NOWOTWORY Journal of Oncology • 2002 • volume 52
Number 6 • 509–511

Varia statistica

Methodology to predict long-term cancer survival from short-term data using Tobacco Cancer Risk and Absolute Cancer Cure models

Richard F. Mould

Introduction

This paper [1] published in November 2002 in the journal *Physics in Medicine & Biology* describes two new prediction models which were devised by R.F. Mould by extending and modifying the concept of the original lognormal model [2,3], which was reviewed in Nowotwory in 2001 [4]. The epidemiological concept underlying the Tobacco Cancer Risk model was the correlation of tobacco smoking with cancers of the larynx and lung. Thus a laryngeal cancer patient who after being cured of this neoplasm, can several years later also suffer from a second primary lung cancer. The validation test data used for this study is *unique* in that all ~ 1000 patients with localised [T1N0, T2N0 or T3N0] cancer of the larynx were treated [1944-68] by the same radiotherapist: Dr Manuel Lederman of the Royal Marsden Hospital, London. Follow-up was available to 1988.

In addition, data for ~ 6000 patients from Connecticut and from Metropolitan Detroit Cancer Registries with cancers of the breast, prostate, cervix uteri, thyroid, tongue and bladder treated 1973-1977 with follow-up to 1999 were obtained from the US Surveillance, Epidemiology and End Results [SEER] Program of the National Cancer Institute. Results showed that the methodology for the SLN, TCR and ACC models could be extended from cancer of the larynx to other cancers: thus making these models more powerful for cancer research.

The work in this paper showing that cancer can be totally cured sends a positive message on cancer to patients, physicians and surgeons for a disease which all too often, even if mistakenly where early stage tumours are concerned, seems to imply in the public view, all doom and gloom.

Table 12

Complete life analysis to 15-years post-treatment for 294 stage T1N0 and 145 stage T2N0 cases where all patients are excluded who have died with cancer of the larynx present or with cancer of the lung present. For stage T1N0 and T2N0 and survival time T=10-20 years agreement, is less than or equal to 1 standard error of the observed T-year survival fraction calculated using the Kaplan-Meier method [5]. These results show that the Haybittle [7] definition of cancer cure as understood by the man-in-the-street has been achieved: "The complete elimination of disease, so that a patient cured of cancer would be one whose subsequent medical history and length of life was completely unaffected by their having had cancer".

Survival time T years	Ca. Larynx Stage T1N0 Survival fractions Observed SF	Expected SF	Standard error of Observed SF	Ca. Larynx Stage T2N0 Survival fractions Observed SF	Expected SF	Standard error of Observed SF
5	0.86	0.81	0.02	0.87	0.82	0.03
10	0.65	0.63	0.03	0.62	0.64	0.04
15	0.46	0.45	0.03	0.46	0.48	0.04
20	0.34	0.31	0.03	0.30	0.34	0.04
25	0.23	0.19	0.03	0.27	0.23	0.04

The Abstract from *Physics in Medicine & Biology is* reproduced below together with Figures 1, 2 and 5 and Table 12 from this *PMB* paper.

Abstract

Three parametric statistical models have been fully validated for cancer of the larynx for the prediction of long-term 15, 20 and 25 year cancer-specific survival

To date it has been generally assumed for early stage disease that although for some 5-10 years after treatment the survival experience of this patient subgroup might be no different from that expected in the matched group, thereafter the death rate of this subgroup becomes lower than that of the matched group. This implies that surviving cancer patients cured of their disease tend to die of other conditions at a higher than normal rate as they become older, and therefore cancer is never *totally* cured. Our conclusion is that at least for cancer of the glottic larynx, the answer to the question "Can cancer *totally* be cured?" is "Yes to at least 15-years post-treatment and also probably to 25 years".

