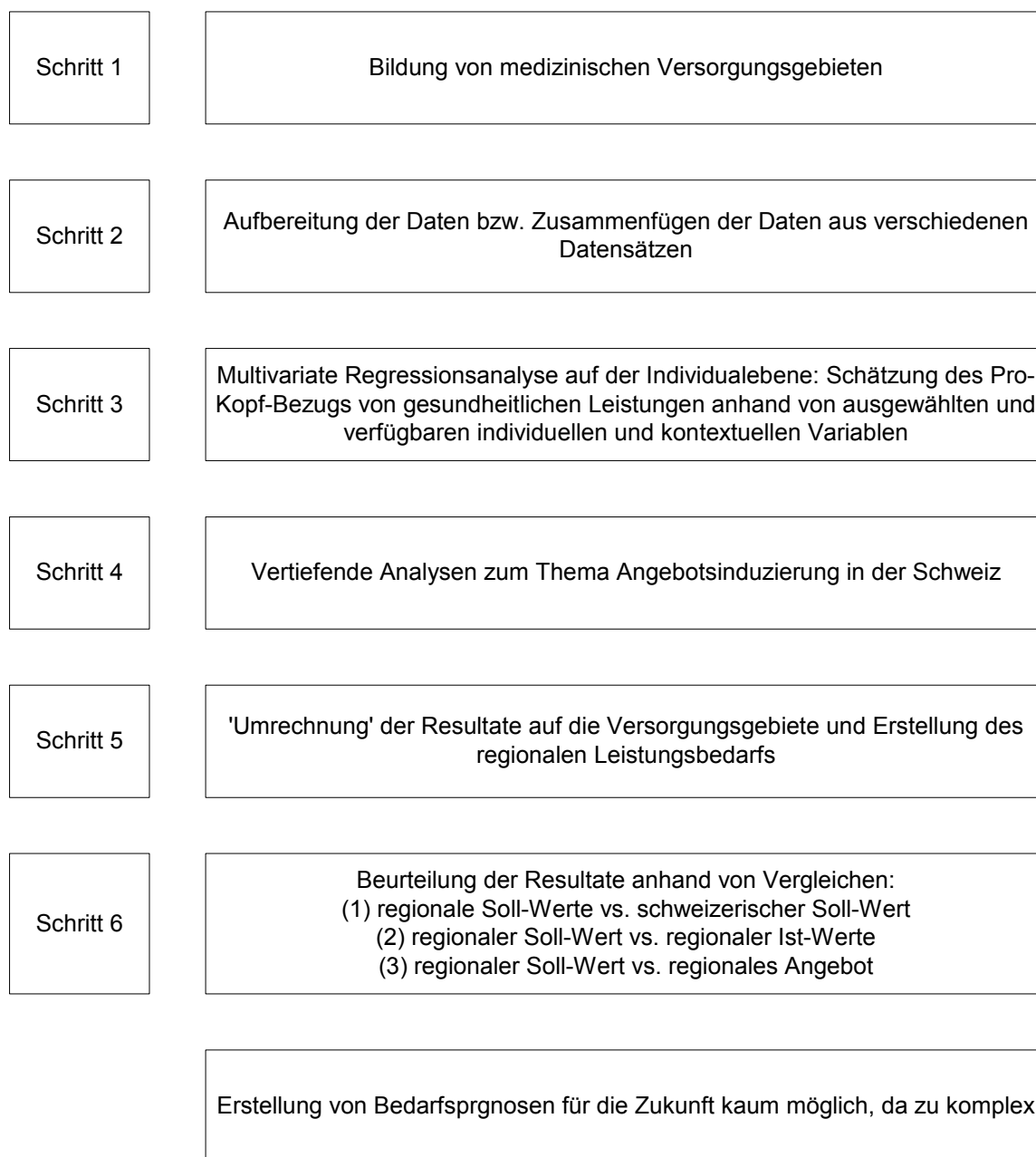
Die vorgeschlagene Mehrebenenanalyse ist noch differenzierter als die Varianten 1 und 2, da nicht auf verhaltenswissenschaftliche Begründungen aus verschiedenen wissenschaftlichen Kontexten verzichtet werden muss. Demgegenüber stehen als Nachteile die noch höheren Anforderungen an die Daten und der noch grössere Aufwand aufgrund der schwierigeren Datenlage. Die «Umrechnung» der Resultate auf Individualebene auf die Ebene der Versorgungsregionen ist dann problematisch, wenn nicht alle berücksichtigten Variablen auf der Ebene der Versorgungsregionen vorhanden sind. Es handelt sich bei diesem Verfahren um ein ähnlich anspruchvolles Verfahren wie dasjenige von Variante 2. In Bezug auf die Verständlichkeit der Resultate im Vergleich zu Variante 1 gilt ähnliches wie für Variante 2. Eine Erstellung von Prognosen für die Zukunft ist aufgrund der höheren Komplexität des Modells noch schwieriger als mit dem Verfahren von Variante 2.

