

Anmerkungen zur Methodik und Qualitätsbewertung der amtlichen Diagnosestatistik

Bömermann, Hartmut; Hauswirth, Angelika

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Bömermann, H., & Hauswirth, A. (2002). Anmerkungen zur Methodik und Qualitätsbewertung der amtlichen Diagnosestatistik. *Berliner Statistik / Monatschrift*, 8, 318-322. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-61257-9>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-ND Lizenz (Namensnennung-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nd/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-ND Licence (Attribution-NoDerivatives). For more information see: <https://creativecommons.org/licenses/by-nd/4.0>

Hartmut Bömermann
Angelika Hauswirth

Anmerkungen zur Methodik und Qualitätsbewertung der amtlichen Diagnosestatistik

Vorbemerkung

Die Krankenhausstatistik ist seit 1990 eine Bundesstatistik; sie ist thematisch und erhebungstechnisch dreigeteilt. Mit den Grunddaten (Teil I) und dem Kostennachweis (Teil III) werden die sachlich-personelle Ausstattung und wichtige Eckdaten des Leistungsgeschehens sowie die Kostenstruktur der Wirtschaftseinheit einrichtungsbezogen erhoben. Als untere Gliederungsebene wird bei der Grunddatenerhebung die Fachabteilung der Einrichtung zugrunde gelegt; der Kostennachweis bezieht sich auf das gesamte Krankenhaus als ein Aggregat. Detaillierter wird die Leistungserbringung des Krankenhauses in der Diagnosestatistik (Teil II) erfasst; sie basiert auf dem einzelnen Behandlungsfall. Aufgrund von Ressourcenbeschränkungen wurde die Diagnosestatistik zeitversetzt gegenüber den beiden anderen Teilsystemen der Krankenhausstatistik erstmalig 1993 durchgeführt. Mit ihr stehen für alle Bundesländer fallbezogene Morbiditätsdaten zur Verfügung. In Berlin waren das für den Berichtszeitraum 1999 über 690 000 Datensätze. Der kumulierte Datensatz 1993 bis 1999 umfasst 4,6 Millionen Datensätze.

Die Auswertung der Diagnosedaten wird für Analysen der Krankenhausversorgung wie für epidemiologische Fragestellungen verwendet.

Aus dem Wissenschaftsbereich ist die Nachfrage – ganz entgegen der Erwartung – eher gering. Im Gutachten der „Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik“ (KVI) vom März 2001 wird im Beitrag über das Informationsfeld Gesundheits- und Sozialwesen die Diagnosestatistik nur am Rande erwähnt, während gleichzeitig die schlechte Datenlage beanstandet wird.

Mit dem Berichtsjahr 2000 hat bei der Verschlüsselung der Hauptdiagnose ein Klassifikationswechsel von der ICD 9 zur ICD 10 stattgefunden. Der Zeitraum 1993 bis 1999 bildet gewissermaßen einen klassifikatorisch einheitlichen Raum und ist deshalb für eine längsschnittliche Betrachtung gut geeignet. Die Daten des ersten Berichtsjahres bleiben allerdings ausgeklammert, da mit der Einführung einer Statistik gesonderte Probleme verbunden sind, die hier nicht näher erörtert werden sollen.

Zunächst werden im folgenden Beitrag einige wichtige allgemeine Aspekte der amtlichen Statistik angesprochen. Anschließend wird die Diagnosestatistik in ihren Grundzügen vorgestellt. Die Darstellung ausgewählter Themen zur Messung, Bewertung und Verbesserung der Datenqualität des Materials schließen sich an.

1.1 Rechtsprinzip der amtlichen Statistik

Von der Nachfrager- oder Kundenseite werden oftmals fehlende Merkmale bemängelt oder es wird eine andere Abgrenzung des Berichtskreises gewünscht. Die amtliche Statistik ist in diesen Fragen jedoch nicht autonom. Sie gründet auf dem Legalitätsprinzip, d. h. Statistiken benötigen eine Rechtsgrundlage. Der Gesetzgeber – und nicht die Regierung, die Wissenschaft oder ein beratendes Gremium – beauftragt die amtliche Statistik mit der Vorbereitung von Bundesstatistiken. Mit der Einführung des Rechts auf informationelle Selbstbestimmung durch das Bundesverfassungsgericht im Volkszählungsurteil von 1983 wurden dem staatlichen Auskunftsbegehren engere Grenzen gezogen und es wurden besondere Vorkehrungen für die Organisation und Durchführung der Erhebungen verlangt.

Der Mindestinhalt der Rechtsgrundlage umfasst nach dem Bundesstatistikgesetz (BStatG) von 1987 die Erhebungsmerkmale, die für die Durchführung der Erhebung erforderlichen Hilfsmerkmale, die Art der Erhebung, den Berichtszeitraum, den Berichtszeitpunkt der Datenlieferung an das Statistische Landesamt, die Periodizität der Erhebung und den Kreis der Befragten (BStatG, § 9 Abs. 1).

Eine Folge des Urteils war eine erhebliche Verstärkung der Regelungsdichte, die eine Statistik bis in Details fest schreibt.

Durch die Verrechtlichung der amtlichen Statistik – eine Statistik kann nicht ohne Rechtsgrundlage durchgeführt werden – wird die Kontinuität von Erhebungen, für die oft erhebliche finanzielle Mittel bereitgestellt werden müssen, gesichert. Sie ist auch die Voraussetzung für eine methodische Besonderheit der amtlichen Statistik, denn anders als bei Erhebungen wissenschaftlicher oder privater Auftraggeber kann eine Auskunftspflicht, die in die verbrieften individuellen Freiheitsrechte eingreift, für die Befragten durch Gesetz erlassen werden. Mit der Auskunftspflicht – darunter ist die vollständige, wahrheitsgemäße, fristgerechte und unentgeltliche Auskunft zu verstehen – schafft sich die amtliche Statistik eine Schlüsselvoraussetzung für eine hohe Datenqualität. Sie ist ein tragendes Element in der Statik des Qualitätsbegriffs der amtlichen Statistik (vergl. Chlumsky, Wiegert u. a. 1993, S. 8).

Eine Anpassung der Erhebungsprogramme an veränderte Realitäten macht Änderungen der rechtlichen Grundlagen erforderlich. Dies ist die mögliche Kehrseite der Verrechtlichung – hier ist mit Verrechtlichung insbesondere die hohe Regelungsdichte und deren Detailliertheit angesprochen –, da der nötige Entscheidungsaufwand zu einem Verlust an Anpassungsfähigkeit führen kann.

Die Rechtsgrundlage für die Krankenhausstatistik ist die Krankenhausstatistikverordnung (KHStatV) von 1990. Im Krankenhausfinanzierungsgesetz (KHG) von 1985 wird in § 28 die Bundesregierung ermächtigt, durch Rechtsverordnung und mit Zustimmung des Bundesrates jährliche Erhebungen über Krankenhäuser und Vorsorge- und Rehabilitationseinrichtungen als Bundesstatistik anzuordnen. Angeordnet wurde die Bundesstatistik für Krankenhäuser mit der Krankenhausstatistikverordnung von 1990. In der nach der Vereinigung Deutschlands novellierten Fassung des Krankenhausfinanzierungsgesetzes von 1991 werden die zu erfassenden Sachverhalte aufgezählt und die Auskunftspflicht der

Krankenhausträger gegenüber den Statistischen Landesämtern geregelt. Die Sachverhalte sind:

- Art des Krankenhauses und der Trägerschaft;
- im Krankenhaus tätige Personen nach Geschlecht, Beschäftigungsverhältnis, Tätigkeitsbereich, Dienststellung, Aus- und Weiterbildung;
- sachliche Ausstattung und organisatorische Einheiten des Krankenhauses;
- Kosten nach Kostenarten;
- in Anspruch genommene stationäre und ambulante Leistungen;
- Patienten nach Alter, Geschlecht, Wohnort, Erkrankungen nach Hauptdiagnosen und
- Ausbildungsstätten am Krankenhaus.

Die Erhebungs- und Hilfsmerkmale, die Periodizität sowie der Berichtszeitpunkt und die Datenweitergabe sind ebenfalls in der Verordnung geregelt.

1.2 Berichtskreis, Grundgesamtheit und Merkmale

Wird über die Ergebnisse von Statistiken gesprochen, dann werden häufig Kurzbezeichnungen verwendet, um Umständlichkeiten zu vermeiden. Was ist aber z. B. genau mit „Hauptdiagnose“ gemeint, und zwar sowohl inhaltlich – welche Diagnose ist die Hauptdiagnose? – als auch umfänglich – über wen werden Aussagen gemacht? Diese und verwandte Fragen sind sowohl bei der Erhebung wie für die Interpretation der statistischen Ergebnisse genau zu klären.

