

Requirements of tools and techniques to support the entry of structured nursing data

Nicholas R. Hardiker^a, Suzanne Bakken^b

^aSalford Health Informatics Research Environment (SHIRE), University of Salford, UK

^bSchool of Nursing and Department of Biomedical Informatics, Columbia University, New York, NY, USA

Abstract

The benefits of structured data are widely accepted within the nursing informatics community. However, despite the existence of structured data in the form of well-established nursing terminologies, computer-based nursing record systems are yet to achieve widespread adoption and few of the potential benefits have yet to be realized. In this paper we argue the need for tools and techniques to support the entry of structured nursing data into computer-based systems. In the absence of a generally accepted solution, we build on preliminary work carried out at the 2002 Nursing Terminology Summit and analyze the results of other studies in order to identify a preliminary set of requirements or desiderata for such tools and techniques. These requirements are centered on: how structured data is presented to users for selection; how to mediate between a variety of conceptual structures - terminologies, information models, user interface models and models of the clinical process; and how to reduce the considerable modeling burden through re-use of modeling constructs. Further applied research is needed with the ultimate goal of developing a general solution that will benefit nurses, other professionals and ultimately their patients.

Keywords:

Nursing Records, Terminology, Data Collection.

Introduction

Many authors have articulated the need for structured clinical data i.e. clinical data that is in a format suitable for automatic processing. For example van Ginneken argues that 'For benefits that go beyond accessibility and readability, structured data are essential...' ([1], p. 226). Powsner et al. acknowledge the need for structured, coded data and agree with the view that most of the benefits of computer-based records rely on such data (as opposed to free text). [2]

Within the nursing informatics community, this is a widely held view. To counter this view to some extent, Powsner et al. also acknowledge the value of tradition e.g. 'flowing prose to paint an evocative clinical picture' ([2], p. 1618) – it could be argued that this applies particularly to the discursive tradition

of nursing records – and point to outstanding issues such as the time overhead of collecting data via terminologies and the loss of control and flexibility in design. They posit that 'computer-based records will be efficient clinical tools only if more time is saved in record review and report-writing than is taken in entering structured, coded data and operating the record system itself' 1 ([2], p. 1619) and set a challenge to systems developers and purchasers to ensure that the disadvantages of collecting structured, coded data are outweighed by better presentation and more effective automatic processing of that data.

We believe that to minimize those disadvantages, i.e. to improve the collection of structured nursing data in the first place, will be equally beneficial. A recent Institute of Medicine report suggests a causal link between the availability of structured data and the efficiency and accuracy of data collection: 'Having clinical data represented with a standardized terminology and in a machine-readable format would reduce the significant data collection burden at the provider level, as well as the associated costs, and would likely increase the accuracy of the data reported.' [3, p.11] We support this view, and assume that there is a need for such structured data within nursing, in order to support increasingly diverse and complex information requirements. However, we do not believe that the mere existence of structured data is an end in itself. Structured data, embodied within a number of well-established terminologies, has been widely available to nursing at least since the early 1990s [4-6]. And yet, in common with other clinical systems, computer-based nursing record systems are far from universal, and it could be argued that few of the potential benefits have been realized [1].

Our intention within this paper is to explore issues around the collection of structured data that is derived from nursing terminologies. It builds on preliminary work initiated at the 2002 Nursing Terminology Summit that used a number of techniques to inform the provision of support for the entry of struc-

¹ It is interesting to note and is perhaps a reflection on the current 'state-of-the-science' that while Powsner et al. present a number of advantages of computer-based records over paper records, the great majority of these concern aspects of data storage and display, rather than any improvements to data entry.

tured nursing data.[7] While the discussion is focused on nursing, many of the arguments apply equally to other disciplines.

Wyatt and Wright present a list of nearly 20 difficulties faced by clinicians in using patient data. For example, they argue that 'paper records are organized to assist data entry, not retrieval' (p. 1375). We would argue that this provides more than adequate motivation to draw out data entry as a separate component of any computer-based patient record system. There is of course a need to preserve the context around the data, but there is no need to confine the use of data according to how it is entered. Nor is there a need to restrict data entry according to the structure and form of its ultimate presentation e.g. a distinct nursing record or a particular style of table.

We agree that in terms of clinical records a great deal has been written about data recording at the expense of data use. [8] However, we would argue that an adequate and general solution to data entry has not yet been determined and, as data entry has such a profound impact on how the entered data is subsequently used, there is a need for continued research in this area.

