

PD Dr. Jürgen Stausberg/Dr. Nils Lehmann/Dr. Dirk Kaczmarek/  
Markus Stein

# Einheitliches Kodieren in Deutschland: Wunsch und Wirklichkeit

**Einheitliches Kodieren mit den gesetzlich vorgeschriebenen Klassifikationen ICD-10-GM und OPS ist die Voraussetzung für fall-, einrichtungs- und sektorübergreifende Auswertungen zur Ökonomie, zur Behandlungsqualität und zur Systemgestaltung. Die Übereinstimmung von Medizincontrollern, Dokumentationsfachkräften und Medizinstudenten beim Kodieren von Diagnosen wurde daher anhand von Arztbriefen im Rahmen einer Studie untersucht. Deren unbefriedigendes Ergebnis zeigt die Unsicherheit, die eine Kodierung von Diagnosen mit der ICD-10-GM unter Berücksichtigung der Deutschen Kodierrichtlinien mit sich bringt. Dem Verfahren des Medizinischen Dienstes der Krankenversicherung (MDK) fehlt somit eine gesicherte Verfahrengrundlage. Sowohl die Klassifikationen als auch die Kodierrichtlinien müssen deutlich vereinfacht werden, um eine einheitliche Kodierung zu gewährleisten.**

Diagnosen und Prozeduren, kodiert mit den gesetzlich vorgeschriebenen Klassifikationen ICD-10-GM und OPS, sind die wesentlichen Merkmale der Basisdokumentation von Krankenhäusern in Deutschland. Die von Ärzten oder Dokumentationsfachkräften ermittelten Schlüssel (= Codes) werden unter anderem an die Krankenkassen zur Abrechnung übermittelt (§ 301 SGB V), dienen der Ermittlung von G-DRGs und Zusatzentgelten, ermöglichen dem InEK die Kalkulation und Definition der G-DRGs (§ 21 KHEntgG) und weisen im Strukturierten Qualitätsbericht nach § 137 SGB V Anzahl und Art der erbrachten Leistungen aus. Aussagen zur Ökonomie, zur Qualität und zur Systemgestaltung im Gesundheitswesen basieren in hohem Maße auf diesen Angaben. Es verwundert daher nicht, dass die Basisdokumentation in das Spannungsfeld zwischen Leistungserbringer und Kostenträger geraten ist. Im Einzelfall ist es der MDK, der nach § 275 SGB V bei Auffälligkeiten eine Prüfung der ordnungsgemäßen Abrechnung durchführt. Außerhalb von auffälligen Einzelfällen erlaubt § 17 c KHG die Prüfung einer ordnungsgemäßen Abrechnung über Stichproben, ebenfalls durch den MDK. Im Rahmen dieser Prüfungen findet ein Vergleich der an die Krankenkassen übermittelten Angaben mit einer Nachkodierung durch den Medizinischen Dienst statt. Dem MDK fällt dabei die Rolle zu, die „wahren“ Codes festzulegen, an denen die Basisdokumentation des Krankenhauses gemessen wird.

Die beschriebene Nutzung der Dokumentation setzt voraus, dass derselbe stationäre Behandlungsfall immer gleich kodiert wird, unabhängig vom Krankenhaus, von der kodierenden Berufsgruppe und vom kodierenden Mitarbei-

ter. Eine einheitliche Kodierung stellt sicher, dass Abrechnung, Qualitätsvergleiche und Risikobewertungen auf tatsächlichen Unterschieden basieren und nicht nur auf der zufälligen Auswahl unter mehreren möglichen Kodieralternativen. Dazu müssen die Werkzeuge zur Kodierung, die Klassifikationen und deren Anwendungsregeln, also die Deutschen Kodierrichtlinien, verständlich, anwendbar und eindeutig sein. Nur dann ist die Festlegung einer „wahren“ Kodierung für den Einzelfall begründbar. Trotz der hohen Bedeutung und der intensiven Nutzung kodierter Diagnosen und Prozeduren ist nicht bekannt, wie hoch die Übereinstimmung bei der Kodierung desselben Sachverhalts in die ICD-10-GM und den OPS ist.

Die Übereinstimmung bei mehrfacher Messung des gleichen Sachverhalts wird als Zuverlässigkeit oder Reliabilität bezeichnet, die Übereinstimmung einer Messung mit einem Goldstandard (dem „wahren“ Wert) als Richtigkeit oder Validität. Wie hängen nun Reliabilität und Validität zusammen? Ein Beispiel aus der Praxis zeigt Tabelle 1.

In Variante 1 liegt die Reliabilität bei 0 Prozent, in Variante 2 bei 100 Prozent. Zur Bestimmung des „wahren“ Codes dient der Kodierleitfaden der Fachgesellschaft.<sup>1)</sup> Danach waren in beiden Varianten alle Ergebnisse falsch. Die Validität liegt also bei 0 Prozent. Es entspannt sich eine intensive Diskussion zwischen den 3 beteiligten Personen. Hierbei werden der Diagnosenthesaurus, die Systematik der ICD-10-GM, verschiedene Kodiertools und Lehrbücher herangezogen. Pathophysiologische Überlegungen zur Entstehung eines Infarkts bei Stenose präzerebraler Arterien werden diskutiert. Eine konsensuelle Entscheidung zwischen Embolie, Thrombose und unspezifischer Angabe gelingt nicht. Alle Codes erscheinen begründbar und könnten als „wahres“ Ergebnis gelten. Damit würde die

