

令和 4 年 6 月 13 日現在

機関番号：16301

研究種目：基盤研究(C) (一般)

研究期間：2019～2021

課題番号：19K10312

研究課題名(和文)リンパ節転移診断遺伝子アルゴリズムの構築による口腔癌リンパ節転移予測の確立

研究課題名(英文) Establishment of prediction of lymph node metastasis by constructing a genetic algorithm in oral squamous cell carcinoma.

研究代表者

合田 啓之 (Goda, Hiroyuki)

愛媛大学・医学部附属病院・講師

研究者番号：00464371

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：口腔扁平上皮癌における最も重要な予後因子はリンパ節転移の有無である。本研究では、原発組織の遺伝子発現プロファイルに基づいた機械学習によるリンパ節転移予測の可能性について検討した。口腔扁平上皮癌一次症例の原発腫瘍よりtotal RNAを抽出し、マイクロアレイによる網羅的遺伝子発現解析後、サポートベクターマシン(SVM)を用いてリンパ節転移予測モデルを構築した。その精度を評価したところ感度60%、特異度100%、正確度80%であった。口腔扁平上皮癌において原発腫瘍の遺伝子発現プロファイルを用いた機械学習が、頸部リンパ節転移予測に有用となる可能性が示唆された。

研究成果の学術的意義や社会的意義

本研究の目的は、原発組織(生検組織)または血清の網羅的遺伝子解析から頸部郭清術の真の適応症例のみ抽出する遺伝子アルゴリズムを構築することである。厳正な適格症例の選択基準が遺伝子解析によって行われ、センチネルリンパ節生検の手法による適切な郭清領域に対する治療が行われれば、治療成績の向上、過剰医療や不足医療による医療費の削減が見込まれ、標準治療の確立というEBMに基づいたcN0症例の治療方針へのパラダイムシフトにつなげることができると考えられる。また、他領域への適応等による医療界全体への波及効果をも期待できる研究と考えられる。

研究成果の概要(英文)：The most important prognostic factor in oral squamous cell carcinoma is the presence of lymph node metastasis. In this study, we investigated the possibility of predicting lymph node metastasis by machine learning based on gene expression profiles of primary tissues. Total RNA was extracted from primary tumors of primary cases of oral squamous cell carcinoma, and after comprehensive gene expression analysis using microarrays, a lymph node metastasis prediction model was constructed using a support vector machine (SVM). The accuracy of the model was evaluated as follows: sensitivity 60%, specificity 100%, and accuracy 80%. The results suggest that machine learning using gene expression profiles of primary tumors may be useful for predicting cervical lymph node metastasis in oral squamous cell carcinoma.

研究分野：外科系歯学

キーワード：口腔癌 頸部リンパ節転移 遺伝子診断

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

口腔癌扁平上皮癌（以下、口腔癌）患者は本邦では 2015 年に 7,800 人となり世界的にみても増加の一途をたどっている。しかしながら、手術療法、放射線化学療法等の進歩にもかかわらず、治療成績はこの 40 年間改善が認められていない (Int J Cancer. 125:2159-2165, 2009)。この要因としては頸部リンパ節転移の制御が困難なことにあった。現在、口腔癌の最も重要な予後因子は頸部リンパ節転移の有無であることは明らかとなっており、この点の改善なくして早期口腔癌の治療成績向上は見込めない。申請者らはこれまで、センチネルリンパ節生検の診断法を確立してきたが、その煩雑性、手術時間の確保の困難さ、侵襲性（70-80% には不必要）により普及は困難な状況である。この現状を打開するには、リンパ節転移診断精度を向上させ、かつ、低侵襲な手術（郭清領域の診断）が必要となる。そこで、患者サンプルからのリンパ節転移アルゴリズムを構築し、診断陽性の患者に対しては、郭清すべき領域（患側、健側または両側）をセンチネルリンパ節生検にて同定し郭清を行うことで、これらの問題が解決するのではないかという発想に至った。本研究では、原発組織（生検組織）または、血清の網羅的遺伝子解析（マイクロアレイ解析）結果を用いたリンパ節転移診断アルゴリズムを構築し、cN0 症例のなかで、頸部リンパ節転移の可能性が高い症例のみを抽出する、リンパ節転移診断遺伝子アルゴリズムを構築することが目的である。本研究は、アルゴリズムの構築による cN0 症例における治療成績の向上、過剰医療や不足医療による医療費の削減、標準治療の確立という EBM に基づいた cN0 症例の治療方針へのパラダイムシフトにつなげることを最終目標としている。

