

Evaluation of two French SNOMED indexing systems with a parallel corpus

Suzanne Pereira^{a,b,c}, Ph.D, Philippe Massari^a, MD, Antoine Buemi^d, MD, Badisse Dahamna^a, Elisabeth Serrot^c, Michel Joubert^b, PhD, Stéfan J. Darmoni^a, MD-PhD.
^a CISMef, LITIS EA 4108, Institute of Biomedical Research, University of Rouen, France
^b LERTIM, Marseille Medical University, France
^c Vidal, Issy les Moulineaux, France
^d APHP Hospital, Paris, France
suzanne.pereira@chu-rouen.fr

Abstract

Background: In this paper, we developed F-MTI (French Multi-Terminologies Indexer) which is a generic automatic indexing tool able to index documentation in several health terminologies such as SNOMED 3.5 (Internationale Systematized Nomenclature of human and veterinary MEDicine). *Objective:* We compared F-MTI and Snocode (a Canadian commercial tool) on a corpus of 100 discharge summaries. *Results:* The results showed that Snocode and F-MTI indexing are as close as two manual indexing can be. They also provided close results in terms of diagnosis.

Keywords :

Abstracting and Indexing/methods; Systematised Nomenclature of Medicine; medical records; international classification of diseases

INTRODUCTION

France chose in 2006 the SNOMED 3.5, the most exhaustive medical terminology in French, for medical record indexing. SNOMED 3.5 is included at 91% in the SNOMED CT (Clinical Terms). It contains 150,000 concepts distributed among eleven axes. The huge number of codes and the complexity of this terminology accounts for the reluctance of the physicians to index diseases in medical records. A computer support for this time-consuming procedure is then urgently required. We developed F-MTI [1] (French Multi-Terminologies Indexer) that generates a document indexing in all the implemented terminologies (MeSH (Medical Subject Heading) , SNOMED 3.5, ICD10 (Classification of Diseases) and CCAM (French CPT). Then all the terminologies are projected in the terminology(ies) desired by the user with the help of the mappings (most of them are coming from the UMLS). The goal of this study is to compare the SNOMED indexing of F-MTI and Snocode3 (a Canadian commercial tool [2] used in several hospitals in England, Canada and France).

MATERIALS AND METHODS

A corpus of 100 patient discharge summaries manually indexed in SNOMED 3.5 is difficult to obtain due to the complexity of this terminology. Faced with these facts, we projected the SNOMED codes to ICD10 codes that can be manually indexed and which enables to compare the two tools in terms of diagnosis. The projection process of SNOMED into ICD10 was performed by the same mapping. The ICD10 manual Indexing of these documents was taken as the reference. First, the two sets of SNOMED codes performed by F-MTI and Snocode were compared without any reference with simple measures. Then the two sets of ICD10 codes resulting from the projection of the SNOMED codes into ICD10 codes, were compared using an ICD10 manual indexing reference.

RESULTS & DISCUSSION

The results showed a Hooper's measure of 32.9 comparing the two sets of SNOMED codes. With the help of a SNOMED-ICD10 mapping we could compare in terms of diagnosis these two sets with a manual ICD10 indexing. We obtained a precision of 6.1 for Snocode and 4.4 for F-MTI and respectively a recall 24.7 and 27.0. Snocode and F-MTI indexing are as close as two manual indexing can be. They also provided close results in terms of diagnosis.

CONCLUSION

This is encouraging for our project. With some improvements we hope that FMTI will integrate a French electronic patient record system.

References

1. Pereira S., Névéol A., Massari P., Joubert M. and Darmoni S.J. Construction of a semi-automated ICD-10 coding help system to optimize medical and economic coding, MIE2006:845-850
2. SNOCODE® MedSight Informatique Inc. <http://www.medsight-info.com>