**Tabelle 1: Beispiel zum Vergleich von Reliabilität und Validität der Hauptdiagnose**

Fall <sup>1)</sup>	Pat. (67 J.) wird mit akutem Mediateilinfarkt rechts eingeliefert. Dopplersonographie: Ipsilaterale (symptomatische) ACI-Stenose, kontralaterale (asymptomatische 70%-ige) ACI-Stenose.
Variante 1	Person A: I63.2 „Hirnfarkt durch nicht näher bezeichneten Verschluss oder Stenose präzerebraler Arterien“
	Person B: I63.1 „Hirnfarkt durch Embolie präzerebraler Arterien“
Variante 2	Person A: I63.2 „Hirnfarkt durch nicht näher bezeichneten Verschluss oder Stenose präzerebraler Arterien“
	Person C: I63.2 „Hirnfarkt durch nicht näher bezeichneten Verschluss oder Stenose präzerebraler Arterien“
„Wahrer“ Wert <sup>1)</sup>	I63.0 „Hirnfarkt durch Thrombose präzerebraler Arterien“

Validität bei 100 Prozent liegen, unabhängig von der Reliabilität. Im Allgemeinen kann allerdings angenommen werden, dass eine niedrige Reliabilität auch zu einer niedrigen Validität führt.

Bezogen auf Prüfungen der Basisdokumentation durch den MDK stellt sich die Frage, ob angemahnte Unterschiede durch Kodierfehler oder durch eine niedrige Reliabilität aufgrund von Mängeln der Klassifikationen und der Klassifikationsregeln bedingt sind. In einer Untersuchung zur Reliabilität der Kodierung wurde daher die Übereinstimmung bei der Kodierung von Diagnosen aus Arztbriefen in die ICD-10 in 3 Gruppen gemessen:

- Medizinstudenten am Universitätsklinikum Essen im Wintersemester 2003/2004,
- Dokumentationsfachkräfte des Klinikums Ludwigshafen,
- Ärzte im Medizin-Controlling an deutschen Krankenhäusern.

## Studienteilnehmer

Grundlage der Untersuchung waren anonymisierte Arztbriefe einer Inneren Abteilung mit dem Schwerpunkt Nephrologie aus den frühen 90er Jahren. Die Studenten erhielten die Unterlagen zur Kodierung in Papierform. Die zurückgegebenen Bögen wurden zur weiteren Auswertung eingegeben. Dokumentationsfachkräfte und Medizincontroller erhielten Arztbriefe als Dateien im Portable Document Format. Die Erfassung der Codes erfolgte direkt in eine Tabelle von Microsoft® Excel.

### ■ Studenten der Medizin

Nach einer Stunde Kodierübungen im Querschnittsbereich Epidemiologie, medizinische Biometrie und medizinische Informatik erhielt jeder Student im Wintersemester 2003/2004 eine von 15 Diagnosenlisten aus den Arztbriefen zur Kodierung in die ICD-10-SGB-V 2.0 als Hausaufgabe. Eine Unterscheidung in Haupt- und Nebendiagnosen wurde nicht vorgenommen, auch blieben die Deutschen Kodierrichtlinien unberücksichtigt. Als Hilfsmittel zur Kodierung wurden die Studenten auf die WWW-Seite des DIMDI hingewiesen (<http://www.dimdi.de>). Von 118 Studenten liegen Bögen zu 15 Diagnosenlisten vor. Die Anzahl der Bögen je Arztbrief liegt zwischen 6 und 11.

### ■ Medizincontroller

In einer Mailingliste der Arbeitsgruppe Medizin-Controlling der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) e. V. wurde um Teilnahme geworben. 34 interessierte Ärzte im Medizincontrolling erhielten je 5 zufällig ausgewählte Arztbriefe von insgesamt 34 zur Kodierung. Die Arztbriefe wurden so ausgewählt, dass jeder Brief an 5 verschiedene Medizincontroller verschickt wurde. Kodiert wurden Haupt- und Nebendiagnosen unter Berücksichtigung der Deutschen Kodierrichtlinien in die ICD-10-GM 2004. 27 Medizincontroller reichten Ergebnisse ein, sodass insgesamt 135 Bögen von 170 möglichen berücksichtigt werden konnten. Von jedem der 34 Arztbriefe liegen mindestens 2 Bögen vor.

### ■ Dokumentationsfachkräfte

Das Klinikum Ludwigshafen setzt Dokumentationsfachkräfte zur DRG-Abrechnung anhand von Patientenakten ein.<sup>2)</sup> 13 Personen beteiligten sich an dieser Untersuchung. Kodiert wurden Haupt- und Nebendiagnosen unter Berücksichtigung der Deutschen Kodierrichtlinien in die ICD-10-GM 2004. Aus den 34 Arztbriefen der Gruppe Medizincontroller wurden 12 zufällig ausgewählt und an alle 13 Dokumentationsfachkräfte ausgeteilt. Alle Arztbriefe wurden komplett kodiert, sodass insgesamt 156 Bögen vorliegen.

### ■ Überlappung der 3 Gruppen

6 Arztbriefe wurden von allen 3 Gruppen kodiert, alle Arztbriefe der Dokumentationsfachkräfte auch von den Medizincontrollern. Eine Analyse dieser Teilmengen sollte die Verzerrung durch gruppenspezifische Unterschiede der Arztbriefe ausschließen. Im Regelfall werden für eine präzisere Schätzung der Reliabilität innerhalb jeder Gruppe die Ergebnisse der kompletten Untersuchung angegeben. Unterschiede in der Kodierung eines einzelnen Arztbriefs zwischen Teilnehmern aus den verschiedenen Gruppen werden nicht betrachtet.

