

Faktische Anonymisierung der Diagnosedaten der Krankenhauspatienten

Dr. Olaf Schoffer, Sebastian Rolland und Gabriele Philipp

1	Einführung	1
1.1	Einzeldaten für die Wissenschaft	1
1.2	Faktische Anonymisierung der Diagnosedaten der Krankenhauspatienten	2
1.2.1	Datenerhebung und -grundlage	2
1.2.2	Einbeziehung potentieller Nutzer und Vorgehensweise	2
2	Anonymisierung	2
2.1	Auswahl der Variablen.....	3
2.2	Vergrößern der Ausprägungen	3
2.3	Verschieben der Ausprägungsklassen	3
2.4	Entfernen von Beobachtungen	4
2.5	Geheimhaltung von Ausprägungen	4
3	Datenangriffe	4
3.1	Datensatzinterne Angriffe	4
3.2	Datensatzübergreifende Angriffe.....	5
3.3	Angriffe mit Zusatzinformationen.....	5
3.4	Umgang mit deanonymisierten Merkmalsträgern.....	5
4	Fazit	5

1 Einführung

1.1 Einzeldaten für die Wissenschaft

Bei der Novellierung des Bundesstatistikgesetzes (BStatG) im Jahre 1987 wurde der Wissenschaft durch die Schaffung des § 16 Abs. 6 ein privilegierter Zugang zu Mikrodaten der amtlichen Statistik ermöglicht.¹ Danach „dürfen vom Statistischen Bundesamt und den statistischen Ämtern der Länder Einzelangaben an Hochschulen und sonstige Einrichtungen mit der Aufgabe unabhängiger wissenschaftlicher Forschung übermittelt werden, wenn die Einzelangaben nur mit unverhältnismäßig großem Aufwand an Zeit, Kosten und Arbeitskraft zugeordnet werden können“². Das Unverhältnismäßigkeitsgebot impliziert, dass eine Verletzung der Anonymität von Merkmalsträgern nur bei nutzbringenden Zuordnungen gegeben ist.³ Damit wird vom Gesetzgeber keine absolute Anonymität mehr vorausgesetzt, sondern eine faktische wird als ausreichend erachtet.

¹ Vgl. Scheffler, Michael (2005): Ein Scientific-Use-File der Einzelhandelsstatistik 1999, in: WiSta, 3/2005, S. 197.

² § 16 Abs. 6 BStatG.

³ Vgl. Höhne/Sturm/Vorgrimler (2003): Konzept zur Schutzwirkung faktischer Anonymisierung, in: WiSta, 4/2003, S. 287.

Damit faktisch anonymisierte Daten von der Wissenschaft angenommen werden, muss eine Anonymisierung zwei gleichrangigen Herausforderungen gerecht werden: sie muss einerseits einen ausreichenden Schutz der Einzelangaben gewährleisten und andererseits die Analyse-möglichkeiten der anonymisierten Daten bestmöglich erhalten.

1.2 Faktische Anonymisierung der Diagnosedaten der Krankenhauspatienten

1.2.1 Datenerhebung und -grundlage

Im Rahmen der amtlichen Krankenhausstatistik werden seit 1993 Diagnosedaten der Krankenhauspatienten in einer jährlich durchgeführten Vollerhebung erfasst. Die Meldung erfolgt seitens der Krankenhäuser, die ihre Patientendaten einmal im Jahr (bis zum 1. April des Folgejahres) an ihr zuständiges Statistisches Landesamt übermitteln. Darüber hinaus melden die Krankenhäuser einrichtungsbezogene Informationen, wie z. B. die Anzahl ihrer Betten und Fachabteilungen. Obwohl es sich um zwei getrennte Erhebungsbereiche handelt (Grunddaten und Diagnosedaten), können dem patientenbezogenen Diagnosedatensatz einrichtungsbezogene Variablen des Krankenhauses zugespielt werden. Dieser Datensatz dient als Grundlage für die Erstellung des Scientific-Use-Files. Über den Umfang des Ausgangsmaterials informiert die Datensatzbeschreibung im Anhang (siehe Anhang 1).

