



JE編集委員会企画
疫学研究に求められる観察研究の報告ガイドライン

データベース研究

東京大学大学院医学系研究科

臨床疫学・経済学

康永 秀生、森田 光治良

日常的に観察され収集される 医療情報を用いた研究の報告基準 ～ The RECORD Statement ～

The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement

様々な医療ビッグデータ

ライフサイエンス系
ビッグデータ

ヒトゲノムデータ
網羅的分子情報
など

臨床疫学系
ビッグデータ

診療報酬系データ
患者登録型データ
SS-MIX、電子カルテ
政府統計データ
など

健康・予防系
ビッグデータ

ウェアラブルデバイス
SNS
音声解析
など

医療ビッグデータとは

保健医療にかかわる種々の目的のために恒常的に収集・蓄積され、閲覧・検索・統合・集計・分析が可能な形でデジタル化されコンピュータに整理・格納されている医学・医療データの集合体

→共通のフォーマット、標準化された入力コード、匿名化処理、セキュリティ、などのデータマネジメントが必要

臨床疫学系ビッグデータの類型

1 診療報酬明細データ

Diagnosis Procedure Combination (DPC)データ、
全国レセプトデータ (NDB)、介護給付実態調査、など

2 患者登録型データ

がん登録、外科学会NCD、胸部外科学会JACVSD、
消防庁救急蘇生統計 (ウツタイン・データ)、など

3 政府統計

人口動態統計、患者調査、社会保険診療行為別調査、
国民生活基礎調査、国民健康栄養調査、医療施設調査、
医師・歯科医師・薬剤師調査、など

4 その他

電子カルテ、オーダリングシステム(SSMix)、など

[出典]康永秀生. DPCデータによる臨床疫学研究の成果と今後の課題. 医療と社会
2016;26(1):7-14

これまでの臨床研究におけるデータ収集の方法論

研究計画を立て、研究目的に沿ったデータを、
prospectiveにしるretrospectiveにしる、
single centerにしるmulti-centerにしる、
単発で収集するというアプローチ。

