

AUSWERTUNG ELEKTRONISCHER GESUNDHEITSAKTE FÜR FORSCHUNG UND QUALITÄTSMANAGEMENT

Gall W¹, Grossmann W², Dorda W¹

Kurzfassung

Lebensbegleitende Elektronische Gesundheitsakte können zusätzlich zur Unterstützung bei der Patientenbehandlung wertvolle Informationen für Forschung und Qualitätsmanagement liefern. Ausgehend von wissenschaftlichen Auswertungen in einem Krankenhaus werden Anforderungen an Auswertungen in einem Gesundheitssystem diskutiert. Es wird festgestellt, dass Archetypen dabei eine zentrale Rolle einnehmen können. Sie können im Rahmen einer Auswertung bei der Visualisierung, semantischen Verknüpfung und standardisierten Übertragung eingesetzt werden.

1. Einleitung

Die Einführung der *elektronischen Gesundheitsakte (ELGA)* ist ein strategisches eHealth-Ziel in Österreich [1] und in Europa [2]. Die Realisierung dieses wichtigen Vorhabens wird die *Dokumentation*, den *Austausch* und die *Auswertung* patientenbezogener Daten verändern. Auf standardisierten EHRA's (Electronic Health Record Architectures) [3] basierende Zugriffsmethoden sollen nicht nur zur Verbesserung der Behandlung eines individuellen Patienten beitragen, sondern auch Qualitätsimpulse für die Forschung und die Gesundheitspolitik liefern.

Auswertungen können die Gesundheitsakte eines Patienten betreffen oder patientenübergreifend durchgeführt werden. Im ersten Fall wird meist ein Patient selbst oder ein Arzt im Zuge einer Behandlung in einer Patientenakte relevante Daten suchen. Im zweiten Fall wird meist ein Forscher oder Statistiker viele Patientenakte nach gemeinsamen Sachverhalten analysieren.

Der Focus dieses Artikels liegt auf dem zweiten Fall der patientenübergreifenden Auswertung. Es werden Anforderungen an Systeme zur Auswertung standardisierter elektronischer Gesundheitsakte diskutiert.

2. Methoden

Alle Auswertungsanforderungen im Detail zu diskutieren würde den Rahmen dieser Arbeit überschreiten. Wir gehen daher von den wichtigsten Anforderungen bei Auswertungen in einem Kran-

¹ Institut für Medizinische Informations- und Auswertesysteme, Besondere Einrichtung für Medizinische Statistik und Informatik, Medizinische Universität Wien

² Institut für Scientific Computing, Universität Wien

kenhaus-Umfeld (als state-of-the-art) aus und analysieren zusätzliche Anforderungen bei Auswertungen im ELGA-Umfeld.

Den Ausgangspunkt unserer Analyse bildet das wissenschaftliche Auswertesystem ArchiMed [4] am AKH-Wien. Das Ziel unserer Analyse bildet ein auf den Empfehlungen der ELGA-Machbarkeitsstudie [5] basierendes Auswertesystem für das Gesundheitswesen.

Abbildung 1 zeigt das System ArchiMed im Krankenhaus-Umfeld. Mit dem System werden jährlich ca. 2500 Auswertungen durchgeführt.