1.2.2 Einbeziehung potentieller Nutzer und Vorgehensweise

Bei der Entwicklung des faktisch anonymisierten Diagnosedatensatzes standen die Interessen potentieller Nutzer im Vordergrund. Daher begleitete ein wissenschaftlicher Beraterkreis aus Vertretern epidemiologischer Institutionen das Vorhaben. Das Informationsbedürfnis der Wissenschaftler wurde sodann vor dem Hintergrund der Möglichkeiten der Krankenhausstatistik und der Vorgabe faktischer Anonymität überprüft.

Um beiden Herausforderungen, dem Schutz der Einzelangaben einerseits und bestmöglichen Analyse-möglichkeiten andererseits, gerecht zu werden, wurden zwei Datensätze erzeugt, die unterschiedliche wissenschaftliche Fragestellungen bedienen sollen. Ein Datensatz enthält vorwiegend patientenbezogene Daten und soll in erster Linie epidemiologischen Studien dienen. Ein zweiter Datensatz ist dagegen vornehmlich an gesundheitsökonomischen und -politischen Fragestellungen ausgerichtet und beinhaltet vorwiegend krankenhausbazogene Daten. Dabei gibt es Merkmale, die in beiden Datensätzen enthalten sind, wie die jeweiligen Datensatzbeschreibungen im Anhang zeigen (siehe Anhang 2 und Anhang 3).

Durch die Trennung der patienten- von den krankenhausbazogenen Angaben soll von vornherein ausgeschlossen werden, dass Patienten und Krankenhäuser über zusätzliche Informationen identifiziert werden können. Dadurch ist eine tiefere Gliederung der Merkmale in den einzelnen Datensätzen möglich.

2 Anonymisierung

Eine Anonymisierung von Merkmalsträgern ist immer mit einem Informationsverlust der dazugehörigen Daten verbunden.⁴ Um diesen Verlust so gering wie möglich zu halten, sollten diejenigen Merkmalsträger schwächer anonymisiert werden, die einem geringeren Risiko ausgesetzt sind.⁵ Risiko bedeutet in diesem Zusammenhang die Wahrscheinlichkeit, aufgrund der Datenweitergabe negative Konsequenzen zu erleiden.

Für einen Patienten als Merkmalsträger besteht in erster Linie das Risiko, dass der Grund seines Krankenhausaufenthaltes – die Hauptdiagnose – von Personen (z. B. Arbeitgeber) herausgefunden werden kann, die diese Information missbrauchen. Dies kann zu negativen Konsequenzen für den Patienten führen. Für ein Krankenhaus als Merkmalsträger besteht das

⁴ Vgl. Vorgrimler, Daniel/Zwick, Markus (2004): Faktische Anonymisierung der Steuerstatistik (FAST), Wiesbaden 2004.

⁵ Vgl. Vorgrimler, Daniel/Zwick, Markus (2004): Faktische Anonymisierung der Steuerstatistik (FAST), Wiesbaden 2004.

Risiko, dass Informationen über ihre wirtschaftliche Situation bekannt werden, die seitens der Krankenkassen, Konkurrenten, Planungsbehörden usw. verwendet und zu negativen Konsequenzen führen können. Grundsätzlich wird der Schaden, den eine Deanonymisierung für den Merkmalsträger herbeiführen kann, für die Krankenhäuser als höher angesehen als für die Patienten.

Aufgrund dieser Deanonymisierungsrisiken für Krankenhäuser und Patienten kommt der Auswahl der Variablen und ihrer Ausprägungen eine besondere Bedeutung zu.

2.1 Auswahl der Variablen

Das Ausgangsmaterial der Diagnosestatistik enthält sowohl patientenbezogene als auch einrichtungsbezogene Daten. Um das Risiko der Deanonymisierung einzuschränken, werden, wie bereits erwähnt, in einem ersten Schritt zwei Datensätze aus dem Ausgangsmaterial erstellt. Der erste enthält vorrangig patientenbezogene, der zweite krankenhausbezogene Daten.

