

Table 2 Conversations regarding treatment* (Continued)

Conversation 10 (advice on treatment)	user17	Dearest (user21), good morning. You are now being treated with Xeroda. It's been just a few days, right? Sorry if I am wrong but it may take some time for the drug to take effect.
	user21	Dear (user17), good morning! Oh, Xeroda. Well, if left for 2-and-a-half months without chemotherapy, that seems rational. (';ω;') Internal medicine apparently works slowly.(';ω;') It will take time, too. (';ω;')

*Japanese conversations were translated into English.

To better understand how cancer patients influence their followers via Twitter, direct investigation involving the use of a survey of cancer patients with Twitter accounts may be necessary in the future. Examining the distribution of user activities did not reveal any significant differences among the different types of cancer noted in users' profiles. On the other hand, our study showed that a smaller number of extremely active accounts existed for each type of cancer examined (Figure 3). Under the hypothesis that such small numbers of active users serve as the center of the patients' networks on social media, we investigated the connections related to the most active users. As a result, we were able to demonstrate that information was exchanged in real time among patients (Figure 4). Based on this finding, we were able to demonstrate for the first time that an information exchange network among patients via social media had already been established.

Of further interest to us is the content of the tweets exchanged among patients. Most of the examined tweets included details of daily life such as greetings or messages concerning treatments, and it was found that almost no medical information concerning cancer was exchanged; this went against our initial expectation that cancer patients would use Twitter to primarily discuss specific cancer-related news and medical information.

Our findings demonstrate that patients use Twitter as a tool of psychological support by being connected among patients, even though it is not a standard or face-to-face method of discussing such information. This observation may support the notion that Twitter plays a unique role that is different from similar-seeming Internet tools such as hospital websites in which patients primarily obtain medical information [2] or blogs in which patients can share their experiences [1]. We expect that as Twitter usage becomes more widespread in the coming years, there will be an attendant rise in the medium's importance to maintaining—and perhaps improving—public health [25]. However, the dissemination of Twitter among patients in the future may generate various methods of usage, making it necessary to continue careful observation in the future.

Twitter can be used not only with real names but also anonymously, which is often controversial. In our study, 53% of the accounts included the users' real names, and 41% of the accounts included personal pictures. In-

vestigation into the Twitter accounts of physicians revealed that 78% of these accounts displayed the users' real names and personal pictures [12], indicating that anonymity is more preferred among patients than physicians. We expect that this discrepancy can be correlated to the fact that information about individuals' medical conditions is considered personal and confidential, and that revealing a Twitter account user's name could lead to the disclosure of potentially private medical details. Many people consider it necessary to maintain anonymity when sharing information through Twitter and other social media; such anonymity may be linked to Twitter's ability to maintain its relevance among the patient populations that use it.

Limitations

While this study demonstrated that a patient network via Twitter is in the process of being established, there remain several issues to be discussed. First, this study targeted only those Twitter users who described "cancer" either in Japanese Hiragana, Katakana or Kanji letters in their profiles. However, this does not mean that all users who were cancer patients included relevant disease names in their profiles; the absence of cancer details in user profiles could potentially exclude an unknown number of cancer patients from analysis.

Second, because of limitations in search tool performance, we were unable to conduct a large-scale comprehensive qualitative analysis. It is expected that the improvement of search-tool performance will enable larger-scale studies in the future.

Finally, future research into this field of study will need to clarify the types of information most often disseminated via social media. It has been reported that social media often include information that is not necessarily beneficial to the health of media users [26]. Furthermore, Chretien et al. (2011), who studied physicians' accounts on Twitter, stated that there existed, although rarely, some ethically problematic content, which could possibly violate the patient privacy [12].

Twitter and other forms of social media can prove quite useful in permitting the rapid and timely dissemination of health-related information. However, as social media continue to evolve, they will need to find ways to provide relevant health information without obstructing patient privacy or delivering inappropriate content.

Overcoming this point will be an important element in the dissemination of medical information via social media.

Conclusions

Twitter users with a variety of types of cancer have proved influential on their followers, as demonstrated through the information exchange engaged in by account owners and their followers. Twitter represents a timely and low-cost medium for cancer patients and others seeking information about specific medical conditions, but our study found that the majority of the tweets posted by the 51 users with “power accounts” focused on conversational details (e.g., greetings, cancer treatments) and psychological support rather than the expected medical news and information. Furthermore, Twitter will need to evolve further in order for patients to fully embrace the power of this social medium, as many people are reluctant to reveal personal details via their Twitter accounts. Our study has demonstrated that Twitter is a powerful medium capable of connecting cancer patients via the establishment of a patient network.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YS and HN designed the study, provided the study materials, collected and assembled the data and wrote the manuscript. YS, HN, AH, LS, KO and AF analyzed and interpreted the data. All authors reviewed and approved the manuscript.

Acknowledgments

We declare no conflicts of interest in association with this study. We are grateful to Drs. Atsushi Tsuya and Atsushi Tanaka (Graduate School of Science and Engineering, Yamagata University) for providing technical advice and useful discussions. This work was supported by a Research Grant (H24-26, 24200701 to Dr. Narimatsu) from the Ministry of Health, Labour and Welfare of Japan (<http://www.mhlw.go.jp/seisakunitsuite/bunya/hokabunya/kenkyujigyou/>).

Author details

¹Department of Medical Informatics, Graduate School of Medical Science, Yamagata University, Yamagata, Japan. ²Advanced Molecular Epidemiology Research Institute, Faculty of Medicine, Yamagata University, Yamagata, Japan. ³Department of Public Health, Yamagata University Graduate School of Medicine, Yamagata, Japan.

Received: 23 October 2012 Accepted: 21 December 2012
Published: 27 December 2012

References

1. Narimatsu H, Matsumura T, Morita T, Kishi Y, Yuji K, Kami M, Komatsu T, Tanaka Y, Sawa T, Nakata Y: Detailed analysis of visitors to cancer-related web sites. *J Clin Oncol* 2008, **26**:4219–4223.
2. Morita T, Narimatsu H, Matsumura T, Kodama Y, Hori A, Kishi Y, Kusumi E, Hamaki T, Kobayashi K, Yuji K, Tanaka Y, Nakata Y, Kami M: A study of cancer information for cancer patients on the internet. *Int J Clin Oncol* 2007, **12**:440–447.
3. Chen X, Siu LL: Impact of the media and the internet on oncology: survey of cancer patients and oncologists in Canada. *J Clin Oncol* 2001, **19**:4291–4297.
4. Signorini A, Segre AM, Polgreen PM: The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS One* 2011, **6**:e19467.
5. Machi user local (in Japanese); [<http://machi.userlocal.jp/>].
6. Mandavilli A: Peer review: trial by Twitter. *Nature* 2011, **469**:286–287.
7. Lulic I, Kovic I: Analysis of emergency physicians' Twitter accounts. *Emerg Med J* 2012, [epub ahead of print].
8. Heavilin N, Gerbert B, Page JE, Gibbs JL: Public health surveillance of dental pain via Twitter. *J Dent Res* 2011, **90**:1047–1051.
9. Kubben PL: Twitter for neurosurgeons. *Surg Neurol Int* 2011, **2**:28.
10. Opelka FG: Social networking in health care: surgeons and their patients. *Surgery* 2011, **150**:10–12.
11. Rajani R, Berman DS, Rozanski A: Social networks—are they good for your health? The era of Facebook and Twitter. *Q J Med* 2011, **104**:819–820.
12. Chretien KC, Azar J, Kind T: Physicians on Twitter. *JAMA* 2011, **305**:566–568.
13. De la Torre-Diez I, Diaz-Pernas FJ, Anton-Rodriguez M: A content analysis of chronic diseases social groups on Facebook and Twitter. *Telemed J E Health* 2012, **18**:404–408.
14. Desai T, Shariff A, Shariff A, Kats M, Fang X, Christiano C, Ferris M: Tweeting the meeting: an in-depth analysis of Twitter activity at Kidney Week 2011. *PLoS One* 2012, **7**:e40253.
15. Fortinsky KJ, Fournier MR, Benchimol EI: Internet and electronic resources for inflammatory bowel disease: a primer for providers and patients. *Inflamm Bowel Dis* 2012, **18**:1156–1163.
16. Thackeray R, Neiger BL, Smith AK, Van Wagenen SB: Adoption and use of social media among public health departments. *BMC Public Health* 2012, **12**:242.
17. Schneider A, Jackson R, Baum N: Social media networking: Facebook and Twitter. *J Med Pract Manage* 2010, **26**:156–157.
18. Vance K, Howe W, Dellavalle RP: Social internet sites as a source of public health information. *Dermatol Clin* 2009, **27**:133–136. vi.
19. Foundation for Promotion of Cancer Research: Cancer statistics in Japan 2010; [http://ganjoho.jp/public/statistics/backnumber/2010_en.html].
20. 16 (one-six) Profile Search β Version for twitter (in Japanese); [<http://www.16ps.jp/>].
21. Mentionmapp; <http://mentionmapp.com/>.
22. whotwi (in Japanese); [<http://whotwi.com/>].
23. Between (in Japanese); <http://between.com/>.
24. Ministry of Internal Affairs and Communication Research on Telecom Usage Trend 2010 (in Japanese); http://www.soumu.go.jp/main_content/000114508.pdf.
25. Christakis NA, Fowler JH: The spread of obesity in a large social network over 32 years. *N Engl J Med* 2007, **357**:370–379.
26. Sajadi KP, Goldman HB: Social networks lack useful content for incontinence. *Urology* 2011, **78**:764–767.