Verfahrensschritte

Abbildung 8 enthält die Verfahrensschritte für die Realisierung von Variante 3.

³¹ Wenn in ein Schätzmodell neben Variablen auf der Individualebene auch so genannte Kontextvariablen auf einer aggregierten Ebene eingebaut werden sollen, müssen aus methodischer Sicht vier Punkte beachtet werden (vgl. Phillips et al. 1998).
(1) Es muss eine Schätzmethode gewählt werden, welche in der Lage ist, Variablen unterschiedlichen Niveaus zu integrieren. Dafür geeignet sind Multilevelanalysen bzw. hierarchische Modelle.
(2) Werden aggregierte angebotsseitige Variablen berücksichtigt (bspw. Ärztedichte), empfiehlt sich ein zwei- oder mehrteiliges Modell. Diese Modelle sind in der Lage, zwischen Erstkontakt und Leistungsbezug während der Behandlungsphase zu unterscheiden.
(3) Bei Multilevelanalysen bestehen oft Probleme der Endogenität-Exogenität. Individuelle Charakteristiken wie bspw. der Gesundheitszustand beeinflussen die Inanspruchnahme, die Inanspruchnahme aber auch den Gesundheitszustand. Bei der Modellwahl muss dies bspw. mit Strukturgleichungsmodellen berücksichtigt werden.
(4) Kontextuelle Variablen stehen oft in komplexen Beziehungen zu anderen Variablen. Auch dies muss bei der Modellspezifikation mitberücksichtigt werden (bspw. mit Pfadanalysen oder mit theoriegeleitetem hierarchischem Einschluss der Variablen).

Abbildung 8: Verfahrensschritte zur Erstellung von regionalen Bedarfsprognosen für Gesundheitsleistungen. Variante 3



