Secondary Use of Structured Patient Data: Interim Results of A Systematic Review

Riikka VUOKKO1, Päivi MÄKELÄ-BENG6, Hannele HYPPÖNEN6 and Persephone DOUPI1
1National Institute for Health and Welfare in Finland (THL)

Abstract. In addition to patient care, EHR data are increasingly in demand for secondary purposes, e.g., administration, research and enterprise resource planning. We conducted a systematic literature review and subsequent analysis of 85 articles focusing on the secondary use of structured patient records. We grounded the analysis on how patient records have been structured, how these structures have been evaluated and what are the main results achieved from the secondary use viewpoint. We conclude that secondary use requires complete and interoperable patient records, which in turn depend on better alignment of primary and secondary users’ needs and benefits.

Keywords. Data structure; Electronic health records; Information quality; Process quality; Systematic literature review.

Introduction

Patient information is documented and used during care. However, it has value also for other use purposes. As Electronic Health Record systems (EHRs) are implemented in health care organizations and integrated with national level systems, large data bases are formed offering access to patient information contributed by various caregivers. Information in EHRs can be further utilized, e.g., in administration and enterprise resource planning, in health care service planning, in research and development, in decision support systems, in statistics and registers. In spite of wide use of patient information, systematic reviews evaluating EHR structures are rare. By EHR structures we refer to any method utilized to provide structure to the data entered in a patient’s record, ranging from structured forms and NLP-methods applied to free text to classifications and ontologies. In this paper, we study the benefits of structured patient data from the viewpoint of secondary use [1] based on the following three research questions: How have electronic health records been structured and for which secondary use purposes? How have health record structures been evaluated or assessed from the secondary users’ perspective? What are the main results achieved by the use of structured patient information for secondary purposes? To answer these questions, we have conducted a systematic literature review, in which the secondary use of structured patient data was one of the chosen focuses. The other two focuses were nurses’ [2] and clinicians’ viewpoints.

1 Corresponding author.
1. Methods

The systematic literature review on structured patient records was planned as a twelve step research protocol based on the Cochrane reviews protocol [3]. Our method (the review protocol, search strategy, databases searched, exclusion/inclusion criteria, and analytical framework) has been described in detail in another article [4]. The search across 15 databases resulted in 680 unique articles (from years 1975-2010). Of these, 102 articles were included in the secondary use analysis group. After full text review by two researchers, 17 more were excluded, leaving 85 for detailed analysis.

In the analysis framework, main result categories [5, 6] were usability, system quality and reliability, information contents and quality, availability, user acceptance, type and amount of use, clinical process impacts and differences, productivity or time saved, patient safety, cost-efficiency, service quality and patient satisfaction, health impacts, care coordination, patient participation, access to services and secondary impacts, such as facilitation of decision support and knowledge management systems.

2. Results

Of the 85 articles, 53 i.e. over half were written by American authors and one fourth of the articles were of European origin. The articles concerned mostly secondary or tertiary care and only in one fifth the study was conducted in the primary care environment. In two thirds of the articles, patient record structures were still in development or in testing phase, and in clinical use only in one third.

Most of the articles focused on development of structures or EHRs. Other goals for secondary use included enhancing health care management and administration, clinical development or research and evaluating work done by secretaries or coders. Most original studies reported several secondary use goals. In 46 articles, the focus was on evaluation of structures for documentation, in 44 articles on evaluation of structures used for search and retrieval of patient information. A few articles studied both documentation and retrieval.

The most common evaluation viewpoint was that of administration or research and development. In these articles, the evaluation usually consisted of auditing structure quality after implementation. The results indicated that auditing the structures further developed users’ skills. In evaluating the use of structures, common indicators were utilization rate, consistency and completeness. Evaluation focused typically on various topics: 1) the quality of codes or structures (N=17), 2) assessing information quality (N=28), 3) care guidelines vs. structures in use (N=7), 4) service quality when a structure was tested or used in a natural language processing (NLP) system (N=29), 5) impact of structures used in NLP-systems to identify adverse effects and complications in care or identifying specific patient groups (N=5), 6) optimizing the hospital’s reimbursement system (N=2), and 7) evaluating health impacts of structured patient records (N=2).

The most commonly used structures within EHRs were classifications, terminologies and standardized forms in entering patient information and NLP-methods utilizing EHR-structures in data retrieval. The most common structures evaluated in the

---

2 The complete list of articles is available upon request by the authors since the space limitations of this paper do not allow inclusion of the full list.
articles were codes, classifications and terminologies in 22 articles, most commonly ICD 9-10 and SNOMED. An information or conceptual model was evaluated in three and documentation standards or standardized forms in 16 articles. In 15 articles free text was used as comparison or for development purposes and in five articles the structuring method was not identifiable. Further 26 articles combined two or several structuring methods, e.g., codes and information models in four, codes and structured forms in three. 19 articles concerned combinations of free text and structured or coded data. The methods to apply or use the data structures were documenting and entering the data in 38 articles, various NLP-based systems (mostly for retrieval of data) in 32, decision support systems in five, resource management systems in four and unidentifiable in eight articles.