## Methoden

Die Codes wurden in einer Microsoft® Access 2003-Datenbank zusammengeführt, Analysen mit Microsoft® Excel 2003, Access und dem Statistical Analysis System® (SAS) in Version 8.2. vorgenommen. Vor der Bestimmung der Reliabilitätsmaße wurden unplausible Daten gelöscht oder korrigiert. So wurden erkennbare Rechtschreibfehler (Leertaste im Code) korrigiert, angehängte Stellen gestrichen sowie nicht endständige Codes um die Ziffern für „nicht näher bezeichnet“ erweitert. Doppelte Codes auf einem Bogen wurden zusammengeführt, ggf. wurde der als Nebendiagnose aufgeführte Code gestrichen. Die optionale Seitenlokalisierung blieb, falls angegeben, unberücksichtigt. Zusatzcodes (Stern, Ausrufezeichen) wurden wie reguläre Nebendiagnosen behandelt. Als Bogen wird das komplette Kodierergebnis eines Teilnehmers zu einem Arztbrief bezeichnet. „Diagnose“ und „Code“ werden synonym benutzt.

Als anerkanntes Maß der Reliabilität wurde der Kappa-Wert für die Übereinstimmung bei der Hauptdiagnose bestimmt.<sup>3)</sup> Der Kappa-Wert berücksichtigt die höhere erwartete Übereinstimmung bei häufig verwendeten Kategorien und kann sich daher von der absoluten Übereinstimmung deutlich unterscheiden. Kappa nimmt Werte zwischen -1 und 1 an. 1 zeigt eine vollständige Übereinstimmung, 0 eine zufällige Verteilung und Werte unter 0 eine noch geringere Übereinstimmung an. Für den Kappa-Wert liegen qualitative Interpretationen vor (Tabelle 2).

Für den Vergleich von Bögen mit mehreren Diagnosen existiert kein anerkanntes Verfahren. Hierzu wurde daher von den Autoren Stausberg und Lehmann ein einfaches Reliabilitätsmaß  $p_{om}$  entwickelt.<sup>5)</sup> Dabei wird die Anzahl von Übereinstimmungen zwischen 2 Bögen eines Arztbriefs geteilt durch die Gesamtzahl unterschiedlicher Codes auf diesen Bögen. Das Ergebnis wird zuerst für jeden Arztbrief

**Tabelle 2: Einteilung der Kappa-Werte nach Landis und Koch<sup>1)</sup>**

Kappa-Wert	Bewertung
<0,00	schlecht (poor)
0,00-0,20	gering (slight)
0,21-0,40	ausreichend (fair)
0,41-0,60	moderat (moderate)
0,61-0,80	substantiell (substantial)
0,81-1,00	fast perfekt (almost perfect)

über die Paare von Bögen und anschließend über alle Arztbriefe gemittelt. Eine 0 zeigt hier fehlende Übereinstimmung an, eine 1 völlige Übereinstimmung.

## Ergebnisse

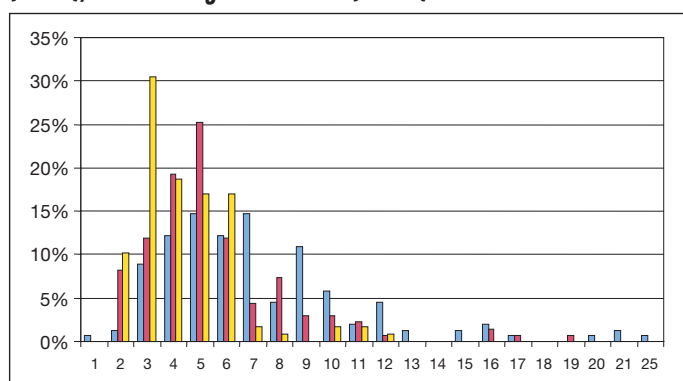
### ■ Anzahl von Diagnosen

Die Medizinstudenten ermittelten 516 Codes bei 118 Bögen. Die mittlere Anzahl von Diagnosen pro Bogen liegt bei 4,4. Häufigste Diagnose ist mit 38 Nennungen I10 „Essentielle (primäre) Hypertonie“ (32,2 Prozent der Bögen). Insgesamt wurden 118 unterschiedliche Codes verwendet. Bei 135 Bögen der Medizincontroller wurden mit 751 Codes im Mittel 5,6 je Bogen verwendet. Häufigste Diagnose ist mit 23 Nennungen E66.0 „Adipositas durch übermäßige Kalorienzufuhr“ (17 Prozent der Bögen). 312 unterschiedliche Codes traten auf. Die Dokumentationsfachkräfte nutzen 1 137 Codes bei 156 Bögen, also im Mittel 7,3 Codes je Bogen. Häufigste Diagnose war mit 51 Nennungen I10.90 „Benigne essentielle Hypertonie, ohne Angabe einer hypertensiven Krise“ (32,7 Prozent der Bögen). Es wurden 247 unterschiedliche Codes verwendet. Abbildung 1 zeigt die Anzahl der Codes je Bogen für alle 3 Gruppen.

