



「日本における傷病名を中心とする レセプト情報から得られる指標のバリデーション に関するタスクフォース」

第2回会議

2016年8月1日(月) PM6-8
学会センタービル地下会議室

日本薬剤疫学会
NPO日本医薬品安全性研究ユニット
久保田潔





電子レセプト匿名化マニュアル

- <http://www.dsrujp.org/> に公開

DSRU Japan

電子レセプト匿名化マニュアル

2016年8月 ウェブ版第1.05版

NPO日本医薬品安全性研究ユニット

〈外部サイトへのリンク、およびコードのコピーが可能なPDF版もご利用になれます。ご利用希望の方はメールの件名に「マニュアルPDF版希望」とご記入のうえ dsruj-adm@umin.ac.jp までご連絡ください（担当：佐藤国靖）〉

電子レセプト匿名化マニュアル



2016/9/9

NPO DSRU Japan

2

電子レセと匿名化

RE,1,1112,42602,久保田 潔,1,3271201,70,,,,, 999,,,,
HO,67138172,27,5165,3,2474,0,,,,,
SY,0000999,4261017,1,,五十肩,
SY,0000999,4251211,1,,打撲(略全身),

- 「メモ帳」などで簡単に見ることができる。
- 氏名・病名の一部は日本語。いやでも目に入る。
 - ． 一部医療機関では匿名化作業を内部で実施することにこだわり？
- 匿名化のための最初の数ステップに関する手順を示す。
 - ． バリデーション研究実施の妨げとなる障害を少しでも和らげる。
- 佐藤泉美先生のバリデーションで使ったSASプログラムを参考(本人のご了解の下)。
 - ． SASではなく「SmoothCSV」(無料ソフト)を利用

レセプトDBと 院内に保存された電子レセ

- レセプトDBは、複数の施設から発行されたレセプトを「名寄せ」して作成される。
 - ・ 複数の医療機関
 - ・ 院外薬局
- 単一の病院/クリニックのレセプトデータからDBデータの完全な「復元」は困難。特に
 - ・ 同時期の他の医療機関のデータが重要な時
 - ・ 外来患者に処方される薬が重要な時
 - ▲ 院内の記録による「補完」などが必要

電子レセの構造

RE,1,1112,42602,久保田 潔,1,3271201,70,,,,, 999,,,,
HO,67138172,27,5165,3,2474,0,,,,,,,,



● 「識別情報」から始まる

- IR: 医療機関情報・1レセ1回
- RE: レセプト共通レコード(1つのレセの始まり)個人情報
- HO: 保険者レコード(患者の属する保険者)個人情報
- KO: 公費レコード(一部負担金発生)の公費)個人情報
- SY: 傷病名
- SI: 診療行為(検査、手術、初診・再診など)
- IY: 医薬品
- TO: 特定器材
- その他: GR(包括評価対象外理由)CO(コメント)SJ(症状
詳記)TR・TS(臓器提供者)GO(請求書レコード)

匿名化作業

- まずレセプト種類(社保・国保・DPC)ごとに研究期間に発行されたレセプトを縦に結合。

- **通し番号**をつける

1, RE,1,1112,42602,久保田 潔,1,3271201,70,,,,, 999,,,

2, HO,67138172,27,5165,3,2474,0,,,,,,,,,

3, SY,0000999,4261017,1,,五十肩,

4, SY,0000999,4251211,1,,打撲(略全身),

レセプト共通レコード(RE)情報抽出

- 以下を取り出す
 - ・ (1) **行番号**
 - ・ (2) **レセプト番号**
 - ▲ 当該月に発行された社保または国保のレセプトに含まれるレセプトに与えられる通番
 - ・ (3) **診療年月**
 - ▲ 4(「平成」1ケタ)+年(2ケタ)+月(2ケタ)
 - ・ (4) 性別(1男、2女)
 - ・ (5) 生年月日
 - ▲ 3(「昭和」1ケタ)または4(「平成」1ケタ)+年(2ケタ)+月(2ケタ)+日(2ケタ)
 - ・ (6) **カルテ番号** <患者を一意に特定・研究用IDと入替>
- 以下を作成
 - ・ (6)[ID入替]+(1)+(2)+(3)→研究者へ
 - ・ (6)[ID入替]+(4)+(5)[左3桁に限れば生年のみ]→研究者へ
 - ・ (6)のカルテ番号-研究用ID対応表(院内に残す)

RE情報の一部を全行に記入

	1	2	3	4	5
1				1	IR
2	999	1	42804	2	RE
3				3	HO
4				4	SY
5				5	SY
6				6	SY
7				7	SY
8				8	SY

	1	2	3	4	5
1				1	IR
2	999	1	42804	2	RE
3	999	1	42804	3	HO
4	999	1	42804	4	SY
5	999	1	42804	5	SY
6	999	1	42804	6	SY
7	999	1	42804	7	SY
8	999	1	42804	8	SY

カルテ
番号

レセプト
番号

診療
年月

行番号

- 全レコードに、カルテ番号、レセプト番号、診療年月を加える。
- カルテ番号は、この後研究用IDと入れ替え

匿名化ファイル

- Mainのファイル:IR、RE、HO、KOを除外したレコード
- RE:2種類
 - ・ カルテ番号(研究用IDとおきかえ)+行番号+レセプト番号+診療年月
 - ▲レコード数=レセプトの数
 - ・ カルテ番号(研究用IDとおきかえ)+性+生年
 - ▲レコード数=患者数
- HO:個人情報(記号・番号)削除
- KO:個人情報(負担者・受給者番号)削除
- IR:「マニュアル」では使用を想定していない
 - ・ 複数の施設からのデータを利用するには有用(例:病院規模、地域?)

対応表

- カルテ番号と研究用IDとの対応表
． 研究終了まで医療機関内に保存

未対応の部分

- 「匿名化」されたMainのファイルは研究期間に発行された全レセプト情報を含む
 - ・ 医療機関(の倫理審査委員会)によっては、さらに「絞り込んだ」集団に関する情報に限定するように要求？
 - ・ その他



2016/9/9

NPO DSRU Japan

12

バリデーション研究の指標

Gold Standard			
		Yes	No
Claims -based definition	Yes	True positive [a]	False positive [b]
	No	False negative [c]	True negative [d]

- 感度 (Sensitivity) = $a/(a+c)$
- 特異度 (Specificity) = $d/(b+d)$
- 陽性的中度 (Positive Predictive Value(PPV)) = $a/(a+b)$
- 陰性的中度 (Negative Predictive Value(NPV)) = $d/(c+d)$

指標の特徴

Gold Standard			
		Yes	No
Claims -based definition	Yes	True positive [a]	False positive [b]
	No	False negative [c]	True negative [d]

- 感度 (Sensitivity) = $a/(a+c)$ と特異度 (Specificity) = $d/(b+d)$ はその施設の当該疾患の患者の割合に依存しない
 - ・ 陽性的中度 (Positive Predictive Value (PPV)) = $a/(a+b)$ と陰性的中度 (Negative Predictive Value (NPV)) = $d/(c+d)$ はその施設の当該疾患の患者の割合に影響される (例: 専門病院)
- 調査対象に含まれる全ての「真」の患者のデータ (例: 院内疾患登録) がないと感度・特異度は求められない
 - ・ PPV だけしか求められない研究も少なくない。
- RR など効果の指標推定では特異度 (PPV) が重要
- 発生 (有病) 割合などを正しく求めたいのなら感度も重要