doi:10.1186/1756-0500-5-699

Cite this article as: Sugawara et al.: Cancer patients on Twitter: a novel patient community on social media. *BMC Research Notes* 2012 5:699.

Submit your next manuscript to BioMed Central
and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



がん患者とツイッター

ーソーシャルメディア上の新しいがん患者コミュニティー

○菅原祐也（山形大学医学部教育研究支援センター）、成松宏人（先端分子疫学研究所）、邵力（公衆衛生学講座）、大谷勝実（公衆衛生学講座）、深尾彰（公衆衛生学講座）

【目的】2006年にインターネット上にTwitter（ツイッター）というマイクロブログサービスが登場した。Twitterのユーザーは「ツイート」と呼ばれる140字以内の短い文を発信する。ツイートは「フォロワー」と呼ばれるそのユーザーのツイートの読者に送信され、フォロワーもツイートを返信することで双方向性のコミュニケーションを迅速に行うことができる。Twitterは、がんに関する情報を求めている患者にとってもその迅速性、双方向性を活かした有用な情報交換ツールになると期待されるが、その利用実態を明らかにした研究はない。そこで、本研究ではTwitterを利用しているがん患者の数とコミュニティの形成について明らかにすることとした。

【方法】Twitter上のがん患者のアカウント数を計測するため、プロフィールにがんの名称が記載されているアカウントの抽出を行った。Twitter上のユーザー同士の「つながり」（ユーザーを指定したツイートの返信を行って会話することと定義する）については、ウェブサイト上の検索ツールを利用し検討した。対象は日本人のTwitterユーザーで性別、年齢、地域は問わないこととした。

【結果】プロフィールに「がん」の記載があるアカウント数は731であった。そのうち患者本人のアカウント数は466であった。疾患別では乳がん147、白血病59、結腸・直腸・大腸がん40、子宮がん39、悪性リンパ腫33、その他148であった（図1）。患者アカウントが最も多かった乳がん患者のアカウントで、有名人のアカウントを除外しフォロワー数が最も多かったアカウント名“user0”（フォロワー：2463）を抽出した。検索を行ったところ、“user0”と複数のがん患者との間で「つながり」が認められた。その「つながり」の相手は、乳がん患者3人、子宮がん患者1人、がん患者と推測されるユーザー1人であった（図2）。

【結論】Twitter上に、多くのユーザーにフォローされ発言に影響力があると考えられるアカウントを中心としたがん患者同士の情報交換ネットワークが存在していることが示された。

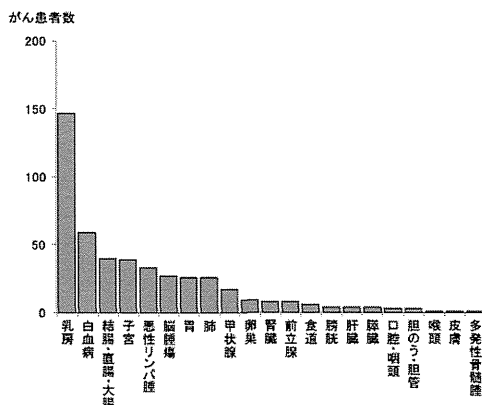


図1 がん別アカウント数

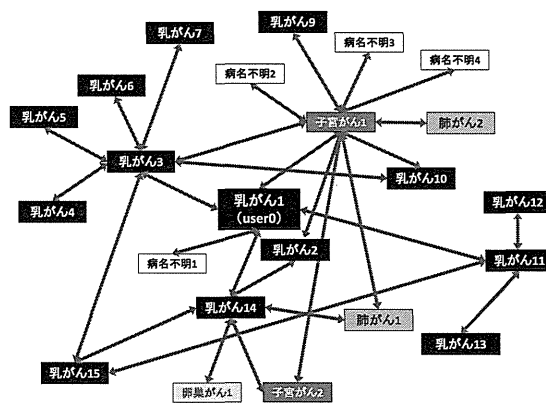


図2 ユーザー同士のつながり

医療機関とツイッター :

ソーシャルメディアを用いた新しい医療広報の可能性

○菅原祐也 (山形大学医学部教育研究支援センター),
成松宏人 (公衆衛生学講座), 深尾彰 (公衆衛生学講座)

【目的】2006年にインターネット上にTwitter(ツイッター)というマイクロブログサービスが登場した。Twitterのユーザーは「ツイート」と呼ばれる140字以内の短い文を発信する。ツイートは「フォロワー」と呼ばれるそのユーザーのツイートの読者に送信され、フォロワーもツイートを返信することで双方向性のコミュニケーションを迅速に行うことができる。医療機関がTwitterを活用すれば、医療情報の広報や市民に対する健康に関する啓発に有用なツールとなることが期待される。例えば、Twitterを利用することで医療機関は直接市民へ情報を発信し、即座に市民から質問などを受けることで、双方向の情報交換が可能となる。この情報交換によって他の診療では得られないような患者・市民の医療への要望、ニーズなどを知ることができよう。しかし、医療機関のTwitter利用は現状では限られており、その実態は明らかになっていない。そこで、本研究では医療機関におけるTwitterの利用実態を知り、医療広報におけるツイッターの役割を明らかにすることとした。

【方法】Twitter上の医療機関のアカウント数を計測するため、プロフィールに「病院」、「医院」、「クリニック」、「診療所」の記載があるアカウントを検索し、その結果から医療機関のアカウントを抽出した。医療機関のアカウントを診療科別に分類した。

【結果】Twitter上の医療機関のアカウントは168あり、そのうち歯科のアカウント73、皮膚科12、美容外科、内科、総合病院がそれぞれ10であった。ツイートの内容は歯科・口腔外科で医学的知識に関するツイート(医学)が8、診療内容に関するツイート(診療案内)が10、美容外科・形成外科・皮膚科では医学2、診療案内14であった。外科系診療科では医学7、診療案内5、内科系診療科では医学4、診療案内3であった。

