Ann. Inst. Statist. Math.
Vol. 45, No. 3, 401–418 (1993)

BAYESIAN ANALYSIS OF LYMPHATIC SPREADING PATTERNS IN CANCER OF THE THORACIC ESOPHAGUS*

AKIFUMI YAFUNE[1], TOSHIKI MATSUBARA[2] AND MAKIO ISHIGURO[3]

[1] *The Kitasato Institute, 5-9-1 Shirokane, Minato-ku, Tokyo 108, Japan*
[2] *Cancer Institute Hospital, 1-37-1 Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan*
[3] *The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan*

(Received November 25, 1991; revised February 3, 1993)

Abstract. For the treatment of patients with cancer of the thoracic esophagus, lymphatic spreading is one important factor to infer how advanced their cancer is. We introduced a one-dimensional scale based on lymphatic spreading patterns, the stage of cancer, to express how advanced their cancer is, and we proposed a method to infer each patient's stage from his lymphatic spreading pattern by applying a Bayesian model. Our Bayesian model was built based on the assumption that lymphatic spreading in cancer could be explained as what was brought about by the advance of stage. In the modeling, we introduced the probability of what stage each patient was in as a prior distribution. We also introduced distribution functions of Weibull distributions to express the relation between the advance of stage and the increase of the probability of metastasis. Our model was applied to the data of nodal involvement obtained from 103 patients with cancer of the thoracic esophagus and the parameters were estimated with the maximum likelihood method. AIC was used to check that the data had enough information to be divided into the stages of a clinically reasonable number. With the estimated parameters, we inferred the probability of metastasis to each lymph node in each stage and calculated by Bayes' theorem with 31 new patients the probability of what stage they were in. The results well represented some characteristics of the lymphatic spreading and suggested the appropriateness of our approach.

Key words and phrases: Cancer of the thoracic esophagus, lymphatic spreading pattern, Bayesian model, Bayes' theorem, Weibull distribution, AIC.

1. Introduction

For the treatment of patients with cancer of the thoracic esophagus, lymphatic spreading is an important factor to infer how advanced their cancer is. We introduce a one-dimensional scale based on lymphatic spreading patterns, the stage

* The present study was carried out under the ISM Cooperative Research Program (91-ISM·CRP-18).

of cancer, to express how advanced their cancer is. The stage of cancer used in this paper is different from the general one defined in the prevailing classification. The lymphatic component included in the prevailing classification of cancer of the thoracic esophagus has many inappropriate points from the clinical point of view. What we are trying is to find a way to define a new "stage" which can be inferred from lymphatic spreading patterns. Note that it is not the aim of our approach to build some models which explain the relation between lymphatic spreading patterns and the prevailing classification. The "real stage of cancer" should be defined by not only lymphatic spreading but the other factors used in the prevailing classification. In cancer of the thoracic esophagus, however, the lymphatic spreading, whose processes are quite complicated, is a particularly important to infer the stage of cancer. Thus, the inference from lymphatic spreading patterns is of clinical interest and we introduce the one-dimensional scale mentioned above, the stage of cancer. Although our approach is basically categorized into the clustering, we use the word "stage" because our clusters have one-dimensional orders.

The probability of metastasis to each lymph node becoming larger according to the advance of stage, the profiles of probabilities are not similar in all lymph nodes. For example, a certain lymph node is apt to be involved in earlier stages and another one is involved in advanced stages only. We should take into account the difference among the profiles to infer the stage of cancer from lymphatic spreading patterns.

In clinical areas, the number of involved lymph nodes has so far been thought as one important factor to infer the stage of cancer in most cases. However, the patients with the same number of involved lymph nodes may be in different stages and the patients with more involved lymph nodes are not necessarily in more advanced stage. Suppose a case of three lymph nodes, L_1, L_2 and L_3, and assume that L_1 and L_2 are apt to be involved in earlier stages and L_3 is involved in advanced stages only. We denote a patient A with metastasis to L_1 as $A(+,-,-)$. Of the three patients, $A(+,-,-)$, $B(-,+,-)$ and $C(-,-,+)$, which one is in the most advanced stage? We sometimes come across such rare cases as C in clinical areas. Since L_3 is involved in C only, it is clear that this patient is in the most advanced stage. Comparing another patient $D(+,+,-)$ with C, which one is in more advanced stage? The number of involved lymph nodes is 1 in C and 2 in D. From the number of involved lymph nodes, D seems to be in more advanced stage. But considering that L_3 is involved in C only, C might be in more advanced stage. Thus, to infer the stage of cancer, we should take into account lymphatic spreading patterns, which are particularly important in patients with cancer of the thoracic esophagus. There has so far been no established method of inference from lymphatic spreading patterns.

The purpose of this paper is to propose a method to infer the stage of cancer of the thoracic esophagus from lymphatic spreading patterns by applying a Bayesian model. Our Bayesian model is built based on the assumption that lymphatic spreading in cancer could be explained as what is brought about by the advance of stage. In the modeling, we introduce the probability of what stage each patient is in as a prior distribution.

We apply our model to the data of nodal involvement obtained from patients

with cancer of the thoracic esophagus and infer the probability of metastasis to a certain lymph node in each stage.

We calculate by Bayes' theorem with new patients the probability of what stage they are in.

We simulate data with our model and compare the results with those reported by Matsubara (1992), the second author.

2. Data

The data are the numbers of involved lymph nodes in 103 patients with cancer of the thoracic esophagus undergoing the systematic dissection of lymph nodes including cervical nodes from January 1985 through August 1990 at Cancer Institute Hospital. The patients underwent no pre-operative treatment whatever, such as radiation therapy. The number of lymph nodes is 45 and we selected 30 clinically important lymph nodes. The numbers of selected lymph nodes are $1 \sim 21$, $23 \sim 29$, 43, 45. The sites of these lymph nodes are as follows.

1. Deep cervical lymph nodes: 1, 2.
2. Para-tracheal lymph nodes: $3 \sim 7$, 12, 45.

Table 1. The frequencies of metastasis to each lymph node in 103 patients. The numbers in parentheses denote the number of patients who were examined.

