

# 地域がん登録資料に基づくがん患者の 治癒確率の推定

伊藤 ゆり<sup>1</sup>・杉本 知之<sup>2</sup>

(受付 2011年1月21日; 改訂 3月10日; 採択 4月27日)

## 要 旨

がん患者の予後を示す指標として5年生存率が用いられてきたが、がんの部位や進行度によって予後は異なるため、一律に5年で判断するのは難しいといえる。一方、がんの5年生存率に代わる指標として治癒率を推定するための方法論の整備がなされてきている。治癒モデルは、以前より提案されていた(Boag, 1949)が、1990年代後半に相対生存率の考え方がとり入れられるまでは、地域がん登録資料による適用は多くなかった。近年、方法論およびソフトウェアの開発により、地域がん登録資料による適用が欧米を中心に盛んになってきており、本論文では、地域がん登録で用いられる治癒モデルとその方法論の説明とともに、我が国の地域がん登録資料に適用した例を紹介した。大阪府がん登録資料より1975-2000年に診断された胃がん患者について、診断時期を5期間に区分し、治癒確率および非治癒生存分布の中央生存時間の推移を評価した。胃がん患者の治癒確率は1975-2000年の25年間で、約20ポイント向上し、約50%以上の患者が治癒すると推定された。また、治癒確率の向上は男女とも20%近くが早期診断の増加による進行度分布の変化により説明できると推定された。進行度別にも治癒確率が大きく向上していることより、治療技術の向上の影響も大きいことが示唆された。

キーワード：治癒モデル、がん登録、相対生存率、胃がん。

## 1. 背景

がん医療の向上に伴い、がんは部位や治療開始時の進行度によっては、長期生存を可能とする疾病となってきた。ただし、がんはなお国内の全死因トップであり、がん患者の長期予後に関する情報とその利用は、がん患者やその家族、がん医療の従事者にとって重要なものである。地域がん登録システムは、地域に発生する全がん患者の情報を登録し、長期間の追跡調査により、生存・死亡を把握することにより、精度の高い母集団生存率の測定を可能とするものである。がん患者の予後を示す指標としては、伝統的に、5年生存率を用いることが多く、このような治療開始後5年の区切りによる推定生存確率は、がんが治癒したとする簡便な目安としての意義をもつ。ただし、がんの部位や種類によって、その性質や予後が異なるため、一般に、一律に5年で判断できると考えるのは難しく、そのため、がんの5年生存率に代わる指標として、もしくは本来考えるべき指標である、治癒率(治癒確率)とその推測は、この領域において

<sup>1</sup> 大阪府立成人病センター がん予防情報センター：〒537-8511 大阪府大阪市東成区中道 1-3-3;  
itou-yu2@mc.pref.osaka.jp

<sup>2</sup> 弘前大学 大学院理工学研究科：〒036-8561 青森県弘前市文京町 3 番地; tomoyuki@cc.hirosaki-u.ac.jp

重要な意義をもっている。

がんの治癒は、個人レベルにおいては、がん細胞の消失を意味するが、この判断は、通常、困難である。Boag (1949) は、集団レベルのがん患者のデータに対して、がんによる死亡とその他の死因とを区別し、統計モデル(治癒率を導入した生存時間モデル)(以降では、治癒モデルと略す)を用いて、治癒率を推測する統計的方法を提案した。その後、多くの著者(例えば、Farewell, 1986; Maller and Zhou, 1996)により、パラメトリック治癒生存モデルについて、さらなる研究がなされた。現在、生存時間データに対し、標準的な回帰解析法を与える Cox モデル(Cox, 1972)とその解法は、治癒モデルとはやや異なる解釈の方向性をもつ。Cox モデルは、臨床試験や臨床研究からのデータに対し、潜在生存分布を特定することなく、限られた観察期間の中で予後への回答を用意しやすい指標である相対ハザード比についての推測を与えるが、ハザード比の指標は治癒率への直接的な解釈を与えるものではない。治癒モデルの推測のために、非治癒の潜在生存分布を特定することなく利用できる回帰モデル、いわゆる、セミパラメトリック治癒モデルが開発され、その解法の研究は進んでいる(Kuk and Chen, 1992; 杉本・後藤, 1999; Peng and Dear, 2000; Sy and Taylor, 2000; Lu and Ying, 2004; Sugimoto et al., 2005; Sugimoto and Hamasaki, 2006)。

地域がん登録資料では、大規模な母集団からデータが収集されるため、治癒モデルの適用がより期待される領域である一方、がんによる死亡か、他の原因による死亡かの特定が難しく、上述のような通常の生存時間解析において開発されてきた解析法を直接的に適用しにくい背景がある。そのため、地域がん登録資料に基づいて、治癒率を検討するために、がん患者の集団の死亡確率が同じ性別、年齢の一般集団の死亡確率とほぼ同じになったとき、がんが治癒したとみなす考え方(相対生存率の考え方)がとり入れられ、死因を特定することなく適用できる、より一般的な治癒確率の推定方法が提案された(Verdecchia et al., 1998; Francisci et al., 2009)。そして、治癒確率を推定する治癒モデルは診断時の年齢や進行度などの共変量の影響を考慮できる回帰モデルへと発展した(De Angelis et al., 1999; Lambert et al., 2007b)。地域がん登録資料を用いた治癒モデルに基づくがん患者の治癒確率の推定の報告は、欧米諸国ではいくつか先行研究がみられるが(Verdecchia et al., 1998; De Angelis et al., 1999; Lambert et al., 2007a; Huang et al., 2008; Shah et al., 2008; Francisci et al., 2009; Woods et al., 2009)、我が国からの研究報告はまだ殆どないといえる。がん患者や医療従事者が真に知りたい情報である治癒確率を、我が国の地域がん登録資料を用いて推定し、治癒モデルを用いて、その傾向を検討し、がん医療評価を行うことが可能になりつつある。本報告では、大阪府がん登録資料より胃がん患者の治癒確率の推移を例に、地域がん登録資料に基づく治癒モデルの方法と適用について紹介する。

## 2. 方法

### 2.1 相対生存率

地域がん登録資料から生存率を推定する場合、「相対生存率」の推定値を報告することが標準となっている。ある時間  $t$  に対して、 $S(t)$  をすべての死因による死亡を対象とする全生存確率(all-cause survival)、 $S^*(t)$  を期待生存確率(expected survival)、相対生存確率を  $S^R(t)$  とすると、これらの関係は

$$(2.1) \quad S(t) = S^*(t)S^R(t)$$

のように表わされる。すなわち、相対生存率  $S^R(t)$  とは、全生存率と期待生存率の比  $S(t)/S^*(t)$  であり、当該疾患による過剰な死亡についての表現できる指標である。期待生存率  $S^*(t)$  は、一般集団の生存確率を意味し、患者の個人のデータからではなく、通常、公的機関などで整備

