

厚生労働科学研究費補助金（がん対策推進総合研究事業）  
（分担研究報告書）

がん患者の個々のニーズに応じた質の高い相談支援の体制整備のあり方に関する研究  
テキストマイニングを用いた相談支援内容の解析

研究分担者 塚本 憲史 群馬大学医学部附属病院（腫瘍センター長）  
研究協力者 角田 明美 群馬大学医学部附属病院（看護師）  
研究協力者 渡辺 恵 群馬大学医学部附属病院（看護師）

研究要旨

がん相談支援センターでの相談支援内容を、テキストマイニング技術を用いて解析した。2020年度のがん相談記録から症例100例を抽出し、個人情報等を除外したCSVファイルを作成、東京医科歯科大学内村先生が開発したソフトを用いて解析した。相談記録には相談内容と相談対応が記載されているが、この解析に必要なものは前者のみであり、多施設でこの解析を行うには、相談記録の記載方法を共通化する必要がある。相談記録をそのまま解析すると不要な情報が多く得られ、解析対象を相談内容に絞り、除外ワードの設定が不可欠であった。除外ワード設定後に再度解析したところ、頻出上位ワードのみならず下位ワードからも必要な情報が抽出される一方、頻出上位ワードに不要な情報も多数得られた。また、相談記録から類似相談内容が抽出できないかをサンプルデータ10例を用いて検討し、特定のキーワードから類似相談内容の抽出は可能で、相談内容、対応内容の幅、違いも可視化可能であった。しかし、この解析を日常業務の合間に行うにはハードルが高いと思われた。

A. 研究目的

がん相談支援センターで蓄積されたデータを利活用するために、相談記録情報を利用し、テキストマイニング技術により疾患別、カテゴリー別の傾向を分析し、頻度の高い相談内容と対応内容を分類する。そして、この解析から得られる情報を用いて、FAQの作成、他施設を含めた情報共有、自施設の相談支援業務の改善に活かさないかを検討する。

B. 研究方法

2020年度のがん相談記録から症例100例を抽出、個人情報等を除外、匿名化し、CSVファイルを作成し、東京医科歯科大学内村先生の解析ソフトを用いてテキストマイニングによる分析および可視化を行った。計量的分析手法として文章を分かち書きし、出現する単語などを集計してFAQへの応用が可能かを検討した。また、相談記録から類似相談内容が特定のキーワードから抽出できないかを、10件のサンプルデータを用いて検討した。

（倫理面への配慮）

患者のプライバシーに関わるもの、個人情報を含むものを除外または匿名化した上で利用した。また、診療内容に対する不満などの相談内容は解析対象外とした。

C. 研究結果

- 1) 相談記録をそのまま解析すると不要な情報が多数抽出され、必要な情報が得られなかった。分析対象を相談内容に絞り、また、除外ワードをあらかじめ設定する必要がある。
- 2) 除外ワードとして、相談内容ではない単語（患者、相談、情報等）、および、動詞、形容詞、副詞、助詞、接続詞、記号（句読点、括弧）など名詞以外を設定した。また、出現頻度が高いワードのほかに、頻度が低い項目の一覧を出力できるようにした。その結果、上位100ワードに具体的な検査、治療法が抽出される一方、下位100ワードにも病期、治療など目的とする情報が抽出された。また、上位100ワードに不要なワード（「こと」「院」「他」など）が抽出され、上位ワードのみに注目すると、必ずしも必要な情報にたどり着けないと思われた。
- 3) 相談記録から類似相談内容が抽出できないかをサンプルデータ10例を用いて検討した。特定のキーワードから類似相談内容の抽出は可能で、相談内容、対応内容の幅、違いも可視化可能であった。

D. 考察

相談記録の内容を大別すると、相談内容と相談対応に分かれる。このうち、今回の解析に必要なものは相談内容のみであり、多施設でこの解析を行うに

は、相談記録の記載方法を共通化する必要がある。

目的とする情報を得るには、できるだけ多くの除外ワードを事前に入力しておく必要があった。また、出現頻度の低いワードにも必要な情報が含まれており、単純に頻度の高いワードを抽出するだけでは不十分と思われ、評価時には注意が必要である。

類似相談内容の抽出、分析することで教育への利用、相談の質保証に応用可能と思われた。一方、解析対象を院内症例にすると、相談内容がその病院の特徴に影響される可能性があり、FAQ作成支援に用いる場合は、院外症例を対象とした方が普遍性の高い情報が得られると思われる。

ただ、この解析を行うにはそれ相当の労力が必要で、日常業務を行いながら行うにはハードルが高いと思われた。

#### E. 結論

必要とする情報がある程度得られるが、労力の割に得られる情報は限定的で、複数の施設で行うには課題が残った。

#### F. 健康危険情報

特になし

#### G. 研究発表

##### 1. 論文発表

1) Yamauchi N, Maruyama D, Choi I, Atsuta Y, Sakai R, Miyashita K, Moriuchi Y, Tsujimura H, Kubota N, Yamamoto G, Igarashi T, Izutsu K, Yoshida S, Kojima K, Uchida T, Inoue Y, **Tsukamoto N**, Ohtsuka E, Suzuki S, Inaguma Y, Ichikawa S, Gomyo H, Ushijima Y, Nosaka K, Kurata M, Tanaka Y, Ueda R, Mizokami M, Kusumoto S. Prophylactic antiviral therapy for hepatitis B virus surface antigen-positive patients with diffuse large B-cell lymphoma treated with rituximab-containing chemotherapy. *Cancer Sci.* 2021; 112(5): 1943-1954.

2) Fujimoto S, Kawabata H, Sakai T, Yanagisawa H, Nishikori M, Nara K, Ohara S, **Tsukamoto N**, Kurose N, Yamada S, Takai K, Aoki S, Masaki Y. Optimal treatments for TAFRO syndrome: a retrospective surveillance study in Japan *Int J Hematol.* 2021; 113(1): 73-80.

##### 2. 学会発表

1) 櫻井麗子、大崎洋平、清水雄大、大曾根勝也、大島宗平、角田明美、塚本憲史。免疫併診プロジェ

クトチームにおける免疫チェックポイント阻害薬の有害事象。第60回日本癌治療学会、2022年10月、神戸

2) 大崎洋平、櫻井麗子、塚本憲史。免疫チェックポイント阻害薬投与時の高齢者評価の有用性。第60回日本癌治療学会、2022年10月、神戸

#### H. 知的財産権の出願・登録状況

(予定を含む)

- |           |    |
|-----------|----|
| 1. 特許取得   | なし |
| 2. 実用新案登録 | なし |
| 3. その他    | なし |