

一般口演 | 医療データ解析

一般口演16

医療データ解析

2019年11月23日(土) 14:40 ~ 16:40 B会場 (国際会議場 2階コンベンションホールB)

[3-B-2-06] MID-NETを用いた脳卒中の検索精度に関する検討

○井上 隆輔¹、大友 千晶¹、中山 雅晴^{1,2} (1. 東北大学病院メディカルITセンター, 2. 東北大学大学院医学系研究科 医学情報学分野)

キーワード : Phenotyping, MID-NET, Stroke

【目的】東北大学病院に導入された医療情報データベース(MID-NET)を用いて、疫学研究の実施に適切な対象者の抽出が可能となるよう抽出条件を設定する。

【方法】脳出血と脳梗塞を対象とし、MID-NETから抽出する条件を設定した。初期条件は、DPC病名があり、頭部CT/MRIを撮影した患者とした。脳出血はくも膜下出血・外傷性出血を含まない。脳梗塞はもやもや病とTIAを含む。該当者は診療録の確認により、真のケースとその他のケースに判定し、陽性的中度(PPV)を算出した。さらに、機械学習によりPPVに影響する要因を検討し、PPVの上昇を目的に改良型アウトカムを作成した。

【結果】脳出血は初期条件により245例抽出され、真のケースは119例、PPVは48.6%であった。その他のケースは、大半が疑い病名や慢性硬膜下血腫であった。機械学習ではAUCは0.94と高く、疑い病名と慢性硬膜下血腫が強く関連した。これらを条件から除いた改良型アウトカムにより、PPVは85%に改善すると見込まれ、感度も保たれた。

脳梗塞は初期条件により990例抽出され、無作為の200例を対象とした。真のケースは63例、PPVは31.5%であった。その他のケースには、疑い病名や脳梗塞に至らないもやもや病が多く含まれた。既往歴としての脳梗塞や陳旧性脳梗塞も含まれた。機械学習ではAUCが0.53であったが、病型や治療法に基づいて病名、検査、治療等を整理し、0.60となった。対象病名を整理し、機械学習の結果を考慮した改良型アウトカムにより、PPVは60%台に改善した。

【結論】脳出血は初期条件のPPVは約50%であったが、対象病名の整理によりPPVの大幅な改善が見込まれた。脳梗塞は初期条件のPPVは低かったが、病名に加え検査や治療等の条件調節によりPPVの改善が認められた。その他の条件の最適化により、よりPPVが改善すると考えられる。

MID-NET を用いた脳卒中の検索精度に関する検討

井上隆輔^{*1}、大友千晶^{*1}、中山雅晴^{*1,2}

*1 東北大学病院メディカル IT センター、*2 東北大学大学院医学系研究科医学情報学分野

Investigation on Retrieval Precision of Stroke with MID-NET

Ryusuke Inoue^{*1}, Chiaki Otomo^{*1}, Masaharu Nakayama^{*1,2}

*1 Medical Informatics Center, Tohoku University Hospital,

*2 Department of Medical Informatics, Tohoku University School of Medicine

[Objective] Using the medical information database (MID-NET) introduced at Tohoku University Hospital, we set extraction conditions so that appropriate subjects can be extracted for conducting drug epidemiological studies.

[Methods] The conditions for extraction from MID-NET were set for cerebral hemorrhage and cerebral infarction. The initial condition was a patient with disease name of DPC and cerebral CT / MRI. Cerebral hemorrhage does not include subarachnoid or traumatic hemorrhage. Cerebral infarction includes moyamoya disease and TIA. Applicable subjects were judged to be true cases and other cases by checking medical records, and positive predictive values (PPV) were calculated. In addition, we investigated the factors affecting PPV by machine learning, and created an improved outcome for the purpose of increasing PPV.

[Results] 245 cases of cerebral hemorrhage were extracted according to initial conditions. Of 245 cases, 119 were true, and PPV was 48.6%. The other cases were mostly cases with suspected disease name or chronic subdural hemorrhage. In machine learning, AUC was as high as 0.94, and suspected disease name and chronic subdural hemorrhage were strongly related. When excluding these conditions, the PPV was expected to improve to 85% and the sensitivity was maintained.