1つのprojectで論文を1～数本書いたらおしまい。
費用対効果の低い方法。
当たれば大きい、外れれば研究費のムダ。

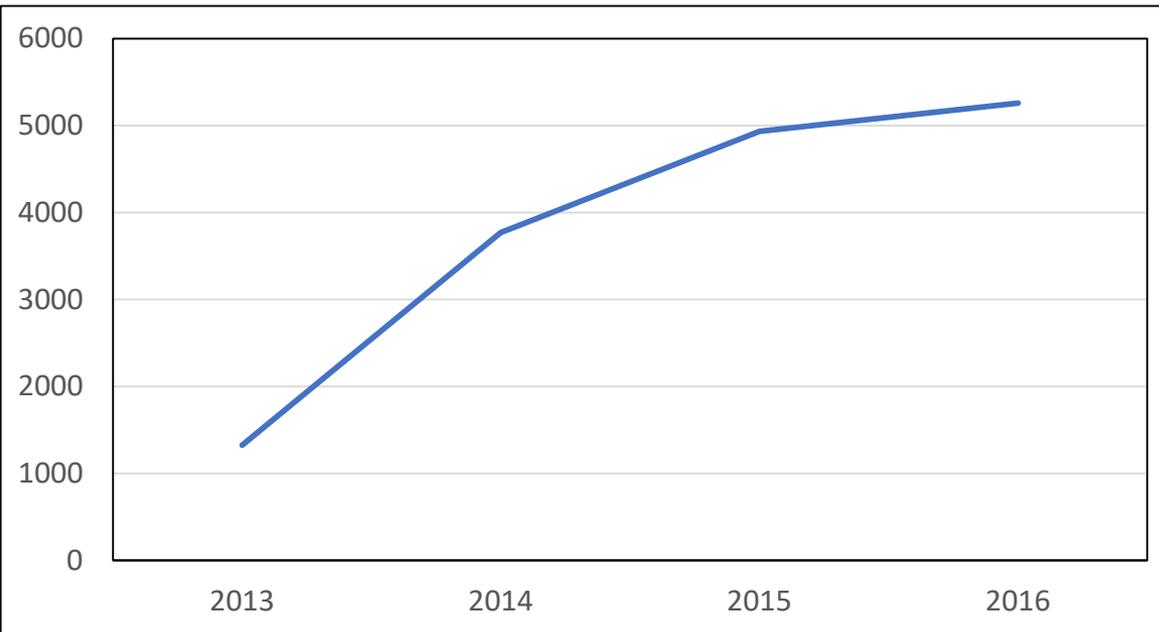
医療ビッグデータの時代

データを多施設から恒常的に収集し、蓄積し、医療ビッグデータを構築。

医療ビッグデータから個々の研究目的に沿ったデータを抽出し、エビデンスを量産し続ける。

背景

- これらのデータベースを用いた観察研究は年々増加中

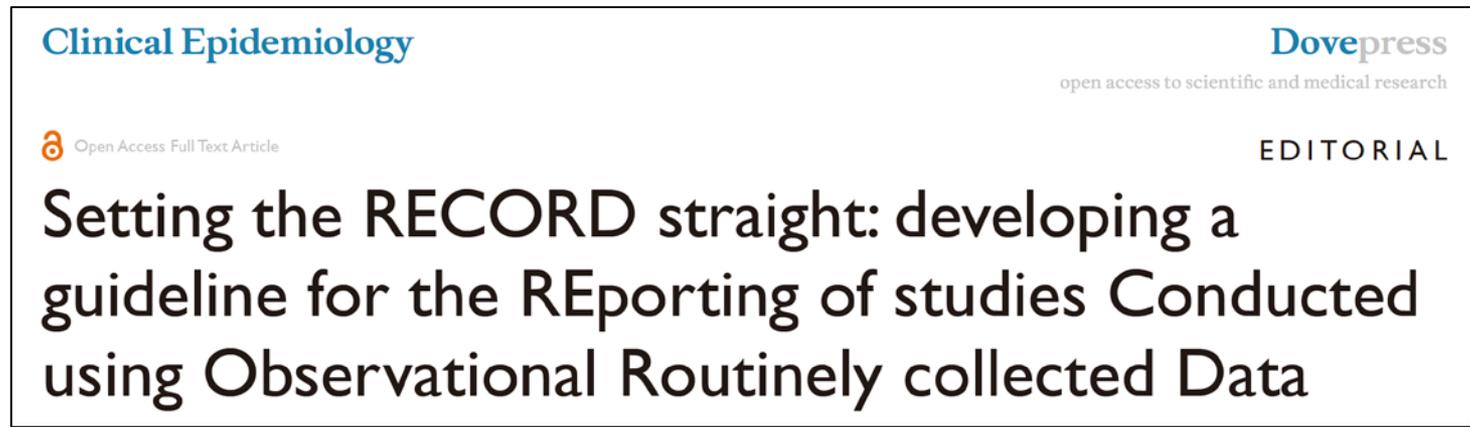


- Search on PubMed
(((((((administrative) OR "electronic medical record") OR registry) OR registries) OR surveillance) OR database) OR "claim data") OR "claims database"
Filters: Observational Study;
Humans; English

- 過去30年でこれらデータベースを用いた研究は増え続けている。

背景

- データベース研究における報告ガイドラインの必要性の高まり



- “ 基本的に Routinely collected data を用いた研究は、観察研究のため STROBE を適応すべきだが、STROBE はデータベース研究に特有の問題に言及できていない”
- “ データベース研究における「強み」、「限界点」、「バイアス」などは不明確”

RECORD声明作成の目的

国際的なコンセンサスを得る形でSTROBEを
拡張し、データベースを用いた研究において

- 質の向上（知識・スキルの向上）、
- 具体的で透明性のある報告の強化、
- 適切な報告を通じ、論文読者（ジャーナル編集者、ピアレビューア、科学者、臨床医、政策立案者を含む）がその仕事を理解し、内的小および外的妥当性を推論できるようになること

に利用できることを目指して作成が開始

STROBE声明

Annals of Internal Medicine

ACADEMIA AND CLINIC

Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration

Ann Intern Med. 2007 ;147(8):W163-94. ほか

観察的疫学研究報告の質改善（STROBE）のための声明：

解説と詳細

Strengthening the Reporting of Observational Studies in Epidemiology (STROBE):
Explanation and Elaboration.

- 120以上の雑誌で参照することが推奨。
- 日本語訳も存在（京都大学翻訳グループ）
- RECORDはSTROBEの代わりではなく、追加となるもの

RECORD声明の内容: Title and abstract

- STROBEに13の項目が追加される形で完成。

STROBE声明		RECORD声明	
タイトルと抄録 (Title and abstract)			
No.1	(a)タイトルか抄録のなかで、研究デザインを示す。 (b)研究で何を行って、何が明らかになったのか十分な情報を含み、かつバランスの良い要約を記載。	1.1	タイトルか抄録のなかで、使用したデータの種類を明示する。可能ならデータベース名を含めるべき。
		1.2	該当するなら、研究が行われた地理的位置と時間的枠組みを明示するべき。
		1.3	データベース間の結合を実施したなら、タイトルか抄録の中でデータ結合について明確に記述すべき。

RECORD声明の内容: Title and abstract

RECORD item 1.1

タイトルか抄録のなかで、使用したデータの種別を明示する。
可能ならデータベース名を含めるべき。

具体的には、“administrative data”, “disease registries”, “primary care databases”, “EHR data”, “population registries”のようなデータタイプを含めることを推奨している。

データベース名(例:DPCやNDB)の提示は、データタイプ明示の代わりにはならないと述べている。

RECORD声明の内容: Title and abstract

RECORD 11

ORIGINAL ARTICLE

Outcomes After Laparoscopic or Open Distal Gastrectomy for

Objective: In a large nationwide administrative database of hospitalized patients, we investigated postoperative outcomes after laparoscopic or open distal gastrectomy in Japan.

Background: The benefits of laparoscopic gastrectomy, such as decreased length of stay and morbidity, have typically been evaluated only with limited data on the basis of small samples.

Methods: Using the Japanese Diagnosis Procedure Combination Database, we identified 9388 patients who were preoperatively diagnosed with stage I and II gastric cancer and underwent laparoscopic (n = 3937) or open (n = 5451) distal gastrectomy between July and December 2010. One-to-one propensity score matching was performed to compare in-hospital mortality, postoperative complication rates, length of stay, total costs, and 30-day readmission rates between the 2 groups.

AD, DPH, †

ウリにはなりたい

RECORD声明の内容: Title and abstract

RECORD item 1.2

該当するなら、研究が行われた地理的位置と時間的枠組みを明示すべき。

(タイトルや抄録のWord数の制限にもよるが…) 調査対象集団を定義するために使用される最大の地理的レベル(例: 国家、県および地域)を報告。

RECORD声明の内容: Title and abstract

RECORD item 1.2

該当するなら、研究が行われた地理的位置と時間的枠組みを明示するべき。

この研究は、日本の全国的な行政データベースを用いて、胃癌手術の地理的位置と時間的枠組みを定義するために実施された。

Objective: In a large nationwide administrative database of hospitalized patients, we investigated postoperative outcomes after laparoscopic or open distal gastrectomy in Japan.

