

Leitlinie zum adaptiven Management von Datenqualität in Kohortenstudien und Registern

Nonnemacher M, Weiland D, Neuhäuser M, Stausberg J

Institut für Medizinische Informatik, Biometrie und Epidemiologie, Universitätsklinikum Essen, Deutschland
michael.nonnemacher@medizin.uni-essen.de

Einleitung und Fragestellung Die Verfügbarkeit elektronisch gespeicherter medizinischer Datenbestände, wie sie u. a. von Kompetenznetzen in der Medizin aufgebaut werden, und der Umfang ihrer möglichen Nutzung nehmen deutlich zu. Damit gewinnt die Frage nach der Qualität der dort erfassten Daten immer größere Bedeutung. Zur Sicherung und Verbesserung der Datenqualität im Rahmen der medizinischen Dokumentation existieren vielfältige Konzepte, die sich auf unterschiedliche Phasen im Lebenszyklus einer Datenerhebung (z. B. Planung, Durchführung, Auswertung) beziehen. Typischerweise beschränken sich solche Konzepte auf eine spezielle Nutzung der Daten (z.B. für zulassungsrelevante klinische Studien, Abrechnungszwecke). Die stärkste Formalisierung solcher Konzepte findet sich zum einen im Bereich von Arzneimittelstudien über die Richtlinien zur Good Clinical Practice [1] sowie im Bereich gesetzlich verankerter Berichtspflichten von Leistungserbringern im Gesundheitswesen über die Vorgabe von Plausibilitätsprüfungen. Im Datenmanagement von Kohortenstudien und Registern gibt es ein vielfältiges Instrumentarium von sinnvollen Maßnahmen; hierzu wurde bereits ein Rahmenmodell entwickelt [2]. Zwei besonders relevante Maßnahmen sind der Originaldatenabgleich (SDV = Source Data Verification) [3] und das Feedback [4]. Bei Kohortenstudien und Registern ist eine komplette Überprüfung der gemeldeten Daten mit der Primärdokumentation der Datenlieferanten, d.h. eine vollständige SDV, zwar wünschenswert, aber wegen Unpraktikabilität und fehlender Finanzierbarkeit nicht umsetzbar. Als Alternative bietet sich ein adaptives Monitoring an, welches eine dynamische Anpassung der Aktivitäten zur Qualitätsprüfung an die spezifischen Bedürfnisse eines Registers/Kohortenstudie und die vorgefundene Datenqualität erlaubt. An dieser Stelle setzt das hier vorgestellte Projekt, gefördert von der Telematikplattform für Medizinische Forschungsnetze (TMF) e. V., an. Ziel ist die Entwicklung einer Leitlinie zum adaptiven Management von Datenqualität in Kohortenstudien und Registern mit den Schwerpunkten SDV und Feedback sowie die Entwicklung eines Software-Werkzeugs zur Unterstützung der Leitlinienanwendung. In einem im Vorfeld des Projekts durchgeführten Workshop "Medizinische Daten in Forschung und Versorgung" im Februar 2005 in Essen wurde die Notwendigkeit eines adaptiven Monitorings eindrücklich betont sowie die fehlende konzeptionelle Basis und der Bedarf nach Ausarbeitung einer Leitlinie und ihrer konkreten Unterstützung durch Software-Werkzeuge deutlich [5].

Material und Methoden Die Vorgehensweise orientiert sich an nationalen Empfehlungen zur Entwicklung von medizinischen Leitlinien [6]. Vertreter der Kompetenznetze Angeborene Herzfehler, Demenzen, HIV/AIDS und Parkinson, sowie Vertreter des Förderers übernahmen Kontrolle und Review der Projektarbeit. Der erste Schritt bei der Entwicklung der Leitlinie war die Identifikation und Analyse der bestverfügbaren Evidenz über eine Sichtung der Literatur und eine Befragung von Experten. Daran schlossen sich die Konsensusfindung zur inhaltlichen Zusammenführung und Ergänzung der Ergebnisse, die Ableitung der Leitlinie und der Entwurf des Software-Werkzeugs an.

