

厚生科学研究費補助金（医療技術評価総合研究事業研究事業）
総括研究報告書

構造化医療用語の実用化に関する研究に関する研究

主任研究者 里村 洋一（千葉大学医学部附属病院医療情報部教授）

研究要旨

診療録の電子化のためには、電子化された医学用語辞書の存在が不可欠である。電子化辞書は個々の用語の標準化ばかりでなく、用語間の関係や用法の情報が統一されている必要がある。本研究では、市中に存在する多くの医学用語辞書を利用して、これらを統合した形で構造化することを目指した。それぞれの用語をその基本概念に分解して、この基本概念の組み合わせを構造の基本として実用的な医学用語集の構築を行った。また、その実用性を検討した。SNOMEDの繁用単語から5,500の基本概念を抽出し、その日本語訳を付与することによって、再帰的にSNOMEDの翻訳を行い、この用語構造が、具体的な実用辞書の編集に利用できるかを検討した。結果として、なお、基本用語の属性や、用語の合成ルールになお不十分な点があるが、有望な手法であることが示唆された。

分担研究者

木村通夫（浜松医科大学）
廣瀬康行（琉球大学）
山崎俊司（琉球大学）
土屋文人（東京大学）
津本周作（島根医科大学）
熊本一郎（鹿児島大学）
石川光一（国立がんセンター）
日紫喜光良（信州大学）

A. 研究目的

診療録の電子化は、医療の効率化と質の向上に貢献し、我が国の医療を改革する上に強力な武器となることが予測されている。しかし、その実現のためには、さまざまな環境の整備が必要とされる。そのもっとも重要なものとして、医学用語の標準化が挙げられる。この研究は電子化した診療記録

を医療機関関係者間で共有するためのデータの標準化を目標として、用語集の整備の手法を開発しようとするものである。申請者等の一部は、これまでの数年間、内外の主要な医学用語集を解析し、診療記録での使用に好適な電子化医学用語集の構築について研究してきた。その結果、部門別または目的別に編集された用語集から医学的に意味を保持する概念の最小単位（ATOM）を抽出し、これらに品詞、領域（部位・臓器、処置・手術、症状、病態、病因、固有名詞など）、同義語、略語、等の属性を与えることによって、構造化された医学用語集を編集できることを証明した。また、このATOMを集めた辞書を用いてSNOMEDの半自動翻訳を試み、見込みのある成果を見た。この研究成果を基盤として、ATOM辞書の充実を試みた上で、医療の各分野で作成された文書の中で使われている

医学用語を、標準用語に置きかえる実験をし、実用上の問題点を検討するのが今回の研究である。

また、同様に英語の用語集の日本語への自動変換にも挑戦する。このシステムの完成によって、電子カルテの設計者が用語の使用で迷うことがなく、電子カルテの利用者が標準用語を意識することなくデータの入力や検索が行える環境が整う。

B. 研究方法

まず、繁用される基本概念 (ATOM) を構築し、SNOMEDの中に出現する頻度の高い単語 (病名、症状、部位と臓器、医療行為など) に関するATOMの抽出をすすめ、それぞれのATOMに、英語とその品詞、対応する日本語とその品詞、使用領域、及び優先度などの属性を与えた。また、それぞれのATOMを含む用語の文脈を定義して、それぞれの文脈における基本語表現の利用法、すなわち命名法のシステム開発を試みた。また、編集を容易にするための類似語検索システムと用例検索システムの開発を行った。

C. 結果

基本概念約5,500について、その英語の品詞、日本語の品詞、使用領域を定義し、品詞と使用領域に応じて日本語表現の優先度を与えた。また、MEID辞書 (日外アソシエーツ) から指定単語を含む表現を検索するツールとSNOMEDの用語から同様に指定した単語を含む表現のリストを作成するツール、更に、SNOMEDの上位コードから、同じカテゴリーに含まれる用語のリストを表示するツールなどを用意し、翻訳に際して随時参照ができるようにした。

この基本語辞書とツール群を利用して、SNOMED III の翻訳システムを構築し

た。このシステムでは、任意のSNOMED用語を、先ず、基本単語に分解し、ついでそれぞれの単語を基本語辞書に従って翻訳する、更に、英語と日本語の語順の違いを克服するために前置詞や形容詞の配置変換ルールを通して、日本語の語順に変える。全ての単語の日本語訳が単一である場合は、用語のエキスパートによってその適否が判断され、正しいと判断された場合には、そのまま正規の翻訳として採用する。また、一つでも基本語に複数の翻訳がある場合は、これを単語別に併記してエキスパートの選択を待つ仕掛けとした。

このシステムを利用して、SNOMED III の用語 (M、T、F、D、Pの各軸に限定) の翻訳試験を行った。その結果、33,441用語のうち、一意に翻訳できたもの (翻訳候補が1つ) が9,438語、複数の候補が示されたものが18,836語、候補を発見できなかったものが5,167語であった。ただし、これらの翻訳結果が正しいかどうかについての全面的な検証は行っていない。