Quelle: Eigene Darstellung

Literaturverzeichnis

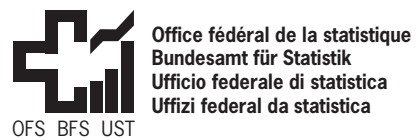
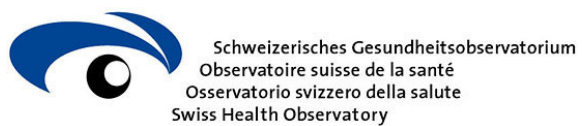
- Andersen H. H. (1992): Themenschwerpunkte und Forschungsfelder der Gesundheitsökonomie. Einführung und Überblick, in: Andersen H. H./Henke, K.-D./ Graf v.d. Schulenburg, J.-M. (Hrsg.): Basiswissen Gesundheitsökonomie. Band 1: Einführende Texte, Berlin, S. 13-37.
- Andersen, H. H. und Schwarze, J. (1997). Angebotsinduzierte Nachfrage bei zunehmendem Wettbewerb? Eine empirische Analyse der Inanspruchnahme ambulanter ärztlicher Leistungen. Veröffentlichungsreihe des Berliner Zentrums Public Health. Berlin, 1-73.
- Andersen, H. H. und Schwarze, J. (2003). Bedarfsprofile in der gesetzlichen Krankenversicherung (GKV). Zur Analyse gruppenspezifischer Unterschiede bei der Inanspruchnahme des Gesundheitsversorgungssystems. Veröffentlichungsreihe des Berliner Zentrums Public Health. Berlin.
- Andersen, R. (1968). A Behavioral Model of Families' Use of Health Services. Research Series. Chicago: Center for Health Administration Studies, The University of Chicago.
- Andersen, R. M. (1995). Revisiting the Behavioral Model and Access to Medical Care: Does It Matter? *Journal of health and social behavior*, 36, 1-10.
- Andersen, R. und Newman, J. F. (1973). Societal and Individual Determinants of Medical Care Utilization in the United States. *Milbank Quarterly*, 51, 95-124.
- Anderson, O. W. (1963). The utilization of health services. *Handbook of medical sociology*. H. Breyer und Zweifel, 1997.
- Bisig, B. und Gutzwiler Felix (Hrsg.) (2004a): Gesundheitswesen Schweiz: Gibt es Unter- oder Überversorgung? Band 1: Gesamtübersicht. Verlag Rüegger. Zürich/Chur.
- Bisig, B. und Gutzwiler Felix (Hrsg.) (2004a): Gesundheitswesen Schweiz: Gibt es Unter- oder Überversorgung? Band 1: Detailresultate. Verlag Rüegger. Zürich/Chur.
- Bodenhöfer 2001: Angebotsinduzierte Nachfrage nach Leistungen niedergelassener Ärzte/Ärztinnen. Gutachten im Auftrag der Gebietskrankenkasse Kärnten. Klagenfurt/Wien.
- Breyer, F. und Zweifel, P. (1997). *Gesundheitsökonomie*. Berlin: Springer.
- Cameron A.C. et al. (1988): A Microeconomic Model. *Review of Economic Studies* 55, 85-106 del of the Demand for Health Care and Health Insurance in Australia.
- Carlsen F. und J. Grytten (1998): More physicians: Improved availability or induced demand: *Health Economics*, 7, 495-508.
- Cassel, D. und Wilke, K. (2000). Das Say'sche Gesetz im Gesundheitswesen: Schafft sich das ärztliche Leistungsangebot seine Nachfrage? Eine Analyse zur angebotsinduzierten Nachfrage in der ambulanten ärztlichen Versorgung anhand deutscher Paneldaten. Dortmund.
- Crivelli, L. et al.: Federalism and regional health care expenditures: an empirical analysis for the Swiss cantons. Quaderno N. 04-07. USI- Istituto di microeconomia ed economia pubblica Mecop. Lugano.
- Cummings, K., Becker, M. und Maile, M. (1980). Bringing the models together: an empirical approach to combining variables used to explain health actions. *Journal of behavioral medicine*, 3, 123-45.
- De Boer, A. G. E. M., Wijker, W. und de Haes, H. C. J. M. (1997). Predictors of health care utilization in the chronically ill: a review of the literature. *Health Policy*, 42, 101-115.
- Delattre, E. und B. Dormont (2003) : Fixed fees and physician-induced demand : A panel data study of French physicians. *Econometric and Health Economics*, 12, 741-754.

