

---

一般口演

## 一般口演5 知識工学

2017年11月21日(火) 08:30 ~ 10:00 G会場 (10F 会議室1006-1007)

---

### [2-G-1-OP5-4] 検査表現と潜在的解釈の相互変換方法とその評価法の設計

山本 英弥, 伊藤 薫, 矢野 憲, 若宮 翔子, 荒牧 英治（奈良先端科学技術大学院大学情報科学研究科）

【背景】ビッグデータ化した医療テキストの二次利用に注目が集まっているが、これまでの多くの研究は名詞、または複合名詞として表現された病名や症状の抽出に留まっていた。しかしながら、複合名詞の範囲を超えた病気や症状の表現も一定量存在し、医療言語処理にとって大きな課題となっている。

【目的】本研究では、名詞でない症状表現のうち、検査表現（検査名と検査値のペア）を疾患を意味する言語表現へ変換することを目指す。

【手法】まず、日本内科学会にて2002年から2015年に報告された約4万5千件の症例報告の本文から検査名、検査値、言語による解釈という3つの要素が共起する文を自動抽出する。これは、検査名の後続の3文字以内から始まる数字列を検査値とし、検査値の後続の25字の中で一番初めに登場した疾患名・症状名を解釈とみなすことで行う。例えば、「CRP12.0と強い炎症が見られ」においては、（CRP,12.0,強い炎症）が（検査名、検査値、言語表現）となる。最後に得られた検査名、検査値、言語表現セットを統計的に処理し、検査名ごとに言語表現のとりうる検査値をモデル化した。

【実験】抽出された言語表現が検査値のガイドラインに記載されているもの（「異常」や「高値」など）は、ガイドラインを参考に評価を行った。また、主観的な表現の場合（「炎症」など）は、数値から言語表現への変換可能性が一意に定まらないことが多く困難である。そこで、医療従事者が、言語表現から逆に検査値を推定する手法で妥当性を検証した。

【結果・考察】検査値の中でも頻出するHbA1cとCRPについて検証を行った。この結果、ほぼガイドラインや医療者の常識の範囲に収まる値を得た。本研究により、検査表現を、疾患を意味する言語表現に変換できる可能性だけでなく、その逆変換もできる可能性も示唆された。

# 検査表現と潜在的解釈の相互変換方法とその評価法の設計

山本英弥<sup>\*1</sup>、伊藤薫<sup>\*1</sup>、矢野憲<sup>\*1</sup>、若宮翔子<sup>\*1</sup>、荒牧英治<sup>\*1</sup>

\*1 奈良先端科学技術大学院大学

## Bidirectional Conversion Method between Inspection Values and Clinical Phenotypes in Medical Text

Hideya Yamamoto<sup>\*1</sup>, Kaoru Ito<sup>\*1</sup>, Ken Yano<sup>\*1</sup>, Shoko Wakamiya<sup>\*1</sup>, Eiji Aramaki<sup>\*1</sup>

\*1 Nara Institute of Science and Technology

The amount of medical texts has been increasing recently; it naturally leads to the situation of that secondary use of medical text data has drawn much attentions. However, there are a lot of problems in extracting medical information from the data written in natural language because natural language allows various ways of expression for one thing. For example, the concept of inflammation may be expressed by simply “inflammation”, but it may also be expressed by using a pair of the inspection name and its value such as “CRP 25”. To link a concept with its various expressions, this paper proposes a method to convert a pair of an inspection name and its value to a clinical phenotype. We used about 45,000 case reports collected from 2002 to 2015 as our materials. We extracted SIE (Sets of Inspection Expression) that consists of keyword, inspection value, and clinical phenotype. In order to evaluate the SIE, we used co-medicals’ opinions and a medical guideline. As a result, SIE becomes similar to human understanding and do not match the guideline. However, we should be aware that SIE depends on who wrote an original medical text and what context diagnoses were taken. We suggest a method to deal with inspection values in medical texts and hope that the method triggers off developing a method for these expressions that machines cannot easily deal with.

