

データ分析の医療への応用
Application to Medical Care of Data Analysis

岩井 幸子¹⁾ 石野 福弥²⁾ 阿万 広大²⁾
Sachiko Iwai Fukuya Ishino Kodai Aman

1. 問題の背景

尿路癌における BCG 療法は有効な治療法のひとつとして用いられ、golden standard な治療法として多くの尿路癌症例に使用されている。しかしながら、BCG 療法は副作用がきわめて強く、重篤な合併症を惹起したり、治療途中で drop out する例も少なくない。BCG 濃度についての検討報告もしばしばみられ、高濃度 (80mg) と低濃度 (40mg) では高濃度での副作用が強いといわれている。しかしながら当初から施行されていた治療マニュアルの高濃度を範として慣習的に施行する施設も多く、再発の点からは低濃度の BCG を使用することには相克がある。本発表では BCG 療法の施行された尿路癌患者の尿細胞診標本を用い、患者毎から得られる 17 項目の (性別、年齢、BCG 濃度、BCG 注入回数、癌の発生部位、癌の組織型、癌の異型度など) のパラメタから統計的解析を用いて再発予測式を作成した。適切な BCG 注入量と注入回数を明らかにすることで過剰な注入を回避し、副作用の軽減を可能にすることを目的とする。医療現場で新しい治療法や改良法を施行するためにはさまざまなプロセスがあり、実際に稼動するためには多くの時間と研究報告例が必要となる。本研究はさまざまな分野で用いられている統計学的分析が医療でも有効に利用され、患者の Quality Of Life の向上に寄与することの端になると期待される。

2. 測定データの収集と分析、結果

1) 再発予測式作成のための症例の収集

BCG 療法後の再発予測式作成にあたり、1990 年から 2004 年まで産業医科大学病院泌尿器科において受診された患者データベースより尿路癌症例を選択し、さらに BCG 療法施行例を検索した。ついで、診療録 650 冊ならびに自然尿細胞診標本 1632 枚より情報を収集し、統計処理に必要なデータが揃う 102 例を抽出した。

2) 得られた情報の種類と事前処理としてのパラメタ設定

目的変数として再発の有無、説明変数として、性別、年齢、BCG 量、BCG 注入回数、細胞診最終陽性回、BCG 注入部位 (膀胱)、BCG 注入部位 (上部尿路)、癌組織型、癌異型度、副作用 (疼痛)、副作用 (発熱)、副作用 (膀胱刺激症状)、副作用 (中止)、予後 (2 年以後の尿路上皮癌の状態)、予後 (2 年以後の尿路上皮癌以外での死亡)、他の治療法との併用、これら合計 17 項目 (表 1) を設定した。

尚、今回使用した症例の膀胱のみでの再発率は 26.2%であるが、(株) 日本 BCG の報告によると BCG 療法適応例中膀胱対象例での全国平均は 7~33%となっていることより採用したサンプルは妥当な値とみなすことができる。

産業医科大学病院 病理部¹⁾、早稲田大学大学院情報生産システム研究科²⁾

表1 設定したパラメタ

変数	実数	ダミー変数 : 0	ダミー変数 : 1
再発の有無		なし	あり
性別		女性	男性
年齢	○		
BCG量	○		
BCG注入回数	○		
細胞診最終陽性回	○		
注入部位(膀胱)		注入なし	膀胱注入
注入部位(上部尿路)		注入なし	上部尿路灌流
癌組織型		単一	重複
癌異型度	○		
副作用(疼痛)		なし	あり
副作用(発熱)		なし	あり
副作用(膀胱刺激症状)		なし	あり
副作用(中止)		なし	あり
予後(2年以後尿路癌の状態)		良好	尿路癌死
予後(尿路癌以外の状態)		良好	他因死
他の治療法との併用		TURあるいは生検のみ	TUR+他の治療法