### ■ Reliabilität bei der Kodierung der Hauptdiagnose

In den Gruppen Dokumentationsfachkräfte und Medizincontroller war auf jedem Bogen ein Code als Hauptdiagnose markiert. Häufigster Code bei den Dokumentationsfachkräften war N18.89 „Chronische nicht-terminale Niereninsuffizienz, Stadium nicht näher bezeichnet“ mit 23 Nennungen bei insgesamt 156 Bögen, bei den Medizincontrollern N18.0 „Terminale Niereninsuffizienz“ mit 7 Nennungen bei 135 Bögen. Tabelle 3 zeigt die Anzahl übereinstimmender

**Abbildung 1: Relative Häufigkeit der Bögen innerhalb einer Gruppe, kategorisiert nach der Anzahl Diagnosen pro Bogen. Dokumentationsfachkräfte blauer Balken (links), Medizincontroller roter Balken (Mitte), Studenten gelber Balken (rechts).**



**Tabelle 3: Reliabilität bei der Kodierung der Hauptdiagnose auf den Aggregationsebenen der ICD. Zugrunde liegen bei den Medizincontrollern 212 und bei den Dokumentationsfachkräften 936 Paare von 2 unterschiedlichen Bögen zum selben Arztbrief.**

	Medizincontroller				Dokumentationsfachkräfte			
	übereinst. Paare		Kappa		übereinst. Paare		Kappa	
	N	%	Mittelwert	95 %-Konfidenzintervall	N	%	Mittelwert	95 %-Konfidenzintervall
Klasse	62	29,2	0,27	0,22-0,32	438	46,8	0,42	0,39-0,46
Dreisteller	128	60,4	0,56	0,50-0,62	641	68,5	0,63	0,59-0,66
Gruppe	145	68,4	0,64	0,58-0,70	707	75,5	0,71	0,68-0,74
Kapitel	166	78,3	0,71	0,65-0,77	756	80,8	0,72	0,68-0,75

Kodierungen sowie den mittleren Kappa-Wert in beiden Gruppen. Auf der Ebene des endständigen Codes (= Klasse) ist das Ergebnis bei den Medizincontrollern als ausreichend und bei den Dokumentationsfachkräften als moderat zu bewerten, auf der Ebene des Kapitels erreichen beide Gruppen einen substantiellen Wert. Bei den Medizincontrollern fällt eine deutliche Verschlechterung bei der Verfeinerung vom Dreisteller zum endständigen Code auf. Bei den Dokumentationsfachkräften überrascht der geringe Abstand zwischen der Gruppe und dem Kapitel der ICD. Einigkeit über das Kapitel führt auch zur Einigkeit über die Gruppe. Insgesamt erzielen die Dokumentationsfachkräfte eine höhere Zuverlässigkeit.

### ■ Reliabilität über alle Diagnosen eines Arztbriefs

Wie Kappa für die Hauptdiagnose wurde  $p_{om}$  für die Aggregationsebenen der ICD in jeder Gruppe getrennt berechnet (Tabelle 4). Dabei wurden doppelte Codes auf den höheren Ebenen für jeden Bogen zusammengeführt. Bei den Konfidenzintervallen zeigen sich große Überlappungen zwischen den Medizincontrollern und den Dokumentationsfachkräften. Ein leichter Vorteil der Dokumentationsfachkräfte auf der Ebene des endständigen Codes kehrt sich auf den höheren Aggregationsebenen der ICD um. Überraschend zeigen sich für die Studenten auf allen Ebenen deutlich höhere Werte im Vergleich zu den beiden anderen Gruppen.

### ■ Überlappung der 3 Gruppen

Die Berechnung des Kappa-Werts der Medizincontroller für die 12 Arztbriefe, die auch von den Dokumentationsfachkräften kodiert wurden, führt zu niedrigeren Ergebnissen auf der Ebene der Gruppe (0,60, 95 Prozent-Konfidenzintervall

**Tabelle 4: Reliabilität bei der Kodierung aller Diagnosen eines Arztbriefs auf den Aggregationsebenen der ICD**

	Medizincontroller		Dokumentationsfachkräfte		Studenten	
	$p_{om}$	95 %-Konfidenzintervall	$p_{om}$	95 %-Konfidenzintervall	$p_{om}$	95 %-Konfidenzintervall
Klasse	0,21	0,17-0,24	0,28	0,22-0,34	0,46	0,39-0,53
Dreisteller	0,40	0,35-0,46	0,39	0,32-0,47	0,58	0,52-0,64
Gruppe	0,50	0,45-0,55	0,46	0,38-0,54	0,73	0,67-0,79
Kapitel	0,64	0,59-0,70	0,60	0,54-0,65	0,87	0,80-0,94

tervall 0,45-0,74) und der Kapitel (0,64, 95 Prozent-Konfidenzintervall 0,46-0,81).  $p_{om}$  ändert sich hingegen kaum. Eine Berechnung von  $p_{om}$  aus den 6 übereinstimmenden Fällen in allen Gruppen führt tendenziell zu niedrigeren Werten der Studenten und der Dokumentationsfachkräfte, zu höheren Werten bei den Medizincontrollern.

