

クランクのレベルでは判断がつかず、医師への確認を要するものがある。(2)WISH<sup>2</sup>の疾病データベースへの変換に支障を来すものがある。例えば、従来は「年月」のみの要求だったのが、「年月(日)」になっている。疾病データベースと異なる場合、年月日のマッピングに支障を来す。例えば、「年月(例 2012/01)」は年月日に直した時に、日はどのように表現するか。不用意に変換することは、データの精度を失い、統計処理上の齟齬を来すことから回避する必要がある。疾病データベース、国立保健医療科学院(以下、「科学院」という。)でホストするデータベースの構造を統一化し、個人票の版管理と同期させることを検討すべきである(考察・提案 4.1)。

#### D. 考察

4.1 難病データベースは厚生労働省と科学院で個別に持つのではなく、一本化すること。

3.5 節で述べたように、データベースの構造の違いは機械的なマッピングでは不可能であり、場合によっては別途に入力を依頼せざるを得ない場合がある。現場にも負担がかかり、また同種のデータベースが複数存在することは不経済である。臨床調査個人票のデータベースは一本化させ、かつ臨床調査個人票入力システム、臨床調査個人票バージョンと整合させた開発体制の構築を提案する。

4.2 疾病データベース上の認定判断ロジックを臨床調査個人票入力システム上で稼働させること。個人票を入力した時点、認定に相当するか判明すれば、患者も保健所も無駄足を運ぶことがなくなる。また、別途にて触れている個人票の記載内容の正確性、ロジックチェックと認定判断ロジックを併せ実施することにより、認定に最低限必要なデータが入力

されることが担保され、個人票のデータ精度が向上することが期待出来る。

4.3 ソフトウェア開発委託は、成果物にソースコードとドキュメントをうけ取ること。

臨床調査個人票の収集は患者の生涯にわたり、システムの維持も長期間にわたる。そのため、企業の継続性のリスクも考え、ソフトウェアの開発委託は成果物としてソースコードとドキュメントを要求すべきである。また、①長期的なサポートの過程で別の会社に委託開発を依頼する可能性がある、②利用しているソフトウェアの商用サポートが停止する可能性がある、ことから、①委託開発したソフトウェア及びソースコードについて科学院側に一定の権利を確保する契約内容に留意すること。②商用サポートの停止リスクに対応するために、オープンソースのソフトウェアもしくは仕様が開示され、代替のソフトウェアが容易に提供されることが予想されるような構成でシステムを構築すること。

#### 5. 仕様要件

本調査であきらかになったことから、今回の臨床調査個人票入力システムの仕様策定において留意すべきことを列挙する。

##### 5.1 患者基本情報の連携

患者基本情報の連携機能は最低の仕様として実装すること。病院によって医療情報システムの多様性があることから、個人票入力システムは、単体版、サーバ版のそれぞれに併せた、いくつかの連携方法について提供すること。

単体版：

1. 所定様式の CSV ファイルのインポート
2. (検討中) ODBC による接続、参照 所定様式のビューが前提

サーバ版：

1. 所定様式の CSV ファイルのインポート
2. ODBC による接続、参照 所定様式のビューが前提

<sup>2</sup> 厚生行政総合情報システム(WISH: Wide-area Information-exchange System for Health and welfare administration)

### 3. HL7 患者基本情報のメッセージファイルインポート

#### 5.2 有・無・不明の入力のデザインの統一

入力においては、未入力・有・無・不明の区別がつくようにデザインを統一すること。ラジオボタン、チェックボックスともに、デフォルトではどれも未入力の状態（チェックされていない）にすること。

#### 5.3 入力されたデータの臨床的整合性の確認

項目のマスタ定義において、簡単なチェックロジックを指定できるようにすること。チェックロジックの記述からアプリケーションでのチェック機能を自動生成できることが望ましい。容易にメンテナンスできる方式で実装すること。なお、このチェックは従来のフィールドの値制約のロジックとは別途に実装すること。

#### 5.4 エラーチェック

エラーのために一時保存を拒否すると下書き保存後の作業が進まなくなるため、エラーの指摘があっても一時保存・保存が申請まで出来るようにすること。なお、保存時にはエラーが発生している場所が一目でわかるように、エラー内容を一括表示する場所を設けること。また、データ申請のための書き出しの時に、書き出す前にエラーがあるかどうかを確認できるようにすること。

## E. 結論

情報（IT）技術は、特に社会、経済の多くの部分を変革してきた。しかしながら、医療分野においては著しく遅れているのが実情である。我々の研究は、これら IT 技術を用いて医療機関と保健行政をつなぐことで効率的な疾病の管理を目指すものである。米国においては、IT 技術の保健医療分野への推進に向けて「大統領府報告書 米国民の保健医療改善に向けた保健情報技術の潜在力の完全開花：これからの道（大統領科学技術諮問委員会

2010 年 12 月）」が取りまとめられているが、この中で「保健医療制度には現在、市場における効果的な保健 IT システムの形成に向けたイノベーション及び活発な競争の意欲をそぐような障壁が複数存在している。」と指摘しており、これはまさしくわが国の現状にも当てはまる[1]。このような問題の解決に向けて、同報告書は、「タグ付きデータ要素手法の一つの重要な利益は、データ要素ごとにセキュリティーのレベルを変えることができる点である。これは一つのシステムで、治療中の主治医が患者の完全な記録にアクセスすること、医療報酬支払者が医師の要約統計にアクセスすること、そして研究者が様々な治療法の臨床研究ができるように識別情報を除いたデータにアクセスすることを許すことができることを意味する。データ・アクセスを層化したシステムによって、保健医療提供者、患者、研究者の用いるアプリケーションはそれぞれが別々の種類のデータを引き出し、用いることができるように役割を付与されるかもしれない。これによって、患者と保健医療提供者は一定目的のためにはデータを共有し、もしかするとそれ以外の目的においては共有しないことに同意することができよう。」と記している。そこで、我々は本研究を進めていく上で、「データ交換基盤構造の形成」と「データ・アクセスの規制」の2つの視点から新たに着手するものである。

#### 1) データ交換基盤構造の形成

最終目標は全国的基盤構造の形成であり、参加者の規模拡大で一人当たりの費用を低減できる。かならずしも集中システムである必要はなく、分散でも良い。政府の強力な牽引力が必要となる。(カナダの Infoway のような連邦政府レベルでの青写真をまず作り、その戦略に沿って地域に分散構築する。日本であれば、道州制を想定した5地域程度。現在総務省などがやっている、県レベルでの、独立

したやり方は全くこれに逆行。将来に禍根を残す。)

## 2) データ・アクセスの規制

従来の紙の記録と異なり、電子記録で「構造化」した情報は、様々な単位で情報のアクセス制御が可能となる。MML (Medical Markup Language) や CDA (Clinical Document Architecture) が有用なのはまさにこの点にある。単なる Messaging である HL7 v.2.X では不可能。医師、患者、保険者、研究者、行政などの様々な立場へのアクセス制御は、EHR における重要な技術要素であり、これを実現することの出来る情報規格を採用することが不可欠である (現時点では、MML, CDA, OpenEHR)。

### [参考文献]

#### 1) 米国大統領府報告書 (2010年10月)

[http://lob.kuhp.kyoto-u.ac.jp/document/20111025PCAST\\_j.pdf](http://lob.kuhp.kyoto-u.ac.jp/document/20111025PCAST_j.pdf)

#### F. 健康危険情報

なし

#### G. 研究発表 (2010/4/1~2011/3/3 発表)

##### 1. 論文発表

- 1) Kimura E, Kobayashi S, Kanatani Y, Ishihara K, Mimori T, Takahashi R, Chiba T, Yoshihara H. Developing an Electronic Health Record for Intractable Diseases in Japan. Stud Health Technol Inform. 2011;169:255-9.

- 2) 金谷泰宏, 木村映善, 小林慎治, 玉置洋, 荻野大助, 吉原博幸, 千葉勉 臨床調査個人票の有効活用及び臨床データベースの構築 (総説) 保健医療科学 2011;60(2):100-104.

