

# 研究実施のお知らせ

作成日：2019年10月25日 ver.1.1

更新日：2023年12月20日 ver.1.3

## 研究課題名

機械学習による退院時要約自動分類システムに関する研究

## 研究の対象となる方

2014年4月から2019年3月の間に、島根大学医学部附属病院に入院され、包括医療費支払い制度方式が採用された方

\* 包括医療支払制度方式が採用されたかどうかは、診療明細書に記載されておりますので、ご参照ください。

## 研究の目的・意義

病院情報システムに蓄積された退院時要約および関連する診療情報のテキストデータに対して、機械学習の手法を適用し、適用症例のDPC(Diagnosis Procedure Combination；診断群分類)※を推定するシステムを開発します。機械学習とは、コンピュータがデータから反復的に学習し、そこに潜むパターンを見つけ出すことです。そして学習した結果を新たなデータにあてはめることで、パターンにしたがって推定することができます。

システムによって正確な推定が行われれば、システムの推定と実際に付与されたDPCとに相違がある場合に、そのDPCが妥当か、あるいは退院時サマリの記述に問題があるかを検討するための指針とできます。たとえば、システムが90%の正答率を上げれば、日に約40例、月に約1200例の退院時要約をはじめとする診療録をチェックしなければならない診療情報管理士の仕事を10%、120例程度にまで削減でき、誤判別例の丁寧な検討を行えます。診療情報管理士の業務は、病歴情報の管理、医師を含め、メディカルスタッフの診療情報の記載をチェックし、内容が正確で、診療行為が十分理解できる診療録を作成できるように、管理することが本来の業務です。しかしながら、現在、診療情報管理士は、1ヶ月毎診療報酬の提出前に、DPCのチェックをしなければならず、これが業務上大きな負担になっています。本研究でDPCを推定するシステムが開発できれば、この業務の負荷が軽減できる可能性があります。当院の診療情報管理士が5人であることを考えれば、本システムの開発の診療情報業務に関する寄与は高いです。

※ 日本における医療費の定額支払い制度に使われる評価方法である診断群分類包括評価では、入院患者の診断群分類(病名)ごとに1日当たりの入院費が決められており、これをもとに計算を行います。診断群分類とは、約500種類の主な疾患(病名)を基本として手術・処置・副傷病名の有無などにより、さらに約5000種類に分類したものです。入院患者は、例外を除いて、この約5000種類のうちのどれか1つに該当することにな

ります。

## 研究の方法

### • 解析の流れ

退院時要約および付随のデータを病院情報システムから抽出します。その退院時要約に患者に適用された DPC のラベルを付与し、テキストデータを単語単位に分解。単語の頻度を集計の上、DPC×単語の形式で集計した頻度表を生成します。その頻度表データに対して対応分析を適用します。対応分析は分析対象である DPC と分析の基準となる単語の関連性を示すことができ、これにより関連性が高い順に単語の順位付けが可能となります。この順位に基づいて解析に用いる単語を選択し、機械学習の手法を適用させ、単語から DPC を推定するシステムを構築します。

### • 評価方法

まず、退院時要約のテキストデータを訓練データとテストデータに分けます。訓練データに本研究で開発した上記の手法を適用して学習させ、テストデータを用いて、正答率(誤判別率)および間違っただけの症例についてデータ出力します。さまざまな機械学習プログラムを使い、プログラムの正答率を評価します。また、各手法で答えが間違っただけの例について、その原因を症例に関する診療情報から検討し、より良い方法の開発に役立てます。

### • 情報の管理

収集したデータは、各研究機関内において外部から容易にアクセスできないサーバに保管します。サーバは、常時施錠されているサーバ室に設置されており、サーバには、セキュリティを設定し、パスワードで使用可能な研究者を制限し、アクセスログを取得します。

共同研究機関へのデータの提供に際しては、匿名化した状態で、さらにファイルに暗号化処理を施します。ファイルは DVD-R ディスクに保存し、郵送にて提供します。

研究に関するデータ及び関連資料は研究の終了を報告してから少なくとも 10 年間保管し、その後匿名化した状態で廃棄（消去）します。

## 研究の期間

2019 年 11 月 22 日～2026 年 3 月

## 研究の公表

この研究から得られた結果は、医学関係の学会や医学雑誌などで公表します。その際にあなたのお名前など個人を識別できる情報を使用することはありません。

## 研究組織

この研究は次の機関が共同で行います。

研究代表者：

島根大学医学部医療情報学講座 津本 周作

共同研究機関

[研究機関]

島根大学医学部附属病院

東京大学

北海道大学

[研究責任者]

津本 周作

和泉 潔

坂地 泰紀

[研究機関の長]

石原 俊治

藤井 輝夫

寶金 清博

### 情報の利用停止

ご自身の情報をこの研究に利用してほしくない場合には、ご本人または代理人の方からお申し出いただければ利用を停止することができます。

なお、利用停止のお申し出は、2023年9月までをお願いいたします。それ以降は解析・結果の公表を行うため、情報の一部を削除することができず、ご要望に沿えないことがあります。

### 相談・連絡先

この研究について、詳しいことをお知りになりたい方、ご自身の情報を研究に利用してほしくない方、その他ご質問のある方は次の担当者にご連絡ください。

研究代表者：津本 周作

島根大学医学部医学科医療情報学講座

〒693-8501 島根県出雲市塩冶町 89-1

電話 0853-20-2171

研究責任者：和泉 潔

東京大学大学院工学系研究科

〒113-8656 東京都文京区本郷 7-3-1 東京大学工学部 8号館 530室

電話 03-5841-1867

研究責任者：坂地 泰紀

北海道大学大学院情報科学研究院

〒060-0814 北海道札幌市北区北 14 条西 9 丁目 北海道大学情報科学研究院棟 8-02 室

電話 011-706-6945