

一般口演

一般口演10

看護情報1（看護記録のICT化とデータ活用）

2018年11月24日(土) 09:00～11:00 C会場 (4F 411+412)

[3-C-1-6] 看護実践用語標準マスターを基にした形態素解析用辞書を用いた看護記録の分析に関する検討

○青木 美和¹, 横田 慎一郎², 篠原 恵美子², 今井 健¹, 大江 和彦¹（1.東京大学大学院医学系研究科, 2.東京大学医学部附属病院）

目的: 本研究の目的は、看護実践用語標準マスター（以下、看護標準マスター）の活用および普及に向けた将来的な用語追加や機能拡張の提案のための探索的調査に用いるための、看護標準マスターを基にした形態素解析用辞書構築に関する検討を行うことである。さらに、看護標準マスターに含まれる用語が看護記録でどの程度使用されているか確認し、異なる記録様式の看護記録で比較することを目的としている。方法: 本研究で対象とする看護記録は、東京大学医学部附属病院（以下、当院）における電子カルテ上の診療日が2014年2月1日～2014年2月7日である、電子カルテに保存された全診療科入院患者の、看護師が記載した経過記録とした。経過記録データ抽出には、当院で独自に構築している SS-MIX2拡張ストレージ検索システムを用いた。本分析では、看護標準マスターに含まれる用語をユーザー辞書として使用した。看護標準マスターを基にした形態素解析用辞書に収載された語が経過記録に記載されている語の中に占める割合（以下、カバー率）を算出し、退院看護サマリーにおけるカバー率との比較を行った。結果: 抽出された経過記録は17109件、患者数は1588人(内、男性836人)、平均単語数は77.7(標準偏差77.2)だった。診療科別にみると、平均単語数は、救急科が最も多く126.8、続いて眼科115.3、老年病科101.3であった。経過記録における全体のカバー率は、24.0%であった。退院看護サマリーにおける全体のカバー率は16.4%であり、経過記録における全体のカバー率のほうが高かった。診療科別にみると、カバー率は、救急科が最も高く27.9%、続いて女性科26.8%、血管外科26.6%であった。考察: 今後、看護標準マスターを基にした形態素解析用辞書の表記揺れの対応およびパラメータの調整、調査対象のデータ抽出期間の検討や経過記録に存在するが看護標準マスターに含まれなかった語の分析を進めていく予定である。

看護実践用語標準マスターを基にした形態素解析用辞書を用いた 看護記録の分析に関する検討

青木美和^{*1} 横田慎一郎^{*2} 篠原恵美子^{*2} 今井健^{*1} 大江和彦^{*1}

^{*1} 東京大学大学院医学系研究科 ^{*2} 東京大学医学部附属病院

Analysis of Nursing Records Using Morphological Dictionary Based on the Standard Terminology for Nursing Observation and Action.

Miwa Aoki^{*1} Shinichiro Yokota^{*2} Shinohara Emiko^{*2} Imai Takeshi^{*1} Ohe Kazuhiko^{*1}

^{*1} Department of Healthcare Information Systems, School of Public Health, Graduate School of Medicine, The University of Tokyo, ^{*2} Department of Healthcare Information Management, The University of Tokyo Hospital

The standardization of data is imperative to reuse data stored in electronic health records. In Japan, however, considerable information is included in unstandardized and unstructured data, such as nursing records in Japan. The purpose of this retrospective study is to develop a morphological dictionary based on the Standard Terminology for Nursing Observation and Action for exploratory investigation as well as confirm how many terms included in the Standard Terminology for Nursing Observation and Action are used in nursing records and compare among them with different types of records. We used 17,109 nursing records in the University of Tokyo Hospital between February 1, 2014, to February 7, 2014. We evaluated the proportion of used standard terms in nursing records using natural language processing technique and compared the proportion among daily nursing records and discharged nursing summaries. We found that the dictionary contained 24% of all nouns included in the nursing records. In addition, the proportion in daily nursing records was higher than that in the discharged nursing summaries. Our overarching objective is to provide some suitable tools to recommend adding new terms and extend to utilize and disseminate to the terminology.

