

厚生労働科学研究費補助金(医療技術実用化総合研究事業(臨床研究・治験推進研究事業))
分担研究報告書

患者数調査のためのデータベースの構築に関する研究

研究分担者：松村泰志

(大阪大学大学院医学系研究科医学専攻情報統合医学講座医療情報学 教授)

研究要旨

治験ネットワークを構築する上で、それぞれの医療機関が診療している各疾患の患者数が把握できると、各医療機関への担当配分数を割り当てる際に有効である。そこで、レセプトデータを利用して患者数を推定する方法について検討した。レセプトの病名は信頼性が低いことに問題がある。そこで、治療内容と病名から治療対象病名を判定し、これを集計することで各疾患の患者数を調べる方法を考案した。今年度は処方・注射データから病名を推定する方法について検討した。病名を大、中、小の分類コードの組みで表現し(疾患コード)、医薬品に対する適用病名のテーブルを作成した。患者病名も疾患コードで表現し、患者毎に集計して最上位の病名とその治療薬を確定として除き、残りについて同じ処理を繰り返し、病名を拾い出す方法とした。患者数 19,921 人(入院 2,031 人 外来 17,890 人)から 16,060 件の疾患名が推測され、うち小分類まで推測出来た疾患名は 326 疾患、10,866 件、中分類までの推測は 40 疾患、1,423 件、大分類までの推測は 16 疾患、3,771 件であった。次年度は、本法の推定精度の評価法を確立し、更に処理法を精緻化する予定である。本法により、ネットワーク下の各病院が診療する各疾患の患者数が推定されることとなり、治験の担当配分数の割り当てに役立つと思われる。

研究協力者

武田理宏(大阪大学大学院医学系研究科)

A. 研究目的

治験ネットワークを構築する上で、各医療機関が診療する各疾患の患者数が分かると、治験の各医療機関への担当配分数を決める際に有効である。この目的のためには、各医療機関から比較的容易に得られるデータの範囲で実施する必要がある。現状では、レセプトオンライン化が制度化され、レセプトデータ、DPC データはコード、形式が標準化されており、かつ、各病院から容易にデータを出すことができる。

レセプトデータは、病名が含まれている

が信頼性に乏しいことが問題である。また、検査結果等の診断の根拠となるデータが含まれていない。しかし、投薬内容や処置、手術等の治療内容については高い精度で含まれている。DPC データは病名を含め信頼性が高いが、外来患者に適用できない問題がある。こうした特性、限界を理解することで、患者数の概数を求めるシステムを構築することができる可能性がある。

本研究では、こうした考えのシステムの実用性を調べるために、大阪大学医学部附属病院のレセプトデータ、DPC データを取得し、解析しやすいデータベースに展開して、患者数の概数を求めるシステムを開発する。本システムが期待通りに機能するこ

とが確認できた場合は、各病院に展開し、統計データをネットワーク事務局に報告するシステムに拡張する予定である。

B. 研究方法

1. ICD-10 コードの疾患コードへの変換

ICD-10 コードは疾患の上位下位概念が明確でない。そこで大分類、中分類、小分類をもつ疾患コードを作成し、ICD-10 コード(3 ケタ)を「疾患コード」に変換した(表1)。病名の推定にあたっては、小分類までの推定を目標とするが、推定が出来ない場合は、中分類、大分類までの推定を行うものとした。

2. 「医薬品と対応病名データ」による候補病名の抽出

レセプトに記載される薬剤については、一般財団法人日本医薬情報センター(JAPIC)が用意する「医薬品と対応病名データ」を活用した。「医薬品と対応病名データ」は、医療用医薬品添付文書の効能効果とこれに対応する ICD10 対応標準病名マスター(MEDIS 標準病名マスター)を関連付けしたもので、使用した薬剤とレセプトに記載される病名の妥当性を評価するために使用されている。「医薬品と対応病名データ」には評価結果が、1: 標準病名として妥当と判断したもの、3: 標準病名として妥当性に判断を要するもの、5: 標準病名として妥当性を欠くもの、の三段階で表記されている(表2)。本研究では「医薬品と対応病名データ」で変換される ICD-10 コードの 3 ケタのみを使用し、「疾患コード」に変換した。

3. 疾患点数による病名の推定

レセプトに記載された薬剤に対し、それぞれ 10 点の点数を与えた。JAPIC の「医薬品と対応病名データ」によって与えられ

る候補疾患(小分類)に対し、「(評価結果:1):(評価結果:3)=2:1」で重みづけをしたうえで、薬剤ごとに与えられた「疾患点数」10 点を配分し、「疾患点数(小分類)」を与えた(表3)。

大分類、中分類が共通する候補疾患については、「疾患点数(小分類)」を加算し、「疾患点数(中分類)」を求めた。また大分類が共通する候補疾患については、「疾患点数(中分類)」を加算し、「疾患点数(大分類)」を求めた。次にある患者に対し複数の薬剤が使用される場合、同一疾患分類ごとに疾患点数を加算し、使用した全薬剤に対する疾患点数(小分類)、疾患点数(中分類)、疾患点数(大分類)を求めた。