Background: The benefits of laparoscopic gastrectomy, such as decreased length of stay and morbidity, have typically been evaluated only with limited data on the basis of small samples.

Methods: Using the Japanese Diagnosis Procedure Combination Database, we identified 9388 patients who were preoperatively diagnosed with stage I and II gastric cancer and underwent laparoscopic (n = 3937) or open (n = 5451) distal gastrectomy between July and December 2010. One-to-one propensity score matching was performed to compare in-hospital mortality, postoperative complication rates, length of stay, total costs, and 30-day readmission rates between the 2 groups.

RECORD声明の内容: Title and abstract

RECORD item 1.2

該当するなら、研究が行われた地理的位置と時間的枠組みを明示すべき。

(タイトルや抄録のWord数の制限にもよるが…) 調査対象集団を定義するために使用される最大の地理的レベル(例: 国家、県および地域)を報告。

RECORD item 1.3

データベース間の結合を実施したなら、タイトルか抄録の中でデータ結合について明確に記述すべき。

“using multiple linked health administrative databases”などの記載や、結合したデータベース名を記載する(例: “DPCと…の結合” など)。

“linked” や “linkage” のような単語を用いることを推奨。

RECORD声明の内容: Title and abstract

Association between Nurse Staffing and In-Hospital Bone Fractures: A Retrospective Cohort Study

Kojiro Morita, Hiroki Matsui, Kiyohide Fushimi, and Hideo Yasunaga

Objective. To determine if sufficient nurse staffing reduced in-hospital fractures in acute care hospitals.

Data Sources/Study Setting. The Japanese Diagnosis Procedure Combination inpatient (DPC) database from July 2010 to March 2014 linked with the Surveys for Medical Institutions.

RECORD声明の内容: Introduction

STROBE声明		RECORD声明
緒言 (Introduction)		
背景/論拠(Background/Rationale) ※		
No.2	研究の科学的な背景と論拠を説明する。	-
目的 (Title) ※		
No.3	特定の仮説を含む研究目的を記述する。	-

RECORDでは、“Introduction”の項目においてSTROBE声明の拡張はない。そのため、通常のSTROBE声明の参照を推奨する。

しかし、データベースを用いた研究では、研究目的がデータの新しい関係を発見すること(データマイニングや仮説生成など)か、仮説検証なのかより明確にすべきである、と述べている。

※以後、※が付く項目はRECORD特有の項目がないもの

RECORD声明の内容: Methods

STROBE声明		RECORD声明
方法 (Methods)		
研究デザイン(Study design) ※		
No.4	研究デザインの重要な要素(Key elements)を論文の最初の部分で示す。	-
セッティング (Setting) ※		
No.5	登録、曝露、追跡、データ収集の期間を含む研究のセッティング、実施場所、基準となる日付について明示する。	-

“Methods” No. 4, 5 の項目において STROBE の拡張はない。しかし、データベースの内容と妥当性、データ収集された元々の理由などの情報を提供する必要に言及している。

また、研究結果を広く母集団に当てはめることができるか判断できるように、選択基準などデータベースの集団と母集団の関係性を記載すべきとしている。

RECORD 声明の内容: Methods

METHODS

Diagnosis Procedure Combination database

The Diagnosis Procedure Combination database is a case-mix patient classification system, which was launched in 2002 by the Ministry of Health, Labour and Welfare of Japan, and linked with a lump-sum payment system.¹⁶ All 82 university teaching hospitals are obliged to adopt this system, but adoption by community hospitals is voluntary. The survey of the participating hospitals is conducted

データベースの内容とデータ収集された元々の理由
研究のセツティンク
日付について明示する。

Ann Surg. 2013;257(4):640-6.

“Methods” No. 4, 5 の項目において STROBE の拡張はない。

しかし、データベースの内容と妥当性、データ収集された元々の理由などの情報を提供する必要に言及している。

また、研究結果を広く母集団に当てはめることができるか判断できるように、選択基準などデータベースの集団と母集団の関係性を記載すべきとしている。

RECORD声明の内容: Methods

STROBE声明

RECORD声明

参加者 (Participants)

No.6	適格基準、参加者の母集団、選定方法を明示する。 (a)コホート研究では、追跡方法についても記述する。 ケースコントロール研究では、ケースとコントロールの選択における論拠を示す。	6.1	研究対象集団の選定方法 (利用したコードやアルゴリズム)を明確に記載すべき。 これができない場合は、説明すべき。
		6.2	対象者選択に用いたコードやアルゴリズムの妥当性研究は、引用すべきである。 もし、妥当性検証について研究の中で実施され、他で公表されていない場合は、詳細な方法と結果を提供すべき。

RECORD声明の内容: Methods

RECORD item 6.1

研究対象集団の選定方法(利用したコードやアルゴリズム)を明確に記載すべき。これができない場合は、説明すべき。

研究の透明性を示すため、研究対象集団を特定するために用いたコードやアルゴリズムの妥当性を報告することが重要としている。これらコードやアルゴリズムを提示することは、他の研究者に対して、研究の外的・内的妥当性を示すことにもつながる。

RECORD声明の内容: Methods

Stroke

JOURNAL OF THE AMERICAN HEART ASSOCIATION



American Heart Association | American Stroke Association®

Im Patient Selection

We identified patients who were admitted with a diagnosis of ischemic stroke (ICD-10 code: I63) and discharged between April 2012 and March 2014. We included the following patients: those aged >20 years, those with independent daily activities before admission (mRS score ≤ 2), those admitted within 3 days of onset, those who received rehabilitation during hospitalization, and those admitted to hospitals receiving at least 10 patients with ischemic stroke per year on average.