## Diskussion

Wunsch bzw. Notwendigkeit einer einheitlichen Kodierung und gemessene Wirklichkeit klaffen deutlich auseinander. Die Zuverlässigkeit bei der Kodierung der Hauptdiagnose aus Arztbriefen durch Ärzte im Medizincontrolling oder durch Dokumentationsfachkräfte ist unbefriedigend. 5 Medizincontroller bzw. 7 Dokumentationsfachkräfte kommen bei einer unabhängigen Kodierung desselben Arztbriefs zu 3 verschiedenen Codes. Bei dem Vergleich aller Diagnosen, also der Haupt- und Nebendiagnosen, zeigt sich eine weitere Verschlechterung ( $p_{om}$  für die Hauptdiagnose hier nicht dargestellt). Die Interpretation der Definition einer Nebendiagnose scheint unsicherer zu sein als die der Hauptdiagnose (Tabelle 5).<sup>6)</sup>

Der Vergleich zwischen den Ergebnissen für Medizincontroller und Dokumentationsfachkräfte ergibt keinen eindeutigen Vorteil zugunsten einer Berufsgruppe. Für die Dokumentationsfachkräfte zeigt sich bei der Kodierung der Hauptdiagnose eine höhere Zuverlässigkeit. Ursächlich ist hier jedoch vor allem eine interne Standardisierung am Klinikum Ludwigshafen zu vermuten. Bei der auch von Roeder et al. gestellten Frage „Arzt oder Koder“<sup>7)</sup> kann die ermittelte Zuverlässigkeit der Kodierung keine Entscheidung für eine der beiden Berufsgruppen begründen. Überraschenderweise zeigt sich, dass Medizinstudenten eine durchweg höhere Übereinstimmung aufweisen.

Die Dokumentationsfachkräfte identifizierten im Mittel je Bogen mit 7,3 mehr Diagnosen als die Medizincontroller mit 5,6. Dies mag für die besondere Sorgfalt dieser Berufsgruppe sprechen. Hingegen differenzierten die Medizincontroller stärker als die Studenten und die Dokumentationsfachkräfte, da sie jeden Code nur 2,4-mal verwendeten, die beiden anderen Gruppen hingegen 4,4-mal bzw. 4,6-mal. Dies könnte in Verbindung mit dem deutlichen Zuverlässigkeitsverlust von der ICD-Gruppe auf den endständigen Code Ausdruck einer spezifisch ärztlichen Interpretation der im Arztbrief aufgelisteten Befunde sein.

### ■ Ursachen der niedrigen Reliabilität

Bei der Erklärung der vorgelegten Ergebnisse sind mehrere Aspekte zu diskutieren. Wesentlich ist jedoch die Unsicherheit, die der ICD-10-GM und die Deutschen Kodierrichtlinien bei der Kodierung von Diagnosen hervorgerufen.

**Tabelle 5: Definition der Haupt- und Nebendiagnose in den Deutschen Kodierrichtlinien<sup>6)</sup>**

<b>Nebendiagnose</b>	Eine Krankheit oder Beschwerde, die entweder gleichzeitig mit der Hauptdiagnose besteht oder sich während des Krankenhausaufenthalts entwickelt.
<b>Hauptdiagnose</b>	Die Diagnose, die nach der Analyse als diejenige festgestellt wurde, die hauptsächlich für die Veranlassung des stationären Krankenhausaufenthalts des Patienten verantwortlich ist.

### Upcoding

Ökonomische Gesichtspunkte können die Dokumentation beeinflussen. Hsia et al. unterscheiden dabei zwischen einer zulässigen Optimierung und einem unzulässigen „DRG creep“.<sup>8)</sup> Das Potenzial der Optimierung gemessen an der mittleren Bewertungsrelation (Casemix-Index) liegt vermutlich bei rund 15 Prozent, der „DRG creep“ deutlich geringer, zwischen 2 Prozent und 4 Prozent.<sup>8)9)10)</sup> Ein „DRG creep“ kann in der durchgeführten Untersuchung weitgehend ausgeschlossen werden, da es sich um eine Kodierung ohne Erlösbedeutung handelte.

### Fehlende Motivation

Der Aufwand für die Dokumentation ist hoch und in der täglichen Routine kaum zu leisten.<sup>11)</sup> Allerdings ist der Anteil der Basisdokumentation hieran recht gering. Dennoch wird dieses Thema immer wieder negativ besetzt.<sup>12)</sup> Eine nachlässige Kodierung im Rahmen der vorgestellten Studie lässt sich ebenfalls nicht völlig ausschließen. Zeit und Ort der Kodierung waren den Teilnehmern ebenso überlassen wie die dafür verwendeten Hilfsmittel. Hinweise auf mögliche Absprachen zwischen Teilnehmern bezüglich der Ergebnisse bestehen jedoch ebenso wenig wie Hinweise auf bewusstes Fehlkodieren.

### Unsicherheit bei der Diagnosestellung

Bei einer Kodierung auf der Basis von Befunden kann es zu unterschiedlichen Einschätzungen hinsichtlich der vorliegenden Erkrankung kommen.<sup>13)</sup> Dies wird sich dann auch in unterschiedlichen Entscheidungen über den richtigen Code widerspiegeln. Die Vorgabe eines Diagnostextes wie bei den Studenten mag im Vergleich zu einem kompletten Arztbrief unterschiedliche Einschätzungen reduzieren.

Die Unsicherheit der Kodierung wird zunehmen, je mehr Informationen über den Patienten vorliegen, weil damit auch der Interpretationsspielraum, insbesondere bezüglich der Zulässigkeit einer Erkrankung als Nebendiagnose, ansteigt. Die Kodierung auf Grundlage der Krankenakte wird daher die Zuverlässigkeit weiter verschlechtern: je detaillierter die Information, desto unsicherer die Kodierung.

### Klassifikationen

Es ist bekannt, dass die ICD fehlerhaft, mehrdeutig, missverständlich und widersprüchlich ist.<sup>14)</sup> Im Anpassungsprozess an die G-DRGs von der ICD-9 über die ICD-10-SGB V 1.3 bis zur ICD-10-GM 2005 kam es zu einer stetigen Verfeinerung von 5 841 bis zu 13 097 Klassen. Zusätzlich wurde die Systematik immer wieder aufgebrochen, so zum Beispiel durch die Einführung von unbelegten Schlüsselnummern, von Prozeduren und von Sonderbedingungen wie dem Ausrufungszeichen. Damit ist der Vorgang der Einordnung einer Krankheit eines Patienten in die richtige Klasse der ICD-10 sehr komplex geworden.

