

希少情報のデータ分析手法の検討

研究分担者 大澤 幸生（東京大学大学院工学系研究科教授）
西原 陽子（東京大学大学院工学系研究科講師）
熊川 寿郎（国立保健医療科学院経営科学部長）

研究要旨

我が国では難病に苦しむ人が多数存在する。一般に、難病は診断・治療が困難で、慢性の経過をたどることが多く、そもそも現段階では名前すらない未分類疾患も存在することが予想される。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するが、未分類疾患の発見方法に関してはほとんど研究が進められていない。そこで本研究では、未分類疾患を分類可能な新しい分類基準を考えるための支援システムを提案する。提案するシステムは疾患名を入力にとり、20個程度をランダムに選んでユーザに示す。ユーザは同じ基準で分類可能な疾患名をまとめてグループを作り、グルーピング後、どのグループに含めるかを迷った疾患名とその理由を記述する。提案システムはユーザにより記述された理由を、未分類疾患を分類可能な新しい分類基準の候補として出力する。評価実験を行った所、眼科の医療者から新しい分類基準を得ることができ、提案システムの有効性を確認した。

A. 研究目的

我が国では難病に苦しむ人が多数存在する。一般に、難病は診断・治療が困難で、慢性の経過をたどることが多く、そもそも現段階では名前すらない未分類疾患も存在することが予想される。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するが、未分類疾患の発見方法に関してはほとんど研究が進められていない。そこで本研究では、未分類疾患を分類可能な新しい分類基準を考えるための支援システムを提案する。

本研究では、未分類疾患を分類する基準は、①医療者が普段考えないような基準や、②患者が認識していない症状の中に存在すると仮定する。②に関しては、難病の患者に疾患に関するエピソードを語ってもらい、疾患の初期症状と思われるキーワードを抽出するシステムの構築を進めており、網膜色素変性の早期診断を実現するようなエピソードが得られ

始めている。しかし、研究の途中段階であるため、本報告では①に限定し、提案したシステムと、その評価実験の結果を報告する。

B. 研究方法

本研究では、単語が書かれたカードを分類させることにより、人の発想能力を評価する「アナロジーゲーム」[中村 08]を改良し、病名分類システムを構築した。また、普段考えないような分類基準を得るために、視線装置を用いて分類時に与える指示を作成するシステムも構築した。

B-1 病名分類システム

病名分類システムの構成を説明する。

B-1-1 アナロジーゲームの説明

アナロジーゲームは単語間の類推を考え、似ているものを分類させるゲームである。アナロジーゲームでは他人が考えないような独

自の分類をするほど発想能力が高いとして評価され、高い点数が与えられるようになっていく。

アナロジーゲームでは、単語が書かれた20枚程度のカードが用意され、ゲームのプレイヤーに与えられる。プレイヤーはマウス操作により、単語カードの移動や同じグループにまとめたことを意味する色づけを行うことができる。分類できるグループの数は5つ以内となっており、グループを表すキーワードを画面下の四角い枠の中に記入することができる。

アナロジーゲームでは単語カードを分類する際に2つのルールを設けている。1つは単語カードを5つ以内のグループに分類すること。もう1つは単語カードを1つも余らせることなく、必ずどこかのグループに分類することである。本研究では、この2つのルールに新しいルールを加えることにより、提案システムを構築した。

B-1-2 病名分類システムの構成

病名分類システムの構成を図1に示す。提案システムは100個程度の疾患名を入力にとり、入力された中から20個程度の疾患名をランダムに選択し、ユーザに示す(図2)。ユーザは同じ基準で分類可能な疾患名をまとめてグループを作り(図3)、グルーピング後、どのグループに含めるかを迷った疾患名とその理由を記述する(図4)。提案システムはユーザにより記述された理由を、難病を分類可能な新しい分類基準の候補として出力する。

病名分類システムが対象とするユーザは医療者とする。ユーザである医療者は、自身が専門とする診療科の疾患名を分類する。分類作業は4回から6回行い、2回ごとに疾患名を入れ替えることとする。

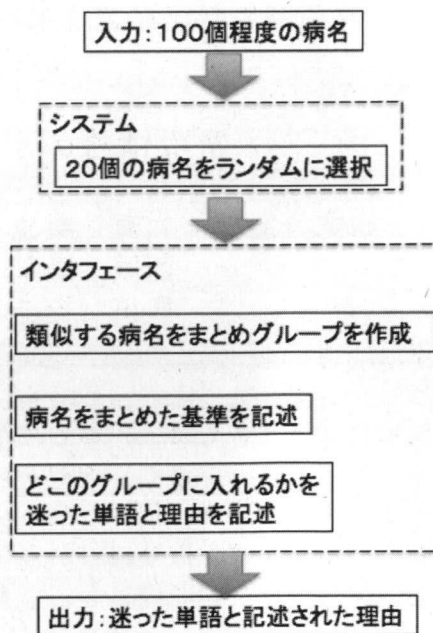


図1 病名分類システムの構成

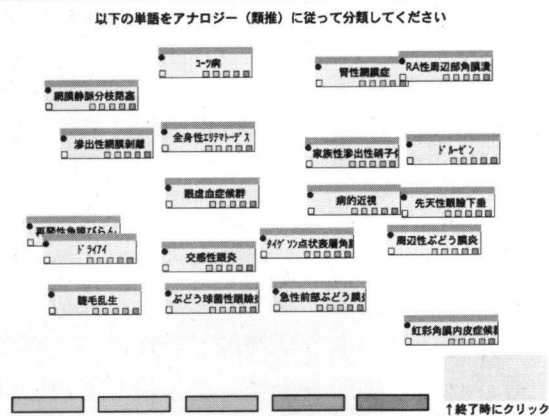


図2 インタフェース初期画面。提案システムがランダムに選んだ20個の病名が表示されている。

めな休息を取入れ、被験者の負担が大きくなりえない実験設定を行った。

B-2 視線装置を用いた分類時指示作成システム

B-2-1 分類時指示作成システムの構成

人間の視線には思考の一部が現れている。アナロジーゲームにおけるプレイヤーの視線を計測したところ、図6のように2つの単語カードをまとめて1つのグループにする直前、2つの単語カードの間を視線が何度も往復する現象が見られた。即ち、視線の往復は、プレイヤーが単語カード間に関連を見いだしたサインだと考えることができる。そこで、本研究では、病名分類システムを使用するユーザの視線を計測し、関連を見いだしている所／見いだしていない所を抽出し、再度分類を行わせるときの指示を作成するシステムを提案する。

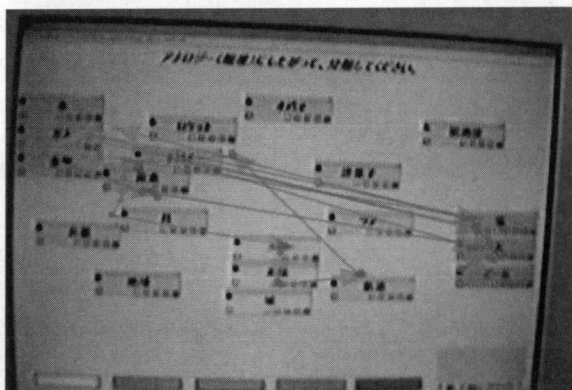


図6 アナロジーゲームを使用するプレイヤーが示した視線の例

指示作成システムでは、病名分類システムを使用するユーザの視線を、視線装置により計測し、各単語間を視線が往復していた頻度を算出する。ユーザが分類を終えた後、再度分類をする際に、次のどちらかの指示を与える。

