

Sonderdruck aus:

# Mitteilungen aus der Arbeitsmarkt- und Berufsforschung

Anne Dreyer-Tümmel, Johann Behrens, Detlef Schulz

Erwerbsverlaufsforschung mit Daten der  
gesetzlichen Krankenversicherung

## **Mitteilungen aus der Arbeitsmarkt- und Berufsforschung (MittAB)**

Die MittAB verstehen sich als Forum der Arbeitsmarkt- und Berufsforschung. Es werden Arbeiten aus all den Wissenschaftsdisziplinen veröffentlicht, die sich mit den Themen Arbeit, Arbeitsmarkt, Beruf und Qualifikation befassen. Die Veröffentlichungen in dieser Zeitschrift sollen methodisch, theoretisch und insbesondere auch empirisch zum Erkenntnisgewinn sowie zur Beratung von Öffentlichkeit und Politik beitragen. Etwa einmal jährlich erscheint ein „Schwerpunkt-heft“, bei dem Herausgeber und Redaktion zu einem ausgewählten Themenbereich gezielt Beiträge akquirieren.

### *Hinweise für Autorinnen und Autoren*

Das Manuskript ist in dreifacher Ausfertigung an die federführende Herausgeberin Frau Prof. Jutta Allmendinger, Ph. D.  
Institut für Arbeitsmarkt- und Berufsforschung  
90478 Nürnberg, Regensburger Straße 104  
zu senden.

Die Manuskripte können in deutscher oder englischer Sprache eingereicht werden, sie werden durch mindestens zwei Referees begutachtet und dürfen nicht bereits an anderer Stelle veröffentlicht oder zur Veröffentlichung vorgesehen sein.

Autorenhinweise und Angaben zur formalen Gestaltung der Manuskripte können im Internet abgerufen werden unter [http://doku.iab.de/mittab/hinweise\\_mittab.pdf](http://doku.iab.de/mittab/hinweise_mittab.pdf). Im IAB kann ein entsprechendes Merkblatt angefordert werden (Tel.: 09 11/1 79 30 23, Fax: 09 11/1 79 59 99; E-Mail: [ursula.wagner@iab.de](mailto:ursula.wagner@iab.de)).

### **Herausgeber**

Jutta Allmendinger, Ph. D., Direktorin des IAB, Professorin für Soziologie, München (federführende Herausgeberin)  
Dr. Friedrich Buttler, Professor, International Labour Office, Regionaldirektor für Europa und Zentralasien, Genf, ehem. Direktor des IAB  
Dr. Wolfgang Franz, Professor für Volkswirtschaftslehre, Mannheim  
Dr. Knut Gerlach, Professor für Politische Wirtschaftslehre und Arbeitsökonomie, Hannover  
Florian Gerster, Vorstandsvorsitzender der Bundesanstalt für Arbeit  
Dr. Christof Helberger, Professor für Volkswirtschaftslehre, TU Berlin  
Dr. Reinhard Hujer, Professor für Statistik und Ökonometrie (Empirische Wirtschaftsforschung), Frankfurt/M.  
Dr. Gerhard Kleinhenz, Professor für Volkswirtschaftslehre, Passau  
Bernhard Jagoda, Präsident a.D. der Bundesanstalt für Arbeit  
Dr. Dieter Sadowski, Professor für Betriebswirtschaftslehre, Trier

### **Begründer und frühere Mitherausgeber**

Prof. Dr. Dieter Mertens, Prof. Dr. Dr. h.c. mult. Karl Martin Bolte, Dr. Hans Büttner, Prof. Dr. Dr. Theodor Ellinger, Heinrich Franke, Prof. Dr. Harald Gerfin, Prof. Dr. Hans Kettner, Prof. Dr. Karl-August Schäffer, Dr. h.c. Josef Stingl

### **Redaktion**

Ulrike Kress, Gerd Peters, Ursula Wagner, in: Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB), 90478 Nürnberg, Regensburger Str. 104, Telefon (09 11) 1 79 30 19, E-Mail: [ulrike.kress@iab.de](mailto:ulrike.kress@iab.de); (09 11) 1 79 30 16, E-Mail: [gerd.peters@iab.de](mailto:gerd.peters@iab.de); (09 11) 1 79 30 23, E-Mail: [ursula.wagner@iab.de](mailto:ursula.wagner@iab.de); Telefax (09 11) 1 79 59 99.

### **Rechte**

Nachdruck, auch auszugsweise, nur mit Genehmigung der Redaktion und unter genauer Quellenangabe gestattet. Es ist ohne ausdrückliche Genehmigung des Verlages nicht gestattet, fotografische Vervielfältigungen, Mikrofilme, Mikrofotos u.ä. von den Zeitschriftenheften, von einzelnen Beiträgen oder von Teilen daraus herzustellen.

### **Herstellung**

Satz und Druck: Tümmels Buchdruckerei und Verlag GmbH, Gundelfinger Straße 20, 90451 Nürnberg

### **Verlag**

W. Kohlhammer GmbH, Postanschrift: 70549 Stuttgart; Lieferanschrift: Heßbrühlstraße 69, 70565 Stuttgart; Telefon 07 11/78 63-0; Telefax 07 11/78 63-84 30; E-Mail: [waltraud.metzger@kohlhammer.de](mailto:waltraud.metzger@kohlhammer.de), Postscheckkonto Stuttgart 163 30. Girokonto Städtische Girokasse Stuttgart 2 022 309. ISSN 0340-3254

### **Bezugsbedingungen**

Die „Mitteilungen aus der Arbeitsmarkt- und Berufsforschung“ erscheinen viermal jährlich. Bezugspreis: Jahresabonnement 52,- € inklusive Versandkosten; Einzelheft 14,- € zuzüglich Versandkosten. Für Studenten, Wehr- und Ersatzdienstleistende wird der Preis um 20 % ermäßigt. Bestellungen durch den Buchhandel oder direkt beim Verlag. Abbestellungen sind nur bis 3 Monate vor Jahresende möglich.

### **Zitierweise:**

MittAB = „Mitteilungen aus der Arbeitsmarkt- und Berufsforschung“ (ab 1970)  
Mitt(IAB) = „Mitteilungen“ (1968 und 1969)  
In den Jahren 1968 und 1969 erschienen die „Mitteilungen aus der Arbeitsmarkt- und Berufsforschung“ unter dem Titel „Mitteilungen“, herausgegeben vom Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit.

**Internet:** <http://www.iab.de>

# Erwerbsverlaufsforschung mit Daten der gesetzlichen Krankenversicherung

Anne Dreyer-Tümmel, Johann Behrens, Detlef Schulz\*

Den Gegenstand der Erwerbsverlaufsforschung bilden individuelle Verläufe des Erwerbslebens. Im Kontext der sozialmedizinischen Erwerbsverlaufsforschung werden gesundheitliche Folgen der Erwerbsarbeit und ihrer Unterbrechungen (Krankheitsursachenforschung), die Folgen gesundheitlicher Beeinträchtigungen für den Erwerbsverlauf (Krankheitsfolgenforschung) sowie Wechselwirkungen zwischen Erwerbskarrieren und gesundheitlichen Beeinträchtigungen untersucht. Die Verlaufsperspektive ermöglicht es, bei diesen Analysen Krankheitsverläufe und Chronifizierungsprozesse zu berücksichtigen, Selektionseffekte (z.B. sogenannte *healthy-* und *unhealthy-worker*-Effekte) zu erkennen und Risikokumulationen sowie Labilisierungs-, Invalidisierungs- und Stabilisierungsprozesse aufzuzeigen, die bei Untersuchungen auf der Basis von Querschnittsdaten unsichtbar bleiben.

Eine gute Datengrundlage für viele Fragestellungen insbesondere der sozialmedizinischen Erwerbsverlaufsforschung stellen die im Verwaltungsablauf kontinuierlich anfallenden, personenbezogenen Versicherten- und Leistungsdaten der gesetzlichen Krankenversicherungen (GKV-Daten) dar. Sie enthalten nicht nur verlaufsbezogene Angaben zu Arbeitsunfähigkeiten, Klinikaufhalten und Diagnosen, sondern darüber hinaus werden auch wesentliche Ereignisse des Erwerbsverlaufs für die Mitglieder von Krankenversicherungen im Längsschnitt taggenau dokumentiert.

In diesem Beitrag werden eine Reihe von Möglichkeiten, Zusammenhänge zwischen Krankheit und Erwerbsverlauf auf der Grundlage von GKV-Daten zu analysieren, aufgezeigt und an einigen empirischen Forschungsergebnissen exemplarisch veranschaulicht sowie Stärken und Schwächen der Angaben in Krankenkassendaten für verschiedene Fragestellungen der Erwerbsverlaufsforschung diskutiert. Ergänzend wird auf einige Parallelen und Unterschiede zwischen GKV-Daten und den Mikrodaten der Beschäftigtenstatistik der Bundesanstalt für Arbeit (BA), insbesondere der IAB-Beschäftigtenstichprobe, als Datenquellen der Erwerbsverlaufsforschung Bezug genommen.

## Gliederung

- 1 Einleitung: Erwerbsverlaufsforschung im Überblick
- 2 Erwerbsverläufe in Krankenkassendaten
  - 2.1 Überblick
  - 2.2 Episoden und Ereignisse
- 3 Erwerbsverläufe und Arbeitsunfähigkeit: Sozialmedizinische Erwerbsverlaufsforschung mit Krankenkassendaten
  - 3.1 Arbeitsunfähigkeiten
  - 3.2 Arbeitsunfähigkeitsdiagnosen
- 4 Unabhängige Variablen und Subgruppen im Kontext der Erwerbsverlaufsforschung: Das Potential in Krankenkassendaten im Überblick
  - 4.1 Soziodemographische Merkmale
  - 4.2 Soziale Schicht, sozialer Status, Einkommen
  - 4.3 Gründe, Motive, Einstellungen und Verhaltensweisen
  - 4.4 Arbeitsmarkt, Betrieb, Arbeitsplatz

## Literatur

\* Dipl.-Soz. Anne Dreyer-Tümmel und Dipl.-SozWiss. Detlef Schulz arbeiten als wissenschaftliche MitarbeiterInnen im Teilprojekt C4 „Abstiegskarrieren und Auffangpositionen“ des Sonderforschungsbereichs 186 „Statuspassagen und Risikolagen im Lebenslauf“ an der Universität Bremen. Prof. Dr. Johann Behrens ist Leiter dieses Projektes sowie von Projekten des Instituts für Supervision, Institutionsberatung und Sozialforschung (ISIS) in Frankfurt. Er lehrt am Aufbaustudiengang „Öffentliche Gesundheit/Gesundheitswissenschaften“ an der Universität Bremen und an der Fachhochschule Fulda. Der Beitrag liegt in der alleinigen Verantwortung der AutorInnen.

## 1 Einleitung: Erwerbsverlaufsforschung im Überblick

Erwerbsverlaufsforschung befaßt sich mit der Untersuchung *individueller Verläufe* des Erwerbslebens. Sie stellt damit einen wichtigen Teilbereich der – seit den späten 60er Jahren stark an Bedeutung gewinnenden – Lebenslaufforschung dar. Ihr Gegenstand können ganze Erwerbskarrieren (von der Ausbildung bis zur Berentung), bestimmte Sequenzen (z.B. Abstiegskarrieren) oder Episoden des Erwerbslebens (z.B. Arbeitslosigkeitsphasen) sowie Ereignisse und Statuspassagen im Erwerbsverlauf (z.B. Arbeitgeberwechsel, Berentungen) sein.

Im Mittelpunkt der *sozialmedizinischen Erwerbsverlaufsforschung* steht die Analyse der Zusammenhänge zwischen Krankheit und Erwerbsverlauf. Es wird nach den gesundheitlichen Folgen der Erwerbsarbeit und ihrer Unterbrechungen (d.h. von Belastungen am Arbeitsplatz, aber z.B. auch von Arbeitslosigkeit) oder – umgekehrt – nach den Folgen gesundheitlicher Beeinträchtigungen für den Erwerbsverlauf (z.B. für das Risiko eines Arbeitsplatzverlustes, die Chance einer erfolgreichen Wiedereingliederung oder die Wahrscheinlichkeit einer Frühberentung) sowie nach Wechselwirkungen zwischen Krankheits- und Erwerbsverläufen gefragt, wie sie z.B. in Labilisierungs- und Invalidisierungsprozessen zum Ausdruck kommen können. Die Verlaufsperspektive ermöglicht es im Unterschied zu Querschnittsuntersuchungen nicht nur, bei diesen Analysen Krankheitsverläufe und Chronifizierungsprozesse zu berücksichtigen (siehe dazu Abschnitt 3), sondern darüber hinaus auch, Abfolgen von Ereignissen und Episoden des Erwerbsverlaufs nachzuzeichnen, am Arbeitsmarkt wirksame Selektionsprozesse zu erkennen und Risikokumulationen sowie Labilisierungs- und Stabilisierungsprozesse aufzuzeigen (siehe dazu unten).

Weitere thematische Schwerpunkte der Erwerbsverlaufsforschung bilden insbesondere die Produktion, Reproduktion und Kumulation sozialer Ungleichheiten im Erwerbsverlauf und die Erwerbs- (und Lebens-)Verläufe von Frauen. Der erstgenannte dieser – sich selbstverständlich untereinander und mit der sozialmedizinischen Erwerbsverlaufsforschung überschneidenden – Themenschwerpunkte umfaßt vor allem Untersuchungen, die sich mit der Bedeutung verschiedener Dimensionen und Determinanten sozialer Ungleichheit (wie z.B. soziale Schicht, Geschlecht, Qualifikation, gesundheitliche Beeinträchtigungen) für Risiken und Chancen im Erwerbsverlauf (z.B. für das Arbeitslosigkeitsrisiko, die berufliche Karriere), sozialen Ungleichheiten (z.B. Einkommensungleichheiten, ungleichen Arbeitsmarktchancen) als Folgen krisenhafter Ereignisse im Erwerbsverlauf (wie z.B. der Verlust des Arbeitsplatzes, längerfristige Arbeitsunfähigkeit) oder entsprechenden Wechselwirkungen und Risikokumulationen befassen. Studien zum letztgenannten Themenschwerpunkt sind vor allem auf Benachteiligungen von Frauen im Erwerbsleben, die (lebensphasenspezifische) Erwerbsbeteiligung von Frauen sowie Fragen der Vereinbarkeit von Familie und Beruf gerichtet. Oft stehen dabei biographische Interpretationen der Frauen im Mittelpunkt des Interesses.

