

厚生労働科学研究費補助金(臨床研究等ICT基盤構築・人工知能実装研究事業)

分担研究報告書

機械学習による予後予測法の開発

研究分担者 馬込大貴 駒澤大学 講師

研究要旨：2005/7/1 から2007/6/30までに登録されたJPOPSコホート1症例を使用し、機械学習技術に基づき、前立腺癌の予後を予測する研究を開始した。初期解析として、JPOPS定義、Phoenix定義のPSA再発の有無、急性期と晩期の尿路系有害事象、直腸有害事象を複数の機械学習手法を用いて予測し、結果を比較した。JPOPSで収集されたデータは世界的に類をみない良質なビッグデータであり、様々な情報を組み合わせて使用することの重要性が示唆された。

## A．研究目的

ヨウ素125シード線源を用いた小線源療法に関する前向きコホート研究(JPOPS, Japanese Prostate Cancer Outcome Study of Permanent I-125 seed Implantation)で収集されたデータは、世界的に類をみない良質なビッグデータである。本研究の目的は、JPOPS研究によって得られたビッグデータを用いて、機械学習手法に基づく前立腺癌の予後予測システムを開発することである。

## B．研究方法

2017年7月28日に臨床研究情報センターにて、研究代表者中村和正、研究分担者馬込大貴と、JPOPS研究の登録データセットをどのように取り扱うか、特許等を取得できた場合の契約等について討議した。

2017年10月19日、2018年1月20日の班会議にて、研究分担者および研究協力者に、本研究の概要、今後の研究計画について、討議を行った。

初期解析として、JPOPSコホート1の解析対象元データ(2339症例)から解析用に2種類のデータセット(Limited datasetとLarge dataset)を作成した。Limited datasetは臨床的に予後との関連があると考えられるデータ項目のみを含み、Large datasetは可能な限り多くのデータ項目を使用した。機械学習手法として、ロジスティック回帰、サポートベクターマシン、ランダムフォレスト、ニューラルネットワークの4種類を使用した。JPOPS定義、Phoenix定義のPSA再発の有無、急性期と晩期の尿路系有害事象、直腸有害事象を予測し、結果を比較した。

さらに、2018年3月28日に臨床研究情報センターにて、研究代表者中村和正、研究分担者馬込大貴と現在までの機械学習による予後予測結果の精度について、討議を行った。

(倫理面への配慮)

本研究はすでにJPOPSで登録され、匿名化された既存データのみを用いる観察研究であり、患者への侵襲は伴わない。

JPOPS研究のコホート1およびコホート2のデータセット原本については臨床研究情報センターにおいて厳重に管理されている。また、駒澤大学へのデータの移送においては、フォルダにパスワードにて暗号化した。

### C . 研究結果

浜松医科大学、駒澤大学にて、本解析について倫理審査で承認を得た後に、匿名化されたJPOPSデータをセキュリティを確保した上で送付し、駒澤大学にて、PSA再発、および有害事象発生の有無の予測解析を開始した。

初期解析の結果として、Limited DatasetよりもLarge datasetを用いた方が予測精度が高い傾向があった。また、機械学習手法としてはランダムフォレストを用いたモデルの予測精度が最も高かった。例として、JPOPS定義のPSA再発予測ではランダムフォレストに基づく予測精度が最も高く、Limited datasetで0.892、Large datasetで0.938であった。

### D . 考察

臨床的に予後との関連があると考えられるデータ項目だけを用いるのではなく、予後と関係しないと考えられているようなデータ項目を含めることで、予測精度が向上した。JPOPSは非常に多くのデータ項目かつ大規模な症例数を有する良質なビッグデータであり、様々な種類の情報を収集することの重要性が示唆された。次年度以降は、本年度に得られた初期結果を研究代表者および他の研究分担者と深く議論し、臨床的に役立つ予測システムの

開発を行う予定である。

### E . 結論

JPOPSコホート1症例について、機械学習技術に基づき、PSA再発および有害事象発生の有無を予測した。

### F . 研究発表

#### 1. 論文発表

なし

#### 2. 学会発表

- 1) 馬込大貴, Radiomics と人工知能技術の放射線治療分野における応用. 第4回浜松放射線治療研究会. 特別講演, 2018.03.16. 浜松プレスタワー, 静岡県
- 2) 馬込大貴, 統計解析、人工知能、機械学習、パターン認識、データ解析. 日本放射線腫瘍学会第6回放射線治療・物理学セミナー. 口頭, 2018.03.03. 兵庫県立淡路夢舞台国際会議場, 兵庫県
- 3) 馬込大貴, 放射線治療分野における機械学習技術を用いた医療ビッグデータの統合解析. 札幌医科大学放射線治療セミナー. 招待講演, 2018.02.13. 札幌医科大学, 北海道
- 4) 馬込大貴, Radiomics と医療ビッグデータの統合解析. 日本放射線腫瘍学会第31回高精度放射線外部照射部会学術大会. 口頭, 2018.02.10. 大阪府立国際会議場, 大阪府
- 5) 馬込大貴, 機械学習による医療ビッグデータ解析の将来ー放射線治療系を主としてー. 平成29年度神奈川県診療放射線技術講習会, 招待講演, 2017.11.12. 神奈川県総合医療会館,

神奈川県

- 6) 馬込大貴, 人工知能を活かした放射線治療. 日本放射線技術学会東京支部, 東京都診療放射線技師会, 第17回合同学術講演会. 特別講演, 2017.09.30. JR 東京総合病院, 東京都
- 7) 馬込大貴, Outcome prediction based on machine learning techniques in radiation oncology. 平成 29 年度第 5 回重粒子線医工連携セミナー. 招待講演, 2017.07.21. 群馬大学重粒子線医学研究センター, 群馬県

#### G . 知的財産権の出願・登録状況

(予定を含む)

1. 特許取得  
なし
2. 実用新案登録  
なし
3. その他  
なし