MEDINFO 2007
HISA Ltd Australia
© 2007 The authors. All rights reserved.

Updating a Computerised Clinical Observation Entry Form According to the Pattern of its Previous use and to the Evolution of Medical Knowledge

Olivier Steichen[a], Christel Daniel-Lebozec[a,b], Marie-Christine Jaulent[a], Jean Charlet[a,c]

[a]INSERM, U872, eq. 20, Paris, F-75006 France; Université Pierre et Marie Curie–Paris6, UMR_S 872, eq. 20, Paris, F-75005 France; Université Paris Descartes, UMR_S 872, eq. 20, Paris, F-75006 France;
[b]AP-HP, Hôpital Européen Georges Pompidou, DIH, Paris, F-75015 France;[c]AP-HP, DSI/STIM, Paris, F-75014 France

Abstract

In order to update a semi-structured medical observation entry form, we studied previously completed forms, revealing physicians' actual needs in clinical settings, and guidelines, reflecting the evolution of medical knowledge. Statistical analysis of completed forms pointed to items almost never used by clinicians and some of them were removed from the form. Terminological analysis of free text answers in previously completed forms and of guidelines revealed concepts relevant for clinical practice. Accordingly, some new items were added in the structured part of the form.

Keywords:

computerised medical records systems,
medical history taking, user-computer interface,
natural language processing

Introduction

Computerised and structured data entry forms can improve the reliability, availability and reuse of clinical data for patient care (including decision support), medical audit, activity coding and clinical research [1,2]. Once an entry form has been modelled, maintenance issues arise with the evolution of medical knowledge and practice.

Natural language processing (NLP) of free-text records has already been used to design or improve clinical data entry forms [3,4]. We present an extended approach, relying on statistical analysis of previously completed forms and terminological analysis of free text answers and clinical guidelines. This procedure has been applied to a computerized clinical record form used since 1975 for hypertension management [5,6].

Material and methods

We analysed 5109 completed entry forms from a single hypertension unit (year 2005), including 176 optional questions, with either structured (Boolean, checklist, numeric) or free-text answers. When a question or a predefined answer was used less than 50 times, its deletion was considered. This threshold, approximately 1% of the analysed forms, was arbitrary chosen as a compromise between signal (identification of items seldom used that should be deleted because they are indeed useless) and noise (identification of items seldom used that should not be deleted because they are nonetheless useful).

In order to find recurrent clinical concepts in free text answers, we used the NLP tools SYNTEX and UPÉRY to perform their terminological analysis [7]. When a concept appeared more than 50 times in the corpus, we discussed the inclusion of a corresponding item in the form. For example, the concept of sleep apnoea syndrome appeared about 80 times in the corpus. It was subsequently added as a predefined answer to the pre-existing checklist question "Respiratory disorders?" in the form. We also separately analysed eight hypertension guidelines. Each new clinical concept found in guidelines was considered for inclusion in the form, regardless of its occurrence rate.

A staff of four senior clinicians took the final decisions to remove old items or add new ones. Besides the statistical and terminological results, they took other factors into account, like the value of the data for medical decision making or clinical research. The physicians explicitly favoured objective questions with reproducible answers. However, in order to restrain the drive toward excessive structuring, they also considered the burden of data entry during the patient visit and the drawbacks of an inflexible, excessively structured, clinical record [8].

Results

Among the 176 questions, 19 were completed less than 50 times and 16 of them were deleted. Among the 203 predefined answers to the 39 checklist questions, 78 were ticked less than 50 times and 47 of them were deleted. Physicians judged the remaining items too important to be deleted.

Terminological analysis of free text answers (respectively guidelines) yielded 103 (respectively 95) concepts not found in the structured part of the entry form. Thirty-two of these concepts were common to both sources. Accord-

ingly, items were added in the form: eight numeric and four free-text questions; six checklist questions and 28 associated answers; and four new answers for pre-existing checklist questions.

The conclusion of the visit was previously recorded in a single free-text question. Some of the new questions contributed to structure a more standardised and informative closing. Four checklist questions on hypertension aetiology, pattern, severity and complications now allow stating the characteristics of the patient's disease. They are followed by a question on the specific blood pressure goals and two checklist questions about the prescribed lifestyle changes and treatment modifications. Additional closing thoughts and decisions can still be recorded in the free-text question.

Discussion

The presented approach allowed us to update a long-standing medical record form. Statistical analysis guarantees that no frequently used item is removed during the process. Terminological analysis of free-text answers and guidelines, critically appraised by senior clinicians, guarantees the pertinence of the new items for patient care, audit and/or clinical research.

Terminological analysis of free-text answers is always possible but will indicate recurrent concepts only if clinical cases exhibit some homogeneity. The approach relies on guidelines to reflect the state of current medical knowledge, but could also use less formal review articles. However, guidelines and review articles only pertain to well circumscribed clinical domains. In all these regards, the context of hypertension management was ideal, since it is limited, highly repetitive and well standardised. To which extent and with which results our approach can be generalized to other clinical settings remains to be investigated, especially for broad specialities like internal or emergency medicine.

Another shortcoming of our approach is that we only involved senior clinicians with an extensive experience of the data entry form. Inputs from naive users would have been helpful to identify ambiguous or vague items and eventually improve their content or phrasing.

Conclusion

Our approach allowed us to revise content and organisation of a medical record form, relying on the analysis of its previous use (reflecting clinicians' actual needs in practice) and on the analysis of recent clinical guidelines (reflecting the evolution of medical knowledge). It should be possible to reproduce this approach in other clinical domains, if they are specialised enough to be covered by guidelines and for free text answers of entry forms to show some recurrent patterns.

Acknowledgements

Didier Bourigault provided the results of terminological analyses with SYNTEX and UPÉRY. This work benefits from a grant of the AP-HP and from a subsidy of the French Society for Hypertension.

References

[1] Powsner SM, Wyatt JC, Wright P. Opportunities for and challenges of computerisation. Lancet 1998;352:1617-1622.
[2] van Ginneken AM. The computerized patient record: balancing effort and benefit. Int J Med Inform 2002; 65:97-119.
[3] Kreis C, Gorman P. Word frequency analysis of dictated clinical data: a user-centered approach to the design of a structured data entry interface. Proc AMIA Annu Fall Symp 1997:724-8.
[4] Wilcox AB, Narus SP, Bowes WA 3rd. Using natural language processing to analyze physician modifications to data entry templates. Proc AMIA Symp 2002:899-903.
[5] Degoulet P, Menard J, Berger C, Plouin PF, Devries C, Hirel JC. Hypertension management: the computer as a participant. Am J Med 1980;68:559-67.
[6] Degoulet P, Chatellier G, Devries C, Lavril M, Menard J. Computer-assisted techniques for evaluation and treatment of hypertensive patients. Am J Hypertens 1990;3:156-63.
[7] Charlet J, Bachimont B, Jaulent MC. Building medical ontologies by terminology extraction from texts: an experiment for the intensive care units. Comput Biol Med 2006;36:857-870.
[8] Ventres W, Kooienga S, Marlin R. EHRs in the exam room: tips on patient-centered care. Fam Pract Manag 2006;13:45-7.

Address for correspondence

Olivier Steichen
INSERM UMR_S872 – SPIM,
15 rue de l'Ecole de Médecine,
F-75006 Paris, France
E-mail: ost@club-internet.fr