## 2. 学会発表

- 1) Kobayashi S, Kimura E, Yoshikawa T, Kanatani Y, Ishihara K, Yasukawa M, Kuroda T, Yoshihara H. Clinical data Modeling for national surveillance of rare diseases in Japan. ICORD 2012 Conference, Feb 2-4, 2012 Tokyo Japan.

## H. 知的財産権の出願・登録状況

(予定を含む。)

### 1. 特許取得

### 2. 実用新案登録

### 3. その他

表 2. 臨床調査個人票の不備内容

疾患名	タブ名	項目	内容	件数
多発性硬化症	基本情報	出生都道府県	空欄(未入力)	-
		発病時在住都道府県	同上	-
		発病年月(必須項目)	同上(月のみ未入力も含む)	1/12
		初診年月日(必須項目)	同上(月日のみ未入力も含む)	2
		保険種別	空欄(未入力)	9
		身障手帳:有無	同上	-
		介護認定:有無	同上	-
		生活状況社会活動	同上	1
		受診状況:通院回数	同上	-
		全身性エリテマトーデス	基本情報	出生都道府県
発病時在住都道府県	同上			17
発病年月(必須項目)	同上(月のみ未入力も含む)			14
初診年月日(必須項目)	同上(月日のみ未入力も含む)			3
保険種別	空欄(未入力)			22
身障手帳:有無	同上			6
介護認定:有無	同上			6
生活状況社会活動	同上			-
受診状況:通院回数	同上			-
関節性多動脈炎	基本情報			出生都道府県
		発病時在住都道府県	同上	1
		発病年月(必須項目)	同上(月のみ未入力も含む)	-
		初診年月日(必須項目)	同上(月日のみ未入力も含む)	-
		保険種別	空欄(未入力)	1
		身障手帳:有無	同上	-
		介護認定:有無	同上	-
		生活状況社会活動	同上	-
		受診状況:通院回数	同上	-
		潰瘍性大腸炎	基本情報	出生都道府県
発病時在住都道府県	同上			2
発病年月(必須項目)	同上(月のみ未入力も含む)			9
初診年月日(必須項目)	同上(月日のみ未入力も含む)			2
保険種別	空欄(未入力)			31
身障手帳:有無	同上			1
介護認定:有無	同上			4
生活状況社会活動	同上			2
受診状況:通院回数	同上			10
発症と経過	前医の詳細			同上
治療	内科的治療	「あり」「なし」両方にチェックあり→カルテ確認要	1	
パーキンソン病	基本情報	出生都道府県	空欄(未入力)	1/43
		発病時在住都道府県	同上	1
		発病年月(必須項目)	同上(月のみ未入力も含む)	20
		初診年月日(必須項目)	同上(月日のみ未入力も含む)	5
		保険種別	空欄(未入力)	36
		身障手帳:有無	同上	5
		介護認定:有無	同上	13
		生活状況社会活動	同上	1
		受診状況:通院回数	同上	8
		原発性胆汁性肝硬変	基本情報	出生都道府県
発病時在住都道府県	同上			-
発病年月(必須項目)	同上(月のみ未入力も含む)			7/13
初診年月日(必須項目)	同上(月日のみ未入力も含む)			3
保険種別	空欄(未入力)			6
身障手帳:有無	同上			-
介護認定:有無	同上			1
生活状況社会活動	同上			-
受診状況:通院回数	同上			3
最近6か月以内の臨床症状	診察日			空欄(未入力)
診断時の臨床症状	診断日	同上	2	

申請する上で、修正が必要な箇所のみをあげています。(前回数分は全て入力するものとした場合も含む)

以下、基本情報タブを全疾患共通としてまとめたもの(上記と重複)

疾患名	タブ名	項目	内容	件数
全疾患共通	基本情報	出生都道府県	空欄(未入力)	22/182
		発病時在住都道府県	同上	21
		発病年月(必須項目)	同上(月のみ未入力も含む)	51
		初診年月日(必須項目)	同上(月日のみ未入力も含む)	15
		保険種別	空欄(未入力)	105
		身障手帳:有無	同上	12
		介護認定:有無	同上	24
		生活状況社会活動	同上	4
		受診状況:通院回数	同上	21
		潰瘍性大腸炎	発症と経過	前医の詳細
治療	内科的治療		「あり」「なし」両方にチェックあり→カルテ確認要	1
原発性胆汁性肝硬変	最近6か月以内の臨床症状	診察日	空欄(未入力)	3/13
	診断時の臨床症状	診断日	同上	2

表 3. 医療クラークの入力ミスの内訳

疾患名	タブ名	項目	訂正内容	件数
全疾患共通	基本情報	① 申請区分	継続のチェック忘れ(新規のままになっている)	15
		② 出生都道府県	不明時も不明を選択せずデフォルトの愛媛県のまま	20
		③ 発病時在住都道府県	同上	20
		同上	転勤・引越し等で2都道府県の入力不可 例:「愛媛～大阪」	1
		④ 保険者番号	特定疾患の受給者番号が入力	5
		記号	記号・番号の数字が逆	5
		同上	記号なしの場合のスペースの入力忘れ	5
		番号	番号の入力間違い	5
		⑤ 身障手帳:有無	不明時に不明を選択せずに空欄	15
		⑥ 家族歴有無	同上	15
全疾患共通	全項目		保存を押さずに閉じたため、途中から入力されていないもの	10
			チェック入力時にクリックですれて、違う箇所にチェックがついている	10
全身性エリテマトーデス	合併症	⑦ その他の合併症	チェックなしを不明にチェックを入れて、消せないためそのまま	5
潰瘍性大腸炎	治療	⑧ 薬剤による副作用	チェックなしをなしにチェックを入れて、消せないためそのまま	5
パーキンソン病関連疾患	基本情報	⑨ 家族内発症者続柄	不明時も不明を選択せずデフォルトの同胞のまま	20
		⑩ Schellong試験:臥位収縮期血圧	114/65と記載(収縮期/拡張期の区別なく入力)	5
		⑪ 塩酸セリグリン	既存の様式と順番が入れ替わっているため、逆に入力	5

表 4. 臨床調査個人票入力システムの問題点

疾患名	タブ名	項目	訂正内容
全疾患共通	基本情報	① 発病年月	Drが未入力時の対応(必須項目のため、00000000入力)
		② 初診年月日	同上入力するも、下書き保存後再度開くと空欄になっている
		③ 受診状況:通院回数	小数点入力すると、パターンと一致しませんとなり、入力不可 → 対処法として、受診状況:その他の内容に手入力
多発性硬化症	臨床所見 治療状況	④ EDSS	0、1.0、...、10等の数値の記載がないので、入力不可
		⑤ (一番初めに)	具体的に記入するスペースがないため、入力不可
		⑥ ステロイド最大投与量の単位	複数選択不可のため、入力不可
全身性エリテマトーデス	臨床症状および所見	⑦ 慢性投与の維持量の単位	→ ステロイド最大投与量、慢性投与の維持量欄に手入力
		⑧ 精神・神経症状	神経障害の項目がないため入力不可
		⑨ 血液系症状	白血球減少、リンパ球減少、血小板減少、抗二本鎖DNA抗体価高値(免疫学的異常) (免疫学的異常)の項目はないが、検査所見欄の検査結果数値のみで可能か(小林先生談)
	治療	⑩ (全体として)	更新時は、「最近1年間・全経過を通じて」の2項目ありますが、データとして蓄積されるので不要ではないでしょうか(小林先生談)
	合併症 (その他)	⑪ (全体として)	同上
		⑫	医療上の問題点の記入欄がないため入力不可
結節性多動脈炎	発症と経過 検査所見 現在の治療状況	⑬ 病理学的検査実施日	現行では年月のみの入力であり、カレンダー入力不可
		⑭ 日付	同上
		⑮ (全体として)	「主に薬物療法、薬物+外科的治療、外科的治療」の項目がないため入力不可
潰瘍性大腸炎	病態 治療	⑯ 最近の罹患部位	複数選択不可のため、入力不可(ほぼ複数の入力です)
		⑰ 手術理由	未実施時はデフォルトの「出血」のまま → 「その他」を選択しましょうか?
		⑱ 日付	各検査施行日の相違時、日付入力不可
パーキンソン病	画像所見	⑲ 排便回数	「〇～〇回」等の入力時、パターンと一致していませんとなり、入力不可
		⑳ CT 撮影日	現行では年月のみの入力であり、カレンダー入力不可
		㉑ MRI 撮影日	同上
	治療 栄養・呼吸の状態	㉒ SPECT 実施日	同上
		㉓ 定位脳手術 実施日	同上
		㉔ 鼻腔栄養開始時期	同上
		㉕ 胃瘻開始時期	同上
	㉖ 気管切開時期	同上	
	㉗ 人工呼吸器開始時期	同上	
原発性胆汁性肝硬変	診断時の臨床症状	㉘ 診断日	同上
	検査	㉙ 検査年月	同上
	診断	㉚ 診断根拠	新規時のみ選択するが、更新時もいずれか選択しないと、申請がかけられない
	薬物療法	㉛ 開始年月日	現行では年月のみの入力であり、カレンダー入力不可
	薬物療法以外の治療	㉜ 開始年月	同上
	㉝ 中止年月	同上	

