

研究の分類・属性

情報発信分野

研究の概要

第3期がん対策推進基本計画では、医療技術や情報端末が進歩し、がん患者の療養生活が多様化する中で、患者やその家族の治療上の疑問や、精神的・心理社会的な悩みも多様化していることから、拠点病院等や小児がん拠点病院のがん相談支援センターが中心となって相談支援の一層の充実をすべきことがあげられた。また、氾濫しているインターネット等での情報には、真偽の確かでない情報が含まれていることがあり、正しい情報を得ることが困難な場合があることも指摘されたように、患者や家族の疑問や不安も増大している。がん情報サービスでは、各種診療ガイドライン等のエビデンスレベルの高い情報を基に情報を作成し、提供しているが、今後も増え続けると考えられるがんや医療に関する情報に対して、患者や家族をはじめとする国民が、がんに関する正しい情報をより迅速に適切に活用できるようにするためのしくみが求められている。

そこで、本研究では、信頼を確保しつつ効率的にがんの情報をより多くの人に届けるために、新しい情報技術を用いた情報提供の自動化について検討する。

将来的には、人工知能による音声対話による相談が理想だが、その情報基盤となるがん患者・家族の情報ニーズは集積されていない。さらに、がんに関する相談は、多種多様であり、医療以外の一般製品のコールセンターでの相談対応のように、容易に一定の質問と回答に集約することは難しく、大量のデータを解析することが必要となるが、その手法は検討されていない。そこで、本研究では、(1)大量のデータより患者・家族の情報ニーズを収集し、解析する手法の検討及び(2)チャットによるがん相談システムの導入手法の検討を目的とする。

前年度までの「わが国におけるがん情報データベースのあり方と評価に関する研究 (29-A-18)」にて、患者・家族の情報ニーズの収集方法として、乳がんの電話相談の自動認識の手法を確立した。本研究では、他の種類のがんの相談についても検討を進める。また、情報ニーズの解析としては、前年度までの手作業での分析データ及び本研究で導入を検討するチャット相談データを、テキストマイニング等の手法を用いて解析し、病院を探しているのか、治療の選択での悩みなのか、不安の相談なのかに応じた的確な情報の提示手法を検討する。

さらに、新しい取り組みとして、米国 National Cancer Institute 等で導入されている、チャットによるがん相談システムの導入手法を検討する。現在、日本でのがん相談は、対面または電話、一部メールにて実施されている。チャットでの相談手法の導入は、増加しているスマートフォンでがんの情報を探す利用者への利便性を高め、相談への敷居を低くし、電話では伝えきれない詳細の情報への誘導が迅速に確実にできるメリットがあり、利用者の裾野を広げることにつながる可能性が期待できる。さらに、利用記録データは、音声による記録データよりも、より質問の意図が明確となり、情報ニーズの抽出も容易である可能性がある。しかし、一方で、チャットでの相談対応では、電話相談での対応とは異なるスキルと留意点があることが想定され、導入前に入念なマニュアルの作成が必要と考えられる。

このチャットシステムでの利用記録データは、電話相談データと共に集積可能となれば、将来的に、大規模データとなり、人工知能による音声対話による相談の情報基盤として、国民に還元する仕組みとなることが期待できる。チャットシステムは、数カ所のがん相談支援センターと連携して、導入の可否を含めて検討を進めていく。

2021 年度研究経費

7,236 千円

研究班の組織

研究者名	所属研究機関名・職名	分担研究課題名

八巻 知香子 (研究代表者) 2022/1/1 から	国立がん研究センターがん 対策研究所 がん情報提供 部患者市民連携推進室・室 長	全体統括
早川 雅代 (研究代表者) 2021/12/31 まで	国立がん研究センターがん 対策研究所 がん情報提供 部情報高度化企画室・室長	全体総括
高山 智子	国立がん研究センターがん 対策研究所 がん情報提供 部・部長	音声認識システム・テキストマイニング の検討、チャットでの相談システムの検 討
内村 祐之	東京医科歯科大学東京医科 歯科大学病院 医療情報 部・特任講師	音声認識システム・テキストマイニング の検討、チャットでの相談システムの検 討
花出 正美	がん研究会有明病院 がん 相談支援センター・看護師 長	チャットでの相談システムの検討

研究の目的と到達目標及び実績要点

全期間

(目的と到達目標)

(目的)

