

バイオマーカーや遺伝子解析を用いたホルモン受容体陽性乳がんに対する 化学療法の適応決定の妥当性の検証

研究分担者

山城大泰 京都大学大学院医学研究科 外科学講座乳腺外科学 非常勤講師

石黒 洋 京都大学医学部附属病院（探索医療センター検証部・外来化学療法部）改革推進講師

稲本 俊 天理よろづ相談所病院 乳腺外科嘱託部長

研究要旨 乳癌全体の約 60%を占めるホルモンレセプター陽性手術可能乳癌患者の術後化学療法の適応を決定することができれば効果の見込めない化学療法を回避し、毒性軽減や医療費削減に寄与することができる。組織グレードや ki-67 陽性細胞率、遺伝子シグネチャーなどのバイオマーカーと化学療法の効果を検討したところ、組織グレード 1 では術前化学療法の著効例はなかった。組織グレード 2、3 では ki-67、遺伝子シグネチャーの結果、術前化学療法の効果の間に一定の関係は認められなかった。組織グレード 2、3 の患者のうち術前化学療法の著効例は ki-67 が 20%以上の症例の中に認められており、ホルモン受容体陽性乳癌では grade 2 以上かつ ki-67 が 20%以上の症例を化学療法の適応と考えるのが妥当と思われる。

A. 研究目的

薬物療法の内容を決定する際のガイドラインとして St. Gallen コンセンサス (2009)¹⁾、NCCN(national Comprehensive Cancer Network)²⁾、乳がん診療ガイドライン (日本乳癌学会編³⁾) などがあるが乳癌に対する術前・術後の化学療法の適応は十分に決定されているとは言えない。ホルモン受容体や HER2 などのバイオマーカーに関していえば術後化学療法では CALGB(Cancer and Leukemia Group B)や Breast cancer Intergroup の試験結果から、Anthracycline やそれに続く Taxane の効果はホルモン受容体陰性乳癌では比べホルモン受容体陽性乳癌に比べてかなり上回っていることが知られており⁵⁾術前化学療法でも、ホルモン受容体

陰性乳癌の pCR (組織学的完全奏効) 率はホルモン受容体陽性乳癌を明らかに上回っている。⁶⁾HER2 については anthracycline や Anthracycline と Taxane 併用に化学療法において、HER2 は術前でも術後でも効果の予測因子となる^{6, 7)}が、ホルモン受容体陽性 (その多くは HER2 陰性) 乳癌に対する化学療法の効果は十分に検証されているとはいえない。このような状況下で 2009 年の St Gallen コンセンサス会議¹⁾では乳癌の治療効果予測性に関して、内分泌療法、抗 HER2 療法、化学療法のそれぞれが適するか否かを判断することが薬物療法の適応決定に重要であるとのコンセンサスが形成された。(表 1) しかし乳癌全体の約 60%を占めるホルモン受容体陽性 HER2 陰性乳癌患者のいわゆる

Luminal タイプ乳癌の治療選択は明確に示されているとは言えない。(表2)組織学的グレードについては全体の約50%がグレード2に分類される上、増殖の指標と考えられる Ki-67 の免疫組織染色や、21 遺伝子シグネチャー^{7, 8)}は日本の保険適応上用いることができない。

本研究の目的はホルモン受容体陽性乳癌の化学療法適応決定に ki-67 の免疫組織染色や遺伝子シグネチャーが及ぼす影響を検討することである。

B. 研究方法

ホルモン受容体陽性 (HR+) かつ HER2 陰性の手術可能乳癌 {T_{1c-3} N₀ M₀ (>1cm) / T₁₋₃ N₁ M₀} に対する補助療法の適応決定に、増殖指標 (Ki-67) や遺伝子シグネチャーから算出される RS (recurrence score) が臨床的に有用であるか後ろ向き (後方視的) に検討した。

(倫理面への配慮)

本研究は、介入を行わない、既存試料を用いた後ろ向き観察研究で、疫学研究の倫理指針に沿って行った。

C. 研究結果

2007年3月から2010年12月まで京都大学医学部附属病院乳腺外科で治療が行われた原発性手術可能乳癌362例について検討した。

検討症例全体の年齢、腫瘍径、リンパ節転移の有無、病期、腫瘍組織グレード、ホルモン受容体と HER2 で分類した腫瘍の分子生物学的表現型 (phenotype) を表3に、表現系別の主要素式グレードを表4に示す。Luminal タイプ乳癌では組織グレードは2を中心に正規分布していたが、その他のタイプでは組織グレードが有意に高い傾向にあった。

Luminal タイプ乳癌において、組織グレードと ki-67、遺伝子シグネチャーの分布を検討した。

(図1) 遺伝子シグネチャーの施行例と非施行例で、組織グレードや ki-67 の染色陽性細胞 (%) の分布には偏りはなかった。組織グレード1で ki-67 が20%未満の症例に high RS の症例はなかった。組織グレード2の症例では ki-67、RS の分布には一定の傾向は認められなかった。

Luminal タイプ乳癌で遺伝子シグネチャーを受けていた症例について、薬物療法の内容と、ki-67、RS の関係を検討した。(図2) RS が18以上の症例で

化学療法を受けていなかった患者は、活動性の B 型肝炎、術前内分泌療法の効果が G2b (化学療法の効果判定基準に準ず)、患者拒否が理由であった。Ki-67 が20%以上の症例で化学療法を受けていなかった患者は術前内分泌療法の効果が G2a 以上で、ki-67 陽性細胞 (%) が術後著明に減少していた症例であった。他方 ki-67 が20%以下且つ RS が18未満で化学療法を受けていたのはいずれも41才と若年で、患者希望により化学療法を受けていた。

Luminal タイプ乳癌において、組織グレードと ki-67、術前化学療法施行例の組織学的治療効果、及び遺伝子シグネチャー施行例ではその RS がどのように分布及していたか検討した(図3) 組織グレード1では術前化学療法は1/2aに限られていた。組織グレード2、3で ki-67 が20%以上の症例においてのみ2b以上の奏効が認められた。組織グレードや RS と治療効果の間に一定の傾向は認められなかった。

D. 考察

Luminal タイプ乳癌の薬物療法の選択は組織グレード、ki-67、RS に基づいて選択されていた。組織グレード1では術前化学療法の効果が2b以上の著効例はなく、大半は化学療法の適応がないものと思われるが ki-67 が20%以上陽性の場合には RS の中間リスク群が含まれてことがある。

組織グレード2、3では ki-67、RS、術前化学療法の効果の間に一定の関係は認められなかった。術前化学療法の著効例は ki-67 が20%以上の症例の中に認められており、luminal type では grade 2 以上かつ ki-67 が20%以上の症例を化学療法の適応と考えるのが妥当と思われる。

E. 結論

乳癌の組織グレード、Ki-67、RSなどを考慮して薬物療法の選択が行われていた。組織グレード2以上かつ ki-67 が20%以上の症例を化学療法の適応と考えられた。今後前向き試験での検討が必要と考えられる。