【結論】Twitterは広告として使われており、かつ、診療科によって活用の程度に違いがある。

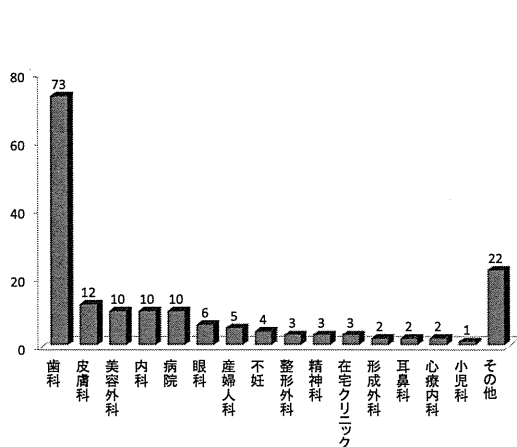


図1 診療科別アカウント数

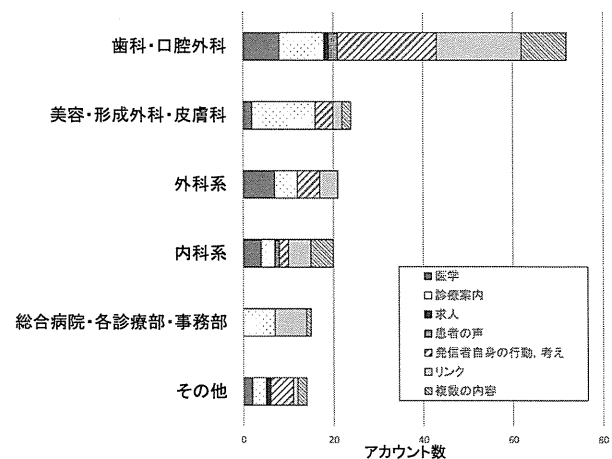


図2 ツイートの内容分類

語の共起ネットワークから見たがん患者の Twitter ツイート

Tweets on Twitter among Cancer Patients from Word Association Network Point of View

○津谷 篤¹, 成松 宏人², 田中 敦¹

○Atsushi Tsuya¹, Hiroto Narimatsu², Atsushi Tanaka¹

¹ 山形大学大学院理工学研究科

¹ Graduate School of Science and Engineering, Yamagata University

² 山形大学大学院医学系研究科

² Graduate School of Medical Science, Yamagata University

Abstract: Twitter is a microblogging service where users can post short messages limited to 140 character length as their tweets. It is said that Twitter users are about 190 million all over the world, and there can be seen some posts by users suffering from some disease. Their tweets include description of struggles against illness, that of medical treatments, their impressions and their own thoughts against the disease. These posts are considered to be informative to users with the same disease. In this paper, concentrating of the users whose profiles include specific disease name, it is analyzed using word association network what kind of tweets are communicated about the disease.

1. はじめに

Twitter はユーザが「ツイート=つぶやき」として 140 文字以内の短い文を投稿するマイクロブログサービスである。Twitter のユーザ数は世界で 1 億 9000 万人と言われており、その中には何らかの病気を患うユーザの投稿も見られる。そのツイート中では投稿者の闘病の様子や、治療の様子、治療を受けた感想、その病気に対する自身の考えが記されている。それら投稿は、同じ病気にかかっているユーザの参考となっていると考えられる。本研究では、プロフィール上に特定のがん名が記されているツイッターユーザのツイート中で、そのがんに関してどの様なツイートがされているかを、語の共起ネットワーク作成を介して明らかにする。

2. 関連研究

がん患者にとって自身の病気に関する情報は重要であり、実際の診療で得られる情報の他、インターネットから情報を得ていることが報告されている。

がん患者のインターネット利用に関する研究として以下のものがある。

カナダのあるがんクリニックを受診している患者を対象とした調査では、調査対象の 50% がインターネットから自身の疾患に関する情報を得ていた[1].

菅原らは Twitter を利用している日本のがん患者の数とコミュニティの形成について調査を行っている[2]. ツイッターのプロフィール検索サイトを用いてがん患者数を調査したところ患者本人であるアカウントは 466 であり、疾患別では乳がん 147, 白血病 59, 結腸・直腸・大腸がん 40, 子宮がん 39, 悪性リンパ腫 33, その他 148 であった。そのうちフォロワー数が 500 以上のアカウントは 52 あり、そのうち最もフォロワー数が多いアカウントは乳がん患者であり、そのアカウントと他のがん患者とのつながりが確認された。

さらに菅原, 津谷, 成松, 田中, 深尾ら筆者らは各がんに関するプロフィール記述があるツイッターアカウントのツイート内容に着目した。まず, 各ア

カウントをノード、ツイート内で同じがんアカウント同士でのリツイートを行った関係をリンクとしてネットワークを作成した。すると、他のネットワークが直線的なネットワーク構造だったのに対し、子宮がん, 乳がん, 白血病アカウントのネットワークは他のネットワークに比べ複雑なネットワーク構造を見せた (図 1)。

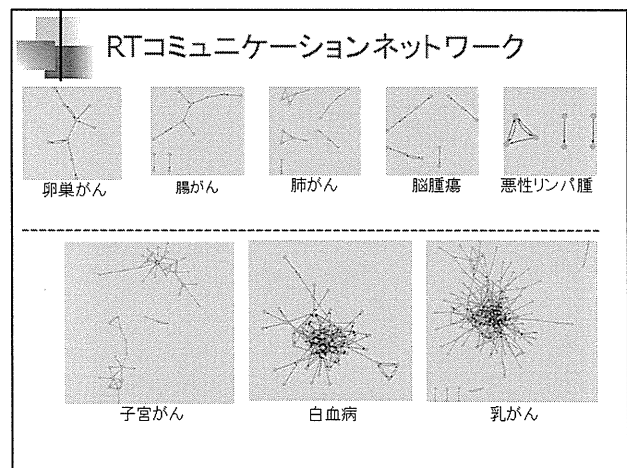


図 1: ツイート内でリツイートを行ったアカウント同士をリンクで結んだネットワーク

続いて各がんアカウントのツイートを形態素分解し、頻出単語のうち医師が選んだ名詞, 形容詞, 未知語の共起頻度を調べ, 各形態素をノード, 共起頻度が高い形態素同士をリンクとして共起ネットワークを作成した。その結果, いくつかのがんで, がんの症状が出る体部位を示すノードにリンクが集中する傾向が見られた[3]. 乳がん共起ネットワークでは, 「副作用」という語が多く他の語とリンクしていた。その理由として, 乳がんには化学療法として「抗がん剤」「ホルモン剤」がよく効くが, それぞれ副作用があり比較されやすいからであることがわかった。そして, 「手」が多くの様々な語とリンクしていた。この理由としては, 自分で「手」で触り乳がんをチェックできること, 乳がん発症後と手術後, 「手」が

上がりにくくなることが挙げられる (図 2). 肺がん共起ネットワークでは「目」が多くの語とつながっていた。実は、瞳孔縮小, 眼の凹み, 目瞼下垂など, 肺がんの症状は目に出ることが確認されている。これが理由となり「目」が多くの語とリンクした (図 3).

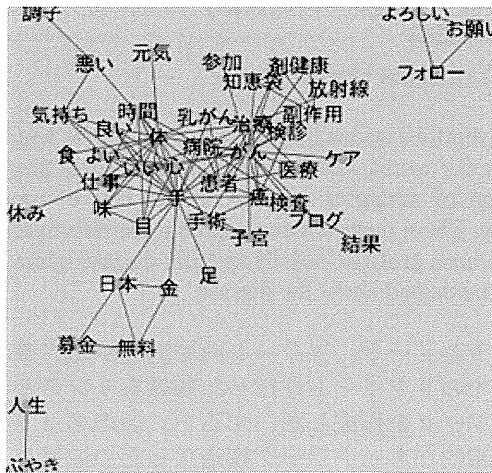


図 2: 乳がんの共起ネットワーク

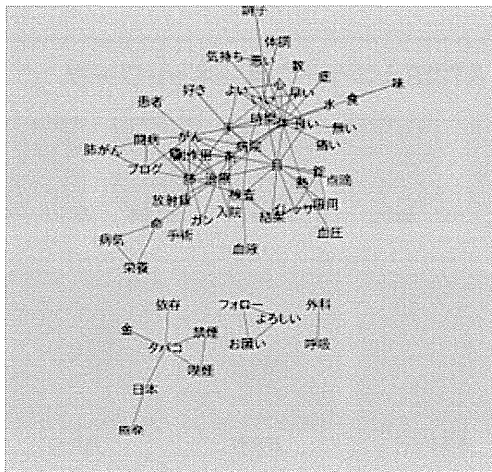


図 3: 肺がんの共起ネットワーク

また, 食道がんの共起ネットワークでは食べ物のネットワーククラスター(かたまり)が見られた。これは, 食道がんになった場合, 食事がうまくとれない, メニューなどに工夫が必要であることが理由として挙げられる (図 4).

