

難病の早期診断を実現する新しい問診項目の抽出手法

*西原陽子⁽¹⁾, 平塚義宗⁽²⁾, 村上晶⁽³⁾, 大澤幸生⁽¹⁾, 熊川寿郎⁽²⁾

(1) 東京大学大学院工学系研究科, (2) 国立保健医療科学院経営科学部,

(3) 順天堂大学医学部

〒113-8656 東京都 文京区 本郷 7- 3- 1

E-mail : nishihara@sys.t.u-tokyo.ac.jp

Abstract: 世界中には難病に苦しむ多くの人が存在する。一般的に難病の治療は困難であるため、患者は病気と付き合っていく中で、自らの症状が将来に向けてどのように変化するかを出来るだけ早期に認識するべきである。そのためには、難病の早期診断が必要となる。本研究では、患者自身も気づいていないような、難病の初期症状に関する情報を抽出する手法を提案する。予備的な実験において、難病の一つである網膜色素変性の患者から得られたエピソードを KeyGraph により可視化することにより、これまで明確にされていなかった初期症状に関連するキーワードを抽出できる可能性が示唆された。

1. はじめに

世界中で難病に苦しむ人が多数存在する。例えば、難病の一つとして知られるパーキンソン病の日本における患者数は、少なく見積もっても 12 万人に上ると見られている[Lozano 05]。一般に、難病は診断、治療が困難で、慢性の経過を辿ることが多い。難病の治療方法に関しては、56 疾患を対象とした国の特定疾患治療研究事業等の大がかりな取り組みが存在し、研究が進められている¹。

一般に難病に罹った患者は、まず病気を受け入れ、徐々に進行していく病状に対応しながら、将来自分の体がどう変化するか、生活が具体的にどう困難になるのかを知り、病気と付き合っていく。そのため、病気の早期発見が重要であり、患者は初期的な症状に気づいたら直ちに受診し、医療者は様々な面から診察し、早期に確定診断を下すことが望ましい。しかしながら、難病の早期発見に関する診断アプローチに関しては、未だ有効な方法が提案されていない。

そこで本研究では、難病の早期発見のための新しい診断アプローチに係る新たな手法として、患者の語るエピソードから病気の初期症状に係るキーワードを抽出する方法を提案する。この方法により、患者や医療者が把握していない初期症状に係る新たなキーワードを得ることが可能となり、それを問診に反映させることにより病気の早期発見・早期診断に貢献できると考えられる。

2. 難病の初期症状に係るキーワードの抽出手法

本手法の処理手順を説明する。初めに、患者に対して難病と診断されるまでに、日常生活で体験したエピソード、特に病気の初期症状を示唆するエピソードを語ってもらう。続いて、複数名の患者から得られたエピソードをテキストに起こし、データ可視化ツールの一つである KeyGraph を用いて、テキスト内のキーワード間の関係を可視化する。最後に、可視化された結果から、低頻度ではあるが重要なキーワードを、病気の初期症状を示唆する可能性が高いキーワ

¹ 難病情報センター (<http://www.nanbyou.or.jp/top.html>)

ードとして抽出する。

KeyGraph はデータ間の関係をグラフの形式で可視化するものであり、頻度の高いデータを黒いノード、共起する頻度の高いデータの集合体を黒いノードのクラスタ、低頻度ではあるが異なるクラスタ間に出現するデータを赤いノードとして可視化する[大澤 92]。KeyGraph で赤いノードとして表現されるデータは、データ間の関係をつなぐ重要なデータとして評価される。本手法では、患者が語るエピソードを **KeyGraph** を用いて可視化し、赤いノードとして評価されたキーワードを、日常生活の中に埋もれていた、病気の初期的な症状を示す可能性が高いキーワードとして抽出する。

3. 評価実験

本手法により、病気の初期的な症状を表すキーワードが得られるかどうかを評価する実験を行った。

3. 1. 評価実験

初めに、患者にエピソードを語ってもらい、エピソードをテキストに書き起こした。続いて、書き起こされたエピソードを一つにまとめ、**KeyGraph** を用いて可視化した。可視化に当たり、黒いノードは最大 75 個、赤いノードは最大 30 個と設定し、可視化された結果から、病気の初期的な症状を表す可能性が高い単語として、赤いノードに相当する単語を抽出した。