**Keywords:** Natural Language Processing, medical text, ICD-10, case report, EHR

### 1 はじめに

電子カルテ(以降、単にカルテと呼ぶ)は、初診から退院までの患者の全体像を記述した情報源であり、患者情報があまねく記録されていると考えられるものの、自然言語処理が困難な非文法的かつ断片化した表現が多く含まれており、利活用は困難であった。そこで、これを二次利用するためカルテから医療用語の自動抽出や病名の自動コーディングを行う研究が近年活発に行われている<sup>1)2)3)</sup>。

しかし、これまでの自然言語処理システムは名詞(もしくは、複合名詞)として表現された表現のみを扱うものが多く、記述された情報をフルに利用するには至っていない。実際に、カルテから、ある症状を抽出する場合、症状が名詞として表現されている場合の他に、検査値が用いられている場合などが考えられる。しかしながら、現状の言語処理の技術では、検査値から症状の情報を抽出するのは困難であり、従来はあまり行われてこなかった<sup>4)</sup>。

例えば炎症に関して、検査値を用いた表現方法には以下のような場合が考えられる。

1. CRP 18.8mg/dL と炎症反応の上昇を認め
2. CRP 18.8mg/dL
3. CRP 18.8 と高値

例1のように、解釈が「炎症反応」という名詞句で明記されている場合は、これまでの言語処理の扱う対象であり問題ないが、例2のように、症状が存在せず、単に数値として表現され

る、または、例3のように数値とその解釈のみが記述される表現も存在する。

そこで、本研究では、例3のような検査値と症状がともに出現する例を収集し、症状を数値表現に対する解釈の教師データとすることで、名詞として表現されていない数値表現に対しても解釈のキーワードを与えることを目指す。

まず、対象となる症例報告データ(2節)から、検査値と解釈のデータセットを構築する(3節)、次に、それを元に、検査値から解釈への変換方法及び解釈から検査値への変換を行う(4節)。最後に、本研究の妥当性を、ガイドラインとの比較及び人間との比較によって検証する(5節)。

### 2 材料

本稿では医療テキストとして日本内科学会症例報告を用いる。本章では、その詳細について述べる。

#### 2.1 症例報告とは

症例報告は、臨床例に考察を加えて報告するものであり、タイトル、著者、キーワード、本文、投稿日、投稿先(学会支部、回数)から構成される。この構成は全報告に共通であるが、本文は自然言語文で書かれており、記述形式は決められていない。

#### 2.2 文書数

今回扱うのは、日本内科学会にて2002年から2015年に報告

されたものであり、北海道から九州まで 10 の地域に分けられ、地域毎に集計が行われ、総数は 44,761 件である。

### 2.3 分析

本稿 1 節で述べた通り、症例報告の本文の記述形式は標準化されておらず、計算機にとって扱いづらい。自然言語で書かれた文章からできるだけ多くの情報を抽出するためには、名詞の範囲を超えた表現を扱う必要がある。例えば、疾患や症状に関する表現を例に絞ると、以下の 4 通りの表現方法に大別される。なお、ここでいう疾患や症状に関連する表現の基準は ICD コードに変換できるかどうかを基準とした。

医療事務経験者 1 名が調査した症例報告 28 件における表 1 の出現頻度は、a,b,c,d の例がそれぞれ 330 件, 76 件, 41 件, 41 件抽出された。

表 1 医療テキストにおける疾患の表現方法

表現方法	例	説明
a. 名詞・複合名詞による表現	咽喉痛	名詞のかたまりで、一般的に疾患名と認識できるもの。
b. 部位名が離れている表現	両肺野に多発結節影	症状だけではコードがつかないが、「部位名+症状」でコードがつく場合、コードがついている症状であるが「部位名」をつけることでコードが変わる場合、「XX の YY」のように助詞を伴うことでコードがつく場合。
c. 検査名・検査値の表現	抗ミトコンドリア抗体が 21	検査名+高値、検査名+低値、検査名+異常、検査名+集積などでコードがつく場合。
d. その他の表現	髄膜刺激徴候	疾患を推測させる表現。

## 3 方法

本研究の基本アイデアは以下である。まず、「CRP 18.8mg/dL と炎症反応の上昇を認め」といった検査名「CRP」、検査値「18.8」と疾患・症状表現「炎症反応」が共起したデータを大量に収集する。ここで疾患・症状表現を検査値の解釈と仮定すると、ある検査について検査値とその解釈のペアが得られ、検査値の統計を得ることができる。この統計から、検査名と検査値のみで出現した場合に、解釈を付与することができる。