これ以外に, 疾病に対し語彙ネットワークを用いた研究として以下の2つがある。

成田らは SNS 患者アカウント書き込み内容に共起ネットワーク図示を用いて摂食障害の調査している[4].

中川らは, がん用語をノード, がん用語間で連想されるもの同士をリンクとして結び, いわば連想ネットワークを作成し, がん用語辞書を作成している [5].

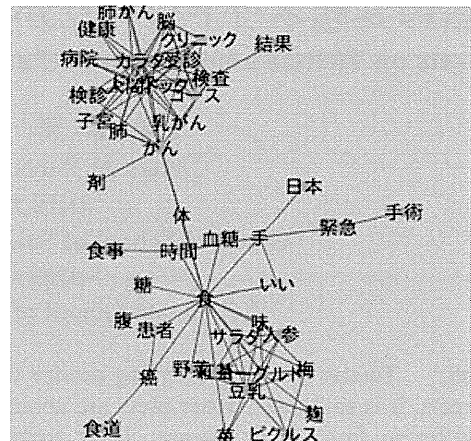


図 4: 食道がんの共起ネットワーク

3. 目的

本研究では[3]の共起ネットワークを用いたがんアカウントツイートの内容把握をさらに進める。ツイートを単語に分解するには形態素解析ソフト茶筌, データ処理にはプログラミング言語なでしこ, ネットワーク描画ソフトウェアとして Cytoscape を用いる。

本研究で共起ネットワークを作成しツイート内容を把握するにあたり現段階では次のような問題点がある。

- ・ 語彙ネットワークに用いる語の選択が医師に委ねられている
- ・ 同義語への対処
- ・ 医療に全く関係ないツイートへの対処

これら問題点にできる限り対処し発表に臨む。

参考文献

- [1] Chen X, Siu LL: Impact of the media and the internet on oncology: survey of cancer patients and oncologists in Canada, Journal of Clinical Oncology, Vol.19, Issue 23, pp4291-4297, 2001
- [2] Yuya Sugawara, Hiroto Narimatsu, Atsushi Hozawa, Li Shao, Katsumi Otani and Akira Fukao: Cancer patients on Twitter: a novel patient community on social media, BMC Research Notes
- [3] 菅原祐也, 津谷篤, 成松宏人, 田中敦, 深尾彰: がん患者とツイッター, 日本感性工学会大会予稿集, 2012
- [4] 成田彩夏, 往住彰文: SNS コメントの大量分析による摂食障害の特徴づけ, 日本感性工学会大会予稿集, 2010
- [5] 中川晋一, 内山将夫, 三角真, 島津明, 酒井善則: コーパスに基づくがん用語集合の作成と評価, 自然言語処理, Vol.16, No.2, pp3-44, 2009

論文

語の共起ネットワークから見たがん患者の Twitter ツイート

○津谷 篤¹ , 菅原 祐也² , 田中 敦¹ , 成松 宏人³

¹山形大学大学院理工学研究科, ²山形大学医学部, ³山形大学大学院医学系研究科

Tweets on Twitter among Cancer Patients from Word Association Network Point of View

○Atsushi Tsuya¹, Yuya Sugawara², Atsushi Tanaka¹, Hiroto Narimatsu³

¹ Graduate School of Science and Engineering, Yamagata University, 4-3-16 Jonan, Yonezawa-shi, Yamagata, Japan

² Yamagata University Faculty of Medicine, 2-2-2 Iidanishi, Yamagata-shi, Yamagata, Japan

³ Department of Public Health, Yamagata University Graduate School of Medicine, 2-2-2 Iidanishi, Yamagata-shi, Yamagata, Japan

Abstract: Twitter is a microblogging service where users can post short messages limited to 140 character length as their tweets. It is said that Twitter users are about 190 million all over the world, and there can be seen some posts by users suffering from some disease. Their tweets include description of struggles against illness, that of medical treatments, their impressions and their own thoughts against the disease. These posts are considered to be informative to users with the same disease. In this paper, concentrating of the users whose profiles include specific disease name, it is analyzed using word association network what kind of tweets are communicated about the disease.

Keywords: Twitter, Co-occurrence, Network analysis

1. はじめに

Twitter はユーザが「ツイート=つぶやき」として 140 文字以内の短い文を投稿するマイクロブログサービスである。Twitter のユーザ数は世界で 1 億 9000 万人と言われており、その中には何らかの病気を患うユーザの投稿も見られる。そのツイート中では投稿者の闘病の様子や、治療の様子、治療を受けた感想、その病気に対する自身の考えが記されている。さらに、その病気の治療にあたる医師の患者に対する情報提供をねらいとしたツイートも存在する。それら投稿は、現在同じ病気にかかっているユーザの参考となっていると考えられるだけでなく、今後の応用を見越したビッグデータとしても貴重なものである可能性がある。本研究では、プロフィール上に特定のがん名が記されているツイッターユーザのツイート中で、そのがんに関してどのようなツイートがされているかを、語の共起ネットワーク作成を介して明らかにする。

2. 関連研究

がん患者にとって自身の病気に関する情報は重要であり、実際の診療で得られる情報の他、インターネットから情報を得ていることが報告されている。

がん患者のインターネット利用に関する研究として以下のものがある。

カナダのあるがんクリニックを受診している患者を対象とした調査では、調査対象の 50% がインターネットから自身の疾患に関する情報を得ていた[1]。

菅原らは Twitter を利用している日本のがん患者の数とコミュニティの形成について調査を行っている[2]。ツイッターのプロフィール検索サイトを用いてがん患者数を調査したところ患者本人であるアカウントは 466 であり、疾患別では乳がん 147, 白血病 59, 結腸・直腸・大腸がん 40, 子宮がん 39, 悪性リンパ腫 33, その他 148 であった。そのうちフォロワー数が 500 以上のアカウントは 52 あり、そのうち最もフォロワー数が多いアカウントは乳がん患者であり、そのアカウントと他のがん患者とのつながりが確認された。

これ以外に、疾病に対し語彙ネットワークを用いた研究として以下の 2 つがある。

成田らは SNS 患者アカウント書き込み内容に共起ネットワーク図示を用いて摂食障害の調査している[3]。

中川らは、がん用語をノード、がん用語間で連想されるもの同士をリンクとして結び、いわば連想ネットワークを作成し、がん用語辞書を作成している[4]。

3. 研究方法

3.1 共起ネットワーク

共起とは、何らかのものと他のものが同時に出現することをさす。ノードとそれらを結ぶリンク（またはエッジ）で構成されたものをネットワークといい、共起したもの同士をリンクで結んだネットワークを共起ネットワークという。この時リンクで結ばれる端のノードは共起したもの同士である。

共起ネットワークとして表現されるものの一つに文章がある。この場合、文章をある単位で区切り、その単位内で同時

に出現した単語をノードとしてリンクを結び共起ネットワークを作成する。区切りとなる単位としての具体例は、文一行、1 ページ、前後の単語などがある。また、一度でも共起したもののノード同士を必ずしもリンクさせ表示するわけではなく、共起頻度という概念がある場合は頻度が高いもののみ表示することもある。表示するリンクをしぼることによってネットワークを把握することが容易になる。

