

電子カルテ自由記述部分からの皮膚疾患における重症度抽出

加藤 由美¹ 平川 聡史² 梶山 晃平³ 堀口 裕正⁴ 狩野 芳伸⁵

静岡大学^{1,3,5} 浜松医科大学² 国立病院機構⁴

{ykato¹, kkajiyama³}@kanolab.net, hirakawa@hama-med.ac.jp², horiguchi-hiromasa@hosp.go.jp⁴,
kano@inf.shizuoka.ac.jp⁵

1 はじめに

本研究では、抗がん薬によって起こる代表的な4つの皮膚障害(爪囲炎・ざ瘡様皮疹・皮膚乾燥・亀裂)について、電子カルテから重症度を抽出する。これにより、将来的ながん治療の際の医師の投薬判断のサポートを目指す。

電子カルテの導入によりデジタル処理が行えるようになり、疾患名や薬品名といった固有名詞を抽出する研究[1]や、数値データや画像データを用いて疾患や重症度を推測する研究[2][3]が進められている。薬剤投与や処方により、患者には副作用をはじめとする様々な意図しない好ましくない症状、有害事象が発生する。有害事象は個々の患者により発生頻度や程度が異なる。そのため、医師は患者の有害事象の発生状況を把握し、適切な薬剤の量や種類を決定することが重要となる。

しかし、電子カルテの自由記載部分は構造化されていないため解析が難しく、自由記載部分のデータを基にした疾患の重症度を判定する研究は少ないのが現状である。島井ら[4]は画像レポートにより間質性肺炎を特定するため、フリーテキストデータの処理では各キーワードの出現頻度とかかり受け解析による否定語の情報を用いて特定している。本研究では、疾患の有無でなく重症度を抽出するため、否定語の有無に加えて疾患の状態を表すキーワードも活用した。

2 医師の診断基準

有害事象の重症度(Grade)は有害事象共通用語規準(CTCAE)が世界的な基準として用いられており、5段階である。本稿ではv4.0日本語訳

JCOG版(CTCAE v4.0 - JCOG)に基づく。CTCAE v4.0 - JCOGでは各症状のGradeによる違いが規定されており、皮膚および皮下組織障害については主に以下の観点で記載されている:

- ・ 症状
- ・ 症状が占める体表面積
- ・ 疼痛の有無
- ・ 身の回り以外の日常生活動作への影響の有無
- ・ 身の回りの日常生活動作への影響の有無
- ・ 必要な治療内容
- ・ 生死への影響の有無

ただし、症状が占める体表面積について、医療現場でカルテに記載することは稀である。

3 模擬電子カルテデータの作成

医療データは患者のプライバシーに関わるため、厳しい利用制限が設けられている。そこで本研究の実験データには、共著者の一人である浜松医科大学の平川が作成した模擬カルテを用い、加藤がタグ付けを行った。がん薬物療法を受けている60代男性を想定した46日分の模擬データである。1回の診療で作成されたカルテに<karte>タグを、日付に<data>タグを、自由記載部分に<text>タグを付与した。<karte>タグには、診察日と医師の診断結果(ICDコードとGrade)を属性として付与した。

また、本研究では、清水宏の「皮膚病の診断法なら新しい皮膚科学」[5]を元に、カルテに記載されることが多い視診及び触診の際にポイントとなる、次の9つの発疹の性状に注目した:(1) 皮疹の数、(2) 個疹の形状、(3) 隆起の状態、(4)

表面の状態, (5) 色, (6) 硬度, (7) 配列, (8) 発生部位, (9) 自覚症状. これら 9 つの性状と CTCAE で定められた症状は, カルテデータの作成段階でキーワードとしてタグ付けした.

カルテは問題志向型記録の叙述的経過記録方式(SOAP 形式) で記載されている. 「S」は患者自身から話を聞くことで得られる主観的データ (Subjective data), 「O」は医師が患者を観察・検査することで得られる客観的データ (Objective data), 「A」は「S」と「O」から導き出される見解・アセスメント (Assessment), 「P」は「A」で明らかとなった状態に対し, 今後どのように改善していくかの計画 (Plan) のことを指す. Assessment には重症度が具体的に記載されることもあり, 模擬カルテでは全体の約 7 割で Grade の記載が見られた.

