

一般口演

## 一般口演1

## 病院情報システム1（電子カルテ・文書管理）

2018年11月23日(金) 10:15～11:45 E会場 (5F 501)

## [2-E-1-1] 電子カルテ記載量予測モデル創出を目指した電子カルテ年間記載文字数の分析

○藤田 健一郎<sup>1</sup>, 杉山 治<sup>1</sup>, 平木 秀輔<sup>1</sup>, 岡本 和也<sup>1</sup>, 竹村 匡正<sup>2</sup>, 黒田 知宏<sup>1</sup> (1.京都大学医学部附属病院, 2.兵庫県立大学大学院応用情報科学研究科)

情報システムを設計するに当たり、データ量の見積はその基礎となるものである。電子カルテから地域医療連携基盤、そして全国規模の医療基盤へと医療情報システムの規模が拡大していく中で、電子カルテのデータ量を予測することは、単に必要なストレージを見積もるために必要なだけでなく、表示領域のサイズやデザイン、データの読み込み速度など、電子カルテアプリケーションの User Experience(UX)における前提条件として本来不可欠なものである。そのためには、電子カルテ記載量の予測モデルを構築する必要があると考えられる。そこで、本研究では、予測モデルを創出するための予備的な調査として、電子カルテの記載文字数に関する分析を行った。具体的には、ある大学病院の電子カルテについて、2006年～2016年の11年間における患者毎の年間記載文字数を、外来、入院及び全体のそれぞれについて計測した。その結果、患者毎の電子カルテ年間記載文字数は、外来、入院及び全体のいずれにおいても概ね対数正規分布に従うことが明らかになった。また、記載文字数は継続的に増加していることも明らかになった。次に、予測モデルにおける説明変数の候補として、患者毎の、年間外来受診日数、年間入院日数及び前年の記載文字数について、電子カルテ年間記載文字数との相関関係を検討した。その結果、或る程度の相関はあるものの、いずれも説明変数としては不十分であった。本研究の結果、非常に記載量の多いカルテが存在すること、記載量そのものが増加していることが明らかになった。これらは、電子カルテアプリケーションが取り扱うべきデータ量の前提条件としては重要なものと考えられる。今後は、計測を年単位ではなくより短い月単位、日単位又は入院単位等にした分析を試み、予測モデルの検討を行う。

# 電子カルテ記載量予測モデル創出を目指した電子カルテ年間記載文字数の分析

藤田 健一郎<sup>\*1\*</sup>、杉山 治<sup>\*1</sup>、平木 秀輔<sup>\*1</sup>、岡本 和也<sup>\*1</sup>、竹村 匡正<sup>\*2</sup>、黒田 知宏<sup>\*1</sup>

\*1 京都大学医学部附属病院、\*2 兵庫県立大学大学院応用情報科学研究科

## Analysis for a predictive model for an annual quantity of description of electronic medical records

Kenichiro Fujita<sup>\*1\*</sup>, Osamu Sugiyama<sup>\*1</sup>, Shusuke Hiragi<sup>\*1</sup>, Kazuya Okamoto<sup>\*1</sup>,  
Tadamasa Takemura<sup>\*2</sup>, Tomohiro Kuroda<sup>\*1</sup>

\*1 Kyoto University Hospital, \*2 Graduate School of Applied Informatics, University of Hyogo

**Abstract:** Despite the popularization of electronic medical records, the amount of description of electronic medical records and its changes over time are not clear. In designing an electronic medical records system, prediction of the amount of description is important. As a preliminary survey to create a prediction model, we analyzed the number of characters described in the electronic medical records and examined candidates for explanatory variables in the prediction model. As a result of the analysis, it became clear that the annual quantity of electronic medical records follows the lognormal distribution, and the amount of description is increasing. Although we could not find a strong candidate for explanatory variables, it suggested the possibility of using disease name or billing information. It also suggested the need to analyze the temporal differences in more microscopic units.

