一般口演 | 第40回医療情報学連合大会（第21回日本医療情報学会学術大会） | 一般口演

一般口演18
自然言語処理

2020年11月21日(土) 11:15 〜 12:38 F会場 (イベントホール・特設会場1)

[4-F-1-01] 自然言語処理を用いた WebQAサイトからのがん患者の医療ニーズ抽出

*勘場 大[1]、眞鍋 雅恵[1]、若宮 翔子[1]、荒牧 英治[1] （1. 奈良先端科学技術大学院大学）
*Masaru Kamba[1], Masae Manabe[1], Shoko Wakamiya[1], Eiji Aramaki[1] （1. 奈良先端科学技術大学院大学）

キーワード：Natural Language Processing, Worry Classification, Breast Cancer

【背景】病気に罹った時，何らかの悩みを抱える人は少なくないだろう．実際に， Tobyoというブログのまとめサイトには，2020年6月時点で約1,400疾患，約62,000もの闘病記やブログが収録されている．なかでも治療が長期化しがちな乳がんに関するブログは多く，約6,600件の闘病記およびブログがヒットした．また， QAサイトである Yahoo!知恵袋においても乳がんを検索すると約60,000件の質問がヒットした．これらの中には医療ニーズに関係する悩みが含まれており，解決方法を実用化することは患者にとって有益であると考えられる．しかしながら，その莫大な量のデータから人手で課題をピックアップするのには限界がある．【目的】本研究では様々な自然言語処理の技術を用いて，これらのテキストデータから悩みを抽出し分類する．【材料・方法】悩みを抽出する対象は，乳がんについての Tobyoブログや Yahoo!知恵袋の質問とし，悩みの分類には静岡分類を用いる．静岡分類とはがん体験者にアンケート調査を実施し，悩みや負担をデータベース化したものである．まず (1) ブログや Yahoo!知恵袋の質問のテキストデータを分散表現に変換し，次に (2) 静岡分類で分類されている実際の悩みのテキストデータを分散表現に変換し，(1)と(2)の関連性を COS類似度で悩みを分類した．【結果】2018年1月1日から2020年6月9日の間に Yahoo!知恵袋に投稿された7,524件の質問を分類した結果，最も多い分類項目は「5.2.1.2. 医師の説明と理解 (2,097件)」であった．同様に，ブログを分類した結果についても検討する． 今後，分類の精度の評価及びこの技術を用いて，変わりゆく患者の悩みを捉え，臨床へ役立てたい．

©一般社団法人 日本医療情報学会

自然言語処理を用いたWebQAサイトからのがん患者の医療ニーズ抽出

勘場 大[*1]、眞鍋 雅恵[*1]、若宮 翔子[*1]、荒牧 英治[*1]
*1 奈良先端科学技術大学院大学

Medical Needs for Breast Cancer Patients from WebQA Site using NLP

Masaru Kamba[*1], Masae Manabe[*1], Shoko Wakamiya[*1], Eiji Aramaki[*1]
*1 Nara Institute of Science and Technology

[Background] Many people have some trouble when they get sick. In fact, as of June 2020, a summary site called Tobyo blog contains about 1,400 diseases and about 62,000 stories. Among them, there are many blogs about breast cancer, which tends to have long-term treatment; about 6,600 stories of fight against the cancer were found. A Japanese Q&A site, Yahoo! Chiebukuro, contains about 60,000 questions. These include the concerns related to medical needs. It could be beneficial to patients if solutions to them are put into practice. However, there is a limit to manually picking up such concerns to be solved from the enormous amount of data. [Objective] In this study, natural language processing techniques are used to extract and classify the concerns of patients. [Methods] The targets for extracting the concerns are the questions from Yahoo! Chiebukuro about breast cancer. Furthermore, Shizuoka classification is used to classify concerns. The Shizuoka classification is a database of patients' concerns through a cancer survivor survey. First, we (1) convert the text data of questions from Yahoo! Chiebukuro to vector representation. Second, we (2) convert the text data of narrative data of Shizuoka classification to vector representation. Finally, the relevance of (1) and (2) were classified by cosine similarity. [Results and discussions] As a result of classifying 6,993 questions posted on Yahoo! Chiebukuro from January 1, 2018, to July 31, 2020, the most common class was “Concerned about cancer to subjective symptoms.” In the future, we plan to evaluate the accuracy of the classification and use this method to capture the changing concerns of patients and to use them clinically.