No.*	Patients with metastasis**
1	7(88)
2	4(93)
3, 4	5(51)
5	41(103)
6, 7	24(90)
8	12(56)
9, 10, 11	20(103)
12	8(103)
13	25(103)
14	7(103)
15, 17	10(103)
16, 18	8(103)
19, 20, 21	9(93)
23, 24	27(103)
25	25(103)
26	15(103)
27	16(85)
28, 29	17(92)
43	5(21)
45	4(39)

*Lymph node number.
**The number of patients with metastasis.

3. Middle and lower mediastinal lymph nodes: $8 \sim 11$, $13 \sim 21$.
4. Upper gastric lymph nodes: $23 \sim 29$.
5. Abdominal left para-aortic lymph node: 43.

Note that the "lymph node" used in this paper denotes a group of adjoining lymph nodes.

Lymph nodes $(3,4)$, $(6,7)$, $(9,10,11)$, $(15,17)$, $(16,18)$, $(19,20,21)$, $(23,24)$, $(28,29)$ are clinically dealt with as one group respectively and hence the data were obtained from 20 lymph nodes, or more properly, 20 lymph node groups. We show the frequencies of metastasis to each lymph node in Table 1, where the numbers in parentheses denote the number of patients who were examined. Since all the lymph nodes were not examined with all the patients, the numbers in parentheses are not the same. In some cases, doctors may not examine some lymph nodes if they are almost sure that the lymph nodes are not involved. In such cases, we should deal with lymph nodes not examined as ones not involved. In this paper, we do not deal with the missing data in such a way because the reasons for the missing data are not clear.

3. Method

We suppose that the patients are in one of k stages. It is clinically reasonable to expect that the probability of metastasis to each lymph node increases as the stage advances. The advance of stage means, as it were, the increase of the load by cancer and the metastasis to lymph nodes can be regarded as the damage brought by this load.

Weibull distributions (Johnson and Kotz (1970), Nelson (1982)) are often adopted to express the relation between a load and its damage. We introduce distribution functions of Weibull distributions to express the relation between the advance of stage and the increase of the probability of metastasis.

The probability of metastasis to the j-th lymph node conditioned by *Stage* l is given by

$$P(j \mid l) = 1 - \exp[-(x_l/\alpha_j)^{\beta_j}], \qquad x_l > 0, \quad \alpha_j > 0, \quad \beta_j > 0, \tag{3.1}$$

where α_j and β_j $(j = 1, 2, \ldots, m)$ are the scale and shape parameters of a Weibull distribution respectively and x_l $(l = 1, 2, \ldots, k)$ is the value for *Stage* l. The probability of metastasis to each lymph node is expected to increase as the stage advances and we hence expect that $x_1 < x_2 < \cdots < x_k$.

We denote the data of the i-th patient $(i = 1, 2, \ldots, n)$ by

$$\boldsymbol{h}_i = (h_{i1}, h_{i2}, \ldots, h_{im})^T,$$
$$h_{ij} = \begin{cases} 0 & \text{not involved,} \\ 1 & \text{involved,} \\ \text{NA} & \text{not available,} \end{cases}$$

where $j = 1, 2, \ldots, m$, and NA means that the lymph node is not examined. Considering the grouping of the lymph nodes mentioned in the previous section, n is 103 and m is 20 for the present data.

In the present study, we assume that the data of each lymph node are mutually independent. On this assumption, the probability that the data of the i-th patient in *Stage* l_i show the pattern of $\boldsymbol{h}_i$ is given by

$$\prod_{\substack{j=1 \\ h_{ij}\neq \mathrm{NA}}}^{m} P(j \mid l_i)^{h_{ij}}(1-P(j \mid l_i))^{(1-h_{ij})}.$$

If $h_{ij} = \mathrm{NA}$, the data are not included in this calculation. We mention the missing data in the final section. It is impossible to know previously the stage of each patient. We introduce the probability of what stage each patient is in as a prior distribution into a Bayesian model. The probability that the i-th patient is in *Stage* l_i and his data show the pattern of $\boldsymbol{h}_i$ is given by

$$\prod_{\substack{j=1 \\ h_{ij}\neq \mathrm{NA}}}^{m} P(j \mid l_i)^{h_{ij}}(1-P(j \mid l_i))^{(1-h_{ij})}\omega_i(l_i),$$

where $\omega_i(l_i)$ denotes the probability that the i-th patient is in *Stage* l_i. If l_i were given, we could estimate $P(j \mid l_i)$ by the ordinary maximum likelihood method. Considering that we have no information on l_i, we use the marginal probability that the i-th patient has the pattern of $\boldsymbol{h}_i$ given by

$$\sum_{l=1}^{k}\left\{\prod_{\substack{j=1 \\ h_{ij}\neq \mathrm{NA}}}^{m} P(j \mid l)^{h_{ij}}(1-P(j \mid l))^{(1-h_{ij})}\right\}\omega_i(l).$$

Assuming that the data of each patient are mutually independent, the likelihood of our model is given by

$$\prod_{i=1}^{n}\left[\sum_{l=1}^{k}\left\{\prod_{\substack{j=1 \\ h_{ij}\neq \mathrm{NA}}}^{m} P(j \mid l)^{h_{ij}}(1-P(j \mid l))^{(1-h_{ij})}\right\}\omega_i(l)\right].$$

We further assume that the probability $\omega_i(l)$ is the same for all the patients, that is,

$$\omega_i(l) = \omega(l), \qquad (l = 1, 2, \ldots, k), \tag{3.2}$$

where $\sum_{l=1}^{k}\omega(l) = 1$. Thus, the log likelihood is

(3.3) log likelihood

$$= \sum_{i=1}^{n}\log\left[\sum_{l=1}^{k}\left\{\prod_{\substack{j=1 \\ h_{ij}\neq \mathrm{NA}}}^{m} P(j \mid l)^{h_{ij}}(1-P(j \mid l))^{(1-h_{ij})}\right\}\omega(l)\right].$$