レセプトに記載される病名は、いわゆる保険病名、終了を忘れた病名などがあり信頼性は高くないが、実際に患者が罹患している病名が登録されていることも少なくない。このため、病名の推定に対しては、レセプトに記載された病名も活用すべきである。そこで、レセプトに記載される病名を「疾患コード」に変換し、それぞれの疾患コードに対し 4 点の点数を与えた。使用した薬剤に対する疾患点数と、レセプトに記載された病名に対する疾患点数を加算し、患者に対する疾患点数(小分類)、疾患点数(中分類)、疾患点数(大分類)を求めた。

このように使用した薬剤とレセプトに記載された病名に対し与えられた疾患点数が高いことは、実際に患者がその疾患に罹患している可能性が高いことを意味する。逆に疾患点数が低いことは薬剤情報と病名情報だけでは疾患を推定できないことを意味する。そこで疾患採択閾値=5 とし、疾患点数(小分類)で採択閾値以上で最も高い疾患点数となった疾患を採択疾患とした。採択疾患を候補疾患コードに含む薬剤と病名を除外し、再度、疾患点数の計算を行い、採択閾値以上で疾患点数(小分類)が最も

高い疾患を採択疾患とする作業を繰り返した。全ての疾患点数（小分類）が採択閾値未満となった場合、疾患点数（中分類）が採択閾値=5 以上で最も高い疾患を採択疾患とし、全ての疾患点数（中分類）が採択閾値未満となった場合は、疾患点数（大分類）で同様の検討を行った。全ての薬剤に対し、疾患コードが割り振られるか、疾患点数（大分類）が採択閾値より小さい値となるまで、処理を繰り返した。

なお、候補疾患の重みづけ（本研究では「(評価結果:1) : (評価結果:3) = 2:1」、病名に対する疾患点数（本研究では 4 点）、採択閾値（本研究では 5 点）は仮に定めたものであり、得られた確定疾患の信頼度により変更を検証する必要がある。

4. 疾患点数の算出方法

ある患者に対して投与された薬剤 m に対する、候補疾患コード (i, j, k) の重みづけを N_{mijk} とする

薬剤 m に対する候補疾患コード (i, j, k) の疾患点数（小分類） SSP_{mijk} は

$$SSP_{mijk} = 10 \times N_{mijk} / \sum_{i=1} \sum_{j=1} \sum_{k=1} N_{mijk}$$

と表わされる。

全ての薬剤に対する候補疾患コード $(i, j,$

$k)$ の疾患点数（小分類） SSP_{ijk} は

$$SSP_{ijk} = \sum_{m=1} SSP_{mijk}$$

となる。

レセプトに記載された病名に対する疾患点数（小分類）を SSD_{ijk} とすると、ある患者に対して投与された薬剤とレセプト記載された病名による疾患点数 SS_{ijk} は、

$$SS_{ijk} = SSP_{ijk} + SSD_{ijk}$$

と表わされる。

疾患点数（中分類） SM_{ij} 、疾患点数（大分類） SL_i はそれぞれ、

$$SM_{ij} = \sum_{k=1} SS_{ijk} \quad SL_i = \sum_{j=1} SM_{ij}$$

と求められる。

得られた SS_{ijk} の最大値 $Max-SS_{ijk}$ が採択閾値 T 以上である場合、疾患コード (i, j, k) を採択疾患とする（図 1）。疾患コード (i, j, k) を候補疾患として持つ薬剤、病名を計算から除外して、再度、疾患点数の算出を行う。 $Max-SS_{ijk}$ が採択閾値 T 未満である場合、 $Max-SM_{ij}$ 、 $Max-SL_i$ について同様の処理を行う。全ての薬剤に対し採択病名が割り振られるか、 $Max-SL_i < T$ となった場合に処理を終了する。

表 1. 作成した疾患コードと ICD-10 コード (3 ケタ) の対応の実例

大分類	中分類	小分類	ICD10_3 桁
09 循環器系の疾患	03 高血圧性疾患	01 本態性高血圧	I10
		02 高血圧性心疾患	I11
		03 高血圧性腎疾患	I12
		04 高血圧性心腎疾患	I13
		05 二次性高血圧	I15
	04 虚血性心疾患	01 狭心症	I20
		02 急性心筋梗塞	I21
	
		05 肺性心疾患および肺循環疾患	01 肺塞栓症

表 2. 「医薬品と対応病名データ」の実例

薬剤レセコード	告示名称	ICD10_3 桁	評価結果
610444042	ガスター-D 錠 20mg	E16	1
610444042	ガスター-D 錠 20mg	F54	1
610444042	ガスター-D 錠 20mg	K21	1
610444042	ガスター-D 錠 20mg	K25	1
610444042	ガスター-D 錠 20mg	K26	1
610444042	ガスター-D 錠 20mg	K27	1
610444042	ガスター-D 錠 20mg	K28	1
610444042	ガスター-D 錠 20mg	K29	1
610444042	ガスター-D 錠 20mg	K31	1
610444042	ガスター-D 錠 20mg	K63	3
610444042	ガスター-D 錠 20mg	K92	1