Most of the identified benefits of utilizing EHR structured data for secondary purposes concentrated in the areas of system quality and reliability or information content and quality. Only a handful of the individual articles evaluated impact on care processes, productivity and costs, patient safety, care quality or health impacts and in most articles these were mostly discussed as goals of secondary use and less as evidence based impacts from the use of structured patient records for secondary purposes.

Results of assessing quality of codes, classifications and terminologies (N=17): SNOMED enables detailed classification of patient data [7, 8], but still, even 15% clinical data can prove to be unclassifiable with it [9]. The International classification of diseases (ICD) provided reliable classification that could be utilized in secondary use [10, 11, 12]. Challenges to use classifications for documenting patient information were e.g. consistency in classifying and using different codes and choosing suitable code or term [13, 14] or using extensions [15]. Using structured forms to enter data contributed to better consistency [16, 17].

Application method quality (N=29) was evaluated in most articles studying applications and tools of NLP-methods, text mining or data mining which were utilizing various structures. NLP was mostly used to enhance search and retrieval of patient information, but also decision support tools were developed and evaluated. A common solution was to utilize diagnostic codes to identify patients or patient groups [e.g. 18, 19, 20, 21]. Best results were achieved when diagnostic codes were combined with, e.g., operational codes [18], pharmacological information [22] or symptoms [21]. An automated classifier can achieve a more reliable or exact result than manual classification of patient information [23], and classifying can be further enhanced by using synonyms, abbreviations etc. [24]. Using narrative patient information for secondary use purposes requires knowledge of the substance area [25], but it can be used, e.g., with semi-automated rules in the data retrieval to predict morbidity [26].

Information quality (N=17): Necessary skills and degree of using the implemented information structures varied according to professional group, context of use, specialty and administrative support. E.g. nurses were most meticulous in structured documentation which, in turn, was most commonly used in follow-up than in acute care.

By applying NLP-based functions in the EHR, increased compliance to care guidelines and processes was noted. E.g. partly automated problem lists, summaries, or prompts and tool-tips in the user interface guided users’ actions. The process impacts implicated earlier identification of patients and their health issues which led to shorter care durations. Shorter hospital stays had further impact on lowering care costs. Patient safety impacts were another documented evaluation result. E.g. system components for drug related adverse effects and interactions were evaluated in relation to the use of
ICD-codes. The ICD-codes were combined with patient’s demographic and pharmacological data. A standardized drug list decreased drug related adverse effects and potentially harmful use of drugs, and increased overall documentation quality.

3. Discussion

Two main secondary use focuses emerged in our review: the reliable use of clinical classifications and terminologies in documenting patient information, and various NLP-methods utilizing structures in data retrieval. Overall, documentation was more uniform and information quality better (e.g. consistency and completeness) with structured patient records. Structures used in EHRs enabled secondary use, such as, hospital resource planning and invoicing, assessing compliance to care guidelines and developing decision support. NLP-based tools enabled also the use of free text documentation in secondary use. There was indication of these types of benefits in individual articles.

The limitations of our review could be caused by the search strategy and due to the fact that EHR data structuring methods have been described ambiguously and have no clear definition as such. Also the methods for evaluating patient record structures are not well established. Furthermore, only 14 % of the identifiable data structures were in routine use, while the remaining were in development or testing phase. Structuring of patient data has been conducted since 1970’s. Some of the reviews included in the article are old, as we wanted to get a full view of methods’ evolution over time. The results indicate that earlier articles evaluate structures of documentation and discuss structures vs. free text. Later, a shift towards evaluating structures of data retrieval emerges.

Bracketing the structuring interventions per result or impact category proved to be a cumbersome task for 3 reasons. The interventions (both structuring methods and utilizing or applying structures for data retrieval) were manifold. The interventions were applied and assessed in diverse contexts with data from varying specialties and care process phases. Most articles discussed several result categories, and contained more references to potential benefits than actual results based on measurements.

In conclusion, there was no evidence that structured EHRs would result in better care of patients. Structured documentation contributed to more complete and reliable records. It was nevertheless feared that important information would be omitted if free text was not allowed to be used alongside structured data. Care professionals’ motivation for structured documenting as users of EHRs depends on the perceived benefits. Clinical benefits can be gained with, e.g. decision support or other contextual support tools within the EHR. Further studies will give valuable evidence on how EHRs are utilized for secondary purposes, and how documentation structures serve different users’ needs, e.g. administration, statistics and research and development, in addition to clinical purposes.

References