G. 研究発表

1. 論文発表
なし
2. 学会発表
なし。

H. 知的財産権の出願・登録状況

1. 特許取得
なし。
2. 実用新案登録
なし。

3. その他
なし。

表1 St Gallen コンセンサスの薬物治療ガイドライン (文献1より改変)

治療手段	適応	コメント
内分泌療法	わずかでもER染色用性の場合	ER陰性/PgR陽性はおそらくartifact
抗HER2療法	ASCO/CAPガイドラインで適宜されるHER2陽性(免疫組織染色陽性細胞が>30%又はFISH>2.2倍)	各臨床試験の定義をもちいてもよい
化学療法	HER2陽性	化学療法と抗HER2療法の同時もしくは逐次併用療法にエビデンスがある
	Triple Negative	殆どの患者
	ER陽性・HER2陰性	リスクに応じて様々
		ER強陽性/HER2陽性では、化学療法なしの内分泌療法+抗HER2療法が理論的には可能だがその根拠はない
		他に有効な方法は現時点ではない
		表2のように臨床病理学的因子や患者の希望などから検討する

表2 St Gallen コンセンサスの治療ガイドライン
ホルモン受容体陽性乳癌の薬物療法の適応基準 (文献1より改変)

	化学・内分泌療法の相対的適応	決定には役に立たない	内分泌療法単独の相対的適応
ERとPgR	低いレベルのERとPgR		高いレベルのERとPgR
組織学的グレード	グレード3	グレード2	グレード1
増殖	高い	中間	低い
腋窩リンパ節転移	4個以上	1-3個	陰性
腫瘍周囲の脈管侵襲(PVI)	広汎なPVIがある		広汎なPVIがない
病理学的腫瘍径	>5cm	2.1-5cm	≤2cm
患者の選好	使用可能なすべての薬剤を希望		化学療法に伴う副作用は避けたい
遺伝子シグネチャー	高スコア	中スコア	低スコア

表3 患者背景

Age	<=30	3(0.8)	Stage	I	172(47.5)
	31-50	118(32.6)		II	162(44.8)
	51-70	174(48.1)		III	25(7.0)
	71=<	67(18.5)		IV	1(0.3)
	median age	57.5		unknown	2(0.6)
T	T1	187(51.7)	Grade	1	94(17.7)
	T2	144(39.8)		2	166(45.9)
	T3	19(5.2)		3	89(24.6)
	T4	12(3.3)		NE	13(3.6)
N	N0	256(70.7)	Phenotype	luminal	263(72.7)
	N1	92(25.4)		HER2	22(6.1)
	N2	3(0.8)		L-HER2	18(5.0)
	N3	5(1.4)		TNBC	59(16.3)
	unknown	6(1.7)			

表4 表現型別の腫瘍組織グレード

腫瘍 Type	Grade		
	1	2	3
Luminal	92	155	72
Luminal-HER2	0	9	9
HER2	1	9	11
TN	4	17	36
	97	190	128

図1 Luminal type の Ki-67, Grade,RS (初期治療が手術の症例を含む)

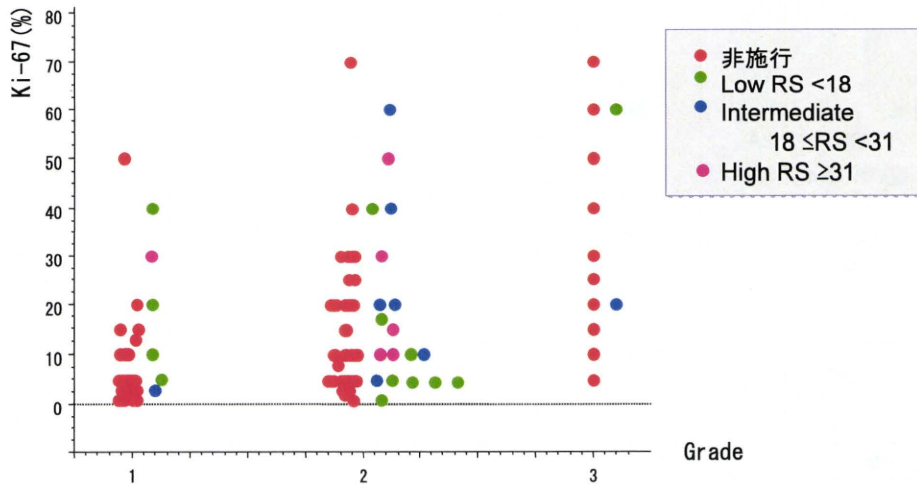


図2 治療法選択と Ki67 index、RS (先行治療が手術の症例を含む、NACEを除く)

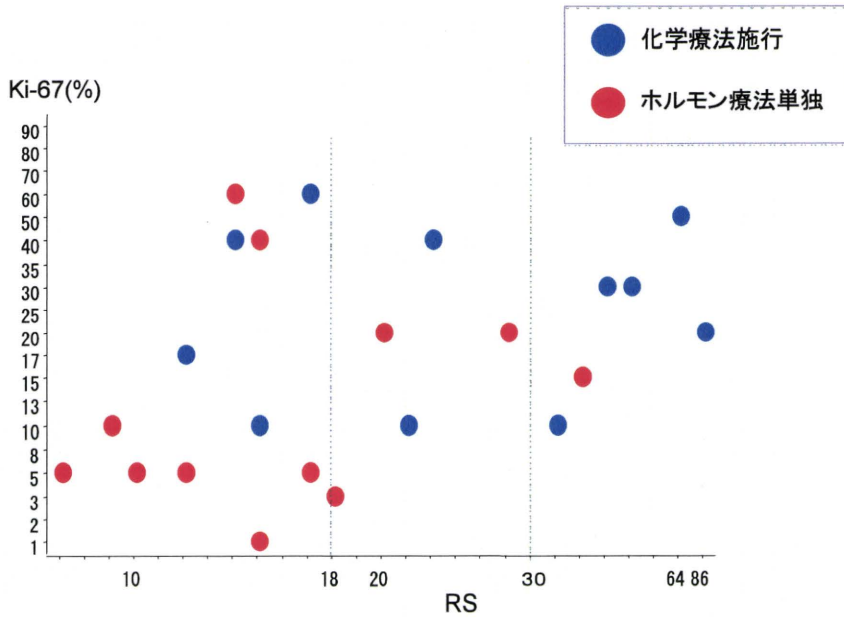
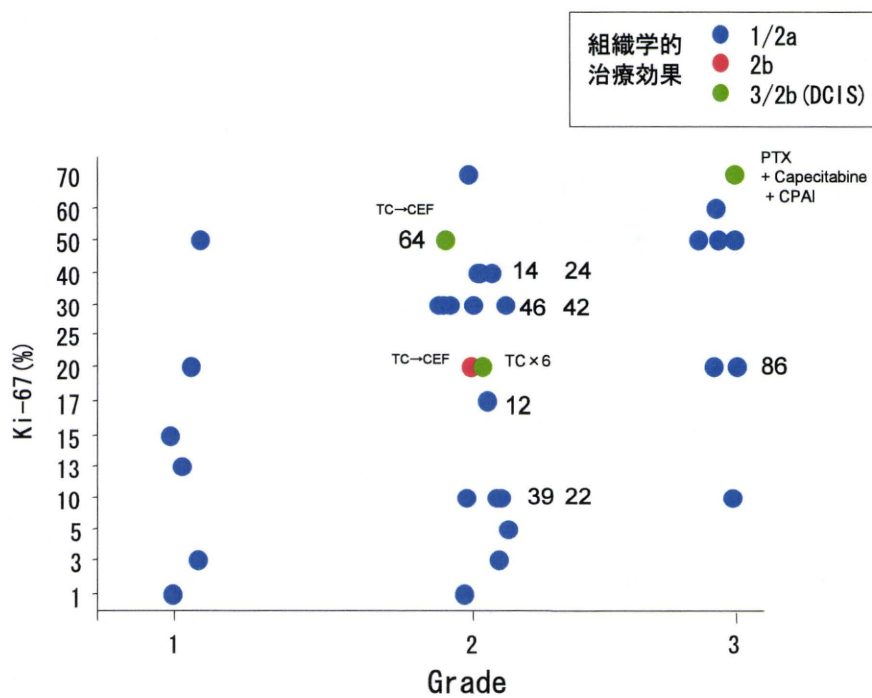


