How to Distinguish Double Documentation from Documentation of Distinct Data

Gert Galster

Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

Abstract

Within health informatics, it has for years been an objective to substitute double documentation of health data with data reuse. The elimination of double documentation potentially reduces unnecessary, additional work, and thereby releases human resources. To meet these potentials it is necessary to have a clear delineation of the involved concepts as well as a method to distinguish between double documentation and documentation of distinct data. With the incidence of double documentation in Danish health registries as a starting point and on the basis of a qualitative literature review, the present paper explores the meaning of double documentation and proposes a definition. With clinical examples and the proposed definition, key issues for the distinction between double documentation and documentation of distinct data are highlighted, and it is argued that the identification of double documentation necessitates a semantic analysis of the registered data.

Keywords:
Medical Informatics Computing, Medical Record Linkage, Records as Topic, Semantics

Introduction

Denmark is among the countries with most population-based health registration, thereby using significant resources in registration, both in local, regional and national health registries and information systems. In Denmark there are approx. 200 national health registries, and several of these have mandatory reporting[1]. There is only little integration between the various health information systems and registries; hence a considerable degree of double documentation[2].

Through several years of development in Danish health informatics it has been a national intent to substitute double documentation with data reuse. Currently the extent of double documentation is a political issue: In an analysis by the Danish Ministry of the Interior and Health and the National Board of Health[1] it is stated that 30 million double documentations occur annually, corresponding to 50-100 full-time jobs (nurses and/or physicians), and that these double documentations as far as possible should be replaced by data reuse. Moreover it has as a national initiative[3] been indicated that the extent of double documentation in the largest Danish health registries should be reduced by at least half over the next year.

In an effort to substitute double documentation with data reuse on this scale, it is required that the different stakeholders in national health registration do agree on which data are unique and which are interchangeable; what is double documentation and what is documentation of distinct data.

The objective of this paper is – through literature studies and analysis - to create the means for operational distinction between double documentation and documentation of distinct data.

Methods

Present paper was induced by speculations on how the Danish central health authorities will eliminate millions of double documentations within a rather short timespan. Such speculations gave rise to two central questions:

- What is the meaning of "double documentation"?
- What is the difference between double documentation and documentation of distinct data?

In order to uncover the meaning of double documentation a qualitative literature review was conducted. The objective was to analyse the various descriptions of double documentation in such depth that it would be possible to synthesise an operational definition of double documentation.

Basis for the review was a number of official publications from the stakeholders in Danish national health registration[1-6] supplemented with an all-fields search in PubMed for "double documentation", "double registration", "dual documentation" and "dual registration". The PubMed search revealed only 21 papers, and as their descriptions of double documentation was rather vague and casuistic, it was decided to supplement further with papers and student projects, which were known to deal with the concept of double documentation. In this way a total of 31 papers was included in the review.

The qualitative literature review was conducted as an analysis supported by the following questions: How is double documentation perceived? Which pros and cons to double documentation? What are the reasons for double documentation? Which systems are involved?
Results

Different perceptions of double documentation

The term 'double documentation' can mean both an act ('to perform double documentation') and the result of this act ('the registry contains double documentation'). In this study the concept in question is the act.

It is characteristic for the reviewed literature that the perception of double documentation is implicit – definitions are rare and casuistic. Eg. the Danish health authorities in a paper about quality data uses this definition of double documentation:

... data which are mandatory to report to the National Patient Register or other key health registries and which are also reported to the clinical databases[1]

and Lundgren, Hesselbo et.al in a study on double documentation and interdisciplinarity uses this definition:

... that nurses and physicians records information about the same phenomenon in their separate health records.[7]

Such definitions with narrow and case-oriented focus on specific systems or organisational aspects, has limited usability in a broader sense. The different meanings of double documentation was therefore obtained by interpretation of the reviewed papers:

- A general view on double documentation is that the same phenomenon is recorded manually more than once, e.g.[1,2,7-10]. One of these papers[2] explicitly includes transcription as an example of double documentation.

- A few papers makes a distinction between double documentation and transcription of data[6,11] thereby excluding simple manual transfer of information from one system to another.

- Furthermore a few papers demonstrate some very different perceptions of double documentation, namely 'double sampling' (the taking of samples which are differing in time, site or otherwise), e.g. [12,13], and 'double count' (unintentional recount of objects), e.g. [14,15]. These meanings of double documentation are irrelevant to this study and are therefore excluded from the following.

Central to the efforts to eliminate or reduce the extent of double documentation is a perception that double documentation and data reuse can be seen as complementary opposites: When a piece of information is required, it can be obtained either by one or the other. This dualistic view is widespread and is explicitly expressed in several papers, e.g. [7,11] including those representing the view of the stakeholders in Danish national health registration[1,4-6].