された生命表より同一の時代・性・年齢の生存確率を用いて推定することができる。本論文では国立がんセンターにより公表されているコホート生存率表(National Cancer Center, 2009)から、Ederer II法を用いて推定している(Ederer et al., 1961)。期待生存率  $S^*(t)$  は、計測対象として関心のあるがん患者集団からみれば、がん患者集団の全生存率  $S(t)$  の中で、一般集団に曝されているリスクと同じリスクによる死亡(例えば、がんに関係しない死亡)に基づく生存確率に該当するものである。したがって、相対生存率  $S^R(t)$  は、正確な死因情報が得られない状況下においても、他死因による競合の影響が大きくなる長期生存率の推定も可能であり、がんが直接的な原因となった場合だけでなく、間接的な原因となった死亡(例えば、がんの治療に起因した他死因死亡など)についても考慮に入れることができる。これらのことから、相対生存率  $S^R(t)$  は、地域がん登録資料を用いたがん患者の治癒確率の推定において、現在、欠かすことのできない考え方になっている。

$\lambda(t)$  を  $S(t)$  に対応する全ハザード関数、 $\lambda(t) = -d \log S(t) / dt$ 、 $\lambda^*(t)$  を  $S^*(t)$  に対応する期待ハザード関数とする。三つの生存率の関係(2.1)をハザード関数で表現すると

$$(2.2) \quad \lambda(t) = \lambda^*(t) + \lambda^R(t)$$

である。ここに、 $\lambda^R(t)$  は  $S^R(t)$  に対応する相対ハザード関数であり、当該疾患(ここではがん)に関連する過剰ハザード(excess mortality rate)と呼ばれる。 $\lambda^*(t)$  は、外部の資料(生命表)から得られる  $S^*(t)$  の推定値に基づいて推定できる。すなわち、 $\hat{S}^*(t)$  を  $S^*(t)$  の推定値、 $\hat{\lambda}^*(t)$  を  $\lambda^*(t)$  の推定値とすると、ハザードと生存率の関係  $\hat{\lambda}^*(t) \Delta t = 1 - \hat{S}^*(t) / \hat{S}^*(t_-)$  から求めることができる。ここに、 $\Delta t = t - t_-$ 、 $t_-$  は  $t$  のある直前の時間である。ただし、通常、 $\hat{S}^*(t)$  は大規模な集団から推定されるので、 $\hat{S}^*(t) \doteq S^*(t)$  と見做せる。

図1に、相対生存曲線  $S^R(\cdot)$  と治癒確率  $\pi$  とのイメージ図を与える。図1に示すように、 $S^R(\cdot)$  が平らになるとき「治癒した」と考えることができる。つまり、過剰ハザード  $\lambda^R(t)$  がゼロに、すなわち、がん患者の全ハザードが一般集団の全ハザードと同じになるとき  $\lambda(t) = \lambda^*(t)$  を、「完全に治癒した」時点と考えることができる。ただし、この曲線の傾きが数学的に完全に0となる時点は、通常の連続生存時間モデルの上では、 $+\infty$  であり、意味をなさない。そこで、

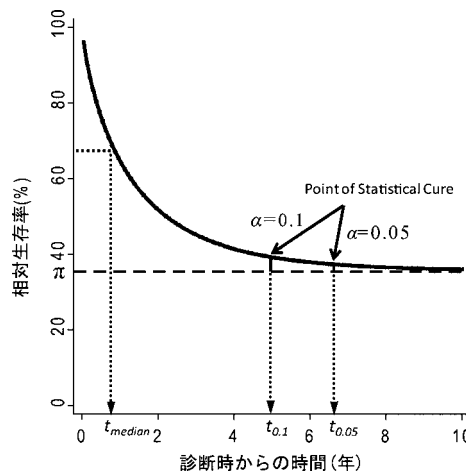


図1. 相対生存曲線  $S^R(\cdot)$  による治癒確率  $\pi$  と非治癒生存分布の中央生存時間  $t_{\text{median}}$  のイメージ図。

非治癒患者のうち  $100(1 - \alpha)\%$  が死亡した地点を「統計的に治癒した時点 (point of statistical cure)」と呼び、推測の対象にすることもできる。

## 2.2 治癒混合モデル

治癒モデルは、大きく二つに分けると、治癒混合モデル (mixture cure fraction model) と治癒非混合モデル (non-mixture cure fraction model) がある。とくに、本論文では、治癒確率の推測の目的において、より簡潔な解釈を与える前者の治癒混合モデルをとり上げることにする。因に、後者の治癒非混合モデルでは、治療直後の治癒を必ずしも仮定することなく (例えば、治療直後には Poisson 分布に従うような個数のがん細胞が残留し、転移の可能性があるという状況を想定する)、治療後の時間経過に伴って生じる治癒を推定することを目的とする (Tsodikov et al., 2003)。そのため、治癒非混合モデルは、地域がん登録データによる治癒確率の推測に対して、モデルのあてはまりという観点で適する場合もあり有用であるが、生物学的には一般に適切でないといわれる (Lambert et al., 2007b)。

地域がん登録データでは、がんによる死亡とその他の死因とを区別するのが困難であるために、De Angelis et al. (1999) は、背景死亡 (background mortality) を考慮し、相対生存率の考え方を取り込んだ拡張版の治癒モデルを導入した。すなわち、がんによる相対生存確率の中に

$$(2.3) \quad S^R(t) = \pi + (1 - \pi)S^C(t)$$

となるような治癒混合モデルを想定する。ここに、 $\pi$  と  $S^C(t)$  は、それぞれ、該当する集団の治癒確率と非治癒患者の生存関数である。したがって、全生存確率  $S(t)$  は、(2.1) と (2.3) により

$$S(t) = S^*(t)\{\pi + (1 - \pi)S^C(t)\}$$

である。相対ハザード関数  $\lambda^R(t)$  は、治癒混合モデル (2.3) のもとで

$$(2.4) \quad \lambda^R(t) = \frac{f^R(t)}{S^R(t)} = \frac{(1 - \pi)f^C(t)}{\pi + (1 - \pi)S^C(t)}$$

となる。ここに、 $f^R(t)$  と  $f^C(t)$  は、それぞれ、 $S^R(t)$  と  $S^C(t)$  に対応する確率密度関数を表す。

## 2.3 治癒混合モデルの推測

### 2.3.1 尤度関数

地域がん登録データから得られる生存情報を、 $D = \{(d_i, t_i, z_i), i = 1, \dots, n\}$  とする。ここに、 $d_i$  は、 $d_i = 1$  であれば死亡、 $d_i = 0$  であれば観察の中途打ち切りを表す個人  $i$  の中途打ち切り指標であり、 $t_i$  は個人  $i$  の生存時間観測値、 $z_i$  は個人  $i$  の共変量である。臨床試験から収集される詳細な個人共変量とは異なり、地域がん登録データで観察される共変量  $z_i$  には、通常、診断時期、性、年齢、進行度が記録される。