The parameters of our model are $\omega(l)$, α_j, β_j and x_l. Note that we can freely take the scale x_l in (3.1) and that we need to fix one of x_l. We fix x_k, the value for the maximum stage. In the modeling, we can freely choose the number of stages k. However, it is clinically meaningless to take k too large. For the present study, we set k equal to 3 or 4. We discuss the selection of k in the final section. With these two values of k, we calculate AIC and select the more appropriate one. AIC (Akaike (1973), Sakamoto *et al.* (1986)) is given by

$$\begin{aligned}\text{AIC} = (-2) \times &(\text{maximum log likelihood of the model}) \\ &+ 2 \times (\text{number of free parameters of the model}).\end{aligned}$$

The AIC of our model is given by

$$\text{AIC} = (-2) \times \left(\sum_{i=1}^{n} \log \left[\sum_{l=1}^{k} \left\{ \prod_{\substack{j=1 \\ h_{ij} \neq \text{NA}}}^{m} \hat{P}(j \mid l)^{h_{ij}} (1 - \hat{P}(j \mid l))^{(1-h_{ij})} \right\} \hat{\omega}(l) \right] \right) + 2 \times q,$$

where $\hat{P}(j \mid l)$ and $\hat{\omega}(l)$ denote the estimated values by the maximum likelihood method. q denotes the number of free parameters, which is equal to $2m + 2(k-1)$.

With the estimated values, $\hat{P}(j \mid l)$ and $\hat{\omega}(l)$, we can calculate the probability $P(l \mid \boldsymbol{h}_s)$ that a patient with data $\boldsymbol{h}_s = (h_{s1}, h_{s2}, \ldots, h_{sm})^T$ is in *Stage* l. From Bayes' theorem, this probability is expressed by

$$P(l \mid \boldsymbol{h}_s) = \frac{P(l, \boldsymbol{h}_s)}{P(\boldsymbol{h}_s)}.$$

With $\hat{P}(j \mid l)$ and $\hat{\omega}(l)$, the probability is given by

$$P(l \mid \boldsymbol{h}_s) = \frac{\left\{ \prod_{\substack{j=1 \\ h_{sj} \neq \text{NA}}}^{m} \hat{P}(j \mid l)^{h_{sj}} (1 - \hat{P}(j \mid l))^{(1-h_{sj})} \right\} \hat{\omega}(l)}{C}, \tag{3.4}$$

where C is the normalizing factor given by

$$C = \sum_{l=1}^{k} \left\{ \prod_{\substack{j=1 \\ h_{sj} \neq \text{NA}}}^{m} \hat{P}(j \mid l)^{h_{sj}} (1 - \hat{P}(j \mid l))^{(1-h_{sj})} \right\} \hat{\omega}(l).$$

4. Result

We estimated the parameters by maximizing the log likelihood (3.3) numerically by *Davidon's method* (Davidon (1968), Ishiguro and Akaike (1989)) and calculated the values of AIC with the two values of k. AIC is 1325.37 for $k = 3$

and 1317.56 for $k = 4$. We select $k = 4$ and assume that the patients are in one of 4 stages. The estimates of x_l, α_j and β_j are listed in Appendix, where x_4 is set equal to 2.0.

We show the estimates of $\omega(l)$ in Table 2 and they indicate that about 80% of the patients are in *Stages* 1 and 2, and only a few patients are in *Stage* 4.

The values of $P(j \mid l)$ calculated from (3.1) with the estimated α_j, β_j and x_l are shown in Table 3, where <0.001 or >0.999 means that the probability is smaller than 0.001 or larger than 0.999 respectively. We plot $P(j \mid l)$ against *Stage* l in Fig. 1, where No. denotes the lymph node number. The results in Table 3 and Fig. 1 show the following characteristics.

• Compared with the other lymph nodes, the probability of metastasis to No. 5 lymph node is remarkably large in *Stage* 1.

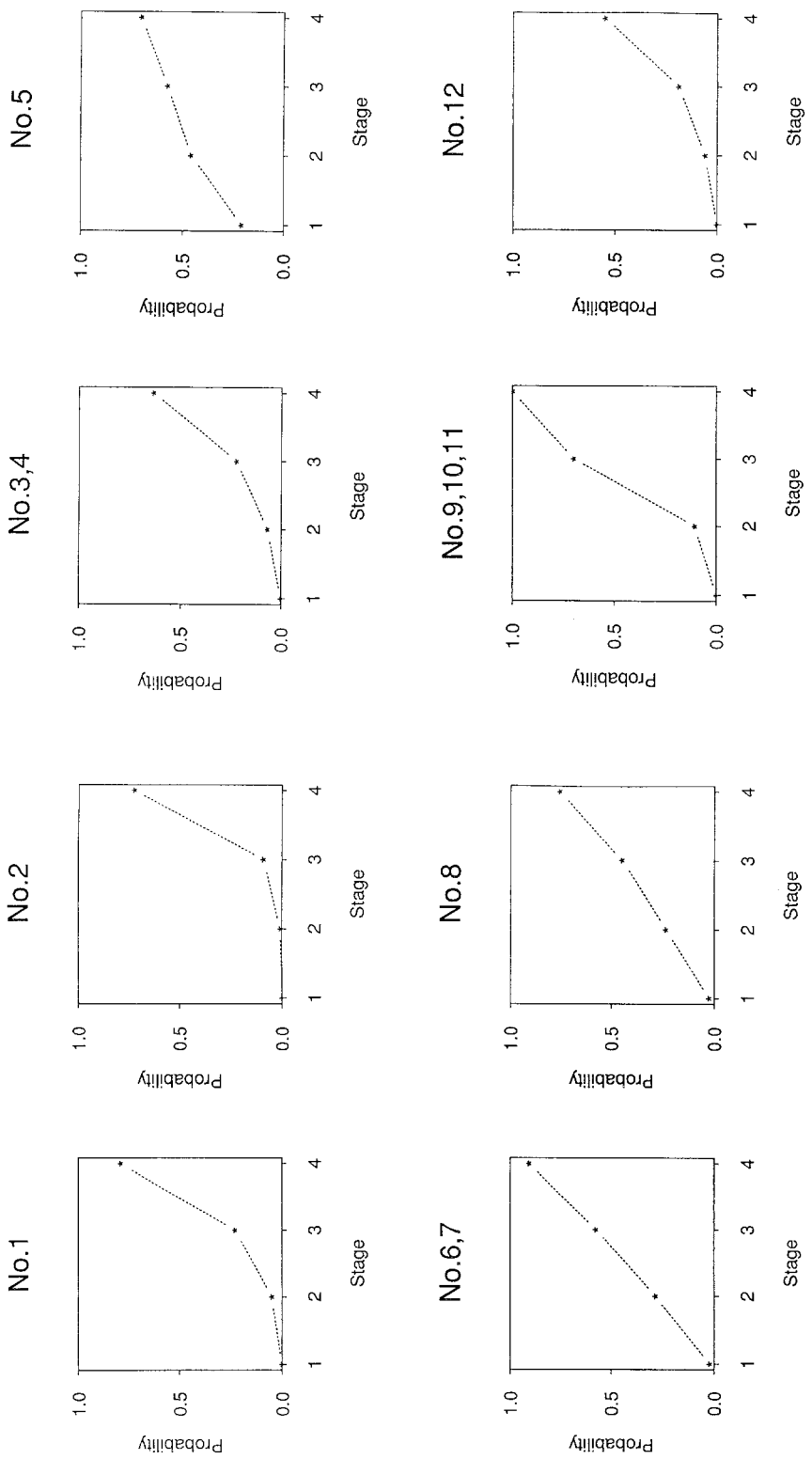
Table 2. The estimates of the prior probabilities $\omega(l)$.

