Evaluation of a Probabilistic Model for Staging of Oesophageal Carcinoma

Linda C. van der Gaag, Silja Renooij,
Department of Computer Science, Utrecht University,
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands
{linda, silja}@cs.uu.nl

Berthe M.P. Aleman, and Babs G. Taal
The Netherlands Cancer Institute, Antoni van Leeuwenhoekhuis,
Department of Radiation Oncology and Gastroenterology,
Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands
{baleman, bgtaal}@nki.nl

Abstract

With the help of two experts in gastrointestinal oncology from the Netherlands Cancer Institute, Antoni van Leeuwenhoekhuis, a decision-support system is being developed for patient-specific therapy selection for oesophageal carcinoma. The kernel of the system is a probabilistic model describing the characteristics of oesophageal carcinoma and the pathophysiological processes of invasion and metastasis. Using data from 185 patients, an evaluation study of the model was conducted. We found that for 86% of the patients, the model established the stage of the patient's carcinoma correctly.

1 Introduction

The Netherlands Cancer Institute, Antoni van Leeuwenhoekhuis, is a specialised center for the treatment of cancer patients. In the hospital, every year some eighty patients receive treatment for oesophageal carcinoma. These patients currently are assigned to a therapy by means of a standard protocol that includes a small number of prognostic factors. Based upon this protocol, some 75% of the patients show a favourable response to the therapy instilled. To arrive at a more fine-grained protocol with a higher favourable response rate, a decision-support system is being developed for patient-specific therapy selection for oesophageal carcinoma. The system is constructed and refined with the help of two experts in gastrointestinal oncology from the Netherlands Cancer Institute, Antoni van Leeuwenhoekhuis, and is destined for general use in clinical practice.

The kernel of our system is a probabilistic model of oesophageal carcinoma. The model describes the various characteristics of an oesophageal carcinoma and the pathophysiological processes underlying its invasion into the oesophageal wall and its metastasis. The model further specifies the sensitivity and specificity characteristics of the various diagnostic tests in use. For prognostication, the model describes the possible effects of the therapeutic alternatives available. Upon a patient's symptoms and test results having been entered, the model establishes the stage of the patient's carcinoma and predicts the most likely outcomes of the various different treatment alternatives.

We conducted an evaluation study of our probabilistic model, using data from 185 patients from the Antoni van Leeuwenhoekhuis diagnosed with oesophageal carcinoma. The study focused on the part of the model that provides for establishing the stage of a patient's carcinoma. This stage summarises the carcinoma's characteristics, depth of invasion, and extent of metastasis and is indicative of the likely outcome of treatment. The study revealed various types of anomaly in the data and served to identify a small number of variables missing from the model. After providing for the anomalies and missing variables, we found that for 86% of the patients the model established the correct stage.

In this paper, we present the results of the evaluation study of our probabilistic model for staging of oesophageal carcinoma. In Section 2, we briefly describe the model; we comment upon its construction in Section 3. In Section 4, we present the results of the study. The paper ends with some concluding observations in Section 5.

2 The probabilistic model

As a consequence of a lesion of the oesophageal wall, for example as a result of frequent reflux, a carcinoma may develop in a patient's oesophagus. An oesophageal carcinoma has various characteristics that influence its prospective growth. These characteristics include the location of the carcinoma in the oesophagus and its histological type, length, and macroscopic shape. An oesophageal carcinoma typically invades the oesophageal wall and upon further growth may invade neighbouring structures such as the trachea and bronchi. In due time, the carcinoma may give rise to lymphatic metastases in distant lymph nodes and to haematogenous metastases in, for example, the lungs and the liver. The characteristics, depth of invasion, and extent of metastasis, summarised in the carcinoma's stage, largely influence a patient's life expectancy and are indicative of the effects and complications to be expected from the various therapeutic alternatives. To establish these factors in a patient, typically a number of diagnostic tests are performed, ranging from multiple biopsies of the primary tumour to gastroscopic and endosonographic examination of the oesophagus and a CT-scan of the patient's chest and liver. The various tests differ considerably with respect

to their sensitivity and specificity characteristics.