他領域では、メラノーマや乳癌、胃癌を含め、cN0 症例に対するセンチネルリンパ節生検がすでに標準治療化されており、その有用性は確立されている。一方で、問題点として、約 70-80% の症例に対しては過剰医療となること、結果による手術時間の予測が困難なことがあり、医療資源や医療費、稼働率等を含め、この点に関する研究成果は皆無である。本研究手法による診断法が確立されたならば、他領域での応用も同様に可能となり、治療成績の向上や医療費の削減等、医療界全体への波及効果をも期待できる研究と考えられる。

2. 研究の目的

本研究の目的は、原発組織（生検組織）または血清の網羅的遺伝子解析から頸部郭清術の真の適応症例のみ抽出する遺伝子アルゴリズムを構築することである。厳正な適格症例の選択基準が遺伝子解析によって行われ、センチネルリンパ節生検の手法による適切な郭清領域に対する治療が行われれば、先述した治療成績の向上、過剰医療や不足医療による医療費の削減が見込まれ、標準治療の確立という EBM に基づいた cN0 症例の治療方針へのパラダイムシフトにつなげることができると考えられる。また、これはすでに標準治療化されているメラノーマや乳癌においても行われておらず、他領域への適応等による医療界全体への波及効果をも期待できる研究と考えられる。

3. 研究の方法

本研究の解析対象は口腔扁平上皮癌の未治療新鮮症例のうち腫瘍の大きさが T2 以下で、造影 CT、造影 MRI、PET-CT、超音波検査等にて明らかなリンパ節転移を認めない cN0 症例とする。解析対象組織は、生検もしくは原発切除組織および術前血清とする。

本研究は、以下の項目にそって研究を進める。

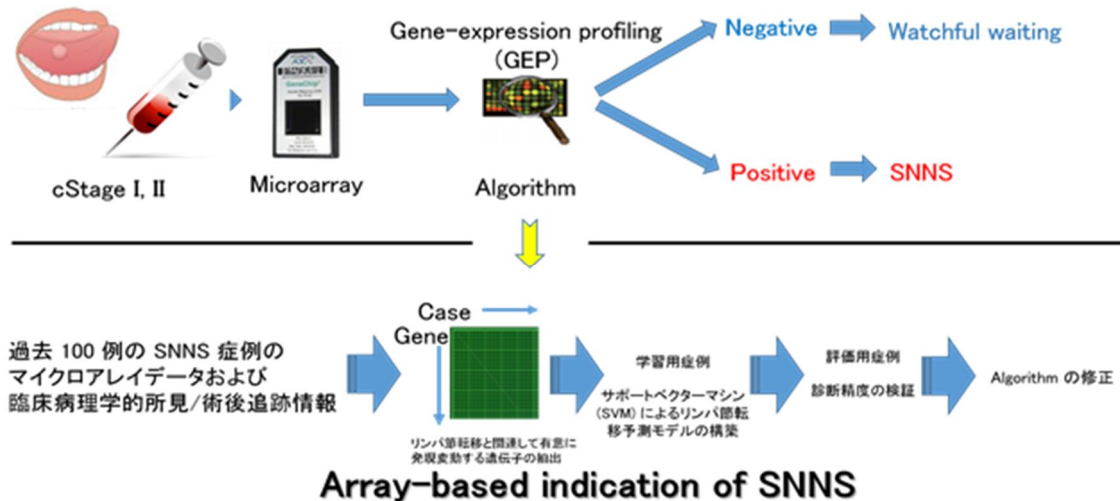
アレイデータの確認

すでに手元にある 2 年以上の経過観察を行った口腔扁平上皮癌 cN0 症例のアレイデータについて、各症例の全発現データを非階層型クラスタリングし、サンプル間での batch effect、性別、年齢、腫瘍径 (T stage)における遺伝子発現プロファイリングに違いがあるかを検討する。スポットの数値化、バックグラウンド補正、色素補正、異常スポットの検出当は Agilent 社の Feature Extraction を用いて行う。

