

ポスター | 医療支援

ポスター3

医療支援

2019年11月23日(土) 09:00 ~ 10:00 ポスター会場2 (国際展示場 展示ホール8)

[3-P2-1-02] 対話型病状判定支援システムによる受診意欲とシステム使用感に関するユーザ評価

○山内 翔大¹、岡本 和也²、平木 秀輔²、杉山 治²、山本 豪志朗²、佐々木 博史²、南部 雅幸²、黒田 知宏² (1. 京都大学, 2. 京都大学医学部附属病院)

キーワード : Interactive decision support system for potential patients, Qualitative evaluation, User experience

自身の体調不良を自覚しながらも受診を怠り、結果として症状が重篤になってしまうケースが存在する。自身の症状への危険意識が希薄な人々を受診させるきっかけを与える手段の1つとして、対話型受診判定支援システムを使用してもらうことが考えられる。しかし、既存のシステムは、病状を特定するために患者へ問う質問数が多くなる傾向にある。我々は病状を特定することに固執せず、ユーザが緊急で受診すべきかどうかを出力することに焦点を当てることにより、質問数が少なく済む対話型病状判定支援システムを提案した。また、提案システムを従来型のシステムと比較して、質問数が削減されていることを観測した。しかし、質問数を削減することで、ユーザに実際に提案システムを使用してもらえるか、また、提案システムを使用したユーザに影響を与え、受診を促すことができるかどうかは不確かである。したがって、提案システムがユーザへ与える影響を評価する必要がある。本研究では、提案システムがユーザに与える影響を評価した。評価方法として、既存システムを模倣したシステムを実装し、提案システムと比較を行った。そのために、既存システムを分析し、1.質問数、2.質問の複雑さ、3.出力要素、の3種類の軸に従って分類し、各分類に対応するシステムを比較対象として構築した。比較方法として、アンケート評価を採用した。作成した各システムに対して同様の使用シナリオを用意し、被験者に用意したシナリオに沿って各システムを使用してもらった。そして、1.信頼性、2.使用の煩雑性、3.受診意欲の促進性、の3種類の定性的指標について比較することで、提案システムの評価を行った。

対話型病状判定支援システムによる 受診意欲とシステム使用感に関するユーザ評価

山内翔大*1、岡本和也*1,2、平木秀輔*1,2、杉山治*2、
山本豪志朗*2、佐々木博史*2、南部雅幸*2、黒田知宏*1,2

*1 京都大学大学院情報学研究科、*2 京都大学医学部附属病院

User evaluation of interactive decision support systems for potential patients in terms of willingness of seeing a doctor and usability

Shota Yamauchi*1, Kazuya Okamoto*1,2, Shusuke Hiragi*1,2, Osamu Sugiyama*2,
Goshiro Yamamoto*2, Hiroshi Sasaki*2, Masayuki Nambu*2, Tomohiro Kuroda*1,2

*1 Graduate School of Informatics, Kyoto University, *2 Kyoto University Hospital

In many situations, people will neglect symptoms to avoid going to the hospital. To prevent such behavior in the case of potentially serious incidents, there are systems which inquiry users about their health conditions and make recommendations to go to the hospital when necessary. Since these systems, however, have not yet been evaluated for their suitability to end users, the lack or inadequacy of system features might make users decide not to use them. Even in the case that the system is used, it might still be ineffective, due to insufficient or incomplete recommendations. In this work, we analyzed existing systems and identified 4 features whose presence might change the user experience, each of them with 2 possible scenarios, resulting in 16 different possible combinations. After excluding combinations that made no sense, we finally constructed 12 different hypothetical systems and compared them, with a user survey. As a result of our evaluation, we concluded that users prefer systems with a small number of simple questions, showing recommendations of guidelines to follow and with detailed information about possible diseases.