### 3.1 検査値・解釈データセットの構築

まず、検査値・解釈データセットの構築方法について述べる。対象のテキストデータを文章ごとに区切り、その各々の文章毎に、キーワード・数値表現・解釈の三つが共起するもののみを抽出する。ここでいうキーワードは、検査名を表しており、数値表現は、各キーワードの後続の 3 文字以内から始まっている数字とした。解釈には、病名抽出器 MEDEX/J<sup>5)</sup> によって抽出されたものの中で、最初に出現したものをを用いた。ただし、解釈については、数値表現の後続 25 字を取り出せないものは取り出せる最大の文字数を取り出し、そこに解釈が含まれている場合のみを扱った。

例えば、「CRP 18.8mg/dl と炎症反応の上昇を認め」というテキストに対しては、(検査名, 数値表現, 解釈) = (CRP, 18.8, 炎症反応)とそれぞれ定まる。

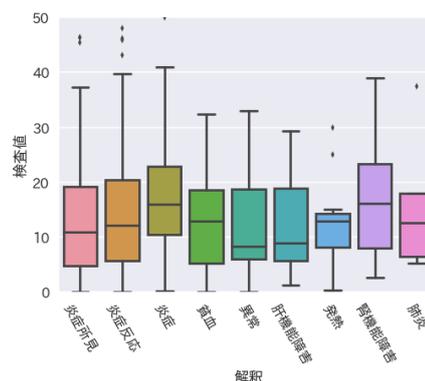
### 3.2 検査値・解釈間の変換

数値表現に後続している解釈の種類毎に、数値の分布を得ることができる。評価にあたっては、1つの検査名に対して 10 件以上得られた解釈のみを扱った。2 章の材料に対して、文献<sup>1)</sup>にて頻出した検査名 10 個(表 2)を対象に、3.1 の方法で抽出した。その結果、Hb に対しては、解釈が 33 件得られた。

表 2: 頻出した検査名上位 10 件

検査名
WBC
CRP
CT
HbA1c
CAG
Hb
血圧
心エコー
Cer
Cre

以下に CRP での例を示す。図の線は上から順に、最大値、



第三四分位数, 中央値, 第一四分位数, 最小値をそれぞれ表している。

図 1: 各解釈に対する検査値の例(CRP)

## 4 実験

前章の方法で抽出したデータの評価をガイドラインとの比較(4.1)と人間との比較(4.2)という二つの方法で行った。

## 4.1 ガイドラインとの比較

ある検査の数値を入力として受け取り、その値が出た場合のもっともらしい解釈を返すタスクを行った。ガイドラインと比較して、どのような違いがあり、どのような合致点があるのかを検討する。

### 4.1.1 比較対象

ガイドラインに掲載されている内容を比較対象とした。検査指標として血圧を選んだ。記載内容は以下の通りである。

表 3: ガイドラインにおける血圧と解釈の対応

血圧(最高血圧)	解釈
180	重症高血圧
160~179	中等症高血圧
140 以上 かつ 最低血圧 90 未満	収縮期高血圧
140~159	軽症高血圧
130~139	正常高血圧
130 未満 かつ 最低血圧 85 未満	正常血圧
120 未満 かつ 最低血圧 80 未満	至適血圧

### 4.1.2 方法

ガイドラインに掲載されている検査値に対する解釈を確認し、今回得られた検査値・解釈の組と比較した。解釈に対する検査値のレンジは、第一四分位数から第三四分位数の範囲とし、同一の検査値に対して複数の解釈が得られる場合は一位に定めず、共起を認めた。また、厳密な意味での血圧は、拡張期血圧(最低血圧)と収縮期血圧(最高血圧)の二種類があるが、医療テキストで記載されるのは最高血圧が多数であると考え、今回得られた検査値は最高血圧として扱った。

### 4.1.3 結果

血圧に関して、血圧の数値のレンジとそれに対応する解釈を図 2 に示す。意識清明は 109~177.5、貧血は 82~118.25、ショック状態は 70~165、高血圧は 90~189 というレンジであり、ガイドラインには出てこないような解釈が出て、ガイドラインでは正常血圧と考えられるレンジにも高血圧の解釈が与えられた。

