

一般口演 | 標準化

一般口演5 標準化

2019年11月22日(金) 09:00 ～ 10:30 |会場 (国際展示場 展示ホール8・特設会場1)

[2-I-1-06] SS-MIX2ストレージに収集されるデータ項目の臨床研究における網羅性の検討

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

キーワード：Standardization, SS-MIX2, SEAMAT

【背景】病院情報システムから診療データを抽出し臨床研究に役立てることに対する期待は大きく、Standardized Structured Medical Information eXchange（SS-MIX2）を用いた事業が一般的である。しかしながら、臨床研究は多種多様の情報を必要とするため、SS-MIX2でどれだけの情報が集まりうるかは一定の見解がない。この研究では、一般的な臨床研究を行う上で、データ範囲の網羅性を検証した。

【方法】一般的な臨床研究で扱うデータ項目を調べるため New England Journal of Medicine誌の原著論文を用い、各研究でエントリーされている患者基本情報からデータ項目を抽出、それらの項目を SS-MIX2標準化ストレージの出力項目がどの程度補えるかを調査した。対象疾患分野は、がん、心臓病、脳卒中、感染症、認知症、糖尿病、呼吸器、血液の7つに絞り、2018年378巻1号から2018年379巻10号までの Original Article論文のうち95報を対象とした。

【結果】標準化ストレージの出力項目で補える可能性が最も高かった疾患は糖尿病（61.8±13.7%）で、次に血液（55.9%±17.1%）、心臓病（48.0±13.0%）、呼吸器（45.0±16.3%）、感染症（43.6±24.9%）、認知症（13.3±13.3%）の順となった。日本循環器標準出力フォーマット（SEAMAT）を用いた場合で検討してみると、心臓病は82.7±0.1%と有意に上昇した。

【結論】SS-MIX2標準化ストレージを中心としたデータ抽出は、糖尿病や血液疾患など採血検査が種となる研究では過半数のデータ項目を自動で抽出できる可能性があるが、その他の診療科特有のデータを必要とする分野では拡張ストレージに収集されるデータを活用する必要があるが示唆された。

SS-MIX2 ストレージに収集されるデータ項目の臨床研究における網羅性の検討

中山 雅晴^{*1,2}、大友 千晶^{*2}、井上 隆輔^{*2}
*1 東北大学大学院医学研究科医学情報学
*2 東北大学病院メディカル IT センター

Completeness of Clinical Data Retrieved from Standardized Structured Medical Information eXchange Storage in Clinical Research

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

^{*1} Medical Informatics, Tohoku University Graduate School of Medicine

^{*2} Medical IT center, Tohoku University Hospital

Background: In Japan, physicians and researchers hold high expectations for the utilization of patient data in hospital information systems for clinical research, and their expectations have recently been realized in many projects using medical databases. Standardized Structured Medical Information eXchange (SS-MIX), including standardized storage and extended storage, is a common means of sharing clinical information from different vendors' hospital information systems. However, little is known about the number of types of data appropriate to research that can be extracted from SS-MIX storage. Method: Thirty-six volumes of the New England Journal of Medicine were examined. Diseases described in original articles were categorized into eight groups (i.e., cancer, cardiac diseases, stroke, infectious diseases, dementia, diabetes, respiratory diseases, and blood disorders). We calculated the proportions of data regarding patient characteristics potentially extracted via SS-MIX2 in each clinical study. Results: Of the 140 articles examined, the proportions varied from 13.3% ± 13.3% (mean ± SD) for dementia to 61.8% ± 13.7% for diabetes. In cardiology, the proportion increased significantly, from 48.0% ± 13.0% to 82.7% ± 0.1%, with the use of data from extended storage. Conclusion: To some extent, data extracted from SS-MIX2 could include basic information regarding patient characteristics in clinical research. More comprehensive data extraction using SS-MIX2 extended storage is required for efficient data usage in clinical research.