図3 組織グレード、Ki67 index、RS と術前化学療法の組織学的効果
図表内の数値はRS



参考文献

- 1) Goldhirsch A, Ingle JN, Gelber RD et al. Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. *Ann Oncol*. 2009 Aug;20(8):1319-29. Epub 2009
 - 2) NCCN ホームページ: <http://www.nccn.org/index.asp>
 - 3) 日本乳癌学会編:乳がん診療ガイドライン 1. 薬物療法. 金原出版. 東京.
 - 4) Berry DA, Cirincione C, Henderson IC, et al. Estrogen-receptor status and outcomes of modern chemotherapy for patients with node-positive breast cancer. *JAMA*. 2006 Apr 12;295(14):1658-67. Erratum in: *JAMA*. 2006 May 24;295(20):2356
 - 5) Toi M, Nakamura S, Kuroi K, et al. ; Japan Breast Cancer Research Group (JBCRG). Phase II study of preoperative sequential FEC and docetaxel predicts of pathological response and disease free survival. *Breast Cancer Res Treat*. 2008 Aug;110(3):531-9.
 - 6) Gennari A, Sormani MP, Pronzato P, et al. HER2 status and efficacy of adjuvant anthracyclines in early breast cancer: a pooled analysis of randomized trials. *J Natl Cancer Inst*. 2008 Jan 2;100(1):14-20. Epub 2007 Dec 25.
 - 7) Kondo M, Hoshi SL, Yamanaka T, Ishiguro H, Toi M.
 - 8) Economic evaluation of the 21-gene signature (Oncotype DX®) in lymph node-negative/positive, hormone receptor-positive early-stage breast cancer based on Japanese validation study (JBCRG-TR03). *Breast Cancer Res Treat*. 2010 Nov 17. [Epub ahead of print]
- Toi, M, Iwata, H, Yamanaka, T, et al. Clinical significance of the 21-gene signature (Oncotype DX) in hormone receptor-positive early stage primary breast cancer in the Japanese population. *Cancer*. 2010 Jul 1;116(13):3112-

乳癌の治療効果予測のための数理アルゴリズム開発法の検討

研究分担者 内藤泰宏 慶應義塾大学環境情報学部 准教授
研究分担者 杉本 昌弘 京都大学大学院医学研究科 講師
研究協力者 高田 正泰 京都大学大学院医学研究科 乳腺外科 大学院生

本研究の目的は、乳癌治療の効果予測の数理モデルを作り、医療経済負担低減につながる診療のアルゴリズムの構築を行うことである。そこで従来の統計学的な視点とは異なる方法で、臨床情報に各種データ・マイニング手法を適応し、臨床情報間に内在するルールを抽出し、特定の治療効果や診断項目の予測を高精度に行う方法論の確立を行った。特に、臨床情報解析で問題となる、欠損値の存在や、施設間での症例発生率の違いを考慮し、耐性のある方法論を開発した。適応例として、前治療のない症例でリンパ節転移の予測を行い、異なる施設の症例を用いて評価試験も行った。本報告書では、この開発した方法論と適応例の結果を報告する。

A. 研究目的

リンパ節(LN)の転移は予後の予測に重要な因子であり、術前・術後の全身療法の選択や腋窩リンパ節郭清の意思決定に大きな影響を及ぼす。現在では、リンパ節転移を低侵襲に評価することができるセンチネルリンパ節(SLN)生検が用いられているが、侵襲性がゼロではなく、コストも要し、高齢患者には合併症の危険性を増大させる問題などもある。このため、SNL生検を実施するまでに得られている画像所見や針生検所見を組み合わせることでLN転移の予測ができれば、SLN生検に加えてリンパ節転移の予測精度を向上させるだけでなく、SLN生検の省略可能な症例の選別、さらには医療費の適正配分につながる可能性がある。

SLNが陽性の症例において、腋窩に残存する非SLN節への転移を予測するモデルは、腋窩リンパ節郭清(ALND)の省略の可能性を予測する目的で、数多く開発されている(表1に、モデルと精度の一覧を示す)。

しかし、本研究と目的を同じとするLN(SLN)転移そのものを予測しているものは、現在のところ、Memorial Sloan-Kettering Cancer Center (MSKCC)のモデル(Bevilacqua et al. Journal of Clinical Oncology, 2007, 25:3670-3679)が開発されているだけである。非SLN節の予測はSLNから得られる所見を主体にモデルが構築されているため、ほぼ全ての症例で統一した所見の情報を入手することができる。一方、LN転移予測ではSLNを行う以前の、より低侵襲な画像所見や病理所見などの情報から予測モデルを構築しなければならず、こうした所見には患者によっては記録していない所見(欠損値)も多く含まれる。

数理モデルの構築方法として、既存のモデルの多くは、多変量解析の一つである多重Logistic回帰分析(MLR)を用いて、各変数と予測すべき変数の関係をグラフィカルにあらわしたNomogram形式のものである。

表 1 乳癌のリンパ節転移に関連した既存のモデルと予測精度

モデル	予測対象	方法 ¹⁾	AUC値 ²⁾
MSKCC nomogram (2007)	SLN転移	MLR	0.754
MSKCC nomogram (2003)	非SLN転移	MLR	0.77
MD Anderson nomogram (2003)	非SLN転移	MLR	N/A
Tenon scoring (2005)	非SLN転移	MLR	0.82
Louisville scoring (2006)	非SLN転移	MLR	0.68
SNUH scoring (2008)	非SLN転移	MLR	0.82
Stanford calculator (2008)	非SLN転移 ^a	MLR+CART	0.74
Cambridge model (2008)	非SLN転移	MLR	0.84
Ljubljana nomogram (2010)	非SLN転移	MLR	0.77-0.79
MD Anderson nomogram (2008)	NAC後の非SLN転移	MLR	0.78
Katz model (2008)	4つ以上のLN転移	MLR	0.81