Keywords: Natural Language Processing, worry classification, web data, Breast Cancer

1. はじめに

病気に罹った時，何らかの悩みを抱える人は少なくないだろう．実際に，TOBYO[8] には，2020 年 6 月時点で約 1,400 疾患にわたる約 62,000 もの闘病記やブログが収録されている．なかでも治療が長期化しがちな乳がんに関するブログは多く，約6,600件の闘病記およびブログがヒットした．また，QA サイトである Yahoo!知恵袋においても，約 60,000 件の乳がんに関する質問がヒットした．これらの中には医療ニーズに関係する悩みも含まれており，解決方法を実用化することは患者にとって有益であると考えられる．さらに，先行研究においては，患者から直接得られた情報（患者報告アウトカム）に注目し，その情報がヘルスケアの変革に役立つ可能性があると述べている[1),2)]．しかしながら，その莫大な量のデータから人手で課題をピックアップするのには限界がある．そこで，自然言語処理の技術を用いて，医療ニーズの特定を試みる．

本研究では自然言語処理の技術を用いて，これらのテキストデータから悩みを抽出し分類する．悩みを抽出する対象は，乳がんについての Yahoo!知恵袋の質問とし，悩みの分類には静岡分類を用いる．ここで，静岡分類とは，2003 年に 7,855 人，2013 年に 4,054 人のがん体験者にアンケート調査を実施し，悩みや負担をデータベース化したものである[4) 5) 6)]．宮部，島本，荒牧 [3)] は静岡分類の大項目に対して分類を行なったが，本研究においては小項目の分類を試みる．さらに，悩みが地域に依存している可能性もあるため，都道府県ごとに悩みの違いがあるかも評価する．

2. 方法

本節では Yahoo!知恵袋のデータを静岡分類に分類するために用いた分類モデル及び分類のアルゴリズムを示す．

2.1 学習データ

本節では，モデルを学習するために用いた3種類のコーパスを示す．

A　静岡分類のコーパス
B　Yahoo!知恵袋のデータを人手で静岡分類に分類したコーパス
C　AとBを統合したコーパス

まず，A の静岡分類のデータは，静岡分類のデータベース[6)] から取得した，静岡分類コードと静岡分類カテゴリ（以降，単にカテゴリと呼ぶこともある）がラベル付けされたがん体験者の悩みのテキストデータである．以降，静岡分類コーパスと呼ぶこととする．

次に，B の Yahoo!知恵袋のデータは，2018 年 1 月 1 日から 2020 年 6 月 9 日までに投稿された乳がんに関する質問からランダムに 1,000 件抽出し，1件につき最大 3 種類の静岡分類カテゴリを人手で割り当てたラベル付きデータである．以降，知恵袋コーパスと呼ぶこととする．実際には 1,000 件のうち 456 件に静岡分類カテゴリを割り当てることができたが，残りの 546 件については該当する静岡分類カテゴリが存在しなかった．最大 3 種類のカテゴリで累積的に分類された質問数の合計は 661 件で，133 種類の静岡分類カテゴリに割り当てられた．表 1 に上位 10 種類の静岡分類カテゴリに分類された質問件数を示す．最も多かったカテゴリは「自覚症状でがんを心配」であり，ラベル付けされたデータの 24.2%がこのカテゴリに該当した．また，「医師に質問や心配事を言い出しにくい」というカテゴリが上位に含まれていることは，医師に言い