Advantages and disadvantages of double documentation

There are situations where double documentation advantageously is used to ensure process and data quality - for example to verify that information about critical medication is correct[10] or to ensure quality in registration of health services[8].

But most studies describe the drawbacks of double documentation - that it constitutes unnecessary extra work, e.g. [1,2,6,7,11], that it negatively affects the clinical staff morale[7,11,16] and patient confidence in staff [7], and that it may lead to ambiguous responsibility for documentation[9,17], inconsistent data and other forms of reduced data quality, e.g. [7,9,11].

Reasons to perform double documentation

While a few papers describe situations, where intentional double documentation is used to ensure process and data quality[7,8,10], the vast majority describes double documentation as an undesirable consequence of technical or organisational circumstances. The main reason for double documentation is lack of (technical) interoperability, e.g. [4-6,9,11], even though several papers also describes a lack of semantic[1,5-7] and even juridical[6] interoperability.

The scope of double documentation

While some papers describe double documentation in a local scope (e.g patient administrative system, health records, observation charts[7-10], others describe it in a regional or national scope (e.g. national health registries and clinical quality databases)[1,5,6,16]. It is obvious from the different descriptions, that the differences in scope is just a matter of aspect.

Discussion

Proposal for definitions

By generalising the different perceptions of double documentation, the common characteristic is, that within an area of information gathering (local, regional or national) redundancy occurs, and this redundancy is introduced by manual registration. Therefore the following definition of double documentation is proposed:

\[
\text{double documentation (synonym: double registration)}
\]

- manual registration of information which - within the relevant sources - already exists

Although this definition applies to health informatics, it is deliberately open-ended in order also to cover double documentation within other domains.

Based on the definition of double documentation and taking into account that data reuse and double documentation appears to be complementary opposites, a working definition of data reuse is proposed:
The finer meaning of double documentation

The proposed definition of double documentation is a generalisation of several sources with slightly different interpretations of the concept. To use the definition in practice, e.g. to determine what is double documentation and what is not, it needs some clarification.

The definition requires that the registration is performed manually. Fully automatic registration (e.g. by use of automatic measuring equipment) or data copying (e.g. by automatic data transfer between registries) is therefore not considered double documentation. The proposed definition does not require that the person doing the registration has a particular function or belong to a particular profession.

The definition does not specify how the manual registration and the subsequent storage of data are implemented. Neither does it specify data format, media or type of information system. Thus, data entries can be double documentation regardless whether data capture is on paper forms or on electronic forms, and regardless whether the documentation system is a health record or a national health registry.

Even though a few papers makes a distinction between double documentation and transcription, it must be assumed, that the latter is far more frequent. Manual transcription is covered by the proposed definition, and is thus considered double documentation.

It is an integral part of the definition that the information is already available - i.e. that double documentation introduces redundancy. The definition does however not require that this redundancy is explicit nor meets particular formalisms. Hence, the entering of data which unambiguously can be derived from existing information, could be considered as double documentation. In this way double documentation could be to enter the patient's age, when date of birth is known, because then age can easily be derived from existing information. Similarly, it could be considered as a double documentation to enter smoking-status (yes/no) when information about tobacco consumption (e.g. g/day) is present.

Especially two aspects of double documentation according to the proposed definition will need clarification in practice:

- Which sources are relevant?
- Is the requested information available within these sources?

In practice both of these two questions for obvious reasons have to be answered on a per-situation basis.

Which sources are relevant?

Denmark is among the countries with most population-based health registration, and as mentioned in the section about the scope of double documentation there is in theory a wide variety of sources to choose from. Which sources are relevant in practice must be determined from a professional assessment of the actual data.

But even if sources are considered relevant in an actual situation, they are not necessarily available for data extraction, and neither data quality nor data structure are necessarily adequate for data reuse. This is illustrated by the Danish Ministry of the Interior and Health, the National Board of Health and the Danish Regions, who jointly through a nearly 10-year period have demonstrated a very broad interpretation of sources relevant to exchange of health data - encompassing information systems within general health care, hospitals, national health registries, social care registries, clinical quality databases, etc.[1,5,6] This continuous statement from the Danish central health authorities represents an ambition, which at present seems neither legally nor technically realisable[6,18].

Does the requested information exist

In the definition of double documentation it is pivotal that the information in question already exists. And this is a crucial point in the quest to eliminate double documentation: It is a challenge to distinguish between double documentation and documentation of distinct data.