表3. ある薬剤に対する候補疾患と疾患点数配分の例

大分類	中分類	小分類	評価結果	重みづけ	疾患点数 (小分類)	疾患点数 (中分類)	疾患点数 (大分類)
01	01	01	1	2	2.5		
01	01	02	1	2	2.5		
01	01	...				5	
01	02	02	1	2	2.5		
01	02	...				2.5	
01					7.5
02	01	02	3	1	1.25		
02	01	03	3	1	1.25		
02	01	...				2.5	
02					2.5
					10	10	10

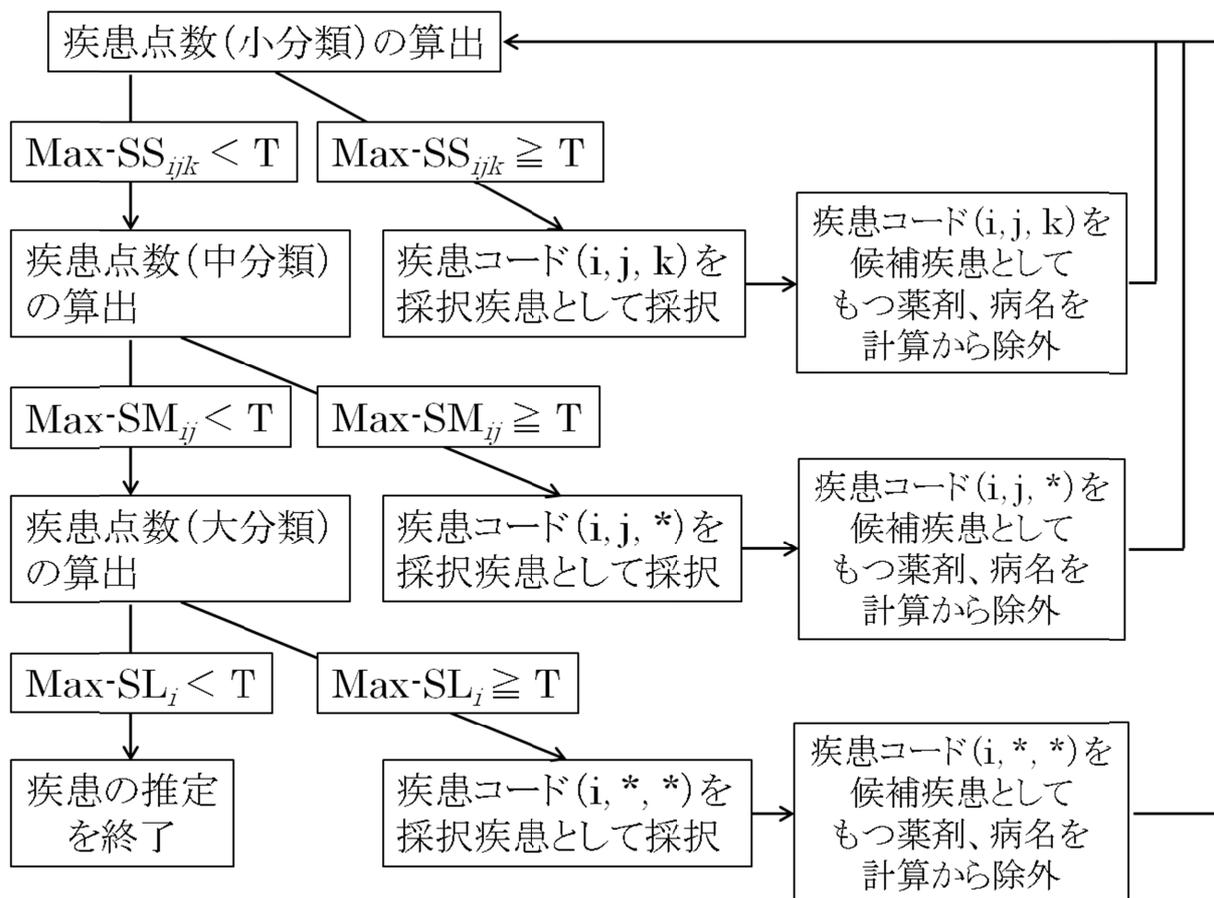


図1. 採択疾患の決定ロジック

C . 研究結果

1. 病名推定の具体例

大阪大学医学部附属病院の実患者 2 名に対して、病名の推定を行った。

1-1. 患者 A

【外来】

- ・「拡張型心筋症」に対して、遮断薬投与中
- ・「心房細動」、「心原性脳梗塞」にてワーファリン投与中
- ・「高度肥満」、「糖尿病」があり、インシュリン治療中
- ・「脂質異常症」に対し、スタチン投与中
- ・下血から貧血が出現した
- ・大腸内視鏡にて「大腸がん」が見つかった