Keywords: SSMIX2, standardized storage, extended storage

1. 背景

病院情報システムから診療データを抽出し臨床研究に役立てることに対する期待は大きく、現在多くのデータベース事業が行われている。その際、異なるベンダーによる病院情報システム間ではデータの互換性が担保されないため、厚労省標準保存形式である Standardized Structured Medical Information eXchange version 2 (SS-MIX2)¹⁾を用いることで複数施設からのデータ共有を可能となる。実際、MID-NET²⁾をはじめとした事業で活用されている。SS-MIX2 のストレージは標準化ストレージと拡張ストレージとに分かれ、前者には構造化されたデータが保存され、患者基本情報、採血検査、処方、注射などが該当する。一方、それ以外の情報は拡張ストレージに格納されるが、標準化コードが未整備であったり正規化が不十分であったりするなど二次利用への適用は困難である。従って、多種多様の情報を必要とする臨床研究を行うには標準化ストレージだけでどれだけの情報が集まりうるかの検討は必要であるが、定量的に示した報告はない。本研究では、一般的な臨床研究を行う上で SS-MIX2 を介したデータ抽出による対象範囲の網羅性を検証した。

2. 方法

一般的な臨床研究で扱うデータ項目を調べるため New England Journal of Medicine 誌の Original Article を用い、多くの研究において Table 1 で示されるような Patient Characteristics で対象となっているデータ項目を抽出した。そして、それらの項目が SS-MIX2 標準化ストレージの出力項目からどの程度補えるかを調査した。各研究を疾患領域ごと

分類し、研究における対象データを分母、SS-MIX2 標準化ストレージから拾える項目を分子とし、その割合の平均と標準偏差を調べた。対象論文は2018年378巻1号から2018年379巻10号までのOriginal Article論文140報とし、対象疾患分野は、がん、心臓病、脳卒中、感染症、認知症、糖尿病、呼吸器、血液の8つに絞り、調査した。また、心臓病の分野に関しては、SS-MIX2 拡張ストレージへの日本循環器学会標準出力フォーマット (Standard Export data format: SEAMAT)³⁾を用いて、抽出できるデータ項目数の増加を調べた。2群間の検定は t-検定で行い、0.05 未満を有意差の条件とした。

表1 糖尿病に関連した臨床研究に対する SS-MIX2 標準化ストレージの網羅性

ACE Inhibitors and Statins in Adolescents with Type 1 Diabetes (N Engl J Med 2017; 377:1755-1765)		
Male — no. (%)		●
Age — yr		●
Age at diagnosis of diabetes — yr		●
Duration of type 1 diabetes — %		×
Blood pressure — mm Hg		×
Glycated hemoglobin level — %		●
Insulin administration — %	Multiple daily injections	●
Renal assessment	Albumin-to-creatinine ratio	●
Estimated glomerular filtration rate — ml/min/1.73 m ²	Based on creatinine equations	●
Retinopathy — %		○
Cardiovascular markers	Asymmetric dimethylarginine — ng/ml	○
	High-sensitivity C-reactive protein — mg/liter	○
Cholesterol level — mg/dl		●
Triglyceride level — mg/dl		●
Ratio of apolipoprotein B to apolipoprotein A1		○
Carotid intima-media thickness — mm		○
Effects of Sirtinib/Orlistat Added to Insulin in Patients with Type 1 Diabetes (N Engl J Med 2017; 377:2337-2348)		
Age — yr		●
Female sex — no. (%)		●
Race or ethnic group — no. (%)		○
Duration of diabetes — yr		○
Glycated hemoglobin — %		○
Fasting plasma glucose — mg/dl		○
Weight — kg		○
BMI		△
Systolic blood pressure — mm Hg		×
Diastolic blood pressure — mm Hg		×
Systolic blood pressure ≥ 130 mm Hg — no. (%)		×
Daily total dose of insulin — IU/day		●
Insulin dose — IU/day	Total	●
Type of insulin therapy — no. (%)	Subcutaneous injections	●
End Points		×
Patients with glycated hemoglobin > 7.0% and no severe hypoglycemia or diabetic ketoacidosis	All patients	●
Patients with glycated hemoglobin > 7.0% and ≥ 1 episode of severe hypoglycemia		●
Patients with glycated hemoglobin > 7.0% and ≥ 1 episode of diabetic ketoacidosis		●