Der sozialpolitische und forschungspraktische Nutzen der Erwerbsverlaufsforschung ist vor allem darin zu sehen, daß durch die Betrachtung der *Verkettung von Ereignissen* im individuellen Verlauf *Wechselwirkungen* und *Risikokumulationen* aufgezeigt und *Prozeßstrukturen* erkannt werden können, die bei einem Denken in Einzelereignissen (Fall- statt Personenorientierung) oft unsichtbar bleiben. Aus den Ergebnissen lassen sich Hinweise auf Bedingungskonstellationen und Scheidepunkte des Erwerbslebens, die für die Krisenhaftigkeit oder Stabilität des weiteren Verlaufs eine zentrale Bedeutung einnehmen, und auf sozialpolitische Steuerungs- und Interventionsmöglichkeiten ableiten, mit denen die Entstehung oder weitere Verschärfung von Krisen des Erwerbsverlaufs vermieden und bereits in Gang gesetzte Labilisierungsprozesse verlangsamt, gestoppt oder sogar umgekehrt werden können. So können auf der Basis individueller Verlaufsdaten beispielsweise typische Wege in die Früh- und in die Altersrente aufgezeigt und damit nicht nur zum Verständnis der Frühberentung als Prozeß beigetragen, sondern auch Hinweise auf (individuelle wie institutionelle) Präventionsmöglichkeiten abgeleitet werden. Sigrid Quack (1991: 778) weist darauf hin, daß eine differenzierte Untersuchung der Hypothese, der Teilzeitarbeit komme bei der Wiedereingliederung in den Arbeitsmarkt nach einer familienbedingten Unterbrechung eine „Brückenfunktion“ zu, den weiteren Erwerbsverlauf der betroffenen Frauen mit einbeziehen muß und somit – so ist zu ergänzen – notwendig individuelle Längsschnittdaten voraussetzt. Gleiches gilt auch für die Prüfung der Frage, ob der im Vergleich zu Erwerbstätigen schlechtere Gesundheitszustand Arbeitsloser primär auf Selektionsprozesse bei Entlassungen und Einstellungen („Selektionsthese“) oder eher auf gesundheitliche Beeinträchtigungen als direkte Folge der Arbeitslosigkeit („Kausationsthese“) zurückzuführen ist (siehe z.B. Elkeles/Seifert 1992; Arrow 1994; Andreß 1996).

Ein weiterer Vorteil individuenbezogener Längsschnittanalysen im Vergleich zu Querschnittuntersuchungen oder Aggregatvergleichen im Zeitverlauf ist darin zu sehen, daß *ökologische Fehlschlüsse* und *Fehlinterpretationen*, die auf die *Nichtberücksichtigung individueller Mobilität* zurückzuführen sind, aufgedeckt und vermieden werden können. So ist wiederholt darauf hingewiesen worden, daß individuelle Mobilität auf der Basis von Querschnittvergleichen unterschätzt

wird und sogar unsichtbar bleiben kann, da gegenläufige Entwicklungstrends trotz eines hohen Maßes an individueller Veränderung im Kollektiv scheinbare Stabilität suggerieren können (z.B. Behrens 1983: 215ff.; Schupp 1991: 790). Beispielsweise haben Helga Krüger und Claudia Born (1991: 145ff.) auf der Grundlage individueller Längsschnittdaten gezeigt, daß weibliche Lebensverläufe durch die sich beim Vergleich aggregierter Daten im Zeitverlauf andeutenden Phasenmodelle nur unzureichend charakterisiert werden und statt dessen durch vielfältige Differenzierungen im Verhältnis von Erwerbs- und Familienarbeit gekennzeichnet sind. Doch auch wenn individuelle Mobilität im Aggregat (zunächst) keine Effekte zeigt, „kann dieser ‘Mikrowandel’ aus der Perspektive der einzelnen z.T. erhebliche biographische Brüche markieren, die kurz- und mittelfristig sichtbare Konsequenzen haben“ (Berger/Sopp 1991: 531).

Fehlinterpretationen als Folge nicht berücksichtigter Erwerbsmobilität sind auch aus der Berufsrisikoforschung bekannt. Ermittelte Zusammenhänge zwischen Arbeitsbedingungen und Gesundheit oder zwischen Berufen und Frühberentungsrisiko können (systematisch) verzerrt sein, wenn in entsprechende Untersuchungen nur die aktuelle Tätigkeit oder der letzte Beruf einbezogen wird und frühere Tätigkeiten sowie Beschäftigungsdauern unberücksichtigt bleiben (siehe z.B. Volkholz/Schwarz 1984: 117; die Stellungnahmen von H.-J. Lange und E. Schach in Hernberg u.a. (Hg.) 1986: 14, 23; Behrens u.a. 1992: 380ff.); Zum einen gibt es nämlich Arbeitsplätze und Berufe, die von gesundheitlich Beeinträchtigten häufig verlassen werden (müssen) – mit der Folge, daß Zusammenhänge zwischen diesen Tätigkeiten und gesundheitlichen Schädigungen bei Untersuchungen auf der Basis der verbleibenden Beschäftigten – der sogenannten „survivor population“ (Hernberg 1984: 288) – verdeckt bleiben können. Dieser *healthy-worker-Effekt* kann durch analoge Selektionseffekte bei Einstellungen noch verstärkt werden. Zum anderen gibt es Positionen, auf die viele ArbeitnehmerInnen gerade wegen gesundheitlicher Beeinträchtigungen wechseln. Dabei kann es sich um gesundheitlich zuträgliche Arbeitsplätze – bekannt ist das Beispiel „Pförtner“ –, aber auch um stark belastende Tätigkeiten (wie z.B. Lagerarbeiten) handeln, die gesundheitlich Vorgeschädigte häufig akzeptieren müssen, wenn sie überhaupt noch einen Arbeitsplatz finden wollen. Bleiben solche *Drift- oder unhealthy-worker-Effekte* unberücksichtigt, kann das mit diesen – oft in der letzten Phase des Erwerbslebens eingenommenen – Positionen verbundene gesundheitliche Risiko überschätzt werden.

Als Datenquellen der Erwerbsverlaufsforschung kommen – je nach Fragestellung – vor allem retrospektiv ausgerichtete Befragungen (wie z.B. die Lebenslaufstudie des Deutschen Jugendinstituts in München), biographische Interviews, Panelerhebungen (wie z.B. das Sozio-ökonomische Panel) und die personenbezogenen Routinedaten der Sozialversicherungsträger (z.B. GKV-Daten) sowie darauf beruhende Datensätze (wie z.B. die IAB-Beschäftigtenstichprobe) in Betracht. Makrodaten (wie z.B. die Bestandsstatistik der BA) und Querschnittdaten (wie z.B. der Mikrozensus) lassen dagegen keine Aussagen über individuelle Verläufe, sondern allenfalls Aggregatvergleiche im Zeitverlauf zu. Unterschiede zwischen den für die Erwerbsverlaufsforschung geeigneten Datenarten bestehen zum einen hinsichtlich des für die Datenerhebung und -aufbereitung erforderlichen Aufwands. Darüber hinaus beziehen sie sich vor allem auf den Erhebungszweck (Forschung versus Verwaltung) und den Entstehungsprozeß (Befragungs- oder Interviewsituation versus Routineerhebung) – und damit auf Inhalte, Vollständigkeit und

Zuverlässigkeit der Daten –, auf die Zeitperspektive bei der Datenerfassung (retrospektiv versus fortlaufend), auf die Aktualität der Daten sowie auf Fallzahlen, Repräsentativität und Beobachtungszeiträume. Die personenbezogenen Routinedaten der Sozialversicherungsträger sind durch – für die Forschung – geringen Erhebungsaufwand, die fortlaufende taggenaue Erfassung einer Reihe von Ereignissen, große Fallzahlen, die ausreichende Zellenbesetzungen auch für Teilpopulationen erwarten lassen, und – sofern sie langfristig archiviert werden – auch durch lange Beobachtungszeiträume gekennzeichnet. Dem steht – neben einem erheblichen Aufwand für die Datenaufbereitung – gegenüber, daß die Angaben in Sozialversicherungsdaten weitgehend auf versicherungspflichtige Personen beschränkt sind und ihre Inhalte in hohem Maße durch gesetzliche Vorgaben und die Anforderungen der Verwaltung bestimmt werden. Die Stärken und Schwächen von Routinedaten der gesetzlichen Krankenversicherungen im Hinblick auf verschiedene Fragestellungen der Erwerbsverlaufs-forschung sollen im folgenden ausführlich diskutiert werden. Ergänzend wird dabei auch auf einige Parallelen und Unterschiede zwischen GKV-Daten und den Mikrodaten der Beschäftigtenstatistik der BA, insbesondere der IAB-Beschäftigtenstichprobe, Bezug genommen.

## 2 Erwerbsverläufe in Krankenkassendaten

### 2.1 Überblick

Auf der Grundlage des am 1.1.1973 eingeführten, in der Datenerfassungs-Verordnung (DEVO) geregelten integrierten Meldeverfahrens zur Sozialversicherung und zur Bundesanstalt für Arbeit haben Arbeitgeber Beginn und Ende von sozialversicherungspflichtigen Beschäftigungsverhältnissen einschließlich ihrer Unterbrechungen sowie Änderungen in der Sozialversicherungspflicht von Beschäftigten – konkret: in den Beitragsgruppen zur Kranken- und Rentenversicherung und zur BA – den zuständigen Krankenversicherungsträgern zu melden (zur Definition sozialversicherungspflichtiger Beschäftigung siehe Cramer 1985: 59f.; zum Meldeverfahren siehe ausführlicher Wermter/Cramer 1988: 471ff.). Diese Meldungen müssen u.a. auch Angaben zur ausgeübten Tätigkeit, zur Stellung im Beruf und zum beitragspflichtigen Bruttoarbeitsentgelt sowie Name, Anschrift und Betriebsnummer des Arbeitgebers enthalten. Neben solchen Arbeitgebermeldungen über sozialversicherungspflichtige Beschäftigungsverhältnisse, die – krankenkassenübergreifend – auch die Grundlage der Beschäftigtenstatistik der BA (siehe dazu z.B. Wermter 1981; Wermter/Cramer 1988) bilden, liegen den gesetzlichen Krankenversicherungen auch entsprechende Meldungen zum Beginn und zum Ende von anderen Sozialversicherungspflicht begründenden Episoden wie z.B. Rehabilitationsmaßnahmen oder – über die Leistungsempfängerdatei der BA auch in die IAB-Beschäftigtenstichprobe eingehende (siehe z.B. Bender/Hilzendege 1995: 183f.) – Arbeitslosigkeiten vor. Darüber hinaus sind in GKV-Daten Angaben über weitere versicherte Personenkreise wie z.B. freiwillige Mitglieder oder RentnerInnen enthalten.

In den Versichertendaten der gesetzlichen Krankenversicherungen liegen also Informationen über eine Reihe von Ereignissen im und Episoden des Erwerbsverlaufs vor, die über die in der Beschäftigtenstatistik der BA und der IAB-Beschäftigtenstichprobe erfaßten Meldetatbestände teilweise noch hinausgehen: Arbeitgeber- und z.T. Berufswechsel, Veränderungen im Versicherungsstatus, Rentenanträge und Berentungen sind taggenau dokumentiert; Beschäftigungsverhältnisse, Arbeitslosigkeiten, Episoden freiwilliger Mitgliedschaft, Rehabilitationsmaßnahmen und Versicherungsverhältnisse nach § 192

SGB V, der den Erhalt der Krankenversicherungsmemberschaft vor allem beim Bezug von Kranken- oder Mutterschaftsgeld regelt, werden mit ihren jeweiligen Dauern abgebildet.

Die Validität der Meldungen zu Beschäftigungszeiten und Zeiten des Bezugs von Lohnersatzleistungen und damit auch der Informationen zu Arbeitgeberwechseln sowie Wechseln zwischen Betrieben und Institutionen der sozialen Sicherung ist als vergleichsweise hoch einzuschätzen, da diese Angaben sich unmittelbar auf die Beitragspflicht des Arbeitgebers oder der jeweiligen Institution auswirken (siehe unter Bezug auf die Beschäftigtenstatistik der BA auch Cramer 1985: 62). Zwar ist bei den An- und Abmeldungen mit verspäteten Dateneingängen zu rechnen, die sich darüber hinaus – wie Stefan Bender und Jürgen Hilzendege (1995: 83) unter Bezug auf die IAB-Beschäftigtenstichprobe ausführen – nicht zufällig verteilen, sondern auf bestimmte Versichertengruppen und Branchen sowie kleinere Betriebe konzentrieren. Von dieser Verzerrung sind aber vornehmlich die jüngsten Jahrgänge der Versichertendaten einer Krankenkasse betroffen. Die Validität und Differenziertheit der weiteren Angaben in GKV-Daten wird im allgemeinen um so höher bewertet, je mehr Bedeutung ein Datum im administrativen Prozeß der Krankenversicherung hat (siehe z.B. die Stellungnahmen von W. Thiele und E. Schach in Hernberg u.a. (Hg.) 1986: 50ff., 80).

Werden die Angaben aus den Meldungen an die Krankenversicherungsträger nicht regelmäßig gelöscht, sondern über längere Zeiträume aufbewahrt oder sogar dauerhaft archiviert, dann liegt damit ein Datenpotential für die Erwerbsverlaufs-forschung vor, das nicht nur die Analyse einzelner *Übergänge im oder Episoden des Erwerbslebens* erlaubt, sondern auch die Auswertung von *Statussequenzen* oder – wenn hinreichend lange Beobachtungszeiträume vorliegen – sogar von *ganzen Erwerbsverläufen* ermöglicht. Das ist nicht nur bei den Mikrodaten der Bundesanstalt für Arbeit und der darauf beruhenden, gegenwärtig den Zeitraum von 1975 bis 1990 umfassenden IAB-Beschäftigtenstichprobe, sondern mittlerweile auch bei den Versichertendaten einiger Krankenkassen der Fall. Allerdings sind die Informationen in den Versichertendaten von Krankenkassen – anders als die krankenkassenübergreifenden, für die Bundesrepublik und einzelne Regionen repräsentativen Daten der Beschäftigtenstatistik der BA und der IAB-Beschäftigtenstichprobe – grundsätzlich auf die *Mitglieder der jeweiligen Krankenversicherung* beschränkt. Für Personen, die die Krankenkasse wechseln – eine für ArbeiterInnen vor der Einführung des sogenannten Kassenwahlrechts zum 1. Januar 1996 nur selten gegebene Option –, liegen in GKV-Daten keine vollständigen Angaben zum Erwerbsverlauf vor. Ebenso sind die Erwerbsverläufe von Personen, die vorübergehend gar nicht Mitglied einer gesetzlichen Krankenversicherung sind – dazu zählen auch mitversicherte Familienangehörige –, in den Versichertendaten der Krankenkassen – in Analogie zu den Erwerbsverläufen vorübergehend nicht sozialversicherungspflichtig Beschäftigter in der IAB-Beschäftigtenstichprobe – nur unvollständig dokumentiert. Verläßt ein Mitglied die Krankenkasse, gibt es zumeist auch keine Informationen über die Gründe des Austritts, d.h. es ist aus den Versichertendaten in der Regel nicht zu erkennen, ob (z.B. als Folge eines Arbeitgeberwechsels oder wegen Überschreitung der Beitragsbemessungsgrenze) ein Krankenkassenwechsel oder ein Wechsel in der regionalen Zuständigkeit erfolgt ist, ein Mitglied in den Status der Mitversicherung wechselte oder der Krankenversicherungsschutz sogar völlig verlorenging.