Subjective data と Objective data のキーワードとなる単語には <symptoms>, <size>, <geometry>, <uplift>, <surface>, <color>, <hardness>, <array>, <bodyregion>, <subjective>, <state>のいずれかのタグを付与した. 図 1 に例を示す.

```
<karte date="2006-7-5" diseases="L030-0 L709-0 L853-1 L309-1"><date date="2006-7-5">7月5日
</date>
<text><sentence> 2回目の入院(FP-Cet 2クール目). </sentence>
<sentence>S <bodyregion>手指</bodyregion>に<symptoms>亀裂</symptoms>が出来ます.
</sentence><sentence>水絆創膏が便利です.
</sentence><sentence>ドレニジ
ンテープは, 濡れると剥がれてしまいます.
</sentence>
<sentence>O <bodyregion>手指</bodyregion>に<symptoms>亀裂</symptoms>が<array>散在
</array>. </sentence>
```

```
</array>. </sentence>
<sentence>A 手指の乾燥・亀裂: Grade 1.
</sentence>
<sentence>P 外用主体に加療継続.
</sentence></text>
</karte>
```

図 1 模擬カルテの例. 2006 年 7 月 5 日の診察で, 爪囲炎・ざ瘡様皮疹が Grade0, 皮膚乾燥・亀裂が Grade1 であったことを表す.

本研究で利用したカルテデータには, 体表面積の記述 (cm, mm 表現) が 0 件であったのに対し, 部位の記述は 71 件であった. なお, 重症度の判断に必要なと思われる記述のみカウントした.

医療関係の専門知識を持たない学生が模擬カルテの Subjective data と Objective data を読み, CTCAE v4.0-JCOG を基に重症度をつけたところ, 正解率は約 7 割であった. 間違いのうち, カルテに疾患の記述がないことによる誤りが約 5 割, CTCAE に評価基準が載っていないことによる誤りが約 3 割であった. 記述がある疾患の重症度はほぼ読み取ることができ, 重症度判断の基準は明確であるといえる.

4 実験

2 節で述べたように, 重症度の診断にはいくつかのキーワードが存在する. 本研究では, キーワードを利用した機械学習手法とルールベース手法を検討し, 性能を比較評価する.

4.1 機械学習

4.1.1 システム構成

図 2 にシステムの構成を示す. まず自由記述部分を抜き出し, 形態素・構文解析を行い, 学習素性を抽出する. 分類器として SVM を用い, 多項式カーネルで学習した. 疾患の重症度をクラス (0~5) とした多値分類を行った. クラス 0 は疾患無しを意味する.

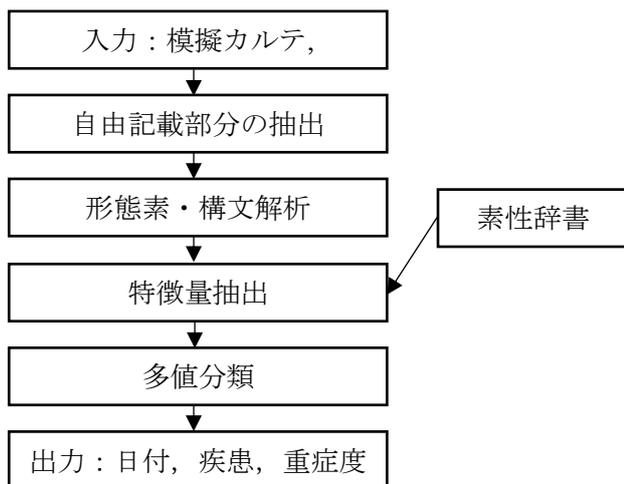


図 2 システム構成

4.1.2 学習素性

各症状の発疹の性状を表す単語の有無を素性とする。疾患を特徴づける症状の単語が現れた文と、その文の後ろ 3 センテンスの間にある単語が外部の語彙辞書と一致したかどうか、の 2 値を単語種類数ぶんの次元とした one-hot ベクトルである。つまり、疾患を特徴づける症状 N 個 \times 語彙 M 個 が次元数となる。

辞書は我々が作成した外部の語彙資源であり、2 節で述べた発疹の性状等を表す語を、医療参考書などを参考に 550 語収録した。

4.2 ルールベース

4.2.1 システム構成

機械学習と同様に形態素および構文解析を行い、以下の手順で推測を行った。同じ入力に対し複数の推測が得られるので、そのなかから最も確信度の高い結果を出力とした。

手順 1：Grade 表現の抽出

「Grade」と数値表現(0~5)が含まれる文の主格が期待する疾患名るとき、疾患と重症度を保持する。確信度は 5 とする。

手順 2：症状の状態による推測

症状の単語の係り先が「治癒」あるいは「消失」などの治癒表現、または否定表現があれば Grade

を 0 とし、確信度を 3 とする。あるいは係り先の「散在」、「残存」、または「軽快」、「軽減」といった発疹の状態に応じて重症度と確信度 2 を定める。症状の単語のみの場合には、Grade を 1 とし、確信度を 1 とする。