D. 考察

今年度の研究は、基本語辞書の整備とSNOMED翻訳支援システムの作成、及びそのテストランに終わったが、このシステムが、高い確率で実用的に利用できる可能性を示し得たと考える。語順のルールはまだ不完全であり、また、複数の概念から構成される英単語をそのままATOMとして解釈しているために、翻訳に際して十分な論理性が備わっていないなどの課題が残った。研究の第2年度に当たる平成13年度には、各医学専門領域毎に、このシステムを用いた翻訳を試み、その翻訳結果の正確さ、語順変換ルールの問題点、基本語辞書の過不足、属性の多様化の必要性などについて検討する予定である。

E. 結論

今年度の計画は、翻訳支援システムを作成しテストすることであったが、その目的をほぼ達成したので、平成13年度に予定している研究に移行できる準備が整った。

F. 健康危険情報

本研究では、用語のみを扱い、生体実験や生体から得られた情報を扱っていない。従って、健康に影響を及ぼす様な要素は全くなかった。

G. 研究発表

1. 論文発表

平成13年度中に「医療情報学」と「Methods of Informatics in Medicine」に投稿し、いずれも掲載予定となっている。

2. 学会発表

第21回医療情報学連合大会に発表予定

TransIT - [TransIT]

ファイル 編集 挿入 表示 印刷 ヘルプ

Sequence: 248 SNOMED参照

SNOMEDTerm: HLA SNOMED中出現頻度: 180

品名1: 品名2: 品名3: MED参照

候補List

行挿入 行削除 削除

	日本語候補	仮名	分類	品名	優先度	頻度
1	HLA	HLA				14
2	HLA抗原	HLAコウゲン				5
3	組織適合抗原	ソシキテキゴウケン				4
4	ヒトリンパ球抗原	ヒトリンパキョウコウ				3
5	ヒトリンパ球抗原	ヒトリンパキョウコウ				3
6	ヒトリンパ白血球抗原	ヒトリンパハクケツキ				2
7	ヒト白血球抗原	ヒトハクケツキョウコ				1
8	ひと白血球抗原	ひとハクケツキョウコ				1
9	組織適合性抗原	ソシキテキゴウセイ				1
10	ヒト白血球抗原	ヒトハクケツキョウコ				1

1/1

TransIT - [TransIT]

ファイル 編集 挿入 表示 印刷 ヘルプ

Sequence: 220 SNOMED参照

SNOMEDTerm: TRICHOSTRONGYLUS SNOMED中出現頻度: 20

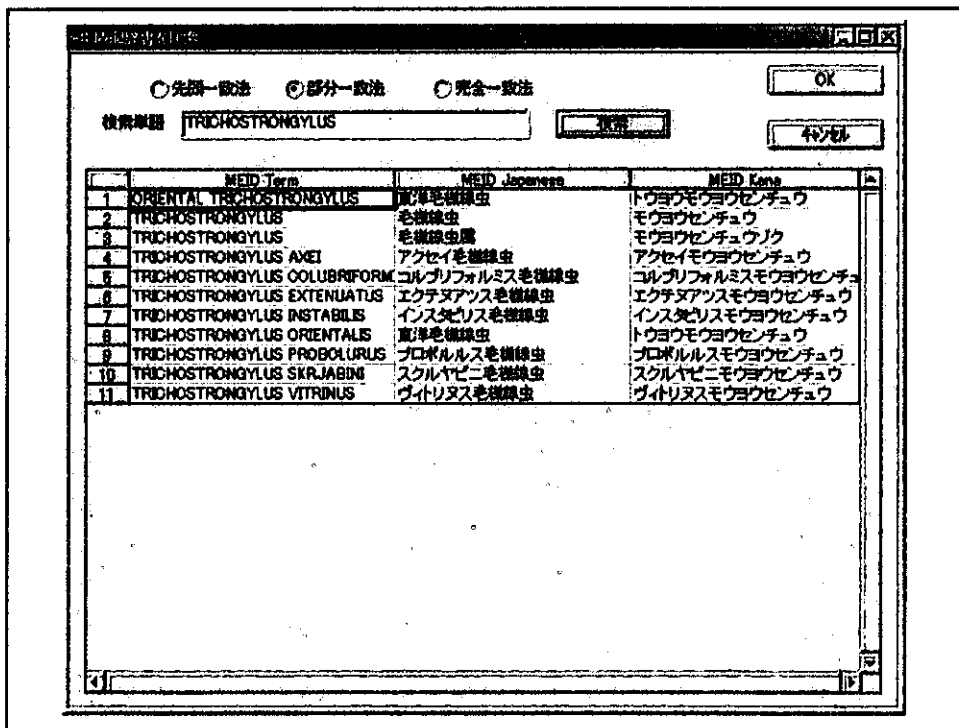
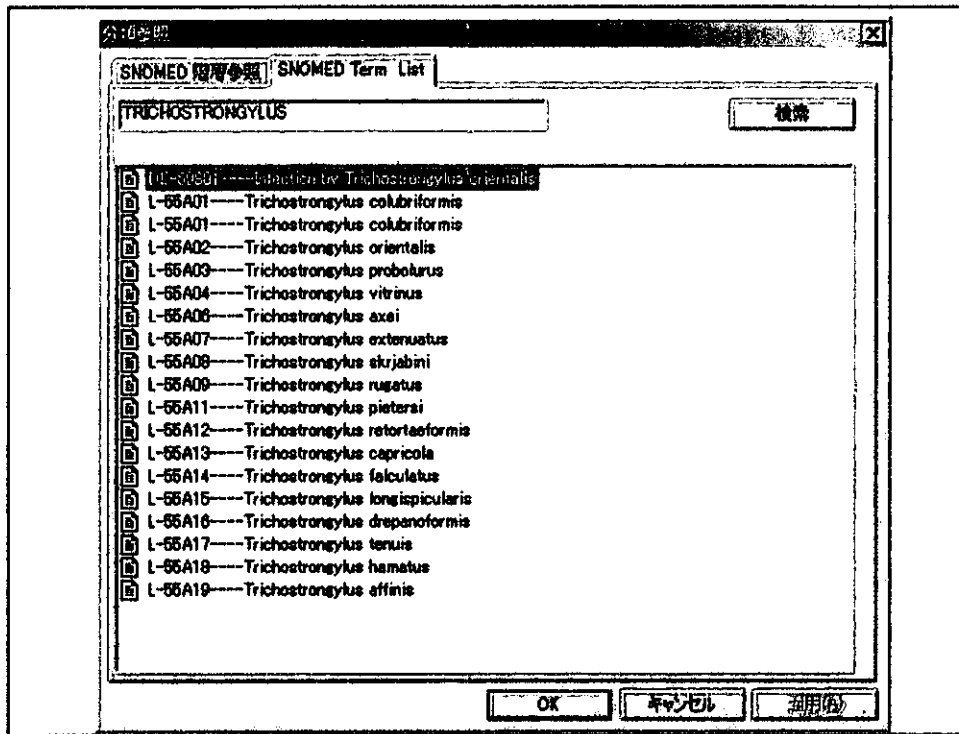
品名1: 1-名詞 品名2: 品名3: MED参照

