

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）

がん相談支援内容の分析と分類に関する研究

研究分担者 内村 祐之 東京医科歯科大学医学部附属病院 医療情報部（特任講師）
研究分担者 早川 雅代 国立がん研究センターがん対策研究所がん情報提供部（室長）

研究要旨

がん相談支援で過去に対応した相談記録情報を利用したテキストマイニング技術による疾患別やカテゴリー別の傾向を分析し可視化する為、2020年度は自由記載で書かれた相談内容や対応内容から単語を集計し相談内容と対応内容に含まれる単語間の繋がりを可視化するプロトタイプを開発した。2021年度はプロトタイプを改良し、単語の出現頻度によるビジュアル化を行い、単語を数値ベクトルに変換してその意味を把握する自然言語処理の手法であるWord2Vecを用い、過去の相談内容から類似度の高い順に類似相談内容を抽出する機能を実現した。更にプロトタイプシステムをWebアプリケーションとしてブラウザのみで動くように改良を行い、がん相談支援を行っている施設での実証試験を開始した。

A. 研究目的

現在、全国のがん相談支援センターでは、一定の基準を満たした上で治療の内容や設備、がんに関する情報提供などを行っており、がん相談支援における相談内容をデータとして管理している。しかし、蓄積されたデータの利活用や相談員間の知識共有は十分に出来ていないのが現状である。将来の相談内容の更なる多様化に備えるためには、蓄積された相談情報を活用し、FAQの作成や相談員間の情報交換など他施設を含めた相談員のノウハウを共有する仕組みを作り、相談員による経験や知識に起因する対応の差を埋めることにより、一定水準以上の相談対応が行えるような相談対応支援の仕組みの構築が必要であると考えられる。仕組みの構築に向けて、がん相談支援で過去に対応した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やカテゴリー別の傾向を分析し可視化する。さらに、現状の相談内容や個々の相談員の対応の傾向を把握した上で、相談内容と対応内容を分類することで、FAQの作成支援や他施設を含めた情報共有の仕組みを検討し、自施設の相談支援業務に活かすことを目的とする。昨年度作成したプロトタイプシステムに改良を加え、実証試験にてシステムの検証を行う。

B. 研究方法

全国の「がん相談支援センター」の相談支援内容は施設により、がん相談支援センター「相談のための基本形式」(https://ganjoho.jp/med_pro/consultation/support/registration_sheet.html)の形式を基として、相談記録の要旨が電子カルテシステムまた

はExcelのシート、FileMakerなどの独自開発システムなどに相談員が入力したデータとして記録されている。この記録を収集し、計量テキスト分析とテキストマイニングにより分析および可視化、内容による分類を行う。2020年度は10件の相談記録「形態素解析による分かち書きで単語を集計」「係り受け解析」「共起ネットワークでビジュアル化」の3種のプロトタイプをプログラミング言語Pythonにより作成した。2021年度はプロトタイプに改良を加え、単語の出現頻度によるビジュアル化（Word Cloud）を行い、単語を数値ベクトルに変換してその意味を把握する自然言語処理の手法であるWord 2Vecを用い、過去の相談内容から類似度の高い順に類似相談内容を抽出する機能を実現した。更にプロトタイプシステムをWebアプリケーションとしてブラウザのみで動くように改良を行い、がん相談支援を行っている施設での実証試験を開始した。

（倫理面への配慮）

プロトタイプ開発の段階ではデータは全て実データと異なるサンプルデータを用いた。実証試験では実証試験施設である群馬大学医学部附属病院の倫理審査委員会の承認（テキストマイニングによるがん相談支援内容の分析と分類に関する研究）を経て、倫理規定等を遵守しながら進めている。