Zunächst ist in jeder Statistik der Berichtskreis abzugrenzen, über den die empirischen Daten gewonnen werden sollen. Der Berichtskreis der Diagnosestatistik bezieht sich auf die Gesamtheit der Krankenhäuser. Diagnosen im niedergelassenen Bereich sind somit nicht Teil der Statistik. Welche Anforderungen muss eine Einrichtung aber erfüllen, um als Krankenhaus zu gelten? In § 107 Abs. 1 des Sozialgesetzbuches (SGB V, S. 2477) – auf den die Krankenhausstatistikverordnung Bezug nimmt – wird ein Kriterienkatalog genannt. Danach muss die Einrichtung

- der Krankenhausbehandlung oder Geburtshilfe dienen,
- unter ständiger ärztlicher Leitung stehen,
- über ausreichende diagnostische und therapeutische Möglichkeiten verfügen,
- nach anerkannten Methoden arbeiten,
- über jederzeit verfügbares ärztliches und nichtärztliches Personal verfügen und
- es müssen Patienten untergebracht und gepflegt werden können,

um als Krankenhaus gelten zu können. Ausgeschlossen sind von der Erhebung die Rehabilitations- und Vorsorgeeinrichtungen und reine Tages- oder Nachtkliniken ohne vollstationäre Betten. Krankenhäuser im Straf- und Maßregelvollzug sowie Polizeikrankenhäuser werden ebenfalls bei der Erhebung nicht berücksichtigt, da sie der Allgemeinheit nicht direkt zugänglich sind.

Berichtspflichtig sind – dies ist dann eher eine erhebungsrechtliche Frage – die Krankenhausträger.

Konstituiert wird die Grundgesamtheit, über die empirische Daten erhoben werden, durch den Merkmalsträger und die identifizierenden Merkmale zur örtlichen und zeitlichen Abgrenzung.

Der *Merkmalssträger* in der Diagnosestatistik ist der Patient. Allerdings nicht als Person, sondern als Behandlungsfall. Der einzelne Behandlungsfall ist die kleinste

Einheit, über die Daten erhoben werden. Die so entstehenden Mikrodaten stehen am Ende des statistischen Aufbereitungsprozesses für Auswertungen und Analysen zur Verfügung. Genauer spezifiziert werden muss, wann ein Patientendatum erhoben wird – es handelt sich schließlich um eine große dynamische Bewegungsmasse – und über welche Patienten Daten erhoben werden. Eingegrenzt ist die Erhebung auf die aus ununterbrochener vollstationärer Behandlung in einem Krankenhaus entlassenen Patienten. Für jeden aus dem Krankenhaus vollstationär Entlassenen, und zwar einschließlich der im Krankenhaus nach einer vollstationären Aufnahme Verstorbenen, wird ein Datensatz erzeugt; gleiches gilt für beurlaubte Patienten, wenn für die Urlaubszeit keine Pflegesätze berechnet werden. Die Zählung wird ausgelöst durch die Entlassung eines Patienten. Bei einem mehrfachen Krankenhausaufenthalt eines Patienten innerhalb eines Jahres wird jeder beendete Aufenthalt als neuer Fall erfasst. Die Zählweise bezieht sich auf den Fall und nicht auf eine Person und deren Behandlungskarriere. Synonym hierzu wird auch von Behandlungs-, Krankenhaus- oder Patientenfall gesprochen.

Die Behandlung einer Krebserkrankung ist häufig nicht nach einer stationären Behandlung abgeschlossen, sondern macht eine oder mehrere stationäre Nachsorgebehandlungen (stationäre Chemotherapie, Strahlentherapie und andere konservative Behandlungsformen) erforderlich. Jede Entlassung schlägt sich als Behandlungsfall in der Statistik nieder. Die Statistik kann aus den Fällen keine Behandlungsidentität rekonstruieren; es sind unverbundene und scheinbar voneinander unabhängige Fälle. Zu noch höheren Fallzahlen kann die Diagnose ICD-Position 780 „Allgemeine Symptome“ (Schwindel, Schlafstörung, Asthenie) führen, wenn die Patienten in speziellen Schlaflabors behandelt werden, da die Patienten zur Diagnostik täglich zum überwachten Schlafen vollstationär aufgenommen und nach Beendigung jeder Untersuchung als ein Fall gezählt werden.

Patienten, die im Berichtszeitraum zwar behandelt, aber nicht entlassen wurden, erscheinen erst in den Daten eines späteren Berichtszeitraumes, und zwar dann, wenn das Ereignis „Entlassung“ eingetreten ist.

Die Kriterien für einen erzeugten Behandlungsfall sind in der Diagnosestatistik:

- Behandlung in einem Krankenhaus (Berichtskreis),
- vollstationäre Unterbringung (nicht ambulante, teil-, vor- und nachstationäre Fälle).

Ausgelöst wird die Zählung durch die Ereignisse:

- Entlassung,
- Tod

Eine Entlassung ist:

- der Abschluss der vollstationären Behandlung.

Die Diagnosedaten werden jährlich von den statistischen Landesämtern über alle in ihrem Zuständigkeitsbereich (Bundesland) liegenden Krankenhäuser erhoben. Der Berichtszeitraum bezieht sich auf ein Kalenderjahr. Bei der Diagnosestatistik wurde als Erhebungstyp die *Totalerhebung* gewählt, alle Krankenhäuser (bis auf die in der Verordnung ausdrücklich ausgegliederten Krankenhausarten) werden einbezogen; es ist weder eine Stichprobe noch gibt es Abschneidegrenzen.

Die Diagnosestatistik umfasst sowohl eine Primärerhebung als auch die statistische Nutzung prozessgenerierter Daten des Krankenhauses, wie sie für die Behandlungsdokumentation und -abrechnung vorhanden sind.

Die wichtigsten Merkmale der Statistik – gemessen an den Auswertungsnachfragen – sind Geschlecht, Alter, Hauptdiagnose, Verweildauer sowie Wohn- und Behandlungsort. Der folgenden Übersicht können alle Erhebungsmerkmale entnommen werden.

Erhebungsmerkmal	Kriterium/Bedingung	Ausprägungen/Wertebereich
Geschlecht		männlich / weiblich
Geburtsdatum		Monat, Jahr
Zugangsdatum		Tag, Monat, Jahr
Abgangsdatum		Tag, Monat, Jahr
Sterbefall		ja / nein
Hauptdiagnose	Diagnose, auf die der überwiegende Teil der Verweildauer zurechenbar ist bzw. den größten Teil der medizinischen Leistung verursacht hat. Abbildung auf die ICD 9 (ab 2000 ICD 10).	ICD-9-Pos. 001 bis ICD-Pos. 998, V-Klassifikation (nichtkranke Zustände)
Operation	in Zusammenhang mit der Hauptdiagnose	ja / nein
Fachabteilung	mit der längsten Verweildauer	35 Fachabteilungen mit Teiluntergliederungen
Wohnort des Patienten		Staat, Bundesland, Regierungsbezirk, Gemeinde, in Berlin Bezirk

Aus den Erhebungsmerkmalen können weitere Merkmale abgeleitet werden, wie die folgende Aufstellung ausgewählter Variablen zeigt:

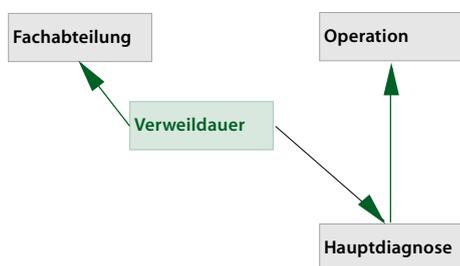
Abgeleitetes Merkmal	Operationalisierung	Ausprägungen
Stundenfälle	am Tag der Aufnahme entlassen (die Verweildauer beträgt null Tage).	ja / nein
Alter	Alter am Tag der Aufnahme	Alter in Monaten, Alter in Jahren
Verweildauer	Differenz zwischen Entlassungs- und Aufnahmedatum	in Tagen
Aufnahme-wochentag		Wochentag (Mo, Di, Mi, Do, Fr, Sa, So)
Entlassungs-wochentag		Wochentag (Mo, Di, Mi, Do, Fr, Sa, So)
Kurzlieger	Verweildauer bis zu drei Tagen	ja / nein

Von den neun Erhebungsmerkmalen sind drei Merkmale durch zusätzliche Verknüpfungen zu einem anderen Merkmal operationalisiert, und zwar die Hauptdiagnose, die Operation und die Fachabteilung. Über die Vergabe der Hauptdiagnose entscheidet die Länge der Verweildauer bzw. der Anteil an der erbrachten medizinischen Leistung. Die Operation ist an die Hauptdiagnose gekoppelt, während die Fachabteilung wiederum einen Bezug zur längsten Verweildauer aufweist. Alle Merkmale weisen einen direkten oder indirekten Zusammenhang mit der Verweildauer auf. Nachfolgend ist die Abhängigkeitsstruktur dargestellt. Die Verweildauer fungiert als Aufwandsindikator.

Die Konstruktion der Merkmale muss bei der Interpretation der Ergebnisse berücksichtigt werden. Das Operationsmerkmal ist kein eigenständiges Merkmal. Vielmehr ergänzt es die Informationen zur Hauptdiagnose. Aus der Verteilung des Merkmals kann nicht die Operationshäufigkeit in Berliner Krankenhäusern ermittelt werden. Die Gültigkeit und Zuverlässigkeit des Merkmals vorausgesetzt – dazu später mehr – könnte nur die Operationshäufigkeit bei einer bestimmten Hauptdiagnose festgestellt werden.

Unter der Hauptdiagnose wird diejenige Diagnose verstanden, auf die der überwiegende Teil der Verweildauer bzw. der erbrachten medizinischen Leistung zurechenbar ist. Durch die Konstruktion des Merkmals „Hauptdiagnose“ wird die mögliche Komplexität der Erkrankung auf einen Schwerpunkt konzentriert. In der Statistik wird häufig so verfahren, um zu einer handhabbaren Erhebung zu kommen. Allerdings können Komorbiditäten (Nebenerkrankungen) oder Multimorbidität damit nicht abgebildet werden.