注釈 1) MLR は多重ロジスティック回帰、NAC は術前化学療法を示す。

注釈 2) AUC 値は受信者動作特性(ROC)曲線以下の面積を示す。

MLR は、欠損値があった場合基本的には計算ができず、様々な補間方法が考案されている(Vach et al. *Statistics in Medicine*, 1995, 14, 1315-1329)が、どの方法がよいかを決める決定的な要素はない(Fung et al. *Biom J*, 1989, 1, 35-47)。現実的な使い方としては、欠損した変数の取りうる値の最小値と最大値を用いて、予測すべき変数の確率を一定の幅で計算せざるを得ない(Bevilacqua et al. *Journal of Clinical Oncology*, 2007, 25:3670-3679)。あるいは、前もって何らかの方法で欠損値を埋め合わせる必要があるが、例えばマイクロアレイの欠損値処理で用いられる k-means 法でクラスタリングして、欠損値は同じクラスタに属する症例の変数と同値のものを用いる方法は、欠損値以外の変数が類似した症例は欠損値も同じ値になるという前提が働いている。

今回対象とする臨床データに用いることが適切かどうかの判断は難しく、このような特殊な前処理を必要とせず、欠損値そのものに耐性のあるアルゴリズムの適応が好ましい。同様の問題は、決定木、CART モデル、サポート・ベクター・マシン、ニューラル・ネットワークなどでも生じる(Troyanskaya et al. *Bioinformatics*, 2001, 17, 520-525)。

また、予測対象とする症状（今回では LN 陽性と LN 陰性）の発生率が施設間で異なる場合、MLR の予測精度が落ちる問題もある。このため、施設間での発生率の違いを補正する方法も提案されているが(Rouzier et al., *Cancer*, 2006, 107, 1459-1465)、開発したモデルを様々な箇所で使う場合には、予めその施設での発生率が情報として必要となり、

非現実的である。Rouzier らは、MLR を構築した施設の症例群で bootstrap 試験を行い、その症例群だけに特異的にモデルがよい成績を出したわけではなく、異なる症例群でも構築した MLR が精度よく予測できると主張しているが、結果としてはこの試験結果だけでは汎化能力の評価が不十分であることを示している。そのため、モデル構築に用いた症例群だけにオーバーフィットせず、どのような発生率であっても精度の劣化が少ない汎化能力の高いモデルの開発が必要となる。

そこで本研究では、欠損値を含み施設間で発生率が異なるような臨床情報でも、高精度に予測することができる数理モデル構築の方法論を開発した。実際に、複数施設から適応症例のデータを収集し、リンパ節(LN)の転移の予測を行った。今回扱った症例の LN 転移確率を開発した方法と MSKCC のモデルを用いて比較検証を行った。

B. 研究方法

B. 1. 症例情報

学習データとして、都立駒込病院の 148

症例(2005 年-2006 年)と京都大学医学部附属病院の 143 症例(2008 年-2009 年)を用いた。

また、評価データとして、大韓民国ソウル大学病院の 174 症例(2010 年 1 月-4 月)を用いた。適応症例の条件は、1) 針生検で浸潤性乳癌であることが確認されている女性患者、2) 前治療なしに SNL あるいは ALND が行われている、3) 腫瘍径が 4cm 以下を選択した。数理モデルの入力変数として用いる臨床情報は年齢、BMI、閉経の有無、身体所見、画像所見(超音波、マンモグラフィ)、組織型、nuclear grade、ER、PgR、Her2 から 25 変数を用いた。予測する変数としては、LN 転移の有無(転移巣の最大径 $\geq 0.2\text{mm}$)を用いた。各施設の主要な変数の集計値を表 2 に示す。

なお、本研究は京大倫理委員会の承認番号：E-652「原発性乳癌の腋窩リンパ節転移予測モデルおよび術前化学療法による治療効果予測モデルの評価に関する研究」UMIN 登録番号：UMIN000002107 にて承認を受けている。

表 2 施設ごとの主要な変数

		駒込病院 (n=148)		京大 (n=143)		ソウル大 (n=174)	
		No	%	No	%	No	%
年齢*	Median(min-max)	55	(31-85)	60	(26-88)	50	(25-74)
BMI*	Median(min-max)	22.9	(16.6-43.2)	22.3	(14.8-31.4)	23.2	(17.8-37)
閉経*	Pre	56	(37.8)	47	(32.9)	84	(48.3)
	Post	82	(55.4)	96	(67.1)	90	(51.7)
腫瘍最大径(mm)*	Median(min-max)	16	(4-37)	16	(5-35)	19	(4-37)
組織型	Ductal	135	(91.2)	129	(90.2)	160	(92)
	Lobular	5	(3.4)	3	(2.1)	7	(4)
	Others	8	(5.4)	11	(7.7)	7	(4)
Nuclear* grade	1	64	(43.2)	43	(30.1)	4	(2.3)
	2	47	(31.8)	63	(44.1)	82	(47.1)
	3	27	(18.2)	36	(25.2)	88	(50.6)
ER*	Positive	119	(80.4)	114	(79.7)	121	(68.5)
	Negative	27	(18.2)	29	(20.3)	53	(30.5)
HER2*	Positive	22	(14.9)	11	(7.7)	48	(28.2)
	Negative	121	(81.8)	131	(91.6)	125	(71.8)
LN転移	Positive	44	(29.7)	44	(30.8)	41	(23.6)
	Negative	104	(70.3)	99	(69.2)	133	(76.4)

注釈 2) 変数に*が付いているものは、施設間で優位差(p<0.05)があったもの

B. 2. 数理モデルの構築

・使用したモデル

数理モデルには Alternative decision tree (ADTree)を用いた(Freud et al, Proceeding of the 16th International Conference on Machine Learning, 1999, 124-133)。これは、一般的な決定木と同様に、ノードに割り当てられた変数の値によってたどる枝を選択するが、全ての到達可能なノードのスコアを足し合わせて合計した値から特定の事象を予測する確率を算出するものである。CD24 や CD44 の幹癌細胞マーカーから術前化学療法の応答性の予測(Horiguchi et al. J. Med. Dent. Sci. 2010, 57, 165-175)や乳癌の生存率予測(Thongkam et al, Proc 2nd Australasian Workshop on Health Data and Knowledge Management, 2008, 55-62)に用いられた例がある。欠損値があった場合でも、比較的狭い範囲の幅で対象とする事象の予測確率を算出することができる特徴がある。

・モデルの学習と評価の方法

モデル作成にあたって最も注意しなければならないのは、学習データへのオーバーフィットである。学習データからモデルを作成した後、モデルの汎化能力を評価する方法として、データを2分割して、片方のデータでモデルを学習させ、残りのデータで評価を行うクロスバリデーション(CV)が一般的である。しかし、CV の評価データの成績を考慮してモデルのパラメータを選択した場合はなおもオーバーフィットの可能性があると指摘されており(Ambroise et al, PNAS, 2002, 99, 6562-6566)、2重CVなど(Boer et al. Lancet Oncol, 2009, 10, 125-134)様々な別の方法が提案されている。本研究では、駒込病院の症例でモデルを学習させ、京都大学の症例でモデルを評価し、最終的に選択されたモデルをソウル大学の症例に適応する方法を実行した。モデル構築と評価のフローを図1に示す。

