

厚生労働科学研究費補助金（循環器疾患・糖尿病等生活習慣病対策総合研究事業）
分担研究報告書

業務データベースを活用した1型糖尿病症例の抽出研究

分担研究者	中島 直樹	九州大学病院メディカル・インフォメーションセンター
研究協力者	山下 貴範	九州大学病院メディカル・インフォメーションセンター
	伊豆倉理恵子	九州大学病院メディカル・インフォメーションセンター
	野尻 千夏	九州大学病院メディカル・インフォメーションセンター
	南 昌江	南昌江内科クリニック
	岡田 朗	岡田内科クリニック
	桶田 俊光	赤坂おけだ糖尿病内.

研究要旨

平成 26 年度の本研究では、九州大学病院の診療業務用データベースを疫学的目的に 2 次利用し、1 型糖尿病推定症例を検出する初期抽出ロジックを開発した。平成 27 年度には、まず、初期抽出ロジックによる抽出結果から、専門医によりロジックを修正した（平成 27 年度暫定抽出ロジック）。次に、専門医により提出された、間違いなく 1 型糖尿病であることを確認した症例リストを入手した。さらに、抽出ロジックにより 2009~2014 年までの 6 年間の症例 DB から抽出された症例、および疑わしいが最終的に抽出されなかった症例など計 866 症例を対象に専門医によるカルテレビューを行い、真の 1 型糖尿病症例、非 1 型糖尿病症例などを確認し、PPV・感度などで表される抽出ロジックの精緻化を試み、平成 27 年度最終抽出ロジック（PPV80.9%、感度 87.0%）とした。このロジックをレセプト項目のみに限定して再抽出し精緻化の修正を試みたところ、平成 27 年度最終レセプト抽出ロジックとして、PPV は 79.9%、感度は 84.8% が得られた。

平成 28 年度以降は、1) 1 型糖尿病でインスリン依存を有する症例の抽出ロジックを作成し、2) これらをさらに精緻化する。これらは NDB などに対して用いることにより 1 型糖尿病（あるいはインスリン依存状態にある 1 型糖尿病）の有病率を測定し、その病態を明らかにすることに有用と考えられる。

A. 研究背景と目的

1 型糖尿病は成人後に比して小児期に発症する頻度が高い疾患であるが、1921 年にバンティングとベストによってインスリンが発見されて以来、発症急性期に死亡することは急激に減少した。つまり現在では、適正なインスリン治療を継続することにより天寿を全うし得る疾患になったのである。

また、近年は成人後発症の 1 型糖尿病症例の存在も広く認められるに至っている。一方で、一生の間インスリン注射を持続することや、罹病が長期にわたるためにどうしても糖尿病合併症の発症が多いことを含めると、健常人に比べて大きな経済的・心理的・社会的な負担となることは容易に推測できる。例えば、インスリン治療を受ける

人の医療費の自己負担分は月額 1.0～1.5 万円といわれている。

小児慢性特定疾患である 1 型糖尿病は、20 歳未満では自己負担分が小児慢性特定疾患治療研究事業により補助されるが、20 歳に達すると同時に補助はなくなる。成人した後の医療費の過度の自己負担や合併症の発症などにより、治療の継続が困難な状況に陥る症例も多いと思われる一方で、これまでに成人以降の 1 型糖尿病の有病率や地域分布、合併症の状況などは方法論が無いために算出できておらず、その実態は不明である。2015 年に、「難病の患者に対する医療等に関する法律」が改正され、指定難病は従来の 56 疾病から 306 疾病へ増加したにも関わらず、1 型糖尿病が指定難病となることは、その不明が故に難しい状況である。

近年、電子化率が 100% に近づいたレセプトデータに加えて、急激に実装されつつある病院情報システムに蓄積した診療データや保険者による特定健診データなどを利用した、データベース疫学とでもいうべき新たな領域が開かれつつある¹⁾。つまり、保健医療（診療）業務で蓄積したデータを 2 次利用して疫学的な目的に利用するものである。例えば厚生労働省と PMDA が推進する「医療情報データベース基盤整備事業（MID-NET 事業）」などがその代表例である。同事業は 10 の協力病院グループの標準的なデータベースから、ある薬剤を投与された症例群にどのような副作用が発生したか、などを抽出口ジックと呼ばれる計算式を用いて算出する事により、薬剤の副作用を検知する。抽出口ジックで Data Base (DB) から導かれた症例数（この場合は副作用