Keywords: Nursing Records, Natural Language Processing, Terminology

1. 緒論

近年、蓄積された医療ビッグデータの利活用の推進が検討されている[1]。看護学分野において、現状の看護の可視化や継続的なケアの質評価のため蓄積されたデータの利活用は肝要である。データを大規模に利活用するためには、データの標準化が重要である。2016年には看護業務における電子的記録に用いる用語集、看護実践用語標準マスター（以下、看護標準マスター）が厚生労働省標準規格として認められ[2]、普及推進の議論がなされている。

2. 目的

本研究の目的は、看護実践用語標準マスターの活用および普及に向けた将来的な用語追加や機能拡張の提案のための探索的調査に用いるための、看護標準マスターを基にした形態素解析用辞書構築に関する検討を行うことである。さらに、看護標準マスターに含まれる用語が看護記録でどの程度使用されているか確認し、異なる記録様式の看護記録で比較することを目的としている。なお、Comejisyo[3]と競合する形態素解析用辞書そのものを開発することは今回の目的ではなく、看護標準マスターの現行の運用について言及するものではない。

3. 方法

3.1 対象看護記録

本研究で対象とする看護記録は、東京大学医学部附属病院（以下、当院）における電子カルテ上の診療日が2014年2月1日～2014年2月7日である、電子カルテに保存された全診療科入院患者の、看護師が記載した経過記録とした。経過記録データ抽出には、当院で独自に構築している SS-

MIX2 拡張ストレージ検索システムを用いた。当院では、電子カルテの非標準化データを蓄積するためのシステムであるSSMIX2 拡張ストレージ[4]には、各種診療記録や各種レポート類をXMLファイルとして格納している。また患者の診療科は別途、参照系電子カルテシステムのデータベースから抽出した。分析環境にプログラム実行結果を逐一確認できるJupyter Notebook、プログラム言語にpython3.6.2を用いて分析を行った。本研究は、東京大学大学院医学系研究科倫理審査委員会の承認を得て実施された（審査番号11845）。

3.2 看護標準マスターを基にした形態素解析用辞書

形態素とは、意味を持つ最小の言語単位である。形態素解析とは与えられた文を形態素単位に区切り、各形態素に品詞などの情報を付与する処理である。本分析において、言語処理のための公開されている形態素解析エンジン MeCab[5]を用いた。MeCabによる形態素解析には辞書が必要となり、システム辞書とユーザー辞書を設定できる。本研究では、MeCab開発者推奨のフリーで公開されているIPA辞書をシステム辞書として使用した。看護標準マスターを基にした形態素解析用辞書に収載された語が記録に記載された語の中に占める割合（以下、カバー率）の算出を行うため、先行研究[6]で使用したユーザ辞書として用いた。

3.3 分析・評価

記録1件あたりの文字数、単語数を算出した。単語数は、MeCabで切り出された重複を含む全形態素数を指す。診療科別記録件数、記録1件あたりの単語数を算出した。またカバー率として、（抽出された看護標準マスター登録語数）/（総

名詞数)を算出した。小数点以下第2位を四捨五入した。

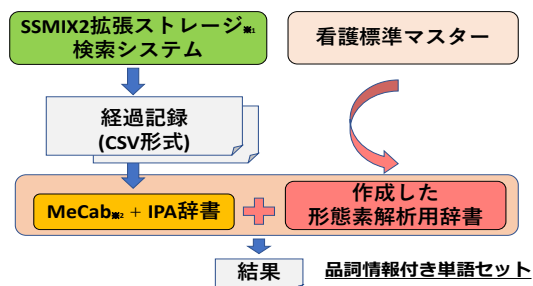


図1. 全体の流れ

4. 結果

4.1 対象データの基本属性

抽出された経過記録は171115件、患者数は1588人(内、男性836人)、平均単語数は77.7(標準偏差77.2)だった。単語数0である経過記録6件を除外した、17109件を分析に用いた。さらに、カバー率の算出では、経過記録中の名詞数が0である7件のデータを除外し17102件を用いた。経過記録1件あたりの平均文字数は155.4(標準偏差(以下、SD)147.0)、平均単語数は77.7(SD77.2)だった。経過記録1件あたりの単語数の最小値は0、最大値は1032であった。診療科別にみると、件数は消化器内科が1373件と最も多く、続いて血液内科1052件、整形外科1051件であった。平均単語数は、救急科が最も多く126.8、続いて眼科115.3、老年病科101.3であった。

4.2 カバー率

全データにたいする経過記録における全体のカバー率は、24.0%であった。診療科別にみると、カバー率は、救急科が最も高く27.9%、続いて女性科26.8%、血管外科26.6%であった。診療科別平均カバー率の上位10科は図2の通りであった。また、記録1件あたりのカバー率と単語数の散布図は図3の通りであった。

