

論文内容要旨

Development of radiomics and machine learning model for predicting occult cervical lymph node metastasis in patients with tongue cancer

(Radiomics と機械学習を用いた舌癌局所治療後の後発リンパ節転移予測モデルの開発)

Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology, in press

主指導教員：永田 靖教授

(医系科学研究科 放射線腫瘍学)

副指導教員：村上 祐司准教授

(医系科学研究科 放射線腫瘍学)

副指導教員：柿本 直也教授

(医系科学研究科 歯科放射線学)

久保 克磨

(医系科学研究科 医歯薬学専攻)

【背景】

リンパ節転移や遠隔転移のない舌癌症例には、標準治療として舌部分切除や小線源治療が行われる。しかし、術後に頸部リンパ節転移再発を来す症例（後発リンパ節転移再発症例）が、2割から5割程度存在する。近年の画像診断技術の向上にも関わらず、これらの症例を術前に同定することは困難で、一部の高リスク症例では予防的な頸部郭清術が考慮される場合もある。ただし、リンパ節転移のない症例においては過大侵襲となり、またその有用性についても positive data、negative data の両方が存在し、意見の分かれるところである。

一方、Radiomics とは、放射線医学の分野において大量の画像情報を系統的に扱う学問である。近年、医用画像を用いた Radiomics と機械学習による予測モデルの開発は大きな関心を集めている。Radiomics と機械学習を用いることで、治療前に同定することが困難な後発リンパ節転移の有無を事前に予測できる可能性がある。

よって、本検討は、Radiomics と機械学習を用いた舌癌局所治療後の後発リンパ節転移再発の予測モデルを開発することを目的とした。

【方法】

本検討は、手術あるいは小線源治療を受けた舌癌症例 161 例を対象とした。症例の適格基準は、舌癌に対する局所治療が施行されていること、頸部郭清や予防的頸部照射、舌癌に対する化学療法歴がないこと、治療前に造影 computed tomography (CT) が施行されていること、十分な経過観察が行っていることであった。後発リンパ節転移再発があった症例では、後発リンパ節転移が病理学的に証明されていることも適格基準に含めた。関心領域の抽出では、各症例において、患側の level Ia、Ib、II、III、IVa の 5 つの regions of interest (ROI) を作成した。また、それらの ROI を足し合わせた total neck node level (患側の level Ia+Ib+II+III+IVa) を作成した。以降の検討では、total neck node level での解析と、患側のそれぞれの level ごとでの解析を行った。ラベリングは、total neck node level の解析の場合は、各症例の total neck node level ごとに後発リンパ節転移の有無を割り当てるため、1 症例に 1 つのラベル、それぞれの level ごとでの解析の場合は、各 Level ごとに後発リンパ節転移有無を割り当てるため、1 症例に 5 つのラベルとなった。画像特徴量の抽出では、first-order statistical features、shape-based features、texture analysis features の 3 種類の特徴量を、z-score normalization を用いて標準化し、least absolute shrinkage and selection operator (Lasso) 回帰分析を用いて、機械学習に有効な特徴量を選択した。モデルの構築では、k 近傍法 (kNN)、サポートベクターマシン (SVM)、classification and regression trees (CART)、ランダムフォレスト (RF)、Ada Boost (Ada) の 5 つのモデルを使用した。モデル構築では、過学習を避けるため、10-fold cross-validation 法を使用した。また、本検討では、不均衡データの対処として、synthetic minority oversampling technique (SMOTE) を併用した。モデルの評価としては、accuracy、precision、recall、receiver operator characteristic (ROC) 曲線、area under the curve (AUC) score を使用した。

【結果】

患者背景

全 161 例中、46 例 (28.6%) で後発リンパ節転移を認めた。後発リンパ節転移の部位は、level Ia が 3 例 (6.5%)、level Ib が 24 例 (52.2%)、level II が 25 例 (54.3%)、level III が 8 例 (17.4%)、level IVa が 3 例 (6.5%) であった (重複あり)。後発リンパ節転移の発生時期は中央値で、初回治療から 3.5 カ月 (幅、0-12 カ月) であった。観察期間の中央値は 56 カ月 (幅、9-131 カ月) であった。

モデルの開発

Lasso 回帰分析の結果、total neck node level ごとの解析、それぞれの level ごとの解析、それぞれで 7 つずつの特徴量が選択された。選択された特徴量を用いて、5 つの機械学習モデルを開発した。それぞれの機械学習モデルで、精度の向上のため SMOTE を適用した。Total neck node level での解析においては、RF+SMOTE のモデルにおいて、accuracy が 0.85、precision が 0.88、recall が 0.82、AUC score が 0.92 と最も高い精度を示した。それぞれの level ごとの解析においては、SVM+SMOTE のモデルにおいて、accuracy が 0.96、precision が 0.96、recall が 0.95、AUC score が 0.98 と最も高い精度を示した。

【結論】

術前の造影 CT を用いた、Radiomics と機械学習による舌癌局所治療後の後発リンパ節転移再発の予測モデルを開発した。特に、SVM+SMOTE のモデルにおいて、accuracy が 0.96、AUC score が 0.98 と最も高い精度を示した。Radiomics と機械学習による舌癌局所治療後の後発リンパ節転移再発の予測モデルは、転移のない舌癌症例の治療方針決定に際し有用な可能性がある。