Das Vorgehen bei der Literatursichtung entspricht den Schritten zur Quellensuche und Quellenbewertung, wie es die BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH für die externe vergleichende Qualitätssicherung nach § 137 SGB V beschreibt [7]. Als relevante Bereiche wurden Qualitätsmessung in Medizin und Industrie, Qualitätsindikatoren, statistische Methoden in der Qualitätssicherung, Qualitätslenkung, Ringversuche sowie Richtlinien und Empfehlungen für die Erstellung und den Inhalt medizinischer Leitlinien identifiziert. Die Quellen umfassten Datenbanken (z.B. Medline über Pubmed), Suchmaschinen (z.B. Google) und Institutionen aus dem Bereich Qualitätsmanagement (z.B. Zentralstelle der Deutschen Ärzteschaft zur Qualitätssicherung in der Medizin). Weiterhin wurden die Literaturverzeichnisse gefundener Quellen und Literaturempfehlungen von Experten in die Suche einbezogen. Das Hauptaugenmerk lag auf der Suche in Medline über www.pubmed.org mit den Schlagwörtern (MeSH Headings) "Quality Control", "Data Collection" und "Feedback". Die Suchergebnisse wurden an Hand ihres Titels und gegebenenfalls eines Abstracts von einem der Autoren auf Relevanz geprüft. Hierbei wurde nach einer ersten Analyse der verfügbaren Evidenz zum Thema Datenmanagement auf die drei Themen Indikatoren für Datenqualität, adaptive SDV und Feedback zur Datenqualität fokussiert. Die als relevant eingeschätzten Artikel wurden im Volltext beschafft, so weit möglich, und qualitativ im Hinblick auf die Projektziele analysiert. Unabhängig von den hier gewonnenen Ergebnissen wurde festgelegt, durchgehend das Glossar der Arbeitsgruppe Qualitätsmanagement in der Medizin der GMDS [8] als Referenz für die Definition qualitätsrelevanter Begriffe zu Grunde zu legen. Dieses Glossar findet sowohl in der Fachöffentlichkeit als auch in der Praxis durchgehend Anerkennung und kann derzeit in Deutschland als De-facto-Standard gelten.

Das Software-Werkzeug soll zum einen die Adaption der Leitlinie an die spezifischen Charakteristika einer Kohortenstudie oder eines Registers ermöglichen, und zum anderen die Anwendung der Leitlinie direkt unterstützen. Die eXtensible Markup Language (XML) wurde als Standardformat für den Datenaustausch festgelegt. Prozessmodellierung und Software-Entwurf werden mit der Unified Modeling Language (UML) durchgeführt.

Ergebnisse Aus der Literatur ließen sich 24 Indikatoren für Datenqualität extrahieren, die wir den Ebenen Plausibilität, Organisation und Richtigkeit zugeordnet haben. Diese Ebenen entsprechen den Ebenen Struktur, Prozess und Ergebnis, die Donabedian für die medizinische Versorgung vorgeschlagen hat [9].

Die SDV wird oft zur Untersuchung der Vollständigkeit oder Genauigkeit bei klinischen Studien, Registern und Elektronischen Patientenakten benutzt. Einzelne Empfehlungen zur Anzahl der Variablen, die in eine SDV eingeschlossen werden sollen, und zur Frequenz der SDV sind nicht durch empirische Untersuchungen mit Vergleich unterschiedlicher Ansätze begründet. In unserem Konzept für eine adaptierte SDV generieren wir einen Qualitätsscore aus den Werten der einzelnen Qualitätsindikatoren und den Ergebnissen der SDV. Die SDV soll dann die aktuelle Qualitätskategorie des Scores bestätigen, wobei jede Kategorie mit einem Erwartungswert für den Anteil der Beobachtungseinheiten mit mindestens einem Fehler assoziiert ist. Durch eine statistische Fallzahlplanung, wie sie in ähnlicher Weise bei der Fehlbelegungsprüfung im Krankenhausbereich genutzt wird, bestimmen wir den Umfang (Anzahl Beobachtungseinheiten) und die Tiefe (Anzahl Merkmale) der SDV. Die geforderte Präzision des gemessenen Anteils soll umso geringer sein, je besser die Datenqualität ist. Dies führt automatisch zu einer höheren Fallzahl bei schlechter Datenqualität.