- ・ 指示A: 視線の往復頻度が高かったが、異なるグループに含まれた2つの単語を選

び、それらの単語が同じグループに含まれるようにする。

- ・ 指示B: 視線の往復頻度が低く、異なるグループに含まれた2つの単語を選び、それらの単語が同じグループに含まれるようにする。

指示Aと指示Bは、新しい分類基準を考えさせる刺激という点で共通しているが、関連を見いだしていなかったという点において、指示Bの方がユーザにとって難しくなっている。

B-2-2 分類時指示作成システムの評価実験

提案したシステムにより、普段考えないような疾患の分類基準が得られるかどうかを評価する実験を行った。

実験手順

以下の①から②の手順を指示を変えて、2回行ってもらった。

- ① 被験者は、病名分類システムを用いて疾患名を分類する
- ② 実験者が、分類時指示作成システムにより得られた指示を与え、被験者は再度病名を分類する

被験者

実験の被験者として、眼科の医療者2名に依頼した。

評価方法

得られた結果に対して、ケーススタディを行った。

倫理面の配慮

本研究は医療者を被験者として実験を行うが、専門的知識の提供が主であり、プライバシーに関わる情報は扱わないため、特に考慮はしなかった。しかし、できるだけ被験者の負担にならないよう、実験の事前説明やこま

めな休息を取入れ、被験者の負担が大きくなり
ない実験設定を行った。

C. 研究結果

C-1 病名分類システムの実験結果

各被験者が記述した分類基準を表1から表
6に示す。

表1 眼科医療者1が記述した基準

1	眼球関連、視神経疾患、角膜関連、斜視、網膜関連
2	眼球関連、角膜関連、神経関連、網膜関連、眼瞼関連
3	眼底疾患、眼圧上昇疾患、視神経疾患、角膜疾患、眼位・眼球運動
4	治療必要なし、再発する疾患、治療困難な疾患、全身性の疾患、外眼部手術の適用
5	網膜関連、虚血性疾患、ぶどう膜炎関連、角膜関連、外部性疾患、眼瞼関連
6	急性疾患、慢性疾患、再発寛解を繰り返す疾患、状態があまり変化しない疾患

表2 眼科医療者2が記述した基準

1	外眼部、視神経、網膜・硝子体、中間透光体、緑内障
2	紹介、ステロイド、緊急入院、経過観察、手術
3	予防できる、遺伝（先天）性、治る、頭部疾患、予後不明
4	文献、老人、予約無し、コンプライアンス

表3 血液内科医療者1が記述した基準

1	白血球、血小板、凝固、赤血球、その他
2	血液悪性腫瘍、病気以外、血液でない病気、その他、良性の血液疾患
3	病巣がリンパ、病巣がその他、その他、病巣が骨髄
4	非常に予後が悪い、抗がん剤以外の治療、日本少なく海外多い、抗がん剤が効く、治療のためにもう少し診断が必要

表4 血液内科医療者2が記述した基準

1	治療方法、二次性・病態、腫瘍性疾患、先天性疾患、自己免疫
2	血清・凝固、血小板、白血球・リンパ、赤血球、幹細胞
3	疾患名（特定）、疾患名（原因）、疾患名（原因不明）、治療名、現象名
4	抗がん剤治療、保存・骨髄入れ替え、ステロイド・鉄剤、治療その他
5	幼児期、青年期、中年期、老年期、いつでも
6	半分位治る、治らない、昔治らず、補充・何とか

表5 血液内科医療者3が記述した基準

1	リンパ腫、血液疾患、感染症、白血病、先天性血液疾患
2	貧血、凝固異常、リンパ性、骨髄性、検査値異常
3	慢性疾患、急性疾患、治療
4	薬物治療、抗がん剤治療、移植
5	日常的、非日常的、血液内科以外
6	腎不全、HIV、白人、アフリカ、ヘアリー細胞白血病

表6 精神科医療者が記述した基準

1	心理的要因、脳機能の障害、本能の障害、依存症、環境要因
2	児童期、青年期、成人期、老年期、不特定
3	薬物治療主体、心理療法主体、環境調整主体、薬物治療と環境調整、薬物治療・心理療法・環境調整・家族療法
4	性行動の亢進、抑うつ状態を起こしやすい、生活指導、身体的管理必要、心理的な反応、性格的变化を起こす

続いて、各被験者が分類に迷った単語とその理由を表7に示し、表8に記述された理由の数と、普段考えないと回答された数を示す。実験において、眼科医から普段考えない分類基準を2つ得ることができた。

表7 医療者が分類に迷った病名と理由

	病名と迷った理由
眼科医 1 4回目	悪性リンパ腫。結膜、眼底、眼球のいずれでも発症するため
眼科医 1 6回目	円錐角膜。稀にデスメ膜破裂を起こすことがあり、その場合は急性疾患になるが、大抵は状態が余り変化しない疾患に含まれるため
眼科医 2 3回目	ドルーゼン。失明する人としていない人がおり、黄斑部にできると失明するため
血液医 2 6回目	血小板減少症。治らないに含めるか、半分位治るに含めるかで迷った
精神科医 4回目	周期性傾眠症、性行動異常。詳しくないため
精神科医 6回目	周期性傾眠症、拘禁反応。詳しくないため

表8 記述された理由の数と、普段考えないと回答された数

	理由の数	普段考えない基準の数
眼科医	3	2
血液内科医	1	0
精神科医	2	0

C-2 分類時指示作成システムの結果

被験者が行った分類結果の例を表9に示す。この結果は指示Bを与えたときに1名の被験者から得られた。

表9 分類時指示作成システムの実験で得られた結果の例

前半	積極的な治療が必要	麦粒腫、角膜ヘルペス、急性網膜壊死
	積極的な治療が必要ではない	眼瞼乳頭腫、ドルーゼン、錐体桿体ジストロフィ
後半	担当すると厄介な病気	角膜ヘルペス、錐体桿体ジストロフィ、急性網膜壊死
	単純な病気	麦粒腫、眼瞼乳頭腫、ドルーゼン

D. 考察

得られた実験結果に対して考察を示す。

D-1 病名分類システムに関する考察

D-1-1 全被験者の分類基準の傾向

表1、表2において、眼科の医療者が記述した基準は、疾患が生じる眼の部位、治療方針、疾患の進行などがあり、両方の医療者で重複していた。この事実に対して眼科医の被験者に質問した所、眼科の医療者が患者を診たときに考えることは、病変が存在する部位がどこにあるのか、どういう原因で起こっているのか、どうすれば治療できるかを順に考える。このように論理的に考えることで、自分の頭の中を整理し、それが結果として患者さんへの納得のいく説明へとつながり、良好なコミュニケーションを得ることができるとの回答が得られた。このため、両方の医療者が重複する基準を記述したと考えられる。

表3、表4、表5において、血液内科の医療者が記述した基準は、病気の細胞、治療方法などがあり、両方の医療者で重複していた。これも先ほどの眼科の医療者と同じ理由で、血液内科の医療者が患者とコミュニケーションを取るためには、血液の中で形が変わってしまった細胞を特定し、症状を緩和する治療方法を探るためとのことであった。

