

分担研究年度終了報告書

医療広報におけるソーシャルメディアの可能性

（生活習慣病とツイッター：がん患者コミュニティとの比較）

分担研究者 田中 敦 山形大学大学院理工学研究科 准教授

研究協力者 津谷 篤 山形大学大学院理工学研究科 研究員

研究要旨

生活習慣病に関するTwitterの利用実態を明らかにするため、Twitter上のアカウント間のコミュニケーションネットワークとツイートに含まれる単語の共起ネットワークを分析した。それらのネットワークはがん患者のものとは異なり、日常的な単語と多く結びついていることが明らかになった。

A．研究目的

近年、糖尿病や高血圧といった生活習慣病が増加している。厚生労働省による平成 23 年の調査では、糖尿病は 270 万人、高血圧は 900 万人の患者がいることがわかっている。こうした患者たちが病気の情報を得る有効な手段としてインターネットの利用があげられる。中でも、2006 年に登場したマイクロブログサービスの Twitter は、利用者数が日本人だけで 2000 万人いると言われており、双方向のコミュニケーションを迅速に行うことができ、さらに匿名での利用も可能で、患者や医師による情報のやり取りも行われている。

本研究では、Twitter における生活習慣病に関するツイートの解析を通じて、Twitter 上での生活習慣病に罹っている人々の生活の様子を発信している情報について考察する。また、これまで行ってきたがん患者におけるツイート解析と比較分析し、生活習慣病とがんのネットワークの違いについて明らかにする。

B．研究方法

本研究で分析対象とする病気は、厚生労働省が主な生活習慣病として挙げている糖尿病、高血圧、脳梗塞、心臓病(心筋梗塞と狭心症)の 4 つである。本研究は次のように進めた。

- (1) 16 プロフィール検索により、各病気に関するアカウントを集める。
- (2) (1)で集めた全アカウントについて、各アカウントから 20 ツイートずつ得る。
- (3) 病気ごとに、全ツイートを単語に分解し、「名詞」「動詞」「形容詞」のみ抽出する。
- (4) 病気ごとに、抽出した全単語の全ツイート中での出現回数を数え、一定数以上出現する単語について各ツイートでの共起回数を数える。
- (5) (4)で一定数以上の共起回数を持つ単語の組み

合わせをネットワークとして図示し、このネットワーク図からそれぞれの病気について特徴を考察する。

ツイートの単語への分解には、形態素解析ソフト「chasen」を用いた。chasen が単語への分解を行う際に使用している辞書には、薬品名や治療方法名、物質名が登録されていないため、これらの単語は「未知語」という品詞として出力された。そこで、「未知語」と出力された単語の中の、薬品名、治療方法名、物質名を辞書に「名詞」として登録し再度ツイートの分解を行うことにした。辞書に登録した単語は全部で 246 語である。

本研究では全ツイート中 10 回以上出現し、特徴的である単語を主観的に選んで共起回数を数えることにした。出現数のトップは 4 種の病気全てで「する」「いる」という単語で、その他「ある」「月」「日」といった単語が上位にあったが、これらの単語は医学的な意味を持たないと思われたため、共起回数のカウントを行う単語からは省いた。

共起回数のカウントを行う対象としては、食品名や「食事」「昼食」などの食生活に関わる単語、「検査」「薬」「病院」といった病気治療を表す単語、「新聞」「テレビ」「ニュース」といった情報源を表す単語、「高い」「悪い」「多い」などの状態を示す単語を選択した。

C．研究成果

本研究で集めたアカウント数、ツイート数は表 1 の通りである。また、表 2 に、各病気ごとの全ての単語の種類数と全ツイート中に 10 回以上出現した単語の種類数、選択した単語の種類数を示す。

表 1：生活習慣病に関する Twitter アカウント情報

病気名	アカウント数(人)	ツイート数(ツイート)
糖尿病	577	11540
高血圧	128	2560
脳卒中	60	1200
心臓病	61	1220

表 2：生活習慣病に関するツイート情報

病気名	全ての単語の種類数(種類)	10回以上出現した単語の種類数(種類)	選択した単語の種類数(種類)
糖尿病	16140	2128	402
高血圧	7115	431	125
脳卒中	4879	245	89
心臓病	4009	191	60

[生活習慣病の共起ネットワーク構造]