3.2 ツイッターツイートから共起ネットワークを作成する場合の流れと問題点

図 1 で示す流れで共起ネットワークを作成した。

まず最初にツイッターのアカウントプロフィールにがんに関しての記述があるアカウントを検索する。ツイッター自体にアカウントプロフィールのみを検索する機能は無い。そこでアカウントプロフィールのみを検索する機能を提供するウェブサービス「16 プロフィール検索」を用いることで該当アカウントを集める。

続いて収集したアカウントのツイートを取得する。ツイートの取得にはツイッターAPI を用いる。ツイッターAPI とはツイッターを運営する組織がツイッターアプリケーション制作者に公式に提供している機能のことである。ツイッターAPI を組み込むことで、アプリケーション制作者は自分が制作しているアプリケーションにツイッター運営組織が提供している様々な機能を加えることができる。ツイッターAPI を用いることで指定アカウントのその時点までのツイート最大 200 ツイートを取得し、全アカウントツイートを 1 つにまとめる。

次に得られたツイート各行を単語（形態素）に分解する。単語への分解には形態素解析ソフト茶笥を用いる。そしてそこから名詞のみを抽出する。文を構成する単語（形態素）の品詞は様々あるがテキストマイニングでは通常、名詞、形容詞、動詞が用いられる。しかし、本研究では 1) 否定文の扱いの難しさ、2) 抽出された語のその品詞の割合の少なさ、から名詞のみを扱うこととする。茶笥には一般的な単語であるならば品詞を判別可能とするデフォルト辞書が備わっているが、今回は医学専門用語も取りこぼすことなく処理したいためウェブサイト「がん情報サービス <http://ganjoho.jp/>」「goo ヘルスケア <http://health.goo.ne.jp/>」の記述から得られた品詞判別不能の単語のうち医師ががんに関係が深いと判定したものを辞書に加え形態素解析を行った。

続いて句点「。」や「?」「!」で区切るために「。」や「?」「!」で改行する。同じ行中に出現する単語の各組合せをカウントし全単語組合せの同じ行中に同時に出現する頻度（共起回数）を調べる。つまり共起頻度の定義は「全アカウントツイート上である単語組が共起した行数」となる。

次に共起回数が多かった組合せ上位 100 位までを求め、それら組合せの単語同士をリンクさせたネットワークを図示する。このとき各単語はノードとして表される。

この時得られた共起ネットワーク 6 つを図 2 に示す。この図 2 では共起ネットワーク上で医学に明らかに関係する内容のものを線で囲んでいるが、共起ネットワークの大部分が医学と関係がない内容だということがわかる。これはがんに関

係するアカウントのツイートが必ずしもがんに関わるものとは限らないことを示している。例えばがん患者のアカウントでも日常的なツイートや趣味に関するツイートが大半を占めているということになる。共起ネットワークに示された医学に関係のない内容の具体例としては、ギャンブル、野球、アダルトサイトなどの趣味の話題が挙げられる。しかし、本研究ではがんに関係するアカウントのがんに関するツイートを取得したい。この目的を果たすには、日常的なツイート、趣味に関するツイートが高頻度であることが問題となる。日常的なツイート、趣味に関するツイートが高頻度であることで、それらが共起ネットワークに出現し、それらほど高頻度に共起しないがんに関するツイートの内容がネットワークに現れないということが起きている可能性があるからである。

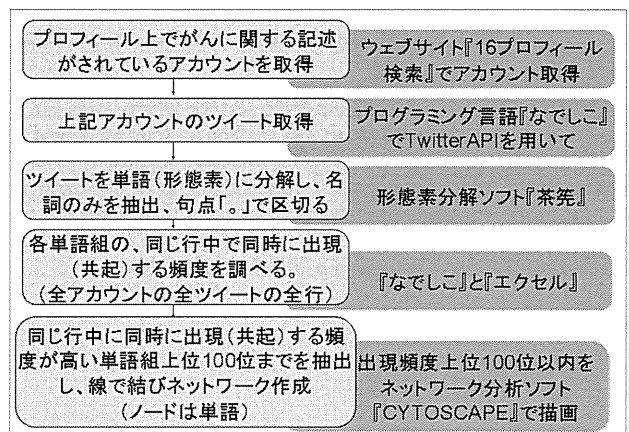


図 1 一般的方法に則したツイッターツイートからの共起ネットワーク作成の流れ

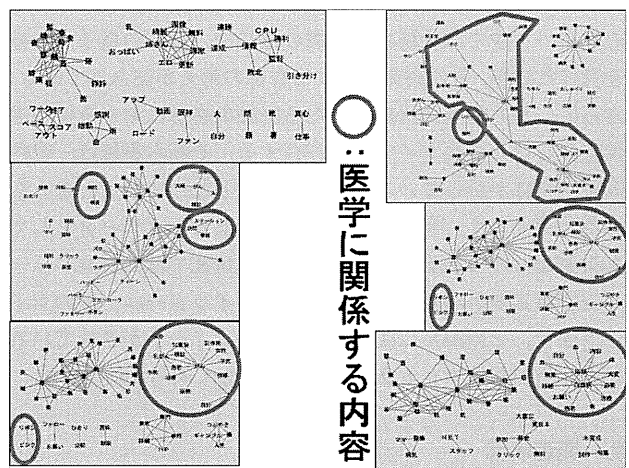


図 2 一般的方法に則して得た共起ネットワーク

3.3 問題点の解決法

高頻度である日常的なツイート、趣味に関するツイートを共起ネットワークから除去する方法として 2 つの方法を採用する。

日常的なツイート除去に関しては、がんに関係がない一般アカウントの日常的なツイートと共通する内容を除去する方法

を採用する。日常的ツイートは一般のアカウントもがんに関係するアカウントも同様ならずである。まず特徴のないアカウント名をいくつか取得し、さらにそれらアカウントのツイートの共通に出現した単語を取得する。そしてその単語を共起ネットワークから除去する。特徴のないアカウントは「16プロフィール検索」サービスで地名をキーワードとして検索して得た。具体的には今回使用したキーワードは「山形県米沢市」である。

趣味に関するツイート除去に関しては、趣味に関する内容を頻繁にツイートするアカウントのツイートを抑制する方法を採用する。趣味に関する内容が共起ネットワークに示されることの原因は同じ趣味を持つアカウントがたくさんいるからではなく、趣味に没頭しているアカウントが趣味に関して頻繁にツイートすることが原因と考えられる。そこで、同一アカウントに複数回ツイートされた共起語は+1しかカウントしないことにする。つまり、3.2における共起頻度の定義は「全アカウントツイート上である単語組が共起した行数」であったが、提案手法における共起頻度は「ツイート内同一行で1度でも共起したアカウント数」となる。3.2で示した共起ネットワークに対してこの問題の解決法を施した方法の流れを図3に、そして得られた共起ネットワーク6つを図4に示す。問題解決法を採用したことにより共起ネットワークの多くが医学に関係する内容になったことがわかる。

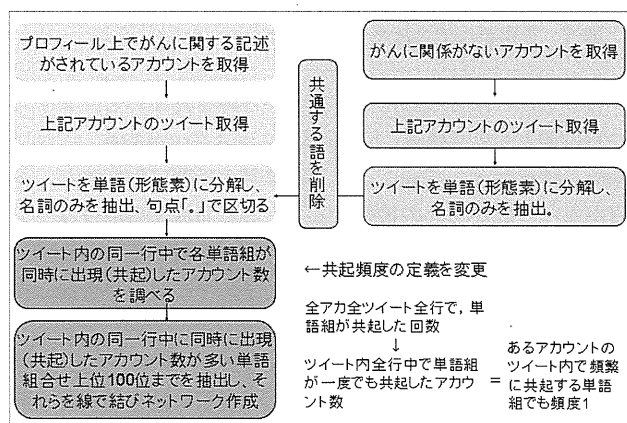


図3 問題解決法を施した共起ネットワーク作成の流れ

4. 共起ネットワーク作成結果

4.1 条件

共起ネットワークを作成したときの条件は以下の通りとなる。

- ・ノード：名詞のみ。ただし数字は含まない
- ・調査するがんの種類および該当アカウント数：胃がん (44人)、腸がん (64人)、子宮がん (134人)、肺がん (87人)、乳がん (313人)、白血病 (158人)。ただし腸がんは大腸がん、直腸がん、結腸がんで、子宮がんは子宮ごとと子宮頸がんで構成されている
- ・ツイート取得時：2012年8月19日0:39~2:52 (胃がん、