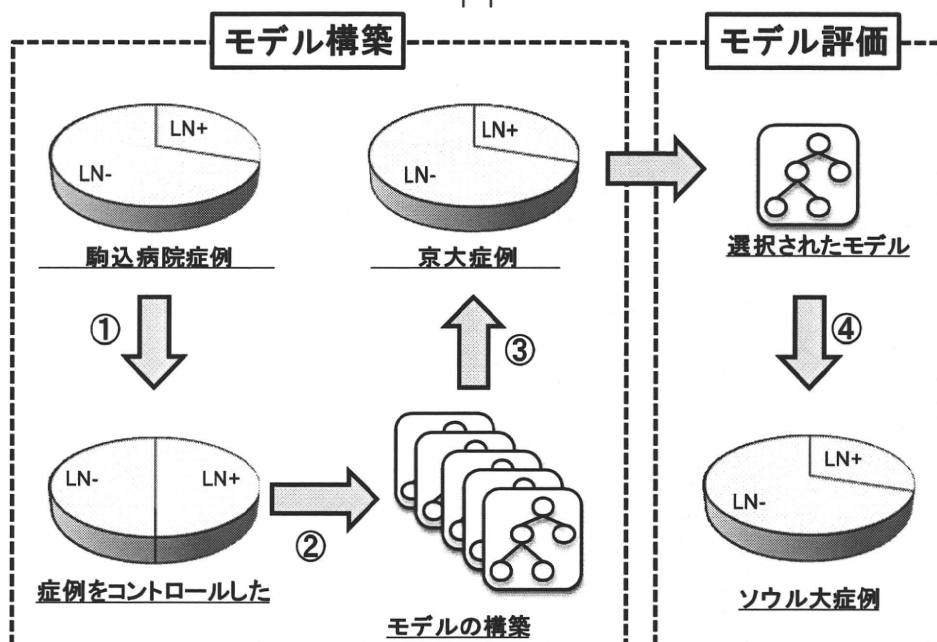


図1 モデル構築と評価のフロー。フロー内の①～④の詳細説明は本文中に記載する

①症例をコントロールした学習データの抽出

また、Rouzierらのモデルで起きているように、学習データの症例に偏りによる汎用性の低下を防ぐために、学習データの症例を全て使わずに、重複を許して症例をランダムに抜き出し、LN-とLN+の症例がほぼ同数になる仮想のデータセットを生成した。

②モデルの生成

①で抽出した学習データに対し、ADTreeモデルを構築した。ADTreeの学習パラメータとしては、ツリー中のノード数(=boosting数)があるため、これを2~20の範囲で変動させた。

③モデルの選択

京大データで最も成績の良かったものを選択し、駒込病院と京大病院の全症例の予測を行った。

④モデルの強化

最終的に開発したモデルの汎化能力や欠損値への耐性を向上させるために、

①の学習データから重複を許してランダムな症例を選択肢、症例数が学習データと同数になるよう仮想的な症例群を作成し、それぞれの症例群に対してモデルを学習させて、各モデルの平均値を最終的な予測とする bagging 法を実施し、予測精度の比較を行った。

⑤モデルの評価

評価データであるソウル大学の症例の予測を行った。

C. 研究結果

開発したモデルの一部を図2に示す。実際には、このツリーと同じノード数のツリーが5つあり、5つの平均スコアを最終的な予測スコアとして、LN転移の確率を計算する。

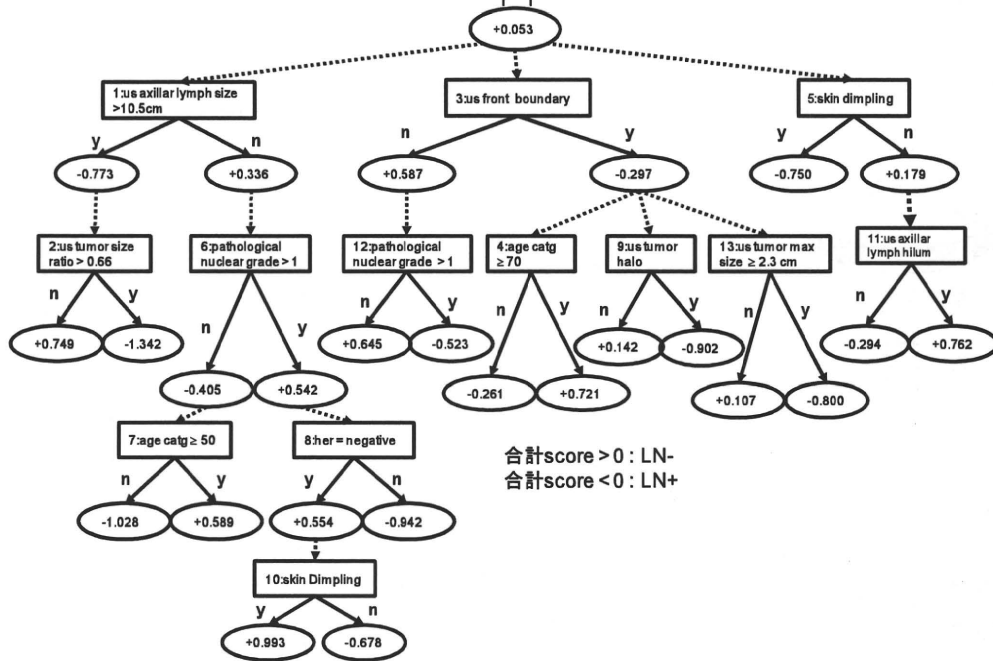


図2 構築したモデルのツリーの一例。一番上の丸ノードから点線のルートは全てたどり各四角のノードでは条件に従い、該当する方へたどる、到達する全ての丸ノードの中のスコアを合算し、合計値が+の場合LN-、-の場合LN+であると予測する。

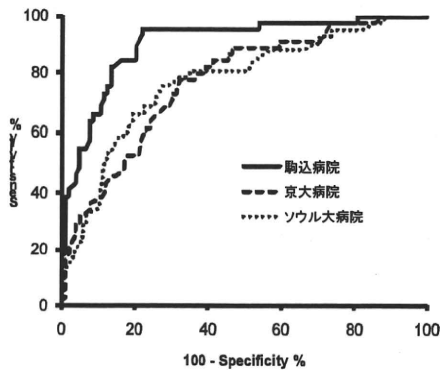


図3 構築したモデルのROC曲線
各施設でのAUC値は、駒込病院で0.917 (95%信頼区間CI: 0.871 - 0.964), $p < 0.0001$ 、京大データでは、0.770 (0.689 - 0.850), $p < 0.0001$ 、ソウル大データでは、0.772 (0.689 - 0.856), $p < 0.0001$

また、ROC曲線を図3に示す。駒込病院データではAUC値0.917と高い値を出し、京大データとソウル大データでは、ほぼ同様のAUC値である、0.770と0.772を出すことができた。1つのツリーに含まれるノード数は13であり、5つのツリーに含まれる変数の数は15種類であった。これをまとめたものを表3に示す。

