

小児がんにおける情報の分類：網羅的情報データベースの構築を目指して新聞記事データからの検討

高尾 憲司 大阪大学大学院医学系研究科保健学専攻博士後期課程

I はじめに

わが国では、がん対策基本計画¹⁾により、がん患者の情報不足感を解消することは、喫緊の課題とされている。中山²⁾は、『国民のがん情報不足感の解消にむけた「患者視点情報」のデータベース構築とその活用・影響に関する研究』のインターネット調査により、がん情報の不足感の実態を明らかにし、山口³⁾は、全国 7,885 人のがん体験者の悩みをまとめたデータベース「静岡分類」を開発し、がん患者に対して情報処方と言われるきめ細やかな情報提供を行っている。

成人のがん患者や生存者は、がんに関するさまざまな情報不足感を持ち、解消するために情報を求めている⁴⁾⁶⁾。また情報不足感が高いほど、不安や抑うつ傾向が高く、生活の質や幸福感にもネガティブな影響を与えていることが明らかとなっている⁵⁾⁷⁾。

一方、小児期に発症するがんである小児がんでは、5 年生存率は 1960 年代においては 3 割以下であったが、現在、治療法が大幅に改善されており、約 8 割となっている。小児がんの治療経過は長く、本人とともに家族全体が、診断当初から治療全体像を理解し、さまざまな問題について前向きに取り組んでいくことが必要である。

小児がんの子どもは情報ニーズが多いほど精神的苦痛が高く、闘病生活における生活の質が低下することが報告されている⁸⁾。また病気の初期段階で、診断や予後について情報提供がされた子どもたちは、不安や抑うつが低いという報告もある⁹⁾。小児がんの子どもは、治療中からあらゆる情報を求めていることが明らかとなっている⁸⁾⁹⁾。

このように、成人のがんでは、患者・家族の情報不足感に対する対策は実施されつつある。小児では、厚生労働省のがん対策推進基本計画¹⁾によると、相談支援や情報提供の充実など、小児がん対策の 1 つとし、今後の取り組みに期待がかかる。小児がんの子どもあるいは生存者は、病気と向き合いながら、その子なりの人生を構築していくことが求められているため、小児がんにおいても体系的な情報提供が可能となるように、小児がん情報提供データベースの構築が必要である。

II 目的

本研究の目的は、小児がん情報提供データベースの構築を目指し、まず既存の小児がん関係情報の網羅的探索するために、毎日新聞社の新聞記事データベース「毎索」を用いて系統的整理を検討することである。

III 用語の定義

小児がんとは、国際小児がん分類の 12 種類とする¹⁰⁾。主に白血病、リンパ腫、脳腫瘍、神経芽腫、網膜芽細胞腫、腎腫瘍、肝腫瘍、骨腫瘍である。

IV 方法

1. 研究対象

毎日新聞社が保有している記事データベース「毎索」を用いて、キーワードを「小児がん」、日付を「全期間：1986 年以降」として検索し、検出された 1871 件の新聞記事。

新聞記事を選定した理由は、①新聞は社会の出来事の報道・解説・論評をすばやくかつ広く伝えるための定期刊行物のため情報源として信頼性が高い。②テキストマイニングソフトによる分析は、まず形態素解析（文章を文節や単語に分割する方法）が必要となり、新聞記事は質の高い校正が行われるため、精度の高い形態素解析が可能となる、からである。そのような新聞の記事データベースを用いることで、過去、現在の情報提供内容を網羅的に把握できる。

2. 調査時期

2015 年 7 月 13 日

3. 調査項目

発行年、記事内容

4. 分析方法

小児がんに関する新聞記事 1871 件から、毎日新聞社が著作権を有していない新聞記事 16 件を除いた 1855 (99.1%) 件を分析対象とした。これらの小児がん情報の内容を網羅的に探索し、系統的に分類するために、テキストマイニング手法を用いる。まず新聞記事のテキストデータは、文字列データのため、計量的に分析ができるように、すべてのデータを Excel ファイルへ入力し、形態素解析を行う。以上のことを踏まえて、以下、分析名、分析機能と目的を表 1 に示す。