990 cases of cerebral infarction were extracted according to the initial conditions, and 200 cases were randomly sampled. There were 63 true cases and PPV was 31.5%. Other cases included suspected disease names and moyamoya disease that did not lead to cerebral infarction. Past cerebral infarction and old cerebral infarction were also included. In machine learning, AUC was 0.53. However, when the disease names, examinations, treatments, etc. were arranged based on disease type and treatment method, AUC became 0.602. PPV improved to the 60% level by improving the outcomes by organizing the target disease names and considering the results of machine learning.

[Conclusion] Although cerebral hemorrhage had an initial PPV of approximately 50%, significant improvement in PPV was expected by organizing the target disease names. In the case of cerebral infarction, PPV of the initial condition was low, but improvement of PPV was observed by adjusting the conditions of examinations and treatments in addition to the disease names. By optimizing other conditions, PPV is expected to improve.

Keywords: Phenotyping, MID-NET, Stroke.

1. はじめに

医薬品の製造販売後の安全対策は従来、対照群を設定しない使用成績調査等評価することで行われてきた。しかし、この方法にはいくつかの課題が知られており、より科学的な根拠に基づく安全対策を実施すべきとの意見もあることや、国際的にも医薬品の安全対策においてデータベースの活用が行われている事を踏まえ、本邦でも医療情報を二次利用可能な環境が整備されつつある。

MID-NET[1]は厚生労働省が主導して整備した医療情報のデータベースであり、レセプト及びDPCの情報に加えて、検査値も利用可能な点が特色である。平成30年度からは製薬企業等もMID-NETを利活用可能となり、利活用が拡大し

つつある。東北大学病院は、このMID-NET事業の協力医療機関として指定されている。

データベースを利用する場合、データベースに含まれる情報を組み合わせて有害事象(アウトカム)を定義する必要がある。データベースを利用した薬剤疫学研究を実施する場合には、より適切なアウトカム定義を作成することが、調査全体の質に影響する。本邦においてアウトカム定義のバリデーションはほとんど行われておらず、本研究では、MID-NETを用いて、アウトカム定義の作成及びその妥当性の評価を効率的に実施する方法について検討した。具体的には、アウトカムとして心不全を定義し、アウトカム定義の妥当性の指標として陽性的中度(PPV: positive predictive value)を算出することを目的として実施した。

2. 方法

本研究では、脳卒中(脳出血、脳梗塞)を対象に、臨床ガイドライン[2]を参考に脳卒中を発現した症例を網羅的に抽出するための条件を作成した。データ期間は 2015 年 4 月 1 日から 2017 年 12 月 31 日とした。初期の検索条件(初期ルール)は、表 1 にある、【病名 a(脳梗塞)/病名 b(脳出血)の開始日、かつ病名 a/b の開始日の前日から終了日まで、検査 a の実施日があること】とした。病名開始日は、DPC の入院日とした。脳出血はくも膜下出血・外傷性出血を含まない。脳梗塞はもやもや病と TIA を含む。

条件名	データソース	条件に含まれる内容
病名 a (脳梗塞)	DPC	G08(静脈洞血栓など)、G45(一過性脳虚血発作など)、I63(脳梗塞など)、I64(明示されない脳卒中など)、I65(脳塞栓など)、I66(脳動脈閉塞など)、I67(もやもや病など)
病名 b (脳出血)	DPC	I61(脳出血など)、I62(非外傷性硬膜下血腫など)、I64(明示されない脳卒中など)、P52(新生児脳出血など)
検査 a	レセプト/DPC	CTまたはMRI撮影(実施)、他医療機関のCTまたはMRI

