

日本における麻疹・風疹・インフルエンザ・COVID-19 の情報探索行動： P03 Yahoo! 知恵袋投稿のテキストマイニングによる比較分析



国際医療福祉大学大学院医学研究科医学専攻 ○熊澤 芽葉恵, 山崎 力

背景

- ◆ 日本では、輸入例を契機とした麻疹(2018)、先天性風疹症候群の患者が報告された風疹(2018-2019)、繰り返す季節性インフルエンザ(毎冬)、未曾有のパンデミックを引き起こしたCOVID-19(2020)が問題となった。
- ◆ 疾患ごとに流行規模・政策対応・ワクチン状況は大きく異なる。
- ◆ こうした違いは、「自分はどれだけ危ないか」「家族は守れるか」「実際に何をすればよいのか」という具体的な不安や意思決定の仕方に影響を与える。
- ◆ 公衆衛生上、専門家による一般市民への適切な情報提供は重要。

出発点

専門家が発信する感染症の情報は、確かに有益

▶▶▶ 一般市民が本当に知りたい情報と一致しているだろうか？



- ◆ 感染症流行時の情報探索行動は、疫学特性だけでなく政策環境や不確実性に左右される。
- ◆ 特に、個人の生活に直結するテーマでは専門家が発信する「正しい情報」と市民が「知りたい情報」との間にギャップが生じる可能性がある。

本当に知りたい情報は？ どうやって知る？

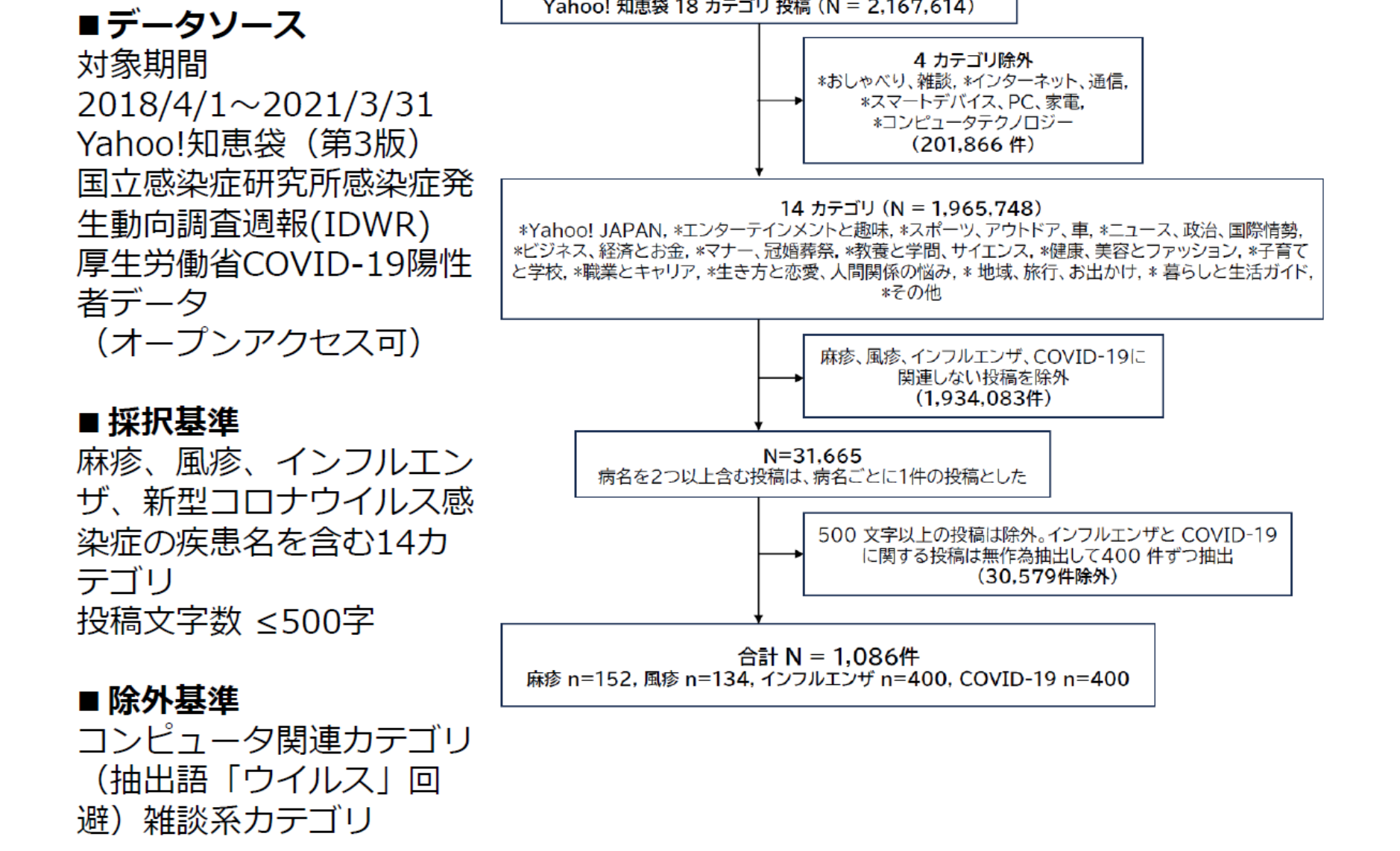
- ◆ 感染症流行時、人々の情報探索行動は、単なる感染者数の推移や政策発表だけでは十分に説明できない。
- ◆ 専門家の発信と市民の関心とのギャップは、定量アンケート調査やインタビューでは捉えにくい。