Abhängigkeitsstruktur bei Merkmalsoperationalisierungen



Weitere Möglichkeiten bietet die Verknüpfung der Falldaten mit Krankenhausmerkmalen, die den institutionellen Kontext des Krankenhausaufenthaltes beschreiben:

Kontextmerkmal	Ausprägungen	Datenquelle
Ort des Krankenhauses (ohne örtliche Einheiten)	Bezirk	Grunddaten
Art des Trägers	öffentlich/ freigemeinnützig/ privat	Grunddaten
Art des Krankenhauses	Hochschulklinik/ Plankrankenhaus/ Versorgungsvertrag nach § 108 Nr. 3 SGB V/ sonstiges Krankenhaus	Grunddaten
Bettengrößenklasse	von ... bis unter ... Betten unter 50 50 - 100 100 - 150 150 - 200 200 - 250 250 - 300 300 - 400 400 - 500 500 - 600 600 - 800 800 - 1 000 1 000 und mehr	Grunddaten

Die in den drei Übersichten genannten Merkmale können beliebig kombiniert werden. Bei der Veröffentlichung von Ergebnissen muss selbstredend der statistischen Geheimhaltung Rechnung getragen werden.

Erhebungs- und Aufbereitungsprozess

Die sich jährlich wiederholenden Erhebungen beginnen mit dem Versand von Informationen und Unterlagen durch das Statistische Landesamt an die berichtspflichtigen Krankenhäuser. Bis zum 30. Juni haben alle Krankenhäuser ihre Diagnosedaten für das abgelaufene Berichtsjahr an die Statistik zu senden. Das Statistische Landesamt wiederum hat bis zur 1. Dezemberwoche die Statistik abzuschließen und zusammengefasste Daten an das Statistische Bundesamt zu übermitteln, damit deutschlandweite Ergebnisse aufbereitet und veröffentlicht werden können. Einen stark schematisierten Überblick über den Erhebungs- und Aufbereitungsprozess der Diagnosestatistik gibt Abbildung 1. Die Darstellung ist vertikal nach den beteiligten Akteuren Auftraggeber bzw. Kunde, Statistisches Landesamt und berichtendes Krankenhaus unterteilt.

Wichtige Stationen sind die Eingangskontrolle, die Plausibilitätsprüfungen, die Schlusskontrolle und die Lieferung an das Statistische Bundesamt. Bei der Eingangskontrolle wird die Vollständigkeit – haben alle Krankenhäuser geliefert? – und die Vollständigkeit – haben sie für das gesamte Berichtsjahr geliefert und wurden alle Erhebungsmerkmale angegeben? – geprüft. Die Prüfungen geschehen möglichst zeitnah zur Lieferung. Gleichzeitig werden die Ergebnisse einer ersten Krankenhausbezogenen Auszählung mit Vorjahresdaten bzw. den Angaben aus der Grunddatenstatistik verglichen, um frühzeitig gravierende und offensichtliche Fehler aufzudecken. In die Prüfung einbezogen werden pro Krankenhaus die Gesamtzahl der Patientenabgänge, die Sterbefälle, die gültigen Fachabteilungen, die Operationen und die Verschlüsselung des Wohnortes. Werden bei diesen Prüfungen Fehler erkannt, wird häufig eine neue Datenlieferung angefordert; die Einzelfallklärung erfolgt erst später und nicht in diesem Prozessabschnitt, der mehr auf das Massengeschäft ausgerichtet ist. Zur Eingangsprüfung gehört auch das so genannte Mahnwesen, da die Auskunftspflicht in einigen Fällen mit einem zusätzlichen Aufwand durchgesetzt werden muss. Daten, die die Eingangsprüfung passiert haben, werden im nachfolgenden Prozessschritt plausibilisiert.

Bei der Plausibilisierung werden entsprechend einer bundeseinheitlichen Verfahrensweise folgende Prüfungen vorgenommen:

- Befinden sich die Angaben im zulässigen Wertebereich? Beispielsweise sind beim Merkmal „Geschlecht“ nur die Angaben „1“ und „2“ sowie ein Schlüssel für eine fehlende Angabe zulässig. Entsprechend werden auch die ICD-Schlüssel und die anderen Merkmale auf ihre formale Korrektheit überprüft.
- Wurden seltene bzw. als besiegt geltende Erkrankungen angegeben? Pest und Cholera gehören in diese Kategorie. Bei diesen Erkrankungen ist jeder Einzelfall signifikant.
- Ist die Kombination von bestimmten Merkmalsausprägungen plausibel? Logisch ausgeschlossen sind beispielsweise Erkrankungen, die im Widerspruch zum Geschlecht des Patienten stehen. Eine Ausschlussbeziehung zwischen Erkrankung und Alter, die auch geprüft wird, ist dagegen weniger trennscharf. Diese Prüfungen geben aber Hinweise auf mögliche Kodierungsfehler. Zu den Kombinationsprüfungen gehört auch die Prüfung der Verschlüsselung des Patientenwohnortes anhand der übermittelten Angaben.

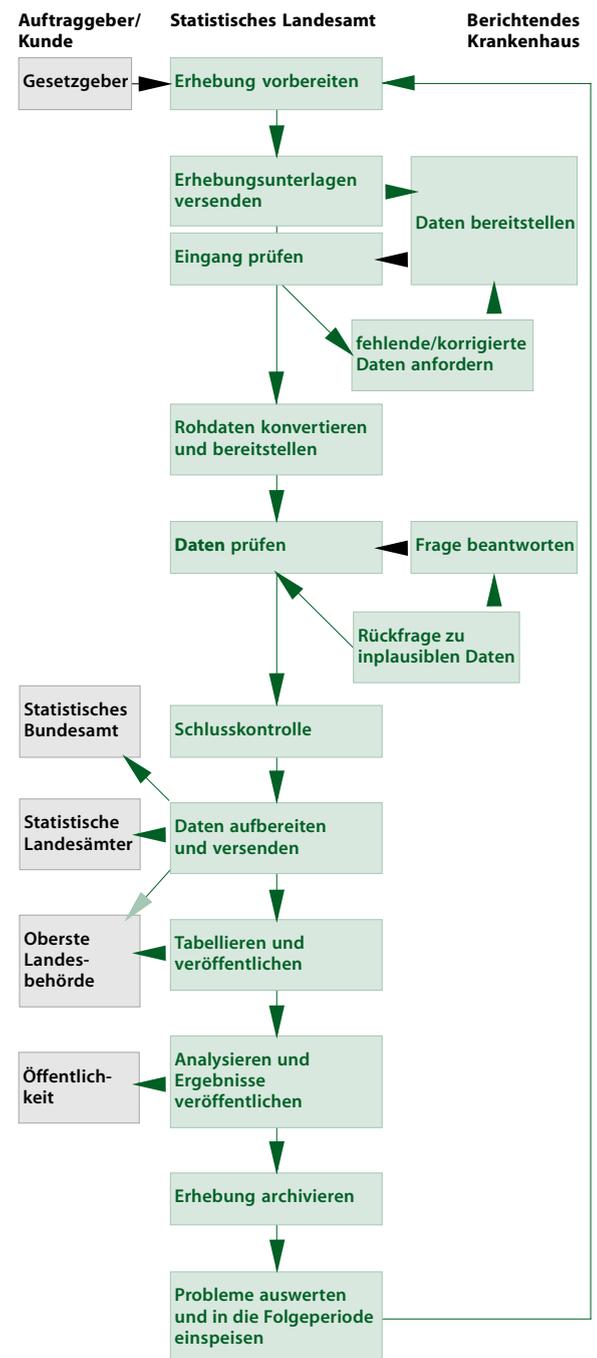
Entdeckte Inplausibilitäten werden durch Rückfragen mit dem Krankenhaus geklärt und entschieden. Ist eine Klärung nicht möglich, werden die betroffenen Merkmale auf unbekannt gesetzt. Diese Arbeiten erfordern viel Erfahrung und ein hohes Verantwortungsgefühl, da die erzielbare Datenqualität hiervon direkt abhängt.

Haben die Daten die Plausibilitätskontrollen durchlaufen, werden sie für eine Schlusskontrolle tabelliert. Nach der Datenfreigabe werden die Ergebnisse der Statistik an das Statistische Bundesamt übermittelt.

Anschließend werden die Diagnosedaten von Patienten aus anderen Bundesländern, die in Berlin behandelt wurden, an das jeweilige Statistische Landesamt elektronisch versandt.

Die Daten werden nun in eine Datenbank eingelesen und auswertungsbezogen abgerufen und verarbeitet.

Abb. 1 Erhebungs- und Aufbereitungsprozess der Diagnosestatistik



Bewertung und Verbesserung der Datenqualität

Die Aufgabe des im vorherigen Kapitel vorgestellten Arbeitssystems ist es, verlässliche Daten über einen Realitätsausschnitt zu gewinnen und zur Verfügung zu stellen. An der Qualität der Daten muss sich jeder Statistikproduzent messen lassen. Dies gilt selbstverständlich nicht nur für die amtliche Statistik (vergl. Sommer, Unholzer, Wiegand 1998). Seit einigen Jahren wird ein erweiterter Qualitätsbegriff (vergl. Quality in Statistics, Eurostat 1998) verwendet, der stärker auf die Relevanz der Konzepte, die Aktualität und Vergleichbarkeit sowie den Zugang zu den Daten abzielt.