表 1 初期ルールに用いた検索条件

MID-NET から当該条件を用いて抽出された症例について、医師が診療録等の記録を確認し、臨床ガイドライン[2]を参考に作成した判定アルゴリズム(図 1、2)に基づいて真のケースを判定し、陽性的中度等を算出した。抽出数が多い場合は、ランダムサンプリングした 200 例を対象とした。さらに、機械学習により PPV に影響する要因を検討し、PPV の上昇を目的に改良型アウトカムを作成した。機械学習アルゴリズムとしては勾配ブースティング(Gradient Boosting Decision Tree, GBDT)を用いた。具体的には、統計解析ソフト R のパッケージ xgboost を用い機械学習を行った。ROC 解析については pROC パッケージを使用し AUC を評価した。

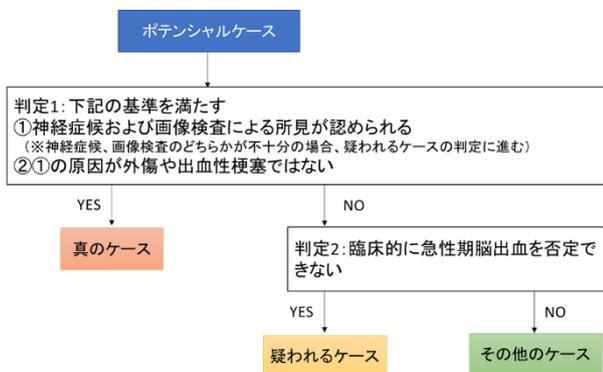


図 1 脳出血の判定アルゴリズム

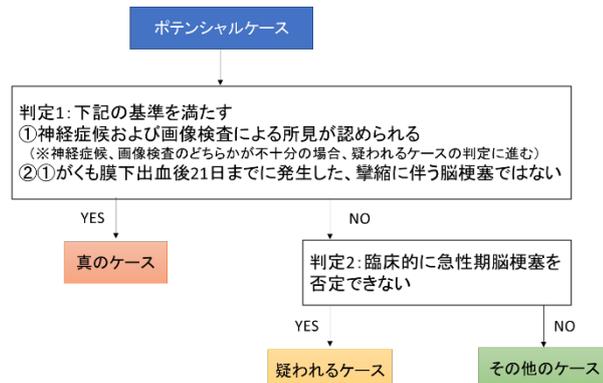


図 2 脳梗塞の判定アルゴリズム

3. 結果

脳卒中は、脳出血と脳梗塞に分類し、それぞれ抽出・評価した。

3.1 脳出血

病名 b に該当した例は 1873 例であった。このうち初期条件に該当した 245 例が抽出され、全例を判定対象とした。真のケースは 119 例、疑われるケース 2 例、その他のケース 124 例であり、PPV は 48.6%であった(表 2)。

脳出血	真のケース	疑われるケース	その他のケース	判定対象ポテンシャルケース	陽性的中度
	件数	件数	件数	件数	割合(%)
	119	2	124	245	48.6
脳出血	真のケース	疑われるケース+その他のケース		判定対象ポテンシャルケース	陽性的中度
	件数	件数		件数	割合(%)
	119	126		245	48.6
脳出血	真のケース+疑われるケース		その他のケース	判定対象ポテンシャルケース	陽性的中度
	件数		件数	件数	割合(%)
	121		124	245	49.4

表 2 脳出血の判定結果

その他のケース 124 例の詳細を調査したところ、疑い病名が 74 例、慢性硬膜下血腫(I620)が 34 例、スクリーニング目的が 17 例、既往歴としての脳出血が 16 例、外傷性脳出血が 4 例であった(重複を含む)。

機械学習の結果、AUC は 0.935 であり、高い AUC が得られた(図 3)。特に病名としての脳出血の疑い、慢性硬膜下血腫が強く関連しており、それらを除外できれば陽性的中度の上昇が見込まれると考えられた。