nwide

ashi

RECORD声明の内容: Methods

RECORD item 6.2

対象者選択に用いたコードやアルゴリズムの妥当性研究は、引用すべきである。もし、妥当性検証について研究の中で実施され他で公表されていない場合は、詳細な方法と結果を提供すべき。

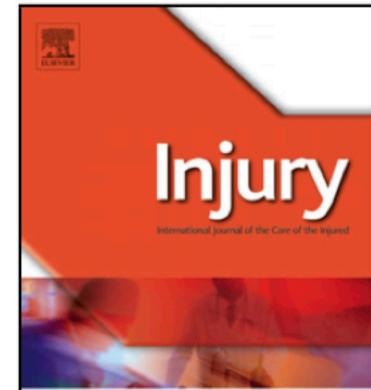
RECORD声明の内容: Methods

RECORD item 6.2

対象者選択に用いたコードやアルゴリズムの妥当性研究は、

Accepted Manuscript

Title: Development and Validation of an ICD-10-Based Disability Predictive Index for Patients Admitted to Hospitals with Trauma



Aut complications arising after admission. These diagnoses are recorded by ICD-10 codes
Hir

and text data written in Japanese. A previous study showed that the validity of diagnosis records in the DPC database was generally high. The sensitivity and specificity of the primary diagnoses were 78.9% and 93.2%, respectively. [13] Although the previous study only evaluated medical diagnoses, and not trauma diagnoses, we believe that the

RECORD声明の内容: Methods

RECORD item 6.2



Journal of Epidemiology

journal homepage: <http://www.journals.elsevier.com/journal-of-epidemiology/>

Original Article

Validity of diagnoses, procedures, and laboratory data in Japanese

Table 3

Frequencies of diagnoses and validity indices for the DPC data-based diagnosis identification.

Diagnoses	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Original Charlson diseases identified by reco				
Myocardial infarction	52.2	99.7	92.3	96.4
Congestive heart failure	68.8	97.5	75.9	96.5
Peripheral vascular disease	34.5	99.3	83.3	93.7
Cerebrovascular disease	50.0	98.9	86.4	93.5
Dementia	37.5	100.0	100.0	96.8

RECORD声明の内容: Methods

RECORD item 6.2

対象者選択に用いたコードやアルゴリズムの妥当性研究は、引用すべきである。もし、妥当性検証について研究の中で実施され他で公表されていない場合は、詳細な方法と結果を提供すべき。

RECORD声明の内容: Methods

RECORD item 6.2

Dental procedures, antibiotic prophylaxis, and endocarditis among people with prosthetic heart valves: nationwide

Methods

Definition of oral streptococcal endocarditis

To evaluate the sensitivity of this algorithm for oral streptococcal infective endocarditis, the study authors (ST, XD) reviewed medical records of patients admitted to a tertiary care hospital between 2008 and 2014 with at least one positive blood culture result for oral streptococci.

Results

Validation of ICD-10 coding for oral streptococcal infective endocarditis

The study definition of oral streptococcal infective endocarditis validated using medical records of 130 participants admitted to hospital with at least one positive blood culture result for oral streptococci had a sensitivity of 54% and a positive predictive value of 100%.