Von einem statistischen Arbeitssystem wird die Einhaltung folgender Qualitätsziele gefordert:

- Genauigkeit
- Vollzähligkeit
- Vollständigkeit
- Zuverlässigkeit (Reliabilität)
- Gültigkeit der Messung (Validität)
- Konsistenz (gleiche Methodik, gleiche Klassifikationen)
- Kontinuität (Längsschnitt)
- Aktualität
- Verfügbarkeit

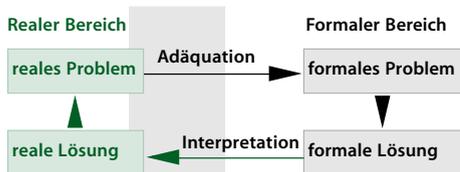
Einige der genannten Punkte sollen im Folgenden genauer betrachtet werden. Das Thema Datenqualität lässt sich in zwei Bereiche unterteilen: Einmal geht es um die Schnittstelle zwischen dem Bereich der Frage- und hypothesengeleiteten Theorie, die auf der Grundlage empirischer Daten bearbeitet werden soll, und zum anderen um den Bereich der Statistik selbst, die durch interne Verfahren die Datenqualität zu sichern hat.

Adäquation zwischen Theorie und statistischen Daten

Die Statistik ist ein Teilsystem im Erkenntnisprozess. Empirische Daten ermöglichen einen Zugang zur wirtschaftlichen, sozialen und politischen Wirklichkeit, der zu überprüfbareren Aussagen führt. Was erhoben wird und wie die Ergebnisse interpretiert werden, ist nicht Teil der Statistik im engeren Sinn.

Grohmann (vgl. Grohmann 1985) verwendet in seiner Erörterung des Erkenntnisprozesses das folgende Schema:

Prozess der Erkenntnisbildung durch Formalisierung einer Fragestellung und Interpretation statistischer Ergebnisse



Der Erkenntnisprozess beginnt mit einem realen Problem, mit theoretischen Begriffen und einer theoretischen Fragestellung, die gelöst werden soll (Realer Bereich). Das reale Problem wird formalisiert, um mit-

tels der Statistik empirische Daten zu gewinnen und geeignete Modelle auszuwählen, die zur Beschreibung des Problems beitragen und zum Testen von Zusammenhängen verwendet werden können (Formaler Bereich). Die Übersetzung zwischen dem realen Bereich und dem formalen Bereich wird als Adäquation bezeichnet, da Entsprechungen zwischen diesen – getrennten – Bereichen gefunden werden müssen. Im Prozess der Adäquation wird versucht, theoretische Vorstellungen in definierte, operable Merkmale und statistische Modelle zu übersetzen. Theoretischer und operabler Begriff müssen so nahe wie möglich kommen. Bei der Adäquation geht es in der Erhebungsphase um das Finden einer geeigneten Entsprechung, die eine gültige (valide) Messung erlaubt.

Die inhaltliche Differenz zwischen dem theoretischen Begriff und dem statistischen Begriff sollte möglichst klein sein. Grohmann ist der Auffassung, dass eine mehr oder weniger große sachliche Diskrepanz bleibt (Grohmann 1985, S.11). Bei dieser Diskrepanz handelt es sich aber nicht um einen messbaren Fehler, sondern um die Übersetzung zwischen zwei begrifflichen Ebenen, die beide unterschiedliche Funktionen im Erkenntnisprozess haben. Die statistischen Begriffe schaffen der Theo-

rie Zugang zur Realität. Für sie erhalten wir Zahlen. Ohne die statistischen Begriffe wäre Theorie bloße vorwissenschaftliche Spekulation.

Das Adäquationsproblem besteht bei der Datengewinnung und bei der Interpretation der erhobenen statistischen Daten.

Zunächst sollen *Probleme der Datengewinnung* vorgestellt werden. Neben der Adäquation werden auch Probleme der statistikinternen Verlässlichkeit angesprochen. Für die Operationalisierung der Hauptdiagnose wurde als Kriterium die längste Verweildauer bzw. der Anteil an der erbrachten medizinischen Leistung festgelegt. Diese Definition folgt dem „Leitfaden zur Erstellung der Diagnosestatistik nach § 16 der Bundespflegegesetzverordnung“ (Bundesministerium für Arbeit und Sozialordnung, 1986), wie sie für Zwecke der Pflegesatzverhandlungen erstellt wird. Aus Gesprächen mit den Zuständigen in den Krankenhäusern wissen wir, dass die Qualifizierung einer Aufnahme-, Einweisungs-, Verlegungs- oder Entlassungsdiagnose als Hauptdiagnose wohl eher nicht der Definition in der Statistik folgt, obwohl im Ergebnis auf die uns mitgeteilte Diagnose in aller Regel die längste Verweildauer bzw. der höchste Ressourceneinsatz entfallen dürfte; allerdings ist der Zusammenhang nicht konstitutiv für das Merkmal. Insbesondere der Bezug zur Verweildauer scheint aus den Unterlagen des Krankenhauses nur durch zusätzlichen Aufwand herstellbar zu sein. Nach den neuen Deutschen Kodierrichtlinien ist die Hauptdiagnose folgendermaßen definiert: „Die Diagnose, die nach Analyse als diejenige festgestellt wurde, die hauptsächlich für die Veranlassung des stationären Krankenhausaufenthaltes des Patienten verantwortlich ist“ (Deutsche Kodierrichtlinien, S. 4). Diese Definition dürfte auch dem Common sense entsprechen.

Die inhaltliche Differenz zwischen statistischer und administrativer Definition führt umweglos zu einem kritischen Punkt der statistischen Erhebung, und zwar zum Grad der Orientierung an den Konzepten, die in den Krankenhäusern Teil der informationellen Infrastruktur sind. Wird eine Statistik als Primärerhebung verstanden, so wie es bei der Diagnosestatistik der Fall ist, dann definiert sie wissenschaftlich autonom den Merkmalskatalog und die erforderlichen Messinstrumente. Nutzt sie dagegen prozessgenerierte Daten, dann setzt sie auf vorhandenen Merkmalen auf. Die Diagnosestatistik ist als Primärerhebung angelegt, tatsächlich entspricht sie mehr einem Abzug prozessgenerierter Daten.

Nach der skizzierten Problematisierung des Konzeptes der Hauptdiagnose und ihrer Validität sollen Fragen erörtert werden, die sich auf den Einsatz des Instrumentes zur Erfassung der Diagnose beziehen. Die Hauptdiagnose wird in den Krankenhäusern nicht als Klartext erfasst, sondern nach der Internationalen Klassifikation der Krankheiten und Todesursachen – ICD 9¹ – klassifiziert und kodiert. Die ICD dient dazu, die diagnostische Vielfalt in einheitlicher und differenzierter Weise zu behandeln, damit Vergleiche – auch internationale – möglich werden. Die richtige Vergabe eines Klassifikationscodes zu einer diagnostizierten Erkrankung ist für die spätere Qualität der Daten entscheidend. Die Klassifikation ist aber nicht ohne weiteres mit dem diagnostischen Schema des

¹ Eine Dokumentation der ICD 9 ist beim Deutschen Institut für Medizinische Dokumentation unter www.dimdi.de abrufbar. Ab dem Berichtsjahr 2000 wird die Hauptdiagnose nach der ICD 10 verschlüsselt.

praktizierenden Mediziners identisch. Neben der Frage der Verlässlichkeit einer Vercodung (würde die Wiederholung zum gleichen Ergebnis führen?) ist die Gültigkeit (verstehen alle Beteiligten das Gleiche darunter?) ein wichtiges Problem.

Ein kleines Beispiel mag das verdeutlichen: Nach einem Hinweis aus dem Bereich Gesundheitsberichterstattung der zuständigen Senatsverwaltung, wo aufgefallen war, dass der überwiegende Teil des deutschlandweiten Behandlungsaufkommens bei „Komplikationen mit der Nabelschnur“ von Berliner Krankenhäusern gemeldet wurde, führte die Prüfung zu einer Bestätigung der Diagnosestellung durch den dafür verantwortlichen führenden Mediziner auf dem Gebiet der Frauenheilkunde und Geburtshilfe. Das heißt, obwohl man von einer unterschiedlichen medizinischen Interpretation ein und desselben Krankheitsgeschehens ausgehen kann, ist die uns übermittelte Diagnose im Erhebungszusammenhang richtig.

Durch die Einführung des diagnoseorientierten Abrechnungssystems nach Fallpauschalen (Diagnosis Related Groups – DRG) gewinnt die standardisierte Datenhaltung und Diagnoseverschlüsselung erheblich an Bedeutung; dies dürfte sich durchaus zugunsten der Statistik auswirken können.

Die ICD ist ein Instrument zur standardisierten Erhebung eines Merkmals. Durch die Standardisierung wird ein festes Schema vorgegeben. Hierbei stellt sich die Frage, wie gut ist es der Aufgabestellung angepasst bzw. wie sorgfältig wird es eingesetzt? In der verwendeten Version umfasst die ICD die Positionen 001 bis 999 sowie die V-Klassifikation (nichtkranke Zustände). Da gesunde Neugeborene in der Statistik keine Berücksichtigung finden, stehen insgesamt 981 Positionen zur Signierung von Krankheiten, Verletzungen, Vergiftungen und nicht-kranken Zuständen zur Verfügung.