表6において、精神科の医療者が記述した基準は、病気の原因、罹る時期、治療方法、病気の症状などであった。治療方法は、眼科も血液内科も記述しており、多くの科の医療者が用いる基準であると考えられる。

以上の考察より、異なる診療科であっても、使用される分類基準は重複するものが多いことが分かった。

D-1-2 追加したルールの効果

被験者が分類に迷ったゲームの回は、4回目から6回目であった(表7参照のこと)。ゲームの1回目や2回目では、普段考えてい

る分類基準を記述できるが、回を重ねるごとに使用可能な分類基準が少なくなっていく。多くの疾患に当てはまる分類基準が少なくなることから、4回目から6回目で迷う被験者が出てきたと考えられる。この結果から、既に用いた分類基準はできるだけ使用しない、ルール③は分類に迷う疾患名や理由を取り出すことに有効である、と分かった。

対して、提案システムが選んだ5つの疾患を別グループに入るよう分類させるルール④については、ルールを適用することにより、迷う被験者と迷わなかった被験者が3人ずつとなった。このことから、ルール④は、分類に迷う疾患名や理由を取り出すことに余り有効でない、と分かった。

D-1-3 普段考えない分類基準の数

表8に示す通り、6人の被験者が合計30回ものゲームを行ったが、迷った疾患名とその理由は6個しか得られなかった。このことから、医療者は疾患名を分類する際、ほとんど迷わないことが推察される。

表8に示された普段考えない分類基準の数を比較すると、眼科医からのみ普段考えない分類基準が得られていた。このことから、眼科医に提案システムを適用すると普段考えない分類基準が得られる割合が高い可能性があるということが分かった。

D-1-4 熟練者と非熟練者の記述する理由の違い

本実験では、眼科医と血液内科医が熟練した医療者、精神科医が熟練ではない医療者として区別し実験を行った。熟練者と非熟練者の記述した理由を見比べると、非熟練者である精神科医が記述した理由は、病名に詳しくないからというものであった。この理由は、熟練者である眼科医や血液内科の様に実際の診断上で迷うことがあるからという理由ではなかった。このことから、熟練した医療者に

提案システムを使用してもらうことが望ましい、と分かった。

D-2 分類時指示作成システムに関する考察

表9に示した結果に対して考察を行う。被験者が前半で行った分類は治療方法によるものであった。これは病名分類システムにおける結果とも重複することであった。

前半で分類作業をしている際に計測された視線の結果から、指示Bを与えた所、表9の後半の結果が得られた。被験者は前半ほど、素早く分類することはできなかったが、考えた末に「担当すると厄介な病気」、「単純な病気」の2つの基準をあげた。これらの基準は、被験者が疾患の治療方法に対して、自分が抱く印象に基づくものであり、これらの基準も医療者が普段考えない分類基準の一種と考えられる。また、疾患名分類後の結果には表れてこないが、分類時の強制的な指示に従うことによって、現実的にはあまり考えることがない疾患の組み合わせや、疾患発生の可能性を想像するという行為が、分類作業中の医療者の思考過程に存在していることがわかった。この部分に焦点を当てることで、今後より効率的に分類基準を得ることができるかもしれない。

以上から、視線装置を用いて分類時に指示を与えることにより、普段考えない分類基準を得ることに有効であることが確認された。

E. 結論

本研究では、未分類疾患を分類するために、医療者が普段考えないような疾患の分類基準を抽出する方法を提案した。病名分類システムにより、眼科の医療者2名から普段考えない分類基準を2つ、抽出することができた。また、分類時指示作成システムが、普段考えないような分類基準を得ることに有効であることを確認し、提案した2つシステムの有効性を確認することができた。また、実験結果

を考察することにより以下の4つの点があった

- ・ ゲームを繰り返し行わせ、可能な限り分類基準を考えさせることが、分類に迷う疾患名と理由を取り出すことに効果を発揮する
- ・ 眼科の医療者に対して提案システムは有効に働き、普段考えない分類基準を取り出すことができる
- ・ 熟練した医療者が提案システムを使用することにより、意味のある分類基準を取り出すことができる
- ・ 視線装置を用いて、関連を見いだしていない病名を1つにまとめてグルーピングさせることにより、普段考えない分類基準を得ることができる

今後は、抽出された基準に関して医療者から情報を集め、未分類疾患を分類する新しい基準となり得るかについて検証すると共に、他の診療科の医療者に対しても提案システムを用いた実験を行う。また、患者に対して行っているインタビューに関しても、分析方法の検討を進める。

F. 研究発表

1. 論文発表

なし。

2. 学会発表

西原陽子・平塚義宗・村上晶・大澤幸生・熊川寿郎、難病早期発見のための新しい診断アプローチの開発, 第24回人工知能学会全国大会, 2010年6月発表予定。

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。

参考文献

[中村 08] 中村潤, 大澤幸生, 概念創造のための類推思考プロセスにおける迷いの効果, 横幹, Vol. 2, No. 1, pp. 40--48, 2008.

分担研究報告書

未分類疾患の分類方法の検討

研究分担者 緒方 裕光（国立保健医療科学院研究情報センター長）
研究協力者 奥村 貴史（国立保健医療科学院研究情報センター情報評価室長）

研究要旨

背景：既存の診断基準等によって特定できない疾患、すなわち未分類の疾患は多数存在している可能性があり、それらの多くはその疾患概念が確立されていない。このような疾患については、症状や治療法などに関する情報が集積されておらず、原因解明、診断技術・治療方法などに関する研究を進めることがきわめて難しい。

目的：本分担研究は、未分類疾患に関する情報を体系的に収集し、未分類疾患を分類するための方法を開発することを目的とする。その一端として、今年度は症候に関する情報に基づいて未分類疾患を見出すための方法について、確率論的な観点から検討を行った。

方法：症候（結果）から疾患（原因）を推定する際に、両者の関係にベイズの定理を適用したベイズ診断の方法がすでに提示されている。この考え方を未分類疾患発見プロセスに応用して、仮想データに基づいて、情報の蓄積と未分類疾患の確率（既分類疾患に分類されない確率）との関係について考察を加えた。

結果：特定の症候パターンに対して特定の疾患名をつけることが難しい場合、その難しさの程度が大きいほどその症候パターンは未分類疾患である可能性が高いこと、一般に情報の蓄積に伴ってその可能性の大きさは減少することなどが示された。

A. 研究目的

既存の診断基準等によって特定できない疾患、すなわち未分類の疾患は多数存在している可能性があり、それらの多くはその疾患概念が確立されていない。このような疾患については、症状や治療法などに関する情報が集積されておらず、原因解明、診断技術・治療方法などに関する研究を進めることがきわめて難しい。

本分担研究は、未分類疾患に関する情報を体系的に収集し、未分類疾患を分類するための方法を開発することを目的と

する。その一端として、今年度は症候に関する情報に基づいて未分類疾患を見出すための方法について、確率論的な観点から理論的な検討を行った。

B. 研究方法

症状から疾患名を診断することは、ある結果（症候）に関する情報を得たときにそれらの結果が特定の原因（疾患）によるものであるかどうかを判定することに相当する。この原因と結果との因果関係に関してベイズの定理を応用すること