表3 構築したモデルに含まれる変数の一覧

変数	取りうる値
1) US: 腫瘍最大径	実数値
2) Dimpling	Y/N
3) US: リンパ節最大径	実数値
4) US: halo	Y/N
5) 年齢	20-/30-/40-/50-/60-/70-
6) US: 腫瘍縦横比	実数値
7) Nuclear grade	1/2/3
8) US: 腫瘍、前方境界線の断裂	Y/N
9) MMG: 石灰化の形状	線状分枝状/多角形/淡く不明瞭/微細円形
10) HER2	陽性/陰性
11) MMG: 石灰化の分布	線状/区域性/集簇性/領域性
12) US: 腫瘍の数	なし/1/2以上
13) BMI	<18.5/<22/<25/<30/>30
14) 乳頭分泌	Y/N
15) US: リンパ門	Y/N

また、施設ごとにADTreeが予測した分布のプロットを図4に示す。どの施設でもボックスの中央の横線(全体の中央値)を見ると、LN-は50%以下、LN+は50%以上に分布している。駒込病院の症例ではLN+は最も小さいもので30%程

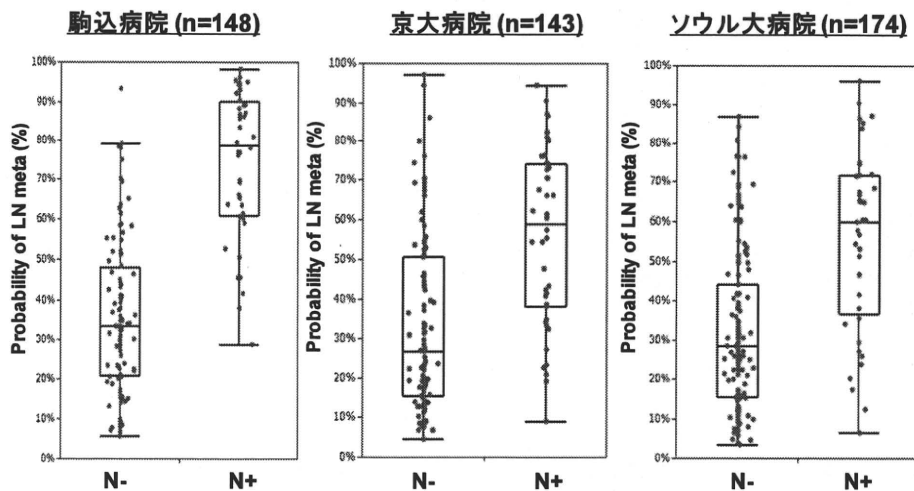


図4 モデルが予測した施設ごとの確率の分布

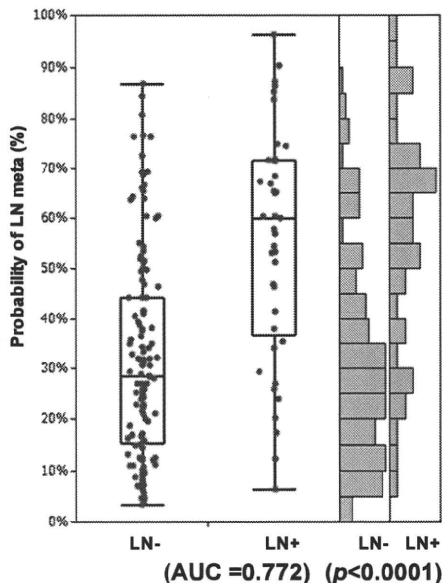
度の転移確率として算出されている。京大データとソウル大データに関しては、LN+の予測確率は広く全体的に広がってしまっているものの、LN-は50%以下に多く固まっていることが確認できる。

評価データとして用いたソウル大の症例を MSKCC の Nomogram に入れて、ADTree と分布の違いを比較した。この結果を図5に示す。ただし、MSKCC のモデルで必要とされるリンパ管侵襲 (LVI) に関しては、手術後に入手できる情報で得られないため、LVI+と LVI-の値を用いて両方でLN+の確率を予測し、平均値を用いた。ADTree を使ったデータのヒストグラムに関して、LN-ではヒストグラム 30%台以下に分布が集中しており、LN+は70%台を中心として広がっていることが確認できる。一方 MSKCC の Nomogram は、LN-、LN+のどちらも20%~80%あたりに予測分布が集中しており、50%近傍でのLN-とLN+の予測分布

のオーバーラップが大きい。AUC 値でも、ADTree では 0.772 ($p < 0.0001$)、Nomogram では 0.64 ($p = 0.013$) と、Nomogram の予測精度はかなり低い結果であった。

ADTree で算出した予測確率に関して、実際のLN転移確率と比較したキャリブレーションプロットを図6に示す。予測値と実測値が一致した場合、 $Y=X$ の45度の直線状にのる。京大データでは40%程度まではほぼ直線状にプロットが乗っているが、60%を超えたあたりから下に乖離しており、LN+を過剰に予測している。ソウル大データでは、全般的にLN+過剰予測の傾向がある。これは、最初にLN-とLN+の症例数を制御して学習データを作成した際、元々の症例セットではLN+の症例が少ないために、LN+を増やすようにして学習データが作成され、LN+の予測が過剰になっていると考えられる。

a) ADTreeのソウル大データ(n=174)



b) MSKCC Nomogram (n=161, defect=13)

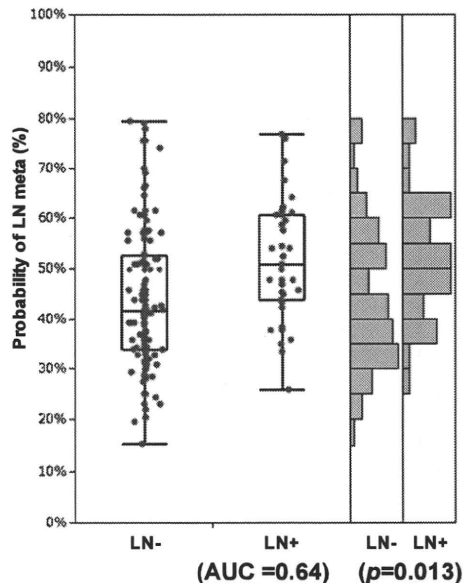


図5 評価データに関するモデルの予測確率分布とヒストグラム

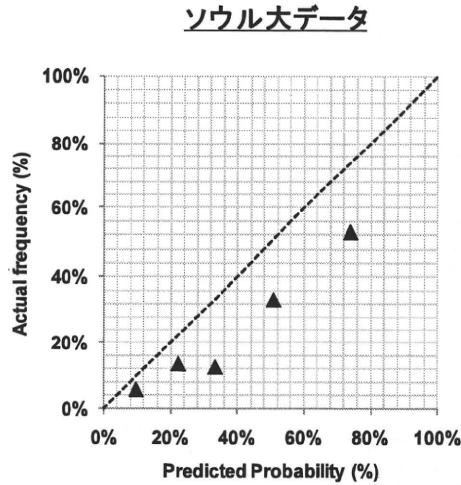
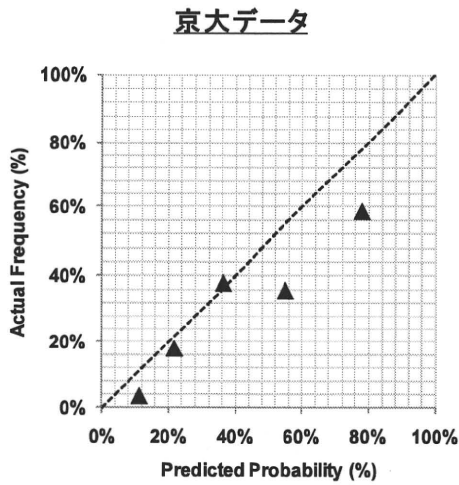