候補List

行挿入 行削除 削除

	日本語候補	仮名	分類	品名	優先度	頻度
1	毛様線虫	モウヨウセンチュウ	L-生物	1-名詞	1-優先語	11
2	毛様線虫属	モウヨウセンチュウシ	L-生物	1-名詞	1-優先語	1
3						
4						
5						
6						
7						
8						
9						
10						

1/1



についたユニーク番号です
 特別な意味はありません

用語番号を
 前後に移動

翻訳の実行

この用語の翻訳が
 確定したら登録する

SNOMED用語

現時点での翻訳結果
 この行の上で編集ができる

単語毎の翻訳結果
 青色の部分が翻訳
 結果に採用されて
 いる

単語の翻訳を追加し
 たい場合の呼び出し

SNOMEDやMEID辞
 書で同じ単語を含むも
 のや、同じカテゴリの用
 語を検索できる

	A	B	C	D	E	F
1	and	subcutaneous	tissue	of	Disease	
2	皮膚	皮下組織	の	疾患	疾病	病状
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						

語の選択結果
 全体の訳に直
 接反映される

複数の単語の訳語の中か
 ら適切な訳語を選択する

	A	B	C	D	E	F
1	bone	development	or	growth	of	
2	骨	発達	または	成長	の	
3		停止		生長		
4		阻害		発育		
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						

Number: 65394

Object Str: Abnormal nucleic acid sequen

Japanese: 異常核酸配列

Katakana: イシカヌクレオサンハイレン

Nucleic に翻訳がない(英語のまま)なのでこの英語の部分を
クリックして選択し、基本語登録を呼び出す

項目	A	B	C	D	E	F	G	H	I	J
1	Abnormal	nucleic	acid	sequence						
2	異常	核酸		配列						
3	異常の			配列						
4				配列						

この場合はnucleic
という単語の翻
訳をするより、
nucleic acidを登
録した方が良いの
でacidを追加入力
して検索

分類がよくわからない場合には
SNOMEDのリストを参照できる

Term: NUCLEIC ACID

SNOMED 用語集: SNOMED TERM LIST

NUCLEIC ACID

- F-05A00 --- Nucleic acid
- F-05A10 --- Deoxyribonucleic acid
- F-05A10 --- DNA

Number: 65394

Object Str: Abnormal nucleic acid sequence

Japanese: 異常核酸配列

Katakana: イシカヌクレオサンハイレン

Object Term: NUCLEIC

品名: F-形薬名

分類: 核酸

基本語登録

日本語の訳候補

項目	日本語候補	品名	分類	品目	優先度	備考
1	核酸	核酸	核酸			
2	核酸の	核酸の				

仮名は通常自動的に
ふられますが、間違
いがあれば修正して
ください

基本語登録では分類や
品目、優先度なども入力
してください

クリックしてノートを入力

Number: 00334

Object String: Altered nucleic acid sequence

Japanese: 異常核酸塩基配列

Katakana: イソウワカクサンカイレイシツ

基本語入力にあとに、「翻訳ボタン」をクリックすると、この例のように、未翻訳の部分が満たされます。

翻訳が不適当な場合はその行の中で編集してください。単語の訳に関わらず、ここに表示されているのが最終的な翻訳結果となります。

Object Term: NUCLEO

品目: 2-形態性

日本語の訳候補

日本語候補	原語	分類	品目	優先度	確定
1. 核の	カクノ	F-部位、部位など	2-形態性	優先度 0	
2. 核	カク	F-部位、部位など	1-名詞	優先度 0	
3.					
4.					
5.					
6.					
7.					
8.					
9.					