本研究では、信頼を確保しつつ効率的にがんの情報をより多くの人に届けるために、新しい情報技術を用いた情報提供の自動化について検討することを目的とする。具体的には、(1) 大量のデータより患者・家族の情報ニーズを収集し、解析する手法の検討及び(2) チャットによるがん相談システムの導入手法の検討を行う。将来的に、これらの仕組みで集積するデータが大規模となり、患者や家族等への人工知能による音声対話による相談に向けた基盤となることを目指す。

(到達目標)

1) 大量のデータより患者・家族の情報ニーズを収集し、解析する手法の検討

先行研究班「わが国におけるがん情報データベースのあり方と評価に関する研究 (29-A-18)」において取得した7がん種(乳がん、肺がん、大腸がん、血液がん、すい臓がん、胃がん、前立腺がん)の電話相談の音声のテキストデータおよび、2)でのチャットによるがん相談により蓄積される同がん種のテキストデータから、機械学習(AI技術の1種; Wordvec2を使用)による抽出型要約技術を用いて、大量のデータより患者・家族の情報ニーズを抽出し、病院探しをしているのか、治療の選択で悩んでいるのか、不安の相談なのかなどに応じた的確な情報の提示ができるための糸口を探索する。

また、解析データを用いて、媒体別の相談利用者を考慮して的確な情報を提示するための素材としての Frequent asked question (FAQ) を大量に作成するための手法の検討を行う。

2) チャットによるがん相談システムの導入手法の検討

チャットによるがん相談システムとは、相談員(人)がPC上で入力して相談対応するシステムを想定している。システム要件を洗い出し、市販のチャットシステムの導入とカスタマイズを軸に仕様を検討する。がん情報サービスサポートセンター及びがん相談支援センター 数施設にて導入を試みる。

チャットでの相談対応では、電話相談での対応とは異なるスキルと留意点があることが想定されることから、導入前にマニュアルを作成し、実際の導入後にその評価と改良を行う。

また、電話相談（音声相談記録）とチャット相談（テキスト相談記録）の相談の特徴（利用者属性、がん種、相談対応時間、相談内容等）を比較することにより、チャット相談のニーズを把握し、導入の可否の検討につなげる。

【1 年次計画提出後の変更点と変更の理由について】

本研究の1 年目である R2 年は、新型コロナウイルス感染症の影響により、急速にオンラインでの生活環境の導入・整備が国内で進むという事態となった。研究実施主体である国立がん研究センターがん対策情報センターがん情報サポートセンターにおいても、出勤が困難な状況が続き、電話相談対応への影響があった。また今後も同様な状況が発生することも懸念された。以上より、1) 大量のデータより患者・家族の情報ニーズを収集し、解析する手法の検討については、計画立案時当初（R2 年1 月）の、電話相談をもとにした音声相談記録を主とする情報ニーズの蓄積を図る手法の検討から、本研究計画の2) において検討を進めるチャット相談による大量のデータ蓄積を行う方針へ計画変更を行った。この方針変更に伴い、前年度継続（開発費 29-A-18）の乳がん他全7 がん種の相談記録とチャット相談記録を解析対象として、大量のデータより機械的に情報ニーズを抽出する手法を探るとともに、電話相談（音声相談記録）とチャット相談（テキスト相談記録）の相談の特徴（利用者属性、がん種、相談対応時間、相談内容等）を明らかにすることとした。また、解析データを用いて、媒体別の相談利用者を考慮して的確な情報を提示するための素材となりうる Frequent asked question (FAQ) を大量に作成するための手法の検討を行うこととした。

以上により、下線部にあげるように、到達目標の変更が生じた。

第2 年次

（到達目標）

1) 大量のデータより患者・家族の情報ニーズを収集し、解析する手法の検討

【電話相談（音声相談記録）とチャット相談（テキスト相談記録）の相談の特徴解析】

電話相談及びチャット相談各 100 件の相談記録の解析により特徴を比較解析する。

【Frequent asked question (FAQ) の作成方法の検討】

乳がん他全7 がん種の電話相談の音声のテキストデータおよび、2) でのチャットによるがん相談により蓄積される同がん種のテキストデータを用いて、Frequent asked question (FAQ) を大量に作成するための手法の検討を行う。