Keywords: Interactive decision support system for potential patients, Qualitative evaluation, User experience

1. 序論

自身の体調不良を自覚する人々が、症状を軽視してしまい病院に行くことを怠り、結果として症状が重篤になってしまうケースが存在する。このケースを防ぐために、受診を促す仕組みとして、症状や生体情報などから病状を判定する対話型病状判定支援システムが存在する。対話型病状判定支援システムは、ユーザに質問を行うことでユーザの情報を引き出し、収集された情報からユーザの病状を特定する。例えば、Singla はあらかじめ決められたルールに則り疾病を抽出するシステムの構築を目指した[1]。また、株式会社モドレーが提供する「症状チェッカー」¹ や Babylon Health 社が提供する「Ask Babylon[2]」など、一般に公開されているシステムが存在する。

しかし、各対話型病状判定支援システムには各々の特徴が存在し、どのようなシステムが最適なのかは十分に評価されていない。

本研究の目的は、対話型病状判定支援システムの特徴となる構成要素を抽出し、どのような構成要素をもつシステムが最適なのかを検証することである。

本研究では、既存の対話型病状判定支援システムから抽出した構成要素を組み合わせることで複数のシステムを作成し、「システム使用の煩雑性」、「判定結果の煩雑性」、「システム使用後の意思決定に与える影響」、「判定結果の信頼性」の4種類の項目についてのアンケート調査を実施する。

2. 対話型病状判定支援システム

本研究における対話型病状判定支援システムの定義は、以下のようなフローを経てユーザの病状を判定するシステムである。

1. ユーザが入力した自覚症状を受信する。

¹<https://medley.life>

² https://www.hokendohjin.co.jp/web_service/iphone_appli03/

2. 受信した入力症状から病状を推定する。
3. 2で病状を推定できない場合には4へ、推定できた場合には7へ進む。
4. ユーザに症状の有無を質問する。
5. ユーザの回答を受信する。
6. 2へ戻る。
7. 推定結果を出力する。

2.1 要素抽出

本研究で参考にした対話型病状判定支援システムは以下のものである。

1. Babylon Health 社が提供する「Ask Babylon[2]」
2. 株式会社モドレーが提供する「症状チェッカー」¹
3. 保健同人社が提供する「症状チェック」²
4. [3]にて設計された緊急度と情報利得を用いた質問設定を行う対話型病状判定支援システム

上記システムを実際に使用し、表1に示す構成要素において、システム間に差異があることが観測された。

表1:システム間の差異

病状判定に必要なとする質問数		
(ア)	①	疾病が特定されるまで質問を繰り返すのか
	②	医者にすぐ見てもらわなければならない(緊急性)が判定されるまで質問を繰り返すのか
(イ)	ユーザの回答により候補疾病から除外すべき疾病を一度だけ候補疾病に残すかどうか	
一度の質問に用いる症状が複数あるか		
詳細な出力をするかどうか		

本研究では上記の構成要素をそれぞれ

- 要素 1-(ア):「メイン出力」

- 要素 1-(イ):「疾病ストック」
 - 要素 2:「複数症状」
 - 要素 3:「詳細出力」
- と定義する。

2.2 システム構築

2.2.1 構成要素の組み合わせ

表 2 は 4 種類の構成要素の組み合わせの結果考えられる合計 16 種類のシステムの構成である。ただし、ストックが○、かつ、複数症状が○であるシステム、すなわちシステム 1、2、9、10 については同じ質問を 2 回繰り返す無意味なシステムとなってしまうため本研究の調査対象外とした。

表 2 システムの構成

	メイン出力	疾病ストック	複数症状	詳細出力
システム 1	疾病	○	○	○
システム 2	疾病	○	○	×
システム 3	疾病	○	×	○
システム 4	疾病	○	×	×
システム 5	疾病	×	○	○
システム 6	疾病	×	○	×
システム 7	疾病	×	×	○
システム 8	疾病	×	×	×
システム 9	緊急性	○	○	○
システム 10	緊急性	○	○	×
システム 11	緊急性	○	×	○
システム 12	緊急性	○	×	×
システム 13	緊急性	×	○	○
システム 14	緊急性	×	○	×
システム 15	緊急性	×	×	○
システム 16	緊急性	×	×	×