$\lambda_i^*(t)$ 、 $S_i^*(t)$ 、 $\lambda_i^R(t)$ 、 $S_i^R(t)$  を、それぞれ、個人  $i$  の期待ハザード関数、期待生存関数、相対ハザード関数、相対生存関数とすると、(2.1) と (2.2) と同様に

$$S_i(t) = S_i^*(t)S_i^R(t) \quad \text{と} \quad \lambda_i(t) = \lambda_i^*(t) + \lambda_i^R(t)$$

である。また、独立中途打ち切りが仮定できるとする。このとき、生存情報  $D$  から得られる対数尤度は

$$(2.5) \quad l(\theta) = \log \prod_{i=1}^n \{\lambda_i^*(t_i) + \lambda_i^R(t_i)\}^{d_i} S_i^*(t_i) S_i^R(t_i)$$

である。ここに、 $\theta$  は、 $\lambda_i^R(t)$  と  $S_i^R(t)$  の中に含まれるモデル・パラメータを表す。治癒混合モデルの適用において、 $\lambda_i^R(t)$  と  $S_i^R(t)$  には、(2.3) と (2.4) の個人単位としての関係

$$S_i^R(t) = \pi_i + (1 - \pi_i)S_i^C(t) \quad \text{と} \quad \lambda_i^R(t) = \frac{(1 - \pi_i)f_i^C(t)}{\pi_i + (1 - \pi_i)S_i^C(t)}$$

が成り立つ。ここに、 $\pi_i$  は個人  $i$  の治癒確率、 $S_i^C(t)$  と  $f_i^C(t)$  は個人  $i$  が非治癒であるときの生存関数と確率密度関数、すなわち、個人  $i$  の  $S^C(t)$  と  $f^C(t)$  を表す。したがって、(2.5)において、モデル・パラメータ  $(\theta)$  に影響しない  $\log \prod_{i=1}^n S_i^*(t_i)$  の寄与は対数尤度から外すことができる。2.1 節で述べたように、 $S_i^*(t_i)$  と  $\lambda_i^*(t_i)$  は、個人  $i$  の性・年齢・暦年が一致する生命表の生存確率に基づいて求める(推定する)ことができる。

### 2.3.2 パラメータ推定

治癒混合モデルのパラメータ推定は、他の統計モデルと同様に、対数尤度  $l(\theta)$  を  $\theta$  の上で最大化することに基づく。パラメータの推定標準誤差は、最尤推定値で評価される Fisher 情報行列の逆行列の平方根によって得られる。Lambert et al. (2007b) は、対数尤度  $l(\theta)$  の最大化のために、数値微分を用いる Newton-Raphson 法を考え、Stata に導入されている (Lambert, 2007)。ただし、そのような機械的な推定方法では、このモデルに内在する推定量の性質やモデルの拡張を調査するための直観は働かない。そこで、本節では、EM アルゴリズム (Dempster et al., 1977) を用いた推定方法を、研究ノートとして書き留める。

個人  $i$  に対して、 $E_i = 1$  であれば、がんに関連した死亡を表し、 $E_i = 0$  であればそれ以外の死亡であるとする。時間  $t_i$  で死亡観測された ( $d_i = 1$ ) もとで、個人  $i$  の死因ががんに関連した死亡である確率は

$$v_i = v_i(\theta) = \Pr(E_i = 1 | t_i \text{ で } d_i = 1 \text{ を観測}) = \frac{\lambda_i^R(t_i)}{\lambda_i^*(t_i) + \lambda_i^R(t_i)}$$

である (Perme et al., 2009 を参照)。この  $v_i(\theta)$  を与えたもとで、対数尤度  $l(\theta)$  に対する(治癒モデルを導入する前の)EM 対数尤度は

$$l_1^{EM}(\theta|v) = \sum_{i=1}^n [d_i v_i \log \lambda_i^R(t_i) + \log S_i^R(t_i) + d_i v_i \log \lambda_i^*(t_i) + \log S_i^*(t_i)]$$

となる。次に、相対生存率  $S_i^R(t)$  に対して、治癒混合モデル(2.3)を導入し、 $l_1^{EM}(\theta|v)$  の最大化を考えると、再び EM アルゴリズムが適用できることがわかる。すなわち、通常の治癒混合モデルの推定(杉本・後藤, 1999; Peng and Dear, 2000; Sy and Taylor, 2000)と同様に、時間  $t_i$  で中途打ち切りが観測された ( $d_i = 0$ ) もとで、個人  $i$  が治癒していない確率は

$$w_i = w_i(\theta) = \frac{(1 - \pi_i) S_i^C(t_i)}{\pi_i + (1 - \pi_i) S_i^C(t_i)}$$

である(例えば、Sugimoto et al., 2005 を参照)。したがって、この  $w_i(\theta)$  を与えたもとで、対数尤度  $l_1^{EM}(\theta|v)$  に対応する EM 対数尤度は

$$l_2^{EM}(\theta|w, v) = \sum_{i=1}^n \{ [d_i v_i + (1 - d_i v_i) w_i] \log(1 - \pi_i) + (1 - d_i v_i) w_i' \log \pi_i \\ + d_i v_i \log f_i^C(t_i) + (1 - d_i v_i) w_i \log S^C(t_i) + d_i v_i \log \lambda_i^*(t_i) + \log S_i^*(t_i) \}$$

となる。ここに、 $w_i' = 1 - w_i$  である。

上記のような EM 対数尤度を用いると、2 重の EM アルゴリズムによりパラメータ推定が実行されることが分かる。簡単のため、 $v_i^{(l)} = v_i(\theta^{(l)})$ 、 $w_i^{(m)} = w_i(\theta^{(m)})$ 、 $G^{(m,l)}(\theta) = l_2^{EM}(\theta|w^{(m)}, v^{(l)})$ 、 $\epsilon$  を十分小さい値として、このアルゴリズムを以下に要約する：

1. 初期パラメータ  $\theta^{(0)}$  を設定し、 $l = 0$ 、 $m = 0$  とする。
2.  $G^{(m,l)}(\theta)$  を最大化し、 $\theta^{(m+1)} = \operatorname{argmax}_{\theta} G^{(m,l)}(\theta)$  とする。  $|G^{(m,l)}(\theta^{(m+1)}) - G^{(m,l)}(\theta^{(m)})| < \epsilon$  であれば、 $\theta^{(l+1)} = \theta^{(m+1)}$  として 3 に進む。そうでなければ、 $m \leftarrow m + 1$  として再び 2 を実行する。

3.  $|\theta^{(l+1)} - \theta^{(l)}| < \epsilon$  であれば終了する. そうでなければ,  $l \leftarrow l+1$  として 2 に戻る.