3. 結果

7つの領域に該当する対象論文数は95報となり、それぞれの内訳はがん(26)、心臓病(8)、脳卒中(7)、感染症(23)、認知症(2)、糖尿病(5)、呼吸器(11)、血液(13)であった。標準化ストレージの出力項目で補える割合が最も高かった疾患は糖尿病(61.8±13.7%)であり、実際の例を表1に示した。次に血液(55.9±17.1%)、心臓病(48.0±13.0%)、呼吸器(45.0±16.3%)、感染症(43.6±24.9%)、認知症(13.3±13.3%)の順となった(図1)。心臓病における例は表2に示し、日本循環器標準出力フォーマット Standard Export data format (SEAMAT)を用いた場合で検討してみると、心臓病は82.7±0.1%と有意に上昇した(図2)。

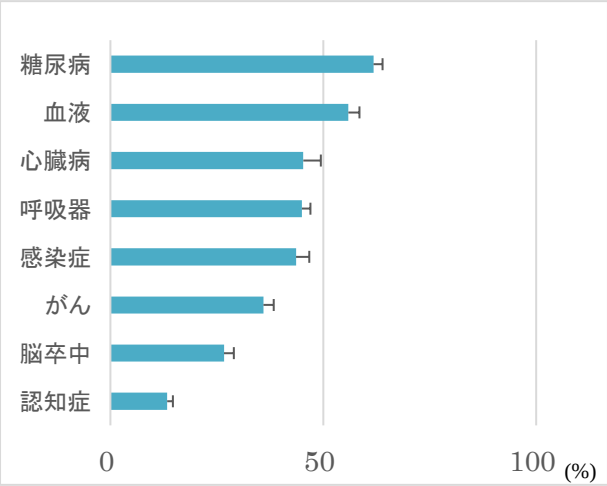


図1 各疾患領域におけるSS-MIX2標準化ストレージから抽出できるデータ項目の割合(平均値±標準偏差)

表2 循環器領域に関連した臨床研究に対するSS-MIX2標準化ストレージの網羅性およびSEAMATの対応

Radial Artery or Saphenous Vein Grafts in Coronary Artery Bypass Surgery (N Engl J Med 2018; 378:2069-2077)		SEAMAT
Age — yr	●	●
Female sex — no. (%)	●	●
Diabetes — no. (%)	○	○
Previous myocardial infarction — no. (%)	△	△
Elective admission — no. (%)	○	○
Renal insufficiency — no. (%)	○	○
Left ventricular ejection fraction <35% — no. (%)	×	×
Target vessel — no. (%)	×	×
No. of grafts	×	×
Proximal anastomosis — no. (%)	×	×
Outcome	×	×
Two-Year Outcomes with a Magnetically Levitated Cardiac Pump in Heart Failure (N Engl J Med 2018; 378:1386-1395)		○
Age — yr	●	●
Male sex — no. (%)	○	○
Race or ethnic group — no. (%)	○	○
Body surface area — m2	○	○
Ischemic cause of heart failure — no. (%)	△	△
History of atrial fibrillation — no. (%)	△	△
History of stroke — no. (%)	△	△
Previous cardiac surgical procedure — no. (%)	×	×
Left ventricular ejection fraction — %	×	×
Atrial blood pressure — mm Hg	×	×
Mean arterial pressure — mm Hg	×	×
Pulmonary capillary wedge pressure — mm Hg	×	×
Cardiac index — liters/min/m2 of body surface area	×	×
Pulmonary vascular resistance — Wood units	×	×
Right atrial pressure — mm Hg	×	×
Serum sodium — mmol/liter	×	×
Serum creatinine — mg/dl	×	×
Estimated glomerular filtration rate — ml/min/1.73 m2	×	×
Intended goal of pump support — no. (%)	×	○
Bridge to transplantation	×	○