Empirisch besetzt waren im Berichtsjahr 1999 in Berlin davon 945 Kategorien. Nicht besetzt waren beispielsweise Ausprägungen wie Pest (ICD-Position 020) oder Cholera (ICD-Position 001). Die Kategorien wurden bei der Diagnosestellung ganz unterschiedlich ausgeschöpft. Werden die Behandlungshäufigkeiten nach der ICD 9 absteigend sortiert, dann entfallen 1999 insgesamt 80 % aller Behandlungsfälle auf nur gerundet 20 % der Kategorien, 90 % der Fälle erstrecken sich über 31 % der Klassifikationsausprägungen und noch für 95 % der Fälle reichen 42 % der Ausprägungen. Somit verteilen sich die fehlenden 5 % aller Behandlungsfälle über 58 % aller Positionen.

Die Verteilung weist deutliche Konzentrationen auf. Am stärksten besetzt ist die Einzeldiagnose ICD-Position 414 – Sonstige Formen von chronischen ischämischen Herzkrankheiten mit 26 284 Fällen. Die häufigste Hauptdiagnose ist folglich eine Restkategorie, die auf der Ebene des 3-Stellers nicht weiter differenziert ist.

Die Ausschöpfung der Kategorien der ICD 9 wird in der Tabelle 1 dargestellt.

Zwischen 1994 und 1999 lässt sich ein Trend ausmachen: Die Ausschöpfung der ICD 9 verengt sich. Während 1994 insgesamt 80 % der Krankenhausfälle durch 20,9 % der ICD-Klassen repräsentiert wurden, waren 1999 nur noch 19,5 % aller ICD-Schlüssel notwendig, um 80 % der Fälle abzubilden. Wird statt des 80 %-Wertes ein höherer Wert genommen, vergrößert sich die Differenz zwischen den Zeitpunkten.

Tab. 1 **Ausschöpfung der ICD 9 durch die Daten der Diagnosestatistik über die aus vollstationärer Behandlung in Berliner Krankenhäusern Entlassenen 1994 bis 1999**

Jahr	... % der kumulierten Fällen entfallen auf ... % der Kategorien			
	80 %	90 %	95 %	99 %
1994	20,9	33,2	45,2	67,7
1995	20,5	32,6	44,1	66,1
1996	19,9	31,8	43,0	64,7
1997	19,6	31,3	42,3	63,7
1998	19,7	31,4	42,3	62,7
1999	19,5	30,9	41,5	62,4

Die starke Besetzung einer Restkategorie kann verschiedene Ursachen haben. Eine mögliche Ursache könnte in der nicht ausreichenden Differenziertheit bzw. Aktualität der Klassifikation begründet sein, da das medizinische Wissen ständig fortschreitet und eine entsprechende Anpassung der Klassifikationen erforderlich macht². Ursächlich könnte auch die konkrete Anwendung der Klassifikation sein. Angabefehler bei der ICD-Kodierung können anhand der Daten nur sehr eingeschränkt überprüft werden.

Das Adäquationsproblem stellt sich auf der interpretatorischen Seite gleichfalls, wie gezeigt werden soll.

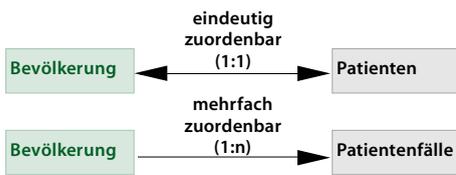
Die empirische Grundgesamtheit umfasst alle im Berichtszeitraum aus vollstationärer Versorgung in einem Krankenhaus entlassenen Patienten. Sie bilden die Beobachtungsgesamtheit. Welche *empirisch zulässigen Aussagen* können auf der Basis dieser Daten gemacht werden? Alle Aussagen, die keinen individuellen Patientenbezug aufweisen, verletzen nicht das in der Erhebung verwendete Fallmodell. Zulässig ist z. B. die Aussage, dass 1999 insgesamt 610 354 Behandlungsfälle einschließlich Stundenfälle (nachfolgend Patienten) aus Berliner Kliniken entlassen wurden; nicht zulässig wäre es aber, wenn behauptet würde, dass 610 354 Berliner aus stationärer Behandlung entlassen wurden, weil die zweite Aussage auf Individuen abstellt.

Aus den Krankheitshäufigkeiten der Diagnosestatistik kann nicht direkt auf die Auftretenshäufigkeit von Erkrankungen in der Bevölkerung geschlossen werden; dies wäre nur möglich, wenn Diagnosedaten einem einzelnen Patienten zuordenbar wären. Ein identifizierendes Merkmal ist – aus Gründen des Datenschutzes – in der Erhebung aber nicht vorhanden. Da auch das Geburtsdatum nur monatsgenau erhoben wird, sind Rekonstruktionen einer Fallidentität unter Verwendung von Plausibilitätsannahmen nicht erfolgreich darstellbar.

Diesen Sachverhalt soll die folgende Darstellung veranschaulichen. In den Diagnosedaten ist die Beziehung zwischen Bevölkerung und Behandlungsfall keine 1:1-Beziehung, sondern eine 1:n-Beziehung. Ein Patient in den Diagnosedaten entspricht nicht einer Person in der Bevölkerung, sondern die Bevölkerung bringt Behandlungsfälle hervor, wobei eine Person mehrere Behandlungsfälle generieren kann. Über den Prozess der Generierung liegen nur wenige Informationen vor, die zudem nicht systematisch sind.

² Mit dem Berichtsjahr 2000 werden die Diagnosen nach der ICD 10 verschlüsselt, die erheblich feiner differenziert ist; die Differenzierungsleistung wird von den Krankenhäusern erbracht werden müssen.

Struktur der logischen Beziehung zwischen Bevölkerung und Patienten bzw. Behandlungsfällen



Eine weitere Fehlinterpretation ergibt sich aus dem Unterschied zwischen Hauptdiagnose und Neuerkrankung. Ein krasses Beispiel – aber leider kein konstruiertes – stellt ein bezirklicher Krebs-

atlas auf der Basis der Diagnosedaten dar. Das wäre nur dann sinnvoll, wenn zumindest Daten über Neuerkrankungen (Inzidenzen) vorliegen würden, darüber sagt die Diagnosestatistik jedoch nichts aus.

Die folgende Tabelle stellt den Effekt einer personenbezogenen und einer fallbezogenen Erhebung auf den arithmetischen Mittelwert des Alters und der Verweildauer gegenüber. Bei einer personenbezogenen Erhebung wird das Durchschnittsalter nicht von der Aufenthaltshäufigkeit beeinflusst; bei einer fallbezogenen Erhebung liegt das Durchschnittsalter im gewählten Beispiel höher und es ist auf die Gesamtmasse der Patientenfälle zu beziehen.

Patient	Personenbezogene Erhebung			Fallbezogene Erhebung		
	Nr. des Aufenthaltes	Alter in Jahren	Verweildauer in Tagen	Patient	Alter in Jahren	Verweildauer in Tagen
A	1	10	10	A	10	10
B	1	100	15	B	100	15
	2		5	B	100	5
Arithmetisches Mittel		55,00	10,00		70,00	10,00

Fehler statistischer Daten

Die Ädaquation bezeichnet die Übergänge zwischen dem Realbereich und dem Formalbereich der Statistik. Innerhalb des Formalbereichs Statistik spielt die interne Qualitätsbewertung und -sicherung traditionell eine zentrale Rolle. Für die Diagnosestatistik liegen die Erfahrungen und die Daten von sieben Berichtsperioden zwischen 1993 und 1999 vor.

Der Wirtschaftsstatistiker Peter von der Lipp schreibt: „Statistiken sollten so genau sein, wie dies für den Auswertungszweck erforderlich und angesichts der Kosten vertretbar ist. Es ist nicht das Problem, ob Statistiken Fehler haben, sondern ob man den Fehlerbereich kennt bzw. abschätzen kann. Eine ‚amtliche‘ Zahl muss von einschätzbarer Genauigkeit sein.“ (von der Lippe 1996, S.42)

Die Daten müssen richtig und verlässlich sein. Es geht um die Vermeidung von Fehlern und dort, wo sie nicht vermieden werden können, um deren Abschätzung.

In der Methodik statistischer Erhebungen wird grundsätzlich zwischen

- Stichprobenfehlern und
- Nichtstichprobenfehlern sowie zwischen
- Zufallsfehlern und
- systematischen Fehlern

unterschieden. Jedes Ergebnis, das auf einer Stichprobe beruht, ist mit einem Zufallsfehler behaftet (Stichprobenzufallsfehler). Über die Genauigkeit des Ergebnisses lassen sich Angaben machen, z. B. dadurch, dass ein Vertrauensbereich berechnet wird, der nur mit einer bestimmten Wahrscheinlichkeit über- bzw. unterschritten wird. Problematischer sind neben den Zufallsfehlern, die nicht vermeidbar, aber behandelbar sind, systematische Fehler, die zu einer Verzerrung der Ergebnisse (Bias) führen.

