A flowgraph model for bladder carcinoma

Gregorio Rubio[1⋆], Belén García-Mora[1], Cristina Santamaría[1], and José Luis Pontones[2]

[1] Instituto de Matemática Multidisciplinar. Universidad Politécnica de Valencia
[2] Departamento de Urología. Hospital Universitario La Fe de Valencia

Abstract. Superficial bladder cancer has been the subject of numerous studies for many years, but the evolution of the disease still remains not well understood. After the tumor has been surgically removed, it may reappear at a similar level of malignancy or progress to a higher level. The process may be reasonably modeled by means of a Markov process. However, in order to fully model the evolution of the disease, this approach is insufficient. The semi-Markov framework allows a more realistic approach, but calculations become frequently intractable. In this context, flowgraph models provide an efficient approach to successfully manage the evolution of superficial bladder carcinoma.

Keywords: flowgraph model, bladder carcinoma, Erlang distribution

1 Introduction

Bladder tumor is a challenge for urology. It supposes an important public health problem because it is biologically very aggressive and presents a high prevalence in the western countries. Approximately 75–85 % of patients with newly diagnosed bladder carcinoma present non muscle-invasive bladder carcinoma (NMI-BC), which can be managed with transurethral resection (TUR), that is a surgical endoscopic technique used to remove the macroscopic tumor from the inner of the bladder. The object of this study is the NMI-BC, that supposes about 70 % of the total expenses of this disease.

Biotechnological advances have allowed us to use different therapeutic procedures (surgery, radiotherapy, chemotherapy, immunotherapy) successfully but still many patients suffer an unfavourable outcome without control of disease. In practice urologists have an important problem: some patients with similar characteristics undergo a different evolution. Consequently, that situation creates a problem with the treatment to be applied. Urologists need tools to accurately predict the real evolution of the disease, that help them to improve treatment modalities and follow-up schemes of non-muscle invasive bladder cancer patients. In this regard an important contribution [1] appeared in European Urology, the official journal of the European Association of Urology. By means of look up tables the probability of recurrence and progression for a patient is provided.

⋆ corresponding author.

However only time to first recurrence is considered, and the analysis is reduced to the Cox proportional hazards regression model. Later works have studied the model validation, finding some limitations [2].

Our team has been working with urologists from University Hospital La Fe for the last ten years. We have developed several models trying to capture different aspects of the disease evolution. Our aim for the near future is to detect the most relevant predictive factors, and also to perform an accurate model of the disease evolution. The first objective includes to investigate at the genetic and molecular level, and the second one could be achieved with a suitable multistate model. While the process may be reasonably modeled by means of a Markov process, in order to fully model the evolution of the disease this approach is insufficient. Specifically, it is possible that time spent in a state influences the future evolution of the process, i.e., it not only depends on the current state. The semi-Markov framework allows a more realistic approach, but calculations become frequently intractable. In this context, flowgraph models provide an efficient approach for the analysis of time-to-event data, since their introduction in this field a few years ago [3]. The present work is a first step in order to explore the evolution of the recurrence progression process by means of this methodology.

The paper is organized as follows: in section 2 we review a few basic concepts of survival analysis and we present the essentials of flowgraph models. Section 3 deals with a simple flowgraph model for the recurrence-progression process in NMI-BC, constructed using a database from La Fe University Hospital of Valencia (Spain). Finally, in section 4 some discussion is given.

2 Survival analysis and flowgraph models

2.1 Survival analysis

Survival analysis techniques deal with the analysis of data in the form of times from well-defined *time-origin* until the occurrence of some particular event or *end-point*.

To summarize survival data there are two key functions: the Survival Function and the Hazard Function. Let T be the random variable associated with the survival time (time until the ocurrence of the event).

The Survival Function is

$$S(t) = \mathrm{P}(\mathrm{T} \geq t) = 1 - F(t)$$

where $F(t)$ is the distribution function of T. It expresses the probability that an individual survives from the time origin to some time beyond t.

The Hazard Function is given by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\mathrm{P}(\, t \leq \mathrm{T} < t + \Delta t \,|\, \mathrm{T} \geq t)}{\Delta t},$$

what expresses the hazard rate or the instantaneous event rate.

In survival analysis data are frequently censored [4], what means that the event of interest has not been observed. The time of follow–up of those patients must be taken into account, because it informs us of the fact that the individual was free of event until the present moment. For instance we started with 620 patients, of whom 247 underwent a recurrence, 16 a progression, and 358 had censored times, what means that at the time of their last revision they had no recurrence or progression.