2.2.2 質問選定アルゴリズム

本研究で提案する質問選定アルゴリズムについて述べる。「メイン出力」が疾病、「複数症状」が×であるシステム、すなわちシステム 3、4、7、8 については以下のように動作する。

1. ユーザの症状入力を受け付ける。
2. ユーザの入力を元に疾病テーブルから候補疾病を絞り込む。ただし、システム 3、4 においては、この時、絞り込みにより候補疾病から除外される疾病について、要素 1-(イ)のルールが未適用ならば、候補疾病群から削除しない。
3. 「候補疾病が 1 つだけである」、「これ以上質問によって候補疾病群を分割できない」という 2 つの条件のうち、少なくとも 1 つの条件を満たしているかどうかを判定する。
4. 3 の 2 つの条件をどちらも満たしていないなら候補疾病群をできるだけ均等に 2 分割するような質問をユーザに行う。
5. 質問に対するユーザの回答を受信する。
6. 3 の 2 つの条件のうち、少なくとも 1 つの条件を満たすまで 2~5 を繰り返した後、ユーザの疾病を出力する。

「メイン出力」が疾病、ストックが×、「複数症状」が○であるシステム、すなわちシステム 5、6 については以下のように動作する。

1. ユーザの症状入力を受け付ける。
2. ユーザの入力を元に疾病テーブルから候補疾病を絞り込む。
3. 候補疾病リストを、緊急性があるものが前に来るようにソートする。
4. 候補疾病リストの先頭要素を取り出し、取り出した疾病に対応する症状全ての有無をユーザに質問する。
5. 質問に対するユーザの回答を受信する。
6. ユーザの回答が「はい」なら疾病が確定、出力する。「いいえ」なら 4 に戻り、次の要素について同様の処理を繰り返す。

「メイン出力」が緊急性、「複数症状」が×であるシステム、すなわちシステム 11、12、15、16 については[3]における緊急度と情報利得を用いたアルゴリズムを使用した。このアルゴリズムは、各疾病に緊急性があるかどうかのラベル付けを行い、質問に対するユーザの回答から、候補疾病群の緊急性ラベルの割合を用いた情報利得計算に基づく質問選定を行うというものである。そして、候補疾病群に緊急性がある疾病しかない、もしくは緊急性がある疾病が存在しない状態になった場合に救急車を呼ぶべきかどうかを出力する。また、システム 13 やシステム 14 では質問中の症状の数を 2 とし、2 つの症状を両方持つかどうかの質問に対するユーザの回答から、候補疾病群の緊急性ラベルの割合を用いた情報利得計算に基づく質問選定を行う。

2.2.3 詳細出力

詳細出力が×になっているシステムでは、それぞれのメイン出力だけが出力される。具体的には、システム 4、6、8 では疾病名のみが、システム 12、14、16 では救急車を呼ぶべきかどうかのみが出力される。

詳細出力が○になっているシステムではメイン出力だけでなく、システムが特定した疾病や、可能性のある候補疾病名の情報が出力される。具体的には、システム 3、5、7 では疾病名に加えてその疾病の緊急度や詳しい解説が出力され、システム 11、13、15 では候補疾病群中のいくつかの疾病の緊急度や詳しい解説が出力される。

3. 評価

3.1.1 シナリオ

アンケートにおけるシナリオについて述べる。

本アンケートは、ユーザが心筋梗塞であるという設定で実施された。ユーザに潜在的な心筋梗塞患者になりきってもらうために、アンケートへ進む前に表 3 に示す内容を含む文章を読んでもらった。

表 3: アンケートシナリオ

ユーザの心境	
(ア)	被験者は意識が遠のく症状がある。
(イ)	被験者は、救急車を呼ぶほどではないと感じているが、少し不安である。
(ウ)	被験者は、病状を判定してくれるロボットを使おうとしている。
アンケートの目的 ロボットの質問や判定結果の違いが以下の 4 項目についてどのような影響を与えるのかを調査する。	
①	システム使用の煩雑性
②	判定結果の煩雑性
③	システム使用後の意思決定に与える影響
④	判定結果の信頼性