によって、この診断のプロセスを定式化することができる（ベイズ診断）。このベイズ診断を実際に行うためには、特定の疾患によって特定の症候が現れる確率を情報として必要とする。通常、疾患と症候は複数あり、それぞれの組み合わせに対応した確率の表（疾患－症候行列）が必要となる。

本分担研究では、この疾患－症候行列が得られているという前提で、特定の症状のパターンが見られた場合に、その症候パターンが未分類疾患である可能性を数量的に表現することを試みる。方法の概略は以下のとおりである。なお、以下に示すプロセスのうち2)～4)はすべて1)の議論を応用することによって計算が可能である。

1) 未知の原因がない場合（ベイズ診断）

D_1, D_2, \dots, D_n を疾患（ある症候の原因）とするとき、症候 S が見られたときその原因が D_i である確率は、ベイズの定理により

$$P(D_i | S) = \frac{P(D_i)P(S | D_i)}{\sum_{k=1}^n P(D_k)P(S | D_k)}$$

と表わされる。ただし、 D_1, D_2, \dots, D_n がすべての原因を網羅していると仮定する。ここで、 $P(D_i)$ は各疾患の事前確率、 $P(S|D_i)$ は疾患 D_i によって特定の症候 S が現れる確率である。通常、症候は複数あるので、疾患－症候行列は $P(S_j|D_i)$ と表わされ、それぞれの疾患 S_j に関してその原因が D_i である確率は $P(D_i | S_j)$ と表わされる。

実際には、ある特定の疾患 D_i に関して、その発生に関する統計的頻度が分かっている場合にはその値を確率に置き換えて $P(D_i)$ とおく。また、その疾患に関してどのような症候がどの程度の割合で現れるかが把握されている場合は、その割合

を確率に置き換えて $P(S_1|D_k), P(S_2|D_k), P(S_3|D_k), \dots$ などとおく。

2) 未知の原因がある場合

上記の1)において、 D_1, D_2, \dots, D_n がすべての原因を網羅していない場合、 D_1, D_2, \dots, D_n 以外のすべての原因を U とおく。この未知の原因 U には、「未分類疾患」、「現時点で情報が入手されていない既分類の疾患」、「疾患ではない原因」が含まれる。症候 S が現れたときその原因が U である確率 $P(S|U)$ を求める（注1）。複数の症候 S_i に関しては、 $P(S_i|U)$ と表現される。

3) 症状がないことを情報として含める場合

症候 S について、症状があるとき $a=1$ 、ないとき $a=0$ とおくと、 a が現れたときその原因が U である確率 $P(U|a)$ を求める。症候が複数ある場合には、症候 S_i に関して a_i とおき、同様に症状があるとき $a_i=1$ 、ないとき $a_i=0$ として、症候 (a_1, a_2, \dots, a_m) が現れたときその原因が U である確率 $P(U|a_1, a_2, \dots, a_m)$ を求める（注2）。

4) 仮想データによる検討

疾患－症候行列について仮想データを設定し、このデータを用いて上記の計算を行い、未分類疾患の発見プロセスについて考察を加える。本報告では、仮想データとして疾患が10（すなわち、 D_1, D_2, \dots, D_{10} ）、症候が5（すなわち、 S_1, S_2, \dots, S_5 ）の場合の疾患－症候行列を想定した。表1にその疾患－症候行列を示す。なお、表1の仮想データは任意に設定されたものであるが、このデータを利用して得られる結論に関して、少なくとも本報告の議論の中では一般性を失うことはない。

(倫理面への配慮)

今年度の本分担研究では、倫理面に関して問題となる事項は含まれない。

表1 疾患－症候行列(仮想データ)

疾患	$P(D)$	症候 $P(S_j D_i)$				
		S_1	S_2	S_3	S_4	S_5
D_1	0.100	0.01	0.49	0.50	0.01	0.00
D_2	0.081	0.10	0.50	0.50	0.02	0.01
D_3	0.005	0.30	0.60	0.10	0.20	0.10
D_4	0.001	0.10	0.20	0.70	0.30	0.10
D_5	0.027	0.20	0.50	0.30	0.15	0.05
D_6	0.005	0.10	0.40	0.50	0.01	0.01
D_7	0.001	0.20	0.70	0.10	0.65	0.10
D_8	0.018	0.50	0.48	0.02	0.30	0.65
D_9	0.001	0.10	0.45	0.45	0.22	0.44
D_{10}	0.054	0.40	0.55	0.05	0.25	0.25

C. 研究結果

以下では、表1の疾患－症候行列に基づき試算した結果を示す。

1) 未知の原因を含めない場合

D_1, D_2, \dots, D_{10} がすべての原因を網羅していると仮定して、 S_1, S_2, \dots, S_5 のうちいずれか1つの症候が現れたとき、 D_1, D_2, \dots, D_{10} のうちのいずれかを原因とする確率 $P(D_i|S_j)$ を計算した。その結果を表2に示した。

表2 症状 S_j から推定される疾患の確率

疾患	$P(D_i S_j)$				
	S_1	S_2	S_3	S_4	S_5
D_1	0.021	0.332	0.472	0.036	0.000
D_2	0.171	0.274	0.382	0.058	0.028
D_3	0.032	0.020	0.005	0.036	0.018
D_4	0.002	0.001	0.007	0.011	0.004
D_5	0.114	0.091	0.076	0.146	0.047
D_6	0.011	0.014	0.024	0.002	0.002
D_7	0.004	0.005	0.001	0.023	0.004
D_8	0.189	0.059	0.003	0.194	0.410
D_9	0.002	0.003	0.004	0.008	0.015
D_{10}	0.455	0.201	0.025	0.486	0.473

当然ながら、発生頻度の高い疾患に関して頻繁に見られる症候で、かつその症候が他の疾患ではあまり見られない場合、その症候から疾患(原因)を特定できる確率は高くなる(例えば、 S_1 が現れた時の D_{10} など)。

2) 未知の原因を含めた場合

原因の中に未知の原因(U)を含め、症候(S_1, S_2, \dots, S_5)に関してそれぞれ症状の有無を情報として考慮した場合(症候 S_j について症状があるとき $a_j=1$ 、ないとき $a_j=0$ とおく)、 a_j が見られたときその原因が U である確率 $P(U|a_j)$ を求めた。その結果を表3に示した。「症状がなかった」という情報よりも、「症状があった」という情報の方が、未知の原因であると判断される確率を小さくしている。

表3 未知の原因(U)を考慮し、かつ症状がないことを情報として含めた場合に推定される原因の確率

疾患	$P(D_i a)$ $a_j=0$ (症状なし), $a_j=1$ (症状あり)									
	S_1		S_2		S_3		S_4		S_5	
	0	1	0	1	0	1	0	1	0	1
D_1	0.10	0.01	0.06	0.27	0.06	0.37	0.10	0.02	0.10	0.00
D_2	0.08	0.10	0.05	0.22	0.05	0.30	0.08	0.04	0.08	0.03
D_3	0.00	0.02	0.00	0.02	0.01	0.00	0.00	0.02	0.01	0.02
D_4	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00
D_5	0.02	0.06	0.02	0.07	0.02	0.06	0.02	0.09	0.03	0.05
D_6	0.01	0.01	0.00	0.01	0.00	0.02	0.01	0.00	0.01	0.00
D_7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
D_8	0.01	0.11	0.01	0.05	0.02	0.00	0.01	0.12	0.01	0.40
D_9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.02
D_{10}	0.04	0.26	0.03	0.16	0.06	0.02	0.04	0.31	0.04	0.46
U	0.73	0.44	0.82	0.20	0.78	0.23	0.72	0.36	0.73	0.04