Auf Grundlage der im Ausgangsmaterial vorhandenen Merkmale hat der wissenschaftliche Beraterkreis prioritäre Variablen für das Scientific-Use-File festgelegt. Für die Wissenschaftler sind das Alter der Patienten sowie deren Verweildauer nach Hauptdiagnose (vierstelliger ICD-Schlüssel) von zentraler Bedeutung. Daher sind diese drei Zielvariablen in beiden Datensätzen enthalten.

Im patientenbezogenen Datensatz (DS I) werden sie ergänzt um Geschlecht und Wohnort des Patienten. Im krankenhausbezogenen Datensatz (DS II) werden die Zielgrößen erweitert um Behandlungsgebiet, Fachabteilung mit der längsten Verweildauer sowie Trägerschaft und Bettengrößenklasse des Krankenhauses.

Merkmale, die nicht in den beiden Scientific-Use-File berücksichtigt sind, können grundsätzlich über andere Zugangswege wie eine On-Site-Nutzung am Gastwissenschaftler-Arbeitsplatz oder über eine kontrollierte Datenfernverarbeitung in den Forschungsdatenzentren (FDZ) genutzt werden.

2.2 Vergrößern der Ausprägungen

Für etliche Merkmale ist eine Vergrößerung, d.h. eine Gruppierung von Merkmalsausprägungen, gegenüber dem Ausgangsmaterial erforderlich, um die faktische Anonymität zu wahren.

So wird die Verweildauer des Patienten bis zum 42. Tag in ganzen Tagen ausgewiesen. Aufgrund der geringer werdenden Fallzahlen bei hohen Verweildauern erfolgt jedoch ab dem 43. Tag eine Klassierung, wobei die Klassenbreite mit steigender Verweildauer zunimmt.

Eine Vergrößerung durch Klassierung erfolgt auch beim Alter. Hier werden 5-Jahres-Altersgruppen gebildet. Der Wohnort des Patienten wird nur bis auf die Kreisebene abgebildet.

Was für die Auswahl der Variablen gilt, trifft auch auf die Vergrößerung zu: die im Rahmen der Anonymisierung des Einzeldatenmaterials notwendigen Einschränkungen können durch andere Nutzungswege (z. B. On-Site-Nutzung) aufgehoben werden.

2.3 Verschieben der Ausprägungsklassen

Grundsätzlich besteht die Möglichkeit, dass der patientenbezogene und der krankenhausbezogene Datensatz über Variablen, die in beiden Datensätzen vorkommen (z. B. Verweildauer), miteinander verknüpft werden können. Dies erhöht das Risiko der Deanonymisierung einzelner Merkmalsträger (Patient bzw. Krankenhaus). Um diese Gefahr einzudämmen, wurden bei den Variablen, die in beiden Datensätzen vorkommen, die Ausprägungsklassen gegeneinander verschoben. So wurde im patientenbezogenen Datensatz die Altersklassierung gegenüber dem krankenhausbezogenen Datensatz um zwei Jahre verschoben (vgl. Datensatzbeschreibungen). Das Gleiche gilt für die Verweildauer ab dem 43. Tag.

2.4 Entfernen von Beobachtungen

Sofern bei den Merkmalen Hauptdiagnose, Verweildauer, Alter oder Geschlecht die Merkmalsausprägung unbekannt ist, wird die Beobachtung komplett gelöscht. Da die drei Merkmale Hauptdiagnose, Verweildauer und Alter in beiden Datensätzen vorkommen, wäre über die Ausprägung „unbekannt“ grundsätzlich eine Zuordnung und damit eine Verknüpfung beider Datensätze möglich. Dies wird zudem durch die geringe Anzahl unbekannter Ausprägungen begünstigt.

Die Entfernung der Beobachtungen mit unbekannter Ausprägung führt dazu, dass die Ergebnisse von jenen der Publikationen der statistischen Ämter abweichen.