【入院】

- ・消化器外科にて切除術を施行した。
- ・手術に際して、ワーファリンからヘパリン置換を行った。

投与薬剤：

【外来】コニール錠、アテレック錠、ディオバン錠、アーチスト錠(2.5mg)、ジゴキシン錠、バイアスピリン錠、ワルファリンカリウム錠、リピトール錠、ラシックス錠、タケプロン OD 錠、ヒューマリン R 注

【入院】レンドルミン錠、カロナール錠、プルゼニド錠、生理食塩液、ビーフリード輸液、ヘパリンナトリウム注、フェンタニル注射液

レセプトに記載された病名：表 5 に記載

投与薬剤に対して採択された疾患名を表 4 に示す。全投与薬剤 18 剤に対して採択疾患が割り振られた。うち 4 薬剤は中分類までの推定であった。網掛けされた疾患コードは病名として登録されていなかった疾患を表すが、主治医により疾患名としては妥当と判断された。病名として妥当と医師

が判断したもののうち、採択疾患として採択されなかったのは、(03,01,01)鉄欠乏性貧血、(04,08,09)脂質異常症、(09,06,19)心房細動であった(表 5)。鉄欠乏性貧血については前月は鉄剤を投与されており、採択疾患となったと考えられるが、対象月は鉄剤の投与がなかったため採択疾患とならなかった。脂質異常症については(04,02,05)糖尿病として、心房細動については(09,07,04)脳梗塞として疾患採択されたため、採択疾患から漏れた。

1.2. 患者 B

【外来】

- ・「拡張型心筋症」に対して、遮断薬投与中
- ・「慢性関節リウマチ」、「混合性結合組織病(MCTD)」に対してステロイド、リウマトレックス、オレンシア投与中
- ・「肺高血圧症」に対し、トラクリア、ドルナー、アンブラーグ投与中
- ・「鉄欠乏性貧血」に対し、フェロミア投与中
- ・「脂質異常症」に対し、クレストール投与中

投与薬剤：

【外来】アンブラーグ錠、ドルナー錠、フェロミア錠、リウマトレックスカプセル、ラシックス錠、ミカルディス錠、クレストール錠、トラクリア錠、アルダクトン A 錠、メドロール錠、生理食塩水、オレンシア点滴静注用

レセプトに記載された病名：表 7 に記載

投与薬剤に対して採択された疾患名を表 6 に示す。全投与薬剤 12 剤に対して採択疾患が割り振られた。うち 1 薬剤は中分類までの推定であった。網掛けされた疾患コードは病名として登録されていなかった疾患を表すが、生理食塩水に対する疾患名は

妥当と判断されなかった。生理食塩水については、オレンシア点滴静注用の希釈液として使用したもので、そもそも疾患に対する投与でなかったためである。メドロールは非常に多くの候補疾患があったが、(13,02,02)慢性関節リウマチを正しく推測できた。ラシックス錠、ミカルディス錠、アルダクトン錠は(09,06,21)心不全に対して投与されていたが、(09,03,01)高血圧に対する投与と推定された。また、(13,02,02)慢性関節リウマチに点数を取られたため、(13,05,05)全身性硬化症、(13,05,06)その他全身性結合組織疾患が推定されるべき疾患から漏れた。

2. 構築した患者数推定プログラムについて

構築した患者数推定プログラムに対し、2013年12月のレセプト電算、EFデータより評価データを作成し、データ内容の確認を実施した。

2-1. 評価前提条件

ハード：Intel core i3-3240 3.40GHz

Memory 3.38GB (実装 4.0GB)

OS：Windows7pro sp1 x86

処理対象データ

EFファイル

入院-156,296 レコード

外来 169,500 レコード

レセ電算

入院：7,976 レコード

外来：134,140 レコード

処理データ概要

患者数：入院 2,031 人、外来 17,890 人

2-2. 処理結果

EFファイル取込

入力データ：EFファイル、レコード数 325,796

作成レコード：取込済み、入院-156,296

外来-169,500

処理時間：3分20秒×2ファイル(入外)

レセ電ファイル取込

入力データ：レセ電算ファイル レコード数-142,116

作成レコード：取込済み、入院-134,140
外来-7,976

処理時間：18分13秒×4ファイル(国保・社保、入外)

疾病評価ファイル

入力データ：取込済EFファイル レコード数-325,796

作成レコード：入院-6,991 外来-58,779
処理時間：2時間39分48秒

集計表ファイル

入力データ：疾病評価ファイル レコード数-65,770

作成レコード：入院-1,229 外来-13,979
処理時間：15秒

各マスタ

レセ電マスタ、薬剤変換マスタ、ICD変換マスタの取込は各20分程度

2-3 得られた集計表ファイル

患者数 19,921 人(入院 2,031 人 外来 17,890 人)から 16,060 件の疾患名が推測された。うち疾患コードの小分類まで推測出来た疾患名は 326 疾患、10,866 件、中分類までの推測は 40 疾患、1,423 件、大分類までの推測は 16 疾患、3,771 件であった。小分類まで推定された疾患で件数が多かった 20 疾患を表 8 に示す。最も件数が多かったのは(04,02,05)詳細不明の糖尿病で 558 件であった。