厚生科学研究費補助金（医療技術評価総合研究事業）

分担研究報告書

構造化医療用語の実用化に関する研究

（分担研究項目：臨床検査用語に関する用語の整備とシステム検証）

分担研究者 木村通男（浜松医科大学附属病院医療情報部 教授）

研究要旨

臨床検査項目の中でもとりわけ画像検査項目コードの分野での、用語の標準化を検証した。

A. 研究目的

臨床検査項目の中でも、検体検査分野における用語の整備は十分に考慮をし基盤整備は進められてきた。しかし、画像検査分野では、古くから画像検査機器メーカーの提供する RIS (Radiology Information System) の普及により、最小限の機能が提供されてきたこともあり、HIS (Hospital Information System) と RIS、あるいは画像検査機器間における標準的なデータ交換基盤が整備されてこなかった。画像自体の伝達手段としては、国内でも DICOM 規格がたいへん普及しているが、撮影に伴う情報、すなわち検査予約の管理、および実施済み検査情報については標準化が行われてこなかった。DICOM 規格では、別に MWM (Modality Worklist Management) および MPPS (Modality Performed Procedure Step) を定めており、上記の解決策を提示しているが、日本国内における固有の事情、実際の運用を考慮し、本研究における ATOM への追加概念として検証する。

B. 研究方法

HIS-RIS 相互接続ワーキンググループとしてすでに、保健医療福祉情報システム工業会と、(社) 日本画像医療システム工

業会とから委員の派遣を受け、浜松医大の木村を委員長とした、JJ1017 委員会の場がもうけられ、上記 MWM, MPPS の国内利用について検討が進められていた。この作業の中で、ATOM 辞書としての精度の向上に貢献する用語の検証を浜松医大にて行った。

C. 研究結果

アメリカでは、放射線科医の裁量の範囲が多く、どういった検査をするかを詳細に決めるのは放射線科医であり、依頼医からのオーダーですでにそのような詳細に言及する、ということは少ない、という背景がある。そのため、DICOM 規格 (MWM, MPPS) で定義されている用語が、撮影依頼時に詳細な情報を求める日本での運用には即していなかった。そこで、検査種別コード、検査部位コード、検査方向コード、およびこれらの将来の拡張性の担保について検討の上、「JJ1017 画像検査コード」として定めた。これらのコードは、修正の上、ATOM の構成用語として1つの領域を構成すると判断した。

D. 考察・結論

画像検査に関する実施運用の日米の差や、日本においてのみ法令で求められてい

る照射録のための項目等に注目すると、DICOM 規格のこの部分をそのまま日本で用いることが適当でない、と考えられた。その例としては、日本では必須とされる照射録のための項目が、DICOM 規格では必須にはなっていない、あるいは、日本ではオーダー内容の詳細度が高いので、DICOM で揃えられた項目では十分でない、といった状況が挙げられる。

また一方で、全国共通の、画像検査用詳細コードを作成する試みも過去にはなされてきたが、画像機器の進歩が早く、制定そのものが困難であった。だからといって、これをローカルコードに任せたのでは、折角の規格のメリットが半減するだけでなく、今後は、画像検査部門での各種データ解析が重要性を増すことが予想され、施設間比較などに支障を来す可能性がある。そこで、検査（手技）、部位、方向について、共通語となる基本分類をコードとして作成し、更に詳細なローカルコードが必要な場合に、その基礎となりうるようにした。

E. 研究発表

1. 論文発表 日本医療情報学会誌に投稿した（平成13年3月）。

2. 学会発表 2001年日本医療情報学会シンポジウムにて発表の予定。

厚生科学研究費補助金（医療技術評価総合研究事業）
分担研究報告書

構造化医療用語の実用化に関する研究

分担研究者 廣瀬 康行 琉球大学医学部附属病院 教授

研究要旨 : 診療記録記載用語を標準化・構造化するための前提として、本研究でいうところの ATOM（意味最小単位）や TERM の収集と整理が不可欠である。本研究では特に外科、麻酔科領域における ATOM や TERM の収集と整理を実施して、用語間の意味ネットワークや概念関係の表現方法と、診療録記述の際の用語データベースの実用性等について考察した。

A. 研究目的

研究計画における分担研究者の役割から、外科・麻酔科領域における ATOM 辞書を充実させることが目的である。

B. 研究方法

外科・麻酔科領域から集められた用語を ATOM に分解し、これに適切な属性を定義したのち、ATOM の再分類が必要かどうか検討した。さらに ATOM の同じ領域においても文脈に依存する場合があるかどうかを検討した。

また、このようにして収集登録した用語群が、電子化診療録システムの環境において十二分に活用されるかについて、若干の考察を試みた。

C. 研究結果

C. 1 属性の追加：

前年度までの研究から次のような ATOM 属性が見出されている。すなわち、品詞、領域（部位・臓器、処置・手術、症状、病態、病因、固有名詞など）、同義語、略語である。今回は 優先度を属性として加えた。ただし、この属性をどのように辞書に反映させるかについては議論のあるところである。