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

臨床調査個人票の有効活用及び臨床データベースの構築

研究分担者 奥村 貴史（国立保健医療科学院 研究情報支援研究センター  
特命上席主任研究官）  
緒方 裕光（国立保健医療科学院 研究情報支援研究センター長）  
研究協力者 種田 憲一郎（国立保健医療科学院 医療・福祉サービス研究部）  
建石 由佳（国立保健医療科学院 研究情報支援研究センター 客員研究員）

研究要旨

平成 21 年度 厚生労働科学研究費補助金 難治性疾患克服研究事業「未分類疾患の情報集約に関する研究」において、診断に悩む臨床医や患者を支援しつつ症例情報を効率的に集積しうる可能性を示した。今年度の本分担では、そうした臨床データベースを実現するために、診断支援システムと症例データベースを低コストに構築、運用していくための手法について研究を進めた。前者については、コストの多くが疾患とその症状に関する膨大な情報の編纂に費やされる。そこで、疾患知識ベースを高い精度でかつ低コストに整備するために、精度を重視した手動編纂と疾患のカバー範囲を重視した自動編纂とを組み合わせたハイブリッド型の疾患知識ベースを構想し、その実現に向けた自然言語処理研究を行った。後者については、症例データベースを安価に提供していくために、クラウド技術を用いた研究班支援システムの実現について検討を行った。これらの研究成果を用いることにより、今後、診断困難症例に悩む臨床医、難病研究に携わる諸研究班を支援する難病研究ポータルサイトを効率的に構築しうる。

A. 研究目的

難病研究には、患者の症例データベースの効率的な整備が欠かせない。しかしながら、臨床医にとって症例情報の提供は負担が大きい。また、詳細な情報を集めようとすればするほど臨床側の負担が増し、情報量を減らせば減らすほど研究上の有用性が低下するという関係にあり、症例情報の効率的な集積手法が求められていた。

そこで我々は、平成 21 年、「厚生労働科学研究費補助金 難治性疾患克服研究事業未分類疾患の情報集約に関する研究」において、従来の手法による情報収集が困難な診断困難症例に関する情報を効率的に収集するための手法について検討を行った。その結果、診断が

付かずに困っている臨床医や患者に対して診断支援システムなどの有用なサービスを提供することで、診断困難な症例の情報を自動的にかつ低コストに収集しうる可能性を示すことが出来た。

この「診断支援システムと一体化した臨床データベース」という新たな手法の有効性を実証するうえでは、実際に臨床医にとって有用な診断支援ツールを提供し、また、難病研究班に対する症例データベースや各種の研究班支援ツールを提供していく必要がある。そこで今年度は、上記のような臨床データベースの実現に向け、診断支援システムの鍵である疾患知識ベースと、症例データベースを低コストに構築、運用していくための手法に関する研究を行った。

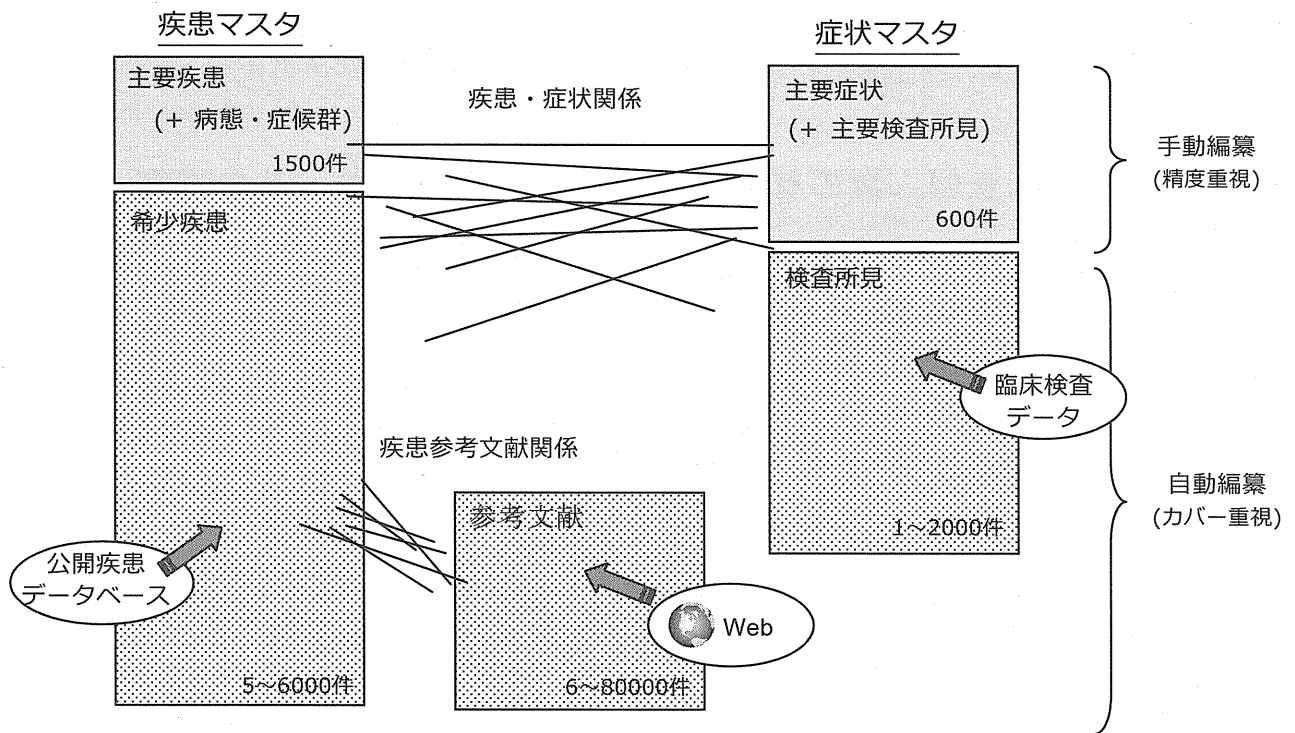


図1 ハイブリッド型疾患知識ベースの概念図

## B. 研究方法

診断困難症例を診察する臨床の現場に役立つ診断支援システムを構築するためには、診断を行う高性能なアルゴリズムに加えて、既知の疾患をくまなく網羅した疾患知識ベースが不可欠である。しかしながら、既知の疾患を網羅するためには5000~7000と言われている膨大な量の疾患情報を収集し、機械可読な形式に編集しなおす必要がある。さらに、この編集に際しては医学知識が不可欠であるため、十分な疾患数と精度を両立した疾患知識ベースの編纂には、監査を医師に委託する膨大なコストが求められることになる。一方、診断アルゴリズム側は、基本的には小さなプログラムであり、開発コストは限られている。したがって、診断支援システムの開発に要するコストの大半は、疾患知識ベースの整備に費やされることになる。

しかしながら、我々の現有する疾患知識ベースの登録疾患数は1500疾患前後に留まり、

新たな症例情報の集積手法を実証するために数千万に登る開発費用を確保することは困難な状況にある。そこで、我々は、重要疾患や頻度の高い疾患や医学的 중요性の高い疾患等を対象としたデータ精度を重視した手動編纂によるデータと、既知の疾患を網羅的にカバーするために既知の疾患情報から自動的に編纂したデータとを統合することで、知識ベースの整備コストを下げ、また、実用的な精度を保つことが出来るハイブリッド型知識ベースを提唱してきた(図1)。