### 2.3.3 仮定する分布とリンク関数

$S^C(t)$  のモデル化のために, Weibull 分布か対数正規分布を用いることが多い. 本論文では, 非治癒生存関数  $S_i^C(t)$  に一貫して, Weibull 分布

$$(2.6) \quad S_i^C(t) = \exp\{-(\lambda_i t)^{\gamma_i}\}$$

を仮定する. 個人  $i$  の治癒確率  $\pi_i$  と Weibull パラメータ  $(\lambda_i, \gamma_i)$  は, 共変量によって変化するようにモデル化する. よく知られているモデル化として, 恒等, 対数, ロジスティック,  $\log(-\log)$  リンクがある. 本論文では, 治癒確率  $\pi_i$  には, ロジスティック・リンク

$$\log(\pi_i/(1-\pi_i)) = \beta_0 + \beta^T z_i$$

を用い, Weibull パラメータ  $\lambda_i$  と  $\gamma_i$  には対数リンクを用いる.

### 2.3.4 治癒モデルに基づくトレンドの評価

ここでは, 大阪府がん登録資料に基づいて診断時期の変化をみるために, 治癒確率と非治癒生存分布の中央生存時間に関する指標の評価方法を記述する.

Weibull パラメータの最尤推定値  $(\hat{\lambda}_i, \hat{\gamma}_i)$  を用いて, 非治癒生存関数  $S^C(t)$  から, 共変量  $z_i$  をもつ非治癒患者群の中央生存時間

$$(2.7) \quad t_{\text{median}} = \{(-\log(0.5))/\lambda_i\}^{1/\gamma_i}$$

を推定することができる(図 1 を参照). この中央値の推定値を非治癒群の要約指標として用いる.

治癒確率の推移について, 診断時期の効果の大きさを多変量モデルにより評価する. 各モデルの共変量  $z_i$  には, 以下のように変数を組み込む:

- モデル 1: 診断時期のみ
- モデル 2: 診断時期+年齢
- モデル 3: 診断時期+進行度
- モデル 4: 診断時期+年齢+進行度

このとき, 診断時期は 1975–80 年, 1981–85 年, 1986–90 年, 1991–95 年, 1996–2000 年の 5 期間, 年齢は 15–39 歳, 40–59 歳, 60–74 歳, 75–99 歳の 4 区分, 進行度は診断時のものとし, 限局(腫瘍が当該臓器に限局している), 領域浸潤(リンパ節転移あるいは腫瘍が隣接臓器まで浸潤している), 遠隔転移の 3 区分を, それぞれ, ダミー変数化して用いる. すなわち, モデル 1 であれば,  $z_i = (\text{診断時期}_2, \dots, \text{診断時期}_5)$  を用いることを意味する. 年齢や進行度などの他の共変量にどの程度影響を受けているかをモデル  $k$  における診断時期  $j$  の推定回帰係数  $\hat{\beta}_j^{(k)}$  とデビアンスを比較することにより検討する. デビアンスは, 対数尤度として  $\log \prod_{i=1}^n S_i^*(t_i)$  の寄与を落としたもの(つまり,  $l(\theta) - \log \prod_{i=1}^n S_i^*(t_i)$ ) を用いた場合の  $-2 \times$  最大対数尤度である. これに基づいて, 説明割合の指標として,  $R^2 = 1 - \text{デビアンス}/\text{帰無デビアンス}$  を算出した.

なお, 進行度不明例が 10–20% 存在したため, multiple imputation を適用して欠測値処理を行った. 補完モデルにおいては, 生存時間(0–1 年, 1–5 年, 5–10 年)と各項目との交互作用項を加え, 予後情報を反映できるモデルとした(Nur et al., 2010).

すべての解析は, Stata MP ver.11.1 により実行した(StataCorp, 2009). また, 2.3.2 節で述べられたパラメータ推定法は, Fortran 言語で書かれ, Stata の数値微分による Newton-Raphson 法の結果と同じ結果を生むことが確認された. 治癒混合モデルは Lambert et al. (2007b) により

提供されたコマンド strsmix (Lambert, 2007) を使用し, multiple imputation は Royston (2006) により提供されているコマンド ice により実行した.

### 3. 結果(大阪府がん登録資料による胃がん患者の解析例)

大阪府がん登録資料より, 1975-2000 年に診断された胃がん患者 (ICD 10 code : C16) 66,032 名を対象とし, 治癒混合モデルを適用し, 治癒確率および非治癒生存分布の中央生存時間を診断時期別に推定し, その推移を評価した (伊藤 他, 2009). 胃がんの生存率には性差があることが知られているため (Sato et al., 2009), 男女別に解析を行った (表 1.1-1.2). 表 1.1 は男性の結果, 表 1.2 は女性の結果である. 全患者についての男女別の治癒確率を横軸に非治癒生存分布の中央生存時間を縦軸にとり, 1975-2000 年を 5 期間に分けた推移を図 2 に示した. 男女別に五つの点は左から 1975-80 年, 1981-85 年, 1986-90 年, 1991-95 年, 1996-2000 年診断患者の値である.

#### 3.1 各診断時期の治癒確率および非治癒患者の生存時間の推移

胃がんの治癒確率は 1975-80 年には男性で 34%, 女性で 27.2% であると推定されたが, 1996-2000 年にはそれぞれ 55.9%, 49.7% と約 20 ポイント以上向上した. 一方, 非治癒生存分布の中央生存時間は男性では 7-8 カ月とあまり変化がなく, 女性では 1986-90 年診断患者では 10 カ月まで延長したものの, その後 7 カ月に減少した. 年齢階級別, 進行度別の治癒確率および非治癒生存分布の中央生存時間の値の結果も表 1.1-1.2 に示した. 年齢階級別では, 15-39 歳と 40-59 歳ではほぼ同様の傾向を示し, 1996-2000 年診断患者においては, 男性では約 60%, 女性では 50-55% 前後の患者が治癒したと推定された. 高齢になるほど, 治癒確率は低下した. 男女差は 60 歳未満で大きく, 概して 10% 程度の差が生じている. 非治癒生存分布の中央生存時間については, あまり大きな変化はなかった.