C. 2 文脈への依存性：

同一の ATOM もしくは TERM は、常に同等の様態で使用されるとは限らないことが判明した。よって、一つの用語の類語もしくは類似語は、広がりをもつこととなる。

従って用語の表わす概念の確定や用語間の類義性にはある程度の長さの「文脈」が影響すること、

となる。逆に言えば、語の意味概念の確定には共起語同定や文脈検出が必要となる。

D. 考察と残された問題

D. 1 属性の追加：

TERM が複数の概念を表わしうる場合には、概念対用語間の優先度が不明確になるという問題を抱えている。このため、同一概念を表わす複数の用語に対して、どの語が優先されるかを ATOM 辞書側に定義する必要がある。一方、自然文を解析するためには TERM 側に優先度の情報が必要である。これらの問題を解決するには、例えば、出現頻度や共起性などによって優先度属性を決定して辞書外に定義するなどのメタ構造が必要になると思われる。

D. 2 文脈への依存性：

本 ATOM 辞書を電子化診療録において有効活用するためには、特に症状や所見用語においては、類語や類似語という観点もさることながら、まさに共起性等に依る語群形成を支援する機能も必要となろう。

もっともこのような語群形成機能には単に共起性のみならず、学術的に合理的な意味ネットワークの存在が検出できる機能を同時に必要とする。よって用語間意味ネットワークや概念関係のメタ表現方法を整え、構造化体系を構築し、またその編集ツールを設計開発等することが、今後に求められているといえる。

E. 結論

本研究は、本研究計画を実施しその目的を果たし

た。今後は前述の考察を礎として、実用レベルのメタ構造やツールを研究設計そして開発していきたい。

F. 学術発表

1) 医療情報学会誌に発表予定 (2001年10月)

以上

厚生科学研究費補助金（医療技術評価総合研究事業）
分担研究報告書

構造化医療用語の実用化に関する研究

分担研究者 山崎 俊司 琉球大学医学部附属病院 講師

研究要旨： 診療記録記載用語を標準化・構造化するためには、本研究でいうところの ATOM（基本概念）や それを表現する TERM の収集と整理が不可欠である。今年度は前年度開発された ATOM 辞書の内科領域の TERM を中心にその属性を検討した。また、本研究では、文脈による依存性についても若干考察を加えた。

A. 研究目的

研究計画における分担研究者の役割から、今年度は、内科領域における ATOM（基本概念）辞書を充実させること目的とした。

B. 研究方法

内科の様々な領域から集められた用語から、ATOM を抽出して、必要に応じて ATOM 辞書に追加した。ATOM を表現する TERM について新たな属性の定義が必要か否かを検討した。さらに個々の TERM に対応する ATOM が、文脈に依存して変化する可能性があるかどうかを検討した。

C. 研究結果

C. 1 属性：

これまでの研究の結果、品詞、領域（部位・臓器、処置・手術、症状、病態、病因、固有名詞など）、同義語、略語などが ATOM の属性として報告されている。今回はさらにその ATOM の上位に位置する概念および下位に位置する概念を収録すべき、と考えた。さらに品詞、略語等の属性は同義語に定義される ATOM に対応する TERM の属性として定義すべき、と考えられた。

C. 2 文脈依存性：

同一 TERM であっても文脈により、異なった ATOM に関連付けられることがある。例えば検査名では、同じ TERM を「検査項目」として用いたり、「検査結果値」として用いたりする。

ATOM は本来文脈に依存しないが、それに属する TERM から、逆に対応付けされる ATOM は文脈に依存する。

D. 考察

D. 1 追加属性：

電子カルテの利用者が標準用語を意識することなくデータの入力や検索を行うためには、利用者が入力しようとする用語を、システムが予測して候補を提示することが現実的な解である。

このためには、ある TERM から予測される類義語、上下位概念に属する TERM を提示しなければならない。そこで、今回は ATOM の属性の中に上下位概念を付け加えることを提案する。

D. 2 文脈依存性：

ある TERM に対応する ATOM の領域を判断するには、TERM の前後の文章より追加情報を得る必要があると考える。

文脈をどのように端的に表現するかは、今後解決すべき問題である。

自由文で診療録を記載する場合、文脈を確定するための語句や限定辞を、自明のこととして、省略することがほとんどである。電子カルテの入力の際にはこのような省略は避けたいが、入力操作が煩雑になる。入力支援ツールとして本辞書を応用するために、用例検索システムの充実が必要であろう。

E. 結論

本研究は、本研究計画を実施しその目的を果たした。今後は前述の考察を礎として、文脈依存性の解決を計りたい。

F. 学術発表

1) 医療情報学会誌に発表予定（2001年10月）

分担研究報告書（医療技術評価総合研究事業）

処置・手術に関する用語の整理とシステム検証に関する研究

（主任研究者 里村洋一 千葉大学医学部附属病院 教授）

分担研究者 熊本一朗 鹿児島大学医学部附属病院 教授

研究要旨：この研究では、病院情報システムあるいは医療機関関係者間で電子化した診療記録を共有するために、とくに処置・手術に関する用語の標準化を目標としている。診療記録に用いられる用語とくに処置・手術に関する用語を収集し整理し、実際のコンピュータシステムにより診療記録を作成するにあたって、統一性のある診療記録データを効率的に作成できるように、処置・手術の用語集を構造化しインテリジェンスのある辞書に編成することをこころみるものである。処置・手術の用語集を網羅的に収集し、標準的な処置・手術のマスタを作成し、用語の構造の分析を行った。とくに処置・手術の用語では、構造的に部位、臓器、病理学的症状、経路、手技内容、機器、時間要素に大きく分類された。