複数の症候(S_1, S_2, \dots, S_5)に関して同時に情報が得られた場合は、その情報は5つの数値(それぞれ0または1)の組み合わせ(a_1, a_2, \dots, a_5)となる。

ここで、(a_1, a_2, \dots, a_5)が現れた場合にその原因が U である確率 $P(U|a_1, a_2, \dots, a_5)$ を求めた。その結果の一部を表4に示した。このとき、(a_1, a_2, \dots, a_5)のパターンは $2^5 = 32$ 通り(P1~P32)あり、それぞれのパターンについて確率 $P(D_i|a_1, a_2, \dots,$

a_5) または $P(U|a_1, a_2, \dots, a_5)$ を求めた。これらの確率の大小に関するパターンはきわめて複雑であるが、症候（症状の有無を含めて）の組み合わせによってはひじょうに高い確率で原因を特定することができる（すなわち未知の原因である確率は低くなる）。

図4 複数の症候の情報が同時に得られた場合に推定される原因の確率

		P1	P2	P3	...	P32
症候	a_1	0	0	0	...	1
	a_2	0	0	0	...	1
	a_3	0	0	0	...	1
	a_4	0	0	1	...	1
	a_5	0	1	0	...	1
疾患	D_1	0.038	0.000	0.014	...	0.000
	D_2	0.027	0.028	0.019	...	0.006

	D_{10}	0.012	0.414	0.139	...	0.566
	U	0.908	0.137	0.720	...	0.000

3) 未分類疾患発見のプロセス

現実には、あらゆる疾患と症候に関して網羅的に情報が得られるわけではなく、通常、疾患－症候行列は不完全なものである。たとえば、原因である D_1, D_2, \dots, D_{10} のうち D_{10} に関する情報が存在していなければ、本来ならば D_{10} が原因である確率は、 U が原因である確率（すなわち未知の原因である確率）に含まれることになる。さらに、症候についても、実際にすべての検査を行うことは不可能である。

一般に、情報は蓄積されていくものであり、この情報の蓄積過程に伴って不完全さの程度は減少していく。したがって、情報の蓄積に伴い原因を特定できる可能性は高くなるはずである。しかし、情報（ここでは疾患－症候行列に関する情報）の蓄積によって原因を特定できる可能性が高くならなければ、その症候パターン

（複数の症候の組み合わせ）は、未分類疾患を原因とする可能性が大きいことになる。

ある症候パターンが見られたときにそれが未知の原因によるものである確率について、疾患に関する情報を追加していった場合（発生率すなわち $P(D_i)$ の高い順に情報を追加していった場合）の変化を図1に示した。ここで、疾患に関する情報が追加されるということは、疾患－症候行列の要素である確率が埋められていくことを意味する。症候パターンによっては、情報が蓄積されても未知の原因によるものである確率が0に近づかない場合がある（たとえば、 $(a_1, a_2, a_3, a_4, a_5) = (0, 0, 1, 1, 0)$ ）。

また、症候に関する情報が蓄積されていく場合（ここでは一例として症候 $a_1 \sim a_3$ の情報があるときに、さらに $a_4=0, a_5=1$ が追加されていく場合を示した）についても、同様に未知の原因である確率の変化を図2に示した。この場合も、図1と同様に情報の蓄積にともなって未知の原因である確率が必ずしも0に近づかない場合がある。

D. 考察

本来、既分類疾患に関する情報がすべてそろっていなければ、ある症候が未分類疾患であるかどうかを判断することはできない。しかし、一般に、世界中に存在している既分類の疾患に関する情報を完全に網羅的に集めることは技術的に限界がある。さらに、完全な情報を維持管理していくためには、時間の経過に伴う情報量の増加や内容の変更に対応しなければならない。したがって、われわれが持っている既分類疾患に関する情報はそもそも完全ではないという前提で、未分類疾患の分類や発見を行う方法を考える方が合理的である。

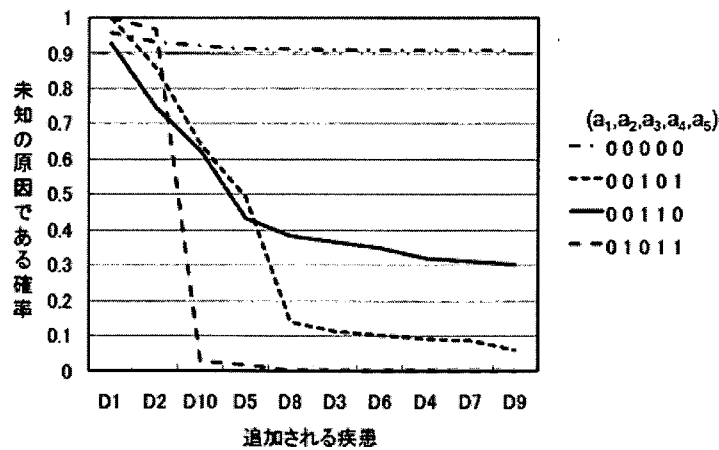


図1 疾患に関する情報が追加されていく場合(罹患率の大きい順)

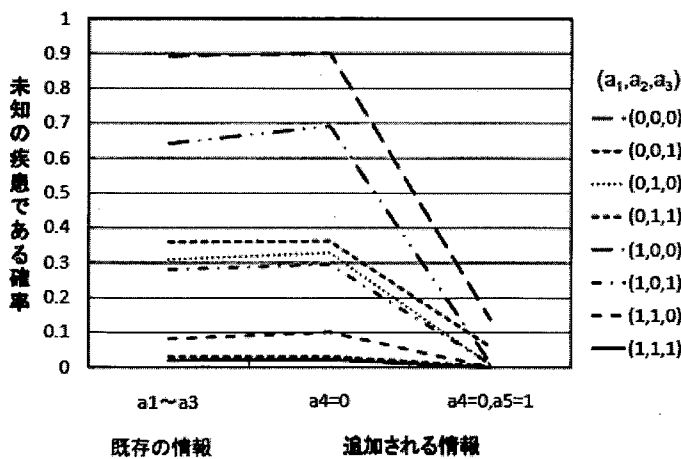


図2 症候に関する情報が追加されていく場合

医療の現場で通常行われている診断プロセスにおいて、症状からほぼ完全に疾患名が特定できる場合もあれば、ある程度の不確実性を持って疾患名が判定される場合がある。未分類疾患についても同様で、既分類疾患の情報が完全でない限りは、未分類疾患であるか否かという判

断は、未分類疾患である可能性が大きいかどうかという判断をすることと同じである。その可能性の大きさを定量的に表現するために確率を用いることは1つの合理的な方法であろう。

また、本研究では、疾患に関する情報として疾患-症候行列を中心に検討した。

しかし、疾患に関する情報には、その他にも臨床現場からの事例報告や、医学研究者による研究論文など様々な形式の情報が存在する。未分類疾患の分類や発見のためにこれらの情報をどのように合理的に活用するかは今後の課題である。