Ein elementares Fehlermodell könnte wie folgt aussehen:

$$\text{Angabewert} = \text{wahrer Wert} + \text{systematischer Fehler} + \text{Zufallsfehler}$$

Der Angabewert der empirischen Daten ergibt sich aus dem wahren Wert und den Komponenten Zufallsfehler und systematischer Fehler.

Bei der Diagnosestatistik treten die klassischen Stichprobenzufallsfehler nicht auf, da sie eine Totalerhebung aller entlassenen Patienten ist und keine zufällige Auswahl. Dennoch sind die Diagnosedaten selbstredend nicht fehlerfrei. Ein nicht unerheblicher Teil des Aufwandes, den die Statistik betreibt, geht in die Vermeidung und Beseitigung der vorhandenen *Nichtstichprobenfehler*.

In der folgenden Übersicht sind entlang der Prozessschritte für die Erstellung der Diagnosestatistik mögliche Fehlerarten aufgelistet. Die Aufstellung beginnt mit der Erhebungsvorbereitung und damit dem Aufbau des Berichtskreises. Zu den Erhebungsvorbereitungen gehört auch das Erstellen der Erhebungsunterlagen. Ein besonders kritischer Punkt sind hierbei die Definitionen der Merkmale und die Anleitung für die Datenbereitstellung. Der Abschnitt über die Erhebungsfehler umfasst die Probleme des Rücklaufs. Hier verfügt die amtliche Statistik mit der Auskunftspflicht über ein vergleichsweise scharfes Instrument. Für die Qualität der gewonnen Daten ist die Vermeidung von Angabefehlern in dieser Phase entscheidend.

In der sich an den Dateneingang anschließenden Plausibilisierung wird das Material nach formalen und inhaltlichen Kriterien geprüft. Die Plausibilisierung des Materials ist eine operative Kontrolle, da die Daten verändert werden (vergl. Federal Committee on Statistical Methodology, Statistical Policy Working Paper 18). Zweifelsfragen werden in wichtigen Fällen durch Rückfrage bei den Befragten geklärt.

Nach der Plausibilisierung des Materials stehen die Daten für die Auswertung zur Verfügung.

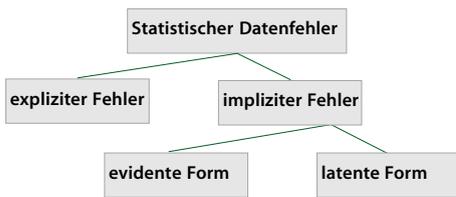
Prozessschritt	Fehler	Führt zu ...
Erhebungsvorbereitung	fehlerhaft aufgebauter Berichtskreis	<ul style="list-style-type: none"> • fehlenden Einheiten (undercoverage) • irrtümlich aufgenommenen Einheiten (overcoverage)
Erhebung	falsch oder unzureichend operationalisierte Merkmale	• Inhaltsfehlern
	falsch oder unzureichend beschriebene Merkmale in den Erhebungsunterlagen	• Inhaltsfehlern
	Komplettausfälle beim Rücklauf durch Verweigerung	<ul style="list-style-type: none"> • fehlenden Einheiten (Unit-Nonresponse)
	Teilausfälle beim Rücklauf durch unvollständige Angaben	<ul style="list-style-type: none"> • fehlenden Merkmalen (Item Nonresponse)
Aufbereitung	fehlerhafte Eingaben	• Angabefehlern
	fehlerhafte Datenaufbereitung beim Befragten	<ul style="list-style-type: none"> • Angabefehlern • fehlenden Fällen • doppelten Fällen
	Dateimanagement- und Übermittlungsfehler	<ul style="list-style-type: none"> • fehlenden Daten • Doppellieferung • Daten aus falschem Berichtszeitraum
Auswertung	Dateneingangsfehler durch falsche Zuordnung von Dateien	<ul style="list-style-type: none"> • fehlenden Fälle • doppelten Fällen
	Programmfehler der Aufbereitungsprogramme	<ul style="list-style-type: none"> • Datenfehlern • unentdeckten Inplausibilitäten • Datenfehlern
Auswertung	Plausibilitätsfehler	• Datenfehlern
	Adäquationsfehler bei Merkmalsauswahl	• inhaltlichen Fehlern
	Adäquationsfehler bei der Wahl der Auswertungsmethode	• inhaltlichen Fehlern

Eine Zusammenfassung der Wirkungen der Fehler führt zu folgender Zusammenstellung:

- fehlende Einrichtungen
- irrtümlich erhobene Einrichtungen
- falsche Berichtszeiträume
- Dubletten (Doppelerfassungen)
- fehlende Merkmalsangaben
- Angabefehler
- Inhaltsfehler

Strecker (Strecker, S. 402) unterscheidet die Fehler nach ihrer Erkennbarkeit in explizite Fehler, das sind erkennbare (Explikation = Sichtbarmachen) und quantitativ abschätzbare, und implizite Fehler, die zwar in den Daten vorhanden sind, aber nicht quantitativ dargestellt werden können. Für die Bestimmung des expliziten Fehlers werden Daten größerer Genauigkeit benötigt, die durch Controllerhebungen gewonnen werden

Statistische Datenfehler nach ihrer Erkennbarkeit



müssten. Strecker wechselt die Perspektive, statt einer Betrachtung der Fehlerkomponenten liegt das Augenmerk hier auf der datenorientierten Sichtbarkeit von Fehlern.

In der Praxis treten nahezu ausschließlich implizite Fehler auf. Die impliziten Fehler werden in evident und latent unterteilt. Evident sind sie, wenn der Fehler an den Daten qualitativ sichtbar hervortritt oder durch Plausibilitätsprüfungen sichtbar gemacht, jedoch quantitativ nicht bestimmt werden kann. Latent ist ein Fehler, wenn er nicht sichtbar hervortritt, weder in den Daten selbst, noch an Maßzahlen oder grafischen Darstellungen. „Diese Form der Fehler in den statistischen Daten kommt in der Praxis am häufigsten vor“ (Strecker, 420).

Die Erkennung impliziter Fehler ist eine der großen Herausforderungen der Datenprüfung. Anhaltspunkte können längsschnittliche Vergleiche ergeben, wenn Veränderungen gefunden werden, für die keine sachliche Begründung gegeben werden kann.

Durch *Plausibilitätsprüfungen* können viele Fehler erkannt und eliminiert werden. Für eine erste Bewertung der Diagnosedaten nach der Plausibilisierung sollen die fehlenden Werte der Jahre 1994 bis 1999 betrachtet werden. *Fehlende Werte* können durch nicht vorhandene Angaben im Liefermaterial entstehen und durch entdeckte Angabefehler, die sich nicht klären ließen und dann auf den Schlüssel für fehlende Angaben gesetzt wurden.

Tabelle 2 gibt die fehlenden Werte für die Erhebungsmerkmale wieder. Deutliche Verbesserungen wurden bei der Verschlüsselung nach der ICD 9 und der Fachabteilung erzielt. Die Verbesserungen sind eine direkte Folge der sehr engen Kontakte zu den Kliniken und der systematischen Rückkopplung lokalisierter Fehler zu Beginn einer neuen Berichtsperiode über klinikspezifische Fehlerhinweise, die als personalisierte Schreiben versandt werden (Abbildung 2).

Die fehlenden Werte sind ein wichtiger Indikator für die Qualitätssicherung der Daten. Allerdings sagt die Häufigkeit von fehlenden Werten nur etwas über die Vollständigkeit der Daten aus; sie ist für die Beurteilung der Zuverlässigkeit und Gültigkeit aber nicht ausreichend. Daten können vollständig sein und dennoch Fehler enthalten. Ein Vergleich mit den Vorjahresergebnissen und mit den Grunddaten, die zwar zum System der Krankenhausstatistik gehören, aber unabhängig von den Diagnosedaten gewonnen werden, kann weitere Aufschlüsse geben (Tabelle 3). Für den Vergleich ist es nicht so wichtig, welche Daten verlässlicher sind; entscheidend ist, ob die Daten Differenzen aufweisen und wie groß diese Differenzen sind. Da die Statistiken überlappend bearbeitet werden, können Abweichungen zwischen ihnen sowohl zu Korrekturen der Diagnosedaten wie der Grunddaten führen.

Die Diagnosedaten enthielten bis auf das Berichtsjahr 1996 mehr Entlassungen als die Grunddaten. Möglicherweise wurden teilstationäre Fälle gemeldet. Die geringsten Abweichungen traten 1996 und 1999 auf. Die Betragsdifferenz bei den Sterbefällen und den Stundenfällen nimmt nahezu konstant ab. Die Reihe mit der durchschnittlichen Verweildauer zeigt 1996 einen Ausreißer, in den anderen Jahren war die Differenz stets kleiner als ein halber Tag. Gravierender waren die Unterschiede bei den Stundenfällen; erst mit dem Berichtsjahr 1997 konnte eine deutliche Verbesserung erreicht werden. Bezogen auf die Gesamtzahl der Fälle sind die Differenzen – mit Ausnahme der Stundenfälle bis 1996 – als sehr gering zu betrachten.

Bei dem Vergleich mit den Grunddaten werden die Daten auf einem stark aggregierten Niveau verglichen. Unterhalb der Insegsamtebene können sich erhebliche Abweichungen verbergen, die durch die Zusammenfassung überdeckt werden. Im Folgenden sollen *Daten auf der Ebene der Einrichtungen* betrachtet werden. Untersucht werden sollen die ausgewählten Merkmale „Operation“, „Sterbefall“ und „Stundenfall“.

