

**Prof. Dr.-Ing. Holger Hinrichs** ist Professor für Informatik an der Fachhochschule Lübeck. Parallel leitet er den Bereich Forschung & Entwicklung der TIQ Solutions GmbH. E-Mail: holger.hinrichs@tiq-solutions.de

## Datenqualitätsmanagement

# Gesunde Prozesse im Krankenhaus

*Da, wie in vielen anderen Branchen, auch im Gesundheitswesen fast alle kritischen Prozesse auf elektronisch verfügbare Informationen angewiesen sind, hat die Qualität dieser Informationen einen entscheidenden Einfluss auf die Qualität der Prozessausführung sowie auf die Qualität der auf Basis der Informationen getroffenen taktischen und strategischen Entscheidungen. Dies betrifft sowohl die Verarbeitung von hochsensiblen Patientendaten als auch die Abrechnung von Leistungen, die Kommunikation zwischen Abteilungen sowie ökonomisch und wissenschaftlich motivierte Datenauswertungen. Der Artikel zeigt die Verbesserungspotenziale auf.*

## Fallbeispiel Krankenhaus-IT

Wenngleich sich Krankenhäuser in verschiedener Hinsicht von herkömmlichen mittelständischen Unternehmen unterscheiden, zum Beispiel in der Individualität der Behandlung und der Unplanbarkeit des Patientenaufkommens, so gibt es doch zahlreiche Gemeinsamkeiten, was die Aufbau- und Ablauforganisation betrifft. So setzen alle größeren Kliniken Informationssysteme zur Verwaltung von Personal, Geräten, Räumen, Medikamenten, Patienten, Diagnosen und Behandlungen ein. Hinzu kommen spezielle IT-Systeme zur Ansteuerung von medizinischen Geräten, etwa in der Kar-

diologie, die Eingabedaten aus den Informationssystemen benötigen und Ergebnisse für diese liefern (siehe Abb. 1). Krankenhäuser sind also hochgradig datenintensive Organisationen. Folglich hat eine mangelhafte Qualität der erfassten und verarbeiteten Daten signifikante Konsequenzen, wie etwa ineffiziente, nicht kostenoptimierte Beschaffungsprozesse oder massive Einnahmeverluste durch fehlende oder falsche Kodierungen im Rahmen des G-DRG-Systems (siehe unten). „Sand im Getriebe“ der internen Prozesse, insbesondere beim Datenaustausch zwischen Abteilungen, kann durchaus die Qualität der Patientenversorgung beeinträchtigen. Jeder unzufriedene Patient, der das Krankenhaus verlässt, gibt seine Erfahrungen an sein Umfeld weiter. Selbst wenn man berücksichtigt, dass Patienten nicht immer selektiv in ein Krankenhaus kommen und oft eine regionale Bindung zum „nächsten“ Krankenhaus vorliegt, sind die negativen Konsequenzen interner Qualitätsmängel nicht zu vernachlässigen.

## Datenqualitätsmängel im Krankenhaus

Beispiele für Datenqualitätsmängel im Krankenhaus sind:

- ◆ Daten über einen Patienten werden unvollständig oder fehlerhaft erfasst.
- ◆ Abteilung A gibt Daten unvollständig oder falsch an Abteilung B weiter.
- ◆ Abteilung B interpretiert die erhaltenen Daten anders als Abteilung A.
- ◆ Ein und derselbe Patient wird mehrfach im System geführt – keine dieser Akten ist in sich vollständig.
- ◆ Diagnosen und Prozeduren werden falsch, unvollständig oder unspezifisch kodiert.
- ◆ Lieferanten- und Preisangaben zu Medikamenten sind veraltet.
- ◆ Die tatsächlichen Medikamentenbestände stimmen nicht mit denen im IT-System überein.

Derartige Mängel in der Datenqualität wirken sich nicht nur auf operativer und taktischer Ebene negativ aus, sondern insbesondere auch auf strategischer Ebene. So leiden Datenanalysen zur strategischen Entscheidungsunterstützung unter dem schon sprichwörtlichen „Garbage in, garbage out“-Effekt. Selbstverständlich kann eine nachhaltige Optimierung von Behandlungsprozessen im

Hinblick auf Effizienz und Qualität nur dann erfolgen, wenn eine möglichst genaue Zuordnung von Kosten zum Einzelfall stattfindet.