Stage 1	Stage 2	Stage 3	Stage 4
0.344	0.459	0.167	0.030

Table 3. The conditional probabilities $P(j \mid l)$.

No.*	Stage 1	Stage 2	Stage 3	Stage 4
1	< 0.001	0.050	0.233	0.790
2	< 0.001	0.009	0.091	0.726
3, 4	0.002	0.070	0.226	0.633
5	0.211	0.460	0.574	0.704
6, 7	0.021	0.284	0.576	0.908
8	0.026	0.236	0.449	0.757
9, 10, 11	< 0.001	0.108	0.697	> 0.999
12	0.002	0.060	0.190	0.547
13	0.008	0.241	0.605	0.969
14	0.002	0.054	0.163	0.468
15, 17	0.001	0.065	0.276	0.828
16, 18	< 0.001	0.021	0.240	0.987
19, 20, 21	0.011	0.102	0.212	0.432
23, 24	0.032	0.299	0.554	0.862
25	0.002	0.214	0.692	0.999
26	< 0.001	0.086	0.469	0.995
27	0.004	0.163	0.489	0.942
28, 29	0.011	0.182	0.423	0.807
43	0.004	0.116	0.343	0.794
45	0.004	0.074	0.198	0.501

*Lymph node number.


No.1
No.2
No.3,4
No.5
No.6,7
No.8
No.9,10,11
No.12
Probability
Stage
0.0
0.5
1.0
1
2
3
4