今年度は、さまざまな医学文献から自動的に疾患情報を抽出するために、自然言語処理の研究者である工学院大学 建石由佳 准教授との共同研究を進めた。具体的には、公開されている希少疾患データベースである Online Mendelian Inheritance in Man (OMIM) より疾患に関する医学文献を取得し、それぞれの疾患に生じる症状や臨床検査結果の情報抽出を試みた。また、数多くの臨床検査とその結果の解釈を機械可読とするために、「臨床検査データブック」を対象とした情報抽出

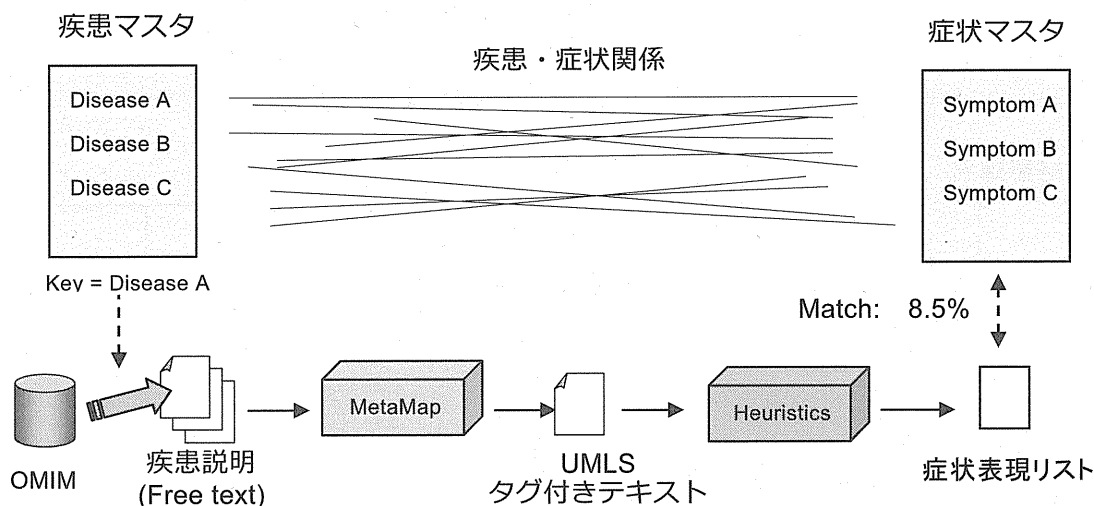


図 2 MetaMap+ヒューリスティクスによる表現抽出

を行い、臨床検査の可能な結果(たとえば、GOT 上昇)を症状マスタへと登録すると共に、『臨床検査データブック』に記載されている検査結果と可能な疾患(たとえば、肝炎)との対応を「疾患・症状関係」として整理した。さらに、それぞれの疾患について、診断支援システム利用者に適切な情報提供を行うために、診療に有用な医学文献を自動提示するための手法を検討した。

(倫理面への配慮)

本研究では、倫理面の配慮を必要とする患者に関するデータを扱っていない。

## C. 研究結果

### 1) 医学文献からの情報抽出

疾患と対応する症状や臨床検査結果の情報抽出には、疾患に関して自由記載された様々な症状から、発熱や頭痛といった臨床所見に関する表現を抽出し、症状マスタを走査し症状 ID を確定する必要がある。

これは一見、「頭痛」という表現に対して、頭痛の ID を検索し付与する単純な処理のように見える。しかしながら、たとえば、「精

神遅滞」という症状が ID を付与されていたとしても、医学文献中で「学業成績が極めて悪く」のように表記されている場合、これらを同一の概念と自動的に判断することは技術的に難度が高い。そこで、我々は、NLM (National Library of Medicine) の公開している MetaMap というツールを活用し、独自の解析手法を組み合わせ精度を向上させる手法について検討を行った。

図 2 に、この情報抽出手法の概要を記す。自由記載された疾患の説明文に対して、MetaMap を用いて UMLS (Unified Medical Language System) によるタグ付けを行い、さらにこのタグ列に対してヒューリスティクスを適用することにより、Recall (感度) 89%、Precision (特異度) 64% という高い性能での症状表現の抽出が可能となった (Okumura 2012)。

その上で、抽出した症状表現に症状 ID を振るために、症状マスタとのマッチングを行った。単純な検索手法では、この効率は 8.5% 程度に留まるために、疾患症状関係の効果的な自動生成に向けて、今後、入力に用いる文献数を増やすか、精度を改善していく必要が示された。



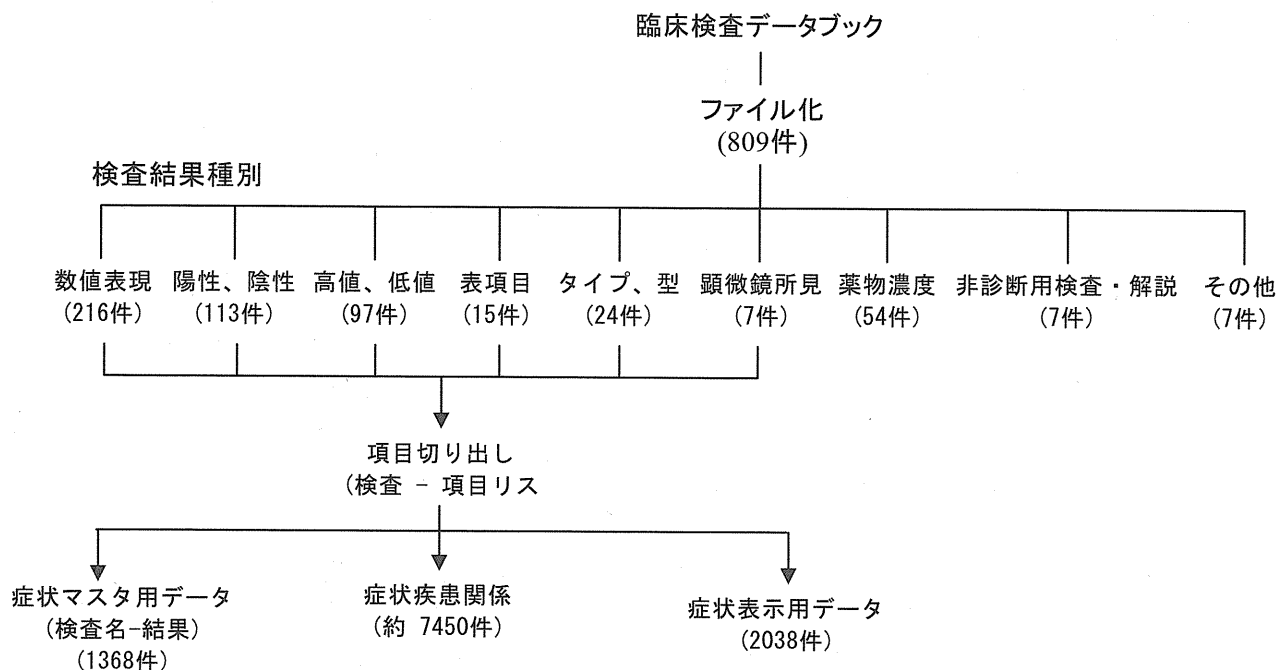


図 3 臨床検査情報のデータ化概要

## 2) 臨床検査情報からの情報抽出

臨床現場では、診察により得られる様々な所見の他に、血液検査や検体検査など様々な検査が行われているため、効果的な診断支援システムの開発に際しては、そうした臨床検査の結果と可能な疾患の関係についてのデータ化が必要である。そこで次に、「臨床検査データブック」より、臨床検査に関する情報の自動抽出に取り組んだ。

まず、データ化の対象となる情報の概要について、目視による統計情報の取得を行った(図3)。その結果、収載されている臨床検査は、検査結果が数値で表されるのか、陽性・陰性で表されるのか、血液型のように型で表現されるのか等、検査結果の表現により9種類に分類しうることが明らかとなった。そのうえで、診断支援システムにて活用しうる6種類の検査についてのデータ化を試みた。その結果、臨床検査名と結果の対(たとえば、GOT・上昇)が1368件、症状疾患関係が7450件得られた。また、手作業で2038件の診断支援システムの表示用データを生成した。