出しにくいことから，Yahoo!知恵袋に投稿されたと考えられる．

最後に，C は静岡分類コーパスと知恵袋コーパスを合わせたラベル付きデータである．以降は統合コーパスと呼ぶこととする．

表 1 人手による Yahoo!知恵袋の質問の静岡分類カテゴリへの分類結果（上位 10 種類）

静岡分類コード	静岡分類カテゴリ	件数
16.3.1.1.	自覚症状でがんを心配	160
16.2.1.1.	がん検診に関すること	85
12.2.4.1.	がんに対する知識不足による不安	42
16.3.2.1.	がんの疑いに関すること(その他)	39
9.1.2.2.	医師に質問や心配事を言い出しにくい	17
3.2.2.2.	結果やその動向が心配	17
12.1.1.1.	再発・転移するかもしれない不安	14
3.2.1.6.	検査の受診に関すること(その他)	11
3.1.1.1.	治療選択の迷い	10
3.2.2.3.	検査の受診に関すること(その他)	9

2.2 分類対象データ

2018 年 1 月 1 日から 2020 年 7 月 31 日までに Yahoo!知恵袋に投稿された 7,993 件の質問のうち，2.1 節で示した知恵袋コーパスの 1,000 件のラベル付き質問（学習データ）を除外した 6,993 件を静岡分類カテゴリの分類対象データとした．

2.3 分類アルゴリズム

本節では 2.1 節の学習データを用いて，2.2 節の分類対象データを静岡分類カテゴリに分類するアルゴリズムを示す．

1. 2.1 節で示した 3 種類のコーパスを静岡分類カテゴリごとに TF-IDF で重み付けした単語ベクトルに変換する
2. 2.2 節で示した分類対象データを Yahoo!知恵袋の質問ごとに TF-IDF で重み付けした単語ベクトルに変換する
3. 1. と 2. の Cos 類似度を計算し，最も関連度の高い静岡分類カテゴリに分類する

ここで，単語ベクトルへの変換の際，形態素解析辞書の mecab-ipadic-NEologd[7] を用いて，名詞，動詞，形容詞を対象に抽出した．ただし，記号や数字は除外した．また，TF-IDF の計算には sklearn.feature_extraction.text モジュールに含まれる TfidfVectorizer 関数を利用した．本関数におけるオプションはデフォルトの設定を利用した．

以降，静岡分類コーパス，知恵袋コーパス，統合コーパスを用いて構築した分類モデルのことをそれぞれ静岡分類モデル，知恵袋モデル，統合モデルと呼ぶことにする．

3. 結果

3.1 Yahoo!知恵袋の月ごとの投稿数

本節では Yahoo!知恵袋のデータを要約した結果を示す．2018 年 1 月 1 日から 2020 年 7 月 31 日の間に Yahoo!知恵袋に投稿された 7,993 件の質問を，月ごとに集計し，その推移を確認した（図 1）．その結果，図 1 より，2020 年 2 月から 2020 年 5 月にかけて急激に質問数が増えていることがわかる．これは COVID-19 の影響によるものと考えられる．