The remainder of this paper takes the preliminary work mentioned previously and begins to form a set of requirements or desiderata for tools and techniques to support the entry of structured data. This set covers among other things:

- Determining how elements are presented to users
- Mediating between a variety of models: terminologies, information models, user interface models, and models of the clinical process
- Promoting and facilitating re-use.

Requirements for supporting tools and techniques

Presentation of elements to users

For structured data to be of value to nursing, it must meet the two fundamental criteria of being both useful (e.g., appropriate for the task) and usable (e.g., match the capacity of its users to use it). [9] The goal therefore with respect to data entry is to present candidate structured data to users in a way that meets their expectations and that fits with clinical practice. In an attempt to achieve this goal, certain existing and emerging terminologies include relatively global notions such as subsets, preferred terms, etc. [10] However, solutions based on terminologies alone can at best be only partial. On the other hand, it would be a mistake to believe that terminologies can be divorced completely from their eventual use. As Rector notes: "many terminology developers explicitly dissociate themselves from clinical pragmatics...consigning it to 'implementation'."([11], p. 242).

There is clearly a dependency between terminologies (and the structured data they embody) and so-called clinical pragmatics. Only a few studies have examined how clinicians interact

with terminologies during the data entry task. [12] The 'implementation' within computer-based record systems of clinical pragmatics includes, but is not limited to, the structures and processes required to support data entry. This is different from broader issues of computer-human interface, such as those addressed by Poon et al. [13] It is more closely related to how structured data from terminology systems is organized and how clinical workflow is represented for the benefit of users. This is the main focus of this paper.

In addition to the challenge presented previously to ensure that the disadvantages of collecting structured, coded data are outweighed by better presentation and more effective automatic processing of that data, Powsner et al. set a further related challenge: to present only data that is relevant, while preserving its original context and permitting access to any remaining data not needed immediately.[2] We would argue that the challenge applies equally in supporting data entry as in supporting other informational tasks such as navigation and interpretation. Thus there is a need, among other things, for some degree of filtering.

van Ginneken sets out a number of recommendations for supporting the entry of structured data:

'These recommendations are:

- Predictable order of labels, entry fields and data
- Headings and subheadings in progress notes
- User-defined views for data entry and look-up' ([1], p. 227)

The same author goes on to define the content of a so-called 'domain model' i.e. the meta-data required to support data entry and presentation within in particular domain. This content includes:

- A list of concepts represented by predefined terms
- Context information to put concepts in a meaningful order and for unambiguous interpretation e.g. whether an observation is a sign or a symptom
- Properties associated with concepts in their context e.g. value type and multiplicity
- Constraints, both simple e.g. maximum numeric value, and complex e.g. systolic blood pressure must be greater than diastolic.

Much of this content had been identified separately in the form of an expanded set of requirements by Hardiker et al. at the 2002 Nursing Terminology Summit through a process of use case analysis, scenario development and prototyping. These requirements include a need to:

- Initiate the dialogue
- Determine what should/should not appear
- Group related but heterogeneous elements
- Make links based on professional knowledge

- Include elements from outside the terminology
- Determine which synonyms should appear
- Determine the level of granularity at which terms should appear
- Determine the order of elements
- Determine which elements are optional and which are mandatory
- Consider users and their tasks to determine the context of the dialogue.

These two sets of findings were arrived at independently but they share a number of characteristics and it is possible to derive our first requirement.

Requirement 1: In terms of how elements are presented to users, tools and techniques to support the entry of structured nursing data should describe and dictate:

- Which elements should appear
- Which possible values should appear for each element
- The form in which they should appear
- The order in which they should appear
- The context in which they should appear.

Mediation between different models

Terminologies

Several authors indicate a clear need to separate tools and techniques to support data entry from the terminologies from which the structured data to be entered is derived. van Ginneken argues that 'a proper domain model for SDE [Structured Data Entry] is content independent and allows independent development of applications for SDE, based on that model.' ([1], p. 234) Further, she states that 'Controlled medical terminologies seem a logical basis for the support of SDE, but they do not represent the information needed to generate intuitive user interfaces that anticipate what is relevant to enter in a given context' (p. 234).

However, while it is all but impossible to conceive of tools and techniques to support the entry of structured data that is entirely independent of the terminologies from which the structured data is derived, it would appear desirable to somehow decouple them.

Requirement 2: Tools and techniques to support the entry of structured nursing data should be de-coupled from the terminologies from which the structured data is derived.