これらのうち、1368件の臨床検査名と結果の対(GOT・上昇)は、新たな「症状」として、症状マスタに追加することになる。一方、その症状が指し示す7450件の症状疾患関係情報は、それぞれシステムに登録されている疾患IDに置き換える必要があり、これが再び多くの手作業を要することが明らかとなった。

たとえば、GOT上昇が「肝炎」や「肝臓癌」を示しているとする、それぞれの疾患に対して、対応する疾患IDを検索していく必要がある。しかしながら、この「疾患」情報は、単一疾患を指し示すケースだけでなく、「〇〇疾患の初期」や「△△の亜種」といった指定がされていることが少なくない。また、「肝炎」といった場合、ウイルス性肝炎もあれば薬剤性肝炎、自己免疫性肝炎など様々な肝炎があり、システム上は独立した疾患IDを有している。したがって、疾患に関する記載を、具体的な疾患IDに手動で展開していく必要がある。

今年度は、予備的解析を行い、疾患群が405件、疾患が1215件、修飾された疾患が291件と、基礎的統計を整理することが出来た。



また、研究班支援ツールを低コストかつ拡張可能な形で提供していくために、昨年度までに準備をしてきた「難病研究ポータル」を、柔軟な情報サービスの構築を可能とする「科学院クラウド」上に実現する準備を行った。これにより、今後、難病研究班支援システムを構築していく上で、クラウド技術を用いたサービス提供が可能となる。

来年度の課題として、以上のサービスを各種の難病研究班に実際に提供していく体制の構築が上げられる。とりわけ、niph.go.jp ドメインでのサービス提供は、一定の品質とサービス継続性が担保された態勢で行わなければ、今後の国立保健医療科学院における難病研究事業の継続そのものに悪影響を及ぼしうる。そのために、今後、研究班本体、あるいは、難病研究事業の今後の展開と密接に連携しながら進める必要がある。また、実際の臨床に役立つ診断支援システムの実現に向けて、疾患知識ベースの自動生成研究を進めるとともに、診断アルゴリズムとシステム部分の改良を行う必要がある。

## E. 結論

今後の難病対策や難病研究事業のあり方を考える際、情報技術の有効活用は欠かすことが出来ない。その際、有用性の高い診断支援システムと、効率の良い研究班支援システムは、難病や希少疾患の情報集約において中心的な役割を担いうる。そこで、今年度は、臨床データベースの実現に向け、診断支援システムと各種の研究班支援システムを低コストに構築、運用していくための手法について研究を行った。

診断支援システムの開発においては、コストの大部分が疾患知識ベースの整備に費やされることになる。そこで、精度とコストを両立させるために、精度を重視した手動編纂と疾患のカバー範囲を重視した自動編纂とを組み合わせたハイブリッド型の疾患知識ベース

を提案し、その実現に向けた自然言語処理研究を行った。

また、症例データベースに代表される研究班支援ツールについては、低コストに各種ツールを提供していくうえでの鍵となるクラウド技術を用いた研究班支援システムの実現について検討と準備を進めた。

本研究分担は、わが国の難病研究や難病行政の効果的な情報化に向け、平成 21 年度より診断支援システム、症例データベース、難病研究班支援システム等の研究成果を蓄積して来た。今年度は、とりわけ、整備コストを軽減しつつ効果的な情報システムを実現する手法についての検討を重ねた。今後、これらの研究成果がわが国の難病研究や難病行政の発展に寄与することを願っている。

## F. 研究発表

1. 論文発表  
なし

2. 学会発表

T. Okumura and Y. Tateishi, "A lightweight approach for extracting disease-symptom relation with MetaMap toward automated generation of disease knowledge base", 1st International Conference on Health Information Science (HIS 2012), Apr 2012.

## G. 知的財産権の出願・登録状況

1. 特許取得  
なし

2. 実用新案登録  
なし

3. その他  
なし

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

難病患者の頻度の推計方法及び疫学的特徴の把握方法の開発

研究分担者 土井 由利子（国立保健医療科学院統括研究官）  
横山 徹爾（国立保健医療科学院生涯健康研究部長）  
高橋 邦彦（国立保健医療科学院政策技術評価研究部主任研究官）  
金谷 泰広（国立保健医療科学院健康危機管理研究部長）

研究要旨

特定疾患のなかで、特に死亡数の少ない疾患における患者の地域比較を行うための統計解析法とその応用について検討を行った。本研究では特定疾患医療受給者証所持者数を用い、都道府県単位、保健所管轄単位での性別、年齢調整を行ったうえで、リスクのベイズ推定ならび集積性の検出を試みた。今年度は複数の疾患でその地域性が異なるかどうか注目し、特に同じ医療機関で診断されると考えられる潰瘍性大腸炎とクローン病での特定疾患医療受給者証所持者数の動向・地域性を比較した。その結果この2つの疾患の受給者証所持者数の地域性が異なる特徴を示していることや、受給者の数自体が近年著しく増加していることが確認できた。

A. 研究目的

疾病の発生頻度や患者数の地域比較を行うことは、その疾病の特徴を捉えるうえで大変重要なことである。難病における死亡数の地域比較については、厚生労働科学研究費補助金の疫学研究班において、行政資料を用いた難病の頻度調査として検討が行われてきた。特に年間死亡数が100例を超えるような疾患を中心に19疾患については、二次医療圏単位で分析が行われ、「難病の死亡統計データブック」（土井・横山編；2006）としてまとめられている。

一方、年間死亡数が100例に満たないような疾患については、小地域での死亡数を見ると件数がゼロの地域が大部分となり、一般的な死亡リスク（SMR）などを指標とすると適切な地域比較が困難となる。そこで本研究ではそのような死亡数が少ない疾患における地域比較を行う際に、従来の方法に代わる適切なデータや統計解析手法について検討することを目的とする。

一般に死亡数が少ない疾患は、死に至る疾患であるが患者が少なく死亡数も少ないものと、患者数はそれなりに多いが死には至らないもののが考えられる。一方で難病において正確な患者数のデータを把握することは困難であり、これまで全国疫学調査、患者調査をはじめとする標本調査による推計、特定疾患医療受給者証所持者数などが患者数の目安として用いられ、検討が行われてきている。そこで本研究では死亡数は少ないものの患者数は多いと考えられる疾患に注目し、その患者数を表すものとして公開されている特定疾患医療受給者証所持者数を用い、その地域比較を試みる。

今年度は複数の疾患でその地域性が異なるかどうか注目し、特に同じ医療機関で診断されると考えられる潰瘍性大腸炎とクローン病の受給者証所持者数の動向・地域性について解析を通して、本研究で検討しているデータや手法の適用の有効性について検討を行った。

## B. 研究方法

難病における正確な患者数は把握できていないため、本研究ではそれに代わるものとして、特定疾患医療受給証所持者数を用いた地域比較を検討した。特定疾患医療受給証所持者数は、現在、厚生労働省衛生行政報告例において、全国（対象疾患別、性別、年齢階級別）および都道府県別（対象疾患別、性別）に集計されたデータが公開されており、難病の各研究班においても、患者数に代わる一つの目安として利用されている。一方、平成13～15年度においては地域保健・老人保健事業報告において、保健所別の統計表が公開されている。そこで本研究ではその中から平成15(2003)年度のデータを用いた地域比較を検討した。なお当該年度のデータでは全国566保健所管轄（ただし京都市は1つの保健所として報告）となっていた。また本年度は特に複数の疾患での地域性が異なるかを検討するため、潰瘍性大腸炎とクローン病の2疾患に注目し、検討を行った。