この文章にはユーザが潜在的な心筋梗塞患者として振舞うために必要な症状や精神状態、システムの使用状況の説明、また、本アンケートの目的が記述されている。

3.1.2 対話フロー

アンケート回答前にユーザに提示される対話フローの例を表 4 に示す。

表 4: 対話フローの例

あなたは、ロボットに対し、「めまい」の症状を訴えました。
↓
ロボットは以下の症状が全て当てはまるかを聞いてきました。 横たわると息苦しくなる、胸痛、冷汗、吐き気、動悸、上腹部痛、胸焼け、顔面蒼白、意識消失、背部痛、呼吸困難、胸の中央の痛み
↓
あなたは、「いいえ」と答えました。
次に、ロボットは、以下の症状が当てはまるかを聞いてきました。 意識消失
↓
あなたは、「いいえ」と答えました。
↓
次に、ロボットは以下の症状が全て当てはまるかを聞いてきました。 血の気が引いて気持ちが悪い、意識障害、無痛、胸痛、冷汗、虚脱感、ショック症状、脱力感、失神、血圧低下、顔面蒼白、痛み、吐き気、呼吸困難、胸の中央の痛み
↓
あなたは、「はい」と答えました。
↓
ロボットは以下の出力をしました。 疾病: 心筋梗塞の可能性あり。 緊急度: 緊急度高。 詳細: 心臓の筋肉に酸素を送る血管(冠動脈)が詰まってしまう、組織が壊死すること。痛み・冷や汗・吐き気などの症状がないこともある。突然死の原因になる。命に関わる疾病の可能性がります。救急車を呼んだ方がいいです。

3.1.3 アンケート項目

DeLone らは、システムの成功指標として、system quality と Information quality とを取り上げた[4]。本研究では[4]を参考にし、system quality の指標として「システム使用の煩雑性」、「判定結果の煩雑性」、Information quality の指標として「システム使用後の意思決定に与える影響」、「判定結果の信頼性」を設定し、これら 4 種類の指標についてのアンケート項目を作成した。具体的に、各指標について表 5 に示す項目を設けた。

表 5: アンケート項目

システム使用の煩雑性	
質問 1-a	ロボットとの対話にめんどくささを感じましたか?
質問 1-b	ロボットの使いやすさについてどう感じましたか?
判定結果の煩雑性	
質問 2-a	ロボットの判定結果は複雑だと感じましたか?
質問 2-b	ロボットの判定結果はわかりやすいと感じましたか?
システム使用後の意思決定に与える影響	
質問 3-a	ロボットとの対話後、救急車を呼びたいと思いませんか?
質問 3-b	ロボットの判定結果を参考にしなくなりそうだと感じましたか?
判定結果の信頼性	
質問 4-a	ロボットの判定結果を信頼できると感じましたか?
質問 4-b	ロボットの判定結果に疑わしさを感じましたか?

また、アンケートには 5 段階のリッカート尺度を採用した。

3.2 結果

本研究では 20 代～40 代の成人男性合計 9 人に対してアンケート調査を実施した。表 6 に、各システムに対する 9 人の各アンケート項目の回答の平均値を示す。ただし、リッカート尺度の中央値は 3 であり、3 より大きいものは質問に対して正の値である。

表 6 各アンケート項目の回答の平均値

	1-a	2-a	3-a	4-a	1-b	2-b	3-b	4-b
システム 3	3.67	3.00	3.67	4.00	2.56	3.67	3.89	2.56
システム 4	3.78	1.44	2.11	2.78	2.67	3.33	2.56	3.00
システム 5	3.22	2.33	3.67	3.33	3.11	3.56	3.44	3.00
システム 6	3.44	2.00	2.00	2.78	2.78	3.33	2.56	3.33
システム 7	1.89	2.67	3.89	3.67	3.89	3.44	3.67	2.56
システム 8	1.78	1.78	2.11	2.56	3.56	3.22	2.56	3.44
システム 11	4.44	2.89	3.67	3.67	2.11	3.67	3.67	2.22
システム 12	4.56	1.56	2.67	2.78	2.22	3.22	2.78	3.33
システム 13	2.78	2.33	4.11	4.00	3.44	4.22	4.00	2.56
システム 14	2.78	1.56	2.67	2.78	3.33	3.00	2.56	3.67
システム 15	2.22	2.44	4.00	3.89	3.78	4.00	3.89	2.78
システム 16	1.78	1.56	2.33	2.33	3.44	3.56	2.00	4.00