図6 京大データとソウル大データのキャリブレーションプロット。LN転移確率を5分割し、実際の確率と予測値の確率をプロットした。

ADTree のモデルモデルに含まれる変数は、Odds 比のように直接変数の重要度を評価できないため、変数の値をランダムに変更して、AUC の劣化度から、モデル内での変数の相対的な重要性を評価した(図7)。リンパ節最大径と超音波所見である halo の重要度が突出しており、dimpling と超音波の画像所見である腫瘍の最大径と前方境界線の断裂の重要度が高い。他の変数は一様に類似した影響度であった。

D. 考察

今回開発した ADTree の選択された 15 の変数は、腫瘍径、年齢、腫瘍の数(多発の有無)、nuclear grade、ER、組織型が MSKCC の Nomogram の変数と一致している。MSKCC のモデルに含まれる PgR は ADTree に含まれなかったが、MSKCC のモデル内でも相対的な重要度は最も小さく、PgR の +/- の違いは Nomogram の 200 ポイント(pt)中、5pt に相当する(LN+の予測確率では 50%近傍の場合、10%程度に相当する)。また、腫瘍の局在に関しても、MSKCC のモデルには含まれ、ADTree モデルには含まれていない

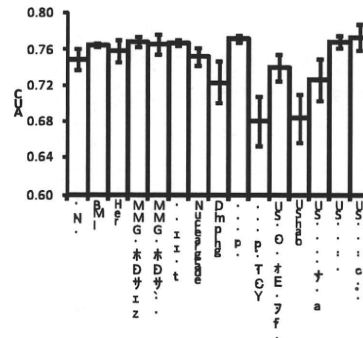


図7 ソウル大データでの感度解析の結果。本来のデータに対し、変数ごとにランダムに取りうる別の値に変更して、AUC 値の計算を 200 回行った。バーは AUC 値の平均値、エラーバーは標準偏差を示す。

が、代わりに他の画像所見を多く含む結果となった。超音波の前方境界線の断裂と halo はいずれも腫瘍の浸潤性に関する所見であり、リンパ節転移の予測因子としては妥当なものである。また、ADTree ではさらにリンパ門の有無や超音波で確認できるリンパ節最大径なども含まれており、変数としてはリンパ節転移影響因子が網羅されている。

MSKCCのモデルではLVIがリンパ節転移を予測したい時点(手術前)では通常得られないため、実際の使用時には欠損値として扱わざるを得ない。しかし200pt中、40ptに相当し、LN+の予測確率では50%近傍の場合、50%程度に相当する。極端な場合はこの値次第で予測がLN+の予測値が25%と75%と大きな範囲で揺らぐことになる。今回作成したADTreeでは、同様に変数の欠損があった場合、モデルそのものもこの予測範囲が比較的小さいものではあるが、更にbagging法を用いることで多数のADTreeで予測を安定化させ、揺らぎを小さくすることに成功しており(図7)、実運用において重要な特徴である。

Nomogramの従来の学習である、学習データで全症例を使ったモデルの構築では、発生率の違う施設での性能劣化が顕著であることが示唆されており、実際に、MSKCCモデルでソウル大の症例を予測した場合もAUC=0.64と極めて低い性能であった。今回行った学習データでの症例数の制御は、この問題を軽減し、京大データとソウル大のデータでほぼ同値のAUC値0.770と0.772を出したが、同時に症例を増やした側(LN+)に過剰予測をする欠点も生じた(図6)。これらはトレードオフの問題であるため、別の指標をもってパレート解中の一つを選択する選択することが必要である。また、本システムを診断支援システムに発展させてゆくためには、どの予測確率であれば、リンパ節転移の有無を確実に予測できるか、あるいは予測できないかの境界を示し、より多くの施設や症例データで追加評価を行う必要がある。

E. 結論

リンパ節転移に関して、欠損値や症例の偏りがある臨床データから転移の確率を高精度に予測し、高い汎化能力を

持つ数理モデルの構築とその評価を行った。学習データの予測精度がMSKCCのNomogramと比べて高かっただけでなく、転移発生率の違う評価データでも今回開発した方法では高い精度で予測することができた。今後、インターネットなどを通して広く利用してもらえるように、Web上でデータの入力と結果の表示を行うツールを構築し、同時に、前向き試験に向けた準備も行っている。今回開発した手法は、リンパ節転移に関して適応したが、適応事例に絞った手法でなく、汎用的な手法であるため、今後は、21遺伝子シグナチャーと臨床情報の組み合わせによる再発リスクの高精度な予測や、ホルモン療法や分子標的治療の効果予測とその因子の抽出を行い、適応範囲を拡大してゆく。

G. 研究発表

1. 論文発表

Horiguchi K, Toi M, Horiguchi S, Sugimoto M, Naito Y, Hayashi Y, Ueno T, Ohno S, Funata N, Kuroi K, Tomita M, Eishi Y. Predictive value of CD24 and CD44 for neoadjuvant chemotherapy response and prognosis in primary breast cancer patients. J. Med. Dent. Sci. 2010, 57, 165-175

2. 学会発表

なし

(発表誌名巻号・頁・発行年等も記入)

H. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

特願2010-099327 予測装置及びその学習装置並びにそれらのコンピュータプログラム

2. 実用新案登録

なし

3. その他

なし

抗がん剤効果及び再発リスク予測による乳がん治療方針決定の医療経済評価

研究分担者 近藤 正英 筑波大学大学院准教授

研究要旨 ホルモンレセプター陽性早期乳がん患者において、再発リスクの評価と化学療法の効果予測に基づいて、術後化学療法の適応を示し、効果の見込めない化学療法を回避することによって、毒性軽減につながる多遺伝子検査（multigene assay）のなかで、国際的に実用化されている21遺伝子シグネチャ（OncotypeDX®）の日本での検証研究に基づいた経済評価を行った。推定された増分費用効果比は、N-，Her2±で¥384,828/QALY、N±，Her2-で¥434,096/QALY、N±，Her2±で¥568,533/QALY、N±，Her2-で¥614,765/QALYであった。これらは費用対効果に優れると解釈することができ、公的医療制度でのこの検査の提供が社会的に受け入れられるものであることが示唆された。

A. 研究目的

ホルモンレセプター陽性早期乳がん患者において、再発リスクの評価と化学療法の効果予測に基づいて、術後化学療法の適応を示し、効果の見込めない化学療法を回避することによって、毒性軽減につながる多遺伝子検査（multigene assay）のなかで、国際的に実用化されている21遺伝子シグネチャ（OncotypeDX®）の日本での経済評価を行った。