まず2つの疾患について、これまで報告されている日本全国における死亡数、年齢調整死亡数、ならびに推計されている患者数について概況をまとめた。同時に特定疾患医療受給者証所持者数をまとめ、それを用いて都道府県別・保健所管轄別に、年齢を調整した Standardized Morbidity Ratio (SMR) のベイズ推定値(EBSMR)を男女・疾患ごとに求めた。次に疾病集積性の検定である flexible scan statistic によって、保健所管轄単位のデータを用いて地域集積性の検出を行った。ここでの解析には FlexScan Ver. 3.1 (Takahashi, Yokoyama and Tango, 2010) を用い、各パラメータの設定として、モデル:Poisson model, 統計量: LLR with Restriction、モンテカルロ検定のための繰り返し数: 9,999回、最大地域数: 500、有意水準:  $\alpha=0.05$  とし、p値が0.05未満のものを有意な集積地域として検出した。

検出された集積地域の様子を地図で視覚的に比べるとともに、保健所管轄単位での EBSMR の相関を求め、その関連について検討を行った。

(倫理面への配慮)

本データで解析に用いたデータは個別のデータではなく集計されたものであり、倫理面で問題が生じることはない。

## C. 研究結果

公開されている地域保健・老人保健事業報告において、潰瘍性大腸炎およびクローン病の全国での2003年度医療受給者証所持者数、ならびに2003年死亡数・年齢調整死亡率(土井・横山編. 難病の死亡統計データブック; 2006)、2002年患者数・受診間隔90日までの者も含めた総患者数(横山・土井編. 平成14年患者調査による難病の受療状況データブック; 2008)はそれぞれ表1の通りであった。また2003～2009年の特定疾患医療受給者証所持者数(性・年齢階級別)は表2-1, 表2-2の通りであった(2004年以降は衛生行政報告例より引用)。いずれの疾患も死亡数は少ないが、医療受給者証所持者数や推計されている患者数でも、実際の患者数は多い疾患であることがわかる。さらに、受給者数は年々増加している様子がわかった。

次に都道府県別のSMR、保健所管轄別のSMR(それぞれベイズ推定値 EBSMR)、ならびに検出された集積地域をそれぞれ図表で示した(図1～4)。いずれの疾患においても男女とも受給者証所持者数の有意な集積が認められ、その地域や分布は疾患によって異なっている様子が観察された。特に都道府県単位ではなく、保健所管轄でみることで、より詳細な分布を観察することができた。

さらに潰瘍性大腸炎とクローン病の保健所管轄単位でのSMR(ベイズ推定値)の相関を見るため、それぞれのEBSMRの値を散布図に表し相関を求めた(図5)。その結果、男性

では相関係数が  $r=0.363$ 、女性では  $r=0.244$  となった。

#### D. 考察

本研究で用いた解析方法によって、疾患毎に医療受給者証所持者数の地域分布が異なることを見出すことができた。本研究では、特に潰瘍性大腸炎とクローン病について検討を行った。この2つの疾患は炎症性腸疾患という点で類似しており、診察・診療も似た病院等で行われると考えられるが、この2疾患において、受給者の地域分布が異なることが観察されたことは興味深い。実際、保健所管轄単位でのEBSMRの相関を観察しても、男女ともそれほど強い関連は見られなかった。これらの解析結果をもとに、今後、臨床研究の専門家らとも連携しながら、難病疾患ごとの地域性についても検討を行っていきたい。

なお、前述したように、特定疾患医療受給者証所持者数は、現在、厚生労働省衛生行政報告例において、全国（対象疾患別、性別、年齢階級別）および都道府県別（対象疾患別、性別）に集計されたデータのみが公開されている。市区町村別の継続したデータがあれば、それをもとに保健所管轄や二次医療圏を単位とした難病の頻度分布の地域比較が可能となり、行政施策・サービスに広く活用できるようになるものと思われる。

本研究で用いた医療受給者証所持者数であり、厳密な意味での患者数ではなく、それらの間には差異があることも考えられる。この点において、これらの結果の解釈の際には十分な注意が必要であろう。しかし表2の結果にもあるように近年、医療受給者証所持者数も年々増加しており、これらのデータを用いた地域比較は有用であると考えられ、これらを活用することは、今後の難病研究推進に寄与できるものと考えられる。今後、関連したデータを用いた、より高度な方法での解析なども検討しながら、これらのデータを有効に活用した解析を行い、その上で、実際の解

析結果の解釈については、疾患ごとに専門家の協力を得ながらより詳細な検討を行うべきであろう。

#### E. 結論

特定疾患医療受給者証所持者数を用い、都道府県単位、保健所管轄単位での性別、年齢調整を行ったうえで、リスクのベイズ推定ならび集積性の検出を試みた。この方法を適用することにより、疾患ごとに地域差を観察することができた。本研究で検討した結果を臨床の専門家に情報提供し難病推進に役立てていきたい。さらに本成果も踏まえ、今後の難病に関するデータ収集ともリンクさせた議論も重要であると考ええる。

#### F. 研究発表

##### 1. 論文発表

なし

##### 2. 学会発表

高橋邦彦、横山徹爾、金谷泰宏、土井由利子. 特定疾患（難病）医療受給者証所持者数の地域比較. 第22回日本疫学会学術総会 2012年1月28日

#### G. 知的財産権の出願・登録状況

(※予定を含む)

なし

表 1: 潰瘍性大腸炎・クローン病における全国の 2003 年度医療受給者証所持者数, 死亡数・年齢調整死亡数, ならびに推計された 2002 年総患者数

	2003 年度医療受給者証所持者数		死亡数 (2003 年) *		年齢調整死亡率 (/100 万) (2003 年) *		総患者数 (2002 年)**		受診間隔 90 日までの者も含めた総患者数 (2002 年)**	
	男	女	男	女	男	女	男	女	男	女
潰瘍性大腸炎	40,212	36,958	7	44	0.08	0.46	24,100	25,000	36,100	39,100
クローン病	15,436	6,904	31	29	0.39	0.34	11,600	4,300	16,400	5,400

\*難病の死亡統計データブック (土井・横山編 ; 2006)

\*\*平成 14 年患者調査による難病の受療状況データブック (横山・土井編 ; 2008)

表 2-1(a) : 特定疾患医療受給者証所持者数 (年齢階級別) 潰瘍性大腸炎・男性

潰瘍性大腸炎	男性 総数	0～9 歳	10～19 歳	20～29 歳	30～39 歳	40～49 歳	50～59 歳	60～69 歳	70歳 以上
2003年	40,212	43	1,437	6,730	9,476	7,850	6,663	4,648	3,365
2004年	41,938	56	1,482	6,611	9,823	8,300	7,042	4,998	3,626
2005年	45,037	52	1,543	6,667	10,436	8,930	7,762	5,525	4,122
2006年	48,153	45	1,531	6,745	11,289	9,603	8,395	5,893	4,652
2007年	51,829	35	1,534	6,839	11,944	10,714	8,821	6,693	5,249
2008年	56,108	32	1,576	7,038	12,587	11,758	9,533	7,737	5,847
2009年	60,775	33	1,674	7,317	13,411	13,031	10,192	8,599	6,518

表 2-1(b) : 特定疾患医療受給者証所持者数 (年齢階級別) 潰瘍性大腸炎・女性

潰瘍性大腸炎	女性 総数	0～9 歳	10～19 歳	20～29 歳	30～39 歳	40～49 歳	50～59 歳	60～69 歳	70歳 以上
2003年	36,958	51	1,112	5,407	8,783	6,927	6,502	4,814	3,362
2004年	37,959	48	1,093	5,267	8,996	7,254	6,688	4,999	3,614
2005年	40,416	37	1,097	5,378	9,495	7,753	7,259	5,386	4,011
2006年	42,474	29	1,151	5,348	10,048	8,236	7,605	5,641	4,416
2007年	45,164	44	1,203	5,467	10,520	9,003	7,780	6,246	4,901
2008年	48,613	47	1,167	5,697	11,123	10,080	8,113	6,974	5,412
2009年	52,531	50	1,246	5,985	11,731	11,331	8,634	7,571	5,983



表 2-2(a) : 特定疾患医療受給者証所持者数 (年齢階級別) クロウン病・男性

クロウン病	男性 総数	0～9 歳	10～19 歳	20～29 歳	30～39 歳	40～49 歳	50～59 歳	60～69 歳	70歳 以上
2003年	15,436	21	609	4,320	5,739	2,781	1,144	550	272
2004年	16,047	24	648	4,195	6,027	3,064	1,252	551	286
2005年	16,994	18	689	4,106	6,420	3,421	1,411	589	340
2006年	17,923	20	733	4,140	6,697	3,760	1,585	621	367
2007年	19,066	21	776	4,249	6,996	4,191	1,736	711	386
2008年	20,473	20	828	4,401	7,309	4,705	1,944	834	432
2009年	21,645	21	900	4,486	7,541	5,189	2,105	913	490