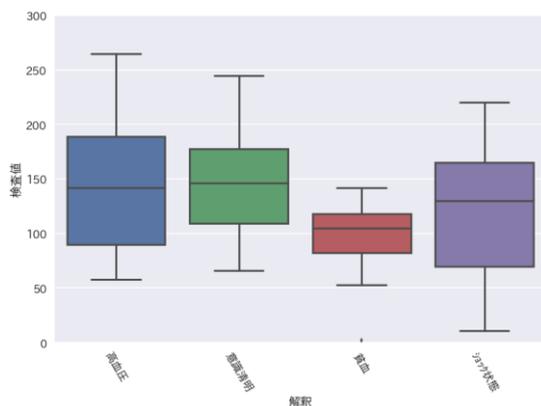


図 2: 各解釈に対する検査値の例(血圧)

表 4: 血圧と解釈の対応

血圧(Q1~Q3)	解釈
109 ~ 177.5	意識清明
90 ~ 189	高血圧
82 ~ 118.25	貧血
70 ~ 165	ショック状態

### 4.1.4 考察

血圧を例にとると、ガイドライン上には、最高血圧が 140 以上の場合に高血圧と解釈するという記述があるが、今回得られた検査値と解釈のペアを確認すると、症例報告ではより低い値でも高血圧という解釈を与えていることがわかる。この例から、高血圧という頻用される病状も、臨床の現場においては 6 幅をもって使用され、検査値に対する解釈はその状況に応じて使われ方が異なるということが考えられる。

ただし、解釈として異常が抽出された際、検査値が高すぎ場合と低すぎる場合の両方を吸収しており、四分範囲を用いたモデルが妥当ではなく、多峰性を扱えるモデルの導入が必要であることもわかった。

## 4.2 人間との比較

医療従事者が、ある一つの検査値の部分のみを空白にした文書全体を読み、検査値を予測するタスクを行った。人間の感覚と比べてどのような違いがあり、どのような合致点があるのかを検討する。

### 4.2.1 比較対象・方法

今回用いた症例報告の原文を、数値の部分のみを隠した状態にし、医療従事者 3 名(内科医 1 名、医療事務経験者 2 名)が本文全体の文脈から検査値の部分の推測したものを 30 件ずつ用意した。検査指標には CRP を選択した。これに用いた文書の例を以下に示す。

北海道 253 20091121 15 ステロイド療法を施行した炎症性腹部大動脈瘤の 1 症例 症例は 68 歳、男性。一週間前より 38 度台の発熱、腹部・背部痛を自覚し市販の消炎鎮痛薬で経過を見ていたが症状の改善を認めなかった。近医受診し CRP 23mg/dl と高値を認め抗生剤を処方されるも症状に変化なく、再診時にも CRP 20mg/dl と依然高値であった為、精査加療目的で当院紹介入院となった。入院時 WBC 8140/ $\mu$ l だが、CRP  $\square$  mg/dl と高度の炎症所見、および造影 CT にて腹部大動脈径の拡大とマンツルサインを認めたことより炎症性腹部大動脈瘤と診断した。炎症巣を含めた瘤径は 35 $\times$ 44mm、大動脈内径 20 $\times$ 18mm であり切迫破裂の危険性は低いと判断し、プレドニン 40mg/日による薬物療法を開始した。ステロイド療法開始後 2 日目より解熱を認め腹痛・背部痛は消失した。ステロイド開始 2 週間後には CRP は陰性化しプレドニンを 30mg/日へと減量したが、炎症巣を含めた瘤径は変化がなく大動脈内径は若干増加傾向にあり慎重に経過観察中である。今後増大傾向認めれば外科的治療を考慮している。<br>これまでの報告では発見時より手術適応の症例が多い一方、瘤径が小さい症例でステロイドによる内科的治療が奏功したという報告も散見されているが、その臨床経過・治療法については不明な点が多い。