表 1 テキストマイニングによる分析と目的

分析名	目的
ビジュアル集計	小児がんに関する新聞記事数の時系列による推移を確認する。
頻度分析	単語頻度や係り受け頻度を集計し、小児がんの新聞記事を概観する。
注目分析	「小児がん」を注目語として、小児がんのイメージ、小児がんに対する行動、話題一般を確認する。
評判分析	「小児がん」が社会にどのような表現で発信されているか確認する。
特徴分析	西暦を属性とし、それぞれの年代の特徴語を抽出する。

5. 倫理的配慮

1) 本調査で使用するメディア情報に関して、研究目的以外での利用は一切行わない。

- 2) データの入力に用いる専用のパソコンを準備し、ネットワークには接続せず、ネットワークから隔離された状態で管理し、パスワードを用いて管理する。
- 3) 解析作業は、部外者にはデータを開示しない。
- 4) 公表する場合は、公表内容に個人情報や著作権への問題の有無について、事前に毎日新聞社へ問合せ、利用の承諾を得る。

V 結果・考察

1. 小児がん情報の新聞記事による社会的変遷

小児がんに関する新聞記事の時系列による推移を確認した（図1）。毎日新聞社の記事データベース「毎索」は、1987年から現在までの新聞記事が収録されている。そのデータベース内では、小児がんに関する新聞記事が掲載されたのは、1988年が最初であった。同年は2件の記事があり、それから5年間、小児がんに関する新聞記事は掲載されていない。次に掲載されたのは、1994年であり、同年は3件であった。翌1995年は、記事が、掲載されていないものの、1996年から記事数が急上昇した。1988年から1995年までの8年間で掲載された記事は6件であった。

1996年から小児がんに関する新聞記事は急上昇した。2005年までの10年間は上昇傾向であり、掲載された記事は1051件（1年平均105.1件）であった。これは1996年から毎日新聞社と毎日新聞社会事業団によって、がんと闘う子供たちの実態を報告した企画記事をきっかけにキャンペーンが始まったからであると考えられる。2006年から現在までは、単年の上昇はみられるものの減少傾向であり、10年間の掲載された記事は、798件（1年平均79.8件）であった。毎日新聞社の「小児がん」に関する新聞記事を西暦によって分類すると、低位期(1988～1995年)、上昇期(1996～2005年)、減少期(2006～2015年)の3つの時期に分類することができた。

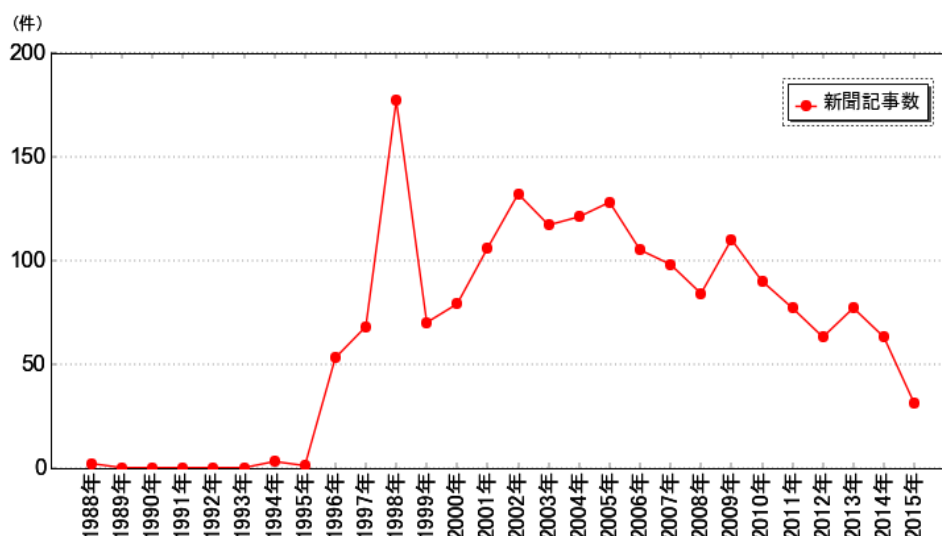


図1 「小児がん」に関する新聞記事数（年別）

2. 単語頻度解析と係り受け頻度解析

小児がんの新聞記事を概観するために、まず単語頻度解析を行った。出力する単語の品詞は、名詞（一般、固有名詞、固有名詞人名、サ変接続、形容動詞語幹）、動詞（自立）、形容詞（自立）として、上位 30 の結果を抽出し、図 2 に示した。