本研究では、未分類疾患を見出すための1つのアプローチとして確率論的な観点から検討した。問題点として、疾病と症候の組み合わせの数は無数にあり現実の計算をどのように行うのか、症状の有無が0と1ではなく連続量で表される場合にどのようにモデル化するのか、といった点が挙げられる。本研究で示した方法は、複雑なパターンを体系的に示すためのシミュレーションとしては十分ではないが、今後いくつかの問題点を整理することによって、多くの情報の中から未分類疾患を見出すための最初のステップとして有効な方法になりうると思われる。

E. 結論

本分担研究では、特定の症候パターンが未分類疾患によるものかどうかという可能性の大きさを確率で表現することを試みた。仮想データを用いた試算を通じて検討した結果、結論として以下のことが言える。

- 1) 特定の症候パターンに対して特定の疾患名をつけることが難しい場合、その難しさの程度が大きいほど未分類疾患である可能性が高い。
- 2) ある症候パターンが未知の原因によるものである可能性は、情報を蓄積することによって一般に減少する。
- 3) 情報の蓄積によっても未知の原因によるものである可能性が減少しなければ、その症候パターンは未分類疾患である可能性が高くなる。

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

なし。

参考文献

- 1) 松原望 (2008) 入門ベイズ統計. 東京図書.
- 2) 水野哲夫 (1973) 臨床・公衆衛生のための意思決定：行動の統計学. 南江堂.
- 3) Gelman A and Meng XL ed. (2004) Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives. Wiley.

.....

注 1)

症候 S が見られたときその原因が U である確率は

$$P(U|S) = \frac{P(U)P(S|U)}{Q + P(U)P(S|U)}$$

ただし、

$$Q = \sum_{k=1}^n P(D_k)P(S|D_k)$$

$$P(U) = 1 - \sum_k P(D_k)$$

$$P(S|U) = P(S) - P(S|\bar{U})$$

である。なお、 $P(S)$ にはその症候が一般に見られる確率を代入する。ここでは、仮想データとして、 $P(S_1)=0.1$ 、 $P(S_2)=0.2$ 、 $P(S_3)=0.15$ 、 $P(S_4)=0.05$ 、 $P(S_5)=0.03$ とした。

注 2)

D_1, D_2, \dots, D_n がすべての原因を網羅していると仮定すると、 a が現れたときその原因が D_i である確率は、

$$P(D_i | a) = \frac{P(D_i)[aP(S | D_i) + (1-a)\{1 - P(S | D_i)\}]}{R}$$

ただし、

$$R = \sum_{k=1}^n P(D_k)[aP(S | D_k) + (1-a)\{1 - P(S | D_k)\}]$$

また、症候 S について症状があるとき $a=1$ 、ないとき $a=0$ とおく。

同様に、 (a_1, a_2, \dots, a_m) が現れたときその原因が D_i である確率は、

$$P(D_i | a_1, a_2, \dots, a_m) = \frac{P(D_i) \prod_{j=1}^m [a_j P(S_j | D_i) + (1-a_j)\{1 - P(S_j | D_i)\}]}{T}$$

ただし、

$T =$

$$\sum_{i=1}^n P(D_i) \prod_{j=1}^m [a_j P(S_j | D_i) + (1-a_j)\{1 - P(S_j | D_i)\}]$$

また、症候 S_i について症状があるとき $a_i=1$ 、ないとき $a_i=0$ とおく。

$P(U | a_1, a_2, \dots, a_m)$ は、注 1 と同様のプロセスで算出できる。

厚生労働科学研究費補助金（難治性疾患克服研究事業）
分担研究報告書

疾患同定支援ツールの検討

分担研究者 種田憲一郎 国立保健医療科学院政策科学部 安全科学室長
分担研究者 高橋 理 聖路加国際病院一般内科 副医長
分担研究者 大曲貴夫 静岡がんセンター感染症科 部長

研究要旨

未分類の疾患を発見するためには、まず既知の疾病であるかどうかの判断が不可欠である。したがって既存の疾病に該当するかどうかの「診断」を支援するツールやデータベースの開発が必要であり、今年度はそれに資する以下のような調査研究を実施した。①医師臨床研修病院において、退院時診断困難例入院患者の症例を利用して、診断に至るプロセスの検討を行った。②また診断困難な事例が多い分野として、不明熱などに代表される感染症が挙げられる。本邦の感染症診療には問題点の一つとして、適切な診療過程が客観的に示されていないことがあると考えられ、感染症の診療の改訂で医療者が抱える問題について要因を分析した。③既存の診療情報を活用する際に手掛かりとなる ICD の成立と内容・改訂過程に関する分析を行い、既存の診断基準等を活用しても従来の疾患概念に該当しないような未分類疾患の診断・疾患概念の確立に際して、疾患分類の発展の過程が一定の疾患像を整理する上で影響をあたえる可能性を検討した。④最後に患者の視点から、診断が適切になされるに至る課題を抽出した。具体的にはネット上の公開されている患者自身の記録から、病気の症状の発症から、病名の確定までの経緯を分析した。

これらの研究から、退院時に診断困難率：1.1%（95%CI: 1.0 - 1.3）、診療科としては、一般内科・アレルギー膠原病内科・感染症科でその頻度が高いことが示された。とくに感染症の診断においては、現場の医師が(1)発熱の原因の質と所在が分からない状態、(2)感染を疑うが問題臓器が同定できない状態、(3)診断違いが想定される状態、の三つで問題を抱えることが明らかになった。また既存の ICD は死因統計への利用とその国際比較が目的であり、その用途・機能は発展してきているが、未分類疾患の情報抽出に資するような理論的基盤に欠いている可能性が示唆された。そして難病に対する情報を医療者だけでなく、一般の人々に広く伝えることが早期の適切な診断につながることを示唆された。

- 研究協力者 -

内山伸（聖路加国際病院呼吸器内科）
富塚太郎（国立保健医療科学院政策科学部）
山本真菜（日本大学大学院文学研究科心理学専攻博士前期課程）
小舘尚文（Risk Programme NIHR King's Patient Safety & Service Quality Research Centre）
Sanjay Saint（Ann Arbor VA Medical Center, University of Michigan Medical School）

A. 研究目的

未分類の疾病を発見するためには、まず既知の疾病であるかどうかの判断が不可欠である。既知の疾病の診断基準には該当せず、また特定の疾病としての概念も確定していないような「未分類」の疾病を発見するた

めに、個々の患者の呈する自覚症状、医師などの診察所見、検査所見などから既存の疾病に該当するかどうかの「診断」を支援するツール・データベースが有用と考えられる。また臨床現場において、診断困難な事例の診断プロセスを支援し、忙しい臨床医にも有用な支援ツールとすることで、実際に活用され、その上で「未分類」と思われる疾病の登録も容易となる。さらに将来的には患者や家族自身が自分自身の疾病を調べることが可能となるようなデータベースの基盤ともなりうると思われる。

したがって今年度の研究では実際の臨床現場における診断困難な事例の抽出と、とくに感染症における課題、そして事例抽出の際にも利用され得る ICD の可能性、さらに患者の視点からみた難病等の診断過程