A 研究目的

医療情報システムの普及により病院施設間のみならず医療機関同士の連携により、医療の効率を高めさらに医療の質を向上をもたらすことが期待されてきている。このためには医療行為自体の情報化とともに、電子化された診療記録の情報の共有化、情報の共通利用性の実現されなくてはならない。情報の交流のための前提としてデータの標準化や通信手法の規格化、セキュリティの確保が重要である。病名の標準化が進められてきているが、さらに医学用語全般、処置・手術の医療行為、臨床検査用語、薬剤に関する用語についても、その構造や用語法について標準化を進め、使いやすい用語処理の方法を利用者に提供する必要がある。

本研究課題は、処置・手術に関する用語を収集し整理し、用語の標準化に必要な要素の抽出と用語の選択、用語の相互関係の構造を分析することである。この成果により、診療記録に記載される情報が統一化され一貫性が得られ、診療情報の電子化を推進するものと期待される。

B 研究方法

本研究では、すでに開発されている標準手術・処置マスタを用いて、用いられている用語を分析しこれらを統合的に利用するための共通的構造とその要素の属性を検討した。この標準手術・

処置マスタは、1) 日本医学会および同分科会の用語集（学会用語集）、2) レセプト電算処理システム用手術・処置マスタ（レセ電算マスタ）、3) 既存の手術・処置マスタ（既存病院情報システムで運用中）の3種類のファイルを統合し、可能な限り網羅性を持たせ基本分類コードにICD・9・CMの手術・処置分類を選定している。

C 研究結果

(1) 標準手術・処置マスタの用語の構造の特徴

電子化した用語集の利用目的としては、とく手術・処置マスタは病院情報システムのサブシステムとしての手術システムで活用されることが多くなってきているが、現在はレセプト請求システムとの連携にて利用されることが多い。そこで上記の標準手術・処置マスタもレセプト電算処理システム用手術・処置マスタ（いわゆる診療報酬請求システムのKコード）を含んでいる。レセプト電算処理システム用手術・処置マスタの用語の構造をみると、なかには複数の処置・手術用語を組み合わせで作成されているものも多い。たとえば「胃悪性腫瘍手術（切除）と胆石摘出術（胆道切開）」などである。これらの用語は要素に分解し、辞書機能により統合して作成し利用しマスタと連携していれば複雑性や曖昧性などの問題は解決される。

(2) 処置・手術の用語の構造

処置・手術の用語の特徴として、用語

を構成する要素は以下のように分類された。すなわち部位、臓器、病理学的症状、経路、手技内容、機器、時間要素に大きく分類された。これらの用語の要素を複数個組み合わせで処置・手術の用語が作成されている。ICD・9・CMの手術・処置用語集では、約半数が3個以上の要素を組み合わせで作成されていた。たとえば「キセノン光凝固法による網膜裂孔の修復」（機器の種類、臓器、病理学症状、手技内容）、「内視鏡的な膵管結石除去」（経路、臓器、病理学症状、手技内容）である。

D 考察と残された問題

本研究では、処置・手術の用語の構造化をめざして、実際に利用されている標準手術・処置マスタの用語を例に構造を分析し検討したが、要素ごとに分解し辞書化し組み合わせることで用語の統一的な利用が可能であることを明らかにした。今後は処置・手術の用語に特異的な用語を網羅した領域別辞書を作成し、実際のシステム検証を行う必要がある。

E 学術発表

- 1) 日本医療情報学会にて発表予定
(2001年11月)

厚生科学研究費補助金(医療技術評価総合研究事業)
分担研究報告書

神経学領域における用語の整備とシステム検証

(分担) 研究者 津本 周作 島根医科大学医学部医学科医療情報学教授

要旨. 階層的構造を含めて、さまざまな知識表現が医学用語に構造を入れるために提案されてきたが、提案された知識表現の正確性、頑健性を評価することが困難であった。本研究では、近年注目をあびつつあるデータベースからの知識発見の手法の1つであるルール生成法を利用することで、知識表現の評価方法を開発した。

A. 研究目的

階層的構造を含めて、さまざまな知識表現が医学用語の構造として提案されてきたが実用化に至っていない。この原因の一つには提案された知識表現の正確性、頑健性を評価することが困難であったことがあげられる。本研究では、近年注目をあびつつあるデータベースからの知識発見の手法の1つであるルール生成法を利用することで、知識表現の評価方法を開発した。

B. 研究方法

ルール生成を用いた知識表現の方法は以下のプロセスで構成される:(1) 知識表現の評価に必要なデータ集合(以下、標準データと呼ぶ。)を用意する。(2) 標準データについてルール生成の手法を適用し、決定論的で無矛盾なルールが生成されることを確認する。(ルール生成を情報の縮約と捉えれば、無矛盾なルールが得られるということは、標準データがルール生成の枠組みの中で無矛盾であることを示唆している。(3) ある知識表現で得られた階層構造をデータの属性を変換するルールの形に翻訳する。(4)この変換ルールを用いて、

