

2020-A-19 がん情報提供の自動化に向けた研究

高山 智子

がん対策情報センター がん情報提供部

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。

2020 年度研究経費

5,705 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
高山 智子 (研究代表者)	国立研究開発法人国立がん研究センターがん対策情報センター がん情報提供部・部長	全体総括
早川 雅代	国立研究開発法人国立がん研究センターがん対策情報センター がん情報提供部・室長	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 医学部附属病院 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
八巻 知香子	国立研究開発法人国立がん研究センターがん対策情報センター がん情報提供部・室長	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て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 1 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

【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。FAQ 抽出方法の文献レビューを行い、研究プロトコルに反映させる。

2) チャットによるがん相談システムの導入手法の検討

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にて導入を試みる。

また、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。

【1 年次計画計画提出後の到達目標の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、大量のデータより情報ニーズを収集・解析する手法の検討について、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。

（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

チャット相談による相談記録のデータ収集と解析を進めるために、研究計画書を作成し、倫理審査申請準備をしているところである。（2020年12月中旬に申請予定）

音声認識データの乳がん他全7がん種の相談対応記録については、機械学習（AI技術の1種）による抽出型要約技術で抽出した重要文と人手で抽出した結果を比較することで、相談対応記録における同技術の利用可能性について検討・検証を行う。現在機械的に行う要約については、業者見積を発注し、年度内に結果を得る予定である。後者の人手で行う抽出については、7がん種中、約半分にあたる4がん種を終了した。残りについては、引き続き次年度に実施する予定である。

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) 導入するチャットシステムの決定：複数のチャットシステムの情報を入手し、導入のしやすさ（初期費用・継続費用）、操作性、情報蓄積（csv等での簡易な情報抽出）、トラブル時バックアップ体制の他、頻出質問対応のカスタマイズ等の側面から、導入システムを決定した。
- (2) チャット相談導入準備：①類似の相談対応をチャットで実施している民間組織（東京メンタルヘルススクエア）より、チャット導入時の準備や実際のチャット対応に関するヒアリングを行い、情報収集を行った。また本研究導入関係者が、同組織が実施・提供するチャット相談対応方法についての研修を受講した。②国立がん研究センターの機関内で使用しているチャットシステム（Office365:Teams）を用いて、複数回の模擬相談を行い、チャット相談対応のイメージづくりと準備内容の抽出を行った。③ ①②を踏まえ、チャット相談導入のマニュアル（ver0.8）を準備した。
- (3) チャット相談窓口の準備：本研究班のホームページを開設し、「がんチャット相談」の案内と相談記録の取扱い等（個人情報を除いた記録を研究目的で使用する等）の相談利用に関する留意事項について記載した。また、国立がん研究センター「がん情報サービス」上の既存の「がん情報サービスサポートセンター（同相談員が対応する電話相談の案内）」から、「がんチャット相談」の案内のリンクを貼った。さらに、全国のがん診療連携拠点病院がん相談支援センターのメーリングリストを通じて、「がんチャット相談」の研究班での開始通知を行った。
- (4) チャット相談開始：2020年11月16日（月）より、平日12-15時で、1回線で相談窓口を開設した。相談対応は、国立がん研究センターのがん情報サービスサポートセンター内で実施することとした。相談対応の仕方、相談の対応困難時のバックアップ体制の取り方等に留意しながら、相談対応および経験を蓄積しているところである。2021年3月31日までを試行導入期間としている。12月1日（火）までの2週間（実質10日間で、12件の相談が寄せられた。その後も一日あたり1～3件の相談を受けており、2021年4月からも引き続き相談を集積予定とした。

研究成果と考察

第1年次評価時点

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

前年度継続（開発費 29-A-18）の乳がん他全6がん種（肺がん、大腸がん、血液がん、すい臓がん、胃がん、前立腺がん）の相談記録をもとに、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、また媒体別の相談利用者を考慮した Frequent asked question (FAQ) を大量に作成するための手法の検討を行うため、研究計画書を

音声認識データの乳がん他全7がん種の相談内容については、機械学習（AI技術の1種）による抽出型要約技術で抽出した重要文と人手で抽出した結果を比較することで、相談対応記録における同技術の利用可能性について検討・検証を行う。機械的に行う要約については、業者に発注を行い、年度内には結果を入手予定である。後者の人手で行う抽出については、7がん種中約半分にあたり4がん種の抽出を終了したところで、次年度についても引き続き抽出作業を実施する予定である。