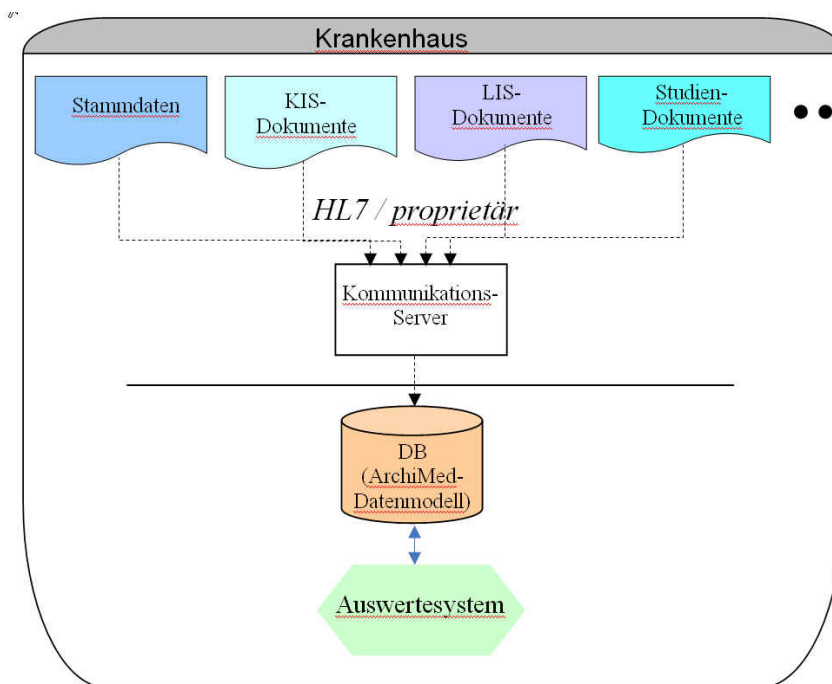


Abbildung 1: Ein Auswertesystem im Krankenhaus-Umfeld

Das System erfüllt die wichtigsten *Anforderungen*, um heterogene medizinische Daten aus unterschiedlichen Routineanwendungen eines Krankenhauses gemeinsam mit Studiendaten patientenübergreifend auswerten zu können:

- a) *Datenintegration* durch Datenüberleitung aus Dokumentationssystemen und Datenzugriff über Views eines zentralen Datenmodells [6].
- b) *Unterstützung der drei typischen Auswerteschritte* von Routedaten, die durch die drei Begriffe *Kollektivbildung* (Suche nach Patienten), *Variablenauswahl* (Selektion der interessierenden Patientendaten und deren Aufbereitung zu einer statistischen Matrix) und *Statistische Analyse* (vorbereitete Statistikfunktionen bis volle Funktionalität des Statistikpaketes SAS) charakterisiert sind.
- c) *Analyse des Zeitverlaufs* von Krankengeschichten [7] durch die Bereitstellung mächtiger Zeitrelationen für Selektionsbedingungen („und innerhalb von 2-4 Monaten danach“).
- d) *Textanalyse* durch den Einsatz von Textoperatoren und Thesauern. Im Vorgängersystem WAREL wurden zusätzlich Algorithmen zur Wortstammzerlegung verwendet [8], wobei sich die Erstellung der fachspezifischen Wissensbasen als zu aufwendig erwiesen hat.

- e) *Unterstützung verschiedenster Benutzergruppen* durch die Bandbreite der Retrievalfunktionalität: von mächtigen Komponenten für Poweruser bis zu vorgefertigten, adaptierbaren Auswertungen für Novizen.
- f) *Intuitive Anfrageformulierung* durch einen grafischen Abfrageeditor (in Testphase).
- g) *Webbasierte Auswertungen* durch eine Webkomponente zur Exekution vorgefertigter Auswertungen.
- h) *Datenschutz* durch ein bis auf einzelne Variable spezifizierbares Berechtigungssystem.

3. Ergebnisse

Auswertungen in einem über Jahre gewachsenem Krankenhaus-Umfeld haben viele Gemeinsamkeiten mit Auswertungen in einem ELGA-Umfeld. Eine unmittelbare Übertragung des Anforderungsprofils für die Kommunikation innerhalb einer großen heterogenen Institution auf das ELGA-Umfeld scheint aber nicht möglich. Wie Abbildung 2 zeigt, ist schon der Datenzugriff mittels standardisierter Austauschformate nicht direkt vergleichbar.

Im Folgenden sollen kurz die Adaptionen und Erweiterungen hinsichtlich der oben genannten Anforderungen für ein ELGA-Auswertesystem diskutiert werden:

- a) *Datenintegration*: die Kommunikation wird auf *Archetypen / Templates* basieren, welche auf einem *EHR-Informationsmodell* aufsetzen [9].
- b) *Unterstützung der drei typischen Auswerteschritte*: für die *Kollektivbildung* und *Variablenauswahl* sind die patientenbezogenen Administrations-Attribute (Patient, Aufenthalt, Dokument) durch *organisatorische* (Art des Gesundheitsanbieters), *räumliche* (Bundesländer) und nicht zuletzt durch *klinische Konzepte* (repräsentiert durch Archetypen) zu erweitern. Die Anforderungen an die Statistik bleiben unverändert.