標準データの属性を変換する。一般に、階層構造の上のレベルでは、一般化した知識が、下のレベルでは、詳細化した知識が記載されているので、上のレベルで変換した場合は汎化した属性が、下のレベルで変換した場合は特殊化した属性が得られる。(5)変換した標準データにルール生成法を適用する。

C. 研究結果

上記の手法を木の形式をもった階層構造をもつ神経学的所見のタキソノミーを例にして適用したところ、以下のことがわかった:(1) もし、上記の階層構造に矛盾がなければ、生成されたルールは決定論的なルールが生成された。(2) もし、確率的なルールのみが得られれば、仮定した知識表現にはあいまいさが含まれていることがわかった。(3) あいまいさを解消するには、あいまいさを引き起こした階層部分の構造を解消することで、(1)の形に修正できた。(4) 過程において、知識表現の近似度が定量化できた。

D. 考察

階層構造の改訂に対しては次のような可能性

を考えなければならない: (i) 知識表現が粗いために、上記のような fuzziness がもたらされた。(ii) 知識表現の問題ではなく、データの方にノイズあるいは誤りがあったために、上記のような fuzziness が現れた。もし、(ii) の場合でない標準的なデータ(データにノイズを含まない)であると仮定すれば、(i) の場合を考えなければならない。つまり、標準データを用いることで、知識表現の粗さを評価できる可能性が示唆された。

E. 結論

本研究では標準データに、概念構造(知識表現)とルール生成法とを組み合わせ、適用することで、概念構造のあいまいさとその表現の近似の程度を評価できる可能性を示すことができた。

F. 健康危険情報

特記すべきことなし。

G. 研究発表

1. 論文発表

Tsumoto S: Knowledge discovery in clinical databases and evaluation of discovered knowledge in outpatient clinic. *Information Sciences*, 124, 125-137, 2000.

Tsumoto S: Automated Discovery of Positive and Negative Knowledge in Clinical Databases. *IEEE BME Magazine*, 19, 56-62, 2000.

2. 学会発表

津本周作: 標準データとルール生成に基づく知識表現の評価、第20回医療情報学会抄録集, 2000.

Tsumoto, S. Fuzzification of attribute

oriented generalization and its application to medicine. *Proceedings of SPIE Aerosense Conference*, 2000.

H. 知的財産権の出願・登録状況

特記すべきことなし。

薬剤に関する用語の整備とシステム検証

分担研究者 土屋 文人 東京大学医学部客員研究員

研究要旨 薬剤に関する用語については副作用の分野でMedDRAの利用が急速に進展しつつある。しかしながら、基本部分ともいえる医薬品名あるいは厚生労働省の承認事項である用法・用量や効能・効果に関する記載については、依然として標準化が行われていないのが現状である。電子カルテでの使用を考慮して、医薬品の標準的コードとしてHOT番号が作成されたことから、医薬品コードと対応関係にある医薬品名についての整備を行った。

A. 研究目的

医薬品名に関しては、販売名、薬価基準告示名、一般名（成分名）等さまざまなものが存在する。臨床上の有用性を考慮して、販売名について、その構成要素を調査し、かつ整備を行う。

B. 研究方法

医療機関の薬務関連業務における電子化の進展状況を把握するとともに、最近薬薬連携という形で施設間の情報交換が行われる場合に関して、その実情を把握する。さらに電子化を阻害する要因について調査し、その解決策についても検討する。また個人認証に関連して薬務関連業務において潜在的な問題についても検討を行う。

C. 研究結果

1 販売名の現状

販売名は表1に示すように規格や剤形あるいは語幹（商標）、企業名等の要素が順番を含め全く統一がとれていないため、オーダーリングシステム等で当該医薬品を選択するための略称をどのように定めるかが恣意的に行われている。またオーダーリングシステム等で医薬品名を登録する際には画面表示上の制約等もあり、各医療機関でそれぞれ修正が行われることから、複数の医療機関間で医薬品名を利用し

表1 販売名の例

2mgセルシン錠
O. S. T. ビタカプセル
O・S・T・ミックス、カプセル
ウチダのジャシヨウシM
総合ビタミン細粒ポボンーS
白色コントミン散
「純正」炭マ
ジュンコウ補中益気湯F Cエキス顆粒医療用
アドナ (AC-17) 錠 (10mg)

てデータのマッチングを行うことは極めて困難な状況にある。医薬品の標準コードが普及することでこれらの問題はある程度克服することができるが、医薬品のコードが定められても、医薬品の名称の標準化が行われていないと臨床での実質的な効果が期待できないことになる。

2 構成要素の分解及び整理

- 1) 販売名を構成する要素を、カタカナ、漢字、数字・英字、記号（「」や<>等）の文字種の違いを利用して販売名を分解する。
- 2) 分解できなかった要素があった場合には編集機能を利用して、「語幹」「剤形」「規格・単位」の3要素について分解する。また、「規格・単位」については「規格」を表す「数字」と、「単位」とを区分する。
- 3) 上記以外の要素については、その内容を確認しながら、意味を有する最小のまとまりに分解する。