3) 種々のパラメタを用いて重回帰分析を行い、再発予測式を作成

重回帰分析では再発の有無を0と1で表わしたことから判定点は0.5とした。重回帰分析で重決定係数の高値を示した独立変数は最終陽性回であり、さらにもっとも高値を示したのは最終陽性回とBCG注入回数をモデルとして用いた場合であり、この2つのパラメタが説明する率としてR2乗値として47.7%を得ており十分な相関が認められる。回帰式全体の有意性は最終陽性回、また最終陽性回とBCG注入回数の2つのモデルでいずれも0.01%水準で有意であった。次に最終陽性回で標準偏回帰係数は0.672で正に影響力を有し、その有意確率は0.000と極めて高かった。また、BCG注入回数の標準偏回帰係数は-0.184で負に有意な影響力を示している。その有意確率は0.013であり、0.05水準で高い。Excel 2003分析ツールならびにSPSS(SPSS 13.0 for Windows)による重回帰分析の結果、次のような再発予測式が得られた。

$$Y = 0.661 + 0.133 * \text{最終陽性回} - 0.092 * \text{BCG注入回数}$$

ここで注目すべき点は、再発率予測の説明変数としてBCG量が現れないことである。

4) 予測式の有効性を検証

統計分析に用いた102例による検証を行い、その後に当院ならびに他院でのBCG療法施行例37例を用い、検証を行った。(水光会総合病院1例、九州労災病院10例、新日鉄記念病院4例、門司労災病院8例、産業医科大学病院14例)統計に用いた症例での再発予予測的中率は80.4%であり、検証例での的中率は86.5%であった。全例では82%の高的中率を得ることが可能となった。検証に用いた症例のうち3例が予測から外れる結果となったが、その原因は後ほど個別に理由付けを行う。

一方、予測判別式を実際の医療現場で使用するには煩雑さが懸念されるため、簡易の再発予測早見表を作成した。この

早見表(表2)を利用した再発予測の的中率は統計分析例で89.2%, 検証例で91.9%, 全例で89.9%と再発予測式での的中率よりも高い結果を得た。

表2 早見表

最終陽性回 BCG 注入回数	0	1	2	3	4	5	6	7	8
1	0.753	0.702	0.835	0.968	1.101	1.234	1.367	1.500	1.633
2	0.577	0.610	0.743	0.876	1.009	1.142	1.275	1.408	1.541
3	0.462	0.518	0.651	0.784	0.917	1.050	1.183	1.316	1.449
4	0.354	0.426	0.559	0.692	0.825	0.958	1.091	1.224	1.357
5	0.248	0.334	0.467	0.600	0.733	0.866	0.999	1.132	1.265
6	0.142	0.242	0.375	0.508	0.641	0.774	0.907	1.040	1.173
7	0.036	0.150	0.283	0.416	0.549	0.682	0.815	0.948	1.081
8	-0.070	0.058	0.191	0.324	0.457	0.590	0.723	0.856	0.989
9	-0.176	-0.034	0.099	0.232	0.365	0.498	0.631	0.764	0.897
10	-0.282	-0.126	0.007	0.140	0.273	0.406	0.539	0.672	0.805
11	-0.389	-0.218	-0.085	0.048	0.181	0.314	0.447	0.580	0.713
12	-0.495	-0.310	-0.177	-0.044	0.089	0.222	0.355	0.488	0.621

5) 他の分析結果

今回の検討は多変量解析のうち要因分析と予測に有効な重回帰分析を用いたがこれは一次の多項式で表わされる。次数の高い多項式を利用することでデータ中の独立変数の動く範囲内での的中率は高くなると考え、重回帰分析の中で二次式、三次式で表わされる多項式回帰でも検討を行った。重回帰式ではP個の変数を使用されるが、多項式回帰では1個の独立変数のP個のべき乗が使用されるという相違がみられる。再発に最も寄与率の高い最終陽性回のみを用い、一次式での単回帰分析を比較対照とした。重決定係数は高次になるほど、僅差ではあるが高値を示した。

表3 重回帰分析—多項式結果

	重相関係数 R	重決定係数 R ²	自由度調整済み重決定 係数 R ²	有意確率
一次式	0.616	0.443	0.438	0
Y = -0.037 + 0.132 × 最終陽性回				
二次式	0.681	0.463	0.452	0
Y = -0.196 + (-0.011) × 最終陽性回 ² + 0.232 × 最終陽性回				
三次式	0.693	0.48	0.464	0
Y = -0.082 + (-0.004) × 最終陽性回 ³ + (0.038) × 最終陽性回 ² + 0.065 × 最終陽性回				