非階層型クラスタリング等により不適当と判断したサンプルは除外する。

後ろ向き研究によるサンプル分類

病歴、追跡データによりセンチネルリンパ節陽性症例および後発リンパ節転移症例群とリンパ節転移陰性群にサンプルをグループ分類する。



グループ間での発現変動遺伝子の抽出 (マーカー遺伝子候補の絞り込み)

上記グループ間での発現変動遺伝子の抽出を training set とし、2 群間での発現変動遺伝子を評価する。統計解析としては、マンホイットニー U テスト等を採用して行う。すべての遺伝子発現データよりマーカー遺伝子の絞り込みを行い、候補を抽出する。

後ろ向き研究による診断精度分析

上記解析にて絞り込みを行ったマーカー遺伝子候補を用いて、統計的に上位の遺伝子から順に学習セットのサンプルを組み、後ろ向きに診断精度の分析を行う。解析手法として、Weighted Vote 法を第一に行う。また、並行して Decision Tree 法、Neural Network 法、Support Vector

Machines 法、Naïve Bayesian 法および Partial Least Squares Discrimination 法での解析も行い、診断精度を最大限に示すマーカー遺伝子群を決定する。

本解析時に最も留意すべき事項として、偽陰性症例を 0% に抑えることを念頭に置く。偽陰性症例はすなわち後発リンパ節転移症例であり、最も回避すべき事項である。本研究の目的は、リンパ節転移診断アルゴリズムの構築であり、先述した画像検査で見落とす、いわゆる潜在転移症例である約 30% の症例を見分け、残りの 70% を除外することにある。

本解析を組織からのアレイデータおよび術前血清からのアレイデータにて別々に解析を行う。

また、サンプリングが正しく行われたか（腫瘍サンプル認定のアルゴリズム）、性別診断のアルゴリズムを併用し、サンプル精度の標準化も同時に行う。

他施設共同研究によるアルゴリズムの診断精度分析

前段階でのマーカー遺伝子群を用いてのいわゆる Validation Set にて多施設共同研究を開始する。対象は、臨床研究倫理審査委員会で承認を受けた本研究の主旨に本人、ご家族を含め同意を頂いた、未治療の口腔扁平上皮癌新鮮症例とし、腫瘍径 4cm 以下の T1-2 症例、画像検査上頸部リンパ節転移を明らかに認めない stage1-2 症例とする。術前の生検組織および血液検査にて血清採取を行い、Validation Set を組み込んだ、アレイチップまたは、マーカー候補によってはカスタムアレイチップを用いてアルゴリズム解析を行う。前述したように、cN0 症例における治療方針は標準化されておらず、各施設における治療方針で頸部の治療方針を決定する。本研究にて患者に不利益が生じる可能性は排除されるべきであり、このような研究計画にて進めることとする。

対象症例のセンチネルリンパ節および他の頸部リンパ節群の病理組織検査結果より頸部リンパ節転移の有無を判別し、その後、少なくとも 2 年間の追跡調査を行うことにより後発頸部リンパ節転移を含めたリンパ節転移の有無を診断し、アルゴリズムの改訂を行う。また、サンプリング時に蛋白の抽出も同時に行い、転移因子の解析も併せて行う。

対象患者群の診断精度の評価

診断精度分析の結果によっては、追加症例を学習用データとして用いることで、アルゴリズムの再構築を行う。追加症例は 10 症例毎とし、上記した偽陰性症例 0% をキープしつつ、cN0 症例の潜在転移率と考えられる約 30% を診断できるアルゴリズムをもって、モデル完成とする。