Datenqualitätsmängel im Krankenhaus können zum Beispiel folgende Ursachen haben:

- ◆ Heterogenität von IT-Systemen
- ◆ Unzureichende Festlegung von Zuständigkeiten und Verantwortlichkeiten
- ◆ Unzureichende Sensibilisierung des Personals hinsichtlich Datenerfassung und -pflege

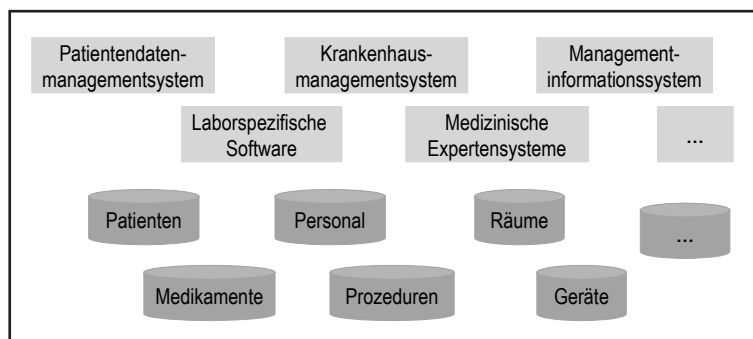


Abb. 1: Daten und IT im Krankenhaus

- ◆ Zeitmangel, Stress
- ◆ Unzureichende softwareseitige Plausibilitätskontrollen
- ◆ Softwarefehler
- ◆ Fehlende oder nicht genutzte Möglichkeiten zur Patientenidentifizierung
- ◆ Unzureichende Dokumentation von Prozessen

Der folgende Abschnitt beschäftigt sich mit Ansätzen zur ursachenorientierten Beseitigung von Datenqualitätsdefiziten.

## Maßnahmen zur Verbesserung der Datenqualität

Aus der Heterogenität von IT-Systemen ergeben sich typische Probleme wie Medienbrüche, redundante Datenhaltung, unterschiedliche Datenrepräsentation etc. Standards wie HL7 [HL7] und DICOM [NEM] helfen, die Interoperabilität zwischen den IT-Systemen zu gewährleisten. Die 1997 gegründete Herstellerinitiative IHE (Integrating the Healthcare Enterprise) arbeitet daran, „lokale Normierungsbestrebungen in die weithin akzeptierten Standards HL7 und DICOM zu integrieren und so die Harmonisierung der beteiligten IT-Systemen weiter voranzutreiben.“ [IHE]

Nun ist Datenqualität ein multidimensionales Konzept, und der Faktor Interoperabilität beeinflusst lediglich einige dieser Dimensionen beziehungsweise Datenqualitätsmerkmale, nämlich die Verfügbarkeit, Einheitlichkeit, Verständlichkeit und Redundanzfreiheit von Daten (vgl. Taxonomie in [Hin02]). Dass Daten, unabhängig davon, über welchen Standard sie ausgetauscht werden, auch Ansprüchen an Kriterien wie Korrektheit, Vollständigkeit, Genauigkeit, Aktualität und Relevanz genügen, kann nur eingeschränkt über technische Regelungen sicher gestellt werden. Als Beispiel seien Eingabekontrollen in Softwaresystemen zur Sicherung der Vollständigkeit und Plausibilität von Informationen genannt.

In vielen Fällen sind Datenqualitätsmängel hinsichtlich der genannten Merkmale auf organisatorische Defizite zurückzuführen. So ist zum Beispiel Datenerfassern oft nicht bekannt, was mit den Daten in späteren Verarbeitungsschritten geschieht (etwa welche Auswertungen gemacht werden), so dass sie unter Umständen Daten gar nicht, falsch oder ungenau erfassen beziehungsweise auf eine regelmäßige Aktualisierung verzichten, was dann in späteren Verarbeitungsschritten, zum Beispiel bei der Abrechnung von Krankheitsfällen, Probleme und damit Kosten verursacht. Organisatorische, ursachenorientierte Maßnahmen wie eine transparente Dokumentation von Prozessen und Zuständigkeiten sowie eine Sensibilisierung aller Beteiligten durch regelmäßige Schulungen versprechen hierbei weitaus nachhaltigere Qualitätsverbesserungen und Kosteneinsparungen als rein symptomorientierte Datenbereinigungsmaßnahmen.