The state-of-the-art knowledge about oesophageal carcinoma is captured in a probabilistic model, also known as a Bayesian network [1]. The model includes a graphical structure composed of statistical variables and their probabilistic interrelationships. Each variable represents a diagnostic or prognostic factor that is relevant for establishing the stage of a patient's carcinoma or for predicting the outcome of treatment. The probabilistic influences between the variables are represented by arcs; their strengths are indicated by conditional probabilities. Figure 1 depicts a small part of the model, showing the prior probability distribution per variable.

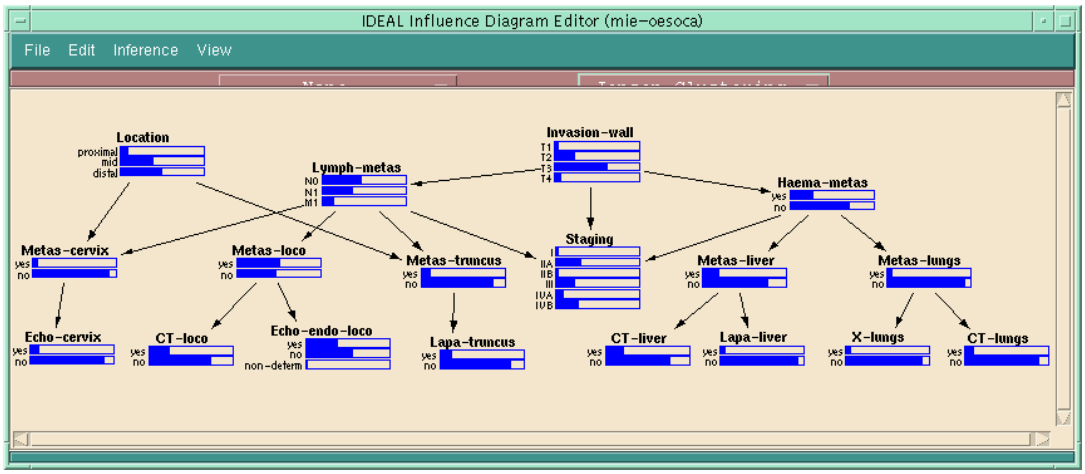

Figure 1: A part of the probabilistic model, pertaining to lymphatic and haematogenous metastases of an oesophageal carcinoma.

Our probabilistic model of oesophageal carcinoma currently includes 70 statistical variables. Of these, 40 variables pertain to the characteristics of an oesophageal carcinoma, to the depth of invasion, the extent of metastasis, and the sensitivity and specificity characteristics of the diagnostic tests in use. The remaining 30 variables model the possible effects and complications of the therapeutic alternatives available. For the statistical variables, a total of 4000 conditional probabilities have been specified.

3 Constructing the probabilistic model

The probabilistic model of oesophageal carcinoma is constructed and refined with the help of two experts in gastrointestinal oncology from the Netherlands Cancer Institute. In a sequence of interviews over a period of two years, the experts identified the relevant diagnostic and prognostic factors to be captured as statistical variables in the model, their possible values, and the influential relationships between them. The relationships between the variables were elicited using the

notion of causality. Typical questions asked during the interviews were "What could cause this effect ?" and "What manifestations could this cause have ?". The elicited causal relationships were expressed in graphical terms by taking the direction of causality for directing arcs between related variables.

Once the graphical structure of the model was considered robust, we focused attention on the assessment of the probabilities required. Various different sources of probabilistic information appeared to be readily available for the task. Neither data nor literature, however, yielded any usable results. As a consequence, the experts involved had to assess the thousands of probabilities required. For this purpose, we used an elicitation method tailored to obtaining a large number of judgemental probabilities in little time. At the heart of this method lies the idea of presenting experts with a separate figure for every probability to be assessed [2]. Figure 2 shows, as an example, the figure pertaining to the conditional probability $\Pr(\mathit{Invasion} = \mathit{T2} \mid \mathit{Shape} = \mathit{polypoid}, \mathit{Length} < 5\mathit{cm})$ for the oesophagus model. On the left of the figure is a fragment of text that transcribes the probability under consideration, thereby circumventing the need for mathematical notation. A vertical scale for marking assessments is depicted to the right of the text