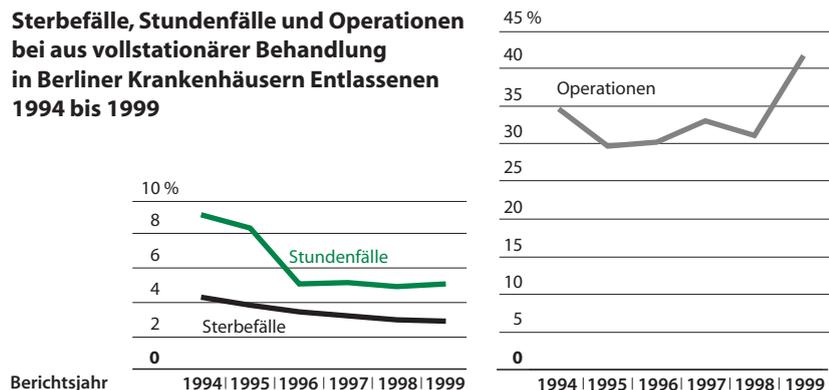
In der Tabelle 4 sind die Mittelwerte und die Standardabweichungen zusammengestellt. Im Verlauf zeigt sich ein sprunghafter Anstieg der Operationen 1998/1999

von 31,0 % auf 41,7 % (siehe auch Abbildung 2). Die Streuung der Operationsanteile war 1996 am stärksten. Die Sterbefälle nahmen zwischen 1994 und 1999 kontinuierlich ab, was einerseits dem Trend in der Bevölkerungsstatistik entspricht und andererseits Reformen im Berliner Gesund-

Tab.2 **Fehlende Werte in den Diagnosedaten über aus vollstationärer Behandlung in Berliner Krankenhäusern Entlassene 1994 bis 1999**

Jahr	Fallzahl insgesamt	Merkmal									
		Geschlecht	Geburtsdatum	Aufnahmedatum	Entlassungsdatum	ICD 9 (3-stellig)	Patientenwohnort		Sterbefall	Operation	Fachabteilung
							Berliner Bezirke	Bundesland, Ausland			
1994	646 830	4	17	24	1	624	1 749	1 193	2	0	744
1995	642 775	0	24	24	0	1 000	44	666	0	0	12 060
1996	639 853	0	0	6	0	168	273	654	0	0	7 160
1997	653 349	0	1	0	0	497	735	496	0	0	0
1998	686 098	0	0	0	0	140	663	158	0	0	2
1999	693 393	3	1	0	0	8	389	1 078	0	0	0

Abb. 2 **Sterbefälle, Stundenfälle und Operationen bei aus vollstationärer Behandlung in Berliner Krankenhäusern Entlassenen 1994 bis 1999**



Tab. 3 **Vergleich ausgewählter Merkmale zwischen Grund- und Diagnosedaten in Berlin 1994 bis 1999 (einschließlich Stundenfälle)**

Jahr	Entlassene insgesamt		Sterbefälle		Stundenfälle		Verweildauer	
	Grunddaten	Diagnose	Grunddaten	Diagnose	Grunddaten	Diagnose	Grunddaten	Diagnose
1994	639 689	646 830	25 665	26 748	39 750	58 871	18,8	18,7
1995	630 715	642 775	24 639	24 043	31 536	53 081	17,8	17,9
1996	639 918	639 853	21 945	21 295	27 623	31 604	14,5	18,0
1997	648 199	653 349	20 178	20 122	31 657	32 875	12,8	13,2
1998	676 990	686 098	19 639	19 631	34 302	33 299	12,0	11,9
1999	693 335	693 393	19 278	19 278	33 599	34 191	11,2	11,2
Abweichung zwischen Grunddaten und Diagnosedaten (Diagnose – Grunddaten)								
1994		7 141		1 083		19 121		- 0,1
1995		12 060		- 596		21 545		0,1
1996		- 65		- 650		3 981		3,5
1997		5 150		- 56		1 218		0,4
1998		9 108		- 8		- 1 003		- 0,1
1999		58		0		592		0,0
Abweichung in %								
1994		1,1		4,0		32,5		- 0,7
1995		1,9		- 2,5		40,6		0,6
1996		- 0,0		- 3,1		12,6		19,4
1997		0,8		- 0,3		3,7		3,0
1998		1,3		- 0,0		- 3,0		- 0,8
1999		0,0		0,0		1,7		0,0

heitswesen geschuldet sein dürfte, da chronisch Kranke und Patienten der Geriatrie nicht mehr zum Sterben ins Krankenhaus kommen, sondern durch eine Vielzahl von Betreuungsangeboten im Sozialbereich aufgefangen werden.

Bei den Stundenfällen gab es 1995/96 einen starken Rückgang von 8,3 % auf 5,0 %; allerdings zeigte sich in diesem Zeitraum eine deutliche Abweichung zu den Grunddaten. Begründet sind diese Abweichungen im hohen Anteil ambulanter Operationen, die nach Recherche zurückliegender Berichtszeiträume durch eine Reihe von Krankenhäusern in der Diagnosestatistik als Stundenfall ausgewiesen wurden. Ab 1996 war nicht nur der Anteil der Stundenfälle deutlich geringer, auch die Streuung wies kleinere Werte auf. Der Anteil der Stundenfälle streute von Krankenhaus zu Krankenhaus nicht mehr so stark.

Die Streudiagramme in der Abbildung 3 stellen die Struktur der Daten auf der Ebene der Einrichtungen dar. In den Streudiagrammen werden die Merkmale als Anteilswerte abgetragen. Verglichen werden immer zwei Zeitpunkte; jeder Punkt in dem Diagramm repräsentiert eine Einrichtung. Wenn die Anteilswerte einer Einrichtung zu beiden Zeitpunkten gleich waren, dann liegt der Punkt genau auf der Diagonale, die ebenfalls eingezeichnet ist. Liegt der Punkt unterhalb der Diagonale, dann ist

der Anteilswert gegenüber dem Vergleichsjahr geringer; liegt der Punkt oberhalb der Diagonalen, dann ist der Anteilswert gestiegen.

In der Grafik wird das Operationsmerkmal in ausgewählten Vergleichszeiträumen dargestellt. Beim Vergleich 1995 zu 1996 zeigt sich, dass einige Einrichtungen, die zuvor hohe Operationsanteile hatten, kaum oder keine Operationen mehr durchführen; im Mittel hat sich dies, wie Tabelle 4 zeigt, zwar nicht ausgewirkt, aber dennoch kommt es auf der Einrichtungsebene zu bemerkenswerten Veränderungen der OP-Anteile. In den Streudiagrammen ordnen sich die abgetragenen Punkte stärker um die Diagonale; die Anteile zeichnen sich durch eine höhere Stabilität im Zeitvergleich aus. Dies würde man auch erwarten, da die Operationshäufigkeit von der Struktur der Einrichtung (OP-Räume und Personal) mitbestimmt sein dürfte und nur in Ausnahmefällen kurzfristige Änderungen zu erwarten sein dürften.

Anknüpfend an diese Annahme soll ein einfaches *Stabilitätsmodell* vorgestellt werden, das weitere Aufschlüsse für die Zuverlässigkeit der Daten liefern soll. Ausgangspunkt ist die Hypothese einer Stabilität der Merkmale im Zeitpaarvergleich. Die Anteile der Operationen, der Sterbe- und Stundenfälle sollte zwischen zwei

Erhebungen in etwa gleich bleiben. Änderungen, z. B. ein Anstieg der Stundenfälle, sollte alle Einrichtungen betreffen. In einem Bundesland mit einem gezielten Wandel im stationären Sektor ist aber durchaus mit Ausnahmen von dieser Stabilitätsregel zu rechnen. Zumindest lassen sich so Veränderungen auf der Einrichtungsebene zeitlich lokalisieren, die auf der übergeordneten Aggregatebene möglicherweise neutralisiert sind. Das zugehörige lineare Regressionsmodell für das Operationsmerkmal sieht wie folgt aus

$$\% \text{ Operationen Zeitpunkt } t_2 = \text{Steigungskoeffizient} * \% \text{ Operationen Zeitpunkt } t_1 + \text{Konstante}$$

Mit den anderen Merkmalen wurde ebenso verfahren. Interpretiert werden soll nur der Determinationskoeffizient R^2 der gewichteten Regression, der die Güte der Anpassung zwischen Regressionsgerade und Beobachtungswerten ausdrückt. Angenommen, die Operationsanteile im Berichtsjahr t_2 entsprächen genau denen des vorherigen Berichtsjahres t_1 , dann betrüge der Determinationskoeffizient 1,00 oder 100 %. Bei einer additiven Änderung (jedes Krankenhaus hat einen Zugewinn von z. B. 5 Prozentpunkten) oder bei einer proportionalen Änderung (jedes Krankenhaus hat einen Zugewinn von 5 %) läge der Determinationskoeffizient ebenfalls bei 1,00. Besteht kein linearer Zusammenhang zwischen den Berichtszeiträumen, dann beträgt der Koeffizient Null.