Tabelle 1: Unbekannte Ausprägungen 2003

Merkmalsausprägung	darunter „unbekannt“
Hauptdiagnose	2 692 Fälle
Verweildauer	14 Fälle
Alter	61 Fälle
Geschlecht	106 Fälle
Insgesamt (einschl. Überschneidungen)	2 859 Fälle

Darüber hinaus werden 62 Fälle gelöscht, deren Kombination Wohnkreis-Alter-Geschlecht eine Deanonymisierung der Hauptdiagnose, des Diagnosekapitels oder der Verweildauer einzelner Patienten erlauben würde.

2.5 Geheimhaltung von Ausprägungen

Im Merkmal Fachabteilung mit der längsten Verweildauer werden neben der Gruppierung der Fachabteilungsschlüssel bestimmte Fälle geheim gehalten. Die resultierende Merkmalsausprägung umfasst sowohl „Fachabteilung geheim gehalten“ als auch „Fachabteilung unbekannt“.

Durch die Geheimhaltung wird die Zuordnung von Fällen zu bestimmten Krankenhäusern anhand selten auftretender Fachabteilungen sehr erschwert und durch die Zusammenfassung von „geheim gehalten“ und „unbekannt“ die Unsicherheit einer eventuellen Zuordnung erhöht.

3 Datenangriffe

Um die aus dem Ausgangsmaterial erstellten Teildatensätze hinsichtlich der Gefahr einer Deanonymisierung zu überprüfen, sind Datenangriffe notwendig. Einerseits besteht die Gefahr, dass durch die Kombination von Merkmalen innerhalb eines Datensatzes einzelne Merkmalsträger identifiziert werden können. Daneben kann durch das Zusammenführen beider Datensätze der Nutzen einer anderweitig erfolgreichen Deanonymisierung erhöht werden. Schließlich kann eine Identifikation durch das Zuspänsuchen von Zusatzinformationen erreicht werden. Die Datenangriffe konzentrierten sich auf diese drei Szenarien.

3.1 Datensatzinterne Angriffe

Für einen Patienten besteht die Gefahr, dass anhand allgemein zugänglicher bzw. beobachtbarer Informationen (z. B. Geschlecht, Alter, Wohnort) er selbst und seine Hauptdiagnose identifiziert werden können. Um diese Gefahr zu unterbinden, ist z. B. zu prüfen, ob es innerhalb einer Altersgruppe nur einen Patienten gibt bzw. ob alle Patienten einer Altersgruppe mit derselben Hauptdiagnose (in derselben Diagnosegruppe, demselben Diagnosekapitel) behandelt wurden. Die Wahrscheinlichkeit steigt, je tiefer die regionale Gliederung erfolgt. Daher wurde für Wohnkreis, Altersgruppe und Geschlecht aus Datensatz I zusätzlich geprüft, ob es Fälle mit Merkmalskombinationen gibt, deren Hauptdiagnosen ausschließlich in ein Diagnosekapitel fallen und dieses dadurch eindeutig identifiziert werden kann. Gemäß Absatz 2.4 wurden alle Fälle gelöscht, die eine solche Identifikation erlauben.

3.2 Datensatzübergreifende Angriffe

Es wurde geprüft, ob der patientenbezogene und krankenhausbezogene Datensatz anhand von Merkmalen zusammengeführt werden können, die in beiden Datensätzen vorkommen. Die Variablen Hauptdiagnose, Alter und Verweildauer sind in beiden Datensätzen enthalten. Durch die Verschiebung der Alters- und Verweildauerklassen wird eine Wiederausführung beider Datensätze wirksam verhindert. Allerdings wird die Verweildauer unter 42 Tagen in Tagen genau ausgewiesen. Daher wurden folgende Fragestellungen untersucht:

- Gibt es Einzelfälle bei der Auswertung der Hauptdiagnose (vierstelliger ICD-Schlüssel), durch die beide Datensätze miteinander verbunden werden können und die dadurch sensible Informationen über den Einzelfall offenbaren?
- Gibt es Einzelfälle bei einer kombinierten Auswertung von Hauptdiagnose (vierstelliger ICD-Schlüssel) und Verweildauer (unter 42 Tage), anhand derer beide Datensätze miteinander verbunden werden können und die dadurch sensible Informationen über den Einzelfall offenbaren?