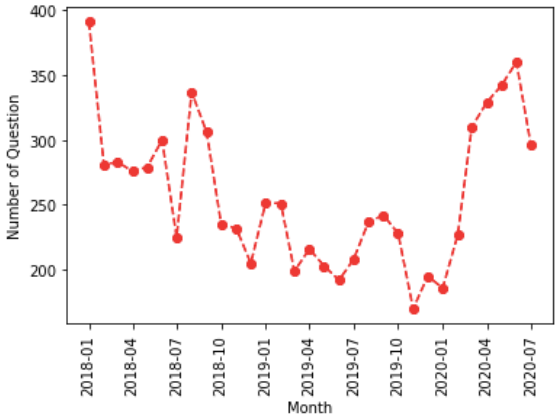

図 1 質問数の 1 ヶ月ごとの推移

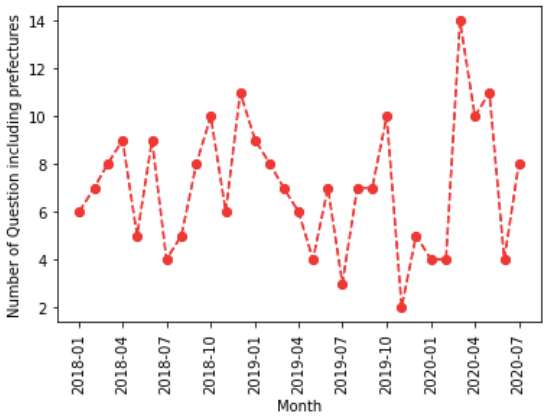

図 2 都道府県名が含まれていた質問数の 1 ヶ月ごとの推移

3.2 分類結果

まず，各分類モデルにより，2.2 節の分類対象データを静岡分類カテゴリに分類した．その結果のうち，分類された質問数が上位 10 カテゴリとその件数をそれぞれ表 2 に示す．最も多くの質問が分類されたカテゴリは，静岡分類モデルでは「過去に受けた検診・診断・治療に対する不満 (370 件)」，知恵袋モデルでは「自覚症状でがんを心配 (2,026 件)」，統合モデルでは「自覚症状でがんを心配 (1,661 件)」であり，知恵袋モデルと統合モデルは同じ分類カテゴリであった．また，全ての分類モデルの結果の上位 10 カテゴリのうち，「がん検診に関すること」，「自覚症状でがんを心配」，「がんの疑いに関すること(その他)」の 3 カテゴリが共通していた．さらに，静岡分類モデルの結果は，知恵袋モデル及び統合モデルの結果と大きく異なったが，知恵袋モデルと統合モデルの結果を比較すると，似たような結果となった．しかし，知恵袋モデルの結果の上位 10 カテゴリには，「漠然とした医療費や経済面に対する不安」，「受けた治療（選択）が正しかったか」，「治療選択の迷い」が含まれていたが，統合モデルの結果には「結果やその動向が心配」，「がんのことが頭から離れない」，「医師の言葉や態度」が含まれており，違いも見られた．

次に，分類モデルによる分類カテゴリと実際の質問の整合性を確認する．表 3 に各分類モデルで分類された質問数が最も多いカテゴリに分類された質問の原文と Cos 類似度を示す．ここでは Cos 類似度が高い順に上位 3 件の質問を示した．この結果，静岡分類モデルにおいて，上位 2 件の質問に対しては正しく分類できているが，3 番目の質問に対しては正しく分類できていないことがわかった．一方，知恵袋モデル及び統合モデルにおいては両方とも正しく分類できていることがわかった．

3.3 モデルの精度

本節では各モデルの精度を評価する．各モデルの精度計算アルゴリズムは以下の通りである．

静岡分類モデルの精度計算アルゴリズム：

1. 静岡分類コーパスを用いて，静岡分類カテゴリごとにがん体験者の悩みのテキストデータを TF-IDF で単語ベクトルに変換する
2. 知恵袋コーパスを用いて，静岡分類カテゴリごとに，知恵袋に投稿された質問を TF-IDF で単語ベクトルに変換する
3. 1. の単語ベクトルと 2. の単語ベクトルの Cos 類似度が最大のカテゴリ，及び，上位 10 カテゴリを求める
4. 3. のカテゴリが，知恵袋コーパスの正解カテゴリ（最大 3 つ）に含まれている割合を計算する

知恵袋モデルでは 5-fold Cross Validation で精度を計算する．具体的には，知恵袋コーパスを学習データと検証データに 5 分割し，それらの Cos 類似度を計算し，静岡分類モデルの精度計算アルゴリズムと同様に静岡分類カテゴリに含まれている割合を計算し，それらの平均，不偏標本標準偏差，中央値を求める．