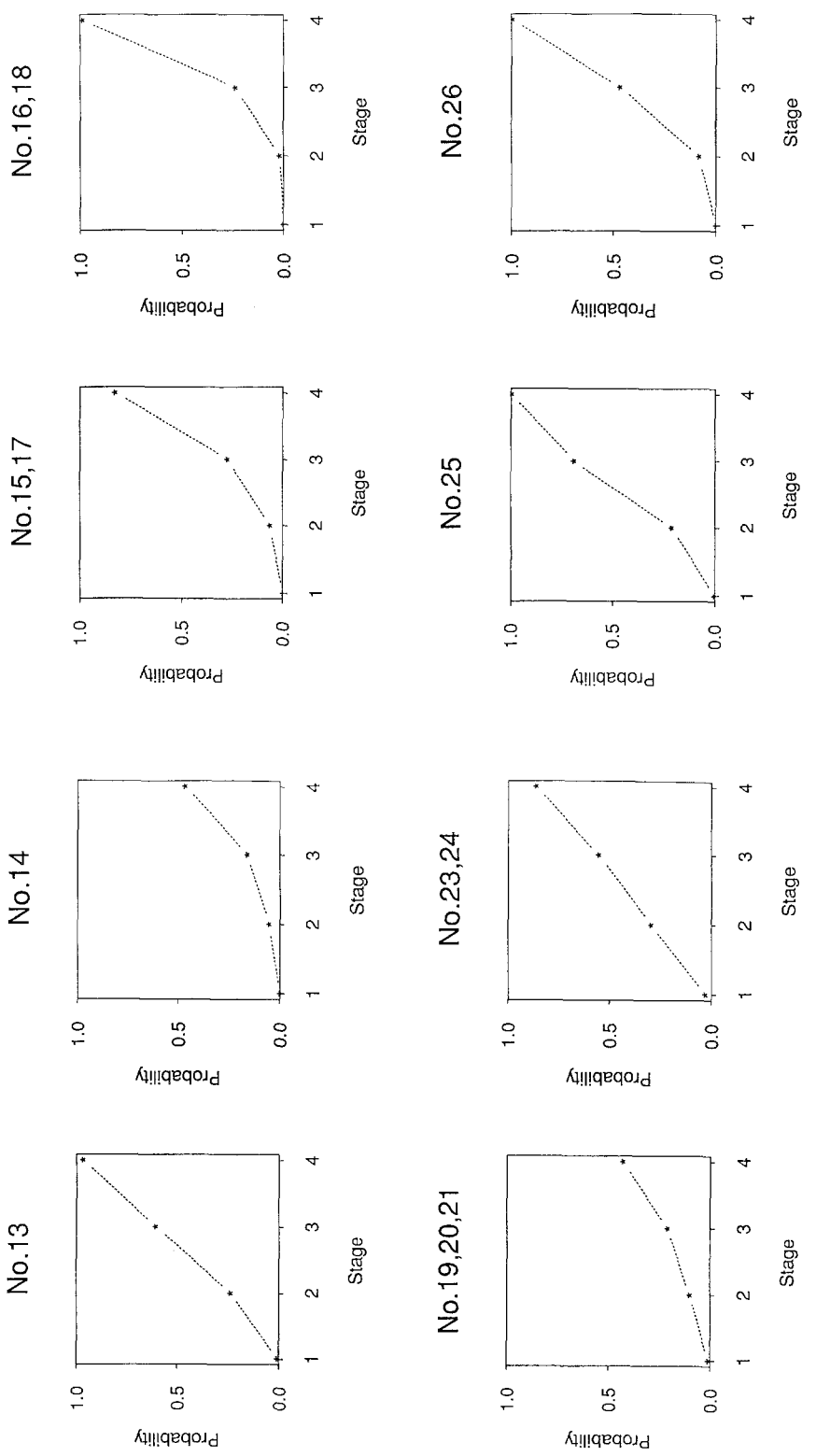


Fig. 1. Conditional probabilities of metastasis to each lymph node for given stages.

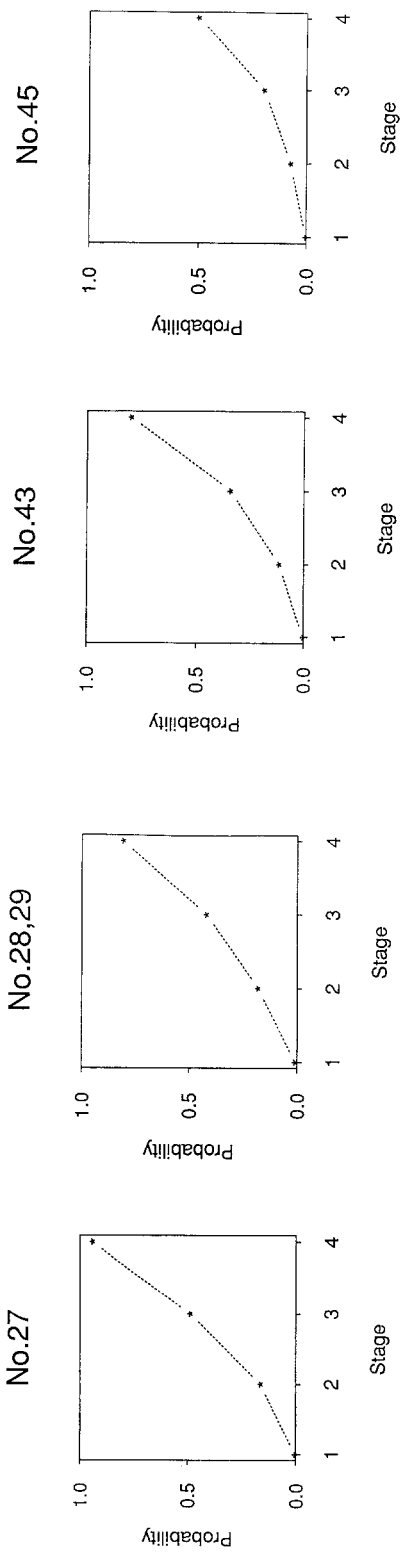


Fig. 1. (continued).

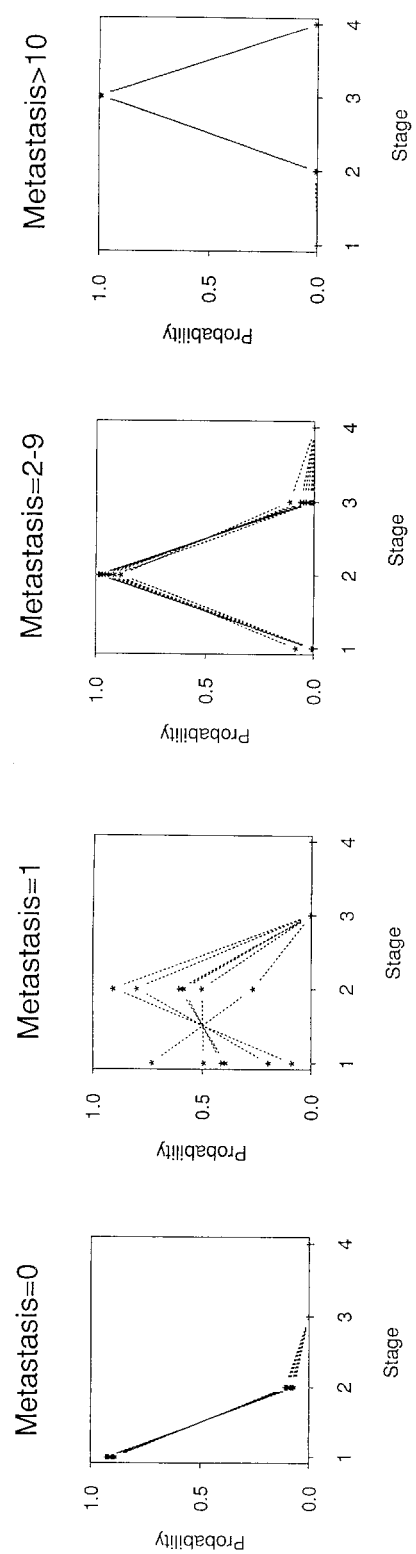


Fig. 2. Posterior probabilities for 31 new patients to be in each stage.

• Compared with No. (6, 7) lymph node, the probability of metastasis to No. 8 lymph node is almost the same in *Stage* 1 and the probabilities are smaller in the other stages.

• Compared with No. (23, 24) lymph node, the probabilities of metastasis to No. 25 lymph node are smaller in the earlier stages, such as *Stages* 1 and 2. In the advanced stages, such as *Stages* 3 and 4, however, the probabilities of metastasis to No. 25 lymph node become larger.

Table 4. The posterior probabilities for 31 new patients to be in each stage.