診断時の進行度により, 治癒確率および非治癒生存分布の中央生存時間は大きく異なった. 1996-2000 年診断患者においては, 腫瘍が胃に局限している局限患者では, 90% 以上が治癒していることになり, これは 1975-80 年診断患者の治癒確率と比べて, 約 15-20 ポイント以上向上

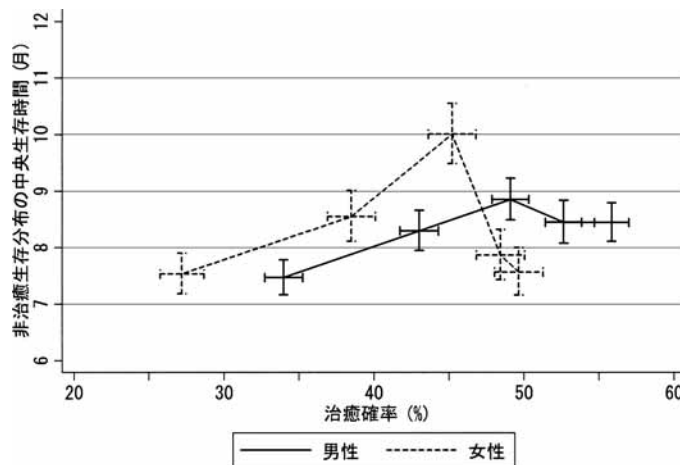


図 2. 大阪府における胃がん患者の治癒確率 (%) と非治癒生存分布の中央生存時間 (月) の推移: 1975-2000 年診断患者, 全進行度.

表 1.1. 大阪府における胃がん患者の5年相対生存率, 治癒率および非治癒患者の中央生存時間の推移: 1975-2000年診断患者, 男性.

	診断時期	5年相対生存率 (%)		治癒率 (%)		非治癒生存分布の中央生存時間 (月)		
		推定値 [95% CI]	推定値 [95% CI]	推定値 [95% CI]	推定値 [95% CI]	推定値 [95% CI]	推定値 [95% CI]	
全患者	1975-1980	29.8 [28.7-30.9]	34.0 [32.7-35.3]	7.5 [7.2-7.8]				
	1981-1985	37.4 [36.3-38.5]	43.0 [41.8-44.3]	8.3 [7.9-8.7]				
	1986-1990	43.0 [42.0-44.0]	49.1 [47.9-50.3]	8.9 [8.5-9.2]				
	1991-1995	45.7 [44.7-46.7]	52.6 [51.4-53.9]	8.4 [8.1-8.8]				
	1996-2000	47.4 [46.4-48.4]	55.9 [54.7-57.0]	8.4 [8.1-8.8]				
年齢別	15-39 歳	1975-1980	44.2 [39.5-48.9]	42.2 [37.4-47.1]	10.7 [9.1-12.6]			
		1981-1985	52.4 [47.5-57.2]	50.4 [45.4-55.3]	10.2 [8.6-12.2]			
		1986-1990	56.4 [50.8-61.9]	55.0 [49.3-60.6]	10.8 [8.7-13.4]			
		1991-1995	60.3 [53.4-67.2]	59.5 [52.3-66.4]	8.9 [6.5-12.2]			
		1996-2000	61.9 [53.9-69.9]	61.9 [53.6-69.6]	7.7 [5.6-10.6]			
	40-59 歳	1975-1980	42.3 [40.4-44.3]	41.3 [39.1-43.5]	10.1 [9.4-10.9]			
		1981-1985	50.4 [48.5-52.2]	50.9 [48.9-52.9]	10.2 [9.5-10.9]			
		1986-1990	55.4 [53.7-57.0]	56.1 [54.3-57.9]	10.6 [9.9-11.3]			
		1991-1995	59.1 [57.4-60.8]	60.0 [58.1-61.8]	9.7 [9.0-10.4]			
		1996-2000	60.1 [58.2-61.9]	62.4 [60.5-64.3]	10.3 [9.6-11.0]			
	60-74 歳	1975-1980	24.8 [23.3-26.3]	30.9 [29.1-32.9]	6.9 [6.5-7.3]			
		1981-1985	34.0 [32.3-35.6]	40.4 [38.3-42.4]	8.4 [7.9-9.0]			
		1986-1990	40.8 [39.1-42.4]	47.1 [45.2-49.0]	9.0 [8.4-9.5]			
		1991-1995	44.2 [42.6-45.7]	50.3 [48.4-52.1]	9.0 [8.5-9.6]			
		1996-2000	48.2 [46.7-49.6]	55.8 [54.2-57.5]	8.8 [8.3-9.3]			
75-99 歳	1975-1980	8.8 [7.0-10.5]	19.1 [16.0-22.7]	4.5 [4.0-4.9]				
	1981-1985	12.5 [10.7-14.3]	24.9 [21.7-28.4]	5.3 [4.8-5.9]				
	1986-1990	16.4 [14.5-18.4]	31.3 [28.1-34.7]	6.1 [5.5-6.7]				
	1991-1995	21.7 [19.8-23.7]	39.2 [35.9-42.7]	5.9 [5.3-6.6]				
	1996-2000	24.8 [22.8-26.8]	42.8 [39.7-45.9]	5.7 [5.1-6.3]				
進行度別	限局	1975-1980	66.0 [63.9-68.1]	74.5 [71.0-77.5]	9.2 [7.0-12.2]			
		1981-1985	75.5 [73.9-77.2]	83.7 [80.1-87.1]	23.7 [14.8-37.8]			
		1986-1990	77.4 [76.0-78.8]	84.2 [80.4-87.2]	29.9 [19.3-46.3]			
		1991-1995	78.9 [77.7-80.2]	89.2 [86.9-90.8]	20.6 [14.7-28.8]			
		1996-2000	79.7 [78.5-80.9]	92.6 [91.6-94.0]	12.7 [10.4-15.6]			
	領域浸潤	1975-1980	22.0 [20.6-23.3]	23.4 [21.8-25.3]	9.5 [9.0-10.0]			
		1981-1985	25.6 [24.1-27.0]	27.3 [25.7-29.2]	10.7 [10.1-11.4]			
		1986-1990	27.6 [26.0-29.1]	30.0 [28.2-31.9]	12.1 [11.5-12.7]			
		1991-1995	26.5 [24.9-28.0]	27.6 [26.4-30.2]	12.0 [11.4-12.7]			
		1996-2000	27.8 [26.2-29.4]	33.0 [30.9-34.7]	12.1 [11.5-12.7]			
	遠隔転移	1975-1980	1.6 [1.0-2.2]	2.5 [1.6-3.7]	5.1 [4.8-5.4]			
		1981-1985	1.6 [1.0-2.2]	2.5 [1.6-3.3]	5.4 [5.1-5.7]			
		1986-1990	1.7 [1.1-2.2]	2.7 [1.8-3.6]	5.3 [5.0-5.6]			
		1991-1995	1.2 [0.7-1.7]	1.9 [1.2-2.8]	4.9 [4.7-5.2]			
		1996-2000	1.6 [1.0-2.1]	2.0 [1.5-2.9]	5.3 [5.1-5.6]			



表 1.2. 大阪府における胃がん患者の5年相対生存率、治癒率および非治癒患者の中央生存時間の推移：1975-2000年診断患者，女性.