の課題について以下のような調査研究を実施した。

A-1. 疾患同定支援ツールの検討

①退院時診断困難例入院患者の検討

退院時に診断困難であったまたは確定できなかった患者の頻度と特徴を調査する。また、その中に、希な疾患や未分類疾患が含まれているかも検討する。

②症例検討会の診断困難例の検討

症例検討会または医学雑誌・テキストなどまとまった症例を利用して、診断困難例について診断に至るまでのプロセスを検討し、その中で診断に最も寄与したと考えられる病歴・身体所見・検査所見を検討する。また、その情報を診断支援ツールに役立てられるかどうか検討する。

A-2. 感染症診療上の問題抽出

研究者は院内にて感染症診療に関するコンサルテーションを行っている。コンサルテーションを行う中で感じることは、相談元の医師達は診療が首尾よく運ばなかった場合に、「何故うまくいかなかったか」について客観的に理解できていないということである。つまりは感染症診療上のどのような点で自身が問題を抱えているかについて客観的に把握できていないがために、結果として次回以降の診療にその反省が活かされにくい。

よって、この診療者自身が気付いていない感染症診療上の問題点を明らかにすることは有用であると考えられる。

A-3. 国際疾病分類に関するレビュー

本研究は、既存の診断基準等を活用しても従来の疾患概念に該当しないような未分類疾患の診断・疾患概念の確立に際して、疾患分類の発展の過程が一定の疾患像を整理する上で影響をあたえる可能性を検討するものである。医師が症状や所見のある患者を適時・適正に疾患として分類させることができないことで治療がなされない、もしくは一群と考えられる疾患が研究の対象となりえない状況が起こる背景の一つとして、疾病分類が医師や研究者の思考枠組みを規定しうる要件になりうるとして着目した。

代表的な国際的疾患分類として、国際保健機関 (WHO) が開発している国際疾病分類 International Classification of Diseases (ICD) があり、過去 9 回の改訂を経て、現在 ICD-10 が最新版として公開され、世界で広く使用されている。わが国において ICD は統計法に基づく政令により告示されている疾患分類であり、統計分類として総務省の告示という形で規定されている。ICD は死因統計や診療録管理等に利用されているが、罹患率が低く従来の疾患概念に該当しないような疾患への体系的情報収集への利用に関しては発展段階にあり、現在進行中の 10 回目の改訂に臨床家の意見を反映し含まれると期待されている。また臨床医においては、2003 年 4 月より特定機能病院などに導入された診断群分類

(Diagnosis Procedure Combination; DPC) を利用した入院医療診療の包括評価により、対象病院の医師は患者入院時に基本 DPC と呼ばれる主要診断群と ICD に基づく傷病名を決定する作業にかかわり始めたことから、ICD コーディングに対する関心も高まっている。

そこで本研究では、ICD の成立と内容・改訂過程に関する分析を目的とした。

A-4. 患者の視点の病名確定上の問題抽出

難病患者にとって、病名をできるだけ早く確定することは重要である。しかし、現状では、様々な理由で病名が何年も確定しないという事例も存在する。したがって本研究では、患者の視点における症状の発症から、病名の確定までの経緯を分析し、病名の確定を遅らせる原因を探ることを目的とする。

B. 研究方法

B-1. 疾患同定支援ツールの検討

①退院時診断困難例入院患者の検討

研究デザインは横断研究。対象は、聖路加国際病院を平成 20 年 11 月から平成 21 年 10 月までに退院した全患者とした。そのうち、緩和医療科を退院した患者を対象から除いた。データは、当院電子カルテから抽出した。抽出した変数は、年齢、性別、入院期間、入院目的、退院科、退院時診断、退院時の転帰、併存症、血液検査・画像検

査・病理検査などの各種結果であった。

退院時診断困例の定義は、退院時のICD10でR69（原因不明および詳細不明の疾患）またはICD9で799.8と、ICD10でR99（診断名不明確および原因不明の死亡）またはICD9で799.9の一部のコードがつけられている症例、また、主病名に疑いが含まれている症例とした。

1年間での全退院患者のうち、退院時に診断困難例であった割合と95%信頼区間を求めた。退院時診断が確定している患者と比較検討し、退院時診断困難例患者の特徴を分析した。連続変数を比較する場合は、t検定、名義変数・順序変数を比較する場合は χ^2 乗検定を用いた。全ての統計分析は、SPSS Ver18 (SPSS, Japan, Tokyo)を用いた。

②症例検討会の診断困難例の検討

当院で定期的に行われているグランドカンファレンスから特に診断に苦慮した例、教育的症例をまとめた参考書を基に、以下の項目を抽出し比較検討する予定である。項目として、主訴、診断名、症状から診断までの期間、入院期間、診断に有効であった病歴、身体所見、検査所見、診断が困難だった理由と診断症例の臨床的意義を2人の研究者が調査票に別々に抽出し、内容が異なった場合は相談し決定することとする。

B-2. 感染症診療上の問題抽出

・研究デザイン：感染症コンサルテーションが行われた症例における担当医・コンサルタントのカルテ上の記載を帰納的にカテゴリー化してその要素を抽出した。
・対象：2009年4月1日から同年4月27日までの間に静岡がんセンター感染症科でコンサルテーションを行った患者34人。
・データ収集・分析方法：当該患者のカルテ上の診療記録をテキストファイルとして抽出し、匿名化した上で質的に分析した。データ分析の方法は、文章のなかから担当医及びコンサルテーションを受けた医師の、感染症診療内容についての記載を抜き出し、データを断片化した。断片化したデータはそのまとまり毎に区分けし、各区分け毎にカテゴリー名をつけた。その後各カテゴリーを更に統合し、それぞれに要素名を決定した。

・患者の同意：本研究は、既存資料等の

みを用いる観察研究であるため、文部科学省・厚生労働省「疫学研究に関する倫理的指針」（平成20年12月1日一部改正）の規程に則り、研究対象者からインフォームド・コンセントを受けていない。

・患者個人情報への配慮

本研究は電子カルテ上に既に記載されている内容のみを用いる観察研究であるが、患者の氏名・属性については匿名化した上で解析を行った。

B-3. 国際疾病分類に関するレビュー

文献およびウェブサイトに関するレビューを用いた。対象は、WHO、厚生労働省の文書とウェブサイト、国内のICD・DPCに関する書籍とした。特にWHOのICD情報

(<http://www.who.int/classifications/icd/en/>)、厚生労働省大臣官房統計情報部人口動態・保健統計課 疾病傷害死因分類調査室

(<http://www.mhlw.go.jp/toukei/sippe/index.html>)、厚生労働省社会保障審議会統計分科会疾病、傷害及び死因分類専門委員会の議事録・資料

(<http://www.mhlw.go.jp/shingi/hosho.html#toukei>)はハンドサーチにて情報を収集した。

B-4. 患者の視点の病名確定上の問題抽出

・研究デザイン：難病患者のテキスト形式の記録のカテゴリー分けを行い、質的研究を実施した。

・対象：インターネット上の難病患者のテキスト記録

データ収集・分析方法：インターネット上の難病患者のテキスト形式の記録を収集し、質的に分析した。検索の際に使用した単語は、「難病」「日記」であった。データ分析の方法は、文章のなかから、症状の発症から、病名確定までの記事を抜き出し、断片化した。断片化したデータはそのまとまり毎に区分けし、各区分け毎にカテゴリー名をつけた。その後各カテゴリーを更に統合し、それぞれに要素名を決定した。