単語頻度解析において、名詞は、「小児がん」が最も多く、対象を表す「子どもたち」、「子ども」、「家族」といった単語が抽出された。他には「病気」、「がん」、「治療」といった医療用語、「小児がん征圧募金」、「募金」、「小児がん征圧キャンペーン」といった寄付関係の単語が多く抽出された。動詞は、「開く」、「生きる」、「戦う」、「話す」、「呼びかける」、「思う」、「受ける」といった単語が抽出された。

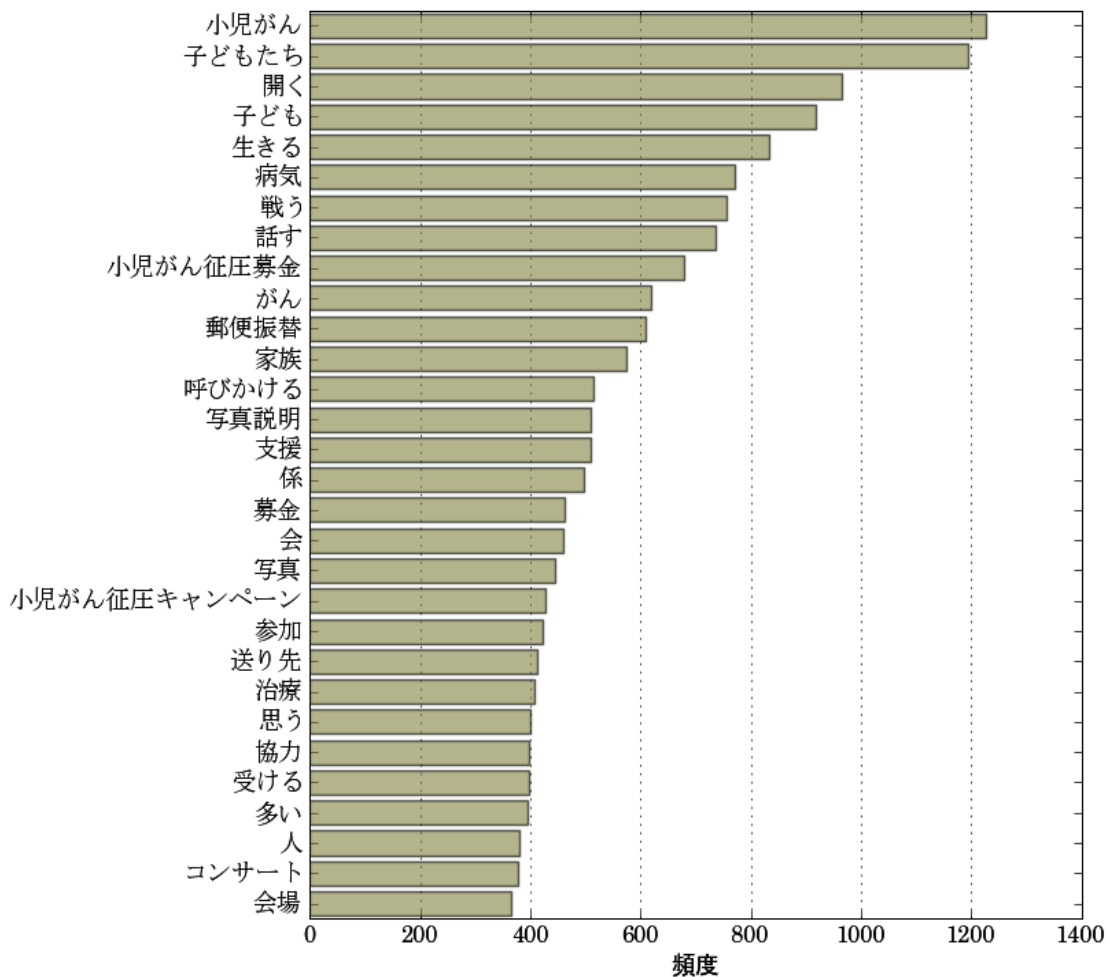


図2 単語頻度解析

次に係り受け頻度解析を行った。出力する単語の品詞は、単語頻度解析同様、名詞（一般、固有名詞、固有名詞人名、サ変接続、形容動詞語幹）、動詞（自立）、形容詞（自立）として、上位 20 の結果を抽出し、図 3 に示した。係り受けは、「戦うー子どもたち」、「小児がんー戦う」、「病気ー戦う」、「病ー戦う」のように「戦う」という行動を表現する単語が抽出されている。また「子どもー守る」、「子どもたちー支援」、「子どもたちー励ます」のように子どもたちに対する行動が抽出された。単語頻度解析と同様に「募金ー呼びかける」、「小児がん征圧募金ー呼びかける」、「小児がん征圧募金ー受け付ける」といった毎日新聞社からの寄付関係の話題が多く抽出されている。単語頻度解析と係り受け頻度解析の結果から、毎日新聞社が社会へ発信している小児がん情報の内容は、小児がんに関わる人々たちへの擁護の姿勢をもち、キャンペーンやチャリティー公演で盛んに募金活動を行い、小児がん征圧を目標としていることがうかがえた。