匿名で自由に質問できるQ&Aプラットフォームは対面調査では表に出にくい葛藤や迷いがそのまま文章化される場
従来の調査では捉えきれない本音の情報ニーズが隠れている

目的

- ◆ 日本最大級Q&AプラットフォームYahoo!知恵袋の投稿から、一般市民が本当に必要としている情報とは何かを明らかにする
- ◆ 疾患ごとの情報探索行動とリスク認知の違いを可視化する
- ◆ 流行時の疾患別リスクコミュニケーション設計に役立つ知見を提示する

方法



テキスト処理と解析環境

- ◆ 解析環境
 - KH Coder 3.02c @2001 (<https://khcoder.net/>)
 - R v.4.4.3 中心媒介性(BC: betweenness centrality)計算
- ◆ テキスト前処理
 - URL・絵文字削除
 - 数字/記号の正規化、℃のみ保持(体温文脈に重要)
 - 名詞・動詞・形容詞を対象
 - 補助語・定型表現はストップワード化
 - 用語の意味保持のため、医学系多語表現は複合語処理

分析手法

- ◆ 各疾患の陽性報告数(感染者数)と投稿数の週次推移
- ◆ 流行曲線と同時期の投稿数を比較する
- ◆ 形態素解析
 - 疾患別に出現した上位語(頻出語)を抽出
 - Lift値を計算 → Lift ≥ 1.5かつTotal ≥ 10を疾患特異語として判定する。
- ◆ 階層的クラスタ分析
 - Ward法 × ユークリッド距離 → 投稿のテーマ構造を把握する。
- ◆ 共起ネットワーク
 - BC値を計算 → 語間の橋渡し語からテーマの繋がりを把握する。
 - ※ BCレンジは疾患間でネットワーク規模が異なるため横比較は困難