C. 研究結果

C-1. 疾患同定支援ツールの検討

①退院時診断困難例入院患者の検討

1年間の全退院患者は15,923人であった。そのうち、緩和医療科で退院した患者数249人を除き、15,674人を対象とした。全患者の平均年齢は51歳(標準偏差:26歳)、男性47.1%であった。在院日数中央値は6日、退院時の死亡数は403人(2.6%)であった。(表1)

退院時に、ICD10でR69(原因不明および詳細不明の疾患)またはICD9で799.8と、ICD10でR99(診断名不明確および原因不明の死亡)またはICD9で799.9の一部のコードがつけられた患者は0であった。主病名に疑いがついていた症例は、177人(1.1%:95%信頼区間:1.0-1.3%)(図1)であった。診断が確定していた症例と比較して、年齢(p=0.7)、性別(p=0.6)は統計学的有意差を認めなかった。転帰では診断困難例のほうが改善した症例の割合が減少し、変化なしの割合が増加する傾向が認められた(p=0.06)。在院日数の中央値は診断困難例の方が統計学的に有意に短かった(p<0.01)(表1)

診断困難例を退院科別に割合を比較した(表2)。割合の高い科は、一般内科(15.6%)、アレルギー・膠原病科(8.3%)、感染症科(4.3%)であり、その後、神経内科(3.8%)、呼吸器内科(3.1%)の順番であった。その中から、アレルギー・膠原病科で退院時の主病名に疑いがついていた症例について、表3にまとめた。今後は、全症例について検討する予定である。

① 症例検討会の診断困難例の検討

現在、30症例を選択した。必要な項目を抽出中である。

C-2. 感染症診療上の問題抽出

1) 対象者の属性

男性21名、女性14名の合計34名。平均年齢は63.7歳であった。

2) データの分類

要素名	カテゴリー名
感染かどうかの判断	発熱の原因が分からない(不明熱) 〇〇感染だと思って治療しているが、どうも違うようだ
感染臓器の同定	感染している臓器はどこか? 感染だと思うのだが、どの臓器の感染かわからない

	血液培養が陽性になったが、どうすればいいか
微生物に関する疑問	感染を起こしている微生物は何か 血液培養が陽性になったが、どうすればいいか MRSAが検出されたがどうすればいいのか? 隔離予防策はどうすべきか? 結核ではないか?
抗菌薬・ワクチンに関する質問	ワクチンの適応は? 周術期予防抗菌薬の選択の相談 抗菌薬の選択は?
経過観察	感染が改善しない

C-3. 国際疾病分類に関するレビュー

1. 国際疾病分類の歴史:成立とその改訂過程

19世紀中ごろより、英国において死亡登録における「死因 causes of death」の記載に統一された病名が記載されず統計学者が登録データを有効に活用することができなかったことを背景に、統一された疾病分類が求められはじめた。

(参考)死因とは、「死亡を引き起こしたか、その一因となったすべての疾病、病態または損傷、およびこれらの損傷を引き起こした事故または暴力の状況」をいう。

ヨーロッパでは1855年の国際統計会議(the International Statistical Congress)において、国際的に統一された疾病分類が検討され、William Farrが作成した疾病分類が採用され、内容の検討が始まったとされる。初めに作成されたものでは、死因は解剖学的に139の見出しに分類された。この分類はJacques Bertillonが議長を務める国際統計機関(the International Statistical Institute)で改訂され、簡約版44項目、99項目、161項目と複数の版を作製し、国際的にもヨーロッパのみならず、北米やメキシコでも利用され始めた。1900年の第1回改訂には26カ国が参加し、内容の検討を行っている。

わが国においては、明治31年の戸籍法の制定によって現在の形式による人口動態統計が確立した際、死因・疾病分類も公衆衛生の新しい立場から検討が行われていたが、1900年(明治32年)に国際的検討に

従いICDを人口動態統計とともに適用した。この時点では、the International List of Causes of Death. という名称であった。

その後も約10年ごとに国際統計機関にて内容の改訂が行われたが、第5回、第7回、第9回改訂に大きな内容の変更が行われた。

1938年の第5回改訂では、保険会社や医療機関等の要請に応えつつ世界各国でばらばらに病気分類が作成されていたことにも鑑み、国際死因分類に加え、疾病分類(International Lists of Diseases)としても使用できるよう配慮され、名称も「疾病・傷害及び死因分類(the International Statistical Classification of Diseases, Injuries, and Causes of Death)」と改められた。死亡に至るわけではない疾患、特に感染症・寄生虫疾患の領域や外傷の分類が重点的に強化された。また一方で、死因統計分類において、表章死因を「原死因」とし、死亡診断書の様式を統一し国際様式を定めた。1948年の第6回改訂からは、WHOが設立されたことより内容の議論の場をWHOに移し、World Health Assemblyで改訂内容が発表されることで、「疾病・傷害及び死因分類」は超国家的な統計分類として発展する新たな段階に到達している。

1955年の第7回改訂では、死産原因、精神障害および循環器疾患の3分野の分類の改正が重点にとりあげられた。

1975年の第9回改訂では、前回の第9版から大きく変化し、より詳細な分類区分が設けられた。①コードの1桁目をアルファベットとした。(桁数を増やさずより多くの分類が可能となった。)②分類項目数が3桁分類項目2,036、4桁分類項目12,159と倍増し、臨床統計としても使用が可能となった。より使いやすさと医学の進歩や疾患の変化に対応できる形式に変更された。③内容例示表の章構成が17章から21章に拡大し、保健サービス等に関する章が加わった。これに伴い、名称も従来の「国際疾病分類」から「疾病及び関連保健問題の国際統計分類」と変更されている。現在はICD-11への改訂が準備されている。

C-4. 患者の視点の病名確定上の問題抽出データの分類

要素名	カテゴリー名
病院へ行くきっかけ	事故などのきっかけで検査を受ける。
	家族・友人などに勧められる。
	ライフイベント(結婚など)をきっかけに。
	家族の病気
	無料の検査があった 人間ドックで異常があった。
病院へ行くことを妨げる要因	治療法がないという情報
	病院へ行くのが怖い
	仕事などで時間がない
どこを受診すべきかの判断	整形外科
	接骨院
	近所の内科
病院での誤診断	年齢的なものだと言われる。
	検査を受けるが異常はないと言われる。
	違う診断をされる。
専門家がわかったきっかけ	自分で調べる(インターネット)
	漢方薬局で大学病院を勧められる。
	小さな病院で大学病院を紹介される。
	たまたま専門家がいた。
	接骨院で病院を紹介される。

D. 考察

D-1. 疾患同定支援ツールの検討

①退院時診断困難例入院患者の検討

我々が知る限りでは、退院時に診断不可能であった症例の頻度については今まで報告がない。この中から希な疾患または未分類疾患が含まれている可能性があり、未分類疾患の情報源として検討すべき集団と思われた。

しかし、いくつかの限界もある。退院患者の中で、主病名に疑いが含まれている集団の中には、退院後に診断が比較的早く明確になる群、誤診の群などが含まれている可能性があり、退院後に一定期間フォローし再検討が必要となろう。しかし、フォロー期間や、ドロップアウトしないでフォローする方法論について検討すべきである。また、過去のデータを利用しているため診断の再検討のために必要なデータの欠損や