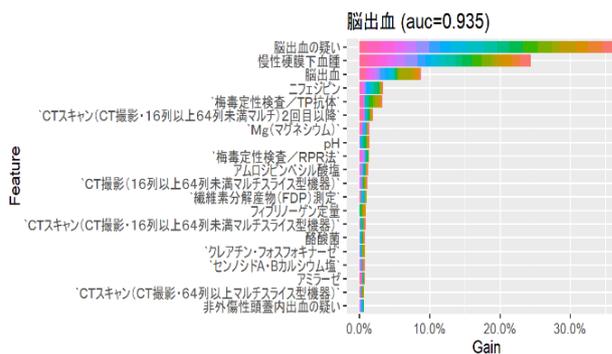


図 3 機械学習の結果(脳出血)

上記の結果から、脳出血の改良型アウトカムとして、「確定病名および画像検査がある」(疑い病名を除き、I620 を含まない)を設定した。予測される PPV は、真のケースを真のみとした場合 85.6%、真のケースを真と疑いとした場合 87.1%と、十分な改善が見込まれた(表 3)。真のケースは減少せず、その他のケースのみ抽出対象外となった。感度、特異度の算出を行ったところ、真のケースを真のみとした場合、感度は

119/119=100%、特異度は 106/126=84.1%、真のケースを真と疑いとした場合、感度は 121/121=100%、特異度は 106/124=85.5%であった。また、正確度、F 値も90%以上であった(表 3)。

改良型アウトカム

真	疑い+その他	全対象	PPV (%)
119	20	139	85.6
真+疑い		その他	全対象
121		18	139
			PPV (%)
			87.1

感度・特異度

	感度	特異度	正確度	F値
真	100.00	84.12	91.83	92.24
真+疑い	100.00	85.48	92.65	93.10

表 3 脳出血検索の改良型アウトカム該当者、感度・特異度

その他のケースとして残った例は、既往歴としての脳出血(リハビリ目的の入院等)や、外傷による出血等であった。

3.2 脳梗塞

病名 a に該当した例は 5802 例であった。このうち初期条件に該当した 990 例が抽出され、200 ケースをランダムサンプリングし、判定対象とした。真のケースは 63 件、疑われるケースは 11 件、その他のケースは 126 件であり、PPV は 31.5%であった(表 4)。

脳梗塞	真のケース	疑われるケース	その他のケース	判定対象ポテンシャルケース	陽性的中
	件数	件数	件数	件数	割合 (%)
	63	11	126	200	31.5
脳梗塞	真のケース	疑われるケース+その他のケース		判定対象ポテンシャルケース	陽性的中
	件数	件数		件数	割合 (%)
	63	137		200	31.5
脳梗塞	真のケース+疑われるケース	その他のケース	判定対象ポテンシャルケース	陽性的中	
	件数	件数	件数	割合 (%)	
	74	126	200	37.0	

表 4 脳梗塞の判定結果

その他のケースに該当した 126 例の詳細を調査したところ、疑い病名が 28 例、既往歴としての脳梗塞・陳旧性脳梗塞が 77 例、スクリーニング目的での病名付与が 36 例、脳梗塞・TIA に至らないもやもや病が 13 例、曖昧な病名(164(脳卒中:脳出血等を含む)、1678(その他の明示された脳血管疾患:放射線性脳壊死など)が 5 例であった(重複を含む)。

機械学習 1 回目として、病名+薬品+検査+画像データを含めた(図 4)。AUC は 0.526 と低値であったため、AUC を改善するため条件を調整して機械学習を繰り返した。

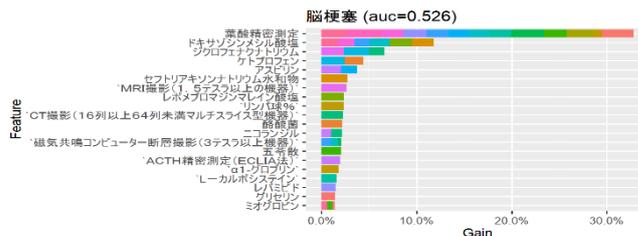


図 4 機械学習の結果 1 回目(脳梗塞)