統合モデルでは知恵袋モデルの学習データに静岡分類コーパスを含めて，知恵袋モデルと同様のアルゴリズムで精度を計算する．

表 4 に上記のアルゴリズムで計算された精度を示す．静岡分類モデルの正解率は約 10%であり，Top 10 での結果でも約 30%であった．一方，知恵袋モデル及び統合モデルでの正解率は平均値及び中央値の両方で約 50%であり，Top 10 での正解率は約 70%であった．

表 4 各モデルの精度

評価指標	統計量	静岡分類モデル	知恵袋モデル	統合モデル
Top 1	平均値	0.1096	0.4891	0.4781
	標準偏差	-	0.017	0.018
	中央値	-	0.4835	0.4725
Top 10	平均値	0.2946	0.6960	0.7062
	標準偏差	-	0.030	0.021
	中央値	-	0.7015	0.7106

3.4 地域間の悩み

本節では Yahoo!知恵袋の質問の地域差の検討結果について示す．質問中に都道府県名が含まれていた質問数の 1 ヶ月ごとの推移を図 2 に示す．都道府県名が含まれている質問は非常に少なかったものの，COVID-19 の流行前後の傾向は図 1 と似た結果となった．

図 3 は都道府県ごとに投稿された質問数（上位 10 都道府県）（左），及び各都道府県の人口で調整した質問数（上位 10 都道府県）（右）を示す．ただし，一つの質問に複数の都道府県名が含まれている場合は，それぞれの都道府県を集計対象として含めた．ここで，図 3 左では，上位に東京都や大阪府など人口が多い都道府県が含まれているが，福島県が 4 番目に多い結果となった．実際に，質問数を人口で割った指標を示した図 3 右においては福島県が最も高い値となった．

表 5 に，図3の上位にある東京，大阪，福島，京都に対して，統合モデルで分類された静岡分類カテゴリ及びその件数を示す．ただし，4 つの都道府県のうちのいずれかの都道府県で 2 件以上の質問が分類されたカテゴリを出力対象とした．この結果，すべてにおいて「がん検診に関すること」に分類された件数が最も多かった．これはがん検診を受診する場所に関する質問であるため，地名と共起することはリーズナブルであると考えられる．また，福島においてはがんの疑い，がんの心配，がんのことが頭から離れない，といったがんに対する心配が多く投稿されていた．

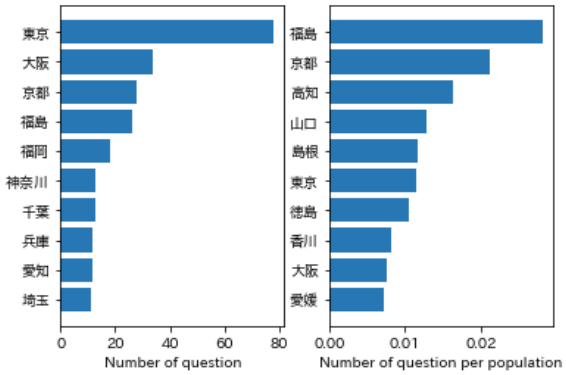

図 3 都道府県毎の質問数（左）と人口あたりの質問数（右）

表 5 各都道府県における質問の静岡分類カテゴリとその件数（2 件以上）

静岡分類カテゴリ	東京	大阪	京都	福島
がん検診に関すること	19	13	5	8
がんに対する知識不足による不安	7	2	3	0
医師に質問や心配事を言い出しにくい	3	3	2	1
自覚症状でがんを心配	6	1	2	1
今後の医療費や生活費等の心配に関すること(その他)	2	1	2	0
がんの疑いに関すること(その他)	3	1	0	4
医師の言葉や態度	3	0	2	0
ホルモン療法等による生理不順・変化	2	1	2	2
病院・医師選択のための情報入手が難しい	2	0	1	0
再発・転移するかもしれない不安	2	0	0	2
今後のがん検診のための病院選択	2	1	0	0
治療法決定時の迷いや選択の難しさに関すること(その他)	2	0	0	0
うつ状態(その他)	2	0	0	0
がんのことが頭から離れない	0	0	0	2