Das hat zur Folge, daß ein Teil der in den Versichertendaten der Krankenkassen, aber auch der in den Mikrodaten der Be-

schäftigtenstatistik der BA und in der IAB-Beschäftigtenstichprobe abgebildeten Erwerbsverläufe in der Frühphase bzw. am Ende abgeschnitten ist (*Links- bzw. Rechtszensierung*) oder *undokumentierte Zeiträume* aufweist. Während zensierte Verläufe und kurze undokumentierte Zeiträume für viele Fragen der Erwerbsverlaufsforschung durchaus hinnehmbar sind, sind Krankenkassendaten für die Bearbeitung anderer Fragestellungen, die auch oder gerade eindeutige Informationen über *Phasen ohne Sozialversicherungspflicht* erfordern, aufgrund dieser Einschränkung meistens nicht geeignet. Das trifft z.B. für Untersuchungen, die auf prekäre Beschäftigungsverhältnisse gerichtet sind, und vor allem für viele Studien über die Erwerbs- und Lebensverläufe von Frauen zu.

Im folgenden wird auf eine Reihe von Ereignissen im und Episoden des Erwerbsverlaufs, die in den Versichertendaten der gesetzlichen Krankenversicherungen abgebildet sind, näher eingegangen. Es sollen Stärken und Schwächen der jeweiligen Angaben diskutiert, Auswertungsmöglichkeiten und -grenzen aufgezeigt und Erfahrungen berichtet werden, die vor allem auf Auswertungen eines Datensatzes der AOK „Küstenstadt“ aus der zweiten Hälfte der 70er Jahre beruhen. Zur Veranschaulichung werden ergänzend einige empirische Forschungsergebnisse, die auf der Grundlage des AOK-Datensatzes „Küstenstadt“ gewonnen wurden, exemplarisch dargestellt. Obgleich sich die Ausführungen auf GKV-Daten beziehen, gelten viele Aussagen zu Möglichkeiten und Problemen allgemein für personenbezogene Sozialversicherungsdaten und treffen somit in analoger Weise auch für die Mikrodaten der Beschäftigtenstatistik der BA und die IAB-Beschäftigtenstichprobe als Datenquellen der Erwerbsverlaufsforschung zu.

## 2.2 Episoden und Ereignisse

*Sozialversicherungspflichtige Beschäftigungsverhältnisse*, zu denen auch – über die Angabe zur Stellung im Beruf (siehe

dazu Abschnitt 4.2) identifizierbare – Ausbildungsverhältnisse zählen, werden in den Versichertendaten der Krankenkassen für den gesamten Zeitraum der Mitgliedschaft in der Regel taggenau erfasst.<sup>1</sup> Nur wenn ein Versicherter bzw. eine Versicherte während eines fortbestehenden Beschäftigungsverhältnisses die Krankenkasse wechselt – eine seit der Einführung des sogenannten Kassenwahlrechts zum 1. Januar 1996 jeweils zum Jahresende für alle ArbeitnehmerInnen gegebene Option –, können Beschäftigungsepisoden in GKV-Daten links- oder rechtszensiert sein. Die Validität der Angaben zum Beginn und zum Ende (und damit zur Dauer) sozialversicherungspflichtiger Beschäftigungsverhältnisse ist aufgrund ihrer unmittelbaren Bedeutung für die Beitragszahlung deutlich höher einzuschätzen als bei retrospektiven Erhebungen, bei denen Erinnerungslücken der Befragten in Betracht gezogen werden müssen. Auch zeitliche Überschneidungen zwischen Arbeitgebermeldungen oder zwischen Arbeitgebermeldungen und sonstigen Meldesätzen sind in der Regel nicht in Datenfehlern begründet, sondern verweisen häufig auf zeitlich parallele Episoden im Erwerbsverlauf der betreffenden Versicherten.<sup>2</sup> So handelt es sich bei sich überschneidenden Meldungen verschiedener Arbeitgeber in den meisten Fällen um *Mehrfachbeschäftigungen*, über die in Krankenkassendaten ebenfalls Informationen vorliegen. *Beschäftigungsverhältnisse ohne Sozialversicherungspflicht* sowie Umwandlungen sozialversicherungspflichtiger in nicht sozialversicherungspflichtige Beschäftigungsverhältnisse (oder umgekehrt) sind in GKV-Daten dagegen nicht dokumentiert.

Neben dem Beginn und dem Ende einer Beschäftigung werden der gesetzlichen Krankenversicherung vom Arbeitgeber auch das beitragspflichtige Bruttoarbeitsentgelt, die Beitragsgruppe zur Kranken- und Rentenversicherung und zur BA sowie ausgeübte Tätigkeit, Stellung im Beruf und Ausbildung „nach den Verhältnissen im Zeitpunkt der Abgabe der Meldung“ (§ 8 (1) DEVO) mitgeteilt. Die Validität und Aktualität der Angaben zum Arbeitsentgelt, die die Grundlage für die Berechnung der Krankenversicherungsbeiträge bilden (siehe z.B. die Stellungnahme von W. Thiele in Hernberg u.a. (Hg.) 1986: 51), und zu den Beitragsgruppen, die eine Unterscheidung zwischen ArbeiterInnen, Angestellten und freiwillig Versicherten erlauben, wird im allgemeinen als hoch eingeschätzt. Bei den Angaben zur Ausbildung (siehe dazu Abschnitt 4.2), die für die Krankenkassen ohne Relevanz sind, erscheinen uns dagegen Zweifel an der Validität angebracht (siehe dazu auch Voges 1994: 115). Ferner gibt es für alle Angaben zur beruflichen Tätigkeit keine Fortschreibungsrichtlinien (siehe z.B. Köster/Finger 1992: 253), so daß der Umfang, in dem entsprechende Änderungen während eines Beschäftigungsverhältnisses der Krankenkasse überhaupt gemeldet werden, und das Ausmaß zeitlich verschobener Änderungsmeldungen hier nicht abschließend beurteilt werden können.

*Arbeitgeberwechsel* sind in den Versichertendaten der gesetzlichen Krankenversicherungen in den meisten Fällen ebenfalls taggenau dokumentiert. Sie können auf der Grundlage des Datenbestandes einer Krankenkasse allerdings dann nicht erfasst werden, wenn Wechsel in und aus Betrieben unbedingt zu einem Krankenkassenwechsel führen (seit Einführung des sogenannten Kassenwahlrechts zum 1. Januar 1996 noch bei einigen Betriebskrankenkassen, vor 1996 häufig z.B. auch bei Innungskrankenkassen), einen Wechsel in der regionalen Zuständigkeit der Krankenkasse zur Folge haben (z.B. bei AOK-Versicherten) oder wenn Versicherte im Zusammenhang mit einem Arbeitgeberwechsel von sich aus die Krankenkasse wechseln. Eine Überschätzung der Häufigkeit von Arbeitgeberwechseln ist ohne Rückgriff auf er-

<sup>1</sup> Stefan Bender und Jürgen Hilzendege (1995: 78f.) sowie Helmut Rudolph (1995: 12) weisen unter Bezug auf die Beschäftigtenstatistik der BA auf die „Sonderrolle“ von Unterbrechungen sozialversicherungspflichtiger Beschäftigungsverhältnisse hin, bei denen diese rechtlich zwar weiter bestehen, faktisch aber die Arbeit ruht und auch kein Entgelt gezahlt und keine Beiträge zur Sozialversicherung entrichtet werden. Das ist z.B. bei Wehr- oder Zivildienst, Krankheit nach Ende der Lohnfortzahlung oder Mutterschafts- und Erziehungsurlaub der Fall (für eine vollständige Auflistung aller Unterbrechungstatbestände siehe Rudolph 1995: 17, Fn.18). Beginn und Ende von Unterbrechungen mit einer Dauer von mindestens einem Monat, die die Sozialversicherungspflicht nicht berühren, müssen den gesetzlichen Krankenversicherungen gemeldet werden und sind über die Angabe zum Abgabegrund einer Meldung in der Regel identifizierbar (zur Ausnahme bei Unterbrechungen zum Jahresende siehe unter Bezug auf die Beschäftigtenstatistik der BA Wermter/Cramer 1988: 478). Wenn Beschäftigungsverhältnisse während einer Unterbrechung enden, ist allerdings zu bezweifeln, daß tatsächlich immer eine Abmeldung bei der Krankenkasse erfolgt (siehe Rudolph 1995: 12, 18). Je nach Differenziertheit des Beitragsgruppenschlüssels können in GKV-Daten über Angaben zum Unterbrechungszeitraum hinaus auch Informationen über den Unterbrechungstatbestand vorliegen. Wie Unterbrechungen bei der Auswertung von GKV-Daten zu bewerten und zu behandeln sind, hängt zum einen von der jeweiligen Fragestellung, zum anderen vom Unterbrechungstatbestand und von der Dauer einer Unterbrechung ab.

<sup>2</sup> Daneben ist bei sich zeitlich überschneidenden Meldungen und ebenso bei undokumentierten Zeiträumen zwischen zwei Meldungen in Einzelfällen allerdings auch die Möglichkeit unpräziser Datumsangaben oder sogar von Datenfehlern in Betracht zu ziehen. Wie sich überschneidende Meldungen (zu Überschneidungen zwischen Meldungen des Arbeitsamtes und nachfolgenden Arbeitgebermeldungen siehe im Text weiter unten) und „Lücken“ zwischen zwei Meldungen zu bewerten und zu behandeln sind, hängt zum einen von der Art der betroffenen Meldungen und der Dauer der Überschneidung bzw. des undokumentierten Zeitraums, zum anderen aber auch von der jeweiligen Fragestellung ab. So sind neben inhaltlich plausiblen Überschneidungen und „Lücken“ für die meisten Fragestellungen der Erwerbsverlaufsforschung z.B. auch Überschneidungen und undokumentierte Zeiträume mit einer Dauer von nur einem oder wenigen Tagen durchaus hinnehmbar.

gänzende Informationen zu den Versichertendaten z.B. dann möglich, wenn einzelne Betriebsteile als selbständige Betriebsstätten eigene Betriebsnummern aufweisen oder wenn Änderungen, z.B. in der Rechtsform eines Unternehmens, zur Zuweisung einer neuen Betriebsnummer führen (siehe dazu ausführlicher Abschnitt 4.4).

*Berufswechsel* – genauer: Wechsel der „ausgeübten Tätigkeit“ – sind in den Versichertendaten der gesetzlichen Krankenversicherungen für sozialversicherungspflichtig Beschäftigte mindestens dann abgebildet, wenn sie mit einem Arbeitgeberwechsel verbunden sind. In welchem Umfang und wie zeitgenau Angaben zur beruflichen Tätigkeit darüber hinaus auch bei innerbetrieblichen Wechseln aktualisiert werden, ist nicht bekannt. Da die Krankenkassen selbst meist kein Interesse an der Zuverlässigkeit und Aktualität der Angaben haben (siehe z.B. die Stellungnahme von Chr. von Ferber in Hernberg u.a. (Hg.) 1986: 85), kann ihre regelmäßige Fortschreibung allerdings angezweifelt werden. Zu *innerbetrieblichen Arbeitsplatzwechseln*, die nicht mit einem Berufswechsel einhergehen, liegen in GKV-Daten grundsätzlich keine Informationen vor. Dagegen werden *Wechsel zwischen Arbeiter- und Angestelltentätigkeiten* über die Beitragsgruppe zur Rentenversicherung auch innerbetrieblich zuverlässig dokumentiert.

Bei der Auswertung von Berufswechseln auf der Basis von GKV-Daten gilt es allerdings, die Validität und die Differenziertheit der – nach den 3-stelligen Berufsordnungen der BA verschlüsselten – Angaben zur ausgeübten Tätigkeit und damit ihre Aussagekraft für die jeweilige Fragestellung zu berücksichtigen (siehe dazu ausführlicher Schulz u.a. 1994; Schulz u.a. 1996): Zum einen sind viele Berufsordnungen sehr heterogen, so daß aus identischen Angaben nicht auf gleiche oder ähnliche Tätigkeiten geschlossen werden kann. So werden z.B. SozialarbeiterInnen und AltenpflegerInnen in einer Berufsordnung zusammengefaßt. Zum anderen sind viele Berufsordnungen aber auch nicht trennscharf, d.h. unterschiedliche Berufsordnungen schließen ähnliche Tätigkeiten nicht aus und Verwechslungen bei der Verschlüsselung sind durchaus zu erwarten. Als Beispiel können Krankenschwestern/Krankenpfleger sowie KrankenpflegehelferInnen angeführt werden, die im amtlichen Schlüsselverzeichnis der BA unterschiedlichen Berufsordnungen zugeordnet sind. Die Heterogenität der Angaben zur beruflichen Tätigkeit in den Versichertendaten der Krankenkassen hat zur Folge, daß Fragestellungen, die die exakte Unterscheidung auch ähnlicher Tätigkeiten voraussetzen, auf der Basis von GKV-Daten nicht beantwortet werden können. So ist z.B. weitgehend unumstritten, daß Rückschlüsse von den Berufsordnungen auf individuelle berufliche Belastungen oder die Bildung belastungshomogener Gruppen nicht mit der für epidemiologische Studien erforderlichen Genauigkeit möglich sind (siehe z.B. die Stellungnahmen von H.-J. Lange und Chr. von Ferber in Hernberg u.a. (Hg.) 1986: 15, 85). Andere Fragestellungen, die stärker auf die Zugehörigkeit zu oder den Wechsel zwischen Berufsgruppen abheben, können dagegen durchaus auf der Grundlage von Krankenkassendaten beantwortet werden. So können bei hinreichendem Beobachtungszeitraum z.B. typische Berufskarrieren aufgezeigt, Beiträge zur Aufdeckung von *healthy- und unhealthy-worker*-Effekten geleistet oder Hinweise auf berufliche Auffangpositionen gewonnen werden.