No.*	Metastasis**	Stage 1	Stage 2	Stage 3	Stage 4
1	0	0.903	0.097	< 0.001	< 0.001
2	0	0.922	0.078	< 0.001	< 0.001
3	0	0.897	0.103	< 0.001	< 0.001
4	0	0.922	0.078	< 0.001	< 0.001
5	0	0.930	0.070	< 0.001	< 0.001
6	0	0.917	0.083	< 0.001	< 0.001
7	0	0.896	0.104	< 0.001	< 0.001
8	0	0.896	0.104	< 0.001	< 0.001
9	0	0.902	0.098	< 0.001	< 0.001
10	0	0.891	0.109	< 0.001	< 0.001
11	0	0.922	0.078	< 0.001	< 0.001
12	0	0.917	0.083	< 0.001	< 0.001
13	1	0.495	0.504	< 0.001	< 0.001
14	1	0.196	0.803	0.001	< 0.001
15	1	0.394	0.605	0.001	< 0.001
16	1	0.412	0.587	0.001	< 0.001
17	1	0.732	0.268	< 0.001	< 0.001
18	1	0.088	0.909	0.003	< 0.001
19	2	0.001	0.981	0.018	< 0.001
20	2	0.005	0.987	0.009	< 0.001
21	2	0.008	0.984	0.007	< 0.001
22	3	< 0.001	0.967	0.033	< 0.001
23	3	0.081	0.916	0.003	< 0.001
24	4	< 0.001	0.887	0.113	< 0.001
25	4	< 0.001	0.938	0.062	< 0.001
26	7	< 0.001	0.953	0.047	< 0.001
27	8	0.006	0.980	0.014	< 0.001
28	11	< 0.001	0.003	0.997	< 0.001
29	18	< 0.001	< 0.001	0.990	0.010
30	18	< 0.001	0.008	0.992	< 0.001
31	23	< 0.001	< 0.001	> 0.999	< 0.001

*Patient number.

**The number of involved lymph nodes.

The differences in the profiles of probabilities between No. (6, 7) and No. 8, and between No. (23, 24) and No. 25 are of clinical interest. The above observations coincide with our clinical impression.

We calculated from (3.4) the values of the posterior probabilities for 31 new patients undergoing the same therapy to be in a certain stage. We show the values of the probabilities in Table 4. We divide the patients into 4 groups as follows.

Group 1: Patients with no involved lymph node.

Group 2: Patients with 1 involved lymph node.

Group 3: Patients with 2 to 9 involved lymph nodes.

Group 4: Patients with more than 10 involved lymph nodes.

For each group, we plot the values of the probabilities against *Stage* l in Fig. 2, where *Metastasis* means the number of involved lymph nodes. The graphic outputs show the following points.

- In *Groups* 1, 3 and 4, all the patients in each group show almost the same pattern of the posterior probabilities.
- The probability is the largest for *Group* 1 at *Stage* 1, for *Group* 3 at *Stage* 2 and for *Group* 4 at *Stage* 3.
- The patients in *Group* 2 show various patterns of the posterior probabilities. It is clinically natural that the probabilities for the advanced stages, such as *Stages* 3 and 4, become larger as the number of involved lymph nodes increases.

The last point mentioned above suggests that the patients with the same number of involved lymph nodes might be in different stages and that the number of involved lymph nodes does not give enough information. We discuss this point in the next section.

5. Discussion

In this paper, we proposed a method to infer the stage of cancer from lymphatic spreading patterns by applying a Bayesian model and analyzed the data of nodal involvement obtained from the patients with cancer of the thoracic esophagus.

In Fig. 2, the patients with 1 involved lymph node show various patterns of the posterior probabilities. Assuming that the patients with 1 involved lymph node in Table 4 are in the stage where their probability is the largest, we find the following points.

- No. 17 patient alone is in *Stage* 1, who has metastasis to No. 5 lymph node. The other patients with 1 involved lymph node are in *Stage* 2. No. 5 lymph node is remarkably involved in *Stage* 1 and we hence expect that the patients with metastasis to this lymph node only are in *Stage* 1.
- No. 18 patient is in *Stage* 2 with the probability larger than 0.9 and No. 13 patient is in *Stage* 2 with the probability almost equal to 0.5. No. 18 and No. 13 patients have metastasis to No. 25 and No. (23, 24) lymph nodes respectively. As mentioned in *Result*, compared with No. (23, 24) lymph node, the probabilities of metastasis to No. 25 lymph node are smaller in the earlier stages, such as *Stages* 1 and 2. In the advanced stages, such as *Stages* 3 and 4, however, the probabilities of metastasis to No. 25 lymph node become larger. This is the reason for the difference between No. 13 and No. 18 patients.

• In No. 15 and No. 16 patients, the probabilities that they are in *Stage* 2 are not remarkably large. The reason is that they have metastasis to No. (6, 7) lymph node which is apt to be involved in the earlier stages compared with the other lymph nodes, except for No. 5 lymph node.

These points suggest that the site of involved lymph node is important for the inference of the stage of cancer in the patients with 1 involved lymph node.

The patients with no involved lymph node have different patterns of lymph nodes not examined. Thus, they have the different profiles of probabilities in Table 4.

We have already discussed the missing data in the section of *Data*. No. 43 and No. 45 lymph nodes' data include particularly many missing data and hence they need careful analysis.

We calculated from (3.4) the values of the posterior probabilities for the 103 original patients to be in a certain stage. We plot the values of the probabilities in Fig. 3. Assuming that the patients are in the stage where their probability is the largest, we find the following points.

• All of the 27 patients with no involved lymph node are in *Stage* 1.

• In the 19 patients with 1 involved lymph node, 11 patients are in *Stage* 1 and all of them have metastasis to No. 5 lymph node. The other 8 patients have metastasis to one of No. (6, 7), No. (23, 24), No. 25, No. 26 or No. (28, 29) lymph nodes. In the 5 patients with metastasis to No. (6, 7) or No. (23, 24) lymph nodes, the values of the probabilities that they are in *Stage* 2 are not remarkably large.

• With the patients with no or 1 involved lymph node, the results are consistent with those of the 31 new patients.

• The 25 patients with 2 to 4 involved lymph nodes are in *Stage* 2 except for one patient.