2) チャットによるがん相談システムの導入手法の検討

チャットによるがん相談システムを構築し、テスト稼働し、テスト参加者によるシステム評価を実施する。

【1 年次計画提出後の2 年次の到達目標の変更の理由について】

変更の理由は、1 年次の到達目標の変更の理由と同様。また1 年次の変更時点に2 年次の到達目標の変更を行った。

（年次評価時点の実績要点）

1) 大量のデータより患者・家族の情報ニーズを収集し、解析する手法の検討

2 年次計画【電話相談（音声相談記録）とチャット相談（テキスト相談記録）の相談の特徴解析】

①2020 年11 月16 月に稼働させたチャットシステムのデータ利用について、2021 年2 月15 日に研究倫理審査の許可を得て、データの活用が可能となった。

②2021 年2 月15 日以降のチャット相談データのうち、初回相談かつ1 往復以上のやりとりのある相談 100 件、がん情報サービスサポートセンターの2021 年2 月の電話相談で5 分以上の相談 100 件の特徴について比較した。

1 年次からの継続課題【音声のテキストデータおよび、2) でのチャットによるがん相談により蓄積される同がん種のテキストデータを用いて、Frequent asked question (FAQ) を大量に作成するための手法の検討】

① a) すい臓がんの電話相談の逐語録 20 件について機械学習による要約作成によって抽出された重要文、b) 同じ電話相談の逐語録から医療専門職によって手動で抽出された重要文の一致度を比較した。

② c) すい臓がんのチャット相談の文面 10 件について機械学習による要約作成によって抽出された重要文、d) 同じチャット相談の文面から医療専門職によって手動で抽出された重要文の一致度を比較した。

③ a) すい臓がんの電話相談の逐語録 20 件について機械学習による要約作成によって抽出された重要文、e) 同じ電話相談から自動音声認識によって作成した逐語録 20 件について機械学習による要約作成によって抽出された重要文を比較する準備を進めている。

2) チャットによるがん相談システムの導入手法の検討

①相談稼働後、毎日の相談件数、個人情報的一切含まない相談者の訴え、相談員の対応困難感等について、相談員が日報として報告し、その報告の特徴について分析した。

研究成果と考察

第2年次評価時点

1) チャットによるがん相談システムの導入手法の検討

※論旨のわかりやすさを優先し、「4 到達目標および実績要点」とは異なる順番で記載しています。

(1) チャット相談のテスト評価

2021年11月16日～2021年7月末までのチャット相談を担当した相談員の日報から、チャット相談の特徴についての記録を抽出・分析した。

チャット相談という手法については、相談員が相談者の質問に答える上でのチャット相談の利点として、「がん情報サービス内該当ページのアドレスを送り、治療の概要図や適応をお知らせできた。視覚的に示すことができ、提供した情報に納得していた」など【情報を視覚的に提示できる】ことが指摘された。また、相談者からの書き込みとして、「電話相談も提案されたがテキストベースの方が頭を整理できるのでチャットにした」など【文字化することで相談者の気持ちの整理につながる】ことや、「海外で不安だったので相談できてよかった」「電車の中から相談ができて良かった」など【時間や場所を問わずに相談できる】という相談者にとっての利点が指摘された。

現行のチャット相談の運用・運営の面では、「相談内容によっては、在宅でも temas で他の相談員と相談しながら対応した。バックアップ体制があることで在宅でも安心してチャット相談ができた」「チャットでは難しいと判断し、サポートセンターの電話相談に誘導した。その後の電話相談につながった」など、チームで電話相談とチャット相談の両機能を担っていることの利点を指摘する【現在の運営形態の利点】が挙げられた。一方、チャットボットに入っただけ、自動応答を経て相談員が直接対応する仕組みであるが、「チャットボットに入っただけの人が重なった場合対応ができないことがある」「待たせている間に、離れてしまったのか対応できなかったものもあった」という報告、ネット回線トラブルの影響を受けた、という【システム運営上の課題】についての報告が見られた。

(2) チャット相談と電話相談の特徴の比較

2021年2月15日以降のチャット相談および電話相談について以下の条件を満たす各100件を比較した。

チャット相談：初回相談かつ1往復以上のやりとりのある相談

電話相談：がん情報サービスサポートセンターの初回の電話相談で、5分以上の相談

相談者の属性は、電話相談では「患者本人(50%)」が多く、チャット相談では「家族・親戚(54%)」が最も多い。電話相談では60代以上が多いことがうかがえる（電話相談では相談内容で年齢が関係した場合のみ把握）のに対し、チャット相談では40歳代がピークとなる傾向があった。相談の開始から終了までの時間は電話相談が平均22.3分(range5.4-8.8)に対し、チャット相談では平均45.6分(range3.2-101.7)とチャット相談の方が長時間になる傾向があった。チャット相談での平均発話数（文面を送信した回数）は概ね10～40回であり、平均23.6回であった。