機械学習 2 回目には、脳梗塞の病型や標準的な治療法を考慮し、複数の類似項目をグルーピングした。病名(アテローム血栓性、塞栓性、ラクナ、TIA、もやもや病、(明示されない)脳梗塞、その他の脳梗塞、疑い病名)、検査(薬酸、PT-INR)、治療薬(抗血小板薬(エイコサペンタエン酸を除く)、ヘパリン(静脈ラインロック用ヘパリン、透析用ヘパリン、皮下注射を除く)、NOAC、t-PA、エダラボン)としたところ、AUC は 0.577 とわずかに上昇した(図 5)。

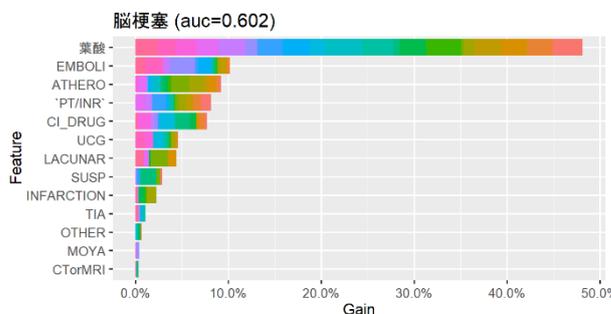


図 5 機械学習の結果 2 回目(脳梗塞)

3 回目には、さらなるグルーピングの項目の追加・整理を行った。病名(アテローム血栓性、塞栓性、ラクナ、TIA、もやもや病、(明示されない)脳梗塞、その他の脳梗塞、疑い病名)、検査(薬酸、PT-INR)、治療薬(以下のいずれかがあれば治療あり、抗血小板薬(エパデールを除く)、ヘパリン(静脈ラインロック用ヘパリン、透析用ヘパリン、皮下注射を除く)、NOAC、t-PA、エダラボン)、超音波検査(心臓超音波検査・経食道エコー)と変更したところ、AUC は 0.602 と上昇が認められた(図 6)。

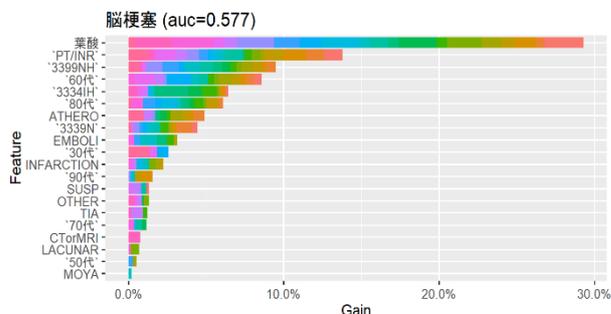


図 6 機械学習の結果 3 回目(脳梗塞)

疑い病名や脳梗塞に至らないもやもや病が多いため、それらを除外できれば陽性的中度的の上昇が見込まれると考えられた。また、機械学習で治療薬、PT-INR、超音波検査は比較

的関連が強かった。そこで、病名を整理し、治療薬、検査を含めた抽出条件により、PPVが60~70%まで上昇すると見込まれた(表5)。一方、もやもや病やTIAを除外することにより、真のケースが8例減少すると考えられた。

抽出条件	真のケース	全症例	PPV
①全対象	63	200	31.5
②:①から、疑い病名、頭部画像なし除く	63	139	45.3
③:②からもやもや病・TIA等除く	55	95	57.9
④:③にINRありを含める	53	84	63.1
⑤:④に薬剤ありを含める	52	85	61.2
⑥:④にUCGありを含める	37	58	63.8
⑦:④にINR・薬剤ありを含める	50	77	64.9
⑧:④にUCG・薬剤ありを含める	35	52	67.3
⑨:④にINR・UCGありを含める	36	55	65.5
⑩:④にINR・薬剤・UCGありを含める	34	50	68.0

表5 脳梗塞抽出の改良型アウトカム

上記の結果から、脳梗塞の改良型アウトカムとして、「「疑い」が付かない確定病名および画像検査および医薬品がある」と設定した。抽出件数が減少し、感度も低下すると考えられるため、検体検査と超音波検査はオプションとした。改良型アウトカムを踏まえたスクリプトの定義1により、111件の抽出が見込まれた。改良型アウトカムを踏まえたスクリプトを実行したところ、定義1により105件が抽出され、他の定義と合わせPPVは40%台後半から60%台と改善が見られた。しかし、疑い病名を完全には除くことはできず、また頭部画像のない症例も除外できないことから、当初の見込みよりPPVは低めとなった(表6)。