疾患コードの小分類まで推定された疾患を含め中分類で最も件数が多かった疾患は(04,02,**)糖尿病で 948 件、ついで(04,08,**)代謝障害が 816 件であり、代謝性疾患が上位を占めた(表 9)

最後に小分類、中分類まで推定された疾患を含め大分類での疾患推定件数を表 10 に示す。最も多かったのは(04,**,**)内分泌、栄養および代謝性疾患で2,087件あり、ついで(11,**,**)消化器系の疾患、が1,952件、(09,**,**)循環器系の疾患が

1,801件と続いた。実際に疾患出現頻度の多い疾患が上位を占めている印象を受ける。一方、(17,**,**)先天奇形、変形および染色体異常、(20,**,**)傷病および死亡の外因の疾患推定件数は0件であった。

表4. 投与薬剤と病名から推定した疾患(患者A)

薬品名称	大分類		中分類		小分類		採択順	点数
	番号	名称	番号	名称	番号	名称		
レンドルミンD錠	6	神経系の疾患	6	挿問性および発作性障害	7	睡眠障害	1	10.00
ブルゼニド錠	11	消化器系の疾患	6	腸のその他の疾患	5	その他の腸の機能障害	1	10.00
ユニール錠	9	循環器系の疾患	3	高血圧性疾患	1	本態性(原発性<一次性>)高血圧(症)	3	8.43
アテレック錠								
ディオバン錠								
ジゴキシン錠								
ラシックス錠								
リビトール錠	4	内分泌、栄養および代謝疾患	2	糖尿病	5	詳細不明の糖尿病	4	7.93
ヒューマリンR注	9	循環器系の疾患	7	脳血管疾患	4	脳梗塞	5	6.86
ワルファリンカリウム錠								
バイアスピリン錠100mg								
ヘパリンナトリウム注								
ワルファリンカリウム錠								
バイアスピリン錠100mg	9	循環器系の疾患	7	脳血管疾患	8	その他の脳血管疾患	5	6.86
ヘパリンナトリウム注	9	循環器系の疾患	6	その他の型の心疾患	21	心不全	7	6.86
アーチスト錠2.5mg								
生理食塩液	4	内分泌、栄養および代謝疾患	8	代謝障害	15	体液量減少(症)	8	5.00
	4	内分泌、栄養および代謝疾患	8	代謝障害	16	その他の体液、電解質および酸塩基平衡障害	8	5.00
タケブロンOD錠	2	新生物	2	消化器	---		10	8.77
カロナール錠	19	損傷、中毒およびその他の外因の影響	20	外科的および内科的ケアの合併症、他に分類されないもの	---		11	6.89
フェンタニル注射液								
ビーフリード輸液	4	内分泌、栄養および代謝疾患	8	代謝障害	---		12	6.86

表5. レセプトに記載された病名と推定した疾患との一致(患者A)

ICD10	大分類		中分類		小分類		病名	病名の推定
	番号	名称	番号	名称	番号	名称		
A53	01	感染症および寄生虫症	05	主として性的伝播様式をとる感染症	04	その他および詳細不明の梅毒		
B18	01	感染症および寄生虫症	12	ウイルス肝炎	04	慢性ウイルス肝炎		
C18	02	新生物	02	消化器	04	結腸の悪性新生物	○	△
C25	02	新生物	02	消化器	11	膵の悪性新生物		
C78	02	新生物	13	部位不明確、続発部位および部位不明の悪性新生物	03	呼吸器および消化器の続発性悪性新生物		
D32	02	新生物	17	良性新生物	23	髄膜の良性新生物		
D50	03	血液および造血器の疾患ならびに免疫機構の障害	01	栄養性貧血	01	鉄欠乏性貧血	○	
E11	04	内分泌、栄養および代謝疾患	02	糖尿病	02	インスリン非依存性糖尿病<NIDDM>		
E14	04	内分泌、栄養および代謝疾患	02	糖尿病	05	詳細不明の糖尿病	○	○
E66	04	内分泌、栄養および代謝疾患	07	肥満(症)およびその他の過栄養<過剰摂食>	02	肥満(症)	○	
E78	04	内分泌、栄養および代謝疾患	08	代謝障害	09	リポたんぱく<蛋白>代謝障害およびその他の脂(質)血症	○	
H35	07	眼および付属器の疾患	05	脈絡膜および網膜の障害	06	その他の網膜障害		
H52	07	眼および付属器の疾患	09	眼筋、眼球運動、調節および屈折の障害	04	屈折および調節の障害		
I10	09	循環器系の疾患	03	高血圧性疾患	01	本態性(原発性<一次性>)高血圧(症)	○	○
I48	09	循環器系の疾患	06	その他の型の心疾患	19	心房細動および粗動	○	
I50	09	循環器系の疾患	06	その他の型の心疾患	21	心不全	○	○
I63	09	循環器系の疾患	07	脳血管疾患	04	脳梗塞	○	○
I67	09	循環器系の疾患	07	脳血管疾患	08	その他の脳血管疾患		○
I82	09	循環器系の疾患	09	静脈リンパ管およびリンパ節の疾患、他に分類されないもの	03	その他の静脈の塞栓症および血栓症		
I84	09	循環器系の疾患	09	静脈リンパ管およびリンパ節の疾患、他に分類されないもの	05	痔核		
K21	11	消化器系の疾患	02	食道、胃および十二指腸の疾患	02	胃食道逆流症		
K74	11	消化器系の疾患	08	肝疾患	05	肝線維症および肝硬変		
L90	12	皮膚および皮下組織の疾患	06	皮膚および皮下組織のその他の障害	11	皮膚の萎縮性障害		
N18	14	泌尿器系の疾患	03	腎不全	02	慢性腎不全		
Z22	21	健康状態に影響をおよぼす要因および保健サービスの利用	02	伝染病に関連する健康障害をきたす恐れのある者	03	感染症のキャリア<病原体保有者>		