Lift 語が特定の投稿にどれだけ偏って出現しているかを表す指標 BC ネットワークの中で語と語の橋渡し役になる語を数値化した指標

$$Lift_d(t) = \frac{Share_d(t)}{baseline_d}$$

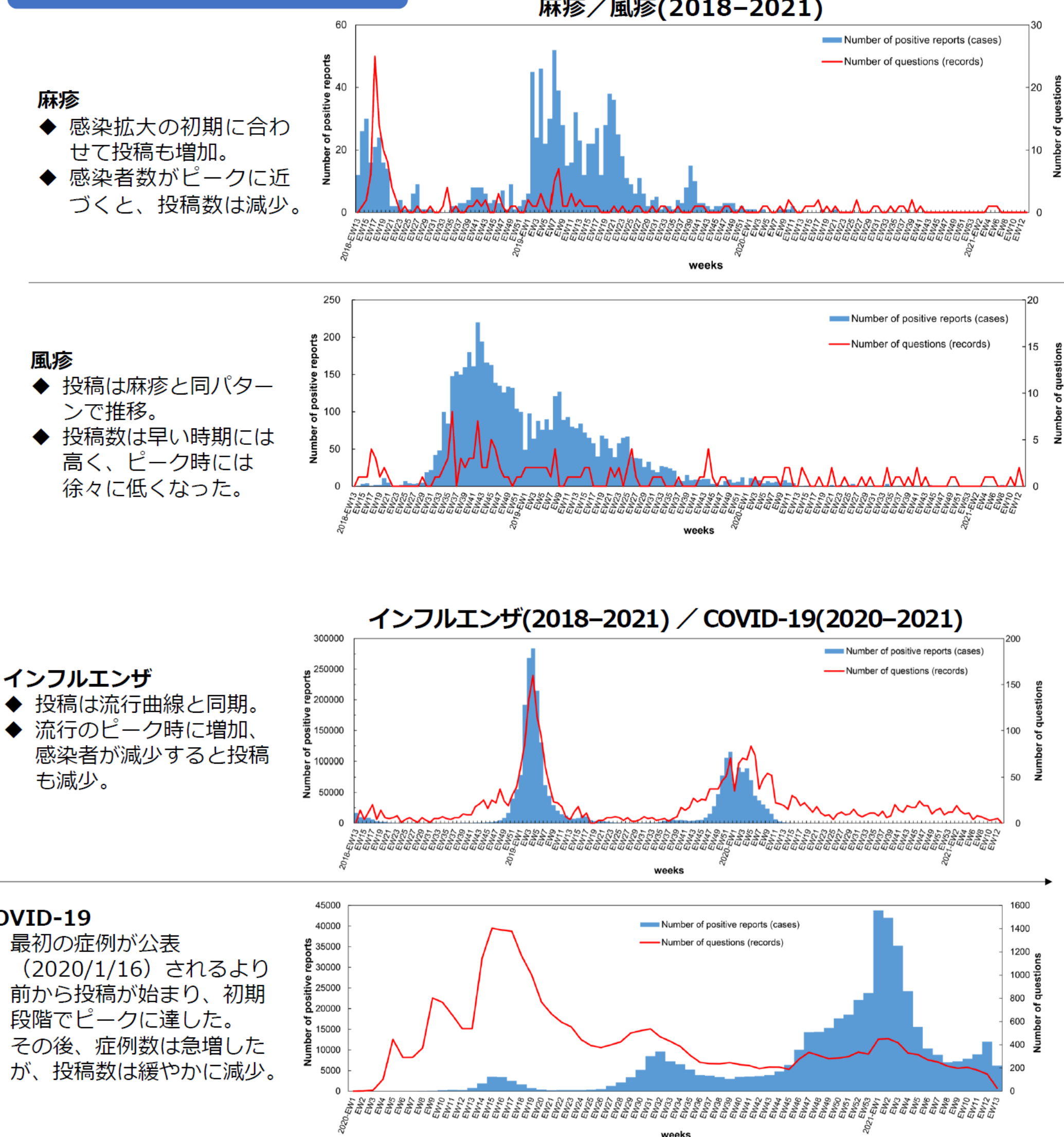
Share_d(t): 語 t が疾患 d の投稿群の中で占める割合
baseline_d: 全投稿(4疾患合計)における「疾患 d の全語数の割合」

$$BC(v) = \sum_{s \neq t} \frac{\sigma_s(v)}{\sigma_t}$$

BCが高い語: 複数の話題をまたいで登場する文章の流れをつなぐ重要語
σ_s(t): 語 s と語 t を結ぶ最短経路の数
σ_t(w): その最短経路のうち、語 v を通る経路の数
すべての語へ a(s,t) について計算し、合計することで v の BC値となる