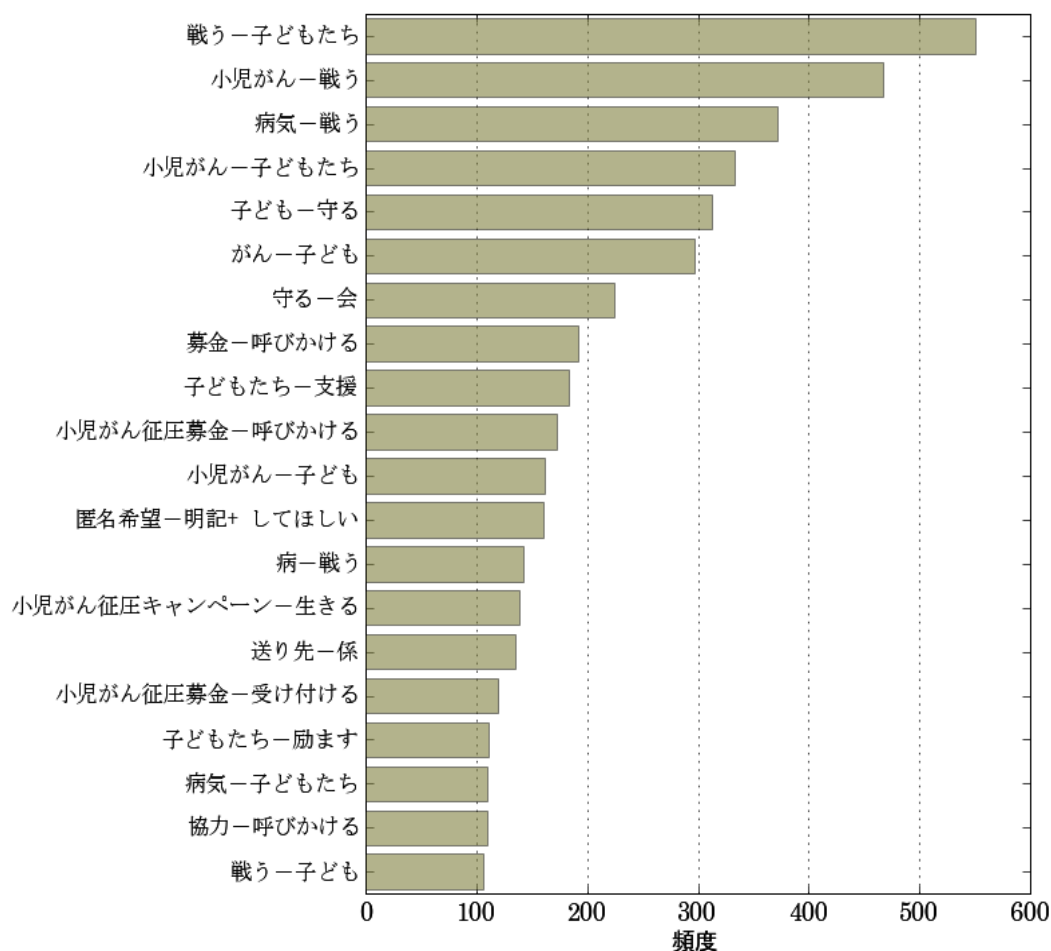


図3 係り受け頻度解析

3. 「小児がん」の注目分析

「小児がん」を注目語として設定し、行単位での抽出とし、最低信頼度 60、出現回数 2 回以上として解析した。「小児がん」を係り元の単語として係り受け頻度解析をした結果を図 4 に示す。「小児がん」と共起した形容詞、形容動詞は、「難病」、であった。また「小児がん」と共起した動詞、サ変接続名詞は、「理解」、「克服」、「治る」、「病気」「告知」、「治療」、「亡くす」、「診断」、「経験」、「発症」、「発病」、「亡くなる」、「かかる」であった。