Für weniger als 0,5 % aller Fälle ist eine Zuordnung zwischen patientenbezogenem und krankenhausbezogenem Datensatz mittels Einzelfällen oder Wertekombinationen mit nur je zwei Fällen möglich, da beide Teildatensätze unterschiedlich sortiert sind.

3.3 Angriffe mit Zusatzinformationen

Durch Abgleich mit anderweitig verfügbaren Daten lassen sich Einzelfälle zu bestimmten Merkmalsträgern (Patienten, Krankenhäuser) mit herausstehenden Merkmalen identifizieren. Zusatzinformationen sind beispielsweise im Krankenhausverzeichnis oder in Melderegistern enthalten. Datenangriffe in diesem Bereich erfolgten gemäß folgenden Fragestellungen:

- Ist ein Patient zu identifizieren, wenn sein Wohnkreis, seine Altersgruppe und sein Geschlecht aus Datensatz I mit dem Melderegister kombiniert werden?
- Ist ein Krankenhaus zu identifizieren, wenn seine Bettengrößenklasse, seine Trägerart und sein Behandlungsgebiet aus Datensatz II miteinander kombiniert werden?
- Ist die Fachabteilung eines Krankenhauses zu identifizieren, wenn dessen Bettengrößenklasse, Trägerart und Behandlungsgebiet mit der Fachabteilung mit der längsten Verweildauer aus Datensatz II kombiniert werden?

Die beschriebene Deanonymisierung von Patienten wird verhindert, da entsprechende Fälle gemäß Absatz 2.4 gelöscht wurden. Einzelne Krankenhäuser oder Fachabteilungen sind auf Grund der geeignet gewählten Vergrößerung (siehe Absatz 2.2) sowie Geheimhaltung (siehe Absatz 2.5) von Merkmalsausprägungen nicht reidentifizierbar.

3.4 Umgang mit deanonymisierten Merkmalsträgern

Sofern Merkmalsträger eindeutig identifiziert werden können, wird die komplette Beobachtung aus dem Scientific-Use-File gelöscht (siehe Absatz 2.4). Dies erschwert oder verhindert eine Deanonymisierung und mithin die Aufdeckung sensibler Informationen einzelner Merkmalsträger.

4 Fazit

Mit der vorgelegten Anonymisierungskonzeption ist es gelungen, erstmalig faktisch anonymisierte Daten aus den Gesundheitsstatistiken zu erstellen, die der Wissenschaft zu Analyse Zwecken zur Verfügung gestellt werden können. Für Fragestellungen, die sich nicht mit dem Scientific-Use-File beantworten lassen, stehen grundsätzlich alternative Zugangswege zu Mikrodaten, die von den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder angeboten werden, zur Verfügung.

Anhang

Anhang 1: Ausgangsdatenmaterial – Diagnosestatistik einschließlich krankenhausbezogener Informationen