21遺伝子シグネチャは、米臨床腫瘍学会（ASCO）や全米総合癌情報ネットワーク（NCCN）の臨床ガイドラインに組み込まれている治療方針決定のための遺伝子診断であるが、費用が高い（日本では45万円）ことが特徴のひとつである。ただし、たとえば、多遺伝子検査によらない従来の診断情報に基づいて、術後化学療法の適応とされた患者について、多遺伝子検査によって費用が高い（最大100万円程度）化学療法が回避される場合もあり、医療財源の効率的使用という観点からは、適切な経済評価が求められる。

B. 研究方法

日本での21遺伝子シグネチャの経済評価としては、研究分担者らが、先行研究を報告している（Kondou M et al (2008) Breast Cancer Res Treat. 112(1):175-87.）。ただし、この研究では、アウトカム推定の根拠として米国での検証研究の結果を用いていた。その後、日本での検証研究の結果を研究代表者らが報告した（Toi et al (2010) Cancer. 116(13):3112-8.）。

そこで、本研究では、先行研究の経済モデルに、日本での検証研究の結果を組み込んだ。

（倫理面への配慮）

本研究は、文献による経済モデリングであり疫学研究の倫理指針や臨床研究の倫理指針にかかるものではない。ただし、経済評価の対象である医療技術は商業的に提供されているものである。そこでその商業的提供者に関して利益相反がないことを確認した。

C. 研究結果

Table 1が日本での検証研究の主な結果である。21遺伝子シグネチャがリンパ節転移陰性（N-）の200例とリンパ節転移陽性を含めた280例（N±）の両方で、多遺伝子検査によらない治療方針決定として日本で標準的に用いられているSt Gallen勧告による再発リスク分類とは異なる分類で、再発リスクを予測できている。St Gallen分類でのNo definitive indicationでの術後化学療法の実施率を50%と仮定すると、21遺伝子シグネチャを用いて治療方針を決定した場合に術後化学療法を実施する割合は、N-では、39.3%から32.5%へ、N±では、40.7%から32.5%へと減少する。

Fig. 1が判断樹とマルコフモデルからなる経済モデルである。またTable 2は、経済モデルに組み込んだ確率とクオリティ・オブ・ライフの調整のための効用値である。さらにTable 3は、費用である。

Table 4が経済評価の結果である。N-での増分費用効果比は、¥384,828/QALY、N±では、¥568,533/QALYであった。

Fig. 2は感度分析の結果を示したトルネード図である。いずれの仮定の変更によっても増分費用効果比が¥1,500,000/QALYを超えることはなかった。

日本での検証研究ではHer2±の症例が含まれていたため、Table 4の結果にはHer2+の症例も含まれたものであるが、21遺伝子シグネチャの適応としては、Her2-に限ることも考えられる。Table 5がこの適応を仮定した場合の経済評価の結果である。N-での増分費用効果比は、¥434,096/QALY、N±では、¥614,765/QALYであった。

D. 考察

医療の技術革新を公的医療制度のもとで普及させる際の医療経済的効率性の基準としては、おおよそ、1QALYを獲得するために500万円程度までは、社会は支払う意思があるとされている。この基準と比較すると、Table 4とTable 5に示された結果は、N-/N±及びHer2+/-に関わらず、費用

対効果に優れ、公的医療財源でその費用をまかなうことが正当化できるということが示された。

E. 結論

国際的に実用化されている21遺伝子シグネチャ (OncotypeDX®) の日本での臨床導入は、日本での臨床的検証研究を踏まえた経済評価によって費用対効果に優れることが示された。この結果は、医療の経済評価の枠組みからは、公的医療制度でのこの検査の提供が社会的に受け入れられるものであることを示唆するものである。

G. 研究発表

1. 論文発表

Kondo M, Hoshi SL, Yamanaka T, Ishiguro H, Toi M.

Economic evaluation of the 21-gene signature (Oncotype DX®) in lymph node-negative/posi

tive, hormone receptor-positive early-stage breast cancer based on Japanese validation study (JBCRG-TR03).

Breast Cancer Res Treat. 2010 Nov 17. [Epub ahead of print]

DOI: 10.1007/s10549-010-1243-y

2. 学会発表

なし。

H. 知的財産権の出願・登録状況

1. 特許取得

なし。

2. 実用新案登録

なし。

3. その他

なし。

厚生労働科学研究費補助金
分担研究報告書

術前化学療法における治療効果と予後に関するCD24とCD44の意義に関する研究

研究分担者 黒井克昌 がん・感染症センター都立駒込病院 外科 臨床試験科 部長
研究協力者 堀口和美 がん・感染症センター都立駒込病院 外科 医師

研究要旨

乳癌幹細胞のマーカーであるといわれている、細胞表面マーカーであるCD44、CD24を術前化学療法施行患者の針生検検体を用いて免疫組織化学的に検索した。これら二つのマーカーと、術前化学療法の治療効果および予後との関連性について解析を行った。

A. 研究の目的

術前化学療法は、乳房温存療法が成功する可能性を増加させ、pathological complete response (pCR)を得られる症例に関してはその予後の延長が得られ、また、術前に行うことによって生物学的なモニタリングが可能であるといったメリットから、化学療法を要する手術可能乳癌患者において施行されている。術前化学療法の治療効果については、未だ、施行前に確実に予測しうる手段はなく、予測ができれば患者利益は勿論のこと、医療経済の観点からも有益であることが予想される。

今回CD44およびCD24と、術前化学療法の治療効果および予後との関連性の有無を検討した。また、従来の統計学的解析方法に加え、Data mining software WEKAをもちいたAlternating Decision Tree解析を行い、治療効果を予測するルールを模索することを目的とした。

B. 研究方法

浸潤性乳管癌の診断目的に施行する針生検検体を用いて、通常の日常臨床で行うホルモンレセプターやHER2といった免疫染色に加えてCD44およびCD24の免疫染色を施行した。CD44およびCD24の発現を negative, 1+ (50%未満), 2+ (50%以上)の3段階に分類し、その結果と、その他の因子との関連性、術前化学療法の治療効果、予後を解析した。統計解析

にはJMP® 8 (version 8. SAS Institute Inc.)を用いた。術前化学療法の治療効果を予測するために、pathological complete responseを予測するツールとしてAlternating Decision Tree解析を用いた。

この研究プロトコルに関してはがん・感染症センター都立駒込病院 倫理委員会においてその承認を得ている。研究対象者に対する人権擁護上の配慮をおこなっていること、研究方法による研究対象者に対する不利益がないこと、危険性は排除されていること に関して文書を用いて説明と同意 (インフォームド・コンセント)を行っており、倫理面の問題がないと判断している。

C. 研究結果

CD24の発現状況と、術前化学療法の治療効果との間には統計学的有意差を以て関連性があることが判明した。また、Alternating Decision Tree解析を用いて、複数の因子から術前化学療法のパथолоジカル完全反応症例を予測できることがわかった (AUC 0.981)。

また、CD44が高発現であることが、無再発生存期間の延長と関連性があることが判明した (p=.0002 Log rank test)。

D. 考察