Hardiker et al. support this view in regards to both so-called 'interface terminologies' and 'reference terminologies' - formal terminology systems, such as GALEN [14] or SNOMED®RT [15], that can support rich descriptions of clinical encounters, data re-use and data comparisons. [7]. They argue that terminologies are designed to perform terminological reasoning,

and as it is difficult to capture the pragmatics of routine practice within such systems, there is a need for a dialogue sub-system (broadly equivalent to van Ginneken's domain model) to manage things that heretofore had been managed by the terminologies themselves e.g. non-terminological organizational groupings or pragmatic associations between concepts. However, they also acknowledge the need for and benefits of some form of interaction or dependency between their dialogue sub-system and the underlying terminology. van Ginneken supports this view: 'These terminologies however, represent a standard, can [sic] be used to enhance the consistency of metadata for data entry, and permit semantically more powerful retrieval'. ([1], p. 235)

Thus there is a potential role for terminologies in the construction of a dialogue sub-system or domain model. There may be a role also for so-called reference terminologies, terminologies that are independent of interface terminologies, in providing at least the major part of the content. Practical experience with the Pen & Pad [16, 17] and Telenurse [18] programs would appear to bear this out. The collaboration between the College of American Pathologists and the UK National Health Service to develop SNOMED®CT marks a shift in emphasis towards a terminology system with reference and interface characteristics to support routine use in computer-based records. [10] However, its success in supporting the entry of structured data has yet to be widely demonstrated.

Requirement 3: Formal terminologies, stable, extensive and independent from interface terminologies should provide the major part of the content of tools and techniques to support the entry of structured nursing data.

Information models

In addition to structuring data, the record itself is a significant factor in data collection. Wright et al. recommend formatting records so that all information can be recorded at the point of elicitation (in order to prevent or at least restrict memory lapses). [19] The same authors also point to a set of principles of information design. These principles are recommended as a foundation to the collection and organization of data and seek to aid the interpretation of medical record data. The authors acknowledge that 'To interpret the data correctly...we also need the data in the right format and language. The format and language depend partly on how clinicians enter data in the record...' (P. 1542) - that is, the presentation of data is dependent in part on how it is entered. The design principles include:

- setting the context
- writing informative headings
- limiting the information under each heading
- including 'signposts and landmarks'
- organizing information for more than one professional group
- making the organization of the material explicit.

While these are important considerations, we agree with van Ginneken that tools and techniques to support the entry of structured nursing data are complementary to information models (models that structure data in order to construct records or for the exchange of data in the form of messages data) i.e. they deal with dialogue or clinical narrative rather than any secondary structuring of data. However, given the dependency identified previously between data entry and data presentation, any support for the entry of structured data must consider the design principles above. Dialogue or clinical narrative depends not just on the content of information models; it depends also on their structure. An obvious example of when this might be an issue concerns the difference between a clinical concept such as 'Diabetes' appearing under the record or message heading 'Diagnosis' and the same concept appearing under the heading 'Family history'. Such issues are subject to negotiation by those responsible for developing terminologies and those responsible for developing record architectures and message structures. The outcome of resolving such issues however might comfortably reside in a dialogue sub-system or domain model.

Requirement 4: Tools and techniques to support the entry of structured data should embody decisions concerning potential conflicts between information models and terminologies.

User interface models

As indicated previously, the focus of this paper is on the tools and techniques required to support the entry of structured data. This is different to the 'bells and whistles' of the user interface such as navigation controls and menus. [13] van Ginneken argues the case for a direct relationship between domain model content and interface content. [1] There may of course be a dependency between interface structures and terminologies e.g. the use of scales requires the identification of a relevant and complete set of possible values for points on the scale and a mapping between them. However, drawing a parallel with the discussion on information models in the previous subsection, while there may be a relationship between the dialogue sub-system or domain model and the *content* of the interface, the relationship to the *structures* of the user interface should be indirect.

Requirement 5: Tools and techniques to support the entry of structured data should be de-coupled from interface structures in which the structured data appears.