改良型アウトカム	真のケース		合計	陽性的中位割合(%)	95%信頼区間 [下限値-上限値]
	件数	疑われるケース+その他のケース			
No.1	57	53	105	49.5	[39.6-59.5]
No.2	44	43	87	50.6	[39.6-61.5]
No.3	33	24	57	57.9	[44.1-70.9]
No.4	28	20	48	58.3	[43.2-72.4]

改良型アウトカム	真のケース+疑われるケース		合計	陽性的中位割合(%)	95%信頼区間 [下限値-上限値]
	件数	その他のケース			
No.1	57	48	105	54.3	[44.3-64.0]
No.2	47	40	87	54.0	[43.0-64.8]
No.3	37	20	57	64.9	[51.1-77.1]
No.4	30	18	48	62.5	[48.8-76.2]

真	感度	特異度	正確度	F値
No.1	52/63=82.5	(137-53)/137=61.3	(52+137-53)/200=68.0	2/(1/82.5+1/49.5)=61.9
No.2	44/63=69.8	(137-43)/137=68.6	(44+137-43)/200=69.0	2/(1/69.8+1/50.6)=58.7
No.3	33/63=52.4	(137-24)/137=82.5	(33+137-24)/200=73.0	2/(1/52.4+1/57.9)=55.0
No.4	28/63=44.4	(137-20)/137=85.4	(28+137-20)/200=72.5	2/(1/44.4+1/58.3)=50.4

真+疑い	感度	特異度	正確度	F値
No.1	57/74=77.0	(126-48)/126=61.9	(57+126-48)/200=67.5	2/(1/77.0+1/54.3)=63.7
No.2	47/74=63.5	(126-40)/126=68.3	(47+126-40)/200=66.5	2/(1/63.5+1/54.0)=58.4
No.3	37/74=50.0	(126-20)/126=84.1	(37+126-20)/200=71.5	2/(1/50.0+1/64.9)=56.5
No.4	30/74=40.5	(126-18)/126=85.7	(30+126-18)/200=69.0	2/(1/40.5+1/62.5)=49.2

表6 脳梗塞検索の改良型アウトカム該当者、感度・特異度

4. 考察

4.1 脳出血

脳出血のDPC病名および頭部画像に基づく抽出では、PPVは50%程度であった。

その他のケースとなった中には、疑い病名としての脳出血が多数含まれていた。意識障害がある例では、ほぼ全例で頭部CTまたはMRIを撮影する。また、頭痛の訴えがある場合、頭蓋内出血の否定のため、しばしば頭部画像を撮影する。さらに、大手術前や治療前のスクリーニングでも頭部画像を撮影する。頭部画像の撮影は多くの局面で行われており、頭蓋内病変が認められなければ、脳出血の疑いの病名が付けられる。

また、本来急性期脳出血ではない、慢性硬膜下血腫も多く抽出された。I620は細分類では硬膜下出血(急性)(非外傷性)だが、慢性硬膜下血腫も同じコードであり、抽出対象となっていた。一方、非外傷性の硬膜下血腫は抽出されなかった。

機械学習でも、脳出血の疑い、慢性硬膜下血腫が突出して強く関連していた。また、脳出血の疑いまたは慢性硬膜下血腫の病名を持つ例には、真のケースは認められなかった。よって、抽出条件から疑い病名およびI620の病名を除外することにより、PPVは大きく改善すると考えられた。実際、改良型アウトカムにより、PPVは85%以上が見込まれ、感度も低下しなかった。