4. 考察

病院情報システムからデータを広く抽出し、データ駆動型の臨床研究を効率的に行うことは、臨床医、疫学研究者、そして医療情報担当者にとって達成すべき目標である。単一の組織であれば可能であっても、多施設からの収集や検証は容易ではない。SS-MIX2が多くの事業で活用され、MIDNET事業における検証作業によりその問題点や改善点が浮き彫りになり、標準化ストレージの活用は現実的なものとなった。しかしながら、各疾患領域の専門家としては各疾患固有の情報こそが重要であり、その具現化が望まれている。

本研究は、筆者の知る限り、SS-MIX2で抽出するデータ項

目が、一般的な臨床研究のデータ項目において、どの程度網羅するかを定量的に調査した最初の研究と思われる。採血検査がデータ収集の中心となる糖尿病や血液分野では標準化ストレージ項目の網羅性は高くなるが、循環器特有の検査、画像診断所見、TNM分類、症状のスコアリング、神経機能、遺伝子検査等々、必要な検査が増加するにつれ、当然ながら網羅性は減少していった。

心電図や心臓超音波検査などで生じる数値情報を、生理検査システムからSS-MIX2拡張ストレージに出力するための出力フォーマットであるSEAMATは、多施設からの循環器情報を集める際に有用であり、既に複数の施設で導入されデータ収集が可能となっている⁴⁾。今回の結果でも飛躍的に網羅性が高まっており、専門性の高い検査の結果収集が望まれることを示唆した。今後、各疾患領域でも同様の試みがなされることにより、臨床研究に資するデータ収集が効率よく進められると推測される。

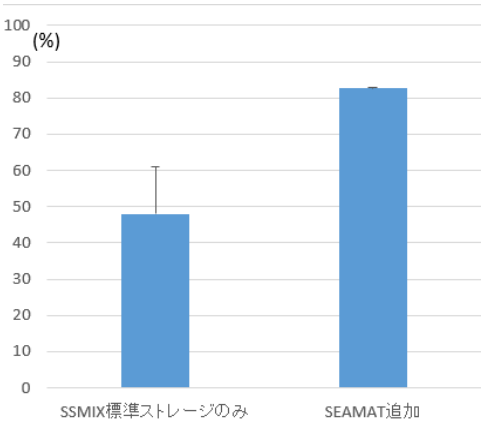


図2 心臓病領域におけるSS-MIX2標準化ストレージから抽出できるデータ項目の割合とSEAMATを用いたデータ項目を追加した場合の割合(平均値±標準偏差)

5. 結論

SS-MIX2標準化ストレージを中心としたデータ抽出は、糖尿病や血液疾患など採血検査が種となる研究では過半数のデータ項目を自動で抽出できる可能性があるが、その他の診療科特有のデータを必要とする分野では拡張ストレージに収集されるデータを活用する必要があると示唆された。

参考文献

- Kimura M, Nakayasu K, et al. SS-MIX: a ministry project to promote standardized healthcare information exchange. Methods Inf Med. 2011;50:131-9.
- Yamada K, Itoh M et al. The utilization and challenges of Japan's MID-NET medical information database network in postmarketing drug safety assessments: A summary of pilot pharmacoepidemiological studies. Pharmacoepidemiology & Drug Safety. 2019;28:601-608.
- SEAMAT [http://www.j-circ.or.jp/itdata/jcs_standard.htm] (accessed 2019-Aug-19).
- Matoba T, Nagai R et al. Architecture of the Japan Ischemic Heart Disease Multimodal Prospective Data Acquisition for Precision Treatment (J-IMPACT) System. Int Heart J. 2019 20;60:264-270.