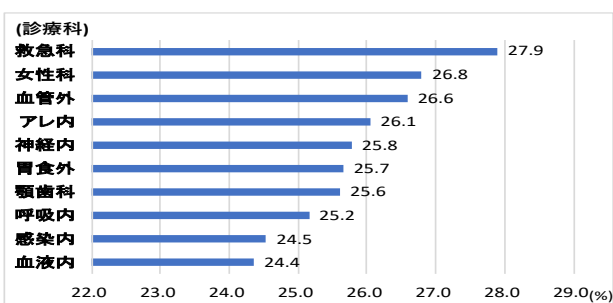


図2. 診療科別平均カバー率上位10科

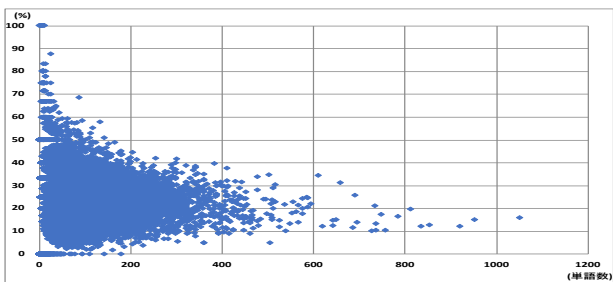


図3. 記録1件あたりのカバー率と単語数

経過記録における全体のカバー率は24.0%、退院看護サマリーにおける全体のカバー率は16.4%であり、経過記録におけるカバー率のほうが高かった。

5. 考察

本研究では、看護師が記載した経過記録を対象看護記録とし、看護標準マスターに含まれる語が臨床の看護記録記載においてどの程度使用されているかを確認した。先行研究[6]より、退院看護サマリーの分析では、日常的に経過記録に記載するような結果、すなわち看護標準マスターで書くような内容がそもそも記録されない可能性を指摘した。今回、経過記録に記載されている可能性を鑑み、分析した結果、カバー率は経過記録のほうが高く、看護標準マスターの用語は退院看護サマリーと比較して経過記録でより使用されていることが明らかになった。しかし、登録した辞書の表層形では記載されていない可能性がある。つまり、概念的には同様のことが書かれているが、表現形の違いがあると考えられる。今後の課題として、今回使用したユーザ辞書のパラメータの調整、表記のゆれへの対応を検討していく必要がある。

6. 結論

本研究では、看護標準マスターに基づいた形態素解析用辞書を用いて看護師が記載した経過記録を分析し、退院看護サマリーのカバー率と比較した。その結果、経過記録における全体のカバー率は24.0%であり、退院看護サマリーにおける全体のカバー率16.4%よりも高かった。

今後、看護標準マスターを基にした形態素解析用辞書の表記揺れの対応およびパラメータの調整、調査対象のデータ抽出期間の検討や経過記録に存在する看護標準マスターに含まれなかった語の分析を進めていく予定である。

利益相反: 開示すべき利益相反はない。

参考文献

- 厚生労働省. 保健医療分野におけるICTの活用. http://www.mhlw.go.jp/file/05-Shingikai-12601000-Seisakutoukatsukan-Sanji_kanshitsu_Shakaihoshoutantou/0000140306.pdf (最終アクセス: 2018年8月24日)
- 一般財団法人医療情報システム開発センター(MEDIS-DC). 医療情報システム開発センターの提供する標準マスターの概要と使い方, 第16版, 平成29年7月.
- Comejisy: <https://ja.osdn.net/projects/comedic/>. (最終アクセス: 2018年8月24日)
- 日本医療情報学会. SS-MIX2 拡張ストレージ 構成の説明と構築のガイドライン Ver.1.2d, 2017. http://www.jami.jp/jami_std/docs/SS-MIX2/SS-MIX2-V1.2d-1/SS-MIX2_Ver.1.2d_released_20170116/SS-MIX2_extstrglVer.1.2d_20170116.pdf. (最終アクセス: 2018年8月24日)
- 工藤拓. MeCab: Yet Another Part-of-Speech and Morphological Analyzer. <http://taku910.github.io/mecab/> (最終アクセス: 2018年8月24日)
- 青木美和, 横田慎一郎, 篠原恵美子, 今井健, 大江和彦. 看護実践用語標準マスターを基にした形態素解析用辞書の構築に関する検討. 第19回日本医療情報学会看護学術大会論文集 2018;7:127-130.