4 . 研究成果

まず、アルゴリズム構築のための学習用症例として、T1-T2 の舌扁平上皮癌 1 次症例 40 例の原発組織より total RNA を抽出し、マイクロアレイを用いて、49,386 プローブにて 2 年以上経過観察を行い、リンパ節転移を認めた 18 症例と転移を認めなかった 22 症例にて転移関連遺伝子を GeneSpring GX にて signal/noise ratio 2 以上にて抽出を行った。リンパ節転移と関連して 2 倍以上の有意な発現変動を示す遺伝子が 329 種類認められ、その中には FN1、LAMC2、TGFB1 の発現亢進や KRT13、IVL、CLDN4 の発現低下などが含まれていた。

Gene symbol	Gene title	Fold change	p
TRIMP63	tripartite motif-containing 63	3.683427	0.010781461
HMGA2	high mobility group AT-hook 2	3.414633	0.010491469
FN1	fibronectin 1	2.819567	0.012240476
LAMC2	laminin, gamma 2	2.439863	0.011090115
KRT8	keratin 8	2.359377	0.026630415
TGFBI	transforming growth factor, beta-induced	2.319371	0.011090115
AMIGO2	adhesion molecule with Ig-like domain 2	2.103128	0.034854025
TIMP3	TIMP metalloproteinase inhibitor 3	2.043434	0.010722678
PDGFC	platelet derived growth factor C	2.018558	0.011090115
VEGFC	vascular endothelial growth factor C	2.013734	0.01681363
KLK6	kallikrein-related peptidase 6	-2.02484	0.010781461
DSG3	desmoglein 3 (pemphigus vulgaris antigen)	-2.14026	0.022585215
ANXA1	annexin A1	-2.17539	0.012319613
CLDN4	claudin 4	-2.98485	0.011090115
IL1F5	interleukin 1 family, member 5 (delta)	-3.08031	0.013822531
IVL	involucrin	-3.58416	0.011976322
S100P	S100 calcium binding protein P	-3.62191	0.011413634
AQP3	aquaporin 3 (Gill blood group)	-4.4222	0.011090115
SPRR1A	small proline-rich protein 1A	-5.34488	0.013690655
KRT13	keratin 13	-8.78773	0.011090115

329 遺伝子のうち、発現上昇を認めた 59 種類、発現低下を認めた 270 種類の遺伝子を用いて、Support vector machine にてリンパ節転移診断アルゴリズムを構築した、感度、特異度、精神率ともに 100% であることを確認した。

続いて、評価用サンプルの 10 症例にて、先ほどの転移予測アルゴリズムの正診率を評価したところ、感度 60%、特異度 100%、正診率 80% の結果であった。

case	gender	N status	Predicted	confidence
TSCC	F	-	[-]	0.17215274
TSCC	M	-	[-]	0.4339309
TSCC	M	+	[+]	0.3311313
TSCC	M	+	[+]	0.25456762
TSCC	M	+	[+]	0.036388267
TSCC	M	+	[-]	0.018365918
TSCC	M	+	[-]	0.43900192
TSCC	M	-	[-]	0.14808789
TSCC	M	-	[-]	0.16422172
TSCC	M	-	[-]	0.24461167

診断アルゴリズムに関しては、機械学習によるニューラルネットワークによる data mining を行うことにより、今後、症例数を蓄積するほど、精度の向上が得られるものと考えられた。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計3件（うち招待講演 1件 / うち国際学会 1件）

1. 発表者名 合田啓之、中城公一、日野聡史、栗林伸行、徳善紀彦、内田大亮
2. 発表標題 口腔癌 NO 症例の治療戦略
3. 学会等名 第 22 回 SNNS 研究会学術集会（招待講演）
4. 発表年 2020年

1. 発表者名 Hiroyuki Goda, Koh-ichi Nakashiro, Norihiko Tokuzen, Nobuyuki Kuribayashi, Satoshi Hino, Daisuke Uchida
2. 発表標題 Sentinel lymph node biopsy with RI method for clinically NO oral squamous cell carcinoma
3. 学会等名 The 58th Annual Meeting of Japan Society of Clinical Oncology（国際学会）
4. 発表年 2020年

1. 発表者名 合田啓之、中城公一、徳善紀彦、栗林伸行、内田大亮
2. 発表標題 早期舌扁平上皮癌の遺伝子発現に基づいた機械学習によるリンパ節転移予測
3. 学会等名 第 64 回公益社団法人日本口腔外科学会総会・学術大会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------