Discussion

妥当性検証論文の引用ではなく、
研究内で妥当性も確認した論文の例

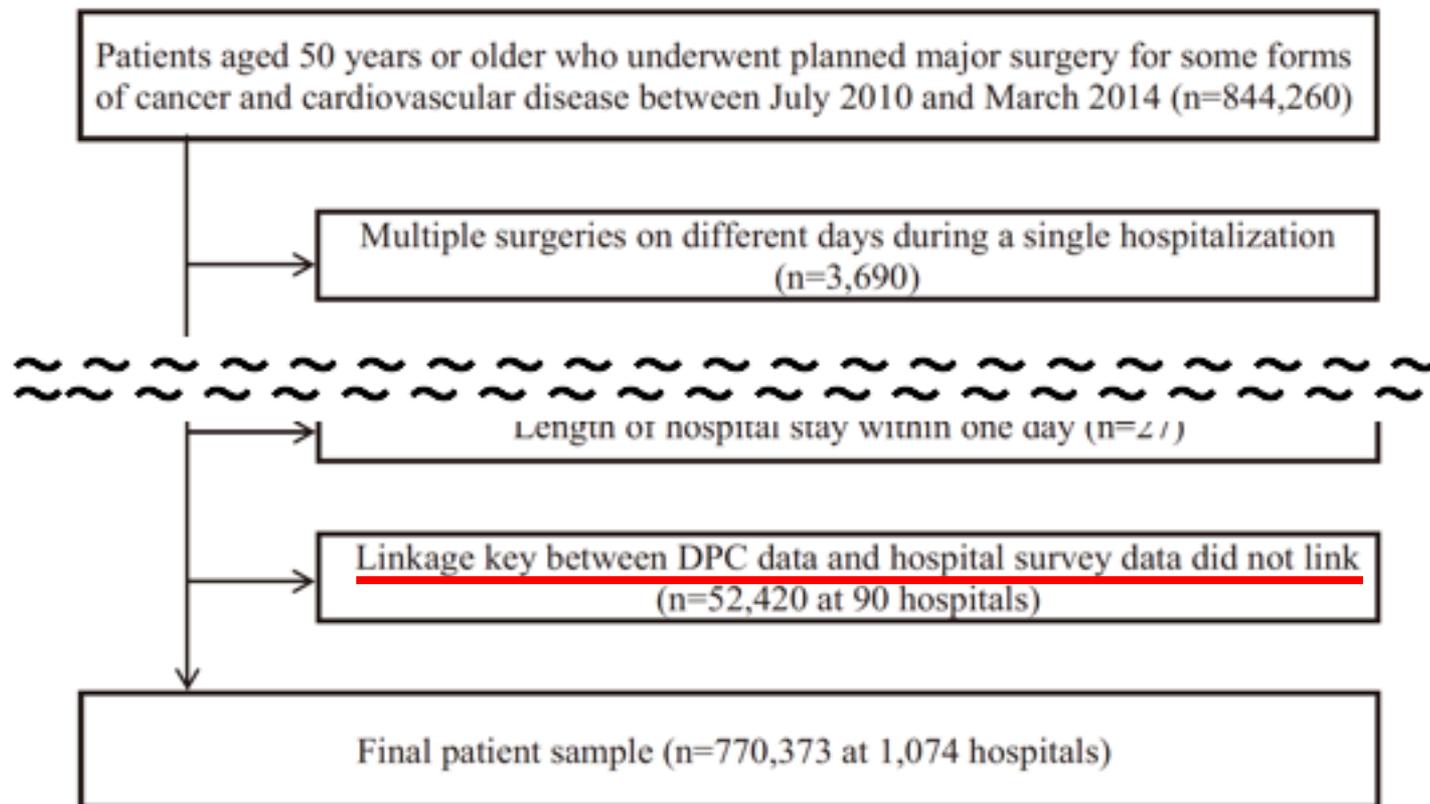
RECORD声明の内容: Methods

STROBE声明		RECORD声明	
参加者 (Participants) つづき			
No.6	<p>(b) マッチング研究の場合 コホート研究では、マッチングの基準、曝露群・非曝露群の各人数を記載する。</p> <p>ケースコントロール研究では、マッチングの基準、ケースあたりのコントロールの人数を記載する。</p>	6.3	<p>もしデータベース同士を結合した研究であれば、データ結合の過程を実証するために、フローダイアグラムまたはその他の図表的表記を用いることを考慮する。</p>

RECORD声明の内容: Methods

8 *HSR: Health Services Research*

Figure 1: Flowchart of Patient Inclusion and Exclusion Criteria



RECORD声明の内容: Methods

STROBE声明

RECORD声明

変数について (Variables)

No.7 全てのアウトカム、曝露、予測因子、交絡因子、効果修復因子を明確に定義する。該当する場合は、診断方法を示す。

7.1 曝露、アウトカム、交絡因子、効果修復因子を分類するために用いたコードやアルゴリズムの完全なリストを提供するべき。

RECORD声明の内容: Methods

RECORD item 7.1

曝露、アウトカム、交絡因子、効果修復因子を分類するために用いたコードやアルゴリズムの完全なリストを提供すべき。

研究の再現性や評価、またほかの研究と比較するために、RECORDの著者らは、研究に用いた全ての診断、手順、薬剤、そのほかのコードのリストを本文および附表、サイトなどで提供することを推奨。

RECORD声明の内容: Methods

MATERIALS AND METHODS

Case Definition

We included all patients with a diagnosis of severe sepsis at admission between July 2010 and March 2013. Severe sepsis was defined as sepsis with failure of at least one organ system. Previous studies estimated the epidemiology of severe sepsis based on ICD-9 Clinical Modification and ICD-10 Australian Modification, using administrative databases (7, 23) In this study, the presence of sepsis was defined as patients who have any bacterial or fungal infection in the admission-precipitating diagnosis using ICD-10 codes used in the previous study (7, 23) (listed in Supplemental Table S1, Supplemental Digital Content 1, <http://links.lww.com/CCM/B742>). To detect a diagnosis of organ dysfunction, ICD-10 codes (7, 23) and Japanese procedure codes listed in Supplemental Table S2 (Supplemental Digital Content 2, <http://links.lww.com/CCM/B743>) were extracted. We

Risks
Patient

Yusuke
Kiyohid
提

目い
axis for

RECORD声明の内容: Methods

STROBE声明		RECORD声明
データソース/測定方法(Data source/Measurement) ※		
No.8	関連する因子に対して、データソース、測定と評価方法の詳細を示す。2つ以上の群がある場合は測定方法の比較可能性を明示する。	-
バイアス (Bias) ※		
No.9	潜在的なバイアスに対処するために行った措置があれば全て明示する。	-
研究サイズ (Study size) ※		
No.10	研究サイズがどのように算出されたか説明する。	-

※RECORDでは、これら項目においてSTROBE声明の拡張はない。

RECORD声明の内容: Methods

STROBE声明		RECORD声明
量的変数(Quantitative variables) ※		
No. 11	量的変数を分析でどのように扱ったか説明する。 該当する場合、集団がなぜ選ばれたのか記載する。	-
統計手法 (Statistical Methods) ※		
No. 12	(a)交絡因子調整に用いた方法を含め、全ての統計学的方法を示す。 (b)サブグループと交互作用の検証に用いた全ての方法を示す。 (c)欠損データをどのように扱ったか説明する。 (d)コホート研究において、脱落例をどのように扱ったか説明する。 ケースコントロール研究において、ケースとコントロールのマッチングをどう行ったか説明する。 横断研究において、サンプリング方式を考慮した分析法について記述する。 (e)あらゆる感度分析の方法を示す。	-

RECORD声明の内容: Methods

STROBE声明	RECORD声明	
データアクセス/データクリーニング方法 (Data access and cleaning methods)		
-	12.1	著者らは、研究対象集団を作成するために用いたデータベースにどの程度アクセス権をもっていたのかを記述すべき。
	12.2	著者らは、研究で用いたデータクリーニング方法の情報を提供すべき。
データ結合(Linkage)		
-	12.3	研究が、個人レベル、施設レベルなのか、また2つ以上のデータベースを結合したデータなのかを明示すること。データ結合の方法は提供すべきである。