Nr.	Merkmal	Ausprägungen
1	Geschlecht	Männlich Weiblich Unbekannt
2	Geburtsmonat	
3	Geburtsjahr	
4	Zugangstag	
5	Zugangsmonat	
6	Zugangsjahr	
7	Abgangstag	
8	Abgangsmonat	
9	Abgangsjahr	
10	Sterbefall	Ja Nein
11	Hauptdiagnose nach ICD-10-Schlüssel	A000 - Z999
12	Operation im Zusammenhang mit der Hauptdiagnose	Ja Nein Unbekannt
13	Fachabteilung mit der längsten Verweildauer	120 Augenheilkunde 153 Gefäßchirurgie 163 Thoraxchirurgie 166 Unfallchirurgie 167 Viszeralchirurgie 169 N.u.Fg./Sonstige Chirurgie 193 Frauenheilkunde 196 Geburtshilfe 199 N.u.Fg./Sonstige Frauenheilkunde und Geburtshilfe 220 Hals-Nasen-Ohren-Heilkunde 250 Haut- und Geschlechtskrankheiten 263 Thoraxchirurgie 269 N.u.Fg./Sonstige und allg. Herzchirurgie 311 Angiologie 313 Endokrinologie 316 Gastroenterologie 319 Hämatologie und internistische Onkologie 323 Kardiologie 336 Klinische Geriatrie 329 Nephrologie 332 Pneumologie 333 Rheumatologie 339 N.u.Fg./Sonstige und allg. Innere Medizin 350 Kinderchirurgie 363 Kinderkardiologie 366 Neonatologie 369 N.u.Fg./Sonstige und allg. Kinderheilkunde 390 Kinder- und Jugendpsychiatrie, -psychotherapie 510 Mund-Kiefer-Gesichtschirurgie 540 Neurochirurgie 570 Neurologie 630 Nuklearmedizin 693 Rheumatologie (Orthopädie) 699 N.u.Fg./Sonstige und allg. Orthopädie 800 Plastische Chirurgie 820 Psychiatrie und Psychotherapie 830 Psychotherapeutische Medizin 870 Strahlentherapie 900 Urologie 930 Sonstige Fachbereiche/Allgemeinmedizin bzw. Krankenhaus ohne abgegrenzte Fachabteilung 999 Fachabteilung unbekannt

Nr.	Merkmal	Ausprägungen
14	Verweildauer in Tagen	
15	Alter in vollen Jahren	
16	Alter in Monaten	
17	Gemeinde des Patienten	
18	Kreis des Patienten	
19	Regierungsbezirk des Patienten	
20	Bundesland des Patienten	
21	Gemeinde des Krankenhauses	
22	Kreis des Krankenhauses	
23	Regierungsbezirk des Krankenhauses	
24	Bundesland des Krankenhauses	
25	Berichtskreis	Krankenhaus Vorsorge- oder Rehabilitationseinrichtung
26	Art der Einrichtung	Hochschulklinik Plankrankenhaus Krankenhaus mit Versorgungsvertrag Sonstiges Krankenhaus Bundeswehrkrankenhaus Vorsorge- oder Rehabilitationseinrichtung (m. Vertrag) Sonstige Vorsorge- oder Rehabilitationseinrichtung
27	Reine Tages-/Nachtambulanz	Ja Nein
28	Ganzjährig geöffnete Einrichtung	Ja Nein
29	Art des Trägers	Öffentlich Freigemeinnützig Privat
30	Zuordnung des Krankenhauses	Allgemeines Krankenhaus Krankenhaus mit aussch. psychiatrischen Betten Krankenhaus mit aussch. psych. und neurol. Betten Bundeswehrkrankenhaus Reine Tages- oder Nachtambulanz
31	Bettengrößenklasse des Krankenhauses	Unter 50 Betten 50 – 99 Betten 100 – 149 Betten 150 – 199 Betten 200 – 249 Betten 250 – 299 Betten 300 – 399 Betten 400 – 499 Betten 500 – 599 Betten 600 – 799 Betten 800 – 999 Betten 1000 und mehr Betten
32	Förderung des Krankenhauses	Gefördert Teilweise gefördert Nicht gefördert
33	Zahl der Fachabteilungen	1 Fachabteilung 2 Fachabteilungen 3 Fachabteilungen 4 Fachabteilungen 5 und 6 Fachabteilungen 7 und mehr Fachabteilungen
34	Anteil der Belegbetten	Unter 25% 25 bis unter 100% 100%
35	Postleitzahl des Patientenwohnorts	

Anhang 2: Teildatensatz 1 – Datensatzbeschreibung des patientenbezogenen Datensatzes (DS I)