Clinical process models

Clinical process models in the context of this paper require little discussion. One of the major motivations behind our desire to specify requirements of tools and techniques to support the entry of structured data is to support the clinical process as it relates to recording practice. Attempts to embody the clinical process in tools to support data entry have been proposed by other researchers. For example, Wilcox et al. describe the use at Intermountain Health Care of text-based templates i.e. proforma notes with placeholders for entering textual data.[20] In its efforts to improve completeness such an approach may, at the point of recording, hide or even confuse context. In re-

sponse to the problems associated with text-based data entry, Brown et al. propose an alternative – template-based entry of structured data. [21] The Brown et al. study demonstrates that for certain forms of data, the entry of structured data is preferred over text entry. However, it would appear from their paper that the structured data is very tightly bound up with, and even appears to form an integral part of, the data entry templates. We believe that this is undesirable, for reasons outlined in a previous section. Other studies have demonstrated the success of succinct, structured approaches that are less closely-coupled to the underlying structured data [1, 16, 17]. Evolving standards for clinical document architectures, templates, and archetypes support such notions. [22, 23]

Requirement 6: Tools and techniques to support the entry of structured nursing data should reflect and embody the clinical process as it relates to recording practice.

Facilitating re-use

In a history of the PROMIS system, perhaps one the first major systematic approaches to guiding clinical data entry, Schultz describes how clinical content was incorporated into, and data entry guided by, frames. [24] In 1967 there were 90 frames within PROMIS covering cardiovascular problems. Seven years later there were approximately 30 000 frames; at the time very close to storage capacity. At its peak the system is reported to have comprised over double this amount [25].

The PROMIS experience, despite its reported success, serves to demonstrate that to expect to define in advance the content of all possible data entry forms is un-realistic as the number of forms will naturally tend to be greater than the number of elements contained therein. For example, elements will appear on more than one form and smaller forms will be combined into larger forms. It is also unrealistic to expect to pre-specify in advance all possible paths among those data entry forms as the number of paths among different forms will tend to be greater than the number of forms themselves. For example clinical practice does not follow a single path and branching and back-tracking should be possible.

Thus, rather than to attempt to specify in advance the content of data entry forms and the paths through those forms, we believe that any tools and techniques to support the entry of structured data, in all but the simplest of applications, should be generative in nature to allow for re-use of individual elements and to allow for multiple paths through the dialogue. We believe that: it should be possible to re-use elements within a form on a more specific topic; it should be possible to embed a form, or the elements therein, within another more general form; and it should be possible to specify a dependency between individual data entry forms or the elements therein. Most importantly, it should be possible to perform these functions on-the-fly, rather than attempting to define them in advance.

Requirement 7: Tools and techniques to support the entry of structured data should be generative allowing for re-use of elements and multiple narratives.

Summary

In this paper we have demonstrated the need, within nursing, for structured data. Experience bears out the fact that the mere existence of structured data is not an end in itself. In order to integrate the use of structured data into nursing practice - to contribute to the widespread adoption of computer-based record systems, and to realize more of the potential benefits of such systems - tools and techniques are required to support data entry. A survey of relevant research in this area, in addition to preliminary work conducted at the 2002 Nursing Terminology Summit, has allowed us to begin to specify a tentative set of requirements for these tools and techniques. These requirements focus on how elements are presented to users, how to mediate between a variety of conceptual structures and how to reduce the modeling burden by promoting and facilitating re-use. Further research is needed in the form of real practical applications; without such research it is clear that nursing will be impoverished.

Acknowledgments

The authors would like to thank participants at the 2002 Nursing Terminology Summit, in particular Amy Coenen, Evelyn Hovenga, Joyce Sensmeier, Annette Smith and Chris Zingo, who helped to formulate some of the initial ideas presented within this paper.

References

- [1] van Ginneken A. Considerations for the Representation of Meta-Data for the Support of Structured Data Entry. *Methods of Information in Medicine* 2003;42:226-235.
- [2] Powsner S, Wyatt J, Wright P. Opportunities for and challenges of computerisation. *The Lancet* 1998;352:1617-1622.
- [3] Institute of Medicine. *Key Capabilities of an Electronic Health Record System*. Washington DC: The National Academies Press; 2003.
- [4] Kim M, McFarland G, McLane A. *Pocket Guide to Nursing Diagnoses*. 5th ed. St. Louis: Mosby-Year Book Inc.; 1993.
- [5] Martin K, Scheet N. *The Omaha System: a pocket guide for Community Health Nursing*. Philadelphia: W.B.Saunders Company; 1992.
- [6] Saba V. *Home Health Care Classification of Nursing Diagnoses and Interventions*. Washington D.C.: Georgetown University; 1994.
- [7] Hardiker N, Coenen A, Hovenga E, Sensmeier J, Smith A, Zingo C. Understanding the dialogue between users and formal terminology systems. In: *Proceedings of Nursing Informatics*; Rio de Janeiro; 2003.
- [8] Wyatt J, Wright P. Design should help use of patients' data. *The Lancet* 1998;352:1375-1378.
- [9] Hoy D, Hardiker N. Representing Nursing: The Puzzle of Language and Classification. In: Mortensen R, editor. *ICNP and Tele-matic Applications for Nurses in Europe*. Amsterdam: IOS Press; 1999.
- [10] Stearns M, Price C, Spackman K, Wang A. SNOMED Clinical Terms: Overview of the development process and project status. In: Bakken S, editor. *Proceedings of American Medical Informatics*