## Hospital Intelligence

Über die Unterstützung des Tagesgeschäfts hinaus besteht in Krankenhäusern zunehmend die Notwendigkeit dispositiver Analysen, um zum Beispiel die Auslastung von Stationen, Ausgaben für Medikamente etc. über die Zeit auszuwerten und daraus entsprechende Managemententscheidungen ableiten zu können. Dies erfordert den Aufbau von Data-Warehouse- und Wissensmanagementsystemen in Krankenhäusern („Hospital Intelligence“). Typischerweise treten Datenqualitätsmängel, die in operativen Systemen bestehen, oft erst bei der Integration der Daten im Zuge der Einrichtung von dispositiven Auswertungssystemen zu Tage. Beispielhaft seien Dubletten genannt, also mehrfach repräsentierte Realweltobjekte (Patienten, Medikamente etc.). Für den Spezialfall der Patientendaten wird die flächendeckende Einführung der elektronischen Gesundheitskarte (eGK) [Bun] Dubletten weitgehend verhindern, ausgenommen Altdaten, denen die Referenz zur Identifikationsnummer der eGK fehlt. Darüber hinaus gilt es jedoch, im Rahmen des ETL-Prozesses (Extraktion aus den meist heterogenen Datenquellen, Transformation in eine einheitliche Repräsentationsform, Laden in das Data Warehouse) umfassende Maßnahmen des Datenqualitätsmanagements (DQM) einzubinden. Der Autor [Hin02] schlägt hierfür ein systematisches, schrittweises Vorgehen vor, das sich heute in vielen auf dem Markt verfügbaren DQM-Werkzeugen wiederfindet (siehe auch Abb. 2).

Dabei wird der Datenbestand zunächst mit Hilfe einfacher statistischer Methoden auf seine Charakteristika und Auffälligkeiten hin untersucht (Wertebereiche, Ausreißer etc.). Im Anschluss an dieses so genannte Data Profiling erfolgt – sofern heterogene, inhaltlich überlappende Datenpools vorliegen – eine Vereinheitlichung der Darstellungsform, um die Vergleichbarkeit der Daten zu gewährleisten. Ausgehend von einer einheitlichen Syntax ist es nun möglich, mit Hilfe medizinischer Expertensysteme die Plausibilität der Daten zu prüfen und inkonsistente Werte zu korrigieren. Anschließend kann automatisiert nach Dubletten innerhalb des Datenbestands gesucht werden. Dabei ist auf eine hinreichende Fehlertoleranz des Abgleichverfahrens zu achten, um durch Heirat, Umzug oder Kassenwechsel verursachte Änderungen in

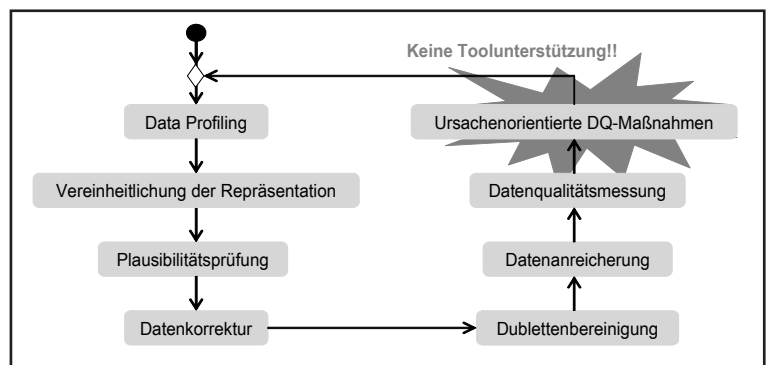


Abb. 2: Regelkreis zum nachhaltigen DQM

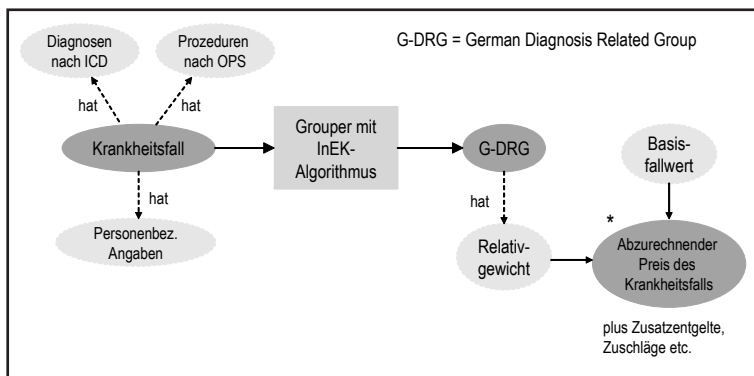


Abb. 3: Bildung von G-DRGs

**Regel nach DKR, Version 2007**

*Entbindung nach vollendeter 41. Schwangerschaftswoche √ Kind zeigt deutliche Übertragungszeichen*  
 ⇒ Kodierung als O48 „Übertragene Schwangerschaft“

**Beispiel eines gemäß der obigen Regel kodierten Falles**

Eine Patientin entbindet ein Kind in der 42. Schwangerschaftswoche.

