

## 医薬品添付文書情報の解析とその応用

○富樫 秀夫 1) 栗原 勝 1) 折井 孝男 2) 鎌田 志乃ぶ 2)

アイティコーディネート株式会社 医用コンテンツ開発部 1)

NTT東日本関東病院 薬剤部 2)

Analysis for contents data of package leaflet in medicinal supplies and How to make use of them practically

Hideo Togashi1) Masaru Kuribara1) Takao Orii2) Shinobu Kamata2)

IT Coordinate,Inc. Tokyo,Japan1)

Kanto Medical Center NTT EC Tokyo,Japan2)

Abstract: Although many convenient information data is in a medical-supplies appending document, it is not separable as a keyword. Furthermore, those terms are not coded in the computer system. We extracted the medicinal keyword from appending document data using the morphological-analysis technique, and classified the word. These expect contributing beneficially in a hospital information system.

Keywords: drug information, context, disorder, contraindication

---

### 1. はじめに

医療用医薬品添付文書（以下、添付文書）の情報をデータ化し、医療情報システム等で利用することの重要性は認識されているが、添付文書情報のデータ化自体は遅れている。

筆者らは、形態素解析手法による添付文書情報のコンテキスト解析とキーワード抽出を行った。

### 2. 添付文書情報の現状と問題点

#### 2.1 医療用語の標準化と添付文書情報

医療情報のデジタル化が進んでいるが、添付文書情報については、いわゆる「平文」のデータのみが存在している。

○キーワード情報の欠落： 添付文書に含まれる、適応病態、適用・禁忌部位、前提診療行為等のキーワードは抽出されていない。

○標準病名と添付文書中の病態表現： 診断病名等、医療機関内外のシステム上で情報交換するための用語は「標準病名マスター」として、よく整備されている。しかし、添付文書中の効能・効果として記述された病態表現は、「標準病名」とは異なり、直接的なコード変換は行えない場合が多く、病名交換コードを安全性チェック等の共通コードとして使用することは、直接的には出来ない。

#### 例) 標準病名マスター

急性胃炎 病名交換コード RJHE ICD-10コード K291

慢性胃炎 病名交換コード ESKE ICD-10コード K295

#### 例) 添付文章中の表現

## 急・慢性胃炎

診療録等に記述する患者病態表現としては、急性期・慢性期のいずれかの胃炎であろうが、添付文書上では、いずれの胃炎にも有効な医薬品の場合「急・慢性胃炎」と記載される。

### 2.2 添付文書情報の低い利便性

添付文書中のキーワードが抽出されておらず、そのコード体系も制定されていないため、システム利用上の利便性は低い。

#### 2.2.1 一般的なキーワード検索の問題点

「平文」であっても、システムによる文字列検索は可能である。しかし、生殖器疾患について検索する目的で「性」のキーワードで包含一致検索した場合、「平文」中では「てんかん性格」「のう胞性線維症」等、極めて多数の検索用語文字列が該当し、「男子性腺機能不全症」に至るまでに多大の労力を要することになる。

更に、検索によって検出した病名が、効能・効果として有効である保証はない。

例) 否定された用語の扱い： 「躁うつ病」を検索した場合、「統合失調症及び躁うつ病に伴う不眠症を除く」と云う文章においても該当文字列はヒットするが、「統合失調症」「躁うつ病」は、効能・効果として有効ではない。

#### 2.2.2 薬剤投与における安全性チェック

電子カルテ等における薬物療法オーダー機能として、適応病態・禁忌病態に対するチェックは重要である。また、局所麻酔剤を手指・陰茎に投与してはならないケースも多く、「身体部位」に対するチェックも重要である。しかし、これまで適応・禁忌病態の部位文字列を添付文書より実用的に抽出する試みはなされなかった。

以上のように、電子カルテと付随する安全性チェック機能の開発・普及は進んでいるが、システムで利用可能なコンテンツの開発・標準化は遅れている。

### 2.3 日本語特有のコンテキスト解析の困難さ

日本語では、欧米諸言語と異なり、各単語間の区切りが無く、文章が連続文字がして記述されており、システムによりキーワードとなる文字列を抽出することは技術的に難易度が高い。