結果

感染者数と投稿数の週別推移



形態素解析

疾患名	関連テーマ	疾患に特異的な抽出語 Lift ≥ 1.5
麻疹	ワクチン接種、検査関連	風疹(7.14)、打つ(4.08)、接種(3.97)、ワクチン(3.73)、抗体(2.62)、ワクチン接種(2.59)、子ども(2.48)、受ける(2.27)
風疹	ワクチン接種、検査、妊娠関連	麻疹(8.10)、妊婦(8.10)、妊娠(6.24)、抗体(5.13)、結果(4.83)、ワクチン(3.88)、打つ(3.47)、受ける(2.87)、ワクチン接種(2.86)
インフルエンザ	発熱症状関連	℃(2.72)、熱(2.71)、COVID-19(2.72)、風邪(2.72)、症状(2.04)、病院(1.81)、出る(1.56)
COVID-19	感染者数、感染予防関連	人(2.72)、感染者(2.72)、マスク(2.72)、日本(2.72)、感染(1.62)

- ◆ Lift値は疾患ごとの全語数で標準化するため、本当に特徴的な頻出語が分かる。
- ◆ 麻疹: ワクチン関連の語が高頻度で抽出。
- ◆ 風疹: 同様にワクチン関連語が並んだが、妊娠に関連する語が高いLift値だった。(注) 麻疹と風疹は【麻疹風疹混合ワクチン】が形態素に分解されるため、各疾患の上位Liftに現れる。
- ◆ インフルエンザ: 臨床症状に関連した語群の頻度が高い。COVID-19と比較した投稿から抽出されたため【COVID-19】が頻出。
- ◆ COVID-19: 【人】は感染者数の単位として多く抽出された。