**Keywords:** Electronic Medical Records, Prediction Model, Medical Records.

### 1. 背景

カルテを中心とする診療記録の電子化が進んでおり、400床以上の医療機関への電子カルテ導入率は70%に達している<sup>1)</sup>。医師法上の診療録の保存年限は5年<sup>2)</sup>であるが、特定生物由来製品の使用記録の保存期間が20年であること<sup>3)</sup>、医療安全管理上過去の記録を参照する必要が生じること、電子化することにより保存場所の問題が緩和されることなどから、電子化されたカルテの保存期間としては20年以上が想定される。従って、電子カルテのデータは非常に長期間に渡って蓄積されていくと考えられる。

一般に情報システムの設計に際して、保存すべきデータ量は基礎的な指標と考えられる。しかしながら、紙カルテの記載量についての調査・解析<sup>4)</sup>や、電子カルテ化により外来受診毎の記載量が増加するという報告<sup>5)</sup>はあるものの、入院・外来を含めた電子カルテ全体の記載量や、電子カルテ化以後の経時的変化については明らかになっていない。

### 2. 目的

電子カルテから地域医療連携基盤、そして全国規模の医療基盤へと医療情報システムの規模が拡大していく中で、電子カルテのデータ量を予測することは、単に必要なストレージを見積もるために必要なだけでなく、表示領域のサイズやデザイン、データの読み込み速度など、電子カルテアプリケーションの User Experience (UX) における前提条件として本来不可欠なものである。加えて、電子カルテの記載量は記載を行う医療従事者の業務量にも影響する。

電子カルテのデータ量を予測するためには、電子カルテ記載量の予測モデルを構築する必要があると考えられる。そこで、本研究では、予測モデルを創出するための予備的な調査として、電子カルテの記載文字数に関する分析を行った。

### 3. 方法

#### 3.1. カルテ記載量の調査

京都大学医学部附属病院(病床数 1,121 床、一日平均外来患者数 2,834 人、以下対象病院)に導入されている電子カ

ルテを対象システムとした。対象病院では 2005 年 4 月に電子カルテシステムを導入し、2011 年 1 月及び 2016 年 4 月にシステムの更新を実施している。調査対象は 2006 年から 2016 年までに記載された電子カルテとした。また、カルテの各記事における患者の状態としては、外来診療及び入院診療の 2 種類に大きく分類される。

そこで、患者カルテを記載年毎に入院及び外来に分類して記載量を計測した。記載量の単位としては文字数を用い、全角・半角のいずれも 1 文字としてカウントした。以下では、1 患者の 1 年間のカルテ記載文字数を「患者別年間記載量」とする。患者別年間記載量の集計イメージを図 1 に示す。

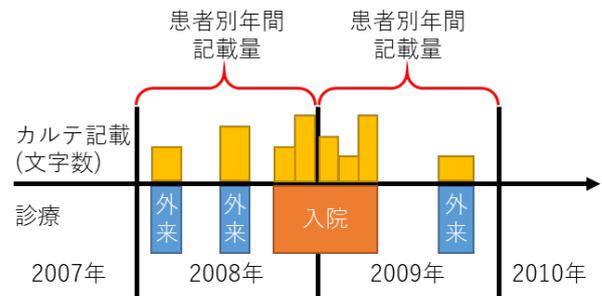


図 1 患者別年間記載量の集計イメージ

#### 3.2. 説明変数候補の検討

次に、予測モデルにおける説明変数の候補を検討した。カルテの記載は診療に伴って発生することから、外来については受診日数、入院については入院日数に比例して記載量が増加すると予想される。また、記載量に疾患や診療科による違いがあっても、患者の疾患や受診診療科が大きく変化しないと仮定すると、前年の記載量から大きく変化することは少ないと推測することが可能である。そこで、説明変数の候補として、年間受診日数、年間入院日数及び前年の記載量について、患者別年間記載量との相関関係を調査した。

## 4. 結果

### 4.1. 全期間の特徴

対象期間全体における患者別年間記載量の集計結果を表1に、患者別年間記載量(10万文字以下)のヒストグラムを図2に、患者別年間記載量の対数ヒストグラムを図3に示す。

表1より、入院と外来において、件数及び記載量の両方に顕著な差が見られる。また、最大値は外来でも約84万文字、全体及び入院では約820万文字であり、100万文字以上のカルテが1,000件以上存在している。

次に、分布について、図2より、患者別年間記載量は正規分布ではないことが予想されるが、図3より、患者別年間記載量は対数正規分布に従うことが予想される。そこで、患者別年間記載量の常用対数に対して、正規分布に対する適合度検定である、コルモゴロフ-スミルノフ検定及びシャピロ-ウィルク検定を行ったところ、いずれの検定においても、全体、外来、入院の全てにおいて有意水準5%で帰無仮説が留保され、患者別年間記載量は対数正規分布に従うことが示された。

表1 対象期間全体における患者別年間記載量

	全体	外来	入院
件数	883,312	881,770	151,600
平均値	25,647	10,401	88,940
中央値	4,285	3,833	28,342
最大値	8,209,843	839,453	8,203,040
100万文字以上の件数	1,577	0	1,454
対数の平均値	8.412	8.167	10.227
対数の中央値	8.363	8.251	10.252
対数のコルモゴロフ-スミルノフ検定	0.772	0.2209	0.8638
対数のシャピロ-ウィルク検定	0.4889	0.6805	0.477