相談員が記録した主な相談内容では、電話相談では「セカンドオピニオン」「不安・精神的苦痛」が多く、チャット相談では「医療機関の紹介」「食事・服薬・入浴・運動・外出など」「介護・看護・養育」が多かった。

以上、(1)(2)より、チャット相談の導入は、相談員・相談者双方にとって、相談方法の選択肢の広がりや、相談内容の質そのものの向上にも寄与しうるものと考えられた。また、チャット相談と電話相談の利用者の属性は異なることがうかがえること、相談内容からは類似しつつも異なったニーズがあると考えられ、チャット相談に十分な利点を見出せる状況であると判断できた。また、電話相談とチャット相談を同一チームで並行して運用していくことは、チャット相談では対応が難しい相談を電話相談に誘導するなど、相互補完的な機能を果たすうえで有効と考えられた。

2) 大量のデータより患者・家族の情報ニーズを収集し、解析する手法の検討

(1) 機械的に、電話相談の逐語録から適切な要約が作成できるか？

すい臓がんの電話相談20件について、人の手で作成した逐語録（以下、人力逐語）について、機械によって抽出された重要文（以下、機械抽出）、医療専門職によって手動で抽出された重要文（以下、人力抽出）を比較した。「人力抽出」を適切な解と想定して、「機械抽出」との一致度を確認した。

「機械抽出」「人力抽出」双方が重要と判断した発話【重要一致】は14.3%、双方が不要と判断した【不要一致】は65.4%、「機械抽出」では重要と判断され、「人力抽出」では不要とされた【ノイズ】は10.2%、「機械抽出」では不要

と判断され、「人力抽出」では不要とされた【見逃し】は10.2%であった。

この結果より、機械抽出が重要と判断した4分の1（重要一致とノイズ計24.5%）のみのデータを確認した場合、本来の重要文の6割を補足できるが、4割を見逃すことになる。また、機械抽出が重要と判断したデータのみを確認しても、そのうちの4割はノイズとなる。現状では、機械抽出した重要文を要約作成に適用しても【見逃し】の割合が高いこと、抽出された重要文の約半数弱が【ノイズ】であることから、機械抽出した重要文をこのままの手法で採用することは効果的とはいえない。

ただし、今後、疾患名や薬剤名等、がん相談において重要な要素となる単語を機械にさらに学習させることにより、重要文抽出の精度を向上させる余地はある。また、適切なFAQを継続的に教師データとして提供し続けることで恒常的な精度向上のサイクルに繋げる可能性もある。さらには、本研究の前身であるがん研究開発費で行った研究では、乳がんの電話相談100件について人力逐語を作成し、複数の研究者による詳細なFAQ作成が実施されたが、そこでは、相談者の多くは冒頭に重要な質問をする割合が極めて高いことがわかっている（早川他、投稿準備中）。相談冒頭部分のやりとりについてはすべて重要文と見做す処理をする等、プログラム上の操作を加えることで【重要一致】の割合を向上させることは可能であることが見込まれる。よって、現時点で機械的に重要文抽出に関する適切な方法が確立したとはいえないが、今後発展可能な手法の手がかりを得ることができた。

なお、このプロセスにおいて明らかになった機械学習の精度向上に有用な単語、重要語を公開し、官民広く利用可能な知識として提供することにより、医療現場で有用な対話型エージェントの開発に寄与できる可能性を持てる。

(2) 機械的に、チャット相談の文面から適切な要約が作成できるか？

すい臓がん10件のチャット相談について、「機械抽出」「人力抽出」を行い、上記電話相談と同様の比較を行った。

【重要一致】は14.0%、【不要一致】は49.3%、【ノイズ】は5.7%、【見逃し】は31.0%であった。「機械抽出」が重要と判断した結果のみのデータを扱った場合、本来の重要文の4割しか拾えておらず、見逃し率が極めて高い。これは、チャット相談では文字でのやりとりという性質上、発話単位が長くなりがちであるため、人力では、重要文とそれに付随する説明を一体として抽出する一方、機械による抽出では単文での判断を行っていることによる可能性もある。