2.2 Flowgraph models

A flowgraph model is a graphical representation of a multistate model, that consists of directed line segments (*branches*) connecting the states, namely, a directed graph. The branches are labeled with *transmittances*, that are the transition probability p_{ij} from state i to state j multiplied by an integral transform $G_{ij}(s)$ of the transition time probability density function (PDF). This transform can be a characteristic function (CF), a moment generating function (MGF), a Laplace transform (LT), or even an empirical transform [5] [6]. Flowgraphs are used to represent semi-Markov processes, given that allowed waiting time distributions go beyond the exponential distribution directly linked to Markov processes.

For instance, Figure 1 shows the flowgraph of the three-state illness-death model, that we will use in this paper.

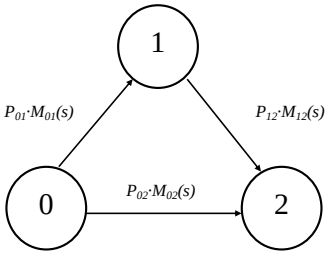

Fig. 1. Three-state illness-death model.

The transmittances are combined according to a systematic procedure (see [7], section 2.5), in order to compute the transforms for the transitions of interest. The final step is to invert these transforms to obtain PDFs.

Flowgraph models for stochastic networks were introduced by Butler and Huzurbazar [3]. An account of the theory developed up to 2005 may be found in [7]. A recent contribution proposing a prognostic model is [8].

3 A flowgraph model for bladder carcinoma

3.1 Data

The database was obtained from La Fe University Hospital of Valencia (Spain). It records clinico-pathological information from 620 patients, followed between January 1995 and January 2010. The mean follow–up time for the entire cohort was 3 years and 11 months. The data record several recurrence times. This means that some patients have no recurrence at all, some have one or more recurrences, and some have progression (directly of after some recurrence). In our model we have considered progressions and one recurrence. As stated above, 246 patients underwent a recurrence, 16 a progression, and 358 had censored times. Then, 44 patients were lost. From the remaining 202 patients, 111 underwent again a recurrence, 8 a progression and 83 were censored.

3.2 Flowgraph model

Our aim in this paper is to test the flowgraph methodology in this particular problem, and so we perform the simple model of Figure 1. In state 0 the patient is free of disease, after the RTU of the primary tumor. State 1 is the first recurrence, and state 2 is progression. Time is given in years.

The first step is to select a suitable distribution for the waiting time in each transition. Our approach will be to compute the empirical distributions and approximate them using mixtures of Erlang distributions. Specifically we use a linear combination of three Erlang proposed in [9]. The coefficients were calculated by means of a non-negative least squares fit (Lawson-Hanson algorithm [10]).

Our aim is to model the overall risk of progression. Therefore we have to calculate the PDF of the overall time to progression. For this we compute the Laplace transform of the global transition from state 0 to state 2. To recover the PDF we use a variant of the inversion algorithm EULER [8]. From this function we obtain the survival function (with regard to progression), that is shown in Figure 5, jointly with the empirical survival function.

4 Discussion

Flowgraph methodology is very flexible. It allows the model to incorporate multiple recurrences, and recently also covariates [5]. We may explore many parametric models looking for the distributions that match the data better. Moreover non-parametric approaches are also avalaible [6].

This versatility, along with the inclusion of molecular biomarkers, allow us to expect to get a very accurate model in a not too distant future.

Acknowledgments. This study has been funded by Vicerrectorado de Investigación de la Universitat Politécnica de Valéncia. Reference 2406.

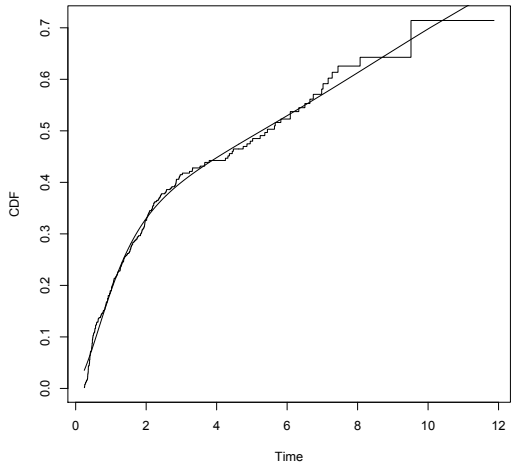

Fig. 2. Erlang mixture (*smooth line*) and empirical distribution (*step function*) for transition 01.

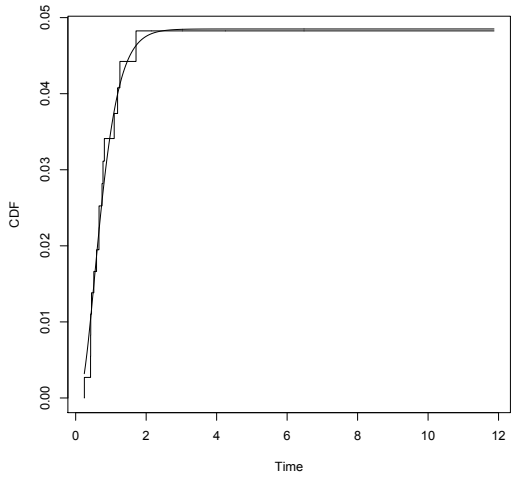

Fig. 3. Erlang mixture (*smooth line*) and empirical distribution (*step function*) for transition 02.