**Hauptdiagnose:** O48 Übertragene Schwangerschaft

**Nebendiagnosen:** Z37.0! Lebendgeborener Einling  
 O09.7! Schwangerschaftsdauer > 41 Wochen

**Prozedur:** 9-261 Überwachung und Leitung einer Risikogeburt

Abb. 4: Exemplarische DKR-Regel und regelkonformer Fall

den Stammdaten adäquat berücksichtigen zu können. Zweifelhafte Dubletten sollten manuell nachbearbeitet werden. Ist der resultierende konsolidierte Datenbestand immer noch unvollständig, kann über einen Abgleich mit externen Referenzdaten eine Datenanreicherung erfolgen. Eine Messung der Datenqualität hinsichtlich der einzelnen Qualitätsmerkmale gibt Aufschluss darüber, inwieweit der Datenbestand a priori festgelegte Qualitätsanforderungen erfüllt. Die bereits erwähnten ursachenorientierten Maßnahmen sichern schließlich die Nachhaltigkeit des DQM, indem sie den Umfang der erforderlichen Datenbereinigung kontinuierlich senken.

## Fallbeispiel Abrechnung

Die Abrechnung von medizinischen Leistungen erfolgt in Deutschland seit 2004 verpflichtend auf Basis der so genannten German Diagnosis Related Groups (G-DRGs) [DRG]. Dabei handelt es sich um ein Klassifikationssystem zur Einteilung von Krankheitsfällen in „Fallgruppen“, wobei als Vergleichskriterien sowohl diagnosebezogene Angaben (insbesondere Hauptdiagnose, signifikante Nebendiagnosen, Prozeduren inklusive Schweregrad) als auch persönliche Informationen (etwa Alter und Geschlecht des Patienten) herangezogen werden. Diagnosen werden in Deutschland heute nach ICD-10 kodiert, Prozeduren

(Operationen, Behandlungen und Untersuchungen) nach dem „Operationen- und Prozedurenschlüssel“ (OPS). Softwarewerkzeuge, sogenannte Grouper, nehmen die Klassifizierung von Krankheitsfällen anhand der genannten Eingabedaten automatisch vor (siehe Abb. 3). Der von den Groupern zu implementierende Algorithmus wurde vom Institut für das Entgeltsystem im Krankenhaus (InEK) veröffentlicht, das auch die Zertifizierung der Werkzeuge vornimmt.

Jede DRG wird vom InEK hinsichtlich ihrer Kosten gewichtet (so genanntes Relativgewicht). Durch Multiplikation des Relativgewichts mit einem „Basisfallwert“, der nach Abschluss der „Konvergenzphase“ bundesweit einheitlich sein soll, ergibt sich der abzurechnende Preis eines Krankheitsfalls. Von Zusatzentgelten für Leistungen mit hohen Sachkosten und Ab- beziehungsweise Zuschlägen, zum Beispiel in Abhängigkeit von der Dauer des Krankenhausaufenthalts, sei an dieser Stelle abstrahiert. Es ist somit leicht einsichtig, dass die Höhe der von einem Krankenhaus abgerechneten Kosten unmittelbar von der Qualität der vorgenommenen Klassifizierung abhängt. Letztere wird jedoch entscheidend von der Qualität der in den Klassifizierungsprozess eingehenden Informationen beeinflusst. Je nach Relativgewicht der aus einer falschen Kodierung resultierenden DRG werden also entweder zu hohe oder zu geringe Kosten abgerechnet. Entsprechend haben sowohl die Leistungserbringer als auch die Kostenträger (insbesondere die Krankenkassen) ein erhebliches Interesse daran, falsche DRG-Gruppierungen aufgrund schlechter Kodierqualität zu vermeiden. Zusätzlich versuchen die Krankenkassen durch das Einschalten des MDK (Medizinischer Dienst der Krankenkassen), Abschläge bei den DRGs durchzusetzen oder Behandlungen aus dem stationären in den ambulanten Bereich zu verlagern.