小児がんの話題一般としてあげられる「告知」、「治療」、「診断」といった単語が抽出された。

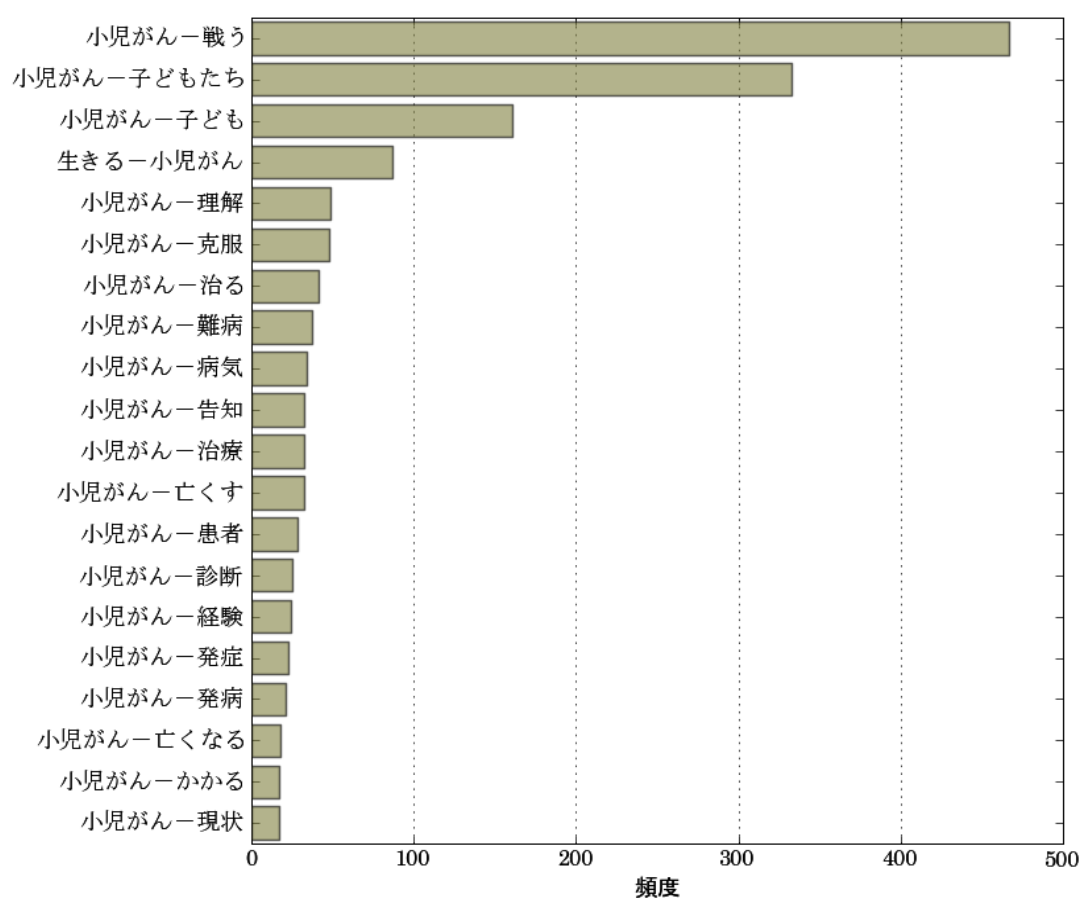


図4 注目語「小児がん」表現情報

「小児がん」のことばネットワーク図を図 5 に示す。小児がんに関するポジティブな単語は、「小児がんキャンペーン活動」は、「熱心」で、「小児がん医療」は「良くしたい」がみられた。ネガティブな単語は、「難しい」、「悪性」、「過酷」、「詳しくない」といった評価がされている。他には、「小児がん治療」は「辛い」、「小児がん患者」は「難しい」といったネガティブな単語もみられた。

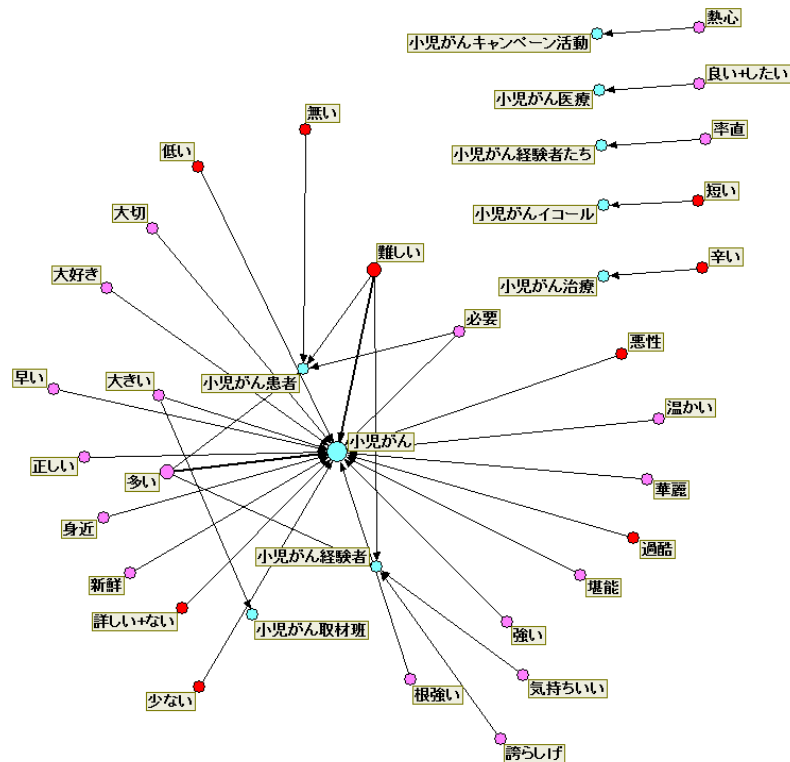


図5 「小児がん」のこトバネットワーク

4. 特徴分析

各年代の小児がんに関する情報の特徴をみるために、小児がんに関する新聞記事が上昇傾向であった1996年から2005年の10年間を「第1期」、2006年から現在までの10年間を「第2期」とした。これらを属性として、それぞれの年代の特徴語を抽出した。出力する単語の品詞は、名詞（一般、固有名詞、固有名詞人名、サ変接続、形容動詞語幹）、動詞（自立）、形容詞（自立）として、抽出指標は、カイ2乗値した。上位20の結果を抽出し、図6、図7に示した。

第1期（1996-2005年）の特徴語は、「骨髄移植」である。小児造血幹細胞移植の歴史は1970年代の同種骨髄移植が始まりとされている。1991年に財団法人骨髄移植推進財団設立、翌1992年には、同財団のドナー・患者登録の受付が開始された。1997年には日本骨髄バンクによる骨髄移植1000例を超え、2003年までに5000例実施され、移植医療が飛躍したことが考えられる。他には、渡哲也、館ひろし、神田正輝といった俳優の人名が抽出された。これは、毎日新聞社が、キャンペーンイベント「生きる-小児がんの子どもたちとともに」を定期的で開催している。その協力団体として株式会社石原プロモーションがあるため、所属の主要俳優陣が抽出されたと考える。

