

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)

総括 研究報告書

ビッグデータからの機械学習による前立腺癌小線源療法の予後予測法の
開発と均てん化への応用

研究代表者 中村 和正 浜松医科大学 医学部 教授

研究要旨：

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) のコホート1の入力項目全291個の中から、生存率、有害事象に関係のない項目を除き、機械学習に用いる解析対象項目を選別した。また、機械学習による治療成績の評価の有効性を確認するため、60Gy以上の外部照射を施行された限局性前立腺癌の匿名化されたデータを用いて、機械学習による予後予測モデルの有用性を確認した。POPSコホート1のデータを用いて、一般的に知られている予後因子 (T因子、治療前PSA値、Gleason score、照射線量、ホルモン療法の有無等が、PSA非再発率に影響するかどうかを調べた。次に、臨床特徴量、治療特徴量とPSA再発、有害事象の関係についてロジスティック回帰、サポートベクターマシン (SVM)、ランダムフォレスト (RF)、ディープラーニング (DNN) の機械学習を用いて解析した。学習データではRFやDNNで予測精度が上がる傾向にあった。今後、コホート2の約3600例を解析対象に加え、さらに症例を増やして研究を進めていく予定である。

研究分担者

斉藤史郎 東京医療センター・医長
萬 篤憲 東京医療センター・医長
伊藤一人 群馬大学・准教授
馬込大貴 駒沢大学・講師
小島伸介 臨床研究情報センター・TRI専門職
菊池 隆 臨床研究情報センター・上席研究員

A . 研究目的

前立腺癌の放射線治療の予後因子としては、T因子、PSA値、Gleason分類などがある。これらによって低・中・高リスクに分類され、大まかな予後予測が可能で、それに沿ったノモグラムが作成されてい

る。しかし、それ以外の患者背景や検査所見、照射線量、治療パラメータなどの多くの因子を網羅的に分析し、治療結果を予測する方法については国内外を含めてほとんど研究が進められていない。

我々は、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) を実施してきた。本研究には全国74施設 (小線源療法を施行する施設の約70%) が参加し、2005年から2010年

までに小線源療法で治療された約7000例（同時期に本邦で小線源療法により治療された症例の約40%）が前向き登録された。調査項目は、患者・家族背景、生活歴、T因子、PSA値、Gleason分類などの腫瘍因子、小線源治療パラメータ、外照射併用の有無等の治療因子、PSA再発の有無、生死、有害事象の有無などで、少なくとも5年以上の経過観察が行われている世界的に類を見ないビッグデータである。2005年から2007年までのコホート1の2339例のデータクリーニングが終了し、2010年までのコホート2の約4600例についても平成28年11月で5年の経過観察が終了し、これらのビッグデータが使用可能となる予定である。

本研究の目的は、JPOPSによって得られたビッグデータを用いて、詳細な臨床情報を機械学習させることにより、新しい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

JPOPS 研究のコホート 1（2005-2007年）で収集された 2339 例のデータのクリーニングはすでに終了している。平成 29 年度は、このデータを用いて、研究を開始した。

1)機械学習に用いる解析対象項目の選別

JPOPS 研究のコホート 1（2005-2007年）の入力項目全 291 個の中から、生存率、有害事象に関係のない項目を除き、機械学習に用いる解析対象項目を選別した。

2)前立腺癌外部照射データでの機械学習の予後予測精度向上の検証

機械学習による治療成績の評価の有効性

を確認するため、1995 年から 2002 年に限局性前立腺癌に対して 60Gy 以上の外部照射を実施した 679 例(Nakamura K, et al. External-beam radiotherapy for localized or locally advanced prostate cancer in Japan: a multi-institutional outcome analysis. Jpn J Clin Oncol. 2008;38(3):200-4.)の匿名化されたデータを用いて、再発の有無の予測を行った。

3)POPS コホート 1 での PSA 非再発率に関する予後因子解析

POPS コホート 1 のデータを用いて、一般的に知られている予後因子（T 因子、治療前 PSA 値、Gleason score、照射線量、ホルモン療法の有無等が、PSA 非再発率に影響するかどうかを調べた。

4)POPS コホート 1 での機械学習による予後予測精度向上の検証

POPS コホート 1 のデータを用いて、ロジスティック回帰、サポートベクターマシン(SVM)、ランダムフォレスト(RF)、ディープラーニング(DNN)の機械学習を用いて、解析対象項目と PSA 再発、有害事象の関係の解析を開始した。

（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。また、JPOPS研究のコホート1およびコホート2のデータセット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。