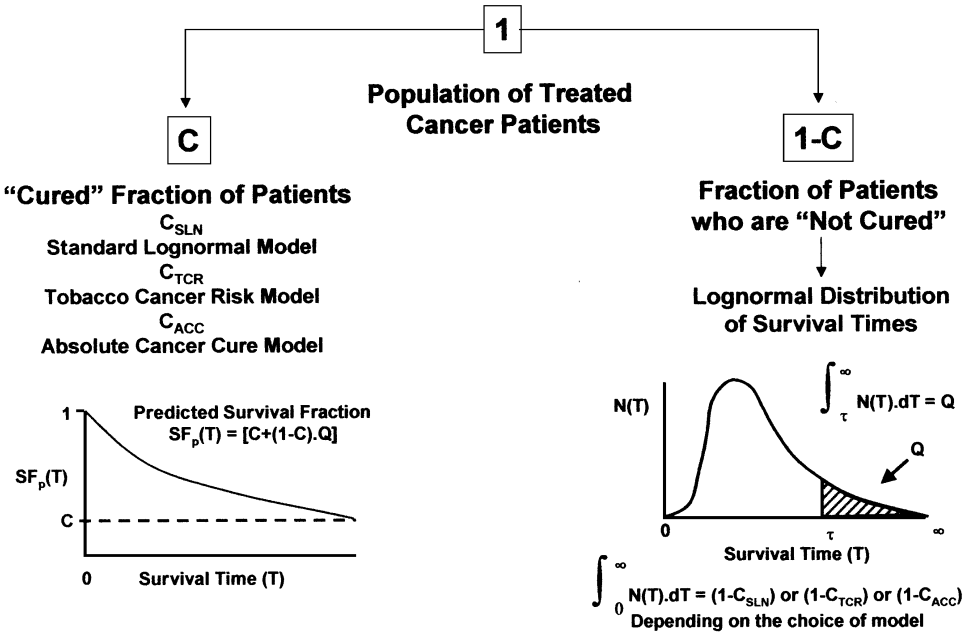

Figure 1. Schematic diagram for the three prediction models studied in which the statistically cured fraction of patients is denoted by C. For a given long-term survival time τ years the tail of the lognormal distribution, denoted by Q, will be the longest for the ACC model and the shortest for the SLN model.

fractions when short-term follow-up data was available for just 1-2 years after the end of treatment of the last patient. In all groups of cases the treatment period was only 5 years. Three disease stage groups were studied, T1N0, T2N0 and T3N0. The models are the *Standard Lognormal* (SLN) first proposed by Boag in 1949 [2] but only ever fully validated for cancer of the cervix, Mould and Boag in 1975 [3] and two new models which have been termed *Tobacco Cancer Risk* (TCR) and *Absolute Cancer Cure* (ACC).

In each, the frequency distribution of survival times of defined groups of cancer deaths is lognormally distributed: larynx only (SLN), larynx and lung (TCR) and all cancers (ACC). All models each have three unknown parameters but it was possible to assume a value for the lognormal parameter *S a priori*. By reduction to two unknown parameters the model stability has been improved.

The material used to validate the methodology consisted of case histories of 965 patients, all treated during the period 1944-1968 by Dr Manuel Lederman of the Royal Marsden Hospital, London, with follow-up to 1988. This provided a follow-up range of 20-44 years and enabled predicted long-term survival fractions to be compared with the actual survival fractions, calculated by the Kaplan and Meier method [5].

The TCR and ACC models are better than the SLN model and for a maximum short-term follow-up of 6 years, the 20 and 25 year survival fractions could be predicted. Therefore the *number of follow-up years saved* are