- Domenighetti, G. (1998): Médecine ambulatoire: Principale Facteur de Coût et d'inflation à la charge directe des assurés. Cahiers de recherches économiques 9801.
- Domenighetti, G. E. Pipitone (2002): Induction de l'offre de prestations médicales par la demande. *Primary Care*, 2, 241-245.
- Domenighetti, G. und C. Limoni: (1994): Effet de la densité médicale sur les dépenses pour la médecine ambulatoire en Suisse. *Médecine & Hygiène* 52E année.
- Duan, N. et al. (1984) : Choosing Between the Sample-Selection Model and the Multi-Part Model. *Journal of Business and Economic Statistics*, 2(3), 283-289.
- Evans, R. G. (1974): Supplier-Induced Demand: Some empirical Evidence and Implications, in: Perlman, M. (Hrsg.) (1974), 162-173.
- Evashwick C, Rowe G, Diehr P & Branch L (1984) Factors explaining the use of health care services by the elderly. *Health Serv Res* 19: 357–382.
- Eve, SB (1988) A longitudinal study of use of health care services among older women. *J Gerontol* 43: M31–39.
- Feldman, R. and F. Solan (1988): Replay from Feldman and Solan, *Journal of Health Politics, Policy and Law* 14, 621-625.
- Fuchs, J. (1995). Beeinflusst Einkommen die Gesundheit? Analysen mit Daten des Sozio-Ökonomischen Panels. *Das Gesundheitswesen*, 57, 746-752.
- Fuchs, V. R. (1978): The Supply of Surgeons and the Demand for Operations. *The Journal of Human Resources* 13, 25-56.
- Gardiol, L. et al. (2003): Separating Selection and Incentive Effects : an Econometric Study of Swiss Health Insurance Claims Data. Working Paper No. 2003-27. DELTA. Paris.
- Gaynor, M. (1994): Issues in the Industrial Organization of the Market for Physician Services. *Journal of Economics & Management Strategy* 3, 211-255.
- Gerdtham 1996: Equity in Health Care Utilisation: Further Tests Based on Hurdle models and Swedish Micro Data. Working Paper Series in Economics and Finance 116. Stockholm School of Economics
- Greene, W.H. (1993), *Econometric Analysis*, 2nd edition, Macmillan Publishing Company. S. 691.
- Grossman M. (1972a) : The demand for health: A theoretical and empirical investigation. New York.
- Grossman M. (1972b) : On the concept of health capital and the demand for health. *Journal of Political Economy*, 80. S. 223-255.
- Grossman M. (2000) in A.J. Culyer and J.P. Newhouse (Hrsg.): *The Human Capital Model*. Handbook of Health Economics Vol. 1. S. 349-408.
- Grytten, J. und RJ Sorensen (2001): Type of contract and supplier-induced demand for primary physicians in Norway. *Journal of Health Economics*, 20, 379-393.
- Hagenbuch R. (2001): Medizin und Ökonomie. «Check up» - Informationsschrift der Ärztekasse, 7, S. 6-7.
- Härtel, U. (1985). Soziale Determinanten des Gesundheits- und Krankheitsverhalten. Ergebnisse und Folgerungen aus der Münchner Blutdruckstudie. Konstanz: Hartung-Görre Verlag.
- Herder-Dorneich, P. (1994): *Ökonomische Theorie des Gesundheitswesens*. Baden-Baden.
- Herschbach, P. (1995). Über den Unterschied zwischen Kranken und Patienten. *Psychotherapie, Psychosomatik und medizinische Psychologie*, 45, 83-89.