表 6. 投与薬剤と病名から推定した疾患 (患者 B)

薬品名称	大分類		中分類		小分類		採択順	点数
	番号	名称	番号	名称	番号	名称		
リウマトレックスカプセル メドロール錠 オレンシア点滴静注用 ドルナー錠 トラクリア錠	13	筋骨格系および結合組織の疾患	2	炎症性多発性関節障害	2	その他の慢性関節リウマチ	1	19.1
ドルナー錠 トラクリア錠	9	循環器系の疾患	5	肺性心疾患および肺循環疾患	2	その他の肺性心疾患	2	11.5
クレストール錠	4	内分泌, 栄養および代謝疾患	2	糖尿病	5	詳細不明の糖尿病	3	6.5
	4	内分泌, 栄養および代謝疾患	8	代謝障害	9	リポたんぱく<蛋白>代謝障害およびその他の脂(質)血症	3	6.5
フェロミア錠	3	血液および造血器の疾患ならびに免疫機構の障害	1	栄養性貧血	1	鉄欠乏性貧血	4	6.0
ラシックス錠 ミカルディス錠 アルダクトンA錠	9	循環器系の疾患	3	高血圧性疾患	1	本態性(原発性<一次性>)高血圧(症)	5 5 5	5.9 5.9 5.9
生理食塩液	4	内分泌, 栄養および代謝疾患	8	代謝障害	15	体液量減少(症)	6	5.0
	4	内分泌, 栄養および代謝疾患	8	代謝障害	16	その他の体液, 電解質および酸塩基平衡障害	6	5.0
アンブラーグ錠	18	症状, 徴候および異常臨床所見・異常検査所見で他に分類されないもの	8	全身症状および徴候	---		7	6.7

表 7. レセプトに記載された病名と推定した疾患との一致 (患者 B)

ICD10	大分類		中分類		小分類		病名	病名の推定
	番号	名称	番号	名称	番号	名称		
B35	01	感染症および寄生虫症	15	真菌症	01	皮膚糸状菌症		
D50	03	血液および造血器の疾患ならびに免疫機構の障害	01	栄養性貧血	01	鉄欠乏性貧血	○	○
E14	04	内分泌, 栄養および代謝疾患	02	糖尿病	05	詳細不明の糖尿病		
E78	04	内分泌, 栄養および代謝疾患	08	代謝障害	09	リポたんぱく<蛋白>代謝障害およびその他の脂(質)血症	○	○
I10	09	循環器系の疾患	03	高血圧性疾患	01	本態性(原発性<一次性>)高血圧(症)		○
I27	09	循環器系の疾患	05	肺性心疾患および肺循環疾患	02	その他の肺性心疾患	○	○
I50	09	循環器系の疾患	06	その他の型の心疾患	21	心不全	○	
I70	09	循環器系の疾患	08	動脈, 細動脈および毛細血管の疾患	01	アテローム<じゅく><粥>状<硬化>(症)		
I73	09	循環器系の疾患	08	動脈, 細動脈および毛細血管の疾患	04	その他の末梢血管疾患		
J00	10	呼吸器系の疾患	01	急性上気道感染症	01	急性鼻咽頭炎[かぜ]		
J84	10	呼吸器系の疾患	07	主として間質を傷害するその他の呼吸器疾患	04	その他の間質性肺疾患		
K21	11	消化器系の疾患	02	食道, 胃および十二指腸の疾患	02	胃食道逆流症		
K85	11	消化器系の疾患	09	胆のう<嚢>, 胆管および膵の障害	05	急性膵炎		
K86	11	消化器系の疾患	09	胆のう<嚢>, 胆管および膵の障害	06	その他の膵疾患		
L02	12	皮膚および皮下組織の疾患	01	皮膚および皮下組織の感染症	03	皮膚膿瘍, せつ<フルンゲル>およびよう<カルブングル>		
M06	13	筋骨格系および結合組織の疾患	02	炎症性多発性関節障害	02	その他の慢性関節リウマチ	○	○
M34	13	筋骨格系および結合組織の疾患	05	全身性結合組織障害	05	全身性硬化症	○	
M35	13	筋骨格系および結合組織の疾患	05	全身性結合組織障害	06	その他の全身性結合組織疾患	○	
M81	13	筋骨格系および結合組織の疾患	12	骨障害および軟骨障害	02	骨粗しょう<鬆>症<オステオポロシス>, 病的骨折を伴わないもの		