RECORD声明の内容: Methods

RECORD item 12.1

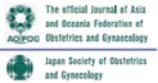
著者らは、研究対象集団を作成するために用いたデータベースにどの程度アクセス権をもっていたのかを記述すべき。

RECORD声明の内容: Methods

RECORD
著者らは
の程度ア

THE JOURNAL OF
Obstetrics and Gynaecology Research

doi:10.1111/jog.13518



J. Obstet. Gynaecol. Res. 2017

Epidemiology of
hospitalization f
exploratory desc
Yasuyuki Okumura a

Cesarean section rates and local resources for perinatal care in Japan: A nationwide ecological study using the national database of health insurance claims

Definition of overdose episode

We identified all hospital admissions from 2012 through September 2013. Over definitive diagnosis of drug poisoning International Classification of Diseases include any overdose caused by

データベースの患者レベルの情報にアクセスできることを示す例

Cesarean section (CS) rates

We applied for permission to use the accumulated data of inpatient insurance claims issued in 2013 from the JNDB, which was subsequently approved by the MHLW. Accumulated data provided included the number of claims with CS codes by prefecture (where

同じデータベースでも集計データのみアクセス権があったことを明記する例

RECORD声明の内容: Methods

RECORD item 12.2

著者らは、研究で用いたデータクリーニング方法の情報を提供するべき。

RECORD声明の内容: Methods

RECORD item 12.2

Low-dose corticosteroid use and mortality in severe community-acquired pneumonia patients

Takashi Tagami^{1,2}, Hiroki Matsui¹, Hiromasa Horiguchi³, Kiyohide Fushimi⁴ and Hideo Yasunaga¹

post-admission complication were not included [28]. The inclusion criteria were: 1) age ≥ 18 years; 2) confirmed diagnosis of CAP (antibiotic therapy initiated on day 0 or 1), not including cases of “suspected pneumonia”; and 3) evidence of severe pneumonia, requiring mechanical ventilation within 7 days of admission [29, 30]. The exclusion criteria were: 1) pregnancy; 2) trauma, haematological malignancy, solid

ICD10コードに対応する日本語テキスト病名に、「疑*」の文字が入力されている患者はデータクリーニングによって除外。

RECORD声明の内容: Methods

RECORD item 12.3

研究が、個人レベル、施設レベルなのか、また2つ以上のデータベースを結合したデータなのかを明示すること。
データ結合の方法は提供すべきである。

データベース間の結合に使った変数の報告以外にも、結合方法(決定的リンケージor 確率的リンケージ)や結合の成功率などに関しても報告を推奨している。

研究前に外部の手によってデータ結合が行われた場合の対応にも言及。

rooms, etc. We linked these survey data to the DPC database using hospital identifiers as the linkage key.

RECORD声明の内容: Results

STROBE声明		RECORD声明	
結果 (Results)			
対象者 (Participants)			
No.1 3	(a)研究の各段階における人数を示す。 (潜在的な適格者数、適格と判断された人数、研究組み入れ人数、追跡完了人数、最終分析人数) (b)各段階での非参加者理由 (c)フローチャート利用を考慮	13.1	研究の対象者選定のプロセスを詳細に記載する。 (これには、データの質に基づくフィルタリング、データの利用可能性と結合を含む) 対象者の選定は本文中とフローダイアグラムの両方またはどちらかにより記載できる。

RECORD声明の内容: Results

STROBE声明		RECORD声明
記述的データ(Descriptive data) ※		
No.14	(a)対象者の特徴と曝露や潜在的交絡因子の情報を示す。 (b)各変数について、データが欠損した対象者数を記載する。 (c)コホート研究において、追跡期間を平均および合計で要約する。	-
アウトカムデータ (Outcome data) ※		
No.15	<ul style="list-style-type: none">● コホート研究において、アウトカムの発生数やsummary measureの数値を経時的に示す。● ケースコントロール研究において、各曝露カテゴリーの数、または曝露のsummary measuresを示す。● 横断研究において、アウトカムの発生数やsummary measureの数値を示す。	-

RECORD声明の内容: Results

STROBE声明		RECORD声明
主な結果(Main results) ※		
No.16	<p>(a)調整前の推定値と、該当する場合は交絡因子での調整後の推定値、そしてその精度を記述する。 どの交絡因子がなぜ調整されたかを明確にする。</p> <p>(b)連続変数がカテゴリー化されているときは、 カテゴリー境界を報告する。</p> <p>(c)相対リスクを意味を持つ期間の絶対リスクに 変換することを考慮する。</p>	-
その他の解析 (Other analysis) ※		
No.17	その他に行われたすべての分析の結果を報告する。	-

RECORD声明の内容: Discussion

STROBE声明

RECORD声明

考察 (Discussion)

極めて重要な結果(Key results) ※

No.18	研究目的に関する極めて重要な結果について要約する。	-
-------	---------------------------	---

限界(Limitations)

No.19	<p>潜在的なバイアスや精度の問題を考慮して研究の限界を議論する。</p> <p>潜在的なバイアスの方向性と大きさを議論する。</p>	19.1	<p>特定のリサーチクエスチョンに答えるために作成または収集されたわけではないデータベースを用いるimplicationについて議論すること。</p> <p>誤分類によるバイアス、未測定の変因、交絡因子、欠損データ、時間によって変化する適格性についての議論を含めること。</p>
-------	---	------	---

RECORD声明の内容: Discussion

RECORD item 19.1

特定のリサーチクエスチョンに答えるために作成または収集されたわけではないデータベースを用いるimplicationについて議論すること。

誤分類によるバイアス、未測定の交絡因子、欠損データ、時間によって変化する適格性についての議論を含めること。

声明内では、潜在的なバイアスの存在は通常の観察研究と同様であるが、データベースを用いた観察研究に特有の問題に言及するように求めている。

- データの質はどうか？
- Administrative dataの場合、コーディングエラーの可能性は？
- 観察期間中に用いられるコードの変更はない？
(例: ICD9→ICD10、診療報酬内容の変更)