4. 考察

知恵袋モデルは Yahoo!知恵袋の質問を分類することに最適化された分類モデルであるが，全ての静岡分類のカテゴリを網羅しているわけではない．そのため，分類精度の向上には知恵袋モデルのさらなる拡張が必要になると考えられる．そこで，本研究においては統合モデルを使った結果で解釈する．実際に 3.3 節で示した精度では，知恵袋モデルと統合モデルに大きな違いは見られなかった．統合モデルの分類結果では，多くが自覚症状でがんを心配しているという悩みであった．

また，日本においては 2 月から COVID-19 に関する悩みも含まれると考えられることから，静岡分類では正しく分類できていない可能性もある．そのため，2 月以降の悩みを分類するためには新たな悩みのカテゴリを定義する必要があると考えられる．

5. 結論

知恵袋モデル及び統合モデルで，多くの質問が「自覚症状でがんを心配」や「がん検診に関すること」，「がんの疑いに関すること(その他)」のカテゴリに分類された．これらは静岡分類の 16 番目の大項目である「がんの予防・がん検診・がんの疑い」に該当する．多くの質問投稿者がこの静岡分類カテゴリの悩みを持っていると解釈することができた．統合モデルにおいてこれらのカテゴリの和は全体の 40.9%を占める．

今後，このような悩みの受け皿となるサービスが必要になると思われる．情報学の観点から模索を継続したい．

さらに，当初は TOBYO のデータにおいても静岡分類カテゴリに分類し，悩みを抽出する予定であったが，予備調査の結果，ブログには悩みとは関連が低い内容も多く投稿されることから，今回の分類対象データからは除外した．ただし，ブログには多くの悩みの情報を含む投稿も存在するため，ノイズを多く含むデータからも適切に分類できるモデルの構築及び手法の検討を今後の継続課題としたい．

謝辞

本研究において，松本妙子氏，友廣公子氏に Yahoo!知恵袋の静岡分類のラベルを設定いただいた．また，矢田竣太郎氏から多大な助言をいただいた．この場を借りて厚くお礼を申し上げ，感謝の意を示す．

参考文献

1) KE Lasch, P Marquis, M Vigneux,et al. Pro development: rigorous qualitative research as the crucial foundation. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation, 2010; Vol. 19, No. 8 : 1087-1096.
2) N Black. Patient reported outcome measures could help transform healthcare. British Medical Journal, 2013; 346:.
3) 宮部真衣，島本裕美子，荒牧英治. Web テキストから医療の悩みを抽出する:がん体験者の悩みの自動分類. 第 13 回情報科学技術フォーラム. 2014;95-96.
4) 「がんの社会学」に関する合同研究班. がん体験者の悩みや負担等に関する実態調査報告書 がんと向き合った 7,885 人の声. 2006.
5) 「がんの社会学」に関する研究グループ. 2013 がん体験者の悩みや負担等に関する実態調査報告書 がんと向き合った 4,054 人の声. 2016.
6) 静岡県立 静岡がんセンター. 静岡分類. [https://www.scchr.jp/cancerqa/start_shizuoka.html(cited 2020-Sep-07)]
7) Toshinori Sato. mecab-ipadic-NEologd : Neologism dictionary for MeCab. [https://github.com/neologd/mecab-ipadic-neologd (cited 2020-Sep-07)]
8) 株式会社イニシアティブ. [https://www.tobyo.jp (cited 2020-Sep-07)]

表 2 各分類モデルによる予測結果 (上位 10 種類)