現時点では機械学習によるチャット相談の文面からの重要文抽出は、有効な方法と言えるかどうかの判断を留保せざるを得ない。しかし、ようやく蓄積できたチャット相談データについて、複数のパターンを用いて重要文の抽出を試行することで機械学習の精度を上げ、実用可能性を高められる可能性はある。チャット相談の特性の分析を進めつつ、特徴をふまえた適切な手順の探索を年度内に進める予定である。

(3) 電話相談の音声データから人的作業を経ずにどれほど重要文の抽出が可能か？

がんの電話相談の録音音声データからどれほど正確な逐語録が作成できるのかについては、前身のがん研究開発費において検討し、現状での直接利用は難しいが、教師データの洗練によって改善できる可能性が示唆されている。この前身研究班で現在も検討中であるが、「プランA：実現可能性が確認できた際に今後の人的労力の削減にどれほど貢献しうるか」、もしくは「プランB：このプロセスの実用可能性が短期間に望めない場合の別ルートの開発可能性があるのか」について検討を進めることとした。

電話相談の録音音声データから、音声認識技術によって作成した逐語録（以下、機械逐語）から「機械抽出」した重要文の該当箇所と、「人力逐語」から「機械抽出」した重要文の該当箇所の一致度を検討する作業を進めている。2022年1月中に結果がまとめられる見込みで、この結果の一致度が高い場合には、「音声認識技術」での実現可能性が低い場合でも、該当箇所をタイムスタンプで切り出し、医療専門職が聞き取ることで解決できる可能性が開ける。

上記(1)(2)(3)より、現時点で人的労力を省き効果的なFAQを多数生み出す手法について確立できたとはいえないが、FAQの効率的な生産に向けた複数の可能性を提示できた。また、これらの知見を公開していくことにより、本研究以外での実装技術の発展に寄与できる見込みが立った。

倫理面への配慮

(1) 遵守すべき研究に関する指針等

- ☐ 再生医療等の安全性の確保等に関する法律
- ☐ 臨床研究法
- ☐ 医薬品の臨床試験の実施に関する基準（GCP）
- ☒ 人を対象とする医学系研究に関する倫理指針
- ☐ ヒトゲノム・遺伝子解析研究に関する倫理指針
- ☐ 遺伝子治療臨床研究に関する指針

☐ 動物実験等の実施に関する基本指針

☒ その他の指針等（指針等の名称： ）

(2) 本研究開発期間中に予定される臨床研究の有無

☐ 有

☒ 無

本研究に関連する、本研究期間中の主な発表論文等

第2年次

(雑誌論文)

・国立がん研究センター研究開発費による成果であることが記載されているもの
記載なし

・国立がん研究センター研究開発費による成果であることが記載はないが、関連するもの

Yamaki C, Takayama T, Hayakawa M, Wakao F. Users' evaluation of Japan's cancer information services: Process, outcomes, satisfaction, and independence. BMJ Open Quality. (in press)

皆川愛、高嶋由布子、八巻知香子、平英司、高山亨太. ろう者を対象にした医療情報の翻訳における課題～がん冊子の手話動画作成を通して～ヘルスコミュニケーション雑誌第13巻第1号（受理・印刷中）

Takayama T, Inoue Y, Yokota R, Hayakawa M, Yamaki C, Toh Y. New approach for collecting cancer patients' views and preferences through medical staff. Patient Preference and Adherence. 2021;15:375-385.

Toh Y, Inoue Y, Hayakawa M, Yamaki C, Takeuchi H, Ohira M, Matsubara H, Doki Y, Wakao T, Takayama T. Creation and provision of a question and answer resource for esophageal cancer based on medical professionals' reports of patients' and families' views and preferences. Esophagus. 2021;18(4):872

(学会発表)

堀抜文香、安藤絵美子、澤井映美、早川雅代、高山智子. 脾臓がんにおいて求められる情報とサポートのあり方の検討: がん相談の記録を手がかりに、ヘルスコミュニケーションウィーク 2021～広島～, オンライン開催, 9月, 2021.

(書籍)

高山智子, 八巻知香子, 早川雅代. がん医療が問いかける新たな医療コミュニケーション. (分担執筆). 徹底研究 患者本位のがん医療. 公益財団法人 医療科学研究所. 2021. 法研. (東京)

早川雅代, 八巻知香子, 高山智子. 患者本位のがん医療の実現に向けた医療コミュニケーション環境整備の課題と展望. 徹底研究 患者本位のがん医療. 公益財団法人 医療科学研究所. 2021. 法研. (東京)

(知的財産権)

記載なし

(政策提言（寄与した指針等）)

記載なし

(その他)

記載なし