### Deutsche Kodierrichtlinien

Die spezifisch deutschen Regeln zur Kodierung waren mit 193 Deutschen Kodierrichtlinien im Jahr 2004 für den Einzelnen nicht zu überblicken.<sup>6)</sup> Die hier vorgelegten Ergebnisse bestätigen, dass die Kodierrichtlinien die Einheitlich-

keit der Kodierung keinesfalls gefördert, sondern eher behindert haben. Als Konsequenz ist daher mit den Deutschen Kodierrichtlinien für das Jahr 2005 ein erster Schritt in Richtung Vereinfachung erfolgt.<sup>15)</sup>

## ■ Prüfungen des MDK

Angesichts der Unsicherheiten bei der Kodierung muss es verwundern, dass in Veröffentlichungen das Ziel beschrieben wird, Abrechnungsdaten des Krankenhauses anhand einer „wahren“ Kodierung durch den MDK auf Richtigkeit zu überprüfen. Wie die vorgelegte Untersuchung belegt, sind Unterschiede bei unabhängiger Nachkodierung durch den MDK nicht nur im Einzelfall, sondern sogar regelhaft zu erwarten. Bislang sind den Autoren keine Untersuchungen des MDK zur Zuverlässigkeit der Kodierung durch deren Gutachter bekannt, obwohl durchaus die Notwendigkeit erkannt wurde, „Intra- und Interrater-Variabilität zu untersuchen“<sup>16)</sup>. Die seitens des MDK festgestellten Unterschiede sind dann auch beträchtlich, wenn auch nicht so ausgeprägt, wie es nach den hier vorgestellten Daten zu erwarten gewesen wäre. Hier mag sich das von Schnabel et al. diskutierte Ziel des MDK widerspiegeln, „offensichtlich in erster Linie einen Ansatz zur Kostenreduktion“ zu verfolgen.<sup>17)</sup> Der MDK führt eine interessengeleitete Nachkodierung unter Kenntnis der Basisdokumentation des Krankenhauses durch. Kodieralternativen ohne potenziellen ökonomischen Einfluss bleiben dann leicht unberücksichtigt. Eine unabhängige Zweitkodierung ist nicht gegeben.

Exakt unter diesen Rahmenbedingungen erfolgte im Jahr 2003 ein Vergleich von 2 011 Fällen des DRK-Krankenhauses Alzey mit einer Nachkodierung durch den MDK Rheinland-Pfalz.<sup>10)</sup> Für Hauptdiagnosen lag die Übereinstimmung insgesamt bei 70,4 Prozent, am schlechtesten in der Gynäkologie mit 59,9 Prozent und in der Inneren Medizin mit 60,8 Prozent. In der Inneren Medizin wurde rund ein Drittel aller Nebendiagnosen bestätigt, welches in etwa dem hier vorgestellten  $p_{om}$  entspricht. Vergleichbare Ergebnisse wurden aus dem Jahr 2002 für das Bürgerhospital Frankfurt am Main veröffentlicht.<sup>9)</sup> Die Übereinstimmung der Hauptdiagnose betrug für die Gynäkologie 64,8 Prozent, für die Innere Medizin 52 Prozent. Insgesamt wurden in der Inneren Medizin 37 Prozent der Diagnosen bestätigt.

### Zusammenfassung

Die Steuerung des deutschen Gesundheitswesens beruht in hohem Maße auf Analysen von Diagnosen und Prozeduren, die seitens der Leistungserbringer mit den gesetzlich vorgeschriebenen Klassifikationen ICD-10-GM und OPS kodiert wurden. Eine einheitliche Kodierung ist die Voraussetzung für fall-, einrichtungs- und sektorübergreifende Auswertungen zur Ökonomie, zur Behandlungsqualität und zur Systemgestaltung. Die Übereinstimmung von Medizincontrollern, Dokumentationsfachkräften und Medizinstudenten bei der Kodierung von Diagnosen aus Arztbriefen wurde daher in einer Studie untersucht. Hierzu wurden die gleichen Arztbriefe von mehreren Personen kodiert, die Übereinstimmung geprüft und Kenngrößen der Reliabilität (= Zuverlässigkeit) bestimmt. Sowohl unter Berücksichtigung aller Diagnosen als auch bei den Hauptdiagnosen ist das Ergebnis unbefriedigend. So stimmten in der Festlegung der Hauptdiagnose nur 30 Prozent der Ergebnisse bei Medizincontrollern und 47 Prozent der Ergebnisse bei Dokumentationsfach-



**FAA**  
FachArztAgentur  
GmbH  
Honorarvertretungen

## Die freundliche Alternative

Freiberufliche Ärzte  
Vermittlung · Organisation · Betreuung

Telefon: 05 21 / 8 97 31 16 · Fax: 05 21 / 8 97 31 17  
Postfach 101109 · 33511 Bielefeld  
www.facharztagentur.de · info@facharztagentur.de