発症者数）が必ずしも 100% 正しいわけではないが、専門医などによるカルテレビューにより陽性的中率（Positive Predictive Value : PPV、抽出口ジックにより副作用発生とされた症例の中で、真にその副作用の疾患/症状が発生した率）や感度（真に副作用の疾患/症状が発生した症例の中で、抽出口ジックによって副作用が発生したとみなされた症例の率）を明らかにし、その PPV・感度の前提の下で副作用の検知を試みるものである。このような手法は、薬剤疫学のみならず、データベースを活用した様々な疫学に展開が可能である^{1,2)}。

本研究では、平成 26 年度には暫定的 1 型糖尿病抽出口ジック作成した。平成 27 年度にはそれをより精緻化した上で、その修正抽出口ジックを検証する目的で九州大学病院の電子カルテを糖尿病専門医の協力を得て、PPV・感度を測定した。さらに、この成果を九大病院のデータベースに留めずに、今後は 1 型糖尿病の有病率測定に用いることを目的として以下を検討した。

MID-NET 事業とともに厚生労働省が主導するもう一つの医療 Big Data 解析に National Data Base (NDB) 事業がある³⁾。日本で発生する全てのレセプトと特定健診結果を匿名化して収集し、公益目的に解析するものである。そこで、平成 27 年度には病院情報システムからの最適な PPV・感度を持つ抽出口ジックを把握した上で、その抽出口ジックに用いたデータ項目をレセプト上の項目に限定した場合の PPV・感度も試算する。この抽出口ジック、PPV・感度は、来年度以降に、NDB で 1 型糖尿病有病率の全国における地域、年齢分布調査など

を実施する際の算出根拠の基礎となることを期待する。

B. 研究方法

1. 1型糖尿病の抽出口ジックのカルテレビュー前修正

平成 26 年度から平成 27 年度前半の本研究会議の議論を経て、平成 26 年度に開発した暫定抽出口ジックを以下のように暫定的に修正した。

・除外薬剤には、SU 剤のみならず、DPP-4 阻害剤、グリニド剤を含める。

・1型糖尿病→除外病名（あるいは除外薬剤処方）の順である場合は、除外するが、逆の場合には、除外しない。

・A 集団の条件「a) 主治医による 1 型糖尿病またはインスリン依存型糖尿病の診断名があること」は、「主治医」の定義が曖昧なため、「a) 1 型糖尿病またはインスリン依存型糖尿病の病名情報があること」とする。

・A 集団の条件「d) 期間を問わずケトアシドーシスの既往があること」は、調査期間の範囲内で検討する他に方法は無いことから、「d)ケトアシドーシスの既往が病名情報から判明すること」とする。

2. ゴールドスタンダード症例の収集

抽出口ジックの感度を調査するために、複数の糖尿病専門医に、確実に 1 型糖尿病で、かつ九州大学病院に受診歴があると考えられる症例（ゴールドスタンダード症例）のリストの提供を依頼した。

3. カルテレビュー

九州大学病院の 2009 年 1 月から 2014 年 12 月までの 6 年間の病院情報データベースから以下の症例を検索した。

・受診歴のあるもので、病名情報に 1 型糖

尿病関連病名があるもの

・受診歴のあるもので、病名情報に 1 型糖尿病関連病名はないが、条件 a、b、c の全てが当てはまるもの

・受診歴のあるもので、病名情報に 1 型糖尿病関連病名はないが、膵島関連自己抗体が陽性であるもの

・受診歴があるもので、上記には含まれないが、提供されたゴールドスタンダード症例に含まれるもの

これら全てをカルテレビューの対象とした。

電子カルテの参照権限を有する糖尿病専門医 3 名が 2015 年 12 月から 2016 年 2 月にわたり、電子カルテをレビューした。

カルテレビューの際には、以下を予め病院情報システムから症例毎に出力し、レビューの参考とした。

- 1) カルテレビュー前の 1 型糖尿病抽出口ジックでは、どのカテゴリーに入るか
- 2) 1 型糖尿病関連病名の有無とその詳細
- 3) インスリン処方の有無と初回処方日
- 4) 基礎インスリンの処方の有無と、その用法、用量
- 5) ケトアシドーシス病名の有無、病名付与日
- 6) 血中 CPR が 0.6ng/ml 以下となった初回日付、および結果値で最終の値
- 7) 自己抗体の有無
- 8) 除外薬剤である SU 剤、グリニド剤、DPP-4 阻害剤の処方回数、最終処方日
- 9) 除外疾患の有無
- 10) 身長、体重、BMI
- 11) HbA1c (NGSP)、血糖
- 12) 尿ケトン体
- 13) 膵移植病名の有無、膵移植病名開始日