表 2 の選択した単語で、共起回数が 5 回以上の単語の組み合わせについてリンクで結んだネットワーク構造を図 1 ~ 4 に示す。

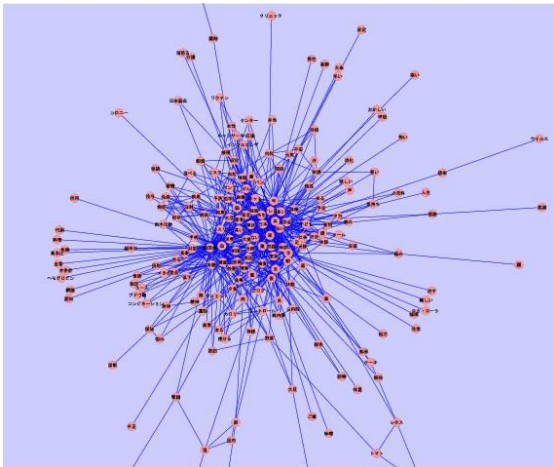


図 1：糖尿病の共起ネットワーク

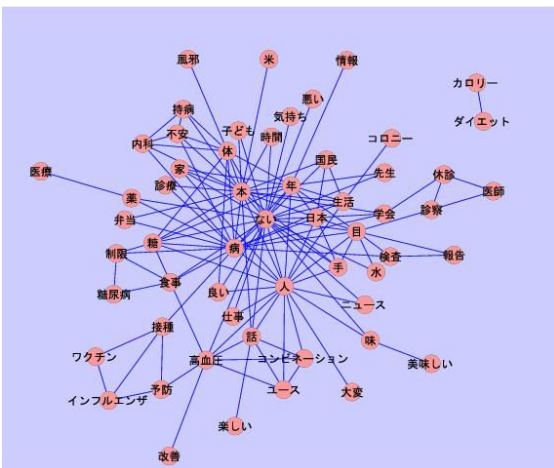


図 2：高血圧の共起ネットワーク

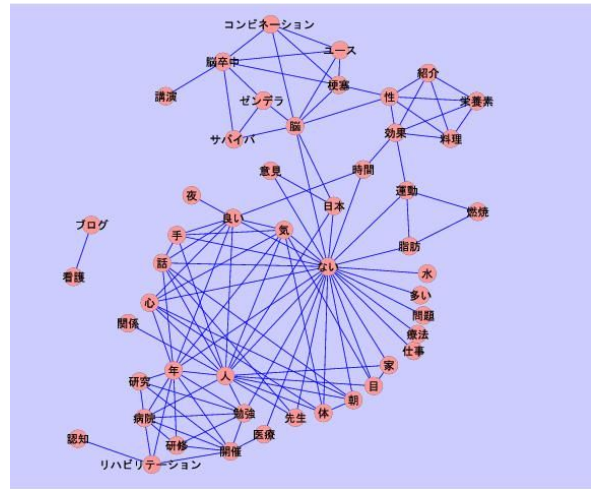


図 3：脳卒中の共起ネットワーク

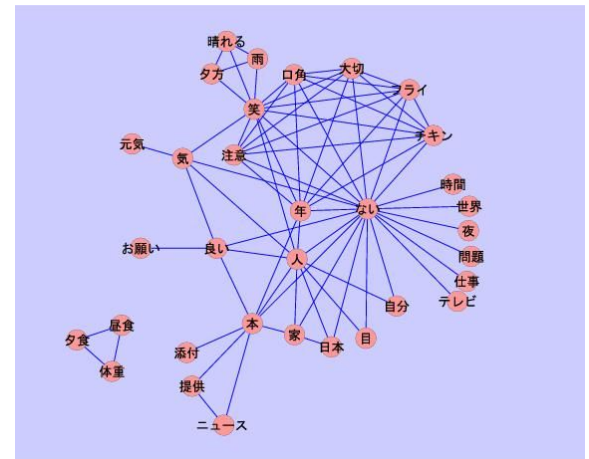


図 4：心臓病の共起ネットワーク

ツイートの中には、病気と関係の無い内容も存在する。生活習慣病とがんのネットワークからこうした病気と関係ない単語を除外するため、病気と関係の無いアカウントのツイートから一般ネットワークを作成した。この一般ネットワークに出現する単語を、ネットワークから除去した。除去後のネットワークのノードとリンクが元のネットワークと比較してどの程度減少したかを計算した結果を表 3 に示す。

表 3：一般語除去の効果

病気名	除去後のノードの割合 (%)	除去後のリンクの割合 (%)
糖尿病	84.2	58.5
高血圧	61.4	34.3
脳卒中	44.9	27.8
心臓病	39.4	23.6
平均	57.5	36.1

D．考察

生活習慣病とがんのネットワークを比較すると、がんでは「手術」「抗がん剤」「放射線」といった治療法を表す単語の種類が多く、全てのネットワークに対して少なくとも1つはこれらの単語が出現しているのに対し、生活習慣病では糖尿病と高血圧にのみ「薬」、糖尿病にのみ「インスリン」「手術」という単語が現れた。脳卒中、心臓病には治療法を表す単語は1つも出現しなかった。このことから、生活習慣病では治療に関するツイートは少なく、生活習慣を含めた一般的な内容のツイートが多いことが考えられる。

また、一般的な単語を除外したネットワークを見ると、糖尿病では「尿」、高血圧では「病」、脳卒中では「性」がハブとなっている。心臓病は細かいコミュニティに分割された。このことから、生活習慣病のツイートでも病気に関する会話がされていることが分かる。しかし、元のネットワーク図と一般の単語を除外したネットワーク図を比較すると、ノード、リンク共に生活習慣病では大きく減少している。

このことから、生活習慣病の患者は病気に関するツイート以上に多くの普段の会話をツイートしていることが考えられる。

E．結論

Twitter上で交される生活習慣病に関する情報は、がん患者による情報と異なり、より生活に密着したものが多くことが明らかになった。生活習慣病を抱えている人は、病気に関することだけではない普段の会話にソーシャルメディアを利用している割合が高いことが推察される。

F．研究発表

- 1.論文発表
なし
- 2.学会発表
なし

G．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