Aus DQM-Sicht ist die genannte Problematik insofern schwierig, als sich nicht jedes Krankheitsgeschehen eindeutig nach ICD klassifizieren lässt. Zwar werden die ICD- und OPS-Kataloge regelmäßig bedarfsorientiert angepasst und erweitert, jedoch müssen in der Praxis nach wie vor zum Teil unspezifische Diagnosecodes als „Notlösung“ vergeben werden. Insbesondere bei Multimorbidität ergibt die Anwendung der ICD-10 unter Umständen keine eindeutige Kodierung. Laut Albrecht Zaiß [Zai07] besteht ein weiteres Problem in „komplett fehlenden Kodierungen – insbesondere bei neuen oder neu entgeltrelevanten Prozeduren.“ Automatisierte Prüfungen in diesem Bereich konzentrieren sich meist auf das Datenqualitätsmerkmal der Widerspruchsfreiheit, das heißt eine Software ermittelt anhand eines implementierten Regelwerks, ob eine gewählte Ko-

dierung zu den übrigen Angaben zum Krankheitsfall passt oder nicht. Essentiell für die krankenhausesübergreifende Vergleichbarkeit von Kodierungen sind weiterhin die Deutschen Kodierrichtlinien (DKR), deren Einhaltung weitestgehend automatisiert überprüfbar ist. Eine exemplarische DKR-Regel zusammen mit einem regelkonform kodierten Fall zeigt Abbildung 4. Da in der Medizin oft keine ja/nein-Entscheidungen getroffen werden können, ist ein interaktives Review der fraglichen Datensätze durch den behandelnden Arzt erforderlich.

Alternativ zu vorgegebenen Regelwerken können auch Data Mining-Verfahren eingesetzt werden, um im Rahmen einer „Ausreißeranalyse“ signifikant abweichende „Case Mixes“ automatisiert zu ermitteln und den Fachexperten zur Beurteilung vorzulegen (siehe auch [GrH01]). Derartige Ansätze können zum Beispiel vom MDK eingesetzt werden, um Fälle von Abrechnungsbetrug durch „Upcoding“, also eine systematische Fehlkodierung zugunsten von Fallgruppen mit höheren Erlösen, aufzudecken. Eine detaillierte Beschreibung einer Vorgehensweise zur Ermittlung vermeintlich „überkodierender“ Krankenhäuser findet sich etwa bei Prof. Norbert Roeder von der DRG-Research-Group der Universität Münster [Roe02].

## Fazit

Moderne Krankenhäuser müssen sich den gleichen Herausforderungen stellen wie herkömmliche Unternehmen. Dazu

gehört die Integration von bisher getrennten Einheiten (etwa ambulante und stationäre Versorgung) genauso wie das Prozessmanagement (Controlling, Optimierung etc.) und die Etablierung von Business-Intelligence-Konzepten. Eine hohe Qualität der verarbeiteten Informationen stellt in allen genannten Bereichen eine unerlässliche Voraussetzung für die erfolgreiche Bewältigung dieser Herausforderungen dar.

## Quellen

- [Bun] Bundesministerium für Gesundheit: Die Gesundheitskarte, <http://www.die-gesundheitskarte.de>.
- [DRG] G-DRG-Homepage, <http://www.g-drg.de>.
- [GrH01] Grimmer, U., Hinrichs, H.: Datenqualitätsmanagement mit Data Mining-Unterstützung, in: Hildebrand, K. (Hrsg.): Business Intelligence, HMD 222, S. 70-80, dpunkt, 2001.
- [Hin02] Hinrichs, H.: Datenqualitätsmanagement in Data Warehouse-Systemen, Dissertation, <http://docserver.bis.uni-oldenburg.de/publikationen/dissertation/2002/hindat02/hindat02.html>, 2002.
- [HL7] HL7-Homepage, <http://www.hl7.org>.
- [IHE] IHE-Homepage, <http://www.ihe.net>.
- [NEM] DICOM-Homepage, <http://dicom.nema.org>.
- [Roe02] Roeder, N., et al.: Die DRG-Abrechnung muss MDK-sicher sein, in: f&w, 1:14-17, 2002.
- [Zai07] Zaiß, A.: Lost (in) Information?, in: Management und Krankenhaus, GIT-Verlag, 02/2007.