4.2 脳梗塞

初期ルールである脳梗塞のDPC病名およびCT・MRI画像の有無に基づく抽出では、PPVは30%程度と低かった。

意識障害や脱力等の症状がある例では、ほぼ全例で頭部CTまたはMRIを撮影するが、所見が認められないことも多い。また、脳出血と異なり、症状がはっきりしないこともあるため、画像所見に依存する部分が多い。さらに、症状がない場合でも、大手術等の治療前にはスクリーニング目的に頭部画像を撮影する。このため、脳梗塞の疑いの病名が付けられる場面は非常に多い。疑い病名の除外により、PPVの改善が見込まれた。

東北大学の脳神経外科は伝統的にもやもや病患者を多く扱っており、多数抽出された。多くは脳梗塞に至らないもやもや病であり、もやもや病の精査あるいは手術目的の入院であった。真のケースとなった例も、TIAがほとんどであり、明確な脳梗塞はごく一部であったため、改良型アウトカムでは抽出対象としないこととした。TIA自体も明確な画像所見が認められることが少ないため、抽出対象外とした。また、初期ルールでは脳卒中のような曖昧な病名も含まれており、脳出血や放射線による脳障害も抽出された。曖昧な病名も対象外とすべきと考えられた。

レセプトやDPCデータにはCT・MRIの撮影部位が含まれないため、頭部画像が存在せず、腹部等の他部位の画像のみが存在する例も抽出された。これらも対象外すべきと考えられた。

機械学習では、単純に個々の薬品や検査を含めてもAUCは0.53程度と高くなかった。そこで、対象を脳梗塞に関連す

ると考えられる項目に絞り込み、さらに病名や薬剤のグルーピングを行うことで、0.6程度までAUCの上昇が認められた。

機械学習の結果を踏まえ、改良型アウトカム定義を作成しスクリプトを実行したところ、PPVは60%台まで改善した。しかし、疑い病名や頭部画像のない症例を完全に除外することはできず、十分な改善は得られなかった。同時に、感度の低下も認められた。検査や治療のある例に限定したのみならず、もやもや病やTIAを除外した影響もあった。

頭部画像のない症例を除外するには、SS-MIX2に移行した放射線オーダー・実施情報を用いることにより、抽出対象を頭部画像のある例に絞り込むことが可能であると考えられる。しかしながら、SS-MIX2の放射線オーダーは標準コード化されていないため、拠点ごとに項目名が異なっている可能性がある。将来的に放射線オーダーの標準コード化も考慮する必要がある。

スクリーニング等によるCTやMRIの撮影でも、偶発的に陳旧性脳梗塞が認められることがあるが、脳梗塞の確定病名が付けられていた。このような場合には無症候性脳梗塞等の病名を選択すべきであり、院内での適切な病名選択を徹底する必要がある。

脳梗塞では、画像所見が特に重要であり、画像所見を評価しないで真のケースを特定するのは限界がある。将来的には、SS-MIX2拡張ストレージを活用し、画像や放射線レポートを取り込んで参照するなどの改良も望まれる。

5. 結論

DPC病名と画像検査に基づく初期ルールではPPVは高くなかった。しかし、脳出血では、疑い病名と慢性硬膜下血腫が多いことが主な原因であり、疑い病名の除外、および対象病名からI620を除いて抽出条件を適正化することにより、十分なPPVの上昇が見込まれた。一般的に脳出血は画像所見に基づいて診断されるが、抽出条件の最適化により、適切な抽出が可能であることが示された。

一方、脳梗塞では病名の整理や機械学習の結果に基づく検査や治療のグルーピングにより、ある程度のPPVの改善が認められたが、十分ではなかった。脳梗塞の診断は脳出血以上に画像所見による部分が大きく、データのみでは限界がある。SS-MIX2拡張ストレージを活用し、画像や放射線レポートを取り込んで参照するなどの改良が望まれる。

参考文献

- [1] MID-NET. <https://www.pmda.go.jp/safety/mid-net/0001.html>. Accessed at August 23, 2019.
- [2] Sacco R, et al. An updated definition of stroke for the 21st Century. A statement for healthcare professionals from the American Heart Association / American Stroke Association. *Stroke*. 2013; 44: 2064-89.