<sup>3</sup> Analog wurde bei der Erstellung einer Gesamtdatenfile aus den drei Basisdatenfiles der IAB-Beschäftigtenstichprobe vorgegangen (siehe Bender u.a. 1996: 74).

Ebenso wie sozialversicherungspflichtige Beschäftigungsverhältnisse werden auch *Arbeitslosigkeitsepisoden* – soweit sie einen Anspruch auf Leistungen nach dem Arbeitsförderungsgesetz begründen – und *Phasen des Bezugs einer Lohnersatzleistung nach § 192 SGB V* in den Versichertendaten der gesetzlichen Krankenversicherungen für den gesamten Zeitraum einer Mitgliedschaft in der Regel taggenau erfaßt (zu Einschränkungen, die sich aus der Einführung des sogenannten Kassenwahlrechts zum 1. Januar 1996 ergeben können, siehe oben). Datenaufbereitungsarbeiten am Datensatz der AOK „Küstenstadt“ haben zwar gezeigt, daß Meldungen des Arbeitsamtes sich in einigen Fällen mit nachfolgenden Meldungen eines Arbeitgebers überlappen. Solche Unstimmigkeiten, die oft darauf zurückzuführen sind, daß eine Beschäftigungsaufnahme dem Arbeitsamt erst verspätet angezeigt wurde, können auf der Basis von Längsschnittdaten durch „Abschneiden“ der Meldung des Arbeitsamtes allerdings weitgehend bereinigt werden (zur Bewertung und Behandlung von sich überschneidenden Meldungen und undokumentierten Zeiträumen zwischen zwei Meldungen siehe allgemein Fußnote 2).<sup>3</sup> Grundsätzlich ist davon auszugehen, daß auch die Angaben zum Beginn und zum Ende (und damit zur Dauer) von Arbeitslosigkeitsepisoden in GKV-Daten im Vergleich zu retrospektiven Erhebungen eine größere Validität aufweisen.

Versichertendaten der Krankenkassen sind somit geeignet, um Fragen zum Arbeitslosigkeitsrisiko (siehe z.B. Volkholz/Schwarz 1984: 76ff.; Arrow/Behrens 1990; Behrens u.a. 1992; Andreß 1996), zu Mehrfacharbeitslosigkeiten (siehe z.B. Volkholz/Schwarz 1984: 79ff.) oder zur Wiedereingliederung im Anschluß an eine Arbeitslosigkeit zu bearbeiten. Zum Beispiel haben Johann Behrens u.a. (1992) auf der Basis des Datensatzes der AOK „Küstenstadt“ den Einfluß von Arbeitsunfähigkeit sowie individuellen und betrieblichen Merkmalen auf das Arbeitslosigkeitsrisiko untersucht. Für 20.259 Versicherte, die im Zeitraum vom 1.1.1975 bis zum 30.9.1979 mindestens eine beginnende Episode als versicherungspflichtig beschäftigte ArbeiterInnen aufwiesen, wurde eine multivariate Analyse der Übergangsrate in Arbeitslosigkeit (Cox-Regression) durchgeführt. Ein zentrales Ergebnis der Studie ist, daß Arbeitsunfähigkeit nicht für alle ArbeitnehmerInnen gleichermaßen ein erhöhtes Arbeitslosigkeitsrisiko zur Folge hat, sondern die sozialpolitische Überbrückung krankheitsbedingter Krisen der Erwerbsfähigkeit für einige Beschäftigtengruppen offenbar eher gelingt als für andere: ArbeiterInnen, die bei Betriebseintritt älter als 40 Jahre waren, sowie Teilzeitbeschäftigte wiesen im Zusammenwirken mit häufiger Arbeitsunfähigkeit eine erhöhte Übergangsrate in Arbeitslosigkeit auf. Darüber hinaus zeigte sich zwischen der Arbeitsunfähigkeitsdauer und der Nationalität ein schwacher – allerdings negativer – Interaktionseffekt.

Die Teilnahme von Mitgliedern an *Rehabilitationsmaßnahmen* wird in den Routinedaten der gesetzlichen Krankenversicherungen – entweder in den Versicherten- und/oder in den Leistungsdaten – in der Regel ebenfalls taggenau dokumentiert (zu Einschränkungen, die sich aus der Einführung des sogenannten Kassenwahlrechts zum 1. Januar 1996 ergeben können, siehe oben). Allerdings liegen die Angaben von Krankenkasse zu Krankenkasse offenbar in unterschiedlich differenzierter Form vor. Der Datensatz der AOK „Küstenstadt“ beispielsweise erlaubt eine Unterscheidung zwischen Krankenkassenkuren und Rehabilitationsmaßnahmen anderer Träger, wobei die letztgenannten in berufliche und medizinische Maßnahmen und nur die erstgenannten in Anschlußheilbehandlungen (AHB) und vorbeugende Maßnahmen untergliedert werden können. Eine weitere Differenzie-

rung der sonstigen Leistungsträger (wie z.B. die Bundesversicherungsanstalt für Angestellte (BfA), Landesversicherungsanstalten oder Berufsgenossenschaften) ist dagegen auf der Basis dieses Datensatzes ebensowenig möglich wie eine Unterscheidung von stationären und ambulanten Maßnahmen zur Rehabilitation. Heilbehandlungsdiagnosen sind im Datensatz der AOK „Küstenstadt“ ausschließlich für AHB-Maßnahmen der Krankenversicherung dokumentiert.

Grundsätzlich können auf der Basis von GKV-Daten also sowohl Wege in die Rehabilitation nachgezeichnet und Fragen der rehabilitativen Versorgung beantwortet als auch Beiträge zur Erforschung der Folgen von Rehabilitation geleistet werden (siehe z.B. Volkholz/Schwarz 1984: 122ff.; Behrens/Dreyer-Tümmel 1994: Kap.4.2). So konnten Behrens u.a. im Rahmen eines Projektes zur Ermittlung von Indikatoren der Rehabilitationsbedürftigkeit aus Routinedaten der gesetzlichen Krankenversicherung auf der Basis des Datensatzes der AOK „Küstenstadt“ u.a. zeigen, daß mindestens 45,8% (n=347) der untersuchten FrührentnerInnen der Berentungsjahrgänge 1977/78 in den letzten 24 Monaten vor der Berentung nicht an einer Rehabilitationsmaßnahme teilgenommen hatten, obgleich ein erheblicher Anteil dieser Versichertengruppe auffällige Arbeitsunfähigkeitszeiten aufwies (siehe Behrens/Dreyer-Tümmel 1994: Kap.4.2; dies. 1995). Einschränkungen der Eignung von GKV-Daten ergeben sich immer dann, wenn für eine spezifische Fragestellung erforderliche Angaben im verfügbaren Datensatz nicht oder nicht in ausreichend differenzierter Form vorliegen. Darüber hinaus hat sich im Rahmen der angesprochenen Untersuchung gezeigt, daß die Erwerbsverläufe einiger Versicherter längere undokumentierte Zeiträume aufweisen und die Frage der rehabilitativen Versorgung für diesen – mehrheitlich Frauen umfassenden – Personenkreis nicht mit hinreichender Sicherheit geklärt werden kann (siehe Behrens/Dreyer-Tümmel 1994: 181ff.; zur Bewertung und Behandlung von undokumentierten Zeiträumen zwischen zwei Meldungen und sich überschneidenden Meldungen siehe allgemein Fußnote 2). Prinzipiell könnte ein Teil dieser Einschränkungen durch die Verknüpfung der Daten einer Krankenversicherung mit den Datenbeständen anderer Sozialversicherungsträger, vor allem der Rentenversicherungen, zwar aufgehoben werden; für einen solchen Daten-Linkage müßte – neben der mit erheblichem Aufwand verbundenen praktischen Realisierung – allerdings der Datenschutz geklärt sein. Zwar steht mit der IAB-Beschäftigtenstichprobe mittlerweile ein krankenkassenübergreifender Datensatz für Forschungszwecke zur Verfügung. Die erwerbsverlaufsbezogenen Angaben dieses Datensatzes umfassen allerdings keine Rehabilitationsmaßnahmen, sondern sind – wie oben dargestellt – auf sozialversicherungspflichtige Beschäftigungsverhältnisse und Arbeitslosigkeits-episoden beschränkt.

*Berentungen sowie Zeiten des Rentenbezugs* (aus einer gesetzlichen Rentenversicherung) werden in den Versicherten-daten der gesetzlichen Krankenversicherungen ebenfalls tag-genau erfaßt. Die Ermittlung des Berentungsdatums auf der Grundlage von GKV-Daten setzt allerdings voraus, daß eine Person bereits vor Beginn des Rentenbezugs als Mitglied bei der betreffenden Krankenkasse versichert war (zur Bewertung und Behandlung von undokumentierten Zeiträumen zwischen zwei Meldungen und sich überschneidenden Meldungen siehe allgemein Fußnote 2). Weitere Angaben zur Rente liegen – ebenso wie jene zu Rehabilitationsmaßnahmen – von Krankenkasse zu Krankenkasse offenbar in unterschiedlich differenzierter Form vor. So wird nur teilweise zwischen Rentenarten – d.h. zwischen eigener und Hinterbliebenenrente einerseits und zwischen *Renten wegen Berufs- oder Er-*

*werbsunfähigkeit* und *Altersruhegeld* andererseits – unterschieden. Auch der Rentenversicherungsträger wird in den Versicherten-daten für die Zeit des Rentenbezugs anscheinend nicht bei allen Krankenkassen dokumentiert. Allerdings lassen sich fehlende Informationen zum Rentenversicherungsträger und zur Rentenart in vielen Fällen zumindest näherungsweise aus anderen Angaben in den Versicherten-daten indirekt erschließen. So erlauben die Angaben zur Beitragsgruppe zur Rentenversicherung in der letzten Phase des Erwerbslebens Rückschlüsse auf den Rentenversicherungsträger zur Zeit des Rentenbezugs. Die Rentenart kann – sofern in den Daten zwischen eigener und Hinterbliebenenrente unterschieden wird – unter Rückgriff auf gesetzliche Regelungen der jeweiligen Anspruchsvoraussetzungen für viele Versicherte näherungsweise aus den Angaben zum Berentungs- und Geburtsdatum, zum Geschlecht und zum bisherigen Erwerbsverlauf abgeleitet werden. So haben Behrens u.a. im Rahmen ihres Projektes zur Ermittlung von Indikatoren der Rehabilitationsbedürftigkeit aus Routinedaten der gesetzlichen Krankenversicherung auf der Grundlage von Berentungsdatum, Geburtsjahr und Geschlecht sowie Angaben zu Arbeitslosigkeiten und undokumentierten Zeiträumen in den 18 Monaten vor der Berentung unter Inkaufnahme einer Restgruppe zwischen Frühberentungen aus eindeutig gesundheitlichen Gründen und Altersberentungen differenziert (siehe Behrens/Dreyer-Tümmel 1994: 19f.). Darüber hinaus wäre auch zur Präzisierung und Erweiterung der Angaben zu Berentungen prinzipiell eine Verknüpfung von Versichertendaten der Kranken- und der Rentenversicherung (zu Daten der gesetzlichen Rentenversicherungen siehe Leibing/Müller-Späth 1981) denkbar. Auf entsprechende Hinderungsgründe wurde im Zusammenhang mit den Ausführungen zu Rehabilitationsmaßnahmen aber bereits hingewiesen. Dagegen stellt die IAB-Beschäftigtenstichprobe auch für Untersuchungen, die Angaben zu Berentungen oder Zeiten des Rentenbezugs erfordern, keine Alternative zu GKV-Daten dar.

Auf der Basis von Krankenkassendaten ist es also sowohl möglich, Erwerbsverläufe vor der Alters- und Frühberentung zu rekonstruieren, als auch, Bestimmungsfaktoren der vorzeitigen Berentung zu ermitteln (siehe z.B. Volkholz/Schwarz 1984: 107ff.; Behrens/Voges 1990; Voges 1994; Behrens/Dreyer-Tümmel 1994: Kap.4.1; dies. 1995). Das Aufzeigen typischer Wege in die Früh- und Altersrente kann – wie bereits ausgeführt – zum Verständnis der Berentung als Prozeß beitragen und Hinweise auf Steuerungs- und Präventionsmöglichkeiten liefern. So wäre es z.B. möglich, Bedingungen aufzuzeigen, unter denen Abstiegskarrieren unterbrochen und Auffangpositionen erreicht werden, die trotz gesundheitlicher Beeinträchtigungen eine Erwerbstätigkeit bis zum Erreichen der regulären Altersgrenze erlauben. Untersuchungen zu Bestimmungsfaktoren der Frühberentung müssen – hinreichend lange Beobachtungszeiträume vorausgesetzt – bei einem Rückgriff auf individuelle Verlaufsdaten der gesetzlichen Krankenversicherung nicht auf Querschnittvariablen und Angaben über die letzte Phase des Erwerbslebens (z.B. die letzte berufliche Tätigkeit vor der Berentung) beschränkt werden. Statt dessen können auch Informationen über frühere Phasen des Erwerbsverlaufs wie z.B. Beschäftigungsdauern, Berufswechsel, Arbeitslosigkeiten oder auch Arbeitsunfähigkeitszeiten bei den Analysen berücksichtigt werden. So haben Johann Behrens und Wolfgang Voges (1990) im Rahmen einer explorativen Untersuchung zu Bestimmungsfaktoren der Frühberentung auf der Basis des Datensatzes der AOK „Küstenstadt“ für die Berentungsjahrgänge 1975 bis 1978 (n=5.814) berufliche, gesundheitliche und sozialpolitische Ereignisse in den letzten 24 Monaten vor der Berentung in die

Analyse einbezogen. Ihre Labilisierungsthese, die zur Erklärung der Frühberentung eine Wechselwirkung zwischen Arbeitsunfähigkeits- und labilisiertem Berufsverlauf unterstellt, erwies sich als mit den Daten vereinbar.