表 8. 疾患コードの小分類で推定された疾患とその件数（上位 20 疾患）

	大分類	中分類	小分類	疾患名	件数
1	04	02	05	詳細不明の糖尿病	558
2	04	08	09	リポたんぱく<蛋白>代謝障害 およびその他の脂(質)血症	526
3	06	06	07	睡眠障害	470
4	11	06	05	その他の腸の機能障害	370
5	13	12	02	骨粗しょう<鬆>症<オステオポロシス>, 病的骨折を伴わないもの	327
6	07	06	01	緑内障	315
7	13	02	02	その他の慢性関節リウマチ	307
8	09	03	01	本態性(原発性<一次性>)高血圧(症)	270
9	10	04	01	血管運動性鼻炎およびアレルギー性鼻炎 <鼻アレルギー>	268
10	04	02	02	インスリン非依存性糖尿病<NIDDM>	217
11	13	02	04	若年性関節炎	212
12	06	06	01	てんかん	199
13	02	07	01	乳房の悪性新生物	195
14	09	06	21	心不全	175
15	04	08	10	プリンおよびピリミジン代謝障害	166
16	06	07	13	その他の多発(性)ニューロパチ<シ>ー	165
17	01	01	10	その他の明示された腸管感染症	162
18	07	11	04	眼および付属器の処置後障害, 他に分類されないもの	158
19	14	07	01	前立腺肥大(症)	155
20	06	04	01	アルツハイマー<Alzheimer>病	154

表9. 疾患コードの中分類で推定された疾患とその件数（上位10疾患）

	大分類	中分類	小分類	疾患名	件数
1	04	02	**	糖尿病	948
2	04	08	**	代謝障害	816
3	11	02	**	インフルエンザおよび肺炎	712
4	06	06	**	挿間性および発作性障害	688
5	13	02	**	炎症性多発性関節障害	658
6	13	12	**	骨障害および軟骨障害	435
7	11	06	**	腸のその他の疾患	424
8	09	03	**	高血圧性疾患	350
9	12	03	**	皮膚炎および湿疹	346
10	07	06	**	緑内障	315

表10. 疾患コードの大分類で推定された疾患とその件数

	大分類	中分類	小分類	疾患名	件数
1	04	**	**	内分泌,栄養および代謝疾患	2087
2	11	**	**	消化器系の疾患	1952
3	09	**	**	循環器系の疾患	1801
4	06	**	**	神経系の疾患	1470
5	13	**	**	筋骨格系および結合組織の疾患	1461
6	05	**	**	精神および行動の障害	1352
7	01	**	**	感染症および寄生虫症	962
8	07	**	**	眼および付属器の疾患	909
9	02	**	**	新生物	762
10	12	**	**	皮膚および皮下組織の疾患	761
11	18	**	**	症状,徴候および異常臨床所見・異常検査所見 で他に分類されないもの	725
12	10	**	**	呼吸器系の疾患	698
13	14	**	**	尿路性器系の疾患	407
14	19	**	**	損傷,中毒およびその他の外因の影響	332
15	03	**	**	血液および造血器の疾患ならびに 免疫機構の障害	241
16	08	**	**	耳および乳様突起の疾患	74
17	15	**	**	妊娠,分娩および産じょく<褥>	34
18	16	**	**	周産期に発生した病態	25
19	21	**	**	健康状態に影響をおよぼす要因および 保健サービスの利用	7
20	17	**	**	先天奇形,変形および染色体異常	0
20	20	**	**	傷病および死亡の外因	0

D. 考察

臨床試験を行う場合、被験者を確保することが成否の鍵を握る。日本の場合、ある疾患を特定の医療機関に集中させるようなコントロールはされていない。このため、個々の医療機関では、十分な数の被験者が得られないことが多い。そこで、治験ネットワークを構築し、複数病院で仮想的に大規模病院を構成し、それぞれの臨床試験に対して十分な数の被験者を確保する体制作りが求められている。

治験ネットワークを構築するためには、様々な課題を克服する必要があるが、ネットワーク事務局のコントロールは重要な鍵を握る。ネットワーク事務局は、治験の契約に際し、ネットワーク下にある複数医療機関でどれだけ症例数を受け入れるかを判断しなければならない。しかし、ネットワーク事務局が、ネットワーク下のそれぞれの医療機関が、どの疾患の患者をどれぐらいの数診療しているかを直感的に把握することは不可能である。

今日では、それぞれの医療機関に病院情報システムが導入されている。電子カルテを導入する医療機関も増えており、あらゆる情報が検索できる状況となりつつある。しかし、このネットワーク事務局のニーズを満たすためには、複数の医療機関で統一的方法で情報検索できなければ意味をなさない。どの医療機関からでも統一された形で出力されるデータは、レセプトデータ以外にない。レセプトオンライン化が推進された結果、レセプトデータは電子化されただけでなく、記載コードもレセ電コードで統一されている。