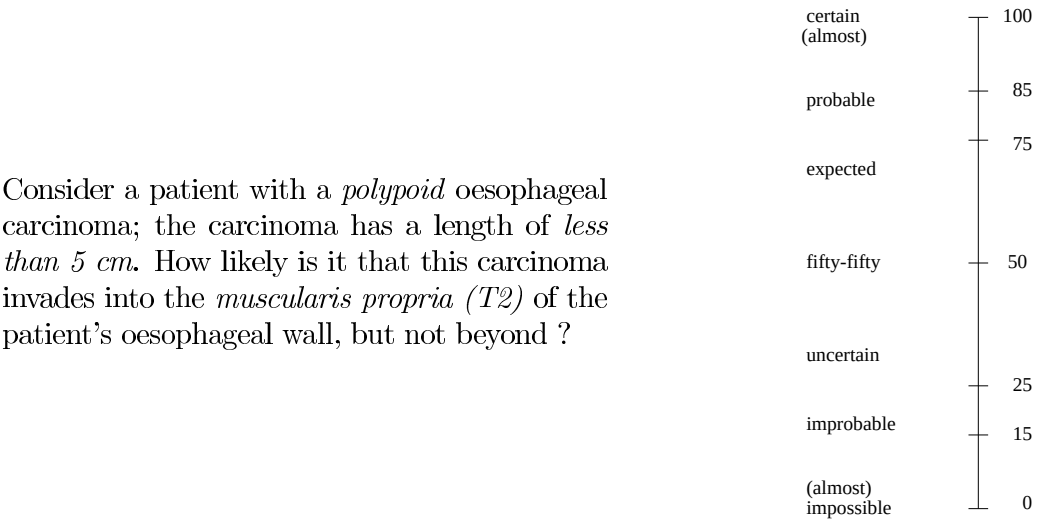

Figure 2: An example fragment of text and scale for probability assessment.

fragment; indicated on this scale are numerical and verbal anchors of uncertainty. The figures pertaining to the various conditional probabilities to be assessed were grouped in such a way that the probabilities from the same conditional distribution could be taken into consideration simultaneously; the figures were presented in groups of two or three on consecutive single-sided sheets of paper. With the method, we elicited from our domain experts the probabilities required for the oesophagus model at a rate of 150 – 200 probabilities per hour in various interviews over a period of one year.

4 Evaluation of the probabilistic model

We conducted an evaluation study of our probabilistic model of oesophageal carcinoma with data from 185 patients from the Antoni van Leeuwenhoekhuis. The study focused on the part of the model that provides for establishing the stage of a patient's carcinoma; the stage of an oesophageal carcinoma can be either I, IIA, IIB, III, IVA, or IVB, in the order of advanced disease. For 29 patients from our data collection the stage of the oesophageal carcinoma was not recorded, leaving us with 156 patients for evaluation.

In the first evaluation of the model, we entered, for each patient from the data collection, all diagnostic test results available. From the model, we then computed the most likely stage of the patient's carcinoma and compared it with the stage recorded in the data. The leftmost table in Figure 3 shows the results from this first evaluation. For 80 of the 156 patients, the stage of the carcinoma recorded in the data matched the stage with highest probability computed from the model. Under the assumption that the stages recorded in the data are correct, therefore, for 51% of the patients the model established the correct stage. We would like to note that this percentage is not uncommonly found in initial evaluations of knowledge-based systems [3].