Die in den Versichertendaten der gesetzlichen Krankenversicherungen abgebildeten Ereignisse im und Episoden des Erwerbsverlaufs können – wie bereits angedeutet – selbstverständlich nicht nur als solche betrachtet, sondern bei Vorliegen eines hinreichend langen Beobachtungszeitraums darüber hinaus auch zu *Statussequenzen* und *Ereignisketten* verknüpft werden. Mithin ist es möglich, auf der Basis von GKV-Daten für eine große Zahl von Versicherten, die kontinuierlich oder zumindest längerfristig Mitglieder der jeweiligen Krankenkasse sind – und damit für Männer häufiger als für Frauen –, *Sequenzen des Erwerbslebens* oder sogar ganze *Erwerbsverläufe* zu rekonstruieren und somit nicht nur Einzelereignisse zu analysieren, sondern auch Abfolgen von (gleichen oder unterschiedlichen) Ereignissen und Episoden sowie komplexe Erwerbskarrieren auswertbar zu machen. Erst auf dieser Basis können z.B. Zusammenhänge zwischen Ereignissen und damit Risikokumulationen im Erwerbsverlauf aufgezeigt, Prozessstrukturen erkannt oder Fehlinterpretationen, die auf die Nichtberücksichtigung individueller Mobilität zurückzuführen sind, aufgedeckt und vermieden werden. Während allerdings für die Analyse einzelner Übergänge mit den seit Ende der 60er Jahre verbreiteten Verfahren der Ereignisanalyse, die eigens auf die spezifische Struktur individueller Verlaufsdaten zugeschnitten sind, ein ganzes Bündel geeigneter Auswertungsmethoden zur Verfügung steht, steht die Entwicklung von Verfahren, die Abfolgen von Ereignissen und Episoden berücksichtigen und der Komplexität von Verläufen auch für große Fallzahlen und viele Variablen gerecht werden, erst am Anfang. So sind Klassifikationsverfahren wie die Korrespondenz- oder die Clusteranalyse unter Bezug auf Längsschnittdaten bisher allenfalls im Ansatz entwickelt.

### **3 Erwerbsverläufe und Arbeitsunfähigkeit: Sozialmedizinische Erwerbsverlaufsforschung mit Krankenkassendaten**

Die besondere Stärke der Erwerbsverlaufsforschung mit Krankenkassendaten liegt bei sozialmedizinischen Fragestellungen und in der Gesundheitsberichterstattung. Anders als in der Beschäftigtenstatistik der BA und der darauf beruhenden IAB-Beschäftigtenstichprobe werden neben umfassenden Informationen zum Erwerbsverlauf in den Leistungsdaten der gesetzlichen Krankenversicherungen nämlich auch Krankenhausaufenthalte, Kuren und Rehabilitationsmaßnahmen (siehe dazu Abschnitt 2) sowie „Krankschreibungen“ taggenau dokumentiert. Mithin liegen für alle Mitglieder einer Krankenkasse Informationen über die Dauer und Häufigkeit von Aufenthalten in Krankenhäusern sowie Kur- und Rehabilitationskliniken und für sozialversicherungspflichtig Beschäftigte sowie – mit Einschränkungen (siehe dazu unten) – für andere Pflichtversicherte darüber hinaus auch verlaufsbezogene Angaben zu Arbeitsunfähigkeiten (AU) und den entsprechenden (Erst-)Diagnosen vor. Personenbezogene Angaben über erbrachte Leistungen wie z.B. ambulante ärztliche Behandlungen und Verordnungen von Medikamenten sowie Heil- und Hilfsmitteln werden von den Krankenkassen gegenwärtig (Stand: 1996) dagegen in der Regel nur in Form von Originaldokumenten wie Kranken- und Überweisungsscheinen oder Rezepten für eine begrenzte Zeit aufbewahrt (siehe dazu im einzelnen Schach 1981) und stehen damit für Auswertungszwecke bisher meist weder auf magnetischen Datenträgern noch im Längsschnitt zur Verfügung.

Auf der Basis von Leistungsdaten der gesetzlichen Krankenversicherungen ist es nicht nur möglich, Arbeitsunfähigkeitsverläufe zu rekonstruieren und AU-Muster zu erkennen. Vielmehr können *Erwerbs- und Arbeitsunfähigkeitsverläufe (einschließlich AU-Diagnosen)* auch *parallel betrachtet und zueinander in Beziehung gesetzt* werden. So wird es möglich, den (statistischen) Effekt von Arbeitsunfähigkeiten und Diagnosen auf Übergangsrisiken und Erwerbsverläufe (d.h. Krankheitsfolgen) oder – umgekehrt – gesundheitliche Folgen von Ereignissen im und Episoden des Erwerbsverlaufs (d.h. Krankheitsursachen) abzuschätzen sowie entsprechende Wechselwirkungen aufzuzeigen. So konnten z.B. Volker Volkholz und Fred Schwarz (1984: 81ff., 88ff.), Jairo Oka Arrow und Johann Behrens (1990), Hans-Jürgen Andreß (1996) und – wie in Abschnitt 2.2 exemplarisch dargestellt – Behrens u.a. (1992) auf der Basis von GKV-Daten zeigen, daß Arbeitsunfähigkeiten das Arbeitslosigkeitsrisiko zumindest für einige Beschäftigtengruppen erhöhen.

Die *gemeinsame Analyse von Erwerbs- und Arbeitsunfähigkeitsverläufen* bildet auch die Voraussetzung dafür, Risikokumulationen – wie z.B. das Zusammenwirken von Arbeitsunfähigkeiten und Arbeitslosigkeitsepisoden bei Invalidisierungsprozessen – oder die Verdichtung von Krisen im individuellen Verlauf aufzuzeigen und krankheitsbezogene Labilisierungs-, aber auch Stabilisierungsprozesse in ihrer jeweiligen Bedingtheit zu erkennen. Ebenso setzen z.B. die Kontrolle von Mobilitätsprozessen, die sogenannte *healthy- und unhealthy-worker*-Effekte zur Folge haben, bei der Analyse des Zusammenhangs von Krankheit und Beruf oder das Aufzeigen betrieblicher Strategien der Externalisierung von Krankheitskosten eine Verknüpfung von Arbeitsunfähigkeits- und Erwerbsverläufen voraus.

Bei der Bearbeitung von Fragen der sozialmedizinischen Erwerbsverlaufsforschung und der Gesundheitsberichterstattung mit Krankenkassendaten gilt es allerdings, nicht nur die Stärken und Schwächen erwerbsverlaufsbezogener Angaben in den Versichertendaten (siehe dazu Abschnitt 2), sondern auch den – vor allem in ihrem Entstehungskontext und ihrem Erhebungszweck begründeten – Gehalt und die Validität der Angaben zu AU-Zeiten und AU-Diagnosen zu beachten. Entsprechende Aspekte, die für die sozialmedizinische Erwerbsverlaufsforschung und die Gesundheitsberichterstattung von zentraler Bedeutung sind, sollen im folgenden kurz dargestellt und in ihren Konsequenzen für die Bearbeitung verschiedener Fragestellungen mit GKV-Daten diskutiert werden.

#### **3.1 Arbeitsunfähigkeiten**

*Arbeitsunfähigkeitsmeldungen* an die Krankenkasse – und damit Angaben zu AU-Zeiten und AU-Diagnosen – liegen in der Regel nur für sozialversicherungspflichtige Mitglieder während der Erwerbsphase (einschließlich freiwillig Versicherter nach Überschreitung der Beitragsbemessungsgrenze) vor, nicht jedoch für sonstige freiwillige Mitglieder (z.B. Selbständige oder BeamtInnen), mitversicherte Familienangehörige und RentnerInnen. Auch für Pflichtversicherte im erwerbsfähigen Alter werden insbesondere kürzere Arbeitsunfähigkeiten für Perioden ohne sozialversicherungspflichtige Beschäftigung, z.B. für Arbeitslosigkeitsepisoden, in den Leistungsdaten der gesetzlichen Krankenversicherungen offenbar nur unvollständig erfaßt (siehe z.B. Volkholz/Schwarz 1984: 110; Andreß 1996: 249f.). So haben Behrens u.a. im Rahmen ihres Projektes zur Ermittlung von Indikatoren der Rehabilitationsbedürftigkeit aus Routinedaten der gesetzlichen Krankenversicherung am Datensatz der AOK „Küstenstadt“ unter Bezug auf die letzten Jahre vor der Berentung ge-

zeigt, daß längerfristig Arbeitslose im Vergleich zu krankenversicherungspflichtig Beschäftigten seltener AU-Meldungen aufweisen (Behrens/Dreyer-Tümmel 1994: 111f.). Für Zeiten sozialversicherungspflichtiger Beschäftigung liegen in GKV-Daten dagegen vollständige Angaben zu „Krankschreibungen“ und damit verlaufsbezogene Informationen zumindest über jene Arbeitsunfähigkeitsepisoden vor, die eine Dauer von (in der Regel) mehr als zwei Tagen haben.

Die Validität der Angaben zum Beginn und zum Ende von „Krankschreibungen“, die nicht nur den Zeitraum legitimer Arbeitsruhe für den Versicherten bestimmen, sondern darüber hinaus unmittelbar die Pflicht der Krankenversicherung zur Zahlung von Krankengeld berühren, ist als hoch einzuschätzen. In den Leistungsdaten der gesetzlichen Krankenversicherungen liegen also zumindest für Zeiten sozialversicherungspflichtiger Beschäftigung verlaufsbezogene Informationen über das *Betroffensein von (ärztlich bescheinigter) Arbeitsunfähigkeit* sowie über *AU-Dauern, AU-Häufigkeiten*<sup>4</sup> und *AU-Sequenzen* vor, deren Grad an Zuverlässigkeit auf der Basis von (retrospektiven) Befragungen wohl kaum erreicht werden kann.

Allerdings dürfen *Arbeitsunfähigkeitsmeldungen an die gesetzliche Krankenversicherung* – trotz (z.T. erheblicher) Überschneidungen der Konzepte – weder mit *Fehlzeiten*, die auch nicht ärztlich bescheinigte Arbeitsunfähigkeiten mit einer Dauer von bis zu zwei Tagen, Fehlzeiten aus anderen legitimierte Gründen sowie Absentismus umfassen, noch mit *subjektivem Krankheitsempfinden* oder mit dem Vorhandensein von *Krankheit im engeren medizinischen Sinne* gleichgesetzt werden. So ist z.B. wiederholt darauf hingewiesen worden, daß Krankheit im medizinischen Sinne nicht notwendig zu einer Arbeitsunfähigkeit oder zu einer „Krankschreibung“ führt, medizinisch begründete Arbeitsunfähigkeit – als die „(f)ür viele Bevölkerungsgruppen und Arbeitnehmersituationen ... einzige gesellschaftliche legalisierte Möglichkeit der Arbeitsruhe“ (W. Thiele in einer Stellungnahme in Hernberg u.a. (Hg.) 1986: 52) – andererseits aber auch zur Rechtfertigung von Fehlzeiten aus anderen, zumindest nicht primär krankheitsbezogenen Gründen herangezogen wird (z.B. Schwefel/Schwartz 1978: 11f.). In diesem Kontext ist zu bedenken, daß die Arbeitsfähigkeit einer Person nicht allein durch ihre somatische und psychische Verfassung determiniert ist, sondern zum einen maßgeblich durch die Art der beruflichen Tätigkeit und die Bedingungen am Arbeitsplatz, daneben aber auch durch ihr subjektives Wohlbefinden und ihren sozialen Kontext mit bestimmt wird, der z.B. Kompensationsmöglichkeiten bereithalten oder zusätzliche Belastungen mit sich bringen kann (zur Relativität von Gesundheit siehe auch Behrens 1994: 112f.; Behrens u.a. 1993: 209ff.). Bei „Krankschreibungen“ sind darüber hinaus auch das – durch Faktoren wie die konjunkturelle Lage, Arbeitsplatzsicherheit und kulturelle wie soziale Milieus beeinflusste – Krankheits- und Krankschreibungsverhalten des bzw. der Versicherten sowie der Interaktions- und Aushandlungsprozeß zwischen Arzt bzw. Ärztin und PatientIn von Bedeutung, die zu einer gegenseitig „hingenommene(n) oder ak-

zeptierte(n) Deutung eines Schmerz- oder Leidenszustands“ gelangen müssen, „die der Gesellschaft und insbesondere den einschlägigen Kontrollorganen gegenüber legitimierbar ist“ (Behrens/Schmidt-Ohlemann 1986: 129).

Bei Arbeitsunfähigkeitsmeldungen an die Krankenkassen handelt es sich also nicht um Angaben über das Vorhandensein von Krankheit im engeren medizinischen Sinne, sondern – mit den genannten Vorbehalten – um Informationen über Zeiten ärztlich bescheinigter und in der Regel subjektiv empfundener, (überwiegend) mit Krankheit verbundener Einschränkungen der Arbeits- und Leistungsfähigkeit von Versicherten in bestimmten Kontexten und Konstellationen. Wenn es vorrangig um diesen Aspekt von Kranksein geht, sind GKV-Daten für die Bearbeitung sozialmedizinischer Fragestellungen mithin gut geeignet. Das ist insbesondere bei Untersuchungen, die Arbeitsunfähigkeit als „maßnahmenauslösendes Ereignis“ (Behrens/Schmidt-Ohlemann 1986: 120) thematisieren und sich z.B. mit den Folgen von AU-Zeiten für die Wahrscheinlichkeit von krisenhaften oder stabilisierenden Statuspassagen und für Erwerbskarrieren befassen, aber auch bei vielen Analysen zu Risikokumulationen und zur Verdichtung krisenhafter Ereignisse im Erwerbsverlauf sowie zu Labilisierungs- und Invalidisierungsprozessen der Fall. Steht dagegen – z.B. bei (nicht ausschließlich deskriptiven) Untersuchungen zu den gesundheitlichen Folgen beruflicher Tätigkeiten – das Vorhandensein von Krankheit im engeren medizinischen Sinne oder die Schwere von Krankheit im Mittelpunkt des Interesses, ist die Aussagekraft der Angaben zu AU-Zeiten in den Leistungsdaten der gesetzlichen Krankenversicherungen deutlich beschränkt.

### 3.2 Arbeitsunfähigkeitsdiagnosen

Arbeitsunfähigkeitsdiagnosen liegen in den Leistungsdaten der gesetzlichen Krankenversicherungen in Form des 3-stelligen Schlüssels der Internationalen Klassifikation der Krankheiten, Verletzungen und Todesursachen (ICD) vor. Ihre Eignung für die sozialmedizinische und vor allem für die epidemiologische Forschung ist umstritten. Kritik an der Aussagekraft und der Validität von AU-Diagnosen bezieht sich vor allem auf zwei Ebenen: auf den Entstehungsprozeß von AU-Bescheinigungen (siehe dazu oben) und AU-Diagnosen einschließlich der Möglichkeit ärztlicher Fehldiagnosen einerseits, auf den mehrstufigen Dokumentations- und Kodifizierungsprozeß der Diagnosen andererseits. In dessen Verlauf können nicht nur Übertragungs- und Kodierfehler auftreten (siehe erläuternd auch von Ferber u.a. 1982: 843f.). Vielmehr können auch die immanenten Schwächen der ICD – z.B. Einschränkungen hinsichtlich der Kodierung topographischer Zuordnungen - sowie fehlende Möglichkeiten, bei der Dokumentation zwischen Verdachts- und validierten Diagnosen oder zwischen dem Neuaufreten und dem Akutwerden sowie zwischen Schweregraden von Erkrankungen zu unterscheiden oder Mehrfachdiagnosen anzugeben, zu Informationsverlusten führen (siehe ausführlicher Dorenburg 1990; ders. 1991: Kap.6).