Merkmal	Ausprägungen	Codierung
Hauptdiagnose	Vierstelliger ICD-10-Schlüssel	A000 – Z999
Verweildauer	0 Tage	0
	1 Tag bis 42 Tage spitz	1 .. 42
	43 bis 49 Tage	50
	50 bis 56 Tage	51
	57 bis 63 Tage	52
	64 bis 70 Tage	53
	71 bis 84 Tage	54
	85 bis 98 Tage	55
	99 bis 140 Tage	56
	141 bis 182 Tage	57
	mehr als 182 Tage	58
Altersgruppe	0 bis unter 5 Jahre	01
	5 bis unter 10 Jahre	02
	10 bis unter 15 Jahre	03
	15 bis unter 20 Jahre	04
	20 bis unter 25 Jahre	05
	25 bis unter 30 Jahre	06
	30 bis unter 35 Jahre	07
	35 bis unter 40 Jahre	08
	40 bis unter 45 Jahre	09
	45 bis unter 50 Jahre	10
	50 bis unter 55 Jahre	11
	55 bis unter 60 Jahre	12
	60 bis unter 65 Jahre	13
	65 bis unter 70 Jahre	14
	70 bis unter 75 Jahre	15
	75 bis unter 80 Jahre	16
	80 bis unter 85 Jahre	17
85 bis unter 90 Jahre	18	
90 bis unter 95 Jahre	19	
	95 Jahre und älter	20
Geschlecht	Männlich	1
	Weiblich	2
Wohnort des Patienten	Erste fünf Stellen des amtlichen Gemeindegeschlüssels (Kreisschlüssel)	
	Ausland	50000
	Unbekannt	99999

Anhang 3: Teildatensatz II – Datensatzbeschreibung des krankenhausesbezogenen Datensatzes (DS II)

Merkmal	Ausprägungen	Codierung
Hauptdiagnose	Vierstelliger ICD-10-Schlüssel	A000 – Z999
Verweildauer	0 Tage	0
	1 Tag bis 42 Tage spitz	1 .. 42
	43 bis 46 Tage	70
	47 bis 53 Tage	71
	54 bis 60 Tage	72
	61 bis 67 Tage	73
	68 bis 77 Tage	74
	78 bis 91 Tage	75
	92 bis 119 Tage	76
	120 bis 161 Tage	77
	mehr als 161 Tage	78
Altersgruppe	0 bis unter 3 Jahre	51
	3 bis unter 8 Jahre	52
	8 bis unter 13 Jahre	53
	13 bis unter 18 Jahre	54
	18 bis unter 23 Jahre	55
	23 bis unter 28 Jahre	56
	28 bis unter 33 Jahre	57
	33 bis unter 38 Jahre	58
	38 bis unter 43 Jahre	59
	43 bis unter 48 Jahre	60
	48 bis unter 53 Jahre	61
	53 bis unter 58 Jahre	62
	58 bis unter 63 Jahre	63
	63 bis unter 68 Jahre	64
	68 bis unter 73 Jahre	65
	73 bis unter 78 Jahre	66
78 bis unter 83 Jahre	67	
83 bis unter 88 Jahre	68	
88 bis unter 93 Jahre	69	
	93 Jahre und älter	70
Behandlungsgebiet	Nord (SH, HH, HB, NI, NW)	1
	Süd (BW, BY, HE, RP, SL)	2
	Ost (BE, BB, MV, SN, ST, TH)	3
Fachabteilung mit der längsten Verweildauer	Chirurgie (153-169)	01
	Frauenheilkunde und Geburtshilfe (193-199)	02
	Hals-Nasen-Ohren-Heilkunde (220)	03
	Innere Medizin (311-339)	04
	Kinderheilkunde (363-369)	05
	Orthopädie (693-699)	06
	Urologie (900)	07
	Psychiatrie (390, 820, 830)	08
	Sonstige (120, 250, 263-269, 350, 510, 540, 570, 630, 800, 870, 930)	09
Fachabteilung geheim gehalten oder unbekannt	99	
Trägerschaft	öffentlich	1
	freigemeinnützig	2
	privat	3
Bettengrößenklasse	unter 100 Betten (ohne öffentliche Krankenhäuser in Ost)	1
	100 bis unter 200 Betten (ohne öffentliche Krankenhäuser in Ost)	2
	200 bis unter 500 Betten (ohne private Krankenhäuser)	3
	500 und mehr Betten (ohne private Krankenhäuser)	4
	unter 200 Betten (für öffentliche Krankenhäuser in Ost)	5
	200 und mehr Betten (für private Krankenhäuser)	6