レセプトデータは、ここに記載されている病名に信憑性が低いことが問題とされている。既に治癒した病名が終了されていなかったり、疑いであるべき病名に疑いの表記が抜けていたりする。しかし、病名の記載もれは査低減となるため、それなりに熱

心に病名が登録されていることも事実である。そこで、本研究では、レセプトに記載されている治療内容に着目し、治療内容から、治療対象病名を推定する方法を検討した。治療されている疾患はアクティブ病名であり、疑い病名でもないと言える。

病名を推定する際に、病名概念の粒度を考慮することが重要である。ある患者の病名について、医師 A は細かな粒度の病名を付け、医師 B はやや粗い粒度の病名を付けたとする。例えば、医師 A は十二指腸潰瘍と診断し、医師 B は同じ患者を消化性潰瘍と診断するなどである。二人が付けた病名は異なっているが、本質的には一致している。これを解決するためには、個々の病名登録に対し上位概念の病名も同時に登録されていると捉える必要がある。この例では、十二指腸潰瘍の病名は、消化性潰瘍でもあり、胃疾患でもある。このようにすれば医師 A と医師 B は基本的には同じ病名を想定していたことが判定できる。

本法では、病名の階層概念をデータ化し、各治療薬について、同一レイヤーの治療対象疾患数で按分する方法をとっている。例えば、抗生剤はあらゆる感染性の疾患が治療対象疾患となるので、このレイヤーでの個々の疾患に対する配点は低くなる。しかし、全てが感染症であるので、感染性疾患の大分類粒度のレイヤーでは、感染性疾患が高得点となる。

この処理方法で特定の疾患が高得点になるのは、治療薬が特定の疾患にしか使われない場合か、複数の薬剤が同一疾患の治療で使われている場合である。

本法は、医師が付けた病名に対して、治療がされている場合は、その病名を治療対象病名として拾い出すことを方針としている。そこで、医師が付けた病名に対しても配点し、病名だけでは基準値を超えず、その疾患に対する治療薬がある場合に、点数が加算され、その最大値をとる病名が最終

的に拾い出す仕組みとなっている。

今回は、単月での処理を行ったが、当該月にたまたま治療が無く落ちてしまった病名もある。今後は、複数月を処理して、その中で病名を推定する方法に改良するべきと考える。

本法で採用した配点基準や採択の基準値は恣意的なものであり、特に根拠があるものではない。この点数を変えることで、おそらく結果が変わってくると思われる。最適となる点数を検討する必要があるが、そのためには、結果の精度を測る必要がある。これらは次年度の課題とする。

今年度は、医薬品を対象として処理をした。手術、処置、検査については、その内容によっては、対象疾患をある程度絞れるものがある。これらを組み入れことで、更に精度を上げることができるかもしれない。これについても次年度の課題とする。

本研究では、病院で診療する患者数を把握することを目的としている。複数の病院を同じ方法で処理することにより、複数病院を比較できることが重要であると考えている。ある臨床試験で、ある疾患を持つ患者が被験者の条件であったとする。本法で得たデータから、その被験者が何人いるかを推定することはできないが、例えば、治験ネットワークの複数病院で 30 例を受け持つ必要があった場合に、それぞれの病院の受け持ち数を、本法で得られた各病院のその病名を持つ患者数で按分して配分することには、それなりの合理性がある。

大阪大学医学部附属病院では電子カルテシステムの全データを含むデータウェアハウスが稼働している。このデータウェアハウスを利用することで、被験者候補をより精度高く検索することができる。そこで、本法で得られた大阪大学での対象病名を持つ患者数と、データウェアハウスで検索した被験者候補の患者数の比を求め、この比を掛けることで、治験ネットワーク下の各病

院の被験者候補の数を推定することができる。同じ疾患であっても、病院によって扱う重症度等が異なるなど、こうして求めた推定値がどれだけ正しいものであるかは分からないが、それなりの根拠のある値が得られることで、それが無い状態よりもはるかに適切に配分数を決めることができるはずである。

E . 結論

レセプトデータの投薬内容と病名から治療対象病名を拾い出し、リストアップするシステムを開発した。サンプル症例からある程度の精度で治療対象病名が拾い出せていると思われる。大阪大学医学附属病院の 1 月分のデータを処理したところ、患者数 19,921 人(入院 2,031 人 外来 17,890 人)から 16,060 件の疾患名が推測され、うち小分類まで推測出来た疾患名は 326 疾患、10,866 件、中分類までの推測は 40 疾患、1,423 件、大分類までの推測は 16 疾患、3,771 件であった。

次年度以後、本法の推定の精度を評価し、パラメータを適切に調整すること、他の治療を含めることなど、処理を更に精緻化する予定である。本法により、病院が診療している各疾患の患者数が推定され、臨床試験を進める上で参考となる指標が得られる可能性がある。

F . 健康危険情報

なし

G . 研究発表

- 1 . 論文発表
なし
- 2 . 学会発表
なし

H . 知的財産権の出願・登録状況

なし