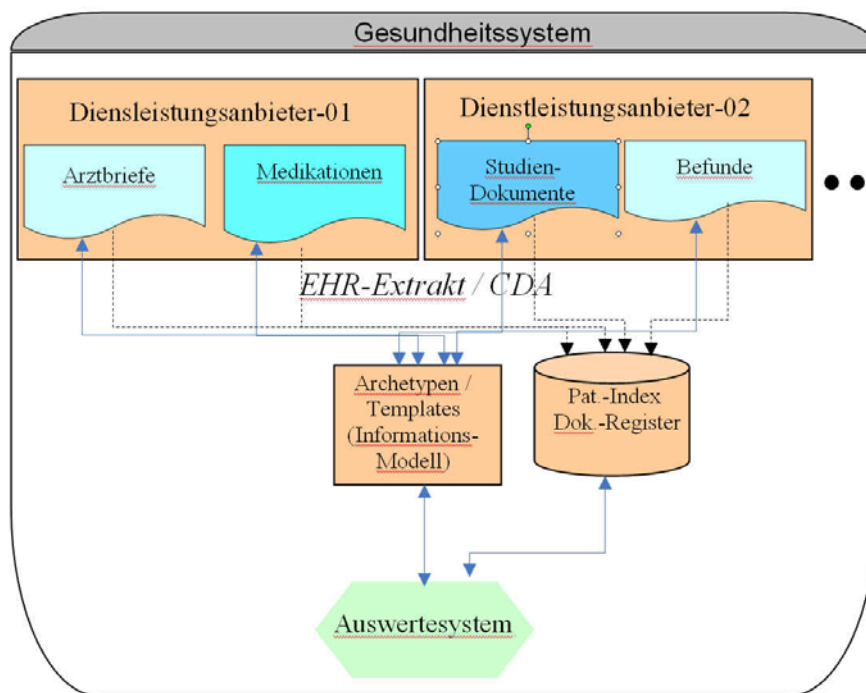


Abbildung 2: Ein Auswertesystem im ELGA-Umfeld

- c) *Analyse des Zeitverlaufs*: die Anforderungen bleiben unverändert (hoch).
- d) *Textanalyse*: Aufgrund unterschiedlichster Datenproduzenten (niedergelassene Ärzte, Kliniken, Labors) steigt die Heterogenität der Daten. Dies erfordert eine komplexe semantische Analyse von unterschiedlichen Begriffen, Schreibweisen oder Abkürzungen. Hier kann der Einsatz von *Ontologien* [10] im Zusammenwirken mit Archetypen zur Problemlösung beitragen.
- e) *Unterstützung verschiedenster Benutzergruppen*: die Bandbreite wird im ELGA-Umfeld nicht größer. Allerdings können durch die räumliche Ausdehnung User kaum mit Schulung und Betreuung unterstützt werden. Die Online-Hilfe muss alleine ausreichen und somit qualitativ höheren Anforderungen genügen.
- f) *Intuitive Anfrageformulierung*: Die größere Komplexität bei der Kollektivbildung und Variableauswahl erfordert eine entsprechende Erweiterung in der Anfrageformulierung. Im ELGA-Umfeld spielen auch der Datenproduzent und der Kontext der Daten, der durch den Archetypen bestimmt wird eine Rolle. Durch den erweiterten Userkreis gewinnen Symbole und Metapher an Bedeutung. Visualisierte Archetypen können zur Übersicht und zur Auswahl von Variablen dienen.
- g) *Webbasierte Auswertungen*: Speziell die grafische Anfrageformulierung auf Basis der Web-Technologien stellt eine technische Herausforderung dar. Hier ergeben sich durch Konzepte des Web2.0 wie Ajax neue Möglichkeiten.
- h) *Datenschutz*: im Prinzip müssen alle Daten, die Gesundheitseinrichtungen verlassen, anonymisiert sein [11]. Werden bei komplexen, mehrstufigen Abfragen (Kollektivbildung Variableauswahl) Identifikatoren benötigt, muss zusätzlich zur Anonymisierung und Pseudonomisierung auf Konzepten der *k-Anonymität* aufgesetzt werden [12]. Bei Extraktionen von Datengruppen muss auf versehentlich mitgelieferte Identifikationsmerkmale geachtet werden.

Eine zentrale Bedeutung wird dem Konzept der Archetypen zukommen. So können bei der Auswahl einzelner Auswertungsvariablen vorhandene komplexe Datenstrukturen durch die *Visualisierung* von Archetypen übersichtlich dargestellt werden. Weiters kann über einen Archetyp durch *semantische Links* – in Zusammenwirken mit Thesauren und Ontologien – die Suche auf andere Archetypen ausgedehnt werden.

Wird beispielsweise bei einer Auswertung von Diabetesdaten nach Patienten mit einem bestimmten Anamnesetext und einem bestimmten Blutzuckerbereich gesucht, könnten zunächst durch die Visualisierung des Archetyps „Diabetes“ die Variablen „Anamnese“ und „Blutglukose“ gefunden werden. Weiters könnte über den Ontologie-Bereich der Variablen „Blutglukose“ in diesem Archetyp die Suche auf den Archetyp „Laborchemie“ ausgedehnt werden, welcher die synonyme Variable „Blutzucker“ enthält.