本研究での判別分析は BCG 療法施行例が再発する群、再発しない群のどちらに属するのかを判別する。判別分析の結果より判別式を作成し、検証例での 86.5%、全例で 82% の予測的中率は、重回帰分析による的中率と同様の結果であった。

ロジスティック回帰分析は、従属変数が 2 値であるときのための重回帰分析であり、独立変数と従属変数の間に直線ではなく S 字形のロジスティック曲線を当てはめる。この方法は計算が煩雑ではあるがとくに本研究のような再発の有無などという病気の発症や死亡の分析などには適した方法といえる。このロジスティック回帰分析では予測的中率 84.3% を得た。これは、直線で表わされる重回帰分析の 80.4% と近似した結果となり、よりの中率は高かった。また、検証例では 91.9% が得られ、全例で 86.3% の高い中率を得ることが可能となった。

予測早見表による予測的中率が高いのは最終陽性回および BCG 注入回数に回数毎に予測値を入力しており、1 回ごとに range を有するためであり、ロジスティック曲線は早見表での再発と非再発の境界域に一致している。

内的構造分析の一方法として主成分分析を行った結果、本研究で用いたパラメタの 70% が再発に関する情報であり、BCG 療法の決定には年齢や性別による影響が窺われた。本研究で収集したパラメタは大きく 3 つの成分から成立しており、再発に関する成分と残り 2 つは BCG 療法内容に関する成分であると考えられる。最終的には 5 つのパラメタ (性別、年齢、BCG 濃度、BCG 注入回数、尿細胞診最終陽性回) で説明可能であった。

37 例中 3 例の誤判定例では 2 例が BCG 療法終了後膀胱全摘術を受けたために再発を考え難い症例であり、摘出していなければ再発していた可能性は高い。1 例は細胞診の誤診断であり、診断者による診断基準のずれがこの結果にあらわれたのだと考える。この点で診断基準のコンセンサスが必要である。これらの誤予測の理由付けを妥当とみなせば、100% の予測的中が得られたことになる。日常業務の中で論者がかねてから感じていた上部尿路癌での再発率の高さに関しても BCG 注入回数と上部尿路の 2 つのパラメタで重回帰分析を行った結果、上部尿路では BCG 注入回数に関係なく、上部の腫瘍の存在そのものが再発の可能性を高いものにすることが確認できた。再発の判別点が 0.5 であり、判別予測式の定数が 0.448 であることから BCG 療法の施行される多くの上部尿路癌での再発の可能性が考えられる。すなわち、上部尿路癌には BCG 療法は適さない可能性が高い。

3. 期待できる効果と今後の展開

本研究で得られた結果より第一に再発の予防としてあるいは副作用の軽減を図る手段として BCG の最適注入回数を見つけることが挙げられる。今回の研究結果よりすべての当該患者に同一の治療法を施行することの矛盾を解決する方法として、再発予測式に則り作成した判別早見表の使用を推奨したい。また、BCG の濃度は 40mg と 80mg のほぼ 2 パタンであるが、80mg でより副作用が強いといわれているものの 80mg は従来から主流であり、濃度の問題は議論の俾仲するところであるが、本研究では 40mg と 80mg の間に有意差がなかった点より論者は 40mg の使用を強く薦めるものである。この研究は最適な BCG 療法マニュアルとして低濃度、多数回の BCG 注入の暗示をするものであり、副作用の軽減と再発予防のため、臨床と連携し、さらなる検討を進める計画である。

今回、論者が専門的に研究している尿路癌の再発予測や副作用についての分析を行ったのであるが、現段階で現在増加傾向にある乳癌の進展予測についても 100 例の症例を用い、統計分析を進めている。ここでも重回帰分析により 72.9% の進展予測的中率が得られており、データ工学の医療への応用が今後の臨床医学に寄与することが窺われる。

(参考文献)

1. 甘利俊一. 多変量解析の展開, 岩波書店:2002
2. 古谷野亘. 多変量解析ガイド, 川島書店:1998
3. Sachiko IWAI. Clinical cytological study of BCG treatments in the urothelial carcinoma using the statistics analysis-Introduction of formulate the recurrence ; Acta Cytologica Japonica : 2006 (program of publication)