腸がん、白血病)と、2012年8月20日14:40~17:24 (子宮がん、乳がん、肺がん)

- ・16プロフィール検索におけるキーワード：「～がん OR ～ガン OR ～癌」(～は部位)。ただし、白血病は「白血病」のみ

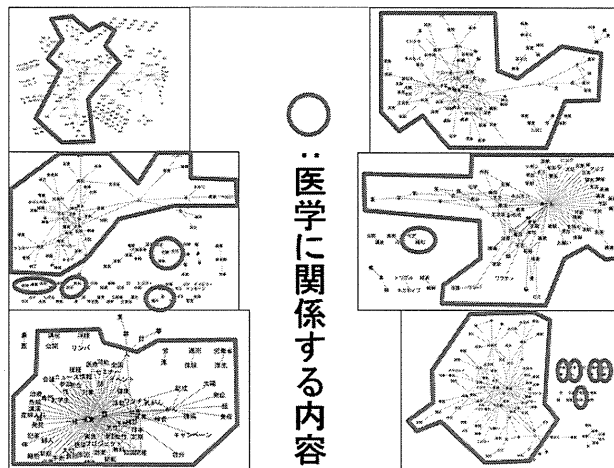


図4 問題解決法を施した共起ネットワーク作成の流れ

4.2 共起ネットワーク作成結果

4.2.1 胃がん

胃がんの共起ネットワークは図5となる。4つの話題に関するクラスターが確認された。それらクラスターは枠で囲んでいる。

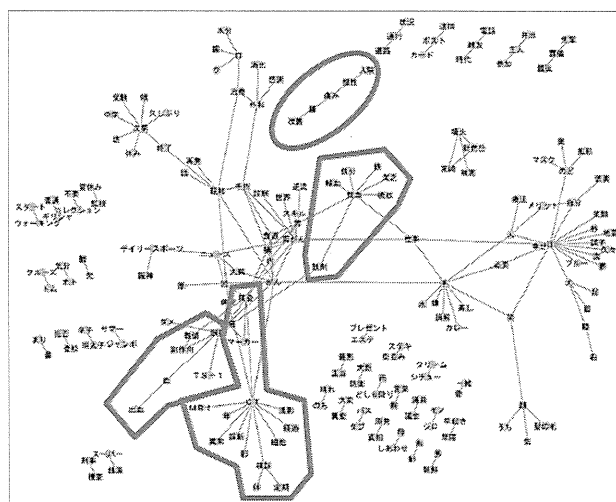


図5 胃がんの共起ネットワーク

枠の上部の高さを基準として上からクラスターを見ていくと、まず1つ目に胃がんの中期、後期に発生する腰痛とその改善策に関するクラスターが確認できる。

続いて2つ目に「貧血」がハブとなっているクラスターがある。ここでは胃の切除からくる貧血とその対処の話題を表している。胃が切除されると胃壁細胞が減少する。赤血球を作るビタミン B12 の吸収にはこの胃壁細胞が作る胃内因子が

必要であるがそれが減少することでビタミン B12 は吸収されにくくなる。そのため赤血球が作られにくくなる。胃が切除されるとさらに胃酸の分泌量も減少し鉄分の吸収も減少するので、これらにより貧血となる。この胃の切除からくる貧血に鉄剤の服用で対処する。

次に3つ目に「CT」をハブとした胃がんの検査クラスターが見られる。胃がんの検査は、CT (X線)、MRI (磁気と電波)、腫瘍マーカー (がん生成物質の有無)、バリウム検査などがある。

4つ目に「TS-1」「服用」「副作用」「出血」という単語が含まれるクラスターが確認できる。抗がん剤 TS-1 とその副作用に関する話題が対応する。TS-1 の副作用として血小板減少があり、それにより、歯茎出血、皮下出血、口内炎が生じる。

4.2.2 腸がん

腸がんの共起ネットワークは図6となる。4つの話題に関するクラスターが確認された。

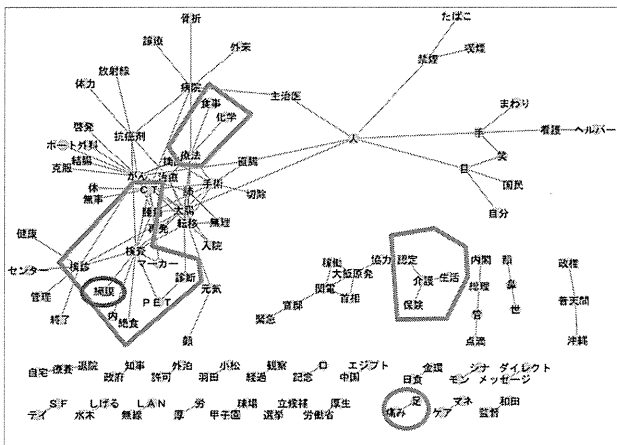


図6 腸がんの共起ネットワーク

枠の上部の高さを基準として上からクラスターを見ていくと、1つ目に療法に関するクラスターが見られる。ここでは抗がん剤を用いる治療である化学療法と食事療法に関する話題が対応する。食事療法とはがん細胞の成長を促進する食べ物の摂取を避ける食事をとるという療法である。食べ物の中にはがん細胞の成長を促進するものが存在する。

続いて2つ目に「検査」という言葉をハブとするクラスターが見られる。このクラスターにはCTやPETといった診断法を使ったがん検査の話題が対応する。腸がんではCT診断のほか、PET診断という検査法がある。このPET診断では、特殊な検査薬でがん細胞に目印を付け、陽電子を用いて撮影する。これにより小さい早期のがんを発見することを可能とする。

3つ目に「介護」「保険」「認定」「生活」という言葉で成り立つクラスターが確認できる。

4つ目に「足」「痛み」というクラスターが確認できるが、それはエルプラットという抗がん剤の副作用に対応していると考えられる。エルプラットは再発や切除不能な大腸がんに効果的な抗がん剤であるが、その副作用として手足などの末梢神経の障害が確認されている。

4.2.3 子宮がん

子宮がんの共起ネットワークは図7となる。この共起ネットワークの中心では2つのハブ「子宮」「頸」が見られる。これら単語は互いにリンクし、さらに、これら単語どちらにもリンクしている単語が非常に多く存在する。これは、「子宮頸」および「子宮頸がん」という名詞が形態素解析ソフト茶筌の辞書に含まれていないことと、「子宮がん」に対し「子宮頸がん」という単語が使われるツイートの比率が非常に高いことが原因といえる。これにより共起ネットワークの構造は非常に複雑になっている。クラスターおよび重要単語と思われるものは枠で囲んである。

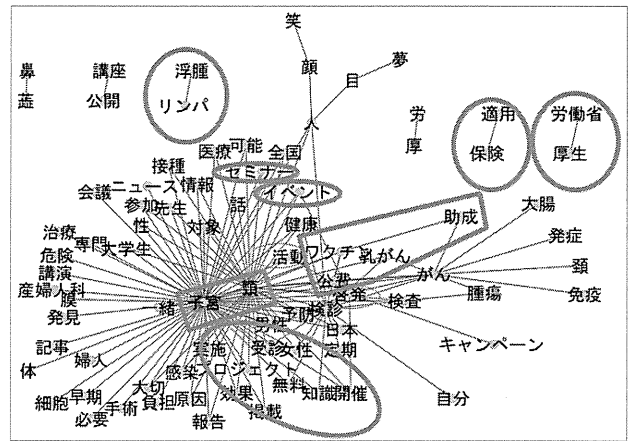


図7 子宮がんの共起ネットワーク

枠の上部の高さを基準として上から枠を見ていくと、まず1つ目の枠として「浮腫」「リンパ」クラスターがある。これは治療の後遺症として起こるリンパ浮腫の話題に対応する。子宮がん (や乳がん) の治療として行われる手術の部位はリンパが心臓に戻る活動に重要なリンパ節群に集中している。このリンパ節群が手術で切除されるとリンパの流れが停滞し脚や腕がむくむリンパ浮腫が起こる。