- 14) 成長ホルモン製剤の処方の有無
 15) ステロイド製剤や免疫抑制剤の処方の有無や内容

カルテレビューに際しては、初診時病歴、サマリー類（入院、外来、手術など）は全て参照することとし、その他、症例に応じて、必要な項目を参照することとした。

参照結果は、1型糖尿病の有無（1; 間違いなく違う、2; たぶん違う、3; 判断不能、4; たぶん正しい、5; 間違いなく正しい、の5段階評価）、インスリン依存性の評価（同様の5段階評価）とし、必要に応じてそれぞれの理由を記載した。

4. カルテレビューの結果による抽出口ジックの修正と、レセプト情報のみによる抽出口ジックの作成
 カルテレビューの結果を分析し、PPV・

感度の両方について最も適正と思われる抽出口ジックを導出した。また、1型糖尿病として抽出された症例からインスリン依存性を有する症例の追加抽出口ジックを検討し、PPV・感度を評価した。さらに、1型糖尿病として抽出された症例をレセプト情報のみで抽出した場合のPPV・感度について検討した。

5. 倫理的配慮

本研究は、九州大学医学研究院・観察研究倫理審査委員会で承認された。また、HISデータベースからの抽出に関しては、九州大学病院の情報公開・個人情報保護委員会でも承認された。なお、観察研究として <http://kenkyu.mic.hosp.kyushu-u.ac.jp/mic-kenkyu2/>に公開している。

C. 研究結果

1. 抽出口ジック

平成 27 年度カルテレビュー前の暫定抽出口ジックは以下となった（表 1、2）。

表 1 . 平成 27 年度暫定抽出口ジック（カルテレビュー前）

<p>$C = A - B$</p> <p>$C + D - (C \text{ AND } D) = 1 \text{ 型糖尿病症例推測値}$</p> <p>A、B、C、D は以下により求める。</p> <p>A 集団：</p> <p>以下の a から d において、a AND (b OR c OR d) または以下の a に関わらず、b AND c AND d</p> <p>a) 1 型糖尿病またはインスリン依存型糖尿病の病名情報があること</p> <p>b) インスリン治療を受けていること</p> <p>c) 一度でも空腹時血中 C-peptide 0.6 ng/mL 未満であること</p> <p>d) ケトアシドーシスの既往が病名情報から判明すること</p> <p>B 集団および C 集団：</p>
--

A 集団から B 集団（除外項目、表 2 参照）を除いたものを C 集団とする。但し、1 型糖尿病病名 → 除外病名（あるいは除外薬剤処方）の時系列である場合には、除外するが、逆の場合には、除外しない。

D 集団：

膵島関連自己抗体陽性は単独で、1 型糖尿病の診断根拠とする。従って、1 型糖尿病関連自己抗体である抗 GAD 抗体、抗 IA-2 抗体が陽性であること。ただし、その値は問わない。

表 2 . 平成 27 年度暫定除外項目（カルテレレビュー前）

ア) 2 型糖尿病		
2 型糖尿病	20050020 など	E11
イ) 遺伝子異常による糖尿病		
インスリン遺伝子異常（プロインスリン症含む）		
インスリン異常症	20051142	E168
HNF4α 遺伝子異常（MODY1）	なし	
グルコキナーゼ遺伝子異常（MODY2）	なし	
HNF1α 遺伝子異常（MODY3）	なし	
IPF-1 遺伝子異常（MODY4）	なし	
HNF1β 遺伝子異常（MODY5）	なし	
NeuroDI 遺伝子（MODY6）	なし	
Kir6.2 遺伝子異常（新生児糖尿病）、SUR1 遺伝子異常（新生児糖尿病）		
新生児一過性糖尿病	20065704	P702
新生児糖尿病	20065789	P702
アミリン遺伝子異常	なし	
ミトコンドリア DNA 異常（MIDD）		
MELAS 症候群	20095359	E888
MERRF 症候群	20095361	E888
MNGIE	20095365	E888
ミトコンドリア病	20094357	E888
ミトコンドリア脳筋症	20084786	G713
インスリン受容体遺伝子異常		
-1 インスリン受容体異常症 A 型		
インスリンレセプター異常症	20051138	E13
-2 妖精症（Leprechaunism）		
ドナヒュー症候群	20052440	E348
-3 Rabson-Mendenhall 症候群		
インスリンレセプター異常症	20051138	E13
ウ) その他の遺伝的症候群で糖尿病を伴う事が多いもの		