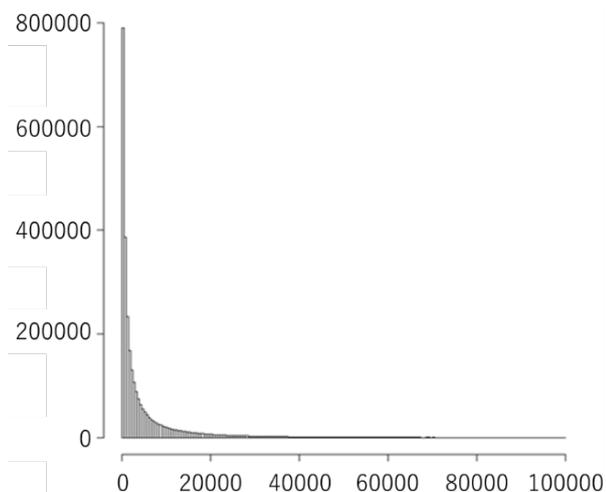


図2 患者別年間記載量(10万文字以下)のヒストグラム

### 4.2. 経時的変化

対象期間の各年毎の患者別年間記載量の平均値及び中央値の推移を図3,4に示す。全体、外来、入院のいずれについても、平均値と中央値の双方で患者別年間記載量が増加し続けている。

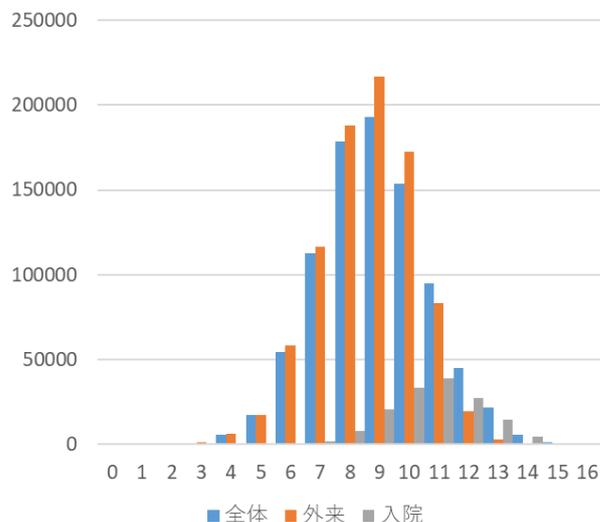


図3 患者別年間記載量の対数ヒストグラム

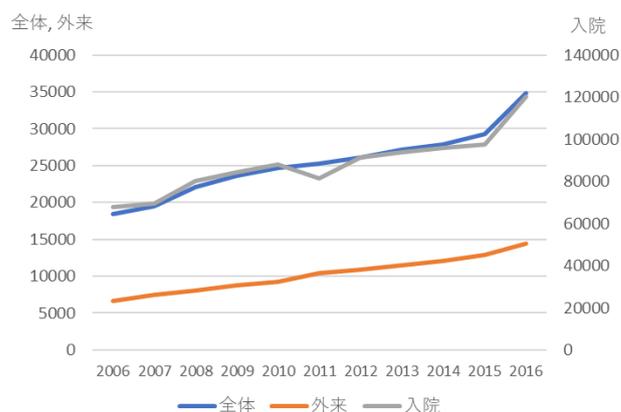


図4 患者別年間記載量の平均値推移

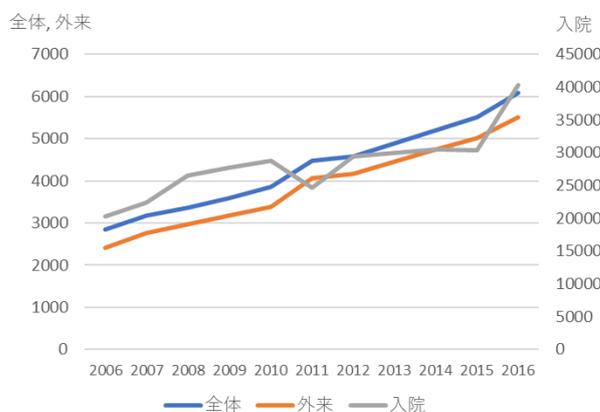


図5 患者別年間記載量中央値推移

### 4.3. 説明変数候補

次に、患者毎の年間受診日数、年間入院日数及び前年の記載文字数と、患者別年間記載量の相関関係について検討を行った。外来の患者別年間記載量と、患者の年間受診日数の散布図を図6に、入院の患者別年間記載量(100万字以下)と、患者の年間入院日数の散布図を図7に、前年の記載量が0文字の患者を除いた患者別年間記載量(100万字以下)と前年の患者別年間記載量との散布図を図8に示す。