Als relativ unumstritten kann vor diesem Hintergrund gelten, daß AU-Diagnosen keine verlässlichen Rückschlüsse auf klinisch-nosologische oder morphologische Krankheitsbilder erlauben und Krankengeschichten i.e.S. – im Gegensatz zu AU-Verläufen – auf der Basis von Arbeitsunfähigkeitsmeldungen in der Regel nicht rekonstruierbar sind (siehe z.B. Dorenburg 1990). Auch Aussagen darüber, „ob ein Individuum eine ganz bestimmte Krankheit zum Zeitpunkt t definitiv hat“ (Dorenburg 1990: 2), dürften allenfalls für einzelne Diagnosen bzw. Diagnosegruppen möglich sein und darüber hin-

<sup>4</sup> Da die Routine vieler ÄrztInnen, Arbeitsunfähigkeit zunächst bis zu einem Freitag oder bis zum letzten Tag vor einem arbeitsfreien Tag zu bescheinigen und dann – wenn die Arbeitsunfähigkeit anhält – vom nächsten Arbeitstag an wieder, dazu führen kann, daß Wochenenden und Feiertage in GKV-Daten als arbeitsfähig verbrachte Tage erscheinen, wurden Arbeitsunfähigkeitsmeldungen, zwischen denen weniger als fünf Tage liegen, bei unseren Auswertungen auf der Basis des Datensatzes der AOK „Küstenstadt“ in der Regel als ein AU-Fall gezählt und die Tage zwischen zwei derartigen AU-Meldungen dann dementsprechend auch der AU-Dauer zugeordnet.

aus einen erheblichen Aufwand für die Validierung durch Hinzuziehung zusätzlicher Informationen – z.B. von Längsschnittangaben zum AU-Verlauf – erforderlich machen (siehe dazu Dorenburg 1990; ders. 1991: Kap.14). Falsch-negative Zuordnungen von Versicherten sind aber selbst unter diesen Voraussetzungen voraussichtlich nicht immer zu erkennen. Auf der Basis von AU-Meldungen zufriedenstellend beantwortet werden kann dagegen – zumindest in Bezug auf (für die jeweilige Fragestellung) geeignete *Diagnosegruppen* – die Frage, „ob ein Individuum mit hinreichender Wahrscheinlichkeit zu einer Gruppe gehört, welche in einem bestimmten Zeitraum spezielle gesundheitliche Probleme hatte, die sich entweder auf eine funktionelle Einheit ... oder auf eine Gruppe ähnlich gearteter Erkrankungen beziehen lassen“ (Dorenburg 1990: 2). Eine entsprechende, unter topographischen und funktionellen Gesichtspunkten erstellte und speziell für verschiedene Fragestellungen der sozialmedizinischen Erwerbsverlaufsforschung entwickelte Diagnoseklassifikation auf der Basis der ICD hat Ulrich Dorenburg (1990) vorgelegt. Auch die Validität solcher Aussagen über das Betroffensein von – einem oder mehreren – bestimmten gesundheitlichen Problemen in einem festgelegten Zeitraum kann durch die Wahl einer zweckdienlichen Operationalisierung, die z.B. Längsschnittinformationen über die Häufigkeit und/oder die Dauer von AU-Diagnosen berücksichtigen könnte, weiter abgesichert werden.

Diagnoseorientierte Auswertungen sind auf der Basis von Routinedaten der gesetzlichen Krankenversicherungen also vor allem dann möglich, wenn der (statistische) Effekt von Arbeitsunfähigkeit auf Übergangsrisiken und Statussequenzen im Erwerbsverlauf – z.B. auf das Arbeitslosigkeitsrisiko oder die Wahrscheinlichkeit einer rehabilitativen Versorgung – diagnose(gruppen)spezifisch untersucht werden soll oder wenn die Zuordnung von Versicherten zu Gruppen von Personen, die mit hinreichender Wahrscheinlichkeit von bestimmten, unter topographischen und funktionellen Gesichtspunkten definierten Diagnosegruppen betroffen sind, ausreicht. Letzteres scheint insbesondere bei Analysen der Fall zu sein, die primär auf die Abschätzung von Größenordnungen oder auf die Ermittlung erster Hinweise auf potentielle Zusammenhänge zwischen Krankheit und Erwerbsverlauf – also z.B. auf Hypothesen über arbeitsbedingte Erkrankungen – gerichtet sind. So haben Behrens u.a. im Rahmen ihres Projektes zur Ermittlung von Indikatoren der Rehabilitationsbedürftigkeit aus Routinedaten der gesetzlichen Krankenversicherung auf der Basis des Datensatzes der AOK „Küstenstadt“ geschlechtsspezifisch untersucht, in welchem Umfang 758 Früh- und 787 AltersrentnerInnen der Berentungsjahrgänge 1977/78 im maximal 4jährigen Zeitraum vor der Berentung von bestimmten Diagnosegruppen betroffen waren. Es zeigte sich, daß sowohl die späteren FrührentnerInnen als auch – auf niedrigerem Niveau – die späteren AltersrentnerInnen am häufigsten wegen Herz-Kreislauf- und Muskel-Skelett-Erkrankungen arbeitsunfähig waren, während die größten Unterschiede zwischen späteren Früh- und Altersberenteten – bei deutlich geringeren Fallzahlen – in Bezug auf das Betroffensein von psychischen Erkrankungen, Neubildungen (böartigen oder unbekanntem Charakter) und – vor allem bei Männern – Alkoholabhängigkeit auftraten (Behrens/Dreyer-Tümmel 1994: Kap.4.1; dies. 1995: 431ff.). Die Ergebnisse solcher deskriptiven und explorativen Untersuchungen können dann u.a. genutzt werden, um weiterführende epidemiologische Studien gezielt anzuregen, vorzubereiten und effektiver zu gestalten. Eine quantitative und qualitative Beurteilung von Kausalzusammenhängen, z.B. exakte Risikoabschätzungen oder definitive Aussagen zu arbeitsbe-

dingten Ursachen bestimmter Erkrankungen, sind dagegen auf der Basis von AU-Diagnosen in den Leistungsdaten der gesetzlichen Krankenversicherungen nicht möglich.

#### **4 Unabhängige Variablen und Subgruppen im Kontext der Erwerbsverlaufsforschung: Das Potential in Krankenkassendaten im Überblick**

Die vorangegangenen Abschnitte dieses Beitrags haben gezeigt, daß in den Routinedaten der gesetzlichen Krankenversicherungen ein beachtliches Datenpotential zur Verfügung steht, um Übergänge im Erwerbsverlauf, Statussequenzen und ganze Erwerbskarrieren zu rekonstruieren und zu gesundheitsbezogenen Einschränkungen der Arbeitsfähigkeit sowie zu Arbeitsunfähigkeitsverläufen in Beziehung zu setzen. Auf der Basis dieser Daten ist es möglich, eine Reihe von Fragestellungen insbesondere der sozialmedizinischen Erwerbsverlaufsforschung und der Gesundheitsberichterstattung zu bearbeiten (siehe dazu Abschnitt 3). Zu einem weiteren Themenschwerpunkt der Erwerbsverlaufsforschung, nämlich den Erwerbs- und Lebensverläufen von Frauen, der neben Angaben zum Erwerbsverlauf in der Regel nicht nur umfassende Informationen über die Lebensumstände von Frauen außerhalb der Erwerbsarbeit, sondern auch über Lebensphasen ohne Erwerbstätigkeit und damit meist auch ohne (Pflicht-)Mitgliedschaft in einer gesetzlichen Krankenversicherung – insbesondere über sogenannte Familienphasen – erfordert, kann auf der Basis von GKV-Daten dagegen kein wesentlicher Beitrag geleistet werden. Gleiches gilt auch für Untersuchungen zum Übergang von der Schule in den Beruf, da über Lebensabschnitte von Jugendlichen vor dem Beginn einer betrieblichen Ausbildung oder Berufstätigkeit in den Versichertendaten der Krankenkassen in den meisten Fällen ebenfalls keine Informationen vorliegen.

Zusätzlich zu verlaufsbezogenen Informationen über Ereignisse im und Episoden des Erwerbsverlaufs sowie gegebenenfalls über Arbeitsunfähigkeiten erfordern viele Fragestellungen der Erwerbsverlaufsforschung weitere Angaben zu den untersuchten Individuen sowie ihren Ressourcen und sozialen Kontexten. So sind Ereignisse wie z.B. Arbeitsunfähigkeiten und Arbeitslosigkeiten bei Untersuchungen zur Produktion, Reproduktion und Kumulation sozialer Ungleichheiten im Erwerbsverlauf erst in jüngerer Zeit ins Blickfeld gerückt. Statt dessen waren und daneben sind für die Bearbeitung vieler Fragestellungen der Ungleichheitsforschung Angaben zu verschiedenen sozialstrukturellen und soziodemographischen Merkmalen wie z.B. der sozialen Schicht bzw. Klasse, dem – z.B. durch Bildung, berufliche Stellung oder Einkommen bestimmten – Status, dem Geschlecht, dem Alter oder der Nationalität erforderlich. Auch die sozialmedizinische Erwerbsverlaufsforschung betrachtet Zusammenhänge und Wechselwirkungen zwischen Erwerbsverläufen und gesundheitlichen Beeinträchtigungen in der Regel nicht isoliert, sondern bezieht in viele Analysen weitere potentielle Einflußfaktoren wie z.B. soziodemographische Merkmale, Einstellungen und Verhaltensweisen der Individuen, ihren sozialen Kontext, Merkmale der beschäftigenden Betriebe oder Arbeitsmarktbedingungen mit ein.

Um die Eignung von Routinedaten der gesetzlichen Krankenversicherungen auch für die angedeuteten und weitere Fragestellungen der Erwerbsverlaufsforschung beurteilen zu können, soll abschließend kurz darauf eingegangen werden, in welchem Umfang und mit welchen Stärken und Schwächen in GKV-Daten über die bereits thematisierten Angaben zum Erwerbsverlauf und zu Arbeitsunfähigkeiten hinausgehende

Informationen vorliegen und welche für verschiedene Fragestellungen der Erwerbsverlaufsforschung zentralen Variablen und Variablengruppen in den Versicherten- und Leistungsdaten der Krankenkassen nicht oder allenfalls in unzureichender Form zur Verfügung stehen. Da die diskutierten Angaben zum überwiegenden Teil auf dem integrierten Meldeverfahren zur Sozialversicherung und zur Bundesanstalt für Arbeit beruhen, können viele der Aussagen zu Möglichkeiten und Problemen in analoger Weise auch auf die Beschäftigtenstatistik der BA und die IAB-Beschäftigtenstichprobe übertragen werden.

#### 4.1 Soziodemographische Merkmale

Als soziodemographische Merkmale sind für die Erwerbsverlaufsforschung vor allem Angaben zum Geschlecht, zum Alter und zur Nationalität von Bedeutung. Da sie auf individuelle Entscheidungen sowie Risiken und Chancen im Erwerbsleben und damit auch auf Erwerbsverläufe einen erheblichen Einfluß haben können, stellen sie eine wichtige Grundlage für die Bildung von Subgruppen dar. Angaben zum Alter und zum Geschlecht sind für einige Fragestellungen der Erwerbsverlaufsforschung darüber hinaus wichtig, um die Erfüllung von Anspruchsvoraussetzungen – z.B. für den Bezug einer Altersrente – beurteilen zu können. Bei der Bearbeitung von Fragestellungen, die eine Differenzierung von Früh- und AltersrentnerInnen erfordern, kann die Altersangabe ferner ein wichtiges Hilfsmittel zur indirekten Erschließung der Rentenart sein (siehe dazu Abschnitt 2.2).

Angaben zum *Geschlecht* stehen in Krankenkassendaten für Mitglieder grundsätzlich zur Verfügung. Sie sind als valide zu beurteilen (siehe z.B. Köster/Finger 1992: 252).

Das *Alter* eines Mitglieds kann für jeden beliebigen Zeitpunkt aus den Angaben zum *Geburtsdatum* errechnet werden. Elisabeth Schach und I. Köster und T. Finger weisen zwar darauf hin, „daß zumindest für Teile der ausländischen Bevölkerung Geburtstag und -monat nicht deren tatsächliche Werte wiedergeben“ (Köster/Finger 1992: 252; siehe auch Schach 1981: 207) und bei AusländerInnen, „die nicht in Deutschland oder der EG geboren wurden, ... auch das Geburtsjahr nicht immer als zuverlässig angesehen werden (kann)“ (Köster/Finger 1992: 252). Diese Ungenauigkeit ist für die Erwerbsverlaufsforschung jedoch ohne Belang, da in bezug auf das biologische Alter kleinere Abweichungen hinnehmbar sind und im Hinblick auf die Erfüllung von Anspruchsvoraussetzungen der Sozialversicherung das im dokumentierten Geburtsdatum zum Ausdruck kommende „offizielle“ Alter ausschlaggebend ist. Eine gewisse Einschränkung stellt demgegenüber dar, daß Geburtstag und Geburtsmonat im Zusammenhang mit der Anonymisierung von Versicherungsnummern mitunter überschrieben werden und dann für Auswertungszwecke nicht mehr zur Verfügung stehen.

Die *Nationalität* ihrer Mitglieder ist den Krankenkassen ebenfalls bekannt. Hinsichtlich der Validität der Angaben, die für den administrativen Prozeß der Krankenversicherung keine Bedeutung haben, sind allerdings Zweifel angebracht. Köster und Finger (1992: 252) weisen ferner darauf hin, daß die Zuordnung von Versicherten, die die Staatsbürgerschaft wechseln, – zumindest, so ist zu ergänzen, wenn das alte Datum überschrieben wird – vom jeweiligen Erhebungszeitpunkt abhängt.

Angaben zum *Familienstand* und zur *Kinderzahl* sind im Kontext der Erwerbsverlaufsforschung vor allem als Informationen über den familiären Kontext interessant, der Bela-

stungen mit sich bringen, aber auch Kompensationsmöglichkeiten bereithalten kann. Auf der Basis von Krankenkassendaten kann in der Regel höchstens ermittelt werden, ob ein Mitglied verheiratet ist und für wieviel Kinder der Kasse gegenüber Ansprüche bestehen. Informationen über Kinder, die nicht bei dem jeweiligen Mitglied mitversichert sind, liegen dagegen nicht vor. Für Fragestellungen, die Informationen zum familiären Kontext von Individuen erfordern, sind GKV-Daten in der Regel also nicht geeignet.