表 2-2(a) : 特定疾患医療受給者証所持者数 (年齢階級別) クロウン病・女性

クロウン病	女性 総数	0～9 歳	10～19 歳	20～29 歳	30～39 歳	40～49 歳	50～59 歳	60～69 歳	70歳 以上
2003年	6,904	14	320	1,787	1,961	1,072	816	551	383
2004年	7,053	13	327	1,729	2,079	1,125	840	535	405
2005年	7,402	9	352	1,730	2,198	1,236	873	565	439
2006年	7,777	10	373	1,736	2,338	1,306	952	591	471
2007年	8,318	17	441	1,771	2,460	1,439	997	664	529
2008年	8,828	16	459	1,766	2,607	1,603	1,014	775	588
2009年	9,246	20	451	1,820	2,654	1,744	1,058	858	641

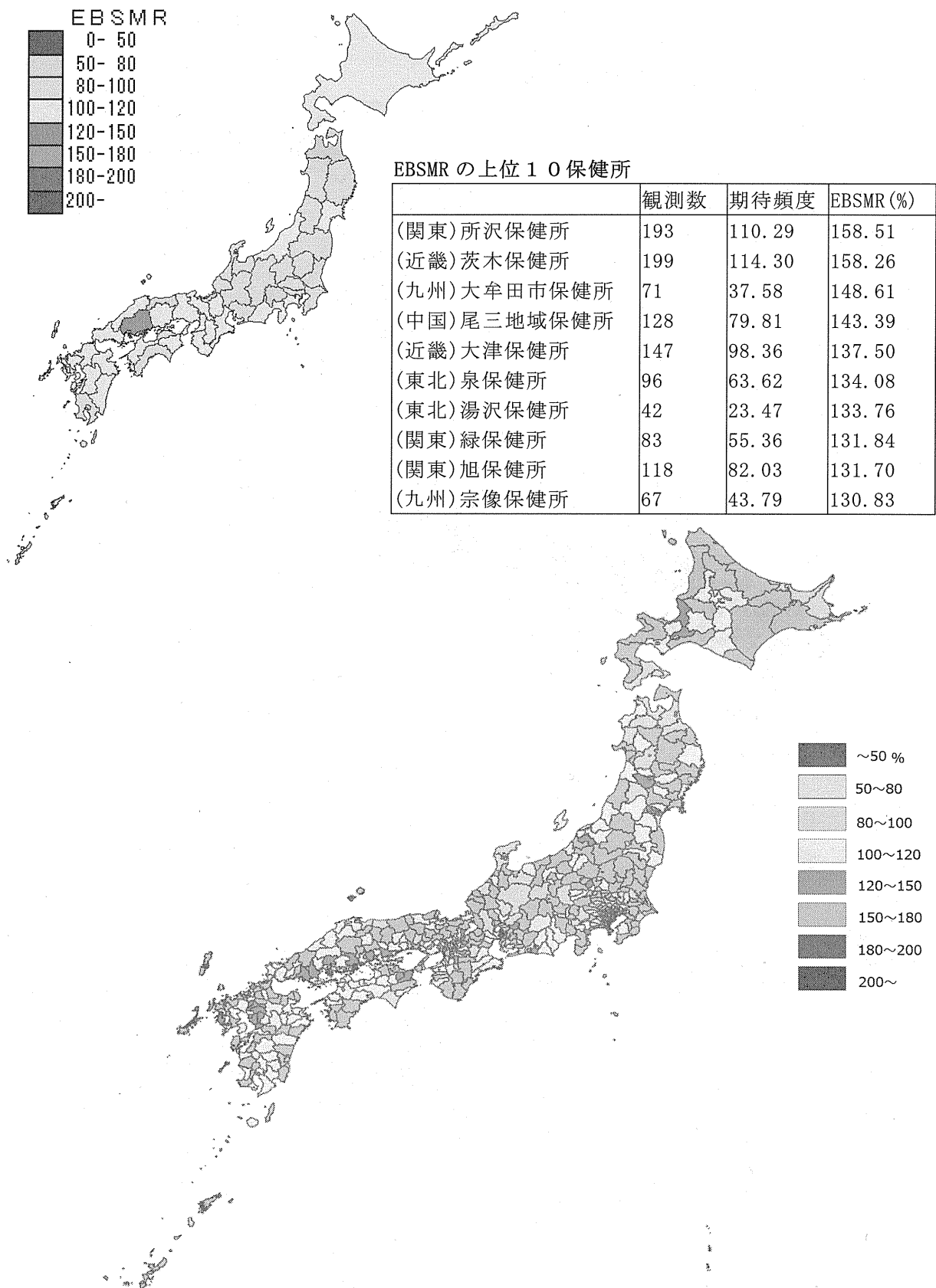
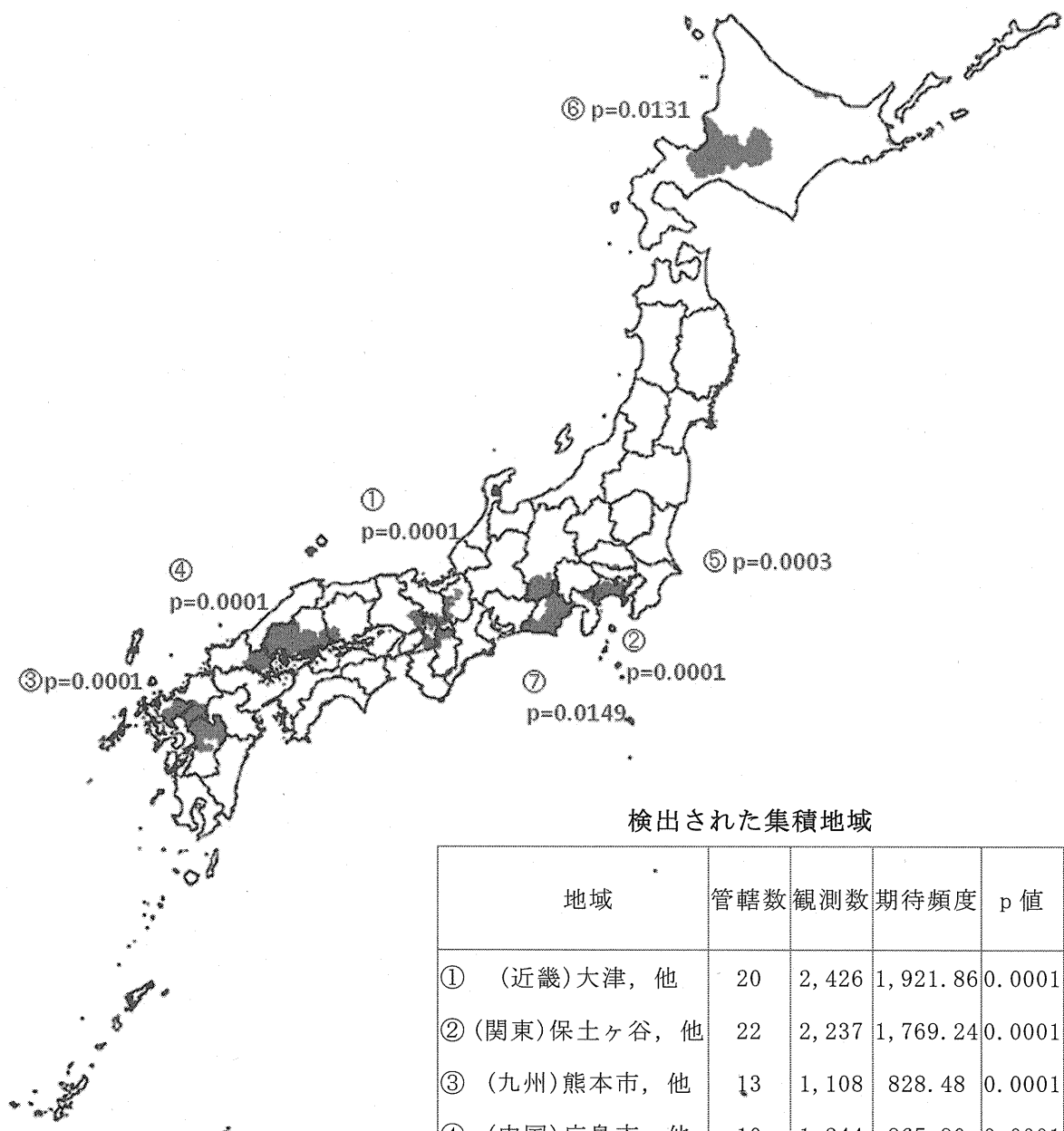