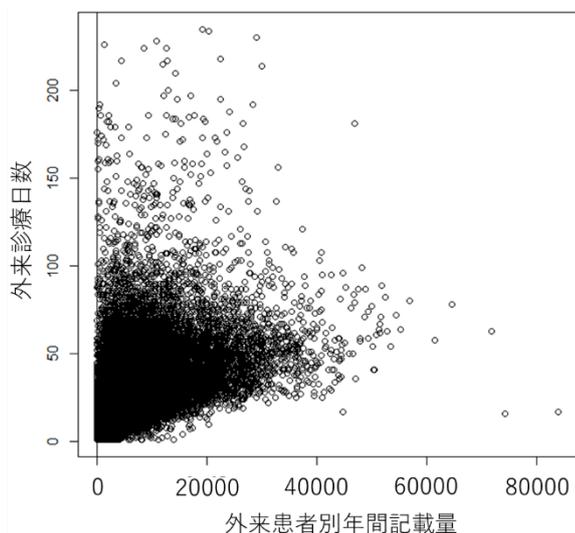


図6 外来の患者別年間記載量と年間受診日数の散布図

受診日数と外来カルテの患者別記載量の関係について、図6より、正の相関が示唆される。相関係数を求めたところ、0.674となり、0.5%水準で有意であった。他方で、診療日数が多いにもかかわらず年間記載量が必ずしも多くないケースが多く存在している。

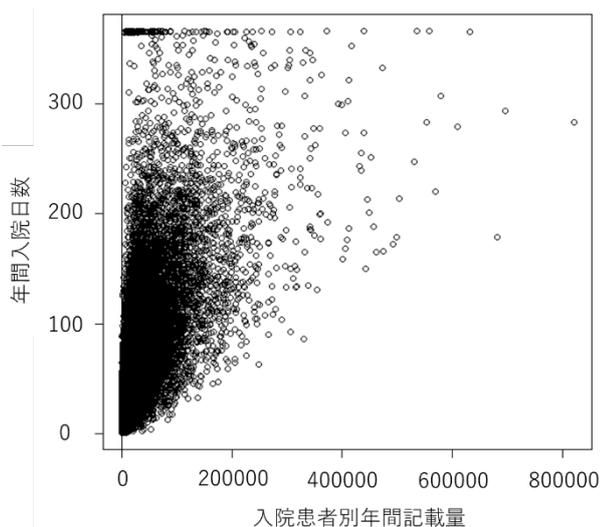


図7 入院の患者別年間記載量と年間入院日数の散布図

入院日数と入院カルテの患者別年間記載量の関係について、図7より、正の相関が示唆される。相関係数を求めたところ、0.718となり、0.5%水準で有意であった。受診日数と同様、入院日数についても入院日数が長いにもかかわらず記載量

が多くないケースが多く存在している。

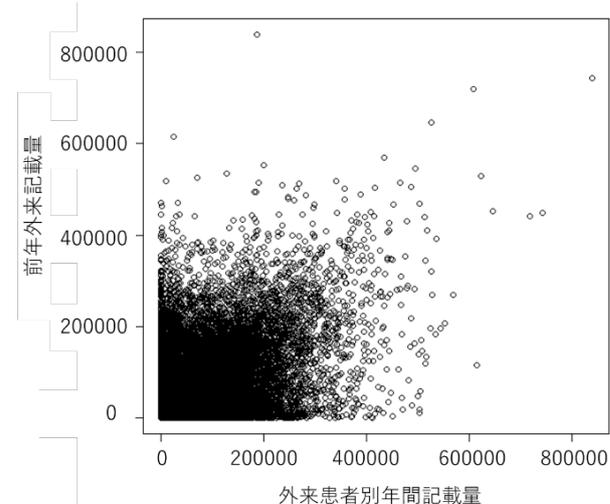


図8 患者別年間記載量(100万字以下)の前年との散布図

前年の記載量と患者別年間記載量の関係について、図8より、弱い正の相関が示唆される。相関係数を求めたところ、0.693となり、0.5%水準で有意であった。しかしながら、昨年の記載量との差が大きい患者も多く存在している。