本年度、JPOPS研究のコホート1のデータを用いて解析を開始したが、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。また、解析用のパソコンにお

いてもパスワード等で厳重に管理されている。

研究の透明性確保のため、研究代表者の所属機関（浜松医科大学）において倫理審査を受け、平成29年7月26日に倫理審査委員会の承認を得た（研究番号17-095）。また、解析を担当する駒沢大学においても同様に平成29年7月21日に倫理審査委員会の承認を得た（通知番号17-18）。

匿名化された既存データのみを用いる観察研究ではあるが、今後も、ハード、ソフトウェアおよびデータの移送のすべてのレベルでセキュリティを強固にする。

C . 研究結果

1) 機械学習に用いる解析対象項目の選別

JPOPS 研究のコホート 1（2005-2007年）の入力項目全 291 個の中から、機械学習に用いる解析対象項目を選別した（資料 1）。特に、臨床的に生存率、有害事象に関係あると考えられる項目を Limited database、その他の項目もすべて含めた Large database に分類した。

2) 前立腺癌外部照射データでの機械学習の予後予測精度向上の検証

1995 年から 2002 年に限局性前立腺癌に対して 60Gy 以上の外部照射を実施した 679 例の匿名化されたデータを用いて、再発の有無の予測を行った。予測には、Logistic 回帰と機械学習の一種である Support vector machine (SVM) を用いた。データ欠損値のある症例を除いた 450 例について、440 例を Training、10 例を Validation として用い、臨床特徴量（PSA、Gleason score 等）のみを用いる場合、治

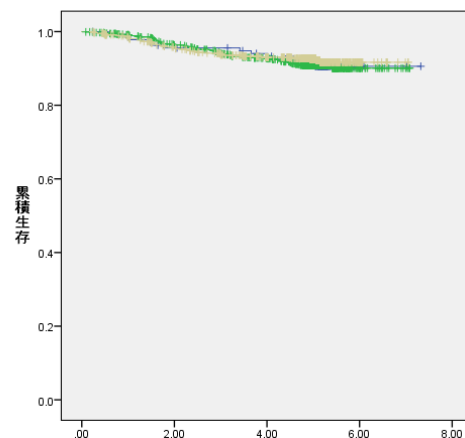
療特徴量（処方線量、照射方法等）のみを用いる場合、すべてを用いる場合の 3 つのモデルを作成し、予測精度を比較した。SVM では有意に予後予測精度が向上しており、機械学習による予後予測モデルの有用性を確認した。

3) POPS コホート 1 での PSA 非再発率に関する予後因子解析

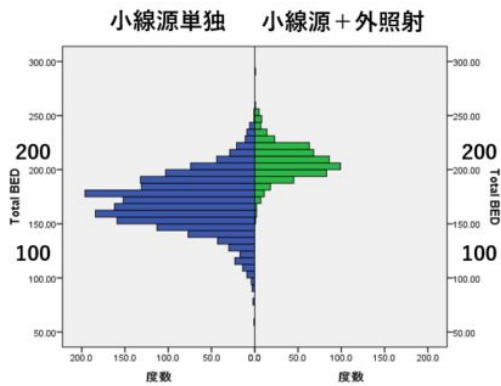
POPS コホート 1 のデータを用いて、一般的に知られている予後因子（T 因子、治療前 PSA 値、Gleason score、照射線量、ホルモン療法の有無等が、PSA 非再発率に影響するかどうかを調べた。

解析対象 2316 例のうち、治療前 PSA、T 因子、Gleason score すべてが入力されており（いずれかが未記入 10 例）かつ D90 が入力されている（未記入 11 例）2295 例で解析を実施した。エンドポイントを Phoenix 定義での PSA 再発とした。年齢中央値 69 歳（45 - 89 歳）、観察期間中央値 5 年（0.09-7.24 年）であった。

NCCN（National Comprehensive Cancer Network）リスク分類（低リスク 1045 例、中リスク 1113 例、高リスク 137 例）の PSA 非再発率に差がなかった。



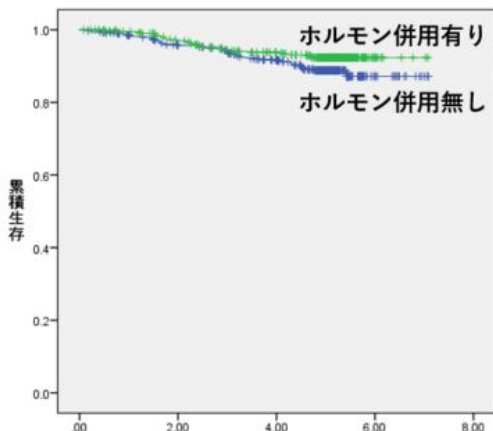
2295 例のうち、小線源単独 1751 例、外照射併用 544 例であった。小線源および外照射での投与線量については、biological equivalent dose (BED) にて換算して比較した。