以上の手順により、販売名の要素は延べで23要素に分解できた。表2に投与区分別の構成要素の有無を示す。そこでこれらの要素を以下の手順で整理した。

- 1) 語幹前修飾語は2つの要素に分解されていたが、これを1つにまとめ、かつ「水溶性」や「複合」「L」等医薬品の性質を表現するために必要最低限のものみに整理し、不必要と思われるものについてはデータを削除した。
- 2) 語幹（商標）については基本的にそのままとした
- 3) 語幹後修飾は徐放性製剤を識別するための記号（L, LA, R等）等同一商標下での薬剤の区別をするために使用されている例が多い。しかしながら、これらの記号を有する医薬品ではその記号を書き落とした場合に他の薬剤の販売名になってしまう例が少なからずあった。この点に

については厚生省（現厚生労働省）の通知により今後このような薬剤が新規に販売される場合については従来の薬剤も名称変更をすることになることから、現状の販売名に対しても同様の改善が必要なので、「_」を付して、何らかの文字が必要であることを記録した。

- 4) 剤形については、承認時に使用される剤形が必ずしも臨床側で考える剤形とは合致しない。例えば「ドライシロップ」は承認時の剤形はあくまで「シロップ」とされているが、調剤を行う場合にはその薬剤は「顆粒」や「細粒」であるため、「散剤」との分類に入ってしまう。これらは外用剤についても同様であり、剤形についてはかなりの整理を要することが判明した。
- 5) 剤形（用途）については、外用剤において、販売名に「眼・耳科用」等の用途を示す用語が使用されているものと、「点眼液」のように剤形を表すものが混在している。従って薬剤の性状を示す「剤形」とその「用途」とを区別することが望ましいと考える。
- 6) 規格（数値）についてはヒューマンエラー防止の点から「倍散」や「%」といった表示は外用剤と一部の注射剤を除き原則使用しないことにした。具体的には希釈された散剤については「○○mg/g」という表示に統一し、また注射剤では現状で「1%2mL」という規格が表示されていて、医師の指示が「10mg」というような場合に計算違いを犯す危険性があることから、当該注射剤に含まれる主薬の全量が明確になるように「○○mg/△mL/A」といった記載方法を採用することにした。
- 7) 配合用規格とは配合剤で複数の規格がある場合の識別用に使用する。
- 8) 容量数値は現行の薬価基準においては、外用剤や注射剤について1容器あたりの量が必ずしも明確でないことから1容器内の量を示すためのものである。
- 9) 容量単位は1容器内の量の単位を示す。
- 10) サイズは外用剤においてサイズを示すM、L等の記号がないものも存在することから、それらの記号とは別にそのサイズをa×b×cのような形式で表現することにした。
- 11) 企業名については製造会社、販売会社、輸入会社等があるが、現状ではこれらが明確ではないことから詳細な情報収集が必要である。
- 12) 付加情報はその他必要な情報を記載する。

表2 投与区分別販売名の構成要素の有無

	内用	外用	注射
前修飾 1	○	○	○
前修飾 2	○	○	○
語幹	○	○	○
後修飾 1	○	○	○
後修飾 2	○	○	○
後修飾 3	○	×	○
後修飾 4	×	×	○
前剤形	○	○	○
剤形 1	○	○	○
剤形 2	○	○	○
規格（数値）	○	○	○
配合用規格	○	○	○
規格（単位）	○	○	○
容量数値	○	○	○
容量単位	○	○	○
サイズ 1	×	○	×
サイズ 2	×	○	×
サイズ 3	×	○	×
企業 1	○	○	○
企業 2	○	○	×
付加情報 1	×	○	○
付加情報 2	×	○	○
付加情報 3	×	×	○

3 情報処理用基本標準名の作成

各要素における問題点を把握する

医薬品を特定するのに最低限必要な「語幹」「剤形」「規格・単位」の3要素を基本とした医薬品名（以下基本標準名という）を作成する。

基本標準名については外用剤、注射剤において情報の「規格・単位」に情報不足のものが存在することにより、他品目との区別が容易に可能となったと考えられる。また散剤や注射剤において、医師等が投与量を決定する際に従来計算を要していた%や倍散といった表示がなくなったことと、基本的には注射剤については1本に含まれる原薬の総量が表示されることにより、与薬時の判断ミスを防ぐ効果もあると考えられる。表3に基本標準名の例を示す。

表3 基本標準名（下段は現行の販売名）

セルシン錠（2） 2mgセルシン錠
コントミン散100mg/g
白色コントミン散
補中益気湯顆粒（ジュンコウ） ジュンコウ補中益気湯FCエキス顆粒医療用
静注用キシロカイン注100mg/5mL 静注用キシロカイン2%（5ml1管）

D. 結論

基本的標準医薬品名を試作することができた。今後はHOT番号を使用する施設において、実証実験を行う予定である。また今回、医薬品名を構成する要素が分解されていることから、医薬品名をタグ付きで構成することも可能になった。今後はこれらの点をふまえ、他の用語の標準化を歯く必要がある。

E. 研究発表

土屋文人；医薬品名の標準化と類似性の検討、医療情報学（投稿中）