プラダー・ウィリー症候群	20052978	Q871
ウェルナー症候群	20051268	E348
ウォルフラム症候群	20090166	Q878
セルロプラスミン低下症		
無セルロプラスミン血症	20076755	E830
脂肪萎縮性糖尿病		
先天性脂肪異栄養症	20079720	E881
筋強直性ジストロフィー		
筋強直性ジストロフィー	20059238	G711
フリードライヒ失調症		
フリードライヒ運動失調症	20052895	G111
Bardet-Biedl 症候群		
ローレンス・ムーン症候群	20053756	Q878
エ) 染色体異常を伴う糖尿病		
Down 症候群		
ダウン症候群	20052260	Q909
Turner 症候群		
ターナー症候群モザイク	20052238	Q96
ターナー症候群核型 4 5 X	20052241	Q960
ターナー症候群 4 6 X Y	20052237	Q963
ターナー症候群モザイク 4 5 X	20052239	Q963
ターナー症候群モザイク 4 6 X X	20052240	Q963
X O 症候群	20050408	Q969
ターナー症候群	20052236	Q969
Klinefelter 症候群		
クラインフェルター症候群	20051631	Q984
オ) 妊娠糖尿病		
2 型糖尿病合併妊娠	20083656	O241
妊娠糖尿病	20072937	O244
カ) 除外薬剤内服		
SU 剤内服者、DPP4 阻害剤、グリニド剤		
キ) ステロイド糖尿病	20052128 など	E13
ク) 膵臓摘出後		
膵全摘後二次性糖尿病	20078972	E891
膵性糖尿病	20078965 など	E891

2009~2014 年の 6 年間に九州大学病院に
受診歴のある症例 DB (約 30 万症例) から

上記の平成 27 年度暫定抽出ロジック(カルテレレビュー前)により、抽出された 1 型糖尿病症例数推測値を以下のように求めた。

A 集団は、
 $[a \text{ AND } (b \text{ or } c' \text{ or } d)] + [b \text{ AND } c' \text{ AND } d] - a \text{ AND } [b \text{ AND } c' \text{ AND } d]$
 $= 412 \text{ 人} + 25 \text{ 人} - \text{重複 } 20 \text{ 人} = 417 \text{ 人}$ であった。

次に、B 集団を求めるために、表 2 の除外項目に該当する集団を A 集団から抽出し、C 集団を求めた。

C 集団 = A 集団(417 人) - B 集団(125 人) = 292 人

また、D 集団は、抗 GAD 抗体は 200 人、抗 IA-2 抗体は 92 人で重複例を除くと 253 人となった。すなわち、

1 型糖尿病症例数推測値 = C + D - (C AND D) = 292 人 + 253 人 - 103 人 = 442 人であった。

2. ゴールドスタンダード症例の収集

2 名の糖尿病専門医から、ゴールドスタンダード症例リストを入手した。合計 60 症例であったが、2009~2014 年に九州大学病院に実際に受診歴があった症例はうち 48 名であった。

3. カルテレレビュー

九州大学病院の 2009 年 1 月から 2014 年 12 月までの 6 年間の病院情報データベースの検索により、

- ・受診歴のあるもので、病名情報に 1 型糖尿病関連病名があるもの 760 名

- ・受診歴のあるもので、病名情報に 1 型糖尿病関連病名はないが、条件 a、b、c の全てが当てはまるもの 5 名

- ・受診歴のあるもので、病名情報に 1 型糖尿病関連病名はないが、膵島関連自己抗体が陽性であるもの 99 名

- ・受診歴があるもので、上記には含まれないが、提供されたゴールドスタンダードにある症例 2 名(48 名中 46 名は上記に含まれていた)の計 866 例が抽出された。

この 866 全件についてカルテレレビューを行い、平成 27 年度暫定抽出ロジック(カルテレレビュー前)の PPV を算出した。また、46 名のゴールドスタンダード症例を用いて同ロジックの感度を算出した(表 3)。その結果、1 型糖尿病(E10)病名のみで抽出した場合の PPV は 54.7%、感度は 100%であった。平成 27 年度暫定抽出ロジック(カルテレレビュー前)の PPV は 74.0%、感度は 82.6%であった。

表 3. カルテレレビューの結果による修正抽出ロジック別の PPV、感度

全体の項目で集計	ロジック分類	母数	1 型 DM		GS	ロジック推定数	感度	
			ロジック推定数	1 型 DM 判定				PPV
(1)平成27年抽出ロジック(カルテレレビュー前)		864	442	327	74.0%	46	38	82.6%
(2)自己抗体陽性でも病名・薬剤の除外対象があれば除外		864	313	262	83.7%	46	34	73.9%
(3)病名・薬剤の除外対象であってもCPR陽性(CPR<0.6ng/ml)であれば除外しない		864	471	342	72.6%	46	40	87.0%
(4)膵移植病名があればロジック推定に含める		864	457	341	74.6%	46	38	82.6%
(5) (1)に(2)~(4)を加えた場合		864	387	313	80.9%	46	40	87.0%