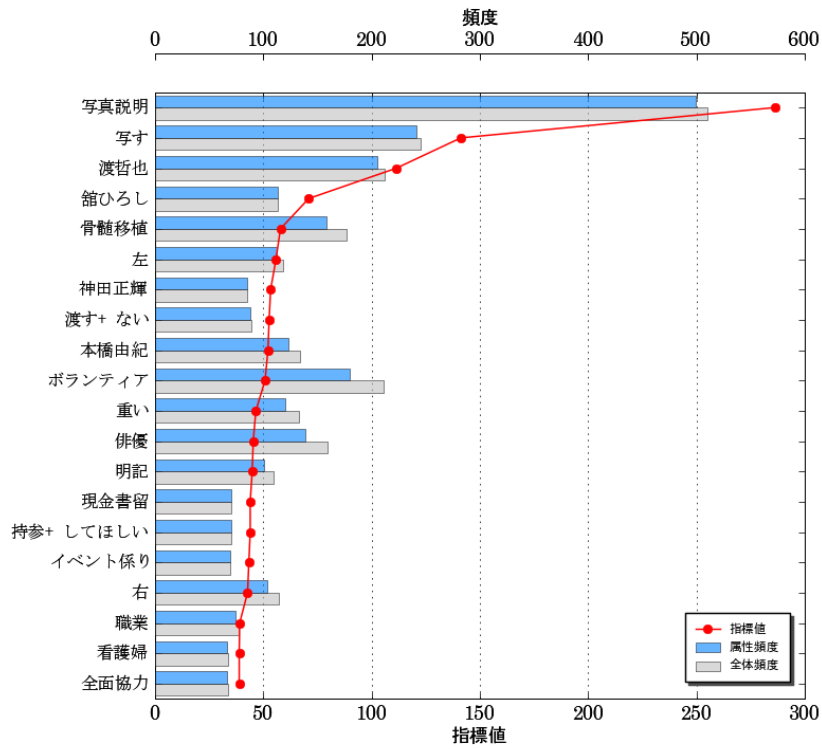


図6 特徴語抽出(1996-2005年)

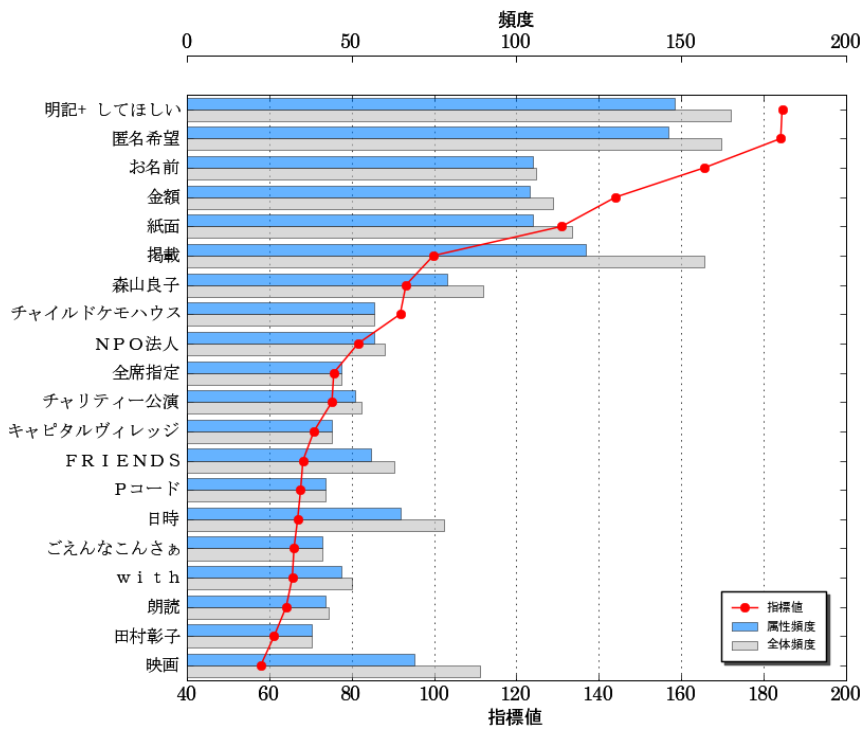


図7 特徴語抽出(2006-2015年)