Abb.3 Anteil der Behandlungsfälle mit Operationen im Zusammenhang mit der Hauptdiagnose in Berliner Krankenhäusern 1995 bis 1999

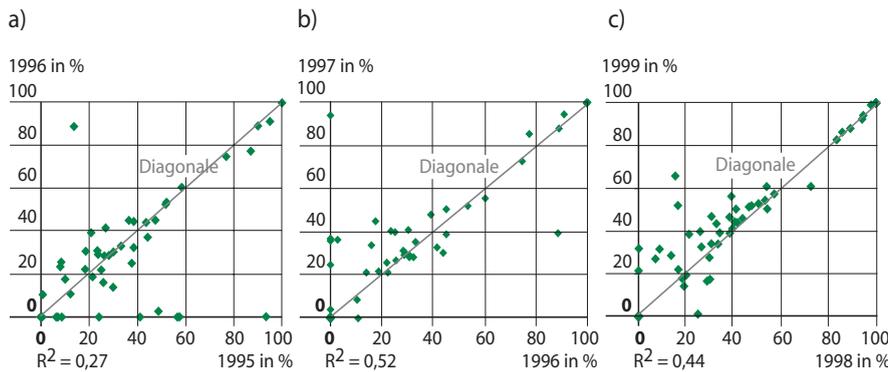
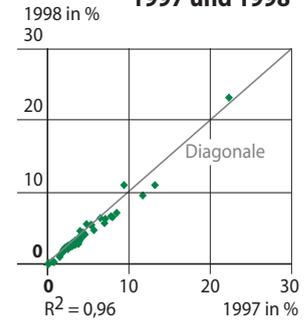


Abb.4 Anteil der Verstorbenen an den vollstationären Behandlungsfällen in Berliner Krankenhäusern 1997 und 1998



Tab.4 Mittelwert und Standardabweichung der Merkmale „Operation“, „Sterbefall“ und „Stundenfall“ bei aus vollstationärer Behandlung in Berliner Krankenhäusern Entlassenen 1994 bis 1999

Jahr	Operation in % aller Behandlungsfälle ohne Stundenfälle		Sterbefall in % aller Behandlungsfälle		Stundenfall in % aller Behandlungsfälle	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
1994	34,6	17,4	4,2	4,2	9,1	7,3
1995	29,7	19,8	3,8	3,9	8,3	6,8
1996	30,2	21,9	3,4	2,1	5,0	2,5
1997	33,0	18,3	3,1	1,7	5,1	2,5
1998	31,0	19,2	2,9	1,7	4,9	2,7
1999	41,7	18,3	2,8	1,6	5,0	3,4

Der Determinationskoeffizient fällt beim Operationsmerkmal überraschend gering aus. Nur beim Berichtsjahrapaar 1997/96 liegt er oberhalb von 50 % (Tabelle 5). Dagegen erreicht das Merkmal „Sterbefälle“ sehr hohe R²-Werte. Bildlich zeigt sich dies auch in Abbildung 4, in der sich die Wertepaare zu einer langgestreckten Wolke gruppieren. Uneinheitlicher sind die Koeffizientenwerte beim Merkmal „Stundenfälle“, hier zeigt sich insbesondere vor 1997 eine große Instabilität.

Die Koeffizientenwerte des Operationsmerkmals weisen auf eine zunächst unerwartet hohe Instabilität hin. Dies spricht gegen die Zuverlässigkeit des Merkmals. Ähnliche Hinweise ergaben sich auch aus Rückfragen der Krankenhäuser, die auf eine große Unsicherheit bei der Angabe des OP-Merkmals schließen ließen. Das Operationsmerkmal sollte nur nach sorgfältiger Abwägung verwendet werden. Eine überarbeitete Spezifizierung des Merkmals ist dringend erforderlich, um zuverlässige und gültige Angaben zu ermöglichen.

Tab.5 Determinationskoeffizient¹ zwischen den Merkmalen „Operation“, „Sterbefall“ und „Stundenfall“ bei aus vollstationärer Behandlung in Berliner Krankenhäusern Entlassenen 1994 bis 1999

In Beziehung gesetzte Berichtsjahre	Opera-tion	Sterbe-fall	Stunden-fall
	in %		
1995/94	33	96	34
1996/95	27	72	02
1997/96	52	85	91
1998/97	37	96	85
1999/98	44	93	65

¹ Determinationskoeffizient der gewichteten linearen Regression der Anteilswerte der Berichtsjahrapaare

Zusammenfassung

Durch die 1993 eingeführte Diagnosestatistik wurde die Datenlage zu gesundheitlichen Themen qualitativ und quantitativ deutlich erweitert. Das komplexe Kranken- und Krankenhausgeschehen stellt an die Datenerstellung sehr hohe Anforderungen. An der Verbesserung der Datenqualität wurde seit Beginn der Statistik gearbeitet. Dabei hat sich der intensive Austausch mit den Nutzern dieser Daten als überaus hilfreich erwiesen, weil durch die analytische Arbeit latente Fehler aufgespürt werden konnten. Die Prozessqualität des Arbeitssystems innerhalb des Statistischen Landesamtes (eingesetzte Programme, Mitarbeiterschulung, Prüfpunkte) wurden innerhalb des praktizierten Total Quality Managements (TQM) kontinuierlich verbessert. Weitere Fortschritte könnten durch die Etablierung der Statistik als integralen Bestandteil der IT-gestützten Geschäftsprozesse in den Krankenhäusern und durch eine stärkere Fehlerbeseitigung schon bei der Datengewinnung (Wein,2002) zu erzielen sein.

Quellennachweis

- Deutsche Kodierrichtlinien. Allgemeine und Spezielle Kodierrichtlinien für die Verschlüsselung von Krankheiten und Prozeduren - Version 2002, hrsg. von der Deutschen Krankenhausgesellschaft (DKG), den Spitzenverbänden der Krankenkassen (GKV) und dem Verband der privaten Krankenversicherung (PKV), URL: www.g-drg.de/deutschesdrg/drg_kodier.htm
- Gesetz für die Statistik für Bundeszwecke (Bundesstatistikgesetz - BStatG) vom 22. Januar 1987 (BGBl. Teil I S. 462, 565, zuletzt geändert durch Artikel 3 des Gesetzes vom 21. Dezember 2000 (BGBl. I S. 1857).
- Gesetz zur wirtschaftlichen Sicherung der Krankenhäuser und zur Regelung der Krankenhauspflegesätze (Krankenhausfinanzierungsgesetz – KHG) in der Fassung der Bekanntmachung vom 10. April 1991 (BGBl. I S. 886), zuletzt geändert durch Gesetz vom 22. Dezember 1999 (BGBl. I S. 2626).
- Sozialgesetzbuch (SGB), Fünftes Buch (V) – Gesetzliche Krankenversicherung - vom 20. Dezember 1988 (BGBl. Teil I S. 2477).
- Verordnung über die Bundesstatistik für Krankenhäuser (Krankenhausstatistikverordnung – KHStatV) vom 10. April 1990 (BGBl. I S. 730).

Literaturverzeichnis

- Bundesministerium für Arbeit und Sozialordnung (Hrsg.): Leitfaden zur Erstellung der Diagnosestatistik nach § 16 Bundespflegesatzverordnung (BPflV), 1986.
- Chlumsky, J., Wiegert, R. u. a. (1993): Qualität statistischer Daten. Beiträge zum wissenschaftlichen Kolloquium am 12./13. November 1992 in Wiesbaden (=Bd. 25 der Schriftenreihe Forum der Bundesstatistik hrsg. vom Statistischen Bundesamt), Stuttgart.
- Federal Committee on Statistical Methodology, Statistical Policy Working Paper 18 - Data Editing in Federal Statistical Agencies, URL: www.fcsm.gov/working-papers/wp18.html
- Gräß, Christopher (1996): Die neue Krankenhausdiagnosestatistik. Ergebnisse für 1993, Wirtschaft und Statistik, H. 2.
- Grohmann, Heinz (1985): Vom theoretischen Konstrukt zum statistischen Begriff. Das Äquivalenzproblem (Vortrag auf der 55. Jahreshauptversammlung der Deutschen Statistischen Gesellschaft am 25. Oktober 1984 in Augsburg), Allg. Statistisches Archiv, Bd. 69, H. 1.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (Hrsg.): Wege zu einer besseren Informationellen Infrastruktur. Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (= KVI-Gutachten), Baden-Baden, 2001.
- Statistisches Amt der Europäischen Gemeinschaften (Hrsg.): Quality in Statistics, Sigma, H. 3, 1998.
- Statistisches Landesamt Berlin (2001): Bericht über die zweite Selbstbewertung nach dem EFQM-Modell 2000/2001.
- Sommer, R., Unholzer, G., Wiegand, E. (1998): Standards zur Qualitätssicherung in der Markt- und Sozialforschung, hrsg. vom Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute, URL: www.adm-ev.de
- Strecker, Heinrich (1995): Ein Beitrag zu Fehlern in statistischen Erhebungen (Der Netto- und Bruttofehler sowie Beispiele für besondere Fehlerursachen), Allg. Statistisches Archiv, Bd. 79, H. 4, Göttingen.
- von der Lippe, Peter (1996): Wirtschaftsstatistik. Amtliche Statistik und Volkswirtschaftliche Gesamtrechnung, 5. Auflage, Stuttgart.
- Wein, Elmar (2002): Datenqualität und Datenprüfungen, Wirtschaft und Statistik