## 5. 考察

### 5.1. 患者別年間記載量の特徴

患者別年間記載量は、対数正規分布に従うことが明らかになった。これは、紙カルテを対象とした先行研究の結果と一致している<sup>4)</sup>。

他方で、多くのカルテでは年間数千~数万文字が入力されているが、いくつかのカルテについては年間数百万文字の入力が発生している。紙カルテにおける記載文字数については、先行研究の計測単位がページであることもあり未調査であるものの、年間の記載量としては数万文字が限度であろうと考えられる。

電子カルテの記載量が非常に大きくなる理由としては、以下の3つが考えられる。

- 1 処方や検査・処置等の指示内容は、オーダーリング機能から自動的に転記されるため、記載作業が省力化されている。
- 2 コピー&ペーストが可能であるため、過去の記録をそのまま引用し、直近の内容を追記する記載スタイルが生じている。
- 3 入力がどの端末からでも可能であり、看護師やコメディカルによる記載が増加している。

カルテの記載量の増大は、情報システム化によるUXの変化とも相まって、カルテの読み方にも影響を与えている可能性もある。

情報システムの観点からは、数千~数万文字のカルテと数百万文字のカルテを同程度のパフォーマンスで処理することはかなり困難であり、正常に処理できない可能性もある。電子カルテシステム的设计においては、記載量に大きな幅があること、記載量が増加し続けることに十分留意する必要があると考えられる。

## 5.2. 患者別年間記載量の経時的変化

患者別年間記載量は、入外のいずれについても増加傾向にある。この理由としては、以下の2つが考えられる。

- 1 オーダリングシステムや電子カルテシステムの導入以降も、ある程度の紙運用は残存しており、それらのシステムへの移行は徐々に進行している。それらの移行分が5.1に挙げた1の理由から自動転記に追加されている可能性がある。
- 2 同じく5.1に挙げた2の理由から、コピー&ペーストを繰り返すことで記載量が徐々に増加している可能性がある。

化により外来カルテ文章の情報量は増加する傾向がある。医療情報学. 2010; 30(Suppl.): 1046-1049. 2010.11.21, 浜松.

## 5.3. 説明変数候補について

患者別年間記載量の説明変数の候補として、年間受診日数、年間入院日数及び前年の記載量を検討したが、ある程度の相関関係にはあるものの、説明変数としては十分ではないと考えられる。

カルテは診療内容及び診療に必要な内容を記載するものであるから、診療の回数としての受診日数及び入院日数にある程度従属すると考えられるが、より詳細には診療行為の件数に従属すると考えられる。他方、個々の患者における、疾患・病態の困難度・複雑度・稀少度にも従属すると考えられるが、単純に同じ患者の過去の記載量と線形な関係にあるわけではないと考えられる。これらの診療の量的側面及び複雑性の側面は、疾患の分類及び診療報酬請求に反映されると考えられる。従って、DPCコードなどの病名・請求情報が説明変数として利用可能である可能性がある。

また、5.1及び5.2で検討したように、過去の記録のコピー&ペーストに加筆するという記載スタイルが存在すると仮定すると、年単位の記載量ではなく、日単位又は記事単位での経時的な差分についてのマイクロな分析が必要と考えられる。

## 6. 結論

本研究では、電子カルテの年間記載量の特徴及び経時的変化について分析し、併せて記載量の予測モデルの説明変数候補について検討を行った。分析の結果、電子カルテの年間記載量は対数正規分布に従うこと、記載量が増加していることが明らかになった。また、入院患者については年間百万文字以上の記載が発生するケースが一定数存在することが明らかになった。

説明変数については有力な候補を見いだすことができなかったが、病名・請求情報の利用可能性が示唆された。また、よりマイクロな単位での経時的な差分の分析の必要性が示唆された。

## 参考文献

- 1) 調査委員会. オーダリング・電子カルテシステム病院導入状況調査報告書[2015年(H27年)調査版]. 保健医療福祉情報システム工業会, 2015. [[https://www.jahis.jp/files/user/\(印刷版\)2015導入調査報告書.pdf](https://www.jahis.jp/files/user/(印刷版)2015導入調査報告書.pdf) (cited 2017-July-10)].
- 2) 医師法. 昭和23年7月30日法律第201号, 2014.
- 3) 医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律施行規則(昭和36年2月1日厚生省令第1号). 厚生労働省, 2018.
- 4) 欧陽寧, 池田充, 徐知行, 山内一信. 診療録における記載量についての解析. 診療録管理 1996; 8(1): 41-43.
- 5) 荒牧英治, 増川佐知子, 山田恵美子, 脇嘉代, 大江和彦. 電子