Eine Studie des Institute of Medicine (IOM), die Daten von 1974 enthält,<sup>18)</sup> kommt zu ähnlichen Ergebnissen. Das IOM hatte 3 301 Entlassungsberichte von 50 Krankenhäusern der USA mit der ICD-8 nachkodiert. Hierbei fand sich für die Hauptdiagnose eine Übereinstimmung von 65,2 Prozent. Im Unterschied zu den Veröffentlichungen des MDK wurde bei den verbleibenden Berichten auch in Betracht gezogen, dass eine definitive Festlegung des „wahren“ Ergebnisses nicht möglich ist. Dies betraf bei der Hauptdiagnose 10,7 Prozent aller Berichte, bei der Hauptprozedur sogar 16,3 Prozent (bei 73,2 Prozent Übereinstimmung). Selbst die Deutschen Kodierrichtlinien räumen ja ein, dass „sofern 2 oder mehr verwandte Krankheiten ... potenziell der Definition der Hauptdiagnose entsprechen, muss ... entschieden werden, welche Diagnose am besten der Hauptdiagnose-Definition entspricht“.<sup>6)</sup> Das Fehlen dieser „Patt“-Situation in den Publikationen des MDK belegt, dass die veröffentlichten Zahlen ein Verhandlungsergebnis zwischen MDK und Krankenhaus darstellen, und nicht die Messung der Basisdokumentation an einem etablierten Goldstandard.

## Fazit

Aus der vorgestellten Untersuchung ergeben sich Handlungsnotwendigkeiten auf 4 Ebenen.

1. Die ICD-10-GM und die Deutschen Kodierrichtlinien müssen weiter vereinfacht werden. International liegen Er-

kräften überein. Überraschend zeigen die Medizinstudenten bei der Berücksichtigung aller Diagnosen eine höhere Zuverlässigkeit. Das Ergebnis verdeutlicht die Unsicherheit, die eine Kodierung von Diagnosen mit der ICD-10-GM unter Berücksichtigung der Deutschen Kodierrichtlinien mit sich bringt. Diese Unsicherheit ist in hohem Maße durch die im Rahmen der DRG-Einführung vorangetriebene Verfeinerung der Klassifikationen und die Ausweitung von Detailregeln bedingt. Die Festlegung einer „wahren“ Kodierung muss daher regelhaft scheitern. Fehlende Übereinstimmung ist dann häufig nicht Ausdruck des Versagens einer Seite, sondern spiegelt alternative und gleichberechtigte Kodieralternativen wider. Dem derzeit üblichen Verfahren des Medizinischen Dienstes der Krankenversicherung zur Überprüfung der Basisdokumentation der Krankenhäuser fehlt somit eine gesicherte Verfahrensgrundlage. Zu fordern ist eine deutliche Entschlackung und Vereinfachung der Klassifikationen und der Deutschen Kodierrichtlinien. Für die Deutschen Kodierrichtlinien ist mit der Version 2005 ein erster Schritt erfolgt.

gebnisse vor, die eine höhere Zuverlässigkeit bei einer deutlich reduzierten Zahl an Klassen bestätigen.<sup>19)</sup> Es ist nicht zu erwarten, dass die Kodierung von Prozeduren mit dem OPS eine höhere Zuverlässigkeit ergibt. Diese Notwendigkeit ist folglich auch auf den OPS zu übertragen, der noch viel stärker als die ICD-10-GM im Rahmen der DRG-Anpassung verfeinert und durch unsystematische Eingriffe in die Struktur kompliziert wurde.

**2.** Prüfungen der Basisdokumentation durch den MDK im Rahmen von § 275 SGB V und § 17 c KHG müssen modifiziert und dem Erreichbaren in der Kodierung angepasst werden. Das unterschwellige Ziel einer Kostenreduktion muss in die Verhandlungen zwischen Krankenhaus und Kostenträger überführt und von der Diskussion um die korrekte Kodierung getrennt werden.

**3.** Die Nutzung von Routinedaten im Gesundheitswesen hat mit größtmöglicher Umsicht zu erfolgen. Nicht nur MDK-Prüfungen der Basisdokumentation sind von der beschränkten Zuverlässigkeit der Kodierung betroffen. Dies gilt gleichermaßen für die Gestaltung des DRG-Systems, für Leistungsvergleiche, zum Beispiel im Rahmen des Qualitätsberichts nach § 137 SGB V, für Qualitätsindikatoren sowie für einen zukünftig morbiditätsorientierten Risikostrukturausgleich. Bereits die Verwendung höherer Aggregationsebenen der Klassifikationen verbessert die Zuverlässigkeit der Daten deutlich. Völlig kontraproduktiv wäre es, weitere parallele Erhebungen für einzelne Verfahren aufzubauen, die ebenfalls unter ähnlichen Problemen leiden werden. Ziel muss es sein, einen Datenbestand mit ausreichender Reliabilität und hoher Validität aufzubauen, der unter Kenntnis seiner Qualität einer mehrfachen Verwendung zugeführt wird.

**4.** Für die Krankenhäuser gibt es keinen Grund, Kodierauseinandersetzungen mit dem MDK zu scheuen. Einrichtungsintern muss in der Basisdokumentation mehr Variabilität in Kauf genommen werden, als es für das Management von Leistungen und Kosten wünschenswert ist. Angesichts der großen Unterschiede selbst zwischen Medizincontrollern und Dokumentationsfachkräften sollte sowohl in ökonomisch als auch in medizinisch wichtigen Kernbereichen Einheitlichkeit angestrebt werden. Hierbei ist dann allerdings zu erwarten, dass sich die individuellen Kodierstandards zwischen Krankenhäusern unterscheiden und einen einrichtungsübergreifenden Vergleich erschweren.