Bei der Verwendung der Leitlinie im Datenmanagement werden die Anwendungsfälle Bestimmung der Datenqualität in der zentralen Studien-/Registerdatenbank, Bestimmung der in der SDV sichtbaren Datenqualität, Planung der SDV und Planung des Feedbacks unterschieden. Akteur ist der Datenmanager der Studie/des Registers. Der Arbeitsablauf umfasst zwei ineinander greifende Zyklen. Der eine Zyklus besteht aus einer periodischen Bestimmung der für die Leitlinienanwendung erforderlichen Parameter durch den Datenmanager, der Software-gestützten Generierung von Empfehlungen für SDV und Feedback gemäß der Leitlinie, und der Durchführung von SDV und Feedback durch den Datenmanager gemäß den Empfehlungen. Ein zweiter Zyklus repräsentiert den Anreiz zur Qualitätsverbesserung. Die Empfehlungen enthalten eine Darstellung der Datenqualität in Form von Qualitätsindikatoren und Qualitätsscore. Diese Information kann sowohl vom zentralen Datenmanager als auch in den meldenden Zentren zur Schwachstellenanalyse genutzt werden, die wiederum zur Implementierung qualitätsverbessernder Maßnahmen führen sollte. Dies wird hoffentlich zu einer Qualitätsverbesserung bei zukünftig gesammelten Daten führen.

Diskussion Es hat sich gezeigt, dass eine empirische Absicherung für Strategien zum Datenmanagement und Monitoring im Bereich von Kohortenstudien und Registern bislang weitgehend fehlen. Ähnliches scheint für klinische Studien zu gelten. Die in der Industrie zum Qualitätsmanagement eingesetzten Verfahren (z.B. Total Quality Management) erscheinen für das gegebene Problem zu komplex.

Die Anwendung der Leitlinie hat einen mehrfachen Nutzen. Sie bietet ein Konzept zur kontinuierlichen Verbesserung von Datenqualität, wodurch die Aussagefähigkeit und Nutzbarkeit der Daten wesentlich verbessert werden können. Die Datenqualität der Erhebungszentren und des zentralen Datenbestandes wird bestimmbar. Eine vollständige und sehr aufwendige SDV wird vermieden. Durch die Anpassung der Maßnahmen an die Datenqualität wird der Aufwand, im wesentlichen für die SDV, dort konzentriert, wo es auf Grund schlechterer Datenqualität besonders nötig ist. Dies erlaubt einen effizienten Einsatz vorhandener Ressourcen.

Die Leitlinie steht in Version 1.0 den Mitgliedern der TMF zur Verfügung. Eine breitere Veröffentlichung ist vorgesehen. Zur Evaluation der Leitlinie befindet sich eine Studie mit mehreren Kompetenznetzen in Planung.

Danksagung Dieses Projekt wurde durch die Telematikplattform für Medizinische Forschungsnetze (TMF) e. V. gefördert. Die Autoren danken den Projektpartnern aus den Kompetenznetzen Angeborene Herzfehler, Demenzen, HIV/AIDS und Parkinson für die Kooperation sowie Frau Katarzyna Domagala für die Unterstützung bei der Literaturrecherche und Herrn Carsten Brannhoff für die Modellierung mit UML.

Literatur

- [1] <http://www.ich.org/>
- [2] Arts DG, De Keizer NF, Scheffer GJ. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. *J. Am. Med. Inform. Assoc.* 2002; 9: 600-11.
- [3] Khosla R, Verma DD, Kapur A, Khosla S. Efficient source data verification. *Indian Journal of Pharmacology* 2000; 32: 180-6.
- [4] De Lusignan S. Using feedback to raise the quality of primary care computer data: a literature review. In: Engelbrecht R, Geissbühler A, Lovis C, Mihalas G, eds. *Connecting Medical Informatics and Bio-Informatics. Proceedings of MIE2005.* Amsterdam: IOS, 2005: 593-8.
- [5] <http://www.ekmed.de/routinedaten/>
- [6] AWMF, Ärztliches Zentrum für Qualität in der Medizin. *Das Leitlinien-Manual von AWMF und ÄZQ. Z. ärztl. Fortbild. Qual.sich.* 2001; 95: Supplement I.
- [7] Mohr VD, Bauer J, Döbler K, Eckert O, Fischer B, Woldenga C, Hrsg. *Qualität sichtbar machen. BQS-Qualitätsreport 2004.* Düsseldorf: BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH, 2005: 60-61.
- [8] GMDS-Arbeitsgruppe "Qualitätsmanagement in der Medizin". *Begriffe und Konzepte des Qualitätsmanagements. Informatik, Biometrie und Epidemiologie in Medizin und Biologie* 2003;34:1-61.
- [9] Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988; 260: 1743-8.