3. 2. 実験データ

本実験では、目の難病の一つである、網膜色素変性の患者が語るエピソードを実験データとして使用した。網膜色素変性は、進行性の夜盲、視野狭窄、視力低下を主症状とする遺伝性の疾患である。日本における患者数は約 23,000 人（1990 年の厚生省班研究による）となっている[疾病対策研究会]。本実験では、5 名の患者に協力を要請し、エピソードを獲得した²。エピソードの一部を表 1 に示す。テキストへと書き起こす際には、表記の統一を行った。

3. 3. 評価方法

評価では、本手法により抽出されたキーワードと、エピソードにおいてキーワードが含まれた文を眼科の医療者に示し、アンケートをとった。アンケートでは、示されたキーワードとエピソード中の文は、今までに医療者の間では知られていないが、網膜色素変性の初期的な症状を示唆する可能性が高いかどうかを質問した。その後、示唆する可能性が高いと評価されたキーワードとエピソードの数を評価した。

表 1 網膜色素変性の患者が語ったエピソードの一部。

それが原因かは知らないけど、3つ4つの頃から暗い所は殆ど駄目なんですよ。例えば、時代が時代ですから、道路でも裸電球 1 個なんですよ。それでも普通の子は皆歩いて、夜になっても行けるんですけど、私は離されてしまうと、駄目だったというのは、自分として分かっていた。学校行っても席は 3 番目から後ろに行かなかったし。職場に入っても、その当時、夜歩くと道路の水溜りに落っこちてしまう。階段は 1, 2 段踏み外すとか、そういう自覚症状があったんだけど。僕は、こういう性格だからか知らないけど、あまり気にしなかったんですよ、しょうがないなと思って。

²本研究は順天堂大学附属順天堂医院の倫理委員会の承認を得た上で実施された。対象患者は、順天堂大学眼科を受診中の網膜色素変性患者であり、エピソード使用に関しては書面により同意を得た。

3. 4. 実験結果

患者のエピソードを可視化した結果を図 1 に示し、図 1 において、赤いノードとして抽出されたキーワードを表 2 に示す。表 2 において、下線が付けられているキーワードは、病気の初期的な症状を示唆する可能性が高いと、医療者が評価したものになる。抽出された 28 個のキーワードの内、13 個が初期的な症状を示唆する可能性が高いと評価された。また、初期的な症状を示唆する可能性が高い文である、と評価されたものは 95 個中 30 個であった。その例を表 3 に示す。

網膜色素変性は遺伝性の病気であるため、病気の症状は早い人であれば乳幼児の頃から出始めるが、遅い人は 20 歳を超えてから出始めるなど、病気の進行に個人差が大きい。この点は語られたエピソードにも現れており、表 3 でも、中学生、高校生、社会人になってからなど様々な時期のエピソードが存在した。このことから、患者の年代に合わせて、質問項目を調整する必要があることが分かった。

4. 考察

抽出された 28 個のキーワードのうち、医療者によって初期的なエピソードに関連すると評価されたものは 13 個あった。このことから KeyGraph により抽出されるキーワードには病気の初期的な症状に係るものが含まれると分かった。