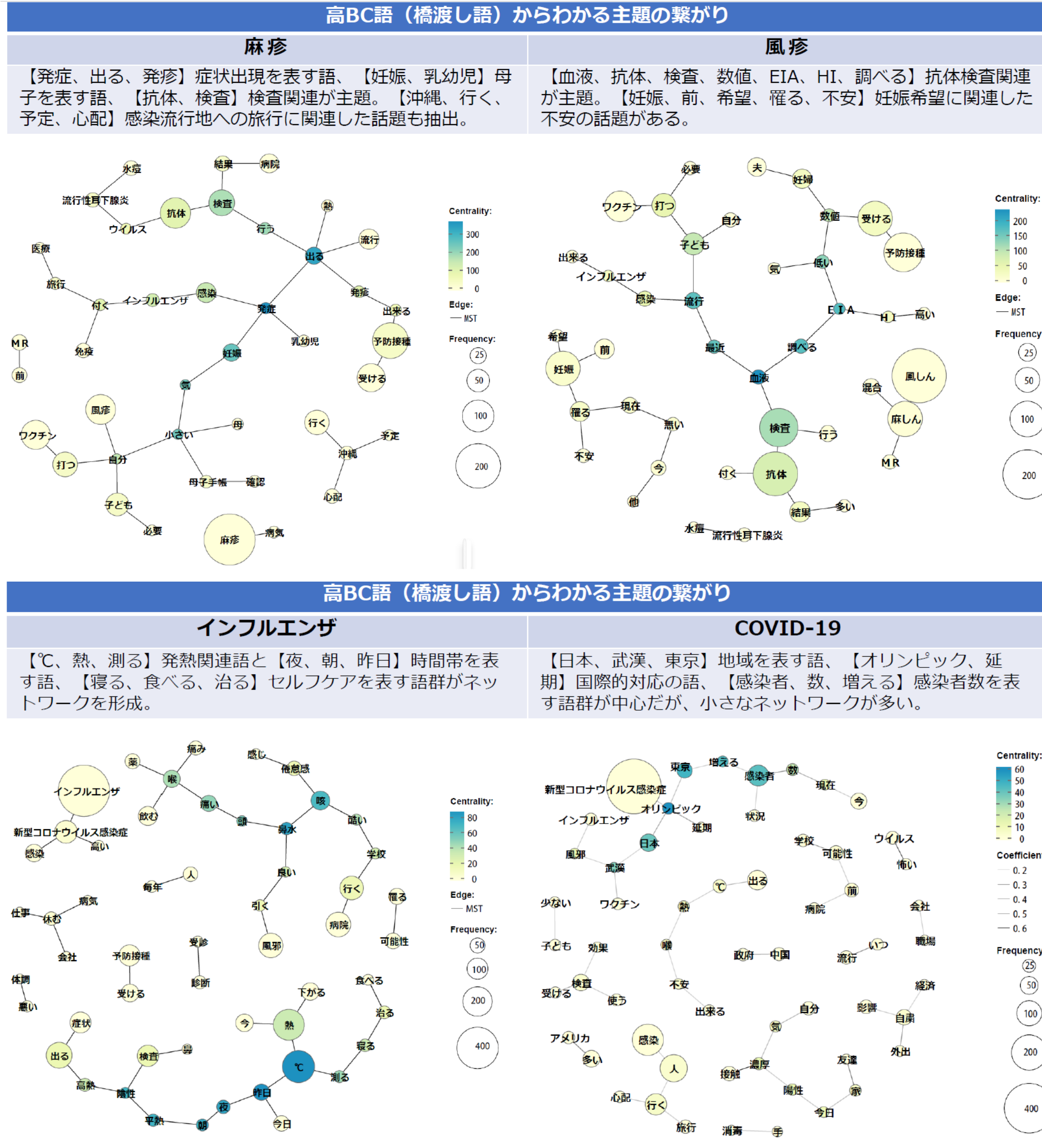
階層的クラスタ分析

併合水準からクラスター数を決定し、形成されたテーマにタイトルを設定。

疾患名	代表的なクラスターテーマ
麻疹	検査・診断、ワクチン接種、接種歴の確認、母子・発症、症状・流行、旅行不安、罹患不安
風疹	検査・診断、ワクチン接種、接種歴の確認、家庭内感染、不妊治療中のワクチン接種、検査法、罹患不安
インフルエンザ	ワクチン接種、感染予防、発熱の日内変動、症状(発熱以外)、症状(発熱)、セルフケア、受診判断、外出判断、体調不安
COVID-19	感染恐怖、国際的対応、感染者数、他者・陽性、受診判断、濃厚接触、外出自粛、感染拡大の不安、経済的影響、感染予防

- ◆ 麻疹・風疹: 母子手帳紛失で接種歴不明、旅行先でアウトブレイク、感染の不安。ワクチン接種により避妊期間が必要となり妊娠タイミングを逃す不安。
- ◆ インフルエンザ: 発熱など症状出現時にどう対応するべきか分からない。
- ◆ COVID-19: 未知の感染症に罹患する恐怖、国内外の感染状況、感染対策に関心がある。

共起ネットワーク



考察

疾患ごとに情報探索行動は違うだろうか

- 麻疹
 - 接種歴・抗体など免疫状態の確認が中心(母子手帳を紛失・記録不明)
 - 流行地域への旅行を続けてよいかどうかを相談
 - 妊娠中・乳幼児など母子に特化したリスク判断が多い
- 風疹
 - 妊娠希望・不妊治療中の接種タイミング相談(打つべきか/いつ打つか)
 - HI・EIA など抗体検査法と判定値への混乱(どちらの検査法を信じるべきか)
- インフルエンザ
 - 発熱が出たときに受診すべきか、様子を見るべきかの判断(セルフケア)
 - 受診の要否・タイミング、解熱後の登校・出勤再開時期の判断
 - その日の行動を決める日常的・実務的な意思決定支援を求めている
- COVID-19
 - 自分が検査の対象か/濃厚接触に当たるかを判定したい
 - 緊急事態宣言・自粛要請・政策的ルールを自分の生活にどう当てはめるか
 - 地域の感染状況・感染者数(例:東京やイベント)に応じた行動可否の相談