		診断時期	5年相対生存率 (%)		治癒率 (%)		非治癒生存分布の中央生存時間 (月)	
			推定値 [95% CI]	推定値 [95% CI]	推定値 [95% CI]	推定値 [95% CI]		
全患者		1975-1980	25.7 [24.4-27.0]	27.2 [25.8-28.7]	7.5 [7.2-7.9]			
		1981-1985	36.2 [34.7-37.6]	38.5 [36.9-40.1]	8.5 [8.1-9.0]			
		1986-1990	42.6 [41.1-44.0]	45.2 [43.6-46.8]	10.0 [9.5-10.6]			
		1991-1995	44.4 [42.9-45.8]	48.4 [46.8-50.0]	7.9 [7.4-8.3]			
		1996-2000	44.3 [42.8-45.8]	49.7 [48.0-51.3]	7.6 [7.2-8.0]			
年齢別	15-39 歳	1975-1980	30.6 [26.7-34.5]	28.1 [24.3-32.2]	9.4 [8.2-10.7]			
		1981-1985	43.7 [39.1-48.4]	41.3 [36.8-46.1]	10.8 [9.3-12.7]			
		1986-1990	48.0 [42.8-53.2]	45.3 [40.1-50.6]	12.5 [10.5-14.8]			
		1991-1995	51.6 [44.6-58.6]	50.1 [43.0-57.2]	9.1 [7.1-11.6]			
		1996-2000	46.5 [38.6-54.3]	46.5 [38.7-54.4]	10.5 [8.4-13.1]			
	40-59 歳	1975-1980	34.5 [32.0-37.1]	33.0 [30.4-35.8]	9.3 [8.6-10.2]			
		1981-1985	46.1 [43.5-48.6]	44.6 [42.0-47.3]	11.0 [10.1-12.0]			
		1986-1990	52.8 [50.4-55.2]	51.3 [48.8-53.8]	11.8 [10.8-12.9]			
		1991-1995	57.1 [54.6-59.5]	56.5 [54.0-59.1]	10.3 [9.3-11.4]			
		1996-2000	54.8 [52.0-57.5]	55.6 [52.8-58.4]	9.0 [8.2-9.9]			
	60-74 歳	1975-1980	24.7 [22.6-26.7]	27.6 [25.2-30.0]	7.7 [7.1-8.2]			
		1981-1985	36.5 [34.1-38.8]	40.1 [37.5-42.8]	8.4 [7.7-9.1]			
		1986-1990	44.1 [41.6-46.5]	46.5 [43.8-49.3]	10.3 [9.4-11.3]			
		1991-1995	46.9 [44.5-49.3]	49.6 [46.9-52.3]	8.5 [7.7-9.3]			
		1996-2000	50.4 [48.0-52.9]	53.7 [51.1-56.3]	8.4 [7.7-9.2]			
75-99 歳	1975-1980	7.1 [5.2-9.0]	13.1 [10.4-16.5]	4.1 [3.7-4.6]				
	1981-1985	13.8 [11.4-16.1]	22.8 [19.4-26.5]	4.9 [4.4-5.5]				
	1986-1990	22.0 [19.5-24.5]	33.4 [29.9-37.1]	6.5 [5.8-7.2]				
	1991-1995	22.9 [20.5-25.3]	34.4 [31.2-37.7]	5.2 [4.7-5.7]				
	1996-2000	25.5 [23.1-27.9]	37.6 [34.3-41.0]	5.5 [4.9-6.1]				
進行度別	限局	1975-1980	62.7 [59.7-65.6]	67.6 [62.7-70.5]	10.2 [8.0-13.0]			
		1981-1985	76.5 [74.2-78.8]	83.4 [80.1-86.2]	15.1 [10.6-21.6]			
		1986-1990	79.9 [78.0-81.8]	85.1 [82.5-88.1]	23.2 [16.6-32.5]			
		1991-1995	82.1 [80.4-83.8]	88.0 [85.8-90.5]	17.0 [11.6-24.9]			
		1996-2000	81.5 [79.7-83.3]	90.9 [89.2-92.8]	11.6 [8.9-15.2]			
	領域浸潤	1975-1980	19.5 [17.9-21.2]	18.7 [17.2-21.1]	9.2 [8.6-9.9]			
		1981-1985	28.2 [26.2-30.1]	28.0 [25.9-30.3]	11.5 [10.7-12.3]			
		1986-1990	29.9 [27.8-32.0]	29.1 [26.9-31.7]	13.8 [12.8-14.7]			
		1991-1995	26.7 [24.6-28.8]	27.6 [25.0-29.9]	11.6 [10.8-12.5]			
		1996-2000	26.2 [23.9-28.4]	29.1 [26.7-31.9]	10.9 [10.2-11.8]			
	遠隔転移	1975-1980	1.4 [0.7-2.0]	1.9 [1.0-2.9]	5.2 [4.8-5.6]			
		1981-1985	1.8 [1.0-2.6]	2.0 [1.3-3.5]	5.4 [5.0-5.8]			
		1986-1990	2.4 [1.5-3.3]	2.9 [2.1-4.5]	5.9 [5.5-6.3]			
		1991-1995	0.8 [0.3-1.3]	1.2 [0.6-2.3]	4.6 [4.3-5.0]			
		1996-2000	1.2 [0.5-1.8]	1.7 [0.9-2.8]	4.8 [4.5-5.2]			

表 2. 大阪府における胃がん患者の治癒率の時代変化について各種モデルによる診断時期の回帰係数と年齢および進行度の影響度.