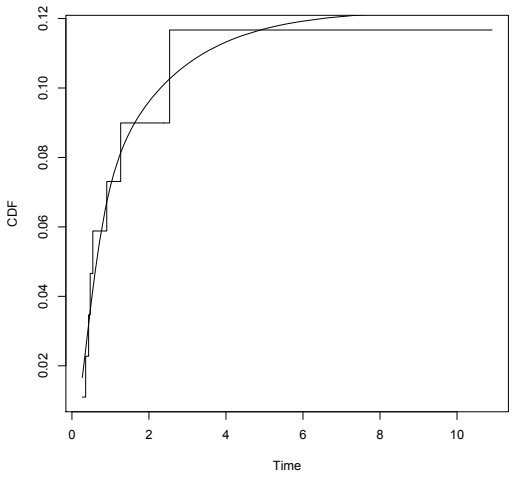

Fig. 4. Erlang mixture (*smooth line*) and empirical distribution (*step function*) for transition 12.

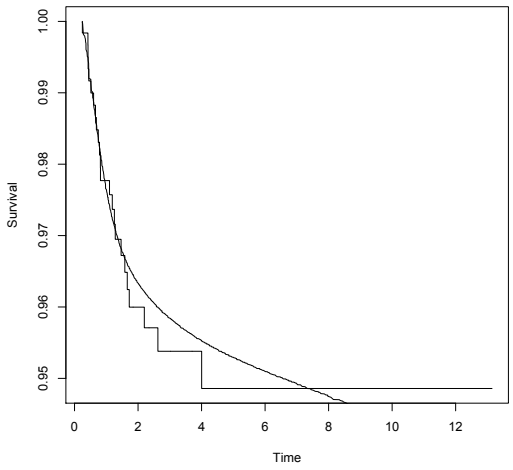

Fig. 5. Survival function model (*smooth line*) and empirical survival function (*step function*).

References

1. Sylvester, R. J., van der Meijden, A. P., Oosterlinck, W., Witjes, J. A., Bouffioux C., Denis, L., Newling, D. W., Kurth, K.: Predicting recurrence and progression in individual patients with stage ta t1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. Eur. Urol. 49, 475–7 (2006)
2. Fernandez-Gomez, J., Madero, R., Solsona, E., Unda, M., Martínez-Pieiro, L., Gonzalez, M., Portillo, J., Ojea, A., Pertusa, C., Rodriguez-Molina, J., Camacho, J., Rabadan, M., Astobieta, A., Montesinos, M., Isorna, S., Muntaola, P., Gimeno, A., Blas, M., Martinez-Pieiro, J. A. : The EORTC tables overestimate the risk of recurrence and progression in patients with non muscle-invasive bladder cancer treated with bacillus calmette-guerin: External validation of the EORTC risk tables. Eur. Urol. 60, 423–430 (2011)
3. Butler, R.W, Huzurbazar, A.V.: Stochastic network models for survival analysis. J. Am. Statist. Assoc.. 92, 246–257 (1997)
4. Klein, J.P., Moeschberger, M.L.: Survival Analysis. Techniques for Censored and Truncated Data, Second Edition. Springer, New York (2003)
5. Huzurbazar, A., Williams, B.: Incorporating Covariates in Flowgraph Models: Applications to Recurrent Event Data. Technometrics. 52, 198-208 (2010)
6. Collins, D.H., Huzurbazar, A.V.: System reliabyility and safety assessment using non-parametric flowgraph models. Proc. IMechE Part O: J. Risk and Reliability. 222, 667–674 (2008)
7. Huzurbazar, A.: Flowgraph Models for Multistate Time-To-Event Data. Wiley, New York (2005)
8. Collins, D.H., Huzurbazar, A.V.: Prognostic models based on statistical flowgraphs. Appl. Stochastic Models Bus. Ind. 28, 141–151 (2012)
9. Pérez-Ocón, R.; Segovia. M.C.: Modeling lifetimes using phase-type distributions. In: Aven, T., Vinnem, J.E. (eds.) Risk, Reliability and Societal Safety, Proceedings of the European Safety and Reliability Conference 2007 (ESREL 2007), Taylor & Francis,3rd edition (2007)
10. Mullen, K.M.; van Stokkum, I.H.: nnls: the Lawson-Hanson algorithm for non-negative least squares (NNLS) `http://www.R-project.org`