初年度にあたる本年度は、チャットがん相談を開始するために、複数のチャットシステムの情報を入手し、導入のしやすさ（初期費用・継続費用）、操作性、情報蓄積（csv 等での簡易な情報抽出）、トラブル時バックアップ体制の他、頻出質問対応のカスタマイズ等の側面から、導入するチャットシステムを決定した。

本研究班のホームページ (<https://plaza.umin.ac.jp/~CanRes/system/>) を開設し、チャット相談窓口の周知の準備を行った。ホームページ上に、「がんチャット相談」の案内と相談記録の取扱い等（個人情報を除いた記録を研究目的で使用する等）の相談利用に関する留意事項について記載した。また、国立がん研究センター「がん情報サービス」上の既存の「がん情報サービスサポートセンター（同相談員が対応する電話相談の案内）」から、「がんチャット相談」の案内のリンクを貼った。さらに、全国のがん診療連携拠点病院がん相談支援センターのメーリングリストを通じて、「がんチャット相談」の研究班での開始通知を行った。

チャット相談は、2020年11月16日（月）より、平日12-15時に1回線の相談窓口を開設した。相談対応は、国立がん研究センターのがん情報サービスサポートセンター内で実施することとした。相談対応の仕方、相談の対応困難時のバックアップ体制の取り方等に留意しながら、相談対応および経験を蓄積しているところである。2021年3月31日までを試行導入期間としている。12月1日（火）までの2週間（実質10日間で、12件の相談が寄せられ、現在のところ大きなトラブルなく、チャット相談の導入が実施できた。今後、相談記録を集積していくとともに、情報基盤の仕組みとしていくために、FAQおよび情報作成の方法についての検討を進める予定である。

☐ 無

本研究に関連する、本研究期間中の主な発表論文等

第1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

Haragi M, Hayakawa M, Watanabe O, Takayama T. An exploratory study of the efficacy of medical illustration detail for delivering cancer information. J of Visual Communication in Medicine, 2020. doi: 10.1080/17453054.2020.1834838. (査読有り)

- ・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

Toh Y, Hagihara A, Shiotani M, Onozuka D, Yamaki C, Shimizu N, Morita S, Takayama T. Employing multiple-attribute utility technology to evaluate publicity activities for cancer information and counseling programs in Japan. J of Cancer Policy (in press). (査読有り)

八巻知香子, 高山智子. 信頼できるがん情報の提供と研究における患者・市民の参画の試み: 国立がん研究センターがん対策情報センター「患者・市民パネル」のこれまでの活動と今後. 科学技術社会論研究. 18; 128-136. (2020)

三輪眞木子・八巻知香子・田村俊作・野口武悟. 視覚障がい者の健康医療情報ニーズの特性と提供の際の課題. 現代の図書館. 58 巻 1 号 p. 46-51. (2020)

(学会発表)

齋藤 弓子, 堀抜 文香, 早川 雅代, 八巻 知香子, 藤 也寸志, 高山 智子. 男性がん患者から求められる性に関する情報と支援についての検討: 医療者を通じて収集した患者の語りから. 第12回日本ヘルスコミュニケーション学会学術集会(東京・オンライン) 2020/9/26.

堀抜 文香, 早川 雅代, 八巻 知香子, 藤 也寸志, 高山 智子. 臓器がん患者や家族が求める情報と環境: 医療者を通じて収集した患者の語りから. 第58回日本がん治療学会学術集会(京都) 2020/10/24.

早川 雅代, 渡部 乙女, 佐野 由美子, 酒井 由紀子, 高山 智子. 患者向け情報資料での的確に、わかりやすく伝えるための文章表現の検討. 第58回日本がん治療学会学術集会(京都) 2020/10/24.

八巻 知香子, 原田敦史, 高橋三智世, 打浪文子, 羽山慎亮, 中山真理, 皆川愛, 柴田昌彦, 平英司. 「がん情報サービス—新型コロナウイルス Q&A」アクセシブル版作成の試み. 第12回日本ヘルスコミュニケーション学会学術集会(2020. 9. 26-27)

皆川 愛, 八巻知香子, 平英司, 高嶋由布子. 言語的マイノリティとしてのろう者を対象にした手話版大腸がん資料の作成. 第12回日本ヘルスコミュニケーション学会学術集会(2020. 9. 26-27)

(書籍)

記載なし

(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)

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: がん対策基本法およびがん対策推進基本計画で進められてきた情報・支援・ネットワークの現状と課題、そして展望. 医療と社会 30(1)9-26, 2020. (査読なし)

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. 医療と社会 30(1)27-41, 2020. (査読なし)