続いて2つ目に「厚生」「労働省」クラスターがある。これはツイート取得前に厚生労働省からの提言に関する内容と対応すると考えられる。HPVワクチン (子宮頸がんワクチン) 接種後に副作用とよばれる症状が見られたことから、厚生労働省はワクチン接種の積極的勧誘を控えるように提言した。

次に3つ目に「適用」「保険」クラスターが確認できる。このクラスターは子宮頸がん検診には健康保険が適用されないことに対応していると考えられる。検診には健康保険が適用されないが、住民検診や職場検診では自治体や健康保険組合が費用を負担してくれるので安価か無料で検診を受けることができる。

続く4つ目、5つ目の「セミナー」「イベント」はこの共起ネットワークの2大ハブである「子宮」「頸」と結びつく重要な単語である。さらに一番下の枠の部分を含めたものも含め、全て子宮がんセミナーに関する情報と対応する。子宮がんに関する無料セミナーが頻繁に開催されている。そのセミナーで伝えられる情報の中でも主要なものの一つは、HPVワクチ

ン（子宮頸がんワクチン）接種によって子宮頸がんは唯一予防できるがんであるといった情報である。

7つ目の、「ワクチン」「助成」「公費」で構成されるクラスターは、子宮頸がんワクチン接種に公費による助成が行われている話題が対応する。

4.2.4 肺がん

肺がんの共起ネットワークを図8に示す。5つの話題に関するクラスターが見られる。

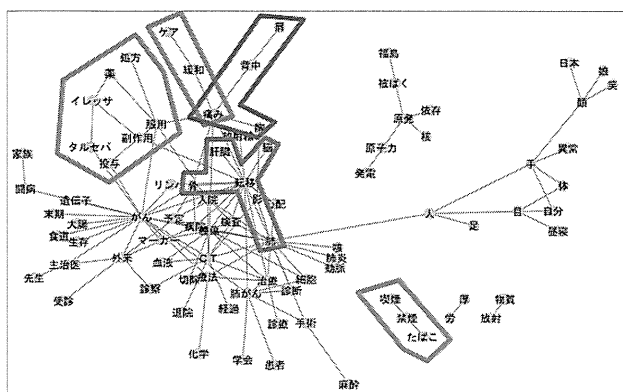


図8 肺がんの共起ネットワーク

上からクラスターを見ていくと、1つ目、2つ目に「緩和」「ケア」「痛み」と「肩」「背中」「痛み」「胸」とクラスターが存在する。これらは肺がんで痛む部位とその緩和ケアの話題に対応する。肺がんはがんの中でも痛みが起る確率が高いがんであることが知られている。肺がんは胸の痛みだけではなく、早期では肩の痛み、進行すると背中や腰の痛みを生じる。骨にがんが転移した場合、強い痛みをとまなう。このようなことから肺がんには緩和ケアが必要となる場合が多い。肺がんにおける痛みの緩和ケアとしては、痛み止め薬、神経ブロック、放射線治療が挙げられる。

続いて3つ目は「イレッサ」「タルセバ」など抗がん剤に関する話題と対応する。化学療法における抗がん剤イレッサは発売当時、肺がんに対し効果が大きく副作用の少ない抗がん剤とされてきた。しかし、イレッサは急速に進行する間質性肺炎を引き起こすことがあることがわかった。

4つ目は転移に関するクラスターである。肺がんは脳、肝臓、骨、（他に副腎）に転移しやすいという話題と対応する。

最後の5つ目のクラスターは「禁煙」「喫煙」に関するクラスターである。

4.2.5 乳がん

乳がんの共起ネットワークを図9に示す。「乳がん」という単語にたくさんの単語がリンクしている。

上から話題毎にクラスターを見ていくと、まず1つ目が「ピンク」「リボン」という「乳がん」とリンクしている単語である。これは乳がん啓発キャンペーンのシンボルであるピンクリボンを表している。乳がんの正しい知識を広め、乳がん検診の早期受診を促進することを目的とした世界規模の啓発キャンペーンが行われている。

続いて2つ目が、「療法」「化学」「ホルモン」「副作用」で構成された療法に関するクラスターである。乳がんには様々な薬物療法が存在する。このクラスターには抗がん剤を使った化学療法とホルモン治療が表されている。女性ホルモン受容体を持っている乳がんは女性ホルモンに合体することで細胞分裂が活発になり増殖する。それに対し、女性ホルモンが受容体に合体することを阻止させる方法や、女性ホルモンの分泌を低下させる方法を用いた療法をホルモン療法という。乳がんのその他の薬物療法として、抗体療法、分子標的療法がある。

次に3つ目が「自分」「手」で構成されたものである。これは自己検診の内容と対応していると考えられる。乳がんは内臓にできるがんと異なり体の表面にできるため、自己検診が可能である。

4つ目が、乳房再建手術に関するクラスターである。近年行われるようになった乳がん手術後の乳房再建手術がこのクラスターに対応する。手術で使用する人工乳房に公的医療保険が適用することになったことも対応している可能性がある。

5つ目が、緩和ケアに関するクラスターである。転移性乳がんを根治する方法はまだ無いが、かなりの延命は可能になってきている。延命治療と同時に緩和ケアも行われるようになった。

6つ目が、リンパ浮腫に関するクラスターである。子宮がんと同様の理由で乳がんもリンパ浮腫が起きる。

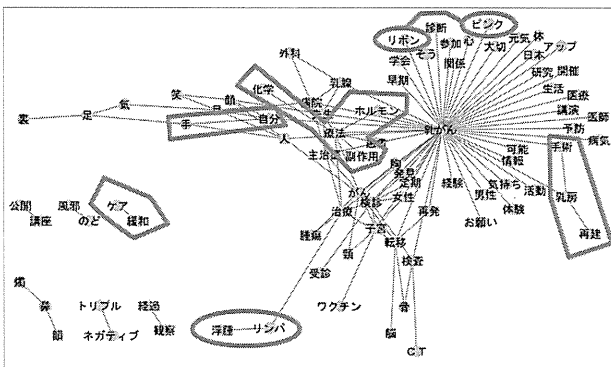


図9 乳がんの共起ネットワーク

4.2.6 白血病

白血病の共起ネットワークを図9に示す。7つの話題に関するクラスターが確認された。

上からクラスターを見ていくと、1つ目に「赤血球」「血小板」「輸血」で構成されるクラスターが確認できる。これは白血病時の赤血球、血小板輸血に関する話題に対応する。白血病の抗がん剤投与後、一時的に血液を作ることができない状態になる。このとき白血球は輸血することはできないが、赤血球、血小板は輸血で補われる。

続いて2つ目（図の右側の分布の真ん中）はAML（急性骨髄性白血病）に関するものである。正常な造血幹細胞はリンパ系造血細胞と骨髄系造血細胞に分かれ成熟していくが、骨髄系造血細胞が腫瘍化し成熟能を失う疾患をAMLという。

3 つ目 (図の右側の分布の右側) は、白血病の兆候の足痛が対応する。白血病の兆候として首やわきの下、足の付け根のリンパ節が腫れる、小児性白血病では足が痛い、骨が腫れるというものがある。

4 つ目 (図の右側の分布の左側) は「肝臓」 「数値」 で、これは肝機能 ADH 数値が高い数値をとる場合に白血病が疑われることに関係する話題といえる。

5 つ目 (図左側) は、抗がん剤ステロイドに関するクラスターである。ステロイドはリンパ系白血病細胞への障害作用を持つため、急性リンパ性白血病に用いられる。

次の 6 つ目は「GVHD」 「移植」 で構成される。GVHD とは移植片対宿主病をさす。通常、拒絶反応は移植片に対し患者の体が起こすが、この GVHD は骨髄移植 (造血幹細胞移植) や輸血といった白血病治療のために移植された移植片に含まれるドナーのリンパ球が、患者の体を異物とみなして攻撃することで起こる。