RECORD声明の内容: Discussion

RECORD item 19.1

特定のリサーチクエスチョンに答えるために作成または収集

outcomes. The third limitation is related to the use of an administrative database. The Diagnosis Procedure Combination database is an administrative claim database, which is similar in some respects to the Nationwide Inpatient Sample in the United States.²⁶ Generally, recorded diagnoses in such databases are less well validated than those in planned prospective cohorts or registries. The definition of complications was not standardized. However, there are several advantages of the data submission processes used in the Diagnosis Procedure Combination database, including diagnoses reported by attending physicians, the requirement of the data entry through a well-specified reporting format, and mandatory submission that has been linked with reimbursement. Fourth, the generalizability of our

- 観察期間中に用いられるコードの変更はない？
(例: ICD9→ICD10、診療報酬内容の変更)

RECORD声明の内容: Discussion

STROBE声明		RECORD声明
解釈(Interpretation) ※		
No.20	目的、限界、解析の多重性、同様の研究で得られた結果やその他の関連するエビデンスを考慮し、慎重で総合的な結果の解釈を記載する。	-
一般化可能性 (Generalizability) ※		
No.21	結果の一般化可能性(外的妥当性)を議論する。	-

RECORD声明の内容: Other information

STROBE声明		RECORD声明	
その他の情報 (Other information)			
財源(Funding) ※			
No.22	研究の資金源、本研究における資金提供者の役割を示す。該当する場合、現在の研究の元となる研究についても同様に示す。		-
プログラムコード・生データ・プロトコルへのアクセス可能性 (Accessibility of protocol, raw data, and programming code)			
	-	22.1	論文著者らは、研究プロトコル、生データ、またはプログラムコードのような追加情報についてどのようにアクセスできるかに関する情報を提供すべきである。

RECORD声明の内容: Other information

RECORD item 22.1

論文著者らは、研究プロトコル、生データ、またはプログラムコードのような追加情報についてどのようにアクセスできるかに関する情報を提供すべきである。

- RECORD声明内では、ハードルが高いことは認めつつ、研究プロトコル、生データの結果、プログラミングコードの公開を推奨するというものになっている。
- 論文の追加資料や、個人・所属機関のウェブサイト、ソーシャルメディア、データレポジトリ、政府の公開データサイトなどを用いての公表を推奨している。

データベース研究の報告内容の現状

A review of routinely collected data studies in urology: Methodological considerations, reporting quality, and future directions

Blayne Welk, MD¹; Justin Kwong, MD²

泌尿器科領域のleading medical journalである
Journal of Urology, European Urology から
2014年に報告されたデータベース研究全56編の
報告の現状について、RECORD声明をどれくらい満たし
ているか調査

Can Urol Assoc J. 2017;11(3-4)

データベース研究の報告内容の現状

Table 1. Proportion of the RECORD items reported in RCD studies from the *Journal of Urology* and *European Urology* in 2014. Each RECORD item is numbered based on the complementary section from the STROBE guidelines

RECORD item	Description	Percentage
1.1	The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included.	83.9%
1.2	If applicable, the time frame within which the study took place should be reported in the title or abstract.	94.6%
12.1	Authors should describe the extent to which the investigators had access to the database population used to create the study population.	87.5%
13.1	Describe in detail the selection of the persons included in the study (i.e., study population selection), including filtering based on data quality, data availability, and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	85.7%
19.1	Discuss the implications of using data that were not created or collected to answer the specific research question(s).	
	c) Include discussion of missing data as they pertain to the study being reported.	82.1%

データソース名の明示(1.1)、調査期間(1.2)、データのアクセス権の範囲(12.1)、対象者選定のプロセスの詳細(13.1)、欠損データについて(19.1)は、約80%超の論文において報告されていた。

データベース研究の報告内容の現状

Table 1. Proportion of the RECORD items reported in RCD studies from the *Journal of Urology* and *European Urology* in 2014. Each RECORD item is numbered based on the complementary section from the STROBE guidelines

RECORD item	Description	Percentage
6.2	Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.	21.4%
6.3	If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	0%
12.2	Authors should provide information on the data-cleaning methods used in the study.	1.8%
19.1	Discuss the implications of using data that were not created or collected to answer the specific research question(s). d) Include discussion of changing eligibility over time as they pertain to the study being reported.	14.3%

用いられたコードの妥当性について(6.2)、データ結合の詳細(6.3)、データクリーニング(12.2)、適格性の経時的変化(19.1)を報告している研究は少なかった。(25%未満)

データベース研究の報告内容の現状

1.2	If applicable, the geographic region within which the study took place should be reported in the title or abstract.	48.2%
1.3	If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	30.8%
6.1	The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided.	60.7%
7.1	A complete list of codes and algorithms used to classify key variables should be provided. If these cannot be reported, an explanation should be provided.	
	a) Exposures	62.5%
	b) Outcomes	46.4%
	c) Confounders and effect modifiers	16.1%
12.3	State whether the study included person level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	64.3%
19.1	Discuss the implications of using data that were not created or collected to answer the specific research question(s).	
	a) Include discussion of misclassification bias as they pertain to the study being reported.	46.4%
	b) Include discussion of unmeasured confounding as they pertain to the study being reported.	73.2%
22.1	Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	50.0%

地理・期間 (1.2)、データ結合に関する抄録記載(1.3)、対象集団の選定方法(6.1)、曝露、アウトカム、交絡因子の定義(7.1)、データ結合の報告(12.3)、データベースを用いるimplication (19.1)、プロトコル、生データの結果、プログラミングコードの公開(22.1)の項目は、30-70%で報告。