BED の分布ヒストグラムでは、外照射併用例で全体的に高線量が投与されていることがわかった。しかし、Total BED を 120 以下 (59 例)、120-160 (541 例)、160-200 (1124 例)、200 より大 (571 例) にて比較したが、PSA 非再発率に差は無かった。

ホルモン療法については、ホルモン療法無し 1159 例、有り 1136 例であったが、PSA 非再発率に差は無かった。

各リスク群でのホルモン療法の有無の影響を調べたところ、中リスク群で、ホルモン療法有りが有意に PSA 非再発率が良好であった (P=0.046)。



外照射の有無では、PSA 非再発率に差はなかった。

また、前立腺体積の中央値は 25.1ml (7-71ml)であったが、40ml 以上群で PSA 非再発率は 100%で、他の群と比較して有意に PSA 非再発率が良好であった。(40ml 以上 vs 20ml 以下 P=0.18、40ml 以上 vs 20-30ml P=0.002、40ml 以上 vs 30-40ml P=0.02)。

有害事象に関しては、Grade 2 以上の有害事象の割合を外照射の有無で比較したところ、外照射は有意に消化器系の晩期有害事象の頻度を上昇させていた (P=0.000)。泌尿器系については、急性期は外照射無しがむしろ高いが (P=0.000)、晩期では差がなかった (P=0.514)。

また、治療計画パラメータと有害事象に関しては、R100 (処方線量の 100%の線量が投与された直腸容積)は Grade 2 以上の消化器系晩期有害事象が見られた群で有意に高かった (有害事象無し 0.4920 ml vs 有り 0.9174 ml P=0.004)。R150 (処方線量の 150%の線量が投与された直腸容積)も同様であった (有害事象無し 0.0514 ml vs 有り 0.1458 ml P=0.004)。

一方、尿路系の有害事象については、UD90 (平等に投与された線量のうち、90%の尿道容積に投与された線量の下限)のみが晩期有害事象と関連していた (有害事象無し 140.28 Gy vs 有り 147.12 Gy P=0.034)。また尿路系有害事象の Grade が高いほど、治療前 IPSS (国際前立腺症状スコア)が高い傾向にあった。

4) POPS コホート 1 での機械学習による

予後予測精度向上の検証

上記解析結果を踏まえ、POPS コホート1のデータを用いて、ロジスティック回帰およびサポートベクターマシン(SVM)、ランダムフォレスト(RF)、ディープラーニング(DNN)の機械学習手法を用いて、PSA再発、有害事象発生率の予測を改善できるかを検討した。入力特徴量は、予後と関連がありそうな特徴量のみ厳選した Limited Database となるべく多くの特徴量を採用した Large Database の2つのデータセットを用いた。評価法は 10-fold cross-validation 法を用いて Accuracy(全症例の内、正しく予測できた症例の割合)、Precision(再発と予測した症例の内、真に再発した症例の割合)、Recall(再発症例の内、真に再発した症例の割合)を評価指標とした。学習データでは RF や DNN で予測精度が上がる傾向にあったが、テスト症例の予測精度は現時点ではいずれも低かった。

D . 考察

平成29年度では、JPOPSコホート1からの有効性解析対象例2316例の臨床特徴量、治療特徴量とPSA非再発率、臨床的非再発率、全生存率、直腸および尿路有害事象発生率との関係に対する機械学習を開始した。現時点では、学習データではRFやDNNで予測精度が上がる傾向にあったが、テスト症例の予測精度はいずれも低かった。今後、DNNの内部構造を工夫してさらに精度を上げる工夫を試みている。また、PSA非再発率だけではなく、臨床的再発率等と特徴量との関連も調べる予定である。

今後、コホート2の約3600例も解析対象に加えることができる予定で、さらに症例を増やして研究を進めていく予定である。

また、治療後のPSAの推移のデータもあり、治療後のPSAの動きを機械学習させ、再発を予測できるかについても新たに研究を進めていく予定である。

E . 結論

JPOPSコホート2316例のデータを用いて、臨床特徴量、治療特徴量を機械学習させることにより、PSA非再発率、臨床的非再発率、全生存率、直腸および尿路有害事象発生率を推測できるかについて解析を開始した。