最後の 7 つ目は「免疫」 「抑制」 で、これは免疫抑制状態の話題に対応する。化学療法により白血球のうちリンパ球が減少している状態を免疫抑制状態といい、この状態では正常な状態ではかからないような感染症にかかりやすくなる。

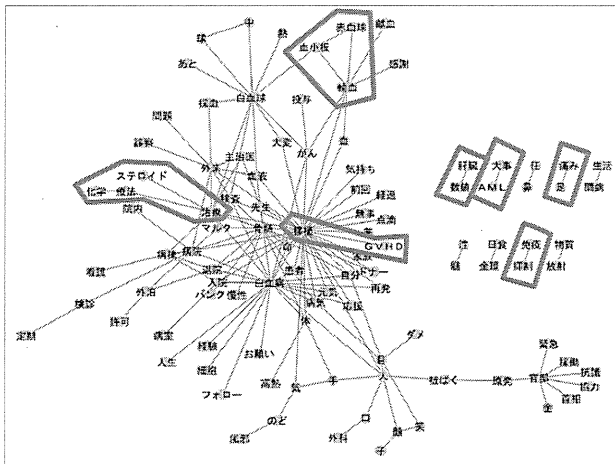


図 10 白血病の共起ネットワーク

4.3 全共起ネットワークのまとめ

6 つのがんの共起ネットワークで見られたクラスターの話題を表 1 にまとめる。各クラスターの話題を「診断」「症状」「治療」「その他」に分け示した。

診断の項目では、CT, MRI, 腫瘍マーカー, PET を主にして自己診断や、内臓数値も話題にされていたといえる。

症状の項目では病状だけでなく抗がん剤の副作用、手術副作用も示した結果、抗がん剤の副作用に関する話題が目立った。

治療の項目では、抗がん剤治療だけでなくがんの細かい種類それぞれに対応した治療法が示されていたほか、治療法の選択肢としていくつか示されていたものもあった。

その他の項目では、介護、健康保険の適用不適用、公費助成、啓発活動といった話題に対応するクラスターが見られた。

各がんの項目毎に見ていくと、子宮がんの共起ネットワークでは診断と治療に関するクラスターが見られないという特徴があった。今回、共起ネットワークにはツイッターツイート上で話題となった頻度が高かったものしか表していない。つまり、子宮がんに関係があるツイッターアカウントは、診断、治療に関する話題をあまりツイートしていないということになり、ツイッター経由では子宮がんの診断と治療に関する情報はあまり得ることができないということになる。この理由として、子宮がん啓発活動が盛んでありすぎること、子宮がん患者のアカウントに対し、医療関係者や、子宮がん啓発活動に関係するアカウントが多い可能性があることが考えられる。

表 1 各共起ネットワークの内容

	診断	症状	治療	その他
胃がん	CT, MRI, 腫瘍マーカー	腰痛, TS-1副作用	抗がん剤TS-1 鉄剤服用	
腸がん	CT, PET	エルプラット副作用	化学療法 食事療法	介護
子宮がん	×	リンパ浮腫	×	・啓発活動 ・検診健康保険適用外 ・ワクチン接種公費助成
肺がん	CT	転移, 肩痛, 背中痛, イレッサ副作用	抗がん剤イレッサ, タルセバ	緩和ケア
乳がん	自分の手で診断可能	転移 リンパ浮腫	化学療法 ホルモン療法	緩和ケア ピンクリボン
白血病	肝臓数値	肝臓数値 足痛 免疫抑制 GVHD	化学療法, ステロイド治療 赤血球輸血 血小板輸血	AML 造血幹細胞移植

また、ネットワーク構造を見ると子宮がんの共起ネットワークで複雑な構造が見られた。「子宮」と「頸」という単語両方に非常にたくさんの様々な語が結びついていた。この理由として「子宮頸がん」という言葉がツイッターツイートで単語に分解する際に用いた形態素解析ソフトの辞書に登録されていないことからノードとして現れず、代わりに辞書に含まれている「子宮」と「頸」両方に結びついたと推測した。そこで、図 11 に「子宮頸がん」という単語を辞書登録した場合の子宮がん共起ネットワークを示す。「子宮」と「頸」という単語両方に結びつく単語の数は大幅に減少し、追加した「子宮頸がん」という 1 単語にたくさんの単語が結びつくといったネットワーク構造が得られた。辞書登録前である図 7 と辞書登録後の図 11 を比較すると、図 11 はネットワークとして非常に見やすくなっていることがわかる。このように辞書登録状況が共起ネットワークの構造に大きく影響する可能性があることがわかった。辞書にある 1 単語を登録することによって共起ネットワークが大幅にシンプルになる可能性がある。このことから逆にネットワーク構造から形態素解析ソフト辞書に登録すべき単語を見いだすことができる可能性が示されたといえる。

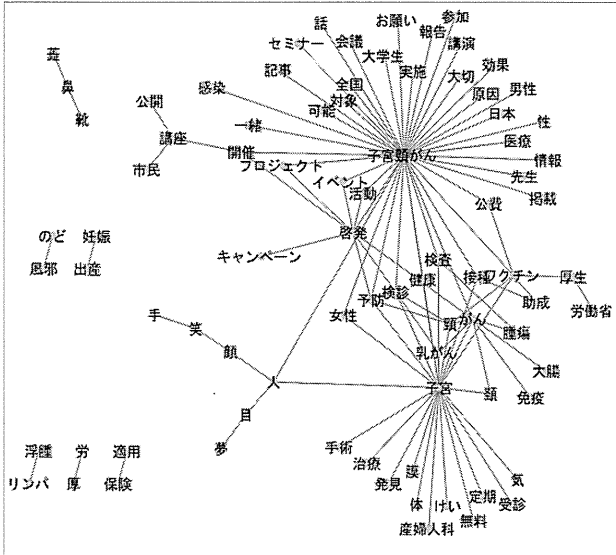


図 11 「子宮頸がん」を辞書登録した場合の子宮がんの共起ネットワーク

5. まとめ

プロフィール上に「胃がん」「腸がん」「子宮がん」「肺がん」「乳がん」「白血病」の記述があるアカウントのツイートを共起ネットワークとして表し、がんに関するアカウントから発せられるツイートのうち頻度が高い話題を明らかにした。共起ネットワーク上で各話題は一つにまとまって表される。共起ネットワーク上で見られた話題として、各がんの「診断」「症状」「治療」に関するもののほかに、介護、健康保険の適用不適用、ワクチン接種公費助成、啓発活動といったものがあつた。ただし、子宮がんでは、高頻度に現れる話題の中に「診断」「治療」の話題が見られなかった。さらに、本来、形態素解析ソフトの辞書に登録すべき単語の未登録が共起ネットワーク構造の複雑さを増すことがあることが示された。逆にいうと、共起ネットワークの構造から辞書登録すべき語を見いだすことができる可能性が示されたといえる。

参考文献

[1] Chen X, Siu LL: Impact of the media and the internet on oncology: survey of cancer patients and oncologists in Canada, *Journal of Clinical Oncology*, Vol.19, Issue 23, pp4291-4297, 2001.

[2] Yuya Sugawara, Hiroto Narimatsu, Atsushi Hozawa, Li Shao, Katsumi Otani and Akira Fukao: Cancer patients on Twitter: a novel patient community on social media, *BMC Research Notes*

[3] 成田彩夏, 往住彰文: SNS コメントの大量分析による摂食障害の特徴づけ, *日本感性工学会大会予稿集*, 2010.

[4] 中川晋一, 内山将夫, 三角真, 島津明, 酒井善則: コーパスに基づくがん用語集の作成と評価, *自然言語処理*, Vol.16, No.2, pp3-44, 2009.