手順 3：部位表現を用いた精度補正

疾患には発生しやすい部位や部位ごとに名称が異なる。手順 2 で得られた文に特定の部位の表現が含まれている場合、確信度を 4 とする。

5 評価

Precision, recall, accuracy, F-measure により評価を行った。

5.1 機械学習

学習性能は 5 分割交差検証で測定した。表 1 に機械学習による疾患ごとの結果を示す。

	爪囲炎	ざ瘡様 皮疹	皮膚乾 燥	亀裂
precision	29.2	38.9	23.7	16.7
recall	67.4	76.0	60.9	50
F-measure	40.8	51.4	34.1	25
Accuracy	67.4	76.1	60.9	50

表 1 機械学習手法の評価結果

5.2 ルールベース

表 2 及び表 3 にルールベースによる疾患ごとの結果を示す。我々は、Grade 表現がないカルテからも重症度を抽出することを目指している。手順 1 を用いず重症度を抽出したところ、表 3 のように精度が全体的に低下した。

	爪囲炎	ざ瘡様 皮疹	皮膚 乾燥	亀裂
precision	45.1	36.2	16.6	29.2
recall	80.4	73.9	50.0	67.4
F-measure	57.8	48.6	25.0	40.8
Accuracy	80.4	73.9	50.0	67.4

表 2 ルールベース (Grade あり) の評価結果

	爪囲炎	ざ瘡様 皮疹	皮膚乾 燥	亀裂
precision	16.7	31.4	14.4	17.9
recall	50.0	69.6	45.7	52.2
F-measure	25.0	43.2	22.0	26.7
Accuracy	50.0	69.6	45.7	52.2

表 3 ルールベース (Gradeなし) の評価結果
色付きのセルは表 1, 表 2 及び表 3 の中で最も良い結果であったことを表す。

6 考察

必要な単語を素性として適切に利用できたかを調査したところ、重症度判断に必要な情報と思われる単語 185 語のうち、132 語を獲得できたことが分かった。また、その他に 450 語の単語を獲得していた。Assessment や Plan の記述から疾患を重複して獲得していた。獲得できなかった単語は、素性辞書に登録されていないことや表記ゆれが原因であった。調査は素性に利用した単語の文字位置と、3 節で述べたキーワードの文字位置を比較した。爪囲炎は重症度によって症状に大きな違いが現れるため、他の情報を考慮せず、症状のみで評価するルールベースの精度が高かった。皮膚乾燥は口唇炎（口の乾燥）を皮膚乾燥と誤って出力しており、確信度が 1 のものは出力しないことで対応可能と考える。また発疹の性状を素性として用いたため、爪囲炎や亀裂のような皮膚の表面の問題ではない疾患の特徴をうまく得ることが出来なかった。Grade3, 4 及び 5 の疾患の記述データが 0 のものがあり、クラスの偏りが大きかったことも機械学習の精度が低下した原因である。

7 おわりに

実験の結果、最も良い結果では人間とほぼ同様の 7 割前後の正確度で重症度を抽出することが出来た。重症度抽出に適したキーワードを選択できたといえる。カルテの自由記載部分は構造化され

ていないが、SOAP や CTCAE などの基準に沿ってカルテが作成されている。自由記述部分からの重症度抽出はある程度の精度で可能であろう。重症度の抽出により、患者の症状悪化の見落としが減り、QOL の向上につながると期待される。医療データは管理が厳しく、扱いが難しい。また Grade3 以上の疾患データは少なく、CTCAE の基準も Grade4 以降の表現が曖昧である。そのため対象データを増やすことが今後の課題である。

8 参考文献

- [1] 島井良重, 武田理宏, 真鍋史郎, 寺本圭, 三原直樹, 松村泰志, 電子カルテ蓄積データを用いた薬剤性間質性肺炎の発生頻度の算出法, 医療情報学, vol.36, no.4, pp.157-165, 2016
- [2] 東山翔平, 関和広, 上原邦昭, 医療用語資源の語彙拡張と診療情報抽出への応用, 自然言語処理, vol.22, no.2, pp.77-105, 2015
- [3] Dinu A.J., Ganesan R, Felix Joseph, Balaji V, A study on Deep Machine Learning Algorithms for diagnosis of diseases, International Journal of Applied Engineering Research, vol.12, no.17, pp.6338-6346, 2017
- [4] Tripoliti E, Papadopoulos T, Karanasiou G, Naka K, Fotiadis D, Heart Failure: Diagnosis, Severity Estimation and Prediction of Adverse Events Through Machine Learning Techniques, Computational and Structural Biotechnology Journal, vol.15, pp.26-47, 2017
- [5] 清水 宏, あたらしい皮膚科学第 2 版, 中山書店, 中山書店, 2011