システム使用の煩雑性に関する質問(1-a, 1-b)については、システム 7,8,15,16 という比較的質問数が少ない方が適切であるという結果となった。

判定結果の煩雑性に関する質問のうち、結果の複雑さを問う質問 2-a については、システム 4,6,8,12,14,16 というメイン出力のみを行うシステムの方が適切であるという結果となった。しかし、結果の分かりやすさを問う質問 2-b については、シス

テム 3,5,7,11,13,15 という詳細の出力を行うシステムの方が適切であるという結果となった。

システム使用後の意思決定に与える影響に関する質問(3-a, 3-b)については、システム 3,5,7,11,13,15 という詳細の出力を行うシステムの方が適切であるという結果となった。

判定結果の信頼性に関する質問(4-a,4-b)については、システム 3,5,7,11,13,15 という詳細の出力を行うシステムの方が適切であるという結果となった。

また、「複数症状」の項目が○であるシステム 5,6,13,14 は「複数症状」が×で他の要素が等しいシステム 7,8,15,16 と比較してシステム使用の煩雑性が若干増すが、他の指標については大きな影響が無いという結果になった。

4. 考察

システム使用の煩雑性について、システム 7、8、15、16 などの質問数が少ないシステムの場合にユーザはシステム使用の煩雑性を感じにくいことが示唆された。

システム 7、8 については、今回扱った、ユーザが心筋梗塞であるという文脈においては質問回数が少なかったが、一般的にシステム 15、16 よりも質問数が増える傾向にあることが先行研究にて示されている[3]。したがって、一般的にはシステム 15、16 の方がシステム使用の煩雑性を感じにくいことが予想される。

判定結果の煩雑性について、「詳細出力」が○であるシステムを使用する方がユーザは複雑さを感じるが理解しやすいという結果が得られた。すなわち、判定結果の複雑さとわかりやすさについては正の相関があることが示唆された。

システム使用後の意思決定に与える影響について、「詳細出力」が○であるシステムを使用する方が意思決定に与える影響は著しく大きくなることが示唆された。また、メイン出力が緊急性、すなわち救急車を呼ぶべきかどうかであるシステムの方が意思決定に与える影響が大きいたことが示唆された。

判定結果の信頼性について、「詳細出力」が○であるシステムを使用する方が判定結果の信頼性が大きくなることが示唆された。

また、複数の症状の有無を質問するシステムは、システム使用の煩雑性が若干増すことが示唆された。

上記の結果をまとめると、質問数が少なく、質問内容がシンプルで、結果として明確な行動指針を示し、付随して詳細な情報を提示するシステムが優れていることが示唆された。

5. 結論

本研究では、既存の対話型病状判定支援システムから特徴となる、構成要素を抽出した。各構成要素を組み合わせたシステムを 12 種類構築し、アンケート調査にて各システムを比較した。

結果として、質問数が少なく、結果として明確な行動指針を示し、付随して詳細な情報を提示するシステムが優れていることが示唆された。

参考文献

- 1) Jimmy Singla. The Diagnosis of Some Lung Diseases in a Prolog Expert System. International Journal of Computer Applications 2013;vol.78, no.15,;pp.37-40.
- 2) Margaret McCartney. General practice can't just exclude sick people. BMJ 2017; vol.359:id.j5190
- 3) 山内翔大, 岡本和也, 平木秀輔, 杉山治, 山本豪志朗, 佐々木博史, 南部雅幸, 黒田知宏.対話型病状判定支援システムにおける質問最適化の試み.システム制御情報学会講演論文集;2019:TS12-2-4

- 4) William H. DeLone, Ephraim R. McLean. The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. J. of Management Information Systems 2003;19(4): 9-30