RECORD 声明利用の現状

翻訳の状況

- ドイツ語、中国語、日本語(チェックリストのみ)に翻訳済み

科学雑誌での承認状況

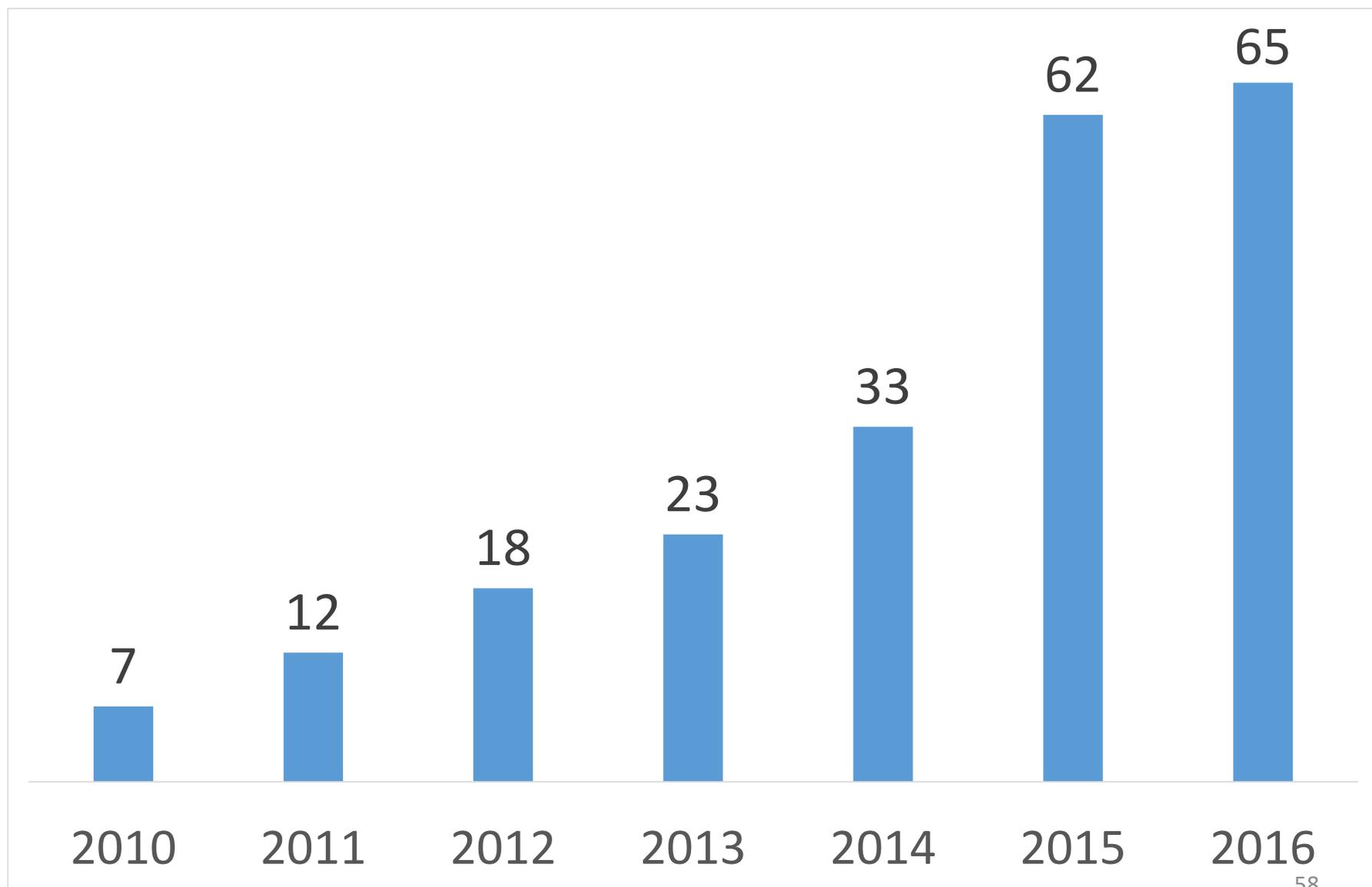
- 15雑誌 (2016年6月)
→21雑誌 (2018年1月現在)

Clin Epidemiol. 2016;8:389-392.

まとめ

- 日々収集される診療情報(主にデータベース)を用いる研究の質を向上させるために作成。
- STROBEに13の項目を追加する形で完成し、2015年公表。(PLoS Med. 2015;12(10):e1001885.)
- データベースを用いた観察研究においては、STROBE声明とRECORD声明の両方を参照することを推奨。

D P C データベース研究 英文原著論文数



Take Home Message

若手研究者大募集！

DPCデータを用いて臨床研究・疫学研究をやってみたい
意欲のある若手の研究者諸君、
康永(yasunagah-ky@umin.ac.jp)に連絡してください。

<条件>
東京大学まで来られること。

研究開始にあたって必要なスキル

ありません。

やる気と根気と折れない心があればOK。
スキルはこれから磨いてください。

注意！！！！

データ容量はテラ・バイト級です

「すべてのデータを取りあえずざっと眺めてから研究テーマを考える」という考えは、金輪際すててください。そんなことは不可能です。

- (1) まずデータベースに存在するデータ項目をよく理解してください。
- (2) 次に、リサーチ・クエスチョンを立てて、入手可能なデータで検証可能かをよく検討し、研究計画書を書いてください。
- (3) 次に、研究の目的・方法に基づいて、データベースから必要なデータを抽出します。

ご清聴ありがとうございました