- Hibbard JH & Pope CR (1986) Age differences in the use of medical care in an HMO. An application of the behavioral model. *Med Care* 24: 52–66.
- Horisberger, B., O. Schmid (2004): Ärztstopp als Kostenbremse – Warum es Zeit ist umzudenken. *Schweizerische Ärztezeitung*, 85: Nr 17.
- Jones A. M. (2000): Health Econometrics, in: Newhouse Joseph P. and A. J. Culyer (2000): *Handbook of Health Economics*, Elsevier, Netherlands, Chapter 6, 265-346.
- Keeler et al. (1985): How free care reduced hypertension in the health insurance experiment. *JAMA* 254, 1926-1931.
- Kern A. O. (2002): Arztinduzierte Nachfrage in der ambulanten Versorgung. *Volkswirtschaftliche Diskussionsreihe Beitrag Nr. 225*. Institut für Volkswirtschaftslehre der Universität Augsburg.
- Kouzis AC & Eaton WW (1998) Absence of social networks, social support and health services utilization. *Psychol Med* 28: 1301–1310.
- Kurth B.-M. et al. (2004): Einflussfaktoren auf die Inanspruchnahme des deutschen Gesundheitswesens und mögliche Steuerungsmechanismen. Robert Koch Institut. Berlin
- Labelle, R.; Stoddart, G.; Rice, Th. (1994): A re-examination of the meaning and importance of supplier-induced demand, in: *Journal of Health Economics* 13, 347-368.
- Labelle, R.; Stoddart, G.; Rice, Th. (1994): Editorial: Response to Pauly on a re-examination of the meaning and importance of supplier-induced demand, in: *Journal of Health Economics* 13, S. 491-494.
- Leu, R. E. und Gerfin, M. (1992). Die Nachfrage nach Gesundheit - Ein empirischer Test des Grossmann-Modells. *Steuerungsprobleme im Gesundheitswesen*. P. Oberender. Baden-Baden.
- Lurie N. et al. (1989) : How fee care improved vision in the health insurance experiment. *Am J Public Health* 79, 640-642.
- Manning W. et al. (1987) : Health Insurance and Demand for Medical Care : Evidence from a randomized experiment. *Am Econ Rev* 77, 251-277.
- Marty, F. E. (1997): Zur Verrechnungspraxis der Ärzte – Eine empirische Untersuchung. *Forschungsbericht für das Nationale Forschungsprojekt* 11 893. Bern.
- MC Guire, TG. Physician agency. *Handbook of Health Economics*. Culyer AJ, Newhouse JP (eds.) 461-536.
- McKinlay, J. B. (1972). Some approaches and problems in the study of the use of services - an overview. *Journal of Health and Social Behavior*, 13, 115-152.
- Mechanic, D. (1976) Sex, illness behavior, and the use of health services. *J Human Stress* 2: 29–40.
- Mechanic, D. (1979). Correlates of Physician Utilization: Why Do Major Multi-Variate Studies of Physician Utilization Find Trivial Psycho-Social and Organizational Effects? *Journal of Health and Social Behavior*, 20, 387-396.
- Muurinen, J. (1982), Demand for Health, in: *Journal of Health Economics*, 1, 5-28.
- Nassiri, A. L Rochaix (2002): Understanding physicians' strategic behavior under expenditure caps: Lesson from a Quebec natural experiment. *Proceedings of the 11th European Workshop on Econometrics and Health Economics*. Sweden.
- Newhouse, J. et al. (1993) : Free for all ? Lessons from the RAND Health Insurance Experiment. *RAND-Studie*, Harvard University, Cambridge (MA).