		モデル 1:	モデル 2:	モデル 3:	モデル 4:
		診断時期	診断時期 + 年齢	診断時期 + 進行度	診断時期 + 年齢 + 進行度
診断時期		$\hat{\beta}_j^{(1)}$ [95% CI]	$\hat{\beta}_j^{(2)}$ [95% CI]	$\hat{\beta}_j^{(3)}$ [95% CI]	$\hat{\beta}_j^{(4)}$ [95% CI]
男性	1975-1980	0.00 (ref)	0.00 (ref)	0.00 (ref)	0.00 (ref)
	1981-1985	0.38 [0.31-0.46]	0.39 [0.32-0.47]	0.39 [0.28-0.51]	0.40 [0.29-0.52]
	1986-1990	0.63 [0.55-0.70]	0.64 [0.56-0.71]	0.51 [0.40-0.63]	0.54 [0.42-0.65]
	1991-1995	0.77 [0.70-0.84]	0.81 [0.74-0.89]	0.52 [0.40-0.64]	0.59 [0.47-0.70]
	1996-2000	0.90 [0.83-0.97]	0.99 [0.91-1.06]	0.72 [0.61-0.83]	0.83 [0.71-0.94]
	デビアンズ ( $R^2$ )	117823 (0.007)	116088 (0.022)	94703 (0.202)	93551 (0.212)
女性	1975-1980	0.00 (ref)	0.00 (ref)	0.00 (ref)	0.00 (ref)
	1981-1985	0.52 [0.42-0.62]	0.54 [0.44-0.64]	0.69 [0.54-0.83]	0.69 [0.55-0.84]
	1986-1990	0.79 [0.70-0.89]	0.84 [0.74-0.94]	0.81 [0.66-0.95]	0.85 [0.71-1.00]
	1991-1995	0.92 [0.82-1.02]	0.99 [0.89-1.09]	0.78 [0.64-0.93]	0.88 [0.73-1.02]
	1996-2000	0.97 [0.87-1.07]	1.08 [0.98-1.18]	0.89 [0.74-1.04]	1.03 [0.88-1.17]
	デビアンズ ( $R^2$ )	61205 (0.010)	60018 (0.03)	48999 (0.208)	47966 (0.225)

している。腫瘍がリンパ節に転移していたり、隣接臓器に浸潤していたりする領域浸潤の患者では、約 30% 以上が治癒していたと推定された。これは、1975-80 年診断患者と比べると約 10 ポイントの向上であった。一方、遠隔転移の患者においては、治癒確率は非常に低く、1-2% 程度と推定された。いずれの進行度においても、非治癒生存分布の中央生存時間は、あまり大きく延長しなかった。

### 3.2 年齢、進行度分布により説明可能な量

全患者における治癒確率は大きく向上していたため、それが、年齢や進行度分布の変化によってどの程度影響していたかを  $R^2$  指標を用いて検討した(これらの結果は表 2 を参照)。治癒確率の向上が年齢分布の変化により説明できる割合は約 1-2% 程度であった。一方、進行度分布については、男女とも約 20% と高く見積もられたことになる。つまり、治癒確率の向上の約 20% が進行度の変化により説明され得ることを表す。

## 4. 考察

全対象者の性別ごとの治癒確率の向上は 20 ポイント以上と著しかった。早期発見の増加に伴う進行度分布の変化は、男女ともに約 20% の寄与率で全体の治癒確率の向上に影響を与えていた。他方、進行度別の結果においても、遠隔転移例を除き、治癒確率が大きく向上したため、治療技術の向上や術後の疾病管理の向上による影響も大きかったと考えられる。治癒確率が大きく向上する一方、非治癒患者の生存時間はあまり延長しなかったといえる。Verdecchia et al. (1998) の結果の解釈の分類によれば、早期診断によるリードタイム・バイアスが生じている場合や化学療法による末期患者への延命治療などの場合、今回の結果とは逆に治癒確率が向上せず、非治癒患者の生存時間のみが延長する。また、治癒確率のみが向上し、非治癒患者の生存時間は延長しないというパターンとしては、死亡のリスクが一般集団と同レベルの患者が患者集団に加入することにより起こるパターンといわれ、過剰診断 (over-diagnosis) の傾向と解釈されることがある (Verdecchia et al., 1998)。しかし、我が国における胃がん検診の受診率が 20% 未満

であるという現状を考えれば、その傾向が該当するとは思えず、また早期胃がんは未治療であれば進行するという先行研究からも否定される。したがって、近年の胃がん治療によって完全に治癒する患者の割合が増加したと考えられる。これには早期診断により、治癒切除可能な症例が増えたことにも起因する。大阪府がん登録資料によれば、胃がん患者における手術患者における治癒切除割合は、1970年代後半では50%程度であったが（大阪府健康福祉部 他, 1987）、2000年では約80%であった（大阪府健康福祉部 他, 2004）。

今回、胃がんの結果においては、治癒確率（ $\pi_i$ ）のみが大きく向上したため、診断時期の効果について、年齢および進行度の影響度の推定は $\pi_i$ についてのみ行い、非治癒患者の生存時間についての検討は行っていない。非治癒生存分布の中央生存時間は、Weibullモデルの場合、二つのパラメータ $\lambda_i$ と $\gamma_i$ により推定されるため、今回の接近法よりも複雑になる。

また、先行研究では、がんの治癒確率を推定する場合に、非治癒生存分布としてパラメトリックなWeibullモデルを用い、リンク関数としてロジスティック関数を適用することが多いため、ここでは、それに準じたが、今後、部位別の治癒確率を検討する際には、他の分布族やリンク関数により、モデルのあてはまりがどのように変化するかを検討することも必要である。非治癒生存分布に、特定の生存時間モデルを仮定しないセミパラメトリック治癒モデルへの方法論の拡張は、2.3.2節の2重EMアルゴリズムと、通常の治癒混合モデルのセミパラメトリック接近法（杉本・後藤, 1999; Peng and Dear, 2000; Sy and Taylor, 2000）に類似する流儀により、容易に展開できることが予想された。

## 5. 結論

地域がん登録資料に基づく治癒混合モデルを用いた治癒患者割合の推定は、がん医療の評価を行う上で重要であり、また医療従事者および患者、その家族にとって、貴重な情報提供となり得る。地域がん登録資料による統計的推測の方法が発展し、比較的容易に治癒率の推定を行うことが可能になってきた。方法論のさらなる精緻化とともに、今後、様々な部位のがんについても適用し、がん医療の評価を行うことが大切である。

## 謝 辞

本研究の遂行において、文部科学省科学研究費・研究活動スタート支援（課題番号：22800094）から研究助成を受けた。本稿改訂にあたり、査読者から有益なコメントをいただいた。ここに感謝の意を表します。

## 参 考 文 献

- Boag, J. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy (with discussion), *Journal of the Royal Statistical Society, Series B*, **11**, 15–53.
- Cox, D. R. (1972). Regression models and life tables (with discussion), *Journal of the Royal Statistical Society, Series B*, **34**, 187–220.
- De Angelis, R., Capocaccia, R., Hakulinen, T., Soderman, B. and Verdecchia, A. (1999). Mixture models for cancer survival analysis: Application to population-based data with covariates, *Statistics in Medicine*, **18**, 441–454.
- Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society, Series B*, **39**, 1–38.