第2期(2006-2015年)の特徴語は、「NPO法人」である。小児がんに関係している主なNPO法人を表2に示した。設立年が2006年度以降の法人が多く、小児がんの子どもや家族を支援するNPO法人の増加傾向であることが考えられる。その中でも「チャイルド・ケモ・ハウス」が上位である理由は、2013年には、小児がん治療中の子どもたちとその家族のQOL(Quality Of Life-生活の質)に配慮した日本で初めての専門治療施設「チャイルド・ケモ・ハウス」が完成したことが考えられる。

他には、チャリティー公演として、歌手森山良子の「生きる2014～小児がんなど病気と闘う子どもたちとともに～森山良子 with FRIENDS」や、女優の竹下景子が中心人物となり活動している「ごえんなこんさあと」が多く抽出されたと考えられる。

VI 結論

今回の毎日新聞社の新聞記事を用いて、テキストマイニングによる分析から得られた小児がんに関する情報は、以下の4点である。

1. 毎日新聞社の「小児がん」に関する新聞記事を西暦によって分類すると、低位期(1988～1995年)、上昇期(1996～2005年)、減少期(2006～2015年)の3つの時期に分類することができた。
2. 1996年以降、毎日新聞社は芸能人などの協力を得て、多数のキャンペーンやチャリティー公演を開催し、小児がん征圧の募金活動を行っている。
3. 「小児がん」に係る単語は、ポジティブ評価とネガティブ評価の双方がみられた。
4. 注目分析によって、「小児がん」に関する一般的な話題を抽出することができた。また西暦の年代別で情報を分類することで、小児がんの子どもや家族を支援するNPO法人や芸能関係に関する情報を得ることができた。

VII 研究の限界と今後の課題

今回は、大手新聞社の1社に限定した結果のため、今回のような分類が他の新聞社でも可能かどうか、また新聞だけでなく、その他のメディアではどのような分類となるか、今後さらに対象を増やして、網羅的な検討が必要であると考えられる。

引用文献

- 1) 厚生労働省 . がん対策基本計画 . 2014 .
http://www.mhlw.go.jp/bunya/kenkou/dl/gan_keikaku02.pdf. (2015年10月13日アクセス可能)
- 2) 中山健夫 . 「がん情報の不足感」実態調査 .
<http://www qlife.jp/cancer/category/anguish/paucity2013>. (2015年10月13日アクセス可能)
- 3) 山口建 他. がん体験者の悩みや負担等に関する実態調査報告書 概要版. 2009.
- 4) Husson O, Mols F, Oranje WA, et al. Unmet information needs and impact of cancer

in (long-term) thyroid cancer survivors: results of the PROFILES registry. *Psycho-Oncology* 2014;23(8):946-52.

- 5) Kent EE, Arora NK, Rowland JH, et al. Health information needs and health-related quality of life in a diverse population of long-term cancer survivors. *Patient Education & Counseling* 2012;89(2):345-52.
- 6) Halkett GK, Kristjanson LJ, Lobb E et al. Information needs and preferences of women as they proceed through radiotherapy for breast cancer. *Patient Education & Counseling* 2012;86(3):396-404.
- 7) Tsuchiya M, Horn SA. An exploration of unmet information needs among breast cancer patients in Japan: a qualitative study. *European Journal of Cancer Care* 2009;18(2):149-55.
- 8) Mico` l E, Gianinazzi, Stefan Essig, Corina S. Rueegg et al. Information Provision and Information Needs in Adult Survivors of Childhood Cancer. *Pediatr Blood Cancer* 2014;61:312-318.
- 9) Last BF, van Veldhuizen AM. Information about diagnosis and prognosis related to anxiety and depression in children with cancer aged 8-16 years. *Eur J Cancer* 1996 ;32A(2):290-4.
- 10) Eva Steliarova • Foucher, et al. *Cancer* 2005;103:1457-67.