• The 17 patients with 5 to 9 involved lymph nodes show various patterns of the posterior probabilities.

• The 12 patients with 10 to 24 involved lymph nodes are in *Stage* 3 except for one patient.

• All of the 3 patients with more than 25 involved lymph nodes are in *Stage* 4.

The patients with the same number of involved lymph nodes show various patterns of the posterior probabilities and they are not necessarily in the same stage. This result suggests that the number of involved lymph nodes does not give enough information and that we should infer the stage of each patient from his lymphatic spreading pattern. Our proposed method makes such an inference possible and this point is one feature of our method.

Matsubara analyzed the data obtained from 110 patients undergoing the same therapy (Matsubara (1992)). His data are about our selected lymph nodes and include our 103 data. He divided the patients into 3 groups by the number of involved lymph nodes N as follows.

Group (*Slight*): $N = 1$.

Group (*Mild*): $2 \leq N \leq 5$.

Group (*Severe*): $6 \leq N$.

He excluded the patients with no involved lymph node from this grouping. We

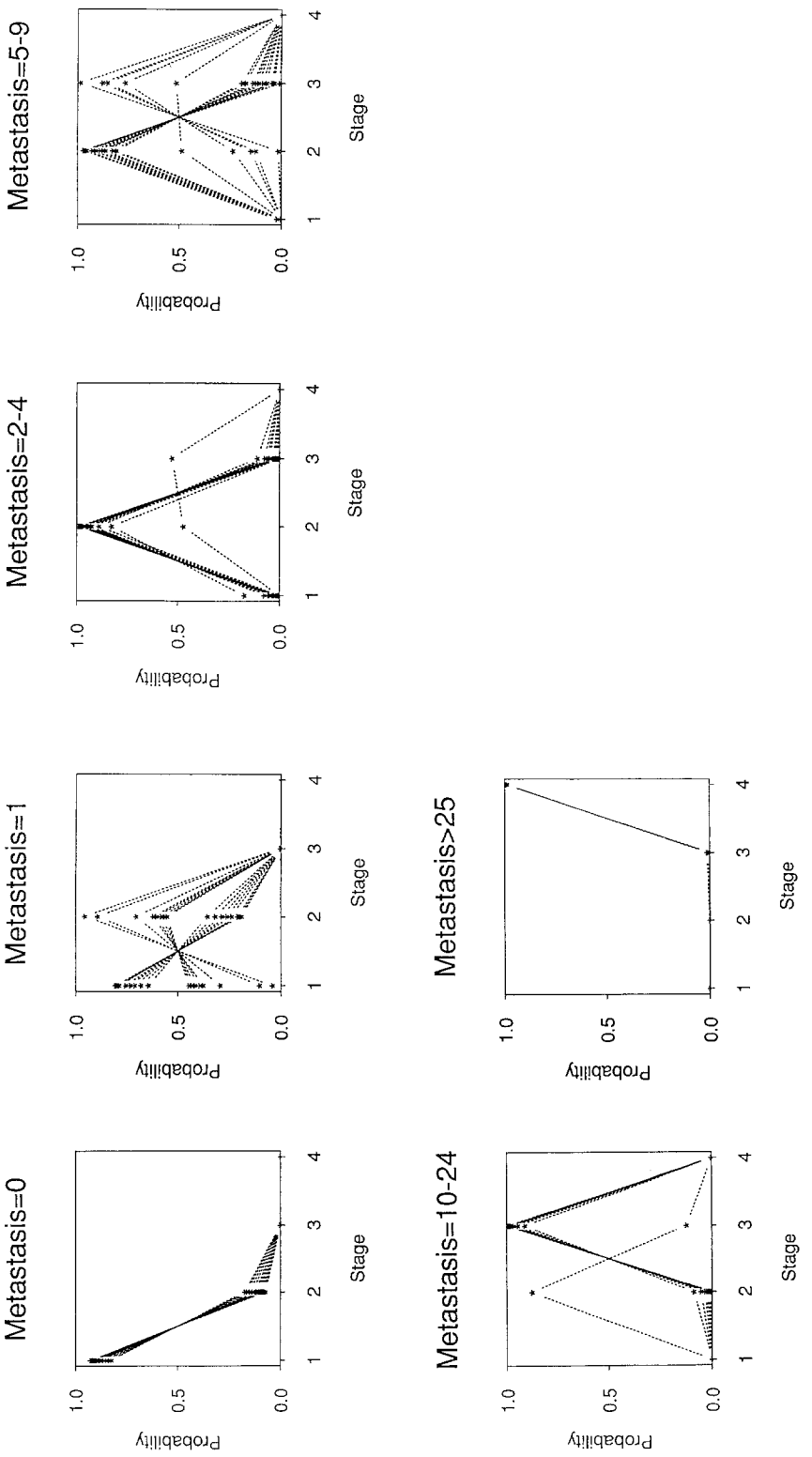


Fig. 3. Posterior probabilities for 103 patients to be in each stage.

show his results in Table 5, where the numbers in parentheses denote the number of patients used for calculating the proportions. Since all the lymph nodes were not examined with all the patients, the numbers in parentheses are not the same. In Table 5, the proportion of the patients with metastasis to No. 5 lymph node is smaller in *Mild* than in *Slight*. This result is not reasonable because we expect that the patients in *Mild* are in more advanced stages than those in *Slight* and that the proportion of the patients with metastasis to No. 5 lymph node in *Mild* is larger than that in *Slight*. Compared with the other lymph nodes, the probability of metastasis to No. 5 lymph node is remarkably large in *Stage* 1 as shown in Table 3, and hence it is natural that in the patients with 1 involved lymph node, the proportion of the patients with metastasis to No. 5 lymph node should be particularly large. This is the reason for the seemingly unreasonable results with No. 5 lymph node shown in Table 2.

Table 5. The proportions of patients with metastasis to each lymph node in each group. The numbers in parentheses denote the number of patients used for calculating the proportions. (Reported by Matsubara ((1992), Table 4).)