## Literatur

- 1) Kodierleitfaden Schlaganfall, hrsg. von der Deutschen Schlaganfall-Gesellschaft, Institut für Integrative Versorgung in der Medizin, Schöling 2005, Münster
- 2) Stein, M.; Haas, S.; Lindner, C.; Graf, V.: Einsatz von „Codern“ – ein Erfahrungsbericht nach einem Jahr. *das Krankenhaus* 2004; 96: 693-695
- 3) Cohen, J.: A coefficient of agreement for nominal scales. *Education and Psychology Measurement* 1960; 20: 37-46
- 4) Landis, J. R.; Koch, G. G.: The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-174
- 5) Stausberg, J.; Lehmann, N.: Kodierübungen im Medizinstudium: Wie gut kodieren Anfänger Diagnosen mit der ICD-10? *GMS Med Inform Biom Epidemiol* 2005; 1:Doc04
- 6) Deutsche Krankenhausgesellschaft, Spitzenverbände der Krankenkassen, Verband der privaten Krankenversicherung, Institut für das Entgeltsystem im Krankenhaus. *Deutsche Kodierrichtlinien Version 2004*. Deutsche Krankenhaus Verlagsgesellschaft 2003, Düsseldorf
- 7) Roeder, N.; Hensen, P.; Fiori, W.; Loskamp, N.; Irps, S.; Bunzemeier, H.; Franz, D.; Fürstenberg, T.; Siess, M.: Arzt oder „Koder“ – wer kodiert Diagnosen und Prozeduren im Krankenhaus? *das Krankenhaus* 2004; 96: 802-810
- 8) Hsia, D. C.; Krushat, W. M.; Fagan, A. B.; Tebbutt, J. A.; Kusserow, R. P.: Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. *N Engl J Med* 1988; 318: 352-355
- 9) Klaus, B.; Ritter, A.; Große Hülsewiesche, H.; Beyrle, B.; Euler, H.-U.; Fender, H.; Hübner, M.; von Mittelstaedt, G.: Untersuchung zur Qualität der Kodierungen von Diagnosen und Prozeduren unter DRG-Bedingungen. *Gesundheitswesen* 2005; 67: 9-19
- 10) Stockdreher, K.; Kuls, G.; Modrack, M.; Schlösser, A.; Walter, S.; Weibler-Villabos, U.; Nordhoff, M.: Parallelkodierung von Krankenhausfällen. *das Krankenhaus* 2004; 96: 437-443
- 11) Blum, K.; Müller, U.: Dokumentationsaufwand im Ärztlichen Dienst der Krankenhäuser. Bestandsaufnahme und Verbesserungsvorschläge. Deutsche Krankenhausverlagsgesellschaft 2003, Düsseldorf
- 12) Fietzek, M.: ICD - Verrat an der ärztlichen Kunst. *Deutsches Ärzteblatt* 2004; 101: 1084-1086
- 13) Friedmann, C. P.; Gatti, G. G.; Murphy, G. C.; Franz, T. M.; Fine, P. L.; Heckerling, P. S.; Miller, T. M.: Exploring the boundaries of plausibility: empirical study of a key problem in the design of computer-based clinical simulations. In: Kohane I, ed. *Proceedings AMIA Annual Symposium*. Philadelphia: Hanley & Belfus, 2002: 275-279
- 14) Surján, G.: Questions on validity of International Classification of Diseases-coded diagnoses. *International Journal of Medical Informatics* 1999; 54: 77-95
- 15) Fahlenbrach, C.; Köhler, N.; Halim, A.; Schlottmann, N.: Anpassung der Deutschen Kodierrichtlinien für 2005. *das Krankenhaus* 2004; 96: 890-899
- 16) Dirschedl, P.; Reichle, M.; Röther, M.: Modellprojekt Kodierqualität. *Gesundheitswesen* 2003; 65: 1-7
- 17) Schnabel, M.; Geks, J.; Schunk, T.; Mann, D.; Gerdes, B.; Schaeg, M.: MDK-Überprüfungen zur Dokumentations- und DRG-Kodierqualität. *das Krankenhaus* 2004; 96: 630-633
- 18) Institute of Medicine. Reliability of hospital discharge abstracts. Report of a study. Washington, National Academy of Sciences, 1997
- 19) Nilsson, G.; Petersson, H.; Åhlfeld, H.; Strender, L.-E.: Evaluation of three Swedish ICD-10 primary care versions: Reliability and ease of use in diagnostic coding. *Method Inform Med* 2000; 39: 325-331

Der Titel des Beitrags geht auf eine Anregung von Dr. Martina Franzkowiak de Rodriguez, Düsseldorf, zurück. Die vorgestellte Untersuchung war nur durch den freiwilligen und nicht vergüteten Einsatz der beteiligten Medizincontroller und Dokumentationsfachkräfte möglich, die sich ebenso wie die beteiligten Medizinstudenten bei der Kodierung sehr engagiert haben – wofür die Verfasser ihnen danken.

### Anschriften der Verfasser:

PD Dr. Jürgen Stausberg/  
Dr. Nils Lehmann,  
Institut für Medizinische Informatik, Biometrie  
und Epidemiologie,  
Universitätsklinikum Essen,  
Hufelandstraße 55, 45122 Essen/  
Dr. Dirk Kaczmarek,  
Sankt Marien-Hospital Buer gGmbH,  
Mühlenstraße 5-9, 45894 Gelsenkirchen/  
Markus Stein,  
Klinikum der Stadt Ludwigshafen gGmbH,  
Bremserstraße 79, 67063 Ludwigshafen ■