疾患ごとにリスク認知はどう異なるか

- 麻疹
 - 子どもの病気の認識から、感染症として重要度は低く見られがち。
 - しかし、妊娠中・流行地域への旅行前にはリスク認知が急上昇。
 - 接種歴が不明なケースが多く、「免疫が分からないこと」自体が不安要因となる。
- 風疹
 - 妊娠希望者・不妊治療中でリスク認知が突出して高い。
 - HI/EIAの違い、抗体価の読み方、接種後の避妊期間など、検査・スケジュールの解釈が難しい点がストレスを増幅。
 - 家族の意見や社会的期待も影響し、接種行動の迷いが生じやすい。
- インフルエンザ
 - リスク認知は、症状の強さやケアの必要性により変動する。
 - 感染そのものより、日常生活への影響(病体・登校・看病)が判断軸になる。
- COVID-19
 - 個人の感染リスクを超えた、社会全体のシステマ的リスクとして認識される。
 - 行動制限、濃厚接触の基準、検査制度など、政策やルールの変化がそのままリスク認知を揺さぶる。

情報ニーズ構造から導くコミュニケーション戦略

- 麻疹・風疹
 - 接種歴・抗体検査・妊娠に関する不安に答えることが必要。
 - 妊娠希望者・妊婦とその家族に対し、
 - ① 抗体検査の意味
 - ② 接種タイミング
 - ③ 避妊期間の考え方を含む、分かりやすいガイダンスの提供
 - いつでも自分の接種歴を確認できる記録システムの整備
- インフルエンザ
 - 発熱や症状の変化に応じた
 - ① 受診の目安
 - ② 登校・出勤再開のタイミング
 - ③ 家庭内での看病方法など、症状ベースの実務的情報を提示。
- COVID-19
 - 緊急事態宣言・濃厚接触の定義・検査基準など、政策・制度の変更を自身の生活にどう当てはめるか示す。

結論

- ◆ 疾患ごとの情報探索行動とリスク認知には、明確な違いがあった
- ◆ 麻疹・風疹: 妊娠関連の不安や免疫確認への関心が強く、流行初期に投稿が集中した。
- ◆ インフルエンザ: 発熱時の行動判断や自宅療養に関する実践的な情報ニーズが顕著であった。
- ◆ COVID-19: 世界情勢、政策変更や社会的制約への対応に関する投稿が持続し、社会システムへの不確実性が反映された。

- ◆ 匿名Q&A投稿は、一般市民の「本音の情報ニーズ」を映す
- ◆ 従来の調査では捉えにくい、迷いや判断の背景が投稿に現れた。
- ◆ 医学的関心と生活上の判断が交差する構造が、共起ネットワーク分析で可視化された。
- ◆ 公衆衛生における実務的示唆
 - 疾患特異的な情報ニーズに応じたリスクコミュニケーションが必要。
 - Q&Aプラットフォームは、感染症流行時の『社会的センサー』として活用可能。

謝辞

本研究では、国立情報学研究所のIDRデータセット提供サービスによりLINEヤフー株式会社のYahoo! 知恵袋データ(第3版)を利用させていただきました。ここに厚く御礼申し上げます。

COI開示

演題発表に関連し、開示すべきCOI関係にある企業等はありません。

倫理的配慮

本研究は国際医療福祉大学研究倫理審査委員会において承認を得た(承認番号22-1g-91)。