4. Diskussion

Die Einführung der standardisierten elektronischen Gesundheitsakte ist ein definiertes strategisches eHealth-Ziel in Österreich und Europa. In dieser Arbeit sind einige zentrale Anforderungen an Systeme zur patientenübergreifenden Auswertung von ELGA-Daten diskutiert worden. Einige dieser Aspekte lassen sich auch auf das Suchen von Daten innerhalb eines Patienten durch den Patienten selbst oder durch einen behandelnden Arzt anwenden.

Die hier diskutierten Anforderungen können natürlich erweitert (z.B. Ergebnisverwaltung) und verfeinert (z.B. Anonymisierungsalgorithmen) werden. Sie wurden auch noch nicht in Hinblick auf den sehr unterschiedlichen Forschungsstand, die Komplexität und die Kosten untersucht.

Eine zentrale Rolle wird den Archetypen zukommen. Dieses Konzept wird von den drei wichtigsten Standardisierungsgremien für EHRAs (CEN, HL7, openEHR) forciert. Hier eröffnet sich ein weites Forschungsfeld und viele Fragen bleiben offen. Zwei wesentliche Fragen sind z.B.: „Wie viele Archetypen sind notwendig?“ oder „Soll es einen Archetyp „ELGA“ als Gesamtview geben?“.

5. Schlussfolgerung

Meist ist der Aspekt der Datenauswertung gegenüber der Datenerfassung ein Stiefkind. Oft wird er erst verspätet, wenn bereits Daten aufwendig erfasst wurden, betrachtet. Zur Berücksichtigung der Anforderungen für die Auswertung ist es dann oft bereits zu spät. Auch in Standards für EHRAs findet man unter dem Punkt „Retrieval“ oft nur die Anmerkung „future work“. Nur wenn auswertorientierte Anforderungen schon bei der Einführung der standardisierten ELGA in Betracht gezogen werden kann diese „Goldgrube“ an Daten auch für Qualitätsmanagement und epidemiologische Forschungen effizient genutzt werden.

6. Referenzen

- [1] Austrian National Council. Healthcare Reform Act 2005. http://www.parlament.gv.at/pls/portal/docs/page/PG/DE/XXII/I/I_00693/fname_030664.pdf (accessed 2007-02-20).
- [2] Commission of the European Communities. e-Health - making healthcare better for European citizens: An action plan for a European e-Health Area. http://eur-lex.europa.eu/LexUriServ/site/en/com/2004/com2004_0356en01.pdf (accessed 2007-02-20).
- [3] Blobel BG. Advanced EHR architectures--promises or reality. *Methods Inf Med* 2006;45(1):95-101.
- [4] Gall W, Sachs P, Duftschmid G, Dorda W. A retrieval system for the selection and statistical analysis of clinical data. *Med Inform Internet Med*. 1999;24(3):201-12.
- [5] Bundesministerium für Gesundheit und Frauen. Machbarkeitsstudie ELGA. <http://www.bmgf.gv.at/cms/site/detail.htm?thema=CH0415&doc=CMS1169796766007> (accessed 2007-02-20).
- [6] Dorda W, Wrba T, Duftschmid G, Sachs P, Gall W, Rehnelt C, et al. ArchiMed: a medical information and retrieval system. *Methods Inf Med*. 1999;38(1):16-24.
- [7] Shaha Y, Combi C. Timing is everything. Time-oriented clinical information systems. *West J Med*. 1998;168(2):105-13.
- [8] Dorda W. Data-screening and retrieval of medical data by the system WAREL. *Methods Inf Med*. 1990;29(1):3-11.
- [9] Beale T. Archetypes and the EHR. *Stud Health Technol Inform*. 2003;96:238-44.
- [10] Smith B, Ceusters W. An ontology-based methodology for the migration of biomedical terminologies to electronic health records. *AMIA Annu Symp Proc*. 2005:704-8.

[11] Austrian Data Protection Commission. Datenschutzgesetz 2000 (DSG 2000), BGBl. I Nr. 165/1999. <http://www.dsk.gv.at/dsg2000d.htm> (accessed 2007-02-20).

[12] Albert J, David D, Lagerfeld C, Keil W, Rey, H, Schöner C, Siebert I. Management-Papier „Pseudonymisierung / Anonymisierung“. 2004-03-16. <http://atg.gvg-koeln.de/xpage/objects/pseudonymisierung/docs/5/files/MP040316.pdf> (accessed 2007-02-20).