静岡分類モデル			知恵袋モデル			統合モデル		
静岡分類コード	静岡分類カテゴリ	件数	静岡分類コード	静岡分類カテゴリ	件数	静岡分類コード	静岡分類カテゴリ	件数
10.2.1.1.	過去に受けた検診・診断・治療に対する不満	370	16.3.1.1.	自覚症状でがんを心配	2026	16.3.1.1.	自覚症状でがんを心配	1661
3.1.3.5.	受けた治療(選択)が正しかったか	359	16.2.1.1.	がん検診に関すること	1493	16.2.1.1.	がん検診に関すること	702
10.1.1.1.	医師の言葉や態度	208	12.2.4.1.	がんに対する知識不足による不安	697	16.3.2.1.	がんの疑いに関すること(その他)	494
16.2.1.1.	がん検診に関すること	188	16.3.2.1.	がんの疑いに関すること(その他)	499	12.2.4.1.	がんに対する知識不足による不安	419
12.2.1.2.	がんの進行による苦痛、治療の副作用や後遺症がでるかもしれない	180	3.2.2.2.	結果やその動向が心配	164	3.1.3.5.	受けた治療(選択)が正しかったか	255
16.3.2.1.	がんの疑いに関すること(その他)	137	12.1.1.1.	再発・転移するかもしれない不安	153	3.2.2.2.	結果やその動向が心配	234
16.3.1.1.	自覚症状でがんを心配	125	12.2.2.1.	漠然とした医療費や経済面に対する不安	134	12.1.1.1.	再発・転移するかもしれない不安	225
12.1.2.1.	初回診断(治療)時、すでに転移していた衝撃とつらさ	105	9.1.2.2.	医師に質問や心配事を言い出しにくい	132	9.1.2.2.	医師に質問や心配事を言い出しにくい	137
11.1.24.4.	ホルモン療法等による生理不順・変化	97	3.1.3.5.	受けた治療(選択)が正しかったか	84	12.3.2.3.	がんのことが頭から離れない	111
12.4.1.4.	なぜ自分(だけ)が、どうして自分なのか	96	3.1.1.1.	治療選択の迷い	75	9.1.1.1.	医師の言葉や態度	93

表 3　各分類モデルで最も多いカテゴリに分類された質問の原文と Cos 類似度 (上位 3 件)