- Ederer, F., Axtell, L. M. and Cutler, S. J. (1961). The relative survival rate: A statistical methodology, *National Cancer Institute Monographs*, **6**, 101–121.
- Farewell, V. T. (1986). Mixture models in survival analysis: Are they worth the risk?, *Canadian Journal of Statistics*, **14**, 257–262.
- Francisci, S., Capocaccia, R., Grande, E., Santaquilani, M., Simonetti, A., Allemani, C., Gatta, G., Sant, M., Zigon, G., Bray, F. and Janssen-Heijnen, M. (2009). The cure of cancer: A European perspective, *European Journal of Cancer*, **45**, 1067–1079.
- Huang, L., Cronin, K. A., Johnson, K. A., Mariotto, A. B. and Feuer, E. J. (2008). Improved survival time: What can survival cure models tell us about population-based survival improvements in late-stage colorectal, ovarian, and testicular cancer?, *Cancer*, **112**, 2289–2300.
- 伊藤ゆり, 井岡亜希子 他(2009). Trends in the proportion cured for stomach cancer between 1975 and 2000 in Osaka, Japan, 第68回日本癌学会総会, 横浜.
- Kuk, A. Y. C. and Chen, C. H. (1992). A mixture model combining logistic regression with proportional hazards regression, *Biometrika*, **79**, 531–541.
- Lambert, P. C. (2007). Modeling of the cure fraction in survival studies, *The Stata Journal*, **7**, 351–375.
- Lambert, P. C., Dickman, P. W., Osterlund, P., Andersson, T., Sankila, R. and Glimelius, B. (2007a). Temporal trends in the proportion cured for cancer of the colon and rectum: A population-based study using data from the Finnish Cancer Registry, *International Journal of Cancer*, **121**, 2052–2059.
- Lambert, P. C., Thompson, J. R., Weston, C. L. and Dickman, P. W. (2007b). Estimating and modeling the cure fraction in population-based cancer survival analysis, *Biostatistics*, **8**, 576–594.
- Lu, W. and Ying, Z. (2004). On semiparametric transformation cure models, *Biometrika*, **91**, 331–343.
- Maller, R. and Zhou, X. (1996). *Survival Analysis with Long-term Survivors*, John Wiley & Sons, New York.
- National Cancer Center (2009). Abridged life tables in Japan for 1962–2008, retrieved 21 Dec, 2010, from <http://ganjoho.ncc.go.jp/professional/statistics/cohort01.html>.
- Nur, U., Shack, L. G., Rachet, B., Carpenter, J. R. and Coleman, M. P. (2010). Modelling relative survival in the presence of incomplete data: A tutorial, *International Journal of Epidemiology*, **39**, 118–128.
- 大阪府健康福祉部, 大阪府医師会, 大阪府立成人病センター(1987). 大阪府におけるがん登録, 第42報, 大阪府健康福祉部.
- 大阪府健康福祉部, 大阪府医師会, 大阪府立成人病センター(2004). 大阪府におけるがん登録, 第67報, 大阪府健康福祉部.
- Peng, Y. and Dear, K. B. G. (2000). A nonparametric mixture model for cure rate estimation, *Biometrics*, **56**, 237–243.
- Perme, M. P., Henderson, R. and Stare, J. (2009). An approach to estimation in relative survival regression, *Biostatistics*, **10**, 136–146.
- Royston, P. (2006). *ICE: Stata Module for Multiple Imputation of Missing Values*, Boston College Department of Economics, Massachusetts.
- Sato, N., Ito, Y., Ioka, A., Tanaka, M. and Tsukuma, H. (2009). Gender differences in stomach cancer survival in Osaka, Japan: Analyses using relative survival model, *Japanese Journal of Clinical Oncology*, **39**, 690–694.
- Shah, A., Stiller, C. A., Kenward, M. G., Vincent, T., Eden, T. O. and Coleman, M. P.

- (2008). Childhood leukaemia: Long-term excess mortality and the proportion 'cured', *British Journal of Cancer*, **99**, 219–223.
- StataCorp. (2009). Stata Statistical Software: Release 11, StataCorp LP, College Station, Texas.
- 杉本知之, 後藤昌司(1999). 退化調整型ハザード・モデルとその推測, *計算機統計学*, **12**, 85–106.
- Sugimoto, T. and Hamasaki, T. (2006). Properties of estimators of baseline hazard functions in a semiparametric cure model, *Annals of the Institute of Statistical Mathematics*, **58**, 647–674.
- Sugimoto, T., Hamasaki, T. and Goto, M. (2005). Estimation from pseudo partial likelihood in a semiparametric cure model, *Journal of the Japanese Society of Computational Statistics*, **18**, 33–46.
- Sy, J. P. and Taylor, J. M. G. (2000). Estimation in a Cox proportional hazards cure model, *Biometrics*, **56**, 227–236.
- Tsodikov, A. D., Ibrahim, J. G. and Yakovlev, A. Y. (2003). Estimating cure rates from survival data: An alternative to two-component mixture models, *Journal of the American Statistical Association*, **98**, 1063–1078.
- Verdecchia, A., De Angelis, R., Capocaccia, R., Sant, M., Micheli, A., Gatta, G. and Berrino, F. (1998). The cure for colon cancer: Results from the EURO CARE study, *International Journal of Cancer*, **77**, 322–329.
- Woods, L. M., Rachet, B., Lambert, P. C. and Coleman, M. P. (2009). 'Cure' from breast cancer among two populations of women followed for 23 years after diagnosis, *Annals of Oncology*, **20**, 1291–1292.

## Estimation of ‘Cure’ Fraction for Cancer Patients Using Population-based Cancer Registry Data

Yuri Ito<sup>1</sup> and Tomoyuki Sugimoto<sup>2</sup>

<sup>1</sup>Department of Cancer Control and Statistics, Osaka Medical Centre for Cancer and  
Cardiovascular Diseases

<sup>2</sup>Department of Biomedical Statistics, Osaka University

Five-year survival has been used as a measure for the prognosis of cancer patients. The prognosis of cancer patients varies among cancer sites; it is unknown whether the five-year follow-up is appropriate for all sites. It is important to estimate the proportion of ‘cure’ based on statistical modelling. A cure fraction model was originally proposed by Boag (1949). Use of population-based cancer registry data, however, did not become common until the late 1990s, when the concept of the relative survival was incorporated to the cure model. Recently, this method has been commonly applied to population-based cancer registry data in the US and European countries, because a methodology and statistical package to estimate ‘cure’ fraction have been developed. We introduced this method and showed an example applying the Osaka Cancer Registry data. We analysed stomach cancer patients diagnosed in 1975–2000. We monitored the time trends of cure fraction and median survival for fatal (uncured) patients for 25 years. The proportion of cure from stomach cancer increased by around 20% during those 25 years, and an estimated 50% of patients were cured. In addition, the increase of cure fraction was partly (around 20%) explained by change of the stage distribution due to earlier diagnosis. A remarkable increase of stage specific cure fraction demonstrated the improvement of treatment for stomach cancer.