Association Annual Symposium; Philadelphia, Pennsylvania: Hanley & Belfus, Inc.; 2001. p. 662-666.

- [11] Rector AL. Clinical Terminology: Why is it so hard? *Methods of Information in Medicine* 1999;38:239-252.
- [12] Cimino J, Patel V, Kushniruk A. Studying the human-computer-terminology interface. *Journal of the American Medical Informatics Association* 2001;8(6):163-173.
- [13] Poon A, Fagan L, Shortliffe E. The Pen-Ivory Project: Exploring user-interface design for the selection of items from large controlled vocabularies of medicine. *Journal of the American Medical Informatics Association* 1996;3(2):168-183.
- [14] Rector A, Rossi-Mori A, Consorti F, Zanstra P. Practical development of re-useable terminologies: GALEN-IN-USE and the GALEN Organisation. *International Journal of Medical Informatics* 1998;48:71-84.
- [15] Spackman K, Campbell K, Cote R. SNOMED RT: A Reference Terminology for Health Care. In: Masys D, editor. *Proceedings of American Medical Informatics Association Annual Symposium*; Philadelphia, Pennsylvania: Hanley & Belfus, Inc.; 1997. p. 640-644.
- [16] Hardiker N. *Applying the GALEN approach to documenting nursing care [MSc]*. Manchester: University of Manchester; 1995.
- [17] Heathfield H, Hardiker N, Kirby J, Tallis R, Gonsalkarale M. The PEN & PAD Medical Record Model: Development of a Nursing Record for Hospital-based Care of the Elderly. *Methods of Information in Medicine* 1994;33:464-72.
- [18] Hardiker N, Roberts A. *Compositional Data Entry Tool (CompET)*. Copenhagen: Dansk Institut for Sundheds og Sygeplejeforskning; 2000. Report No.: TELENURSE ID-ENTITY Deliverable D.07.09.
- [19] Wright P, Jansen C, Wyatt J. How to limit errors in interpretation of data. *The Lancet* 1998;352:1539-1543.
- [20] Wilcox A, Narus S, Bowes III W. Using natural language processing to analyze data entry templates. In: Kohane I, editor. *Proceedings of American Medical Informatics Association Annual Symposium*; Philadelphia, Pennsylvania: Hanley & Belfus, Inc.; 2002. p. 899-903.
- [21] Brown S, Hardenbrook S, Herrick L, St Onge J, Elkin P. Usability evaluation of the progress note construction set. In: Bakken S, editor. *Proceedings of American Medical Informatics Association Annual Symposium*; Philadelphia, Pennsylvania: Hanley & Belfus, Inc.; 2001. p. 76-80.
- [22] Dolin R, Alschuler L, Beebe C, Biron P, Boyer S, Essen D, Kimber E, Lincoln T, Mattison J. The HL7 Clinical Document Architecture. *Journal of the American Medical Informatics Association* 2001;8(6):552-569.
- [23] http://www.openehr.org/downloads/archetypes/templates_and_archetypes.pdf. Accessed: 9/12/2003.
- [24] Schultz J. A history of the Promis technology: an effective human interface. In: White J, Anderson K, editors. *Proceedings of The ACM Conference on The history of personal workstations*; New York, NY: ACM Press; 1996.
- [25] <http://www.wikipedia.org/wiki/PROMIS>. Accessed: 9/12/2003.

Address for correspondence

Nicholas Hardiker
Room PO 42, Brian Blatchford Building
University of Salford
Salford
Greater Manchester

Hardiker N, Bakken S. *Requirements of tools and techniques to support the entry of structured nursing data*. In: Fieschi M, Coiera E, Yu-Chan JL (eds). MEDINFO 2004: Proceedings of the 11th World Congress on Medical Informatics 2004. Amsterdam: IOS Press, 621-625.

M6 6PU
UK.

Phone: +44 (0)161 295 7013
Email: n.r.hardiker@salford.ac.uk