例) 各単語間の区切りの例

に伴う歯科疾患・・・ 「伴う」と「歯科」と云う用言と体言

によるう歯発生・・・ 「による」と「う歯」と云う用言と体言

更に、同一文字・文字列が複数の意味に使用されることも多い。

内性器感染症・・・ 「性」は身体部位根幹表現

うっ血性心不全・・・ 「性」は接頭語の語尾

### 3. 文書情報の解析手法と結果

文書情報の解析は、文章からキーワードとなりうる名詞、及び名詞句の文字列を抽出し、次にキーワードとしての採用の可否を決定する手順で行った。

○文書情報の解析( コンテキスト解析とキーワード抽出 ): 文章に現れる文字列の抽出は、一般的な形態素解析手法を用いた。

但し、名詞・名詞句が、下記の手法により、病態表現・身体部位表現・診療行為のいずれかと判断出来る場合は、主語・目的語・補語・動詞の関係解析において、意味論的品詞要素決定を行っている。

また、解析対象文字列が用言であっても、医学的意味を持つ文字列を含む場合、抽出の対象とした。

○キーワード文字列の解析( コンテキスト解析 ): キーワードのカテゴリズについては、個々のアトム・フラグメントの分離を軸管理手法を用いて、前後関係よりキーワード自体の分類区分を決定した。

キーワード：例)「カフェイン禁断性頭痛」 医学的意味を持つ独立した単語として見なせるフラグメントの集合体

フラグメント：例)「カフェイン禁断性，頭痛」1個以上のアトムからなり、何らかの医学的意味を持つ文字列

アトム：例)「カフェイン，禁断，性，頭，痛」それ以上分解すると本来の意味を失う1文字以上の文字列

○解析結果の検証： 形態素解析により、キーワードと判断された文字列( 名詞・名詞句等の体言、医学的意味を持つ用言 ) は、必ずしも正しい医療用語とは限らない。よって、新規に出現したキーワードにおいて、その構成要素としてのフラグメント・アトムが既知のものではない場合には、複数の医療従事者の判断により、カテゴリズを行った。

○抽出キーワードのコード化： 抽出されたキーワードに対して内部コードを付与し、次にキーワードが標準病名マスターのリードタームと同義であるか、その概念に関連している場合には、病名交換コードとICD-10コードに対するリンク付けを行った。

#### 4. 考察と応用範囲

一般的な形態素解析手法のみを採用した場合、解析結果の検証に負担がかかる。しかし、文書情報のコンテキスト解析により抽出したキーワード候補に対し、キーワード文字列自体のコンテキスト解析を行い、繰り返し文書情報のコンテキスト解析を行うことにより良好な結果が得られることが判明した。

##### 4.1 医療情報システムでの利用

○病名・症状チェック： 処方・注射時の安全性チェックへの応用として、1号用紙の病名又は2号用紙の所見に存在する疾患・症状を基に、投与する医薬品が適応症として有効であるか、あるいは禁忌対象病態であるか等の判定に有用と考えられる。

○投与目的検索： 治療対象の疾患名の他、「消炎・鎮痛」等、薬理作用キーワードを入力することにより、該当する医薬品を検索するシステムにも利用可能と考えられる。

○投与部位チェック： 投与禁忌となる部位を医薬品の塗布・注射時に表示すること等、診療実施時の支援機能にも利用可能と考えられる。

○他の診療行為との関係チェック： 電子カルテシステムにおいては、検査・画像診断・処置・手術等のオーダー情報が存在するが、医薬品投与時に、他の診療行為との前後関係を勘案して注意・禁止事項等を表示するシステムにも応用可能なデータとなる。

#### 4.2 電子辞書・文書入力補助システムでの利用

○医療情報の電子辞書化への応用： 本研究の手法は、診療録等に含まれる所見・主訴等の自由文からのキーワード抽出にも適用可能であり、データマイニングへの応用も考えられる。

○文書入力支援への応用： 抽出されたキーワードとその意味・分類の解析により、医療用語のシソーラス化が可能となった。そのデータを基に、所見・主訴入力や、医学関係論文の執筆時の支援機能にも応用可能と考えられる。