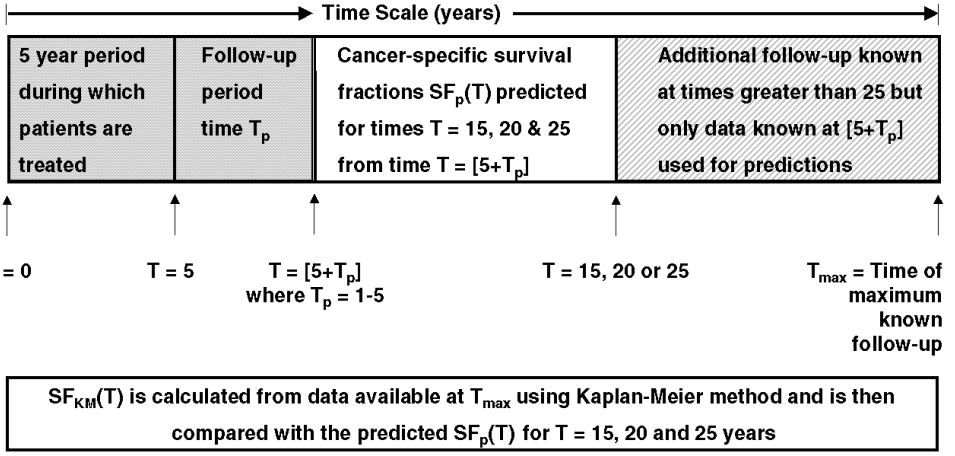

Figure 2. Schematic diagram illustrating the concept of the Phase 2 validation procedure. The Phase 1 validation procedure consists of a minimum chi-squared test for lognormality of the observed survival times for a given model: SLN, TCR or ACC.

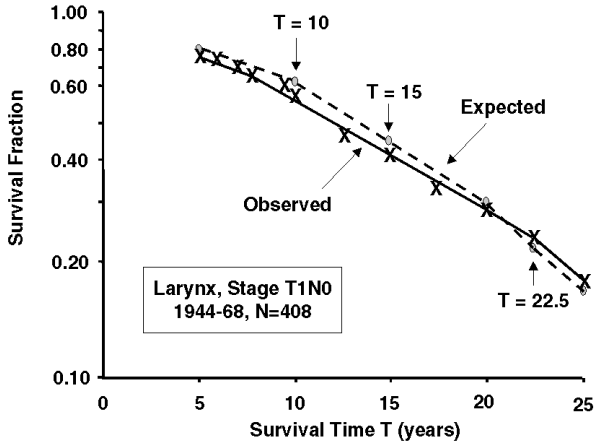

Figure 5. Comparison of observed and expected survival rates on a semi-logarithmic graph plot for stage T1N0 cancer of the larynx patients. This demonstrates the concept of statistical cure proposed by Easson and Russell of the Christie Hospital, Manchester [6] in 1968. This was described in the following terms "If a group of patients subsequently show an annual death rate from all causes which is similar to that of the normal population of the same age and sex structure, then this group can be said to be cured".

respectively 14 years and 19 years. Clinical trial results using the TCR and ACC models can thus be analysed much earlier than currently possible.

Absolute cure from cancer was also studied, using not only the prediction models which incorporate a parameter for a *statistically cured* fraction of patients C_{SLN}, C_{TCR} and C_{ACC}, but because of the long follow-up range of 20-44 years, also by *Complete Life Analysis*. The survival experience of those who did not die of their original cancer of the larynx was compared to the expected survival experience of a population with the same age, birth cohort and sex structure.

Richard F. Mould MSc PhD
41, Ewhurst Avenue
Sanderstead, South Croydon
Surrey, CR2 ODH
United Kingdom

References

1. Mould RF, Lederman M, Tai P, Wong JKM Methodology to predict long-term cancer survival from short-term data using Tobacco Cancer Risk and Absolute Cancer Cure models. *Phys Med Biol* 2002; 47: 3893-3924, Online at stacks.iop.org/PMB/47/3893.
2. Boag JW. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *J Roy Stat Soc Series B* 1949; 11: 15-53.
3. Mould RF, Boag JW. A test of several parametric statistical models for estimating success rate in the treatment of carcinoma cervix uteri. *Br J Cancer* 1975; 32: 529-550.
4. Mould RF. Lognormal modelling for the prediction of long-term survival rates from short-term follow-up data. *Nowotwory J Oncol* 2001; 51: 339-354.
5. Kaplan EL, Meier P. Non-parametric estimation from incomplete observations. *J Amer Stat Assoc* 1958; 53: 457-482.
6. Easson EC, Russell MH. *The curability of cancer in various sites*. London: Pitman Medical; 1968.
7. Haybittle JL. What is cure in cancer? in: Stoll BA, ed, *Cancer treatment: end point evaluation*. New Horizons in Oncology; 2. Chichester: Wiley; 1983, 3-21.