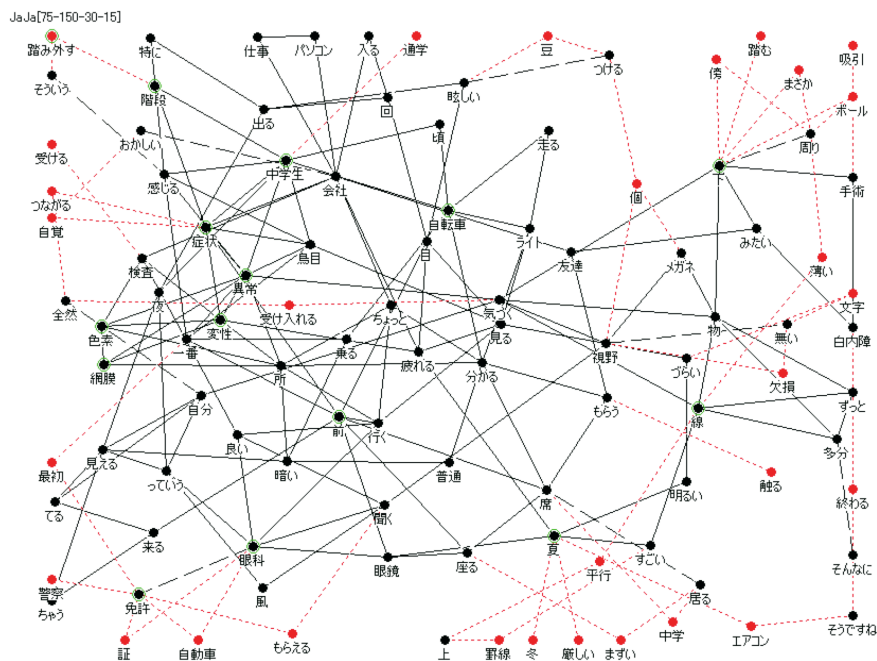


図 1 KeyGraph により可視化された網膜色素変性患者（5 名）が語ったエピソード。

表 1 図 1 において赤いノードとして可視化されたキーワード。

品詞	キーワード
動詞	踏み外す, 受ける, つながる, もらえる, 受け入れる, 触る, 終わる, 踏む
名詞	自覚, 最初, 警察, 証, 通学, 豆, 罫線, 冬, 中学, 平行, エアコン, 欠損, 文字, 個, 傍, 吸引, ボール
形容詞	厳しい, まずい, 薄い
副詞	まさか

表 2 初期的な症状と医療者が判断したキーワードと、エピソード中の文の例。（カッコ内の文は著者の一人が可読性向上のために追加したもの）

キーワード	エピソードに含まれた文
通学	・中学生の頃か自転車通学をしておりますて（、ライトだけでは道路が見えませんでした）.
踏む	・あとは、手術室実習のときに、下が見えなかったのかな、吸引のチューブを踏んでしまつて. ・下に置いてる物に気づかなくて踏んじったりとか.
薄い	・データを取っているときに、すごく薄いんですよ、線が、罫線が薄いんですよ. ・きっと、注意されても（体を曲げて文字を書くことを）止められなかったのは、白い所に緑の薄い線が書いてあるじゃないですか. ・（ノートの罫線の）周り（の枠線）は濃いので良いのですが、これは（罫線は）薄いじゃないですか.
罫線	・データを取っているときに、すごく薄いんですよ、線が、罫線が薄いんですよ. ・罫線が上手く追えなくなっちゃった、1 個上に上がったり、下がったりしちゃったわけです.
平行	・その線が、左側に例えば A 商品があるとして、左をずっと追っかけていって、平行に行に入れないといけない（. それが難しい）. ・今はたまたま線がありましたから、平行に書けましたけど、どんどん上に上がってしまう.

表 2 では学校生活や仕事上の業務において体験した日常的なエピソードがあげられており、キーワード「中学」を含むエピソードに対しては 7 個も得られた。このように、確定診断されるまでに、患者の多くは様々な初期的な症状を体験していることが、今回の実験結果から示唆された。

表 2 で得られたエピソードは、すぐにでも病気の早期発見に活用することができる。例えば、小学校の先生に対しては「体を曲げて文字を書いている児童が居れば、視力が低下している可能性があるので注意して下さい」と依頼することにより、初期症状が出始めている児童を早い段階で眼科専門医に紹介することが可能となる。

5. 結論

本論文では、難病の早期診断を実現するために、新しい診断アプローチに係る手法を提案した。網膜色素変性の患者を対象に実験を行ったところ、初期症状を示唆するキーワードが得られることが確認された。今後は、患者から得られたキーワードを活用して、難病の早期診断につながる新しい問診項目を作成する。

謝辞

この研究は「未分類疾患の情報集約に関する研究」厚生労働科学研究費補助金（難治性疾患克服事業課題番号：H21-難治-一般-167）の助成を得て行われた。

参考文献

- [Lozano 05] Andres M. Lozano et al., New Movement in Parkinson's, Scientific American, 293, 68–75 (2005).
 [大澤 92] 大澤他, KeyGraph: 語の共起グラフの分割・統合によるキーワード抽出, 電子情報通信学会論文誌 D-1, J82-D-1(2), 391–400 (1992).
 [疾病対策研究会] 疾病対策研究会編集, 難病の診断と治療方針, 東京六法出版(2005).