Angaben zum (aktuellen) *Wohnort* von Versicherten liegen den Krankenkassen zwar vor, sie stehen für Auswertungszwecke aus datenschutzrechtlichen Gründen aber nicht oder allenfalls in stark aggregierter Form zur Verfügung. Ferner weisen Köster und Finger (1992: 252f.) darauf hin, daß bei Wohnortwechseln zeitliche Verzögerungen bei der Aktualisierung der Daten auftreten können. Da Angaben zum Wohnort von Individuen im Kontext der Erwerbsverlaufsforschung in der Regel nur erforderlich sind, wenn ergänzende Informationen zum regionalen Kontext – z.B. regionalspezifische Angaben zur konjunkturellen Lage – erhoben werden sollen (siehe dazu Abschnitt 4.4), reichen – sofern die betreffende Krankenkasse nicht ohnehin regional organisiert ist – aggregierte Angaben zum Wohnort aus. Eine Einschränkung ist dagegen darin zu sehen, daß Angaben zum vorhergehenden Wohnort bei einem Wohnortwechsel in der Regel überschrieben werden.

#### 4.2 Soziale Schicht, sozialer Status, Einkommen

Angaben zur *sozialen Schicht* (bzw. Klasse) von Individuen sind im Kontext der Erwerbsverlaufsforschung nicht nur als eine zentrale (vertikale) Dimension sozialer Ungleichheit von Interesse. Darüber hinaus sind Informationen zum (*sozialen*) *Status* einer Person auch erforderlich, um sozialen Auf- oder Abstieg im individuellen Verlauf erfassen und damit die „Richtung“ von Karrieren bestimmen zu können. Als Merkmale, die Rückschlüsse auf die soziale Schicht bzw. den Status eines bzw. einer Versicherten erlauben könnten, stehen in den Versichertendaten der gesetzlichen Krankenversicherungen prinzipiell Angaben zum Versichertenstatus, die Beitragsgruppen zur Renten- und zur Krankenversicherung, Angaben zur ausgeübten Tätigkeit, zur Stellung im Beruf und zur Ausbildung sowie Einkommensangaben zur Verfügung.

Angaben zur *ausgeübten Tätigkeit* (siehe dazu Abschnitt 2.2), zur *Stellung im Beruf* und zur *Ausbildung* sind in den Arbeitgebermeldungen an die Krankenkassen enthalten. Hinsichtlich der Stellung im Beruf wird zwischen Auszubildenden (einschließlich Anlernlingen, PraktikantInnen und VolontärInnen), zwei Kategorien teilzeitbeschäftigter ArbeitnehmerInnen sowie vollzeitbeschäftigten ArbeiterInnen, FacharbeiterInnen, MeisterInnen und PolierInnen, Angestellten sowie HeimarbeiterInnen und Hausgewerbetreibenden unterschieden. Bei der Ausbildung wird nach der Art des Schulabschlusses (Volks-/Hauptschule oder mittlere Reife oder gleichwertige Schulbildung, Abitur, Abschluß einer Fachhochschule, Hochschul-/Universitätsabschluß) und bei den beiden erstgenannten Kategorien darüber hinaus nach dem Vorliegen einer abgeschlossenen Berufsausbildung differenziert. Bei allen Angaben kann die regelmäßige Fortschreibung und damit ihre Aktualität im Anschluß an Veränderungen allerdings angezweifelt werden (siehe dazu Abschnitt 2.2). Vor allem in bezug auf die Ausbildung, die für den Verwaltungsprozeß keine Bedeutung hat, sind darüber hinaus auch Zweifel hinsichtlich der Validität der Angaben angebracht.

Die Angaben zum *Versichertenstatus* und zur *Beitragsgruppe* ermöglichen es nicht nur, die Mitglieder einer gesetzlichen Krankenversicherung nach ihrer Position im Erwerbsleben (z.B. sozialversicherungspflichtige Beschäftigung, Arbeitslosigkeit, Rentenbezug) zu untergliedern (siehe dazu Abschnitt 2), sondern darüber hinaus können auch (sozialversicherungspflichtig beschäftigte) *ArbeiterInnen und Angestellte* sowie *pflchtigversicherte und freiwillig versicherte Mitglieder* unterschieden werden. Die Validität dieser Angaben ist als hoch einzuschätzen (siehe auch Abschnitt 2). Die sehr heterogene Kategorie der freiwillig Versicherten, die nicht nur ArbeitnehmerInnen mit einem Gehalt über der Beitragsbemessungsgrenze, FreiberuflerInnen und Selbständige (von TaxifahrerInnen bis zu GroßunternehmerInnen), sondern z.B. auch SozialhilfeempfängerInnen umfassen kann, ist zur Bestimmung des sozialen Status einer Person und zur Erfassung von sozialem Auf- oder Abstieg im individuellen Verlauf aber allenfalls bedingt geeignet. Diese Einschränkung wiegt weniger schwer, wenn der EDV-geführte Datenbestand einer Krankenkasse eine weitergehende Differenzierung der freiwillig Versicherten erlaubt. So kann beispielsweise auf der Grundlage des Datensatzes der AOK „Küstenstadt“ zwischen freiwillig Weiterversicherten nach § 313 RVO und Versicherungsberechtigten nach § 176 RVO (heute: § 9 SGB V) sowie zwischen BeamtenInnen und sonstigen freiwillig Versicherten unterschieden werden.

*Einkommensangaben* in GKV-Daten beziehen sich für rentenversicherungspflichtige Mitglieder auf das versicherungspflichtige Einkommen bis zur Höhe der Beitragsbemessungsgrenze zur Rentenversicherung. Da das (vom Arbeitgeber gemeldete) Entgelt die Grundlage für die Höhe der Krankenversicherungsbeiträge bildet, werden die Angaben im allgemeinen als valide angesehen (siehe z.B. die Stellungnahme von W. Thiele in Hernberg u.a. (Hg.) 1986: 51; Köster/Finger 1992: 253). Für freiwillig versicherte Mitglieder gesetzlicher Krankenversicherungen liegen dagegen Angaben zum Gesamteinkommen vor (siehe z.B. Schach 1981: 208; Köster/Finger 1992: 253), die ebenfalls zur Berechnung der Beitragshöhe herangezogen werden. Da sie in der Regel auf Selbstangaben der Versicherten beruhen, ist ihre Validität wohl etwas geringer einzuschätzen als bei Pflichtversicherten. Für RentnerInnen sind in GKV-Daten keine Einkommensangaben verfügbar. In den Routinedaten der gesetzlichen Krankenversicherungen liegen also für Erwerbspersonen – je nach Versichertengruppe – individuenbezogene Informationen über das (sozialversicherungspflichtige) Bruttoarbeitsentgelt oder über das Gesamteinkommen vor. Ob diese Angaben für die Erwerbsverlaufsforschung geeignet sind oder ob statt dessen Angaben über das verfügbare (Netto-)Einkommen und/oder über das Haushaltseinkommen erforderlich sind, hängt von der jeweiligen Fragestellung ab.

### 4.3 Gründe, Motive, Einstellungen und Verhaltensweisen

Einige Fragestellungen der Erwerbsverlaufsforschung erfordern nicht nur Angaben über Episoden des und Ereignisse im Erwerbsverlauf, sondern darüber hinaus auch Informationen über *Gründe, Motive, Einstellungen* und *Verhaltensweisen*, deren Resultate die Ereignisse und Episoden sein können. So wäre das Verlassen eines Betriebs im Kontext der Krankheitsfolgenforschung möglicherweise nicht nur je nach Zielzustand (z.B. Arbeitgeberwechsel, Übergang in Arbeitslosigkeit oder Berentung) unterschiedlich zu bewerten, sondern auch in Abhängigkeit vom jeweiligen Grund des Arbeitsplatzverlustes, der z.B. in einer Kündigung durch den Arbeitgeber (z.B. aus Krankheitsgründen oder wegen Arbeitsmangel), einer – wie

auch immer motivierten und begründeten – Kündigung seitens des bzw. der Beschäftigten oder auch in einem befristeten Arbeitsvertrag liegen kann. Bei Untersuchungen zu Bestimmungsfaktoren der Inanspruchnahme von Rehabilitationsmaßnahmen können neben Angaben zu gesundheitlichen Faktoren beispielsweise auch Informationen über die Angebotsstruktur und die Informationsvermittlung, d.h. über strukturelle Gründe, und über die Motive und Einstellungen der Versicherten eine wichtige Rolle spielen. Im Zusammenhang mit Stellungnahmen zur Eignung von Krankenkassendaten zur Berufsrisikoforschung ist wiederholt darauf hingewiesen worden, daß neben Angaben zu Belastungen am Arbeitsplatz auch Informationen über Verhaltensweisen außerhalb der Erwerbsarbeit, z.B. über Ernährungsgewohnheiten, Rauch- und Trinkverhalten oder Freizeitsport, erforderlich sind (siehe z.B. die Stellungnahme von H.-J. Lange in Hernberg u.a. (Hg.) 1986: 15f.; anders: Volkholz/Schwarz 1984: 155f.).

Solche Angaben über Gründe, Motive, Einstellungen und Verhaltensweisen sind in den Versicherten- und Leistungsdaten der gesetzlichen Krankenversicherungen – ebenso wie in der Beschäftigtenstatistik der BA und der IAB-Beschäftigtenstichprobe – nicht enthalten. Es kann allenfalls aus den Resultaten auf potentielle Gründe und Motive geschlossen und damit ein Beitrag zur Hypothesenbildung geleistet werden, deren empirische Evidenz dann in weiterführenden Studien zu prüfen wäre.

### 4.4 Arbeitsmarkt, Betrieb, Arbeitsplatz

Arbeitsmärkte, Betriebe und Arbeitsplätze sind zentrale Kontexte von Erwerbsverläufen. Arbeitsmarktfaktoren, betriebliche Merkmale und Politiken sowie die Bedingungen am Arbeitsplatz können relevante Bestimmungsfaktoren für Risiken und Chancen im Erwerbsleben und für den Verlauf von Erwerbskarrieren sein. Sie stellen somit für viele Fragestellungen der Erwerbsverlaufsforschung wichtige Variablen dar.

Die *Betriebsnummer* und die Anschrift von Arbeitgebern der Versicherten werden den Krankenkassen gemeldet. Wenngleich für Forschungszwecke in der Regel nur chiffrierte Betriebsnummern zur Verfügung gestellt werden, sind damit auf der Basis von GKV-Daten neben individuenbezogenen prinzipiell auch *betriebsbezogene Auswertungen* (einschließlich betrieblicher Gesundheitsberichte) möglich. Allerdings kommt es vor, daß einzelne Betriebsteile eigene Betriebsnummern aufweisen, Betrieben – z.B. bei einer Änderung der Rechtsform – im Zeitverlauf eine neue Betriebsnummer zugewiesen wird, Betriebe nach einer räumlichen Verlagerung ihre bisherige Betriebsnummer beibehalten oder Betriebsnummern nach einer Betriebsschließung neu vergeben werden (siehe dazu unter Bezug auf die IAB-Beschäftigtenstichprobe auch Bender/Hilzendegen 1995: 80, 84). Darüber hinaus ist zu beachten, daß in vielen Fällen nicht alle ArbeitnehmerInnen eines Betriebs bei derselben Krankenkasse versichert sind und somit keine Angaben über vollständige Belegschaften vorliegen. Ausnahmen bildeten vor der Einführung des sogenannten Kassenwahlrechts zum 1. Januar 1996 – bei einer Beschränkung auf sozialversicherungspflichtig Beschäftigte – die Betriebs- und die Innungskrankenkassen. Bei den Allgemeinen Ortskrankenkassen war in Regionen ohne Arbeitersatzkassen vor 1996 davon auszugehen, daß zumindest alle sozialversicherungspflichtig beschäftigten ArbeiterInnen eines Betriebs dort versichert waren. Daß diese Bedingungen – abgesehen von einigen Betrieben mit Betriebskrankenkassen, die sich nicht geöffnet haben – auch zukünftig noch zu treffen werden, ist allerdings zweifelhaft.

Als *betriebliche Merkmale* liegen in GKV-Daten gelegentlich Angaben zur *Branche* vor, die im Datensatz der AOK „Küstenstadt“ nach dem 5stelligen Schlüssel der Systematik der Wirtschaftszweige des Statistischen Bundesamtes vercodet sind. Hinsichtlich der Validität und vor allem der Vollständigkeit und Aktualität der Angaben, die für den Verwaltungsprozeß der Krankenkassen ohne Bedeutung sind, ist allerdings Skepsis angebracht (siehe z.B. Thiele 1985: 64). Weitere betriebliche Merkmale wie z.B. *Betriebsgrößen* oder Angaben zur *Zusammensetzung von Belegschaften* sind in GKV-Daten nicht vorhanden. Für Zeiträume, in denen alle ArbeitnehmerInnen eines Betriebs bei derselben Krankenkasse versichert sind (siehe dazu oben), können sie allerdings – in Analogie zum Vorgehen bei der Bereitstellung von Betriebsmerkmalen für die IAB-Beschäftigtenstichprobe (siehe dazu Bender/Hilzendege 1995: 84) – für jeden beliebigen Zeitpunkt zumindest näherungsweise aus den Versichertendaten berechnet werden.<sup>5</sup> So wurden auf der Grundlage der Versicherten- und Leistungsdaten der AOK „Küstenstadt“ als Proxy für die Betriebsgröße die Zahl der sozialversicherungspflichtig beschäftigten ArbeiterInnen und darüber hinaus Kennziffern zur Zusammensetzung der als ArbeiterInnen beschäftigten Belegschaft nach Alter, Nationalität und Geschlecht, zur Fluktuation und zum betriebsspezifischen Krankenstand ermittelt (siehe dazu Behrens/Dreyer-Tümmel 1994: Anhang A). Aufgrund ihrer Fehleranfälligkeit bei kleinen Fallzahlen sind Anteilswerte zur Charakterisierung der Belegschaften von Kleinbetrieben in der Regel allerdings nicht geeignet. Darüber hinaus ist zweifelhaft, ob die Voraussetzungen zur Ermittlung betrieblicher Merkmale aus den Versichertendaten einer Krankenkasse angesichts der Einführung des sogenannten Kassenwahlrechts zum 1. Januar 1996 auch zukünftig noch gegeben sind.