Taking the results from this first evaluation for a point of departure, we carefully examined the data of the patients for whom the probabilistic model returned a stage different from the recorded one. We identified three major sources of mismatch that could to a large extent be attributed to the data. For 10 patients, the stage recorded in the data was acknowledged by the experts to be incorrect. Various other anomalies in the data constituted the second source of mismatch. For example, during surgery for some patients a deeper invasion of the carcinoma into the oesophageal wall was found than conjectured from endosonographic findings. For these patients, the pre-surgical findings and the post-surgical stage were recorded in the data. As just the findings were entered in our evaluation, a stage different from the recorded one was established by the model. The third major source of mismatch was found in the way findings were entered into the patients' medical records. Often no distinction was made between facts and findings from diagnostic tests. For example, for many patients, the medical record stated the presence of metastases in the cervical lymph nodes without indicating how this

		network, first evaluation							*network, second evaluation*						
		I	IIA	IIB	III	IVA	IVB	*total*	I	IIA	IIB	III	IVA	IVB	*total*
	I	**2**	0	0	0	0	0	2	**2**	0	0	0	0	0	2
	IIA	0	**34**	0	3	0	0	37	0	**37**	0	1	0	0	38
data	IIB	0	3	**0**	3	0	0	6	0	1	**0**	3	0	0	4
	III	1	16	1	**24**	1	1	44	1	11	1	**34**	0	0	47
	IVA	1	9	2	23	**6**	1	42	0	0	0	1	**38**	0	39
	IVB	0	2	0	8	1	**14**	25	0	0	0	3	0	**23**	26
	total	4	64	3	61	8	16	156	3	49	1	42	38	23	156

Figure 3: The results from the evaluation study.

fact was established. Without explicitly stated test results, the model could not establish the presence of these metastases and as a consequence returned an incorrect stage. Explicit test results, in the medical records, for lymph nodes near the truncus coeliacus, on the other hand, led to the conclusion that two variables modeling diagnostic tests to this end were missing from the model.

Building upon the above observations, we performed a second evaluation of the model. For this purpose, we corrected the erroneous stages and various other anomalies in the data in close consultation with the experts. We entered for each patient the test results available as before. In addition, we entered for each patient the facts stated in the medical record for which no test results were recorded. On average, 1.5 additional facts were entered per patient. The results of this second evaluation are shown in the rightmost table in Figure 3. For 134 of the 156 patients, the stage of the carcinoma recorded in the (modified) data matched the stage computed from the model. Once again under the assumption that the stages recorded in the data are correct, in 86% of the patients the model established the correct stage. We would like to note that this percentage of correct staging probably slightly overrates the actual performance of the model as for the additional facts entered the sensitivity and specificity characteristics of the tests concerned were not taken into consideration.

5 Conclusions

A decision-support system is being developed for patient-specific therapy selection for oesophageal carcinoma. The kernel of the system is a probabilistic model describing the characteristics of oesophageal carcinoma and the pathophysiological processes of invasion and metastasis. Using data from 185 patients, the model was evaluated. A first evaluation revealed various sources of mismatch between the stage recorded in the data for a patient's carcinoma and the stage computed from the model. To a large extent, these sources could be attributed to the data. We feel that this is not uncommon in evaluation studies like the present one. The first evaluation in addition served to identify a small number of variables missing from the model. After correcting various anomalies in the data and providing for the missing variables, we found that for 86% of the patients a correct stage was established by the model. Given that the probabilities used are rough initial assessments and the patient data require further clearing out, the results from the study are quite encouraging. We would like to note that in our decision-support system, the depth of invasion and extent of metastasis of a patient's carcinoma are of interest rather than its stage. Focusing our evaluation study on the summarising stage so far has served to yield insight in the diagnostic part of the network. We are now in the process of investigating the ability of the model to predict the outcome of treatment and hope to report the results in the near future.

References

[1] F.V. Jensen (1996). *An Introduction to Bayesian Networks*. UCL Press, London.

[2] S. Renooij, C.L.M. Witteman (1999). Talking probabilities: communicating probabilistic information with words and numbers. *International Journal of Approximate Reasoning*, vol. 22, pp. 169 – 194.

[3] E.S. Berner, G.D. Webster, A.A. Shugerman, J.R. Jackson, J. Algina, A.L. Baker, E.V. Ball, C.G. Cobbs, V.W. Dennis, E.P. Frenkel, L.D. Hudson, E.L. Mancall, C.E. Rackley, and O.D. Taunton (1994). Performance of four computer-based diagnostic systems. *The New England Journal of Medicine*, vol. 330, pp. 1792 – 1796.