分類モデル	原文	Cos 類似度	静岡分類カテゴリ
静岡分類モデル	(前略)半年前に会社の健康診断で乳がんの疑いによる再検査となりました。その再検査でシコリが2つみつかったのですが、1つは大きさや形からまず問題ないだろうということで、もう1つのほうだけ細胞を取って検査しました。結果は良性で一応半年後にもう一度検査することになりました。そして先日半年後の再検査に行ってきたのですが、前の検査で細胞を取らなかった方のシコリが大きくなっていたので、こちらも調べてみましょうとなり、細胞を取って検査した結果、悪性腫瘍と診断されました。(中略)最初の検査で安易に大丈夫だと判断したのは医者の誤診にはならないのでしょうか？追加で質問ですが来週から詳しい検査が始まるのですが、正直担当の先生への不信感があり、調べたところその病院の評判も悪いため他の病院に変わりたいと考えています。この場合、別の病院に移ってもよろしいのでしょうか？(後略)	0.3916	過去に受けた検診・診断・治療に対する不満
静岡分類モデル	(前略)検診を受けた病院で結果説明と超音波検査を受けました。医師（高齢）は市から来た書類を見て、「これなら問題ない」といい、とりあえず超音波検査を受け「異常なし」と言われました。頭から「異常なし」という前提で診ている感じでした。(中略)マンモグラフィの結果は書類のみで画像はありませんでした。検診時から全く触診が無かったので、こちらから念のためお願いしましたが何も言われませんでした。(中略) 高齢医師で意思疎通がうまく行きにくいのとネットで悪評が沢山あったのが気になりました。大型掲示板にも書かれていました。★検査器具注文中ですが、別の病院に行って検査等することも可能でしょうか？★私の為だけに注文なのか不明ですが何か請求されたりしますか？(後略)	0.3851	過去に受けた検診・診断・治療に対する不満
静岡分類モデル	(前略)会社の健康診断の結果が先日返ってきました。乳がん検査の結果の欄に「要精密検査」の文字がありました。右胸に乳腺腫瘤の疑いがあるとのことです。去年乳がん検診(視触診、マンモグラフィ検査、超音波検査)を受けた際は乳腺症と書かれており、調べてみたら多くの方が言われることらしく安心していました。(中略)視触診では 異常なし となっておりました。ネットで調べ、セルフチェックを行いましたがしこり等の自覚症状はありません。健康診断の結果を見て、早急に精密検査の予約をしたのですが、一番早くて1か月も先の 7 月で予約がとれました。(中略)また検診を受けた日が生理の 5 日前で胸が張ったりしていたのですが、このことは診断結果になにか影響があるのでしょうか？ネットに乳がんは 40 代以降の方から発症されることが多いと書いてありましたが、若年性乳がんも全体のほんのわずかですがあるので、とても不安で怖いです。 (後略)	0.3700	過去に受けた検診・診断・治療に対する不満
知恵袋モデル	数日前左胸がちくちくする様な痛みがあり触ってみると乳首下あたりにしこりのようなコリコリしたものがありました。不安になり乳がん検診をしたところ触診とエコーをしました。結果はしこりなどなく乳腺自体がコリコリしているのでそれだろうと言われました。 (中略) エコーでしこりが見つかってない場合乳がんでないと思って大丈夫でしょうか？検診をしたときは生理前でこれも関係していますでしょうか？	0.6156	自覚症状でがんを心配
知恵袋モデル	(前略)私の胸にはしこりがあります。しこりは直径約 6 センチ、厚さ約 3 センチほどの丸いものです。 (中略) これは乳がんなのでしょうか？？それとも普通のことなのでしょうか？？または乳がんではない別の病気なのでしょうか？？(後略)	0.5703	自覚症状でがんを心配
知恵袋モデル	(前略)最近胸が張って痛かったのですが、2, 3 日前に左胸の外側の上部の方にしこりのようなものがあることに気が付きました。(中略)そして、今日お風呂の中で胸を触って確認していると、右胸の下の方にも同じようにしこりがあることに気が付きました。(中略)10 代では乳がんの確率が低いことは理解しています。気にしすぎなのでしょうか？調べているうちに乳腺症かなと思ったのですが、10 代でも乳腺症にはなりますか？(後略)	0.5645	自覚症状でがんを心配
統合モデル	(前略)私の胸にはしこりがあります。しこりは直径約 6 センチ、厚さ約 3 センチほどの丸いものです。ほとんど同じ大きさのものが、両胸どちらにもあります。(中略) これは乳がんなのでしょうか？？それとも普通のことなのでしょうか？？または乳がんではない別の病気なのでしょうか？？(後略)	0.5590	自覚症状でがんを心配
統合モデル	(前略)最初両胸にしこりみたいなものができました。そのしこりとともに胸も大きくなった感じです。(中略)生理前になると痛いです。このしこりはあってもいいものなのですか？みんなこんな感じなのでしょうか？親に聞く勇気もなく、病院にもなかなか行けないのですごく気になっています	0.5449	自覚症状でがんを心配
統合モデル	数日前左胸がちくちくする様な痛みがあり触ってみると乳首下あたりにしこりのようなコリコリしたものがありました。不安になり乳がん検診をしたところ触診とエコーをしました。結果はしこりなどなく乳腺自体がコリコリしているのでそれだろうと言われました。 (中略) エコーでしこりが見つかってない場合乳がんでないと思って大丈夫でしょうか？検診をしたときは生理前でこれも関係していますでしょうか？	0.5400	自覚症状でがんを心配