Angaben zu Faktoren wie *betrieblichen Personalpolitiken*, der *technischen Ausstattung* von Betrieben oder ihrer *Arbeitsorganisation*, die beispielsweise Rückschlüsse auf die Möglichkeiten und Strategien eines Betriebs im Umgang mit leistungsgewandelten ArbeitnehmerInnen erlauben könnten, sind in Krankenkassendaten – ebenso wie in der Beschäftigtenstatistik der BA und der IAB-Beschäftigtenstichprobe – selbstverständlich nicht vorhanden. Ebenso liegen keine Angaben zu *Arbeitsplatzmerkmalen* oder konkreten *Arbeitsbedingungen* vor.

Angaben zu *Arbeitsmarktfaktoren* wie z.B. Arbeitslosenquoten oder der Zahl offener Stellen sind natürlich ebenfalls kein Bestandteil von Routinedaten der gesetzlichen Krankenversicherung. Trotzdem können solche Kontextbedingungen, die der amtlichen Statistik zu entnehmen sind, in Untersuchungen zu Erwerbsverläufen auf der Basis von GKV-Daten einbezogen werden.

## Literatur

Andreß, Hans-Jürgen (1996): Arbeitslosigkeit und Arbeitsunfähigkeit. Eine empirische Analyse langfristiger Folgen von Arbeitslosigkeit mit Krankenkassendaten. In: Johann Behrens/Wolfgang Voges (Hg.), *Kritische Übergänge. Statuspassagen und sozialpo-*

litische Institutionalisierung. Frankfurt/New York: Campus Verlag, S. 227-273.

Arrow, Jairo Oka (1994): The Influence of Health on Unemployment in Germany: A Duration Model. In: Vierteljahrshefte zur Wirtschaftsforschung, Heft 1/2, S. 133-138.

Arrow, Jairo Oka/Johann Behrens (1990): Bridges and Traps: Passages after Health Crises in Employment. Beitrag zur Jahrestagung der Britischen Vereinigung für Soziologie vom 2.-5.4.1990 in London. Manuskript: Universität Bremen.

Behrens, Johann (1983): 'Bedürfnisse' und 'Zufriedenheiten' als Statussymbole und Anrechte. Lehren aus einem Panel für Bedürfnistheorie und Planung. In: Karl Otto Hondrich/Randolph Vollmer (Hg.), *Bedürfnisse im Wandel. Theorie, Zeitdiagnose, Forschungsergebnisse*. Opladen: Westdeutscher Verlag, S. 193-244.

Behrens, Johann (1994): Der Prozeß der Invalidisierung – das demographische Ende eines historischen Bündnisses. In: Christoph Behrend (Hg.), *Frühinvalidität – ein „Ventil“ des Arbeitsmarkts?* Berlin: DZA, S. 105-135.

Behrens, Johann/Matthias Schmidt-Ohlemann (1986): Das Versichertenblatt, ein Arbeitsmittel zur Darstellung von Ereignisabfolgen. In: Fred Schwarz/Volker Volkholz, *„Krankenkassendaten und arbeitsbedingte Erkrankungen“*. Teil A: Berufliche und berufs-spezifische Arbeitsunfähigkeitsquoten im interregionalen und intertemporalen Vergleich. Dortmund: Gesellschaft für Arbeitsschutz- und Humanisierungsforschung, S. 117-169.

Behrens, Johann/Wolfgang Voges (1990): Labilisierende Berufsverläufe und der vorzeitige Übergang in den Ruhestand. In: Werner Dressel u.a. (Hg.), *Lebenslauf, Arbeitsmarkt und Sozialpolitik*. Beiträge zur Arbeitsmarkt- und Berufsforschung 133. Nürnberg, S. 201-219.

Behrens, Johann/Jairo Oka Arrow/Ulrich Dorenburg/Anne Dreyer-Tümmel (1992): Gesundheitsberichterstattung und Beschäftigtenmobilität: Welchen Beitrag kann die multivariate Analyse von GKV-Daten zur Identifizierung der Bedingungen beruflicher Labilisierung leisten? In: Ulrich Laaser/Friedrich Wilhelm Schwartz (Hg.), *Gesundheitsberichterstattung und Public health in Deutschland*. Berlin u.a.: Springer-Verlag, S. 379-392.

Behrens, Johann/Anne Dreyer-Tümmel/Rainer Müller (1993): Überbrückung oder Statuspassage: Krankheitsfolgen als ein sozialpolitischer Indikator. In: Lutz Leisering u.a. (Hg.), *Moderne Lebensläufe im Wandel. Beruf – Familie – Soziale Hilfen – Krankheit*. Weinheim: Deutscher Studien Verlag, S. 201-220.

Behrens, Johann/Anne Dreyer-Tümmel (unter Mitarbeit von Hermann Kocyba, Jairo Oka Arrow und Detlef Schulz) (1994): Indikatoren der Rehabilitationsbedürftigkeit aus Routinedaten der gesetzlichen Krankenversicherung (GKV-Daten). Zusammenfassung ausgewählter Ergebnisse und Schlußfolgerungen für die Anregung von Maßnahmen zur Rehabilitation. In: *Deutsche Rentenversicherung*, Heft 7-8, S. 413-443.

Behrens, Johann/Anne Dreyer-Tümmel (1995): Ermittlung von Indikatoren der Rehabilitationsbedürftigkeit aus Routinedaten der gesetzlichen Krankenversicherung (GKV-Daten). Zusammenfassung ausgewählter Ergebnisse und Schlußfolgerungen für die Anregung von Maßnahmen zur Rehabilitation. In: *Deutsche Rentenversicherung*, Heft 7-8, S. 413-443.

Bender, Stefan/Jürgen Hilzendege (1995): Die IAB-Beschäftigtenstichprobe als scientific use file. In: *MittAB* 1, S. 76-95.

Bender, Stefan/Jürgen Hilzendege/Götz Rohwer/Helmut Rudolph (1996): Die IAB-Beschäftigtenstichprobe 1975-1990. Beiträge zur Arbeitsmarkt- und Berufsforschung 197. Nürnberg.

Berger, Peter A./Peter Sopp (1991): Stabilität und Fluktuation. Theoretische und empirische Argumente für eine „verzeitlichte“ Sozialstrukturanalyse. In: Wolfgang Glatzer (Hg.), *25. Deutscher Soziologentag 1990. Die Modernisierung moderner Gesellschaften*. Sektionen, Arbeits- und Ad hoc-Gruppen, Ausschuß für Lehre. Opladen: Westdeutscher Verlag, S. 530-533.

<sup>5</sup> Stefan Bender und Jürgen Hilzendege (1995: 79) sowie Helmut Rudolph (1995: 12) weisen unter Bezug auf die Beschäftigtenstatistik der BA auf die Problematik hin, daß unterbrochene, aber rechtlich fortbestehende Beschäftigungsverhältnisse dann zu einer Doppelzählung führen, wenn der entsprechende Arbeitsplatz während einer Unterbrechung neu besetzt wird. Zu den Möglichkeiten, Unterbrechungen sozialversicherungspflichtiger Beschäftigungsverhältnisse auf der Grundlage von GKV-Daten zu erkennen und somit prinzipiell bei der Ermittlung betrieblicher Merkmale berücksichtigen zu können, siehe Fußnote 1.

- Cramer, Ulrich (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv 69, S. 56-68.
- Dorenburg, Ulrich (1990): Zum Problem der Diagnoseverschlüsselung in Daten der Gesetzlichen Krankenversicherung. Unveröffentlichtes Manuskript: Universität Bremen.
- Dorenburg, Ulrich (1991): Gesundheitsberichterstattung mit Routinedaten der Sozialversicherungsträger. Bremerhaven: Wirtschaftsverlag NW.
- Elkeles, Thomas/Wolfgang Seifert (1992): Arbeitslose und ihre Gesundheit. Empirische Langzeitanalysen. Veröffentlichungsreihe der Forschungsgruppe Gesundheitsrisiken und Präventionspolitik. Wissenschaftszentrum Berlin für Sozialforschung, Nr. P92-201.
- Ferber, Liselotte von/Holger Friedrich/Gert Grellmann/Jochen Barth (1982): Sinnhaftigkeit von GKV-Diagnoseauswertungen, insbesondere von Krankenscheindiagnosen. Untersuchung und Vergleich von Diagnosen auf Arbeitsunfähigkeitsbescheinigungen und Krankenscheinen. In: Die Ortskrankenkasse, 64. Jg., S. 841-855.
- Hernberg, Sven (1984): Arbeitsbedingte Erkrankungen. Hinweise zur epidemiologischen Methodik aus der skandinavischen und angloamerikanischen Fachliteratur. In: Arbeitsmedizin, Sozialmedizin, Präventivmedizin 19, S. 285-289.
- Hernberg, S./H. Kollmeier/K. Kuhn (Hg.) (1986): Nutzung von Daten der Kranken- und Sozialversicherung zur Darstellung des Zusammenhangs von Arbeitsbedingungen und Gesundheit. Expertisen. Kolloquium am 24.9.1985 in Dortmund. Schriftenreihe der Bundesanstalt für Arbeitsschutz, Tagungsband Nr.45. Bremerhaven: Wirtschaftsverlag NW.
- Köster, I./ T. Finger (1992): Gesundheitsberichterstattung als Instrument der Qualitätssicherung und Qualitätsverbesserung. Repräsentativität von Untersuchungen an Ortskrankenkassendaten. In: Ulrich Laaser/Friedrich Wilhelm Schwartz (Hg.), Gesundheitsberichterstattung und Public health in Deutschland. Berlin u.a.: Springer-Verlag, S. 249-254.
- Krüger, Helga/Claudia Born (1991): Unterbrochene Erwerbskarrieren und Berufsspezifität: Zum Arbeitsmarkt- und Familienpuzzle im weiblichen Lebenslauf. In: Karl Ulrich Mayer u.a. (Hg.), Vom Regen in die Traufe: Frauen zwischen Beruf und Familie. Frankfurt/New York: Campus Verlag, S. 142-161.
- Leibing, Christa/Dieter Müller-Späth (1981): Daten über den Zugang an Berufs- und Erwerbsunfähigkeitsrenten in der deutschen gesetzlichen Rentenversicherung. In: Ralph Brennecke u.a. (Hg.), Datenquellen für Sozialmedizin und Epidemiologie. Berlin u.a.: Springer-Verlag, S. 180-197.
- Quack, Sigrid (1991): Teilzeitarbeit aus der Perspektive des Erwerbsverlaufs. In: Wolfgang Glatzer (Hg.), 25. Deutscher Soziologentag 1990. Die Modernisierung moderner Gesellschaften. Sektionen, Arbeits- und Ad hoc-Gruppen, Ausschuß für Lehre. Opladen: Westdeutscher Verlag, S. 778-781.
- Rudolph, Helmut (1995): Das DEVO/DÜVO-Meldevorhaben als Datenquelle der Beschäftigtenstatistik. Beitrag für den ZUMA-Workshop „Die Beschäftigtenstatistik als Datenbasis für Arbeitsmarktanalysen“ vom 28.-29.9.1995 in Mannheim. Manuskript.
- Schach, Elisabeth (1981): Daten der gesetzlichen Krankenversicherung am Beispiel einer AOK. In: Ralph Brennecke u.a. (Hg.), Datenquellen für Sozialmedizin und Epidemiologie. Berlin u.a.: Springer-Verlag, S. 201-214.
- Schulz, Detlef/Hermann Kocyba/Anne Dreyer-Tümmel/Johann Behrens/Jairo Oka Arrow (1994): Belastungs- und tätigkeitsbezogene Berufsklassifikation auf der Basis von GKV-Daten. Band 2 der Schriftenreihe „Qualitätsmanagement und psychosoziale Kompetenz“ des Instituts für Supervision, Institutionsberatung und Sozialforschung. Frankfurt: ISIS.
- Schulz, Detlef/Anne Dreyer-Tümmel/Johann Behrens (1996): Die Angaben zur beruflichen Tätigkeit in den Daten der gesetzlichen Krankenversicherung – ihr Beitrag zur verlaufsbezogenen Analyse berufsbedingter Belastungen und ihrer Folgen. In: Anne Dreyer-Tümmel u.a., Möglichkeiten und Grenzen der Erforschung von Zusammenhängen zwischen Krankheit und Erwerbsverlauf mit Routinedaten der gesetzlichen Krankenversicherung. Arbeitspapier Nr. 40 des Sonderforschungsbereichs 186 „Statuspassagen und Risikolagen im Lebensverlauf“ der Universität Bremen, S. 29-38.
- Schupp, Jürgen (1991): Temporale Analysen mit Paneldaten am Beispiel des Sozio-ökonomischen Panels (SOEP). In: Wolfgang Glatzer (Hg.), 25. Deutscher Soziologentag 1990. Die Modernisierung moderner Gesellschaften. Sektionen, Arbeits- und Ad hoc-Gruppen. Ausschuß für Lehre. Opladen: Westdeutscher Verlag, S. 788-792.
- Schwefel, Detlef/Friedrich Wilhelm Schwartz (1978): Aussagefähigkeit und Auswertbarkeit von Diagnosen in der ambulanten medizinischen Versorgung – Ein Problemüberblick. In: Friedrich Wilhelm Schwartz/Detlef Schwefel (Hg.), Diagnosen in der ambulanten Versorgung. Aussagefähigkeit und Auswertbarkeit. Eine Expertenumfrage in der Bundesrepublik Deutschland. Köln: Deutscher Ärzte-Verlag, S. 7-34.
- Thiele, Wilhelm (1985): Gesundheitsverschleiß und medizinische Versorgung am Beispiel der Mechaniker. Nutzungsmöglichkeiten von Prozeßdaten der GKV für eine feingliedrige und differenzierte Analyse. In: Wilhelm F. Schröder/Wilhelm Thiele (Hg.), Krankheit und Arbeitswelt. Möglichkeiten der Analyse mit Daten der gesetzlichen Krankenversicherung. BASiG-Schriftenreihe „Strukturforschung im Gesundheitswesen“. Berlin, S. 57-84.
- Voges, Wolfgang (1994): Mißbrauch des Rentensystems? Invaldität als Mittel der Frühverrentung. Frankfurt/New York: Campus Verlag.
- Volkholz, Volker/Fred Schwarz (1984): Annäherung an sozialwissenschaftliche Verlaufsanalysen mit Hilfe von Krankenkassendaten. Arbeitsbericht zum Problem der begrenzten Tätigkeitsdauer. Bremerhaven: Wirtschaftsverlag NW.
- Wermter, Winfried (1981): Die Beschäftigtenstatistik der Bundesanstalt für Arbeit. In: MittAB 4, S. 428-435.
- Wermter, Winfried/Ulrich Cramer (1988): Wie hoch war der Beschäftigtenanstieg seit 1983? Ein Diskussionsbeitrag aus der Sicht der Beschäftigtenstatistik der Bundesanstalt für Arbeit. In: MittAB 4, S. 468-482.