さらにその結果から、PPV・感度を向上させるために、抽出ロジックの修正を様々なに試みた。

平成 27 年度暫定抽出ロジック(カルテレビュー前)に「自己抗体陽性でも病名・薬剤の除外対象があれば除外」「病名・薬剤の除外対象であっても血中 CPR 陽性 (CPR <0.6ng /ml) であれば除外しない」を足し合わせた場合の感度、PPV を表 3 に示すが、それぞれ、PPV あるいは感度が向上するものの、PPV と感度の関係はトレードオフであった。また、「膵移植病名があればロジック推定に含める」などを足し合わせた場合には感度は変化せず PPV がわずかに上昇した(表 3)。

さらに、これらを全て合わせた場合には、PPV は 80.9%、感度は 87.0%のままにまで向上した。

平成 27 年度の検討では、このロジックを平成 27 年度最終抽出ロジックとした。なお、

平成 28 年度には機械学習などを用いて網羅的にどの項目が真の 1 型糖尿病の予測や除外に最も寄与しているかを調査することにより、さらに PPV・感度を向上する予定である。

また、平成 27 年度最終抽出ロジックに使用しているデータ項目の中からレセプト項目のみで抽出してみたところ、表 4 に示すように PPV は 87.1%であったが感度は 73.9%であった。それをさらに向上するべく、ロジックを追加して検討したところ、薬剤のみで除外を実施する、を追加したところ、PPV は 79.9%であったが感度は 84.8%であった。PPV は 7.2%減少したものの、感度が 10.9%と向上した。レセプト項目に限定し、情報が減少した分、不正確な病名情報の影響が大きく、過剰に除外されていたためであろう。

表 4. レセプト項目のみで抽出した場合のカルテレビューの結果による修正抽出ロジック別の PPV・感度

ロジック分類	母数	1型DM			GS	ロジック推定数	感度
		ロジック推定数	1型DM判定	PPV			
(7)レセプト項目のみで集計、膵移植はロジック推定に含める かつ、薬剤・疾患で除外実施	760	294	256	87.1%	46	34	73.9%
(8)レセプト項目のみで集計、膵移植はロジック推定に含める かつ、薬剤・疾患での除外を実施しない(ロジック推定数が増える)	760	395	293	74.2%	46	40	87.0%
(9)レセプト項目のみで集計、膵移植はロジック推定に含める かつ、薬剤のみで除外を実施する	760	348	278	79.9%	46	39	84.8%

D. 考察・結論

平成 27 年度最終レセプト抽出ロジックは、病院のレセプト DB 上あるいは保険者 DB あるいは NDB でも使用可能である。つまり、NDB を用いた解析で、一定の精度を保ちながらに日本全体の 1 型糖尿病の有病率、地域分布、年代分布、あるいはどのような糖尿病合併症などを併発しているかが把握できることとなる。平成 28 年度には、このロジックを用いて NDB あるいは保険者データで厚生労働省戦略研究の共同研究を実施している満武巨裕氏とともに 1 型糖尿病症例の抽出を試みる予定である。

また、インスリン依存の有無は、今年度のカルテレビューを実施した 866 名に関して既に評価を終えており、平成 28 年度には 1 型糖尿病でインスリン依存に陥っている症例の抽出ロジックを考案する。これをレセプト項目のみのロジックへ落とし込むことにより、NDB 上で一定の精度で、1 型糖尿病でインスリン依存に陥っている症例数などが抽出され、指定難病への申請も可能となることが期待される。

E. 研究発表

- | | |
|---------|----|
| 1. 論文発表 | なし |
| 2. 学会発表 | なし |

F. 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 |

G. 参考文献

- 1) 伊豆倉理恵子、山下貴範、野尻千夏、野原康伸、安德恭彰、中島直樹、医療情報データベース基盤事業の本格稼働に向けたデータ検証、第 34 回医療情報学連合大会 34th JCMI, 710-3, 2014.11.
- 2) 中島直樹：国家規模の医療情報データベース事業“MID-NET” 医学のあゆみ, 248(12):927-928, 2014.03.
- 3) 松田晋哉：レセプト情報・特定健診情報によるナショナルデータベース—医療情報化に関するタスクフォースの方向性—。プラクティス, 2013; 30: 715-722