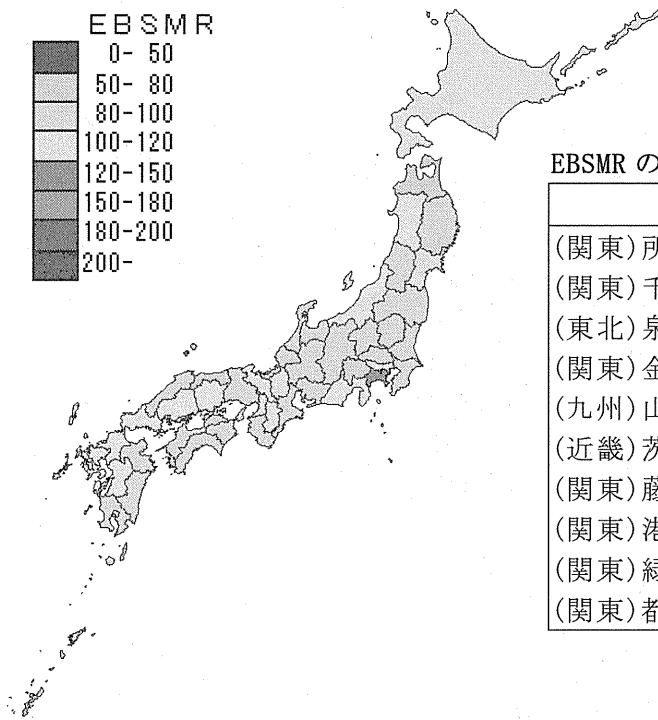
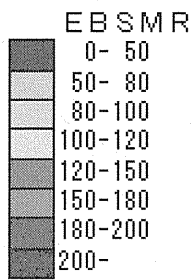
図 1 (a). 潰瘍性大腸炎 (2003 年度 男性)  
都道府県別 EBSMR、保健所管轄別 EBSMR とその上位 10 保健所リスト



検出された集積地域

地域	管轄数	観測数	期待頻度	p 値
① (近畿)大津, 他	20	2,426	1,921.86	0.0001
② (関東)保土ヶ谷, 他	22	2,237	1,769.24	0.0001
③ (九州)熊本市, 他	13	1,108	828.48	0.0001
④ (中国)広島市, 他	10	1,244	965.90	0.0001
⑤ (関東)所沢保健所	1	193	110.29	0.0003
⑥ (北海道)札幌市, 他	5	948	785.00	0.0131
⑦ (東海)浜松市, 他	6	1,014	946.66	0.0149

図 1 (b). 潰瘍性大腸炎 (2003 年度 男性)  
検出された有意な集積地域



EBSMR の上位 10 保健所

	観測数	期待頻度	EBSMR (%)
(関東)所沢保健所	180	97.19	165.84
(関東)千代田保健所	30	12.12	143.41
(東北)泉保健所	97	61.20	139.59
(関東)金沢保健所	96	60.87	138.98
(九州)山鹿保健所	46	25.34	137.82
(近畿)茨木保健所	153	103.33	137.45
(関東)藤沢保健所	166	113.76	136.50
(関東)港南保健所	99	65.34	135.53
(関東)緑保健所	76	48.40	135.47
(関東)都筑保健所	76	48.87	134.64

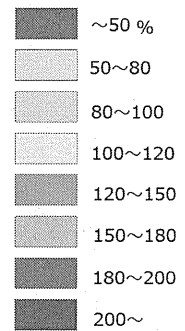
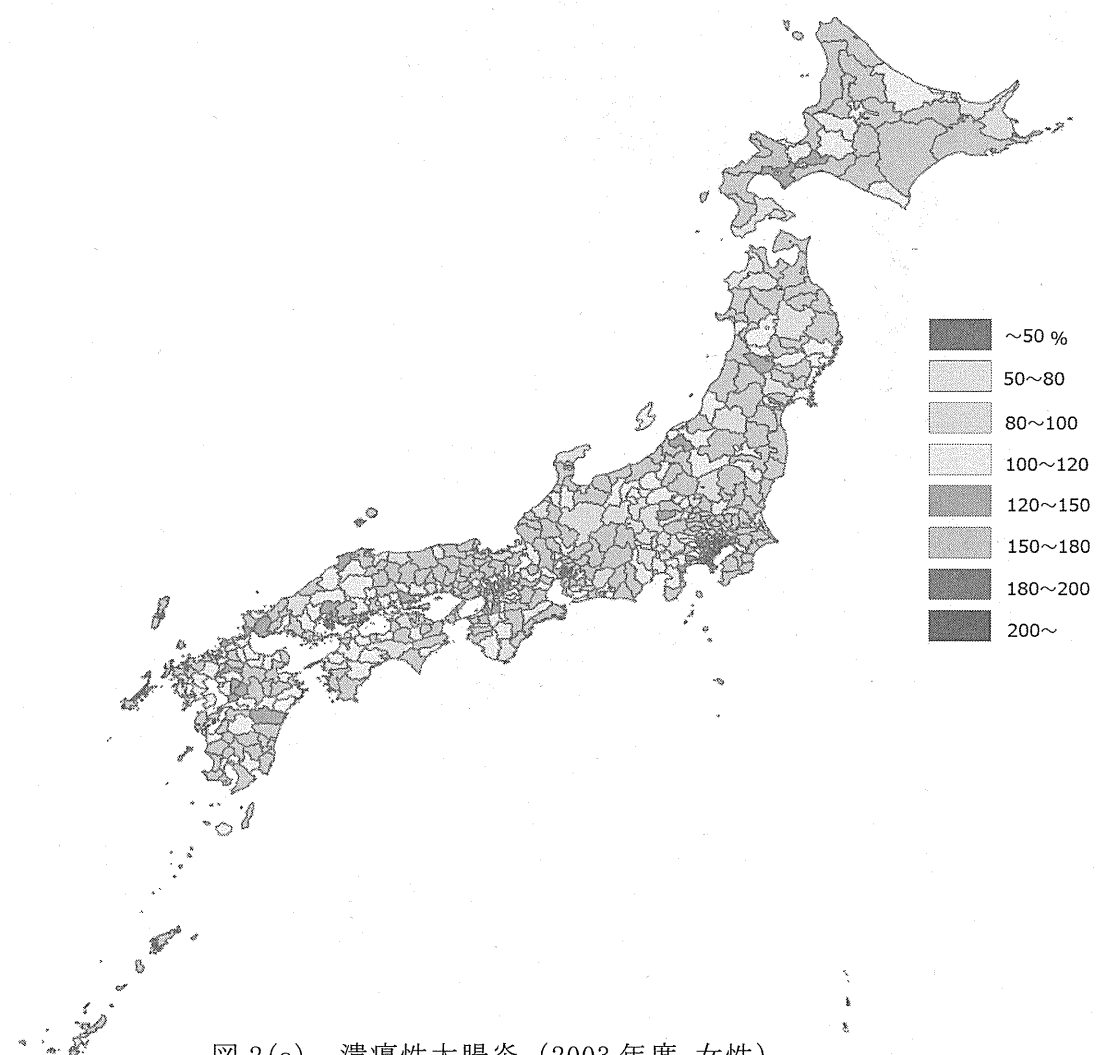


図 2.(a). 潰瘍性大腸炎 (2003 年度 女性)  
都道府県別 EBSMR、保健所管轄別 EBSMR とその上位 10 保健所リスト