本研究の遂行により、従来にない、正確な予後予測アルゴリズムを確立することができれば、以下のような活用が期待できる。

1) 小線源療法にて治療された患者で、治療時点の背景因子、治療パラメータ等を入力することで、PSA非再発率、有害事象発生確率を正確に予測できる。

2) 治療結果が正確に予測されれば、再発リスクの低い症例では経過観察期間を延ばすことが可能となる。また、再発リスクが高いと判断されれば、密な経過観察を行い、再発の時点で早期に治療介入することが可能となる。このように個別化した介入が可能となり、医療費削減効果が期待できる。

3) 治療パラメータを入力することで、再発確率が計算できるため、各施設間の治療成績の予測が可能となる。これにより、治療成績の施設差を明らかとすることができ、放射線治療の質の均てん化に役立てる

ことができる。

4) 現在、前立腺癌の高線量率組織内照射の成績調査、陽子線治療、炭素線治療の前向き臨床試験、強度変調放射線治療の前向き登録が進行しており、研究代表者の中村はそのすべてに分担研究者等として関与している。本研究で機械学習による予後予測法を確立することにより、その研究プロセスを応用して、小線源療法のみならず、前立腺癌放射線治療モダリティすべての予後予測アルゴリズムの確立が期待される。

F . 健康危険情報

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。

G . 研究発表

1. 論文発表

1. Nakamura K, Konishi K, Komatsu T, Sasaki T, Shikama N. Patterns of radiotherapy infrastructure in Japan and in other countries with well-developed radiotherapy infrastructures. Jpn J Clin Oncol 2018;48(5):476-479

2. 学会発表

1. 中村和正. 「QOL から見た各種限局癌の比較 -外照射療法-」パネルディスカッション「QOL からみた各種限局癌治療の比較」第十四回前立腺癌密封小線源永久挿入治療研究会 東京 2018.1.21
2. 中村和正. 「前立腺癌に対するX線外部照射の現状と展望」シンポジウム5 . 最

新エビデンスに基づく前立腺癌診療の現状と今後の展望. 日本放射線腫瘍学会第30回学術大会 2017.11.19 大阪

3. 中村和正. 「前立腺癌外部照射における課題と展望」(シンポジウム)シンポジウム1 前立腺癌治療における課題と展望. 日本泌尿器腫瘍学会第3回学術集会. 2017.10.22-10.23 横浜市
4. 中村和正. 「前立腺癌放射線治療の現状と今後の展開」(特別講演)平成29年度福岡県放射線科医会総会 2017.9.29 福岡市
5. 中村和正. 「前立腺癌の放射線治療根治から緩和まで」(特別講演)遠州・三河放射線治療講演会 2017.9.2 浜松
6. K. Nakamura. Present status and planned prospective studies of radiation therapy for prostate cancer in Japan. 2017 International Academic Forum of Peking on New Technique of Radiation Oncology Process. 2017.8.19, Beijing, China.

H . 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

資料 1)

JPOPS 入力データからピックアップした機械学習解析対象項目

【インプット情報】

1. 患者背景

身長、体重、PS、既往歴、家族歴、喫煙歴、飲酒歴

2. 疾患情報

生検箇所（左右の陽性個数/生検個数）、Gleason 分類、臨床病期、PSA 値

3. 治療情報

治療施設

3 - 1. 外照射治療の実施状況

治療の有無、治療期間、外照射総線量、X 線エネルギー、外照射野、外照射方法、照射体位

3 - 2. 内分泌治療の実施状況

前立腺容積、IPSS スコア、治療の有無、ブロックの有無、治療開始時期、治療期間 内分泌治療開始日、治療内容

3 - 3. 小線源療法

プレプラン

前立腺容積、測定日、小線源治療の完遂、線源強度、線源個数、処方線量、挿入針の数、線源配置法、線量計画法

ポストプラン

前立腺線量 (V100、V150、D90)、直腸線量 (R100、R150)、尿道線量 (D90、D5、U200)

3 - 4. 治療後情報

すべての治療終了日、PS

安全性の評価（脱落線源の有無、脱落した線源の個数、移動線源の有無、移動先の部位）

* 下線：特に生存率、有害事象に関係あると考えられる項目（Limited database）

【アウトプット情報】

治療効果

PSA 再発、臨床的再発、生存の有無、各事象確認日（または最終生存日）

有害事象

血尿、排尿時痛、失禁、頻尿/尿意逼迫、（残尿）/尿閉

直腸炎、直腸出血/血便、疼痛