C. 研究結果

1. WordCloudによるビジュアル化と形態素解析による分かち書きで単語を集計
がん相談内容と対応内容に関するサンプルデー

タを元に頻出語を頻度に比例する大きさと文字を並べたWordCloudによるビジュアル化と単語の出現回数を可視化した。可視化の例を以下に示す。

1-1 がん相談内容のビジュアル化



1-2 がん相談内容の分ち書きによる単語集計

単語	出現回数
治療	8
相談	7
通院中	6
電話相談	5
受け	5
当院	5
がん患者	4
他院	4
言わ	4

1-3 がん相談内容に対する対応のビジュアル化



1-4 がん相談内容に対する対応の分ち書きによる単語集計

単語	出現回数
情報提供	15
お伝え	8
確認	6
特徴	5
お話	4
必要	4
受診	4
いただく	4
よう	4
相談	4
適応	4
通院	4
場合	4
治療	4

2. Word2vecによる類似文章検索

単語を数値ベクトルに変換し意味を把握する自然言語処理の手法であるWord2Vecを用い、過去の相談内容から類似度の高い順に類似相談内容を抽出する機能を実現した。サンプルデータによる抽出結果を以下に示す。

元の文章 (元がんサンプル)	全10データ中最も類似度が高かった文章 (元大サンプル)	類似度
子宮体部の癌性疾患が再発した。都内で実績のある病院を教えてください。	他院通院中の大腸がん患者より相談あり。手術と術後の化学療法を行い、現在は経過観察中である。術後の影響でまた手足の痺れが残っており、なんとか仕事をしている状態。同じような経験をしている人の話が聞きたいが、そうい場所がないかとのご相談。	0.7742452
3年間に乳がんステージIIIで術前化学療法、手術、放射線治療。現在タモキシフェンを内服している。1ヶ月前に頸椎の骨粗鬆症が診断された。今後手術が必要とされている。本日に手術を予定しているが、今後どのような経過を辿るかわからない。術後の経過が心配です。術後の経過が心配です。術後の経過が心配です。	乳癌患者より電話相談あり。浸潤性乳がんの診断を受け、今後手術予定。術後の治療が必要とされている。本日に手術を予定しているが、今後どのような経過を辿るかわからない。術後の経過が心配です。術後の経過が心配です。術後の経過が心配です。	0.5845203
遺伝性貧血症があり内臓を移植したポリープがあり、一応治療するといわれた。がんの可能性をおとすと不安で仕方ない。妻がもう一度出産なので産にも言えず一人である。ローンは組めるか。がん保険は入れるか。	他院通院中の膵臓がん患者の家族より電話で問い合わせあり。膵臓、膵臓がんと診断されたが、転移に転移があり手術はできないと言われた。今は化学療法予定。治療では転移内投与はできないと言われたが、当院で転移内投与ができるかとの質問あり。	0.63119334
癌が末期、昨年8月からケートルダと抗がん剤で治療していたが、副作用が強く、入院治療も現在は状況が安定。治療はできないので、急に緩和ケアがホスピスをすすめられた。ホスピスを出ないといえないのか。父はこれを拒絶することができない。医師に余命を聞いてもよかった。	他院通院中の膵臓がん患者の家族より電話で問い合わせあり。膵臓、膵臓がんと診断されたが、転移に転移があり手術はできないと言われた。今は化学療法予定。治療では転移内投与はできないと言われたが、当院で転移内投与ができるかとの質問あり。	0.7210703
2年前に中程度の膵臓がんが診断された。術後の経過が安定しているが、PSAが上がり、手術が必要とされている。説明がつかないが、前立腺癌に近い部分の一部に転移があったとのこと。来月のPSAでホルモン療法を行うか検討したい。得たことがなかったが、進行してしまつたか、とても心配。	乳癌患者より電話相談あり。浸潤性乳がんの診断を受け、今後手術予定。術後の治療が必要とされている。本日に手術を予定しているが、今後どのような経過を辿るかわからない。術後の経過が心配です。術後の経過が心配です。術後の経過が心配です。	0.57097757

F. 健康危険情報

特になし

G. 研究発表

(発表誌名巻号・頁・発行年等も記入)

1. 論文発表

なし

2. 学会発表

なし

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