- Newhouse, J. und CH. Phelps (1974): Price and Income Elasticities for Medical Care Services. In The Economics of Health and Medical Care: Proceeding of a Conference of the International Economics Association.
- Nocera, S. und P. Zweifel (1998), The Demand for Health: An Empirical Test of the Grossman Model Using Panel Data, in: Zweifel, P. (Hg.), Health, the Medical Profession, and Regulation, Kluwer, Dordrecht, 35-49.
- Pauly M. V. und Satterhwaite (1981): The pricing of primary care physicians' services: A test of the role of consumer information. *Bell Journal of Economics* 12, 480-506.
- Philips, K. A., Morrison, K. R., Andersen, R. und Aday, L. A. (1998). Understanding the Context of Healthcare Utilization: Assessing Environmental and Provider-Related Variables in the Behavioral Model of Utilization. *Health Services Research*, 33, 571-596.
- Pohlmeier, W. und Ulrich, V. (1995). An Econometric Model of the Two-Part Decision Process in the Demand for Health. *Journal of Human Resources*, 30(2), 339-361.
- Reinhardt U. E. (1995): The theory of physician-induced demand – Reflections after a decade, *Journal of Health Economics* 4, 187-193.
- Richardson J. (1991): The effects of Co-payments in Medical Care. National Health Strategy Background Paper No 5. Calgary.
- Rosenstock, I. (1966). Why People Use Health Services. *Millbank Memorial Fund Quarterly*, 44, 94-124.
- Rossiter, L.F. and G.R. Wilensky (1983): A re-examination of the use of physician services: The role of physician-initiated demand. *Inquiry* 20, 162-172.
- Schaufelberger, H. J., Dolder-Zeller, C., Jost, U. und Junker, C. (1989). Determinanten der Entscheidungen für die Inanspruchnahme einer ambulanten ärztlichen Einrichtung. Medizinisches Institut für Ausbildungs- und Examensforschung: Fakultät der Universität Bern. 1, 1-241.
- Schellhorn M. (2002): Auswirkungen wählbare Selbstbehalte in der Krankenversicherung: Lehren aus der Schweiz? *Vierteljahreshefte zur Wirtschaftsforschung* 71, 4, 411-426.
- Schellhorn M. (2004): Wählbare Selbstbehalte in der Krankenversicherung der Schweiz: Nachfragesteuerung oder Selektion? *Diskussionsschriften* 04-02 Volkswirtschaftliches Institut Universität Bern.
- Schneider, U. (2003): Kostenfalle Gesundheitswesen? Ökonomische Herausforderung und Perspektiven der Gesundheitssicherung. *Wirtschaftswissenschaftliche Diskussionspapiere*. Universität Bayreuth. Diskussionspapier 08-03. Bayreuth.
- Schöffski, O. (2004): Gesundheitsmanagement und gesundheitsökonomie. <http://www.gm.wiso.uni-erlangen.de/Materialien/ArtGesmanagoek.pdf> (Februar 2005).
- Schrire, S. (1986) Frequent attenders – a review. *Fam Pract* 3: 272–275.
- Schulenburg, J. (1981): Systeme der Honorierung frei praktizierender Ärzte und ihre Allokationswirkungen. Tübingen.
- Simon, HA (1982): Bounded Rationality. *Behavioral Economics and Business Organization*, London.
- Sorensen, RJ und J. Grytten (1999): Competition and supplier-induced demand in a health care system with fixed fees. *Health Economics*, 13, 497-508.
- Stano, M. (1987): An analysis of the evidence on competition in the physician services markets. *Journal of Health Economics* 4, 197-211.

- Wagstaff, A. (1986): The Demand for Health: Some New Empirical Evidence. *Journal of Health Economics*, 5(3), 195-233.
- Werblow, A. (2002) Alles nur Selektion? Der Einfluss von Selbstbehalten in der gesetzlichen Krankenversicherung. *Vierteljahreshefte zur Wirtschaftsforschung* 71, 4, 427-436.
- Werblow, A. und S. Felder (2002): Der Einfluss von freiwilligen Selbstbehalten in der gesetzlichen Krankenversicherung: Evidenz aus der Schweiz. Institut für Sozialmedizin und Gesundheitsökonomie. Otto-von-Guericke-Universität. Magdeburg.
- Zweifel, P. (1985): The Effect of Aging on the Demand and Utilization of Medical Care. Tilquin, C. (Hrsg.): *Systems Science in Health and Social Services for Elderly and Disabled*. Pergamon Press, 313-318. Toronto.
- Zweifel, P. (1992). Das Individuum als Produzent seiner Gesundheit: Eine stochastische Formulierung. *Steuerungsprobleme im Gesundheitswesen*. P. Oberender. Baden-Baden.



Das Schweizerische Gesundheitsobservatorium (Obsan) ist eine Organisationseinheit des Bundesamtes für Statistik
L'Observatoire suisse de la santé (Obsan) est une unité organisationnelle de l'Office fédéral de la statistique