The Danish central health authorities have in an effort to identify double documentations in national health registries made a search for interchangeable, clinical registrations[1]. Data were regarded as interchangeable, and thereby the result of double documentation, if they had the same label, were of the same type (same dimension) and had the same possible set of values. This approach seems to be rather superficial, as it does not take context into account. As pointed out by several researchers, it is difficult to convey the full meaning of clinical data, when they are represented in a formal way, e.g. in records, because their meaning might vary depending on their appearance in a context[19,20].

On this basis it is of interest to consider, what impact context may have on the meaning of clinical data in connection with double documentation. Is it necessary to match context attributes, when distinguishing between interchangeable and non-interchangeable data, i.e. between double documentation and documentation of distinct data?

To illustrate, what impact context can have on data interchangeability, one can imagine that a patient's body weight was recorded in a database 10 years ago, without additional context information. Now, 10 years later, the patient is having surgery, and the preoperative weight needs to be recorded. Will this be double documentation?

Body weight can be expected to change significantly during the specified period, and body weight then and now must be considered non-interchangeable information. So the actual registration will not be a double documentation. But - in the same situation - if the recorded body weight was not 10 years but just one week old, then the information could be reused and registration would be double documentation. But then again, if the information about body weight was not to be used in preparation for ordinary surgery, but for assessment of dialysis needs, even a week old body weight would be out of
date, and a new registration would not be double documentation.

Time has different impact on other types of information. While e.g. sex and height (for adults) are very stable information types, diuresis and tumour growth has quite another "durability".

This example illustrates that with regard to the identification of double documentation, both data context (e.g. recorded 10 years or one week ago) and usage context (e.g. dialysis or ordinary surgery) can have a crucial impact. It also illustrates that the impact of different types of context (e.g. data context and usage context) may be mutually dependent. Finally the example illustrates that the sensitivity to context (e.g. time of observation) may be dependent of the type of information (e.g. sex or body weight).

In another example one can imagine a health record with several instances of registered blood pressure without additional context information. Now in preparation for medical treatment a resting blood pressure needs to be recorded. Will this be double documentation?

It is known that blood pressure can vary considerably under a number of commonly occurring conditions and situations. Therefore, one should not use random blood pressure measurements as a basis for treatment, and the proposed registration is not double documentation.

The same type of considerations apply to a wide range of information types, for which the circumstances in relation to information gathering is a critical part of the information - e.g. exercise ECG, fasting blood glucose level, resting heart rate, supine chest X-ray, etc. This is not the case for all combinations of circumstance and information type, though; concepts like e.g. fasting height and upright weight has no clinical meaning.

This example compared to the previous, illustrates the existence of yet another important context attribute, namely observation circumstance, which may also have crucial impact with regard to the identification of double documentation. This example also illustrates that the sensitivity to context (e.g. rest, fasting) may be dependent of the type of information (e.g. ECG, sex).

It is illustrated above that context can have crucial impact on the identification of double documentation. So can a lack of contextual information. Maybe all of the blood pressure measurements in the previous example were adequately measured resting blood pressures, where the context was just not fully recorded. If so, it would be the lack of contextual information rather than non-existence of data that would prevent the identification of the former registrations as interchangeable data.

In actual databases and registries the contextual information exists, in varying degrees, as registered meta-data. Considering the multitude of possible context attributes, possible data types (dimensions) and possible combinations thereof, it must be assumed that existing meta-data under given circumstances may prove insufficient. Thus, situations where double documentation is required, because the available data does not include the necessary context, must be expected.

Conclusion

On the basis of a literature review the concept of double documentation is delineated and a formal definition is proposed. This definition offers conceptual clarification and can serve as the basis for distinction between double documentation and documentation of distinct data.

But the definition is not in itself sufficient to operationalise this distinction. Through examples it is demonstrated that it requires concrete semantic analysis to determine, whether a given piece of information can be obtained through data reuse. It is also illustrated that this analysis only to a limited extent can be generalised, because it must take into account:

- the type of information
- an in principle interminable list of context attributes, each with a weight determined both by the context in which the piece of information is currently to be used and the context in which it was originally recorded
- the extent and quality of the associated meta information

The conditional significance of context attributes and need for metadata could be a topic for further study.

In the Danish perspective, it seems likely that the 30 million yearly double documentations reported by the Danish health authorities is somewhat exaggerated, as this result is based on context-ignorant counting. The fact that the necessary semantic analysis must be done on a situational basis, makes it unlikely that double registrations to any significant extent will be substituted by data reuse within a manageable period.

References


Address for correspondence
Gert Galster, Medical Informatics Group, Fredrik Bajers Vej 7 D1, DK-9220 Aalborg, Denmark. Mail: gal(at)hst.aau.dk