図 3: 実験に用いた文書例

が数多く生まれることを期待したい。

## 4.2.2 結果

結果を図4に示す。グラフの横軸が元の文章に入っていた数値、縦軸は各被験者が予測した数値である。直線は、各被験者のデータ毎に最小二乗法で推定した回帰直線である。

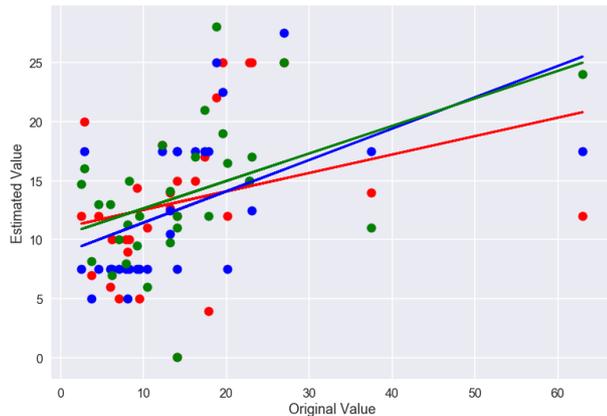


図4: 被験者の推測値と元データの比較

三人のデータの(直線の傾き, 相関係数)はそれぞれ(0.16, 0.28), (0.26, 0.51), (0.23, 0.47)であり, 抽出したデータと人間の感覚には相関がみられた。

## 4.2.3 考察

この実験により, 検査値とその解釈の間には一定の合意があるとみなせる。これは, 提案手法の大前提の妥当性が示されたと考える。また, この実験で用いたデータには CRP の値が60を超えるものが入っており, そのような数値を被験者が予測できず, 結果に影響を与えていることも考えられる。

## 5 考察

本稿では, これまでは扱いが困難とされてきた医療テキスト内の検査値表現を扱う方法として, 検査名と検査値のペアを抽出し, それに対して与えられる解釈を抽出する方法を設計し, それを元に, 検査値から解釈への変換方法及び解釈から検査値への変換方法を提案した。本研究から考えられることは, 症例報告以外の医療テキストに対して本方法を適用した場合, 性質の異なったデータベースが得られるだろうということだ。医師や現場の環境に応じて感覚量が異なると考えられ, 臨床への応用を考える際は, 考慮する必要があると思われる。

## 6 おわりに

これまででは, 医療テキストデータからの情報抽出は名詞の範囲に限定されてきた。本稿では, その範囲を超えた検査名・検査値の表現についても扱える可能性について述べた。本稿の抽出方法は, ガイドラインには出てこないような多様な解釈(例えば, 血圧における貧血やショック状態)を抽出可能である。これによって, 副次的に起こるような疾患を抽出することができる可能性がある。また, 課題としては, 「異常」という解釈の取り扱いが挙げられる。

今後は, 検査表現のみならず, 様々な表記方法を扱う方法

## 謝辞

ご多忙の中, 実験に協力してくださった東京大学大学院医学系研究科の河添悦昌講師には感謝を申し上げる。本研究の一部は, 厚生労働省科学研究費補助金(課題番号: H28-ICT-一般-008)

## 参考文献

- 1) Aramaki, Miura, Tonoike, Ohkuma, Mashuichi, Ohe: TEXT2TABLE: Medical Text Summarization System Based on Named Entity Recognition and Modality Identification, In Proceedings of BioNLP, 185-192, 2009.
- 2) Wendy Chapman, John Dowling and David Chu. 2007. ConText: An algorithm for identifying contextual features from clinical text. Biological, translational, and clinical language processing (BioNLP2007), pp. 81-88.
- 3) C. Friedman, P.O. Alderson, J.H. Austin, J.J. Cimino, and S.B. Johnson. 1994. A general natural language text processor for clinical radiology. Journal of the American Medical Informatics Association, 1(2):161-174.
- 4) 篠原(山田)恵美子, 三浦康秀, 大熊智子, 増市博, 荒牧英治, 大江和彦. 退院時要約に見られる検査定量値と対応する定性表現文字列の自動抽出. 医療情報学 2012; 32(Suppl.): 288-291.
- 5) 矢野憲, 若宮翔子, 荒牧英治. 医療テキスト解析のための事実性判定と融合した病名表現認識器. 言語処理学会 第23回年次大会 2017