No.*	Slight	Mild	Severe	Total
1	0(18)	0.069(29)	0.28(25)	0.095(95)
2	0(19)	0(31)	0.154(26)	0.040(101)
3, 4	0(8)	0(11)	0.267(15)	0.095(42)
5	0.55(20)	0.333(33)	0.733(30)	0.4(110)
6, 7	0.118(17)	0.313(32)	0.615(26)	0.286(98)
8	0(12)	0.167(18)	0.471(17)	0.186(59)
9, 10, 11	0(20)	0.182(33)	0.5(30)	0.191(110)
12	0(20)	0(33)	0.233(30)	0.064(110)
13	0(20)	0.242(33)	0.6(30)	0.236(110)
14	0(20)	0.030(33)	0.2(30)	0.064(110)
15, 17	0(20)	0.152(33)	0.267(30)	0.118(110)
16, 18	0(20)	0.030(33)	0.233(30)	0.073(110)
19, 20, 21	0(19)	0.067(30)	0.269(26)	0.09(100)
23, 24	0.2(20)	0.303(33)	0.533(30)	0.273(110)
25	0.05(20)	0.242(33)	0.533(30)	0.227(110)
26	0.05(20)	0.091(33)	0.4(30)	0.145(110)
27	0(20)	0.212(33)	0.4(30)	0.173(110)
28, 29	0.053(19)	0.25(28)	0.385(26)	0.184(98)
43	0(5)	0(7)	0.556(9)	0.238(21)
45	0(10)	0.111(9)	0.333(15)	0.146(41)

*Lymph node number.

To reproduce this seemingly unreasonable results, we simulated 110 data with the parameters in Appendix and the probabilities in Tables 2 and 3, and divided the simulated data into 3 groups in the same way as in Table 5. We show the

results in Table 6, where the proportion of the patients with metastasis to No. 5 lymph node is smaller in *Mild* than in *Slight*. Our results in Table 3 and Fig. 1 have no such unreasonable results.

Compared with Table 5, the results in Table 6 are slightly different in some lymph nodes. We think, however, that the findings of Table 6 express the features of Table 5 well enough as a whole.

Table 6. The proportions of patients with metastasis to each lymph node in each group. (110 simulated data.)‡

No.*	Slight(17)**	Mild(49)**	Severe(26)**	Total
1	0	0	0.269	0.064
2	0	0.020	0.153	0.045
3, 4	0	0.061	0.192	0.073
5	0.529	0.449	0.692	0.445
6, 7	0.118	0.286	0.692	0.309
8	0.059	0.082	0.5	0.164
9, 10, 11	0	0.142	0.462	0.173
12	0.059	0.061	0.231	0.091
13	0	0.204	0.615	0.236
14	0	0.020	0.231	0.064
15, 17	0	0.041	0.423	0.118
16, 18	0	0.061	0.231	0.082
19, 20, 21	0	0.082	0.231	0.091
23, 24	0.118	0.245	0.580	0.264
25	0	0.224	0.692	0.264
26	0	0.061	0.462	0.136
27	0	0.245	0.5	0.227
28, 29	0	0.143	0.615	0.209
43	0	0.184	0.423	0.182
45	0.059	0.061	0.231	0.091

*Lymph node number.

**The number of patients in each group.

‡The data were simulated with the parameters in Appendix and the probabilities in Tables 2 and 3.

We assumed that the stages were discrete. Strictly speaking, the stages of patients are continuous and not discrete. We can introduce a certain continuous distribution for the stages of patients. This approach is, however, not so practicable from the clinical point of view and hence, we did not take this approach.

In this paper, we statistically chose the number of stages, which should be selected clinically. The AICs suggested that the present data had enough information to be divided into 4 stages. Assume that the clinically reasonable number of stages is equal to 4 and the model with 3 stages has the smaller AIC. This means

that the data do not have enough information from the point of the information criterion and that we need more data for the division into 4 stages. Thus, we think that our approach with AIC can check that the data have enough information to be divided into the stages of a clinically reasonable number. The definition of the clinically reasonable number of stages is left for the future study. With the defined number of stages, we can use AIC to search some structures of the parameters α_j and β_j based on lymphatic spreading patterns. In this approach, we can introduce other factors, such as the local extension and the site of tumor, through the structures.

Our model with Weibull distributions might be alternated with the logistic regression model, where we may easily introduce the other factors mentioned above. By comparing the AICs of the two models, one with Weibull distributions and the other with the logistic regression, we can choose the more appropriate one.

Although we assumed that the data of each lymph node were mutually independent, the processes of metastasis to each lymph node have some correlation to each other. A number of methods for analyzing correlated binary data have been already developed (Ashby *et al.* (1992)). In our approach, the correlation can be introduced through assuming some correlations among the parameters α_j and β_j in building models. Such a modeling is left for the future study.

As already mentioned, the "lymph node" used in this paper denotes a group of adjoining lymph nodes. In the present study, we did not take into account the number of lymph nodes in each "lymph node". It is left for the future study to introduce this number of lymph nodes into models.

Remark. We can introduce other factors, such as the local extension and the site of tumor, in several ways. A simple way is to apply our approach to the lymphatic data stratified based on these factors. Another way is to apply the same type of model as (3.1) also to these factors if their data are dichotomous as our lymphatic data. Even if the data are graded more finely than the simple presence-absence dichotomy, our Bayesian approach can be applied by building some models based on the stage substituting for the model (3.1). The modeling for more than dichotomous data is left for the future study. In any approach, we may need a device to reduce the number of parameters to be estimated.

Acknowledgements

The authors are grateful to the referees and the editors for their helpful suggestions.

Appendix

1. The estimates of x_l of each stage.

Stage 1	Stage 2	Stage 3	Stage 4
0.022	0.303	0.749	2.000

2. The estimates of α_j and β_j of each lymph node.

No.*	α_j	β_j	No.*	α_j	β_j
1	1.56	1.81	15, 17	1.44	1.73
2	1.81	2.65	16, 18	1.19	2.82
3, 4	2.00	1.39	19, 20, 21	3.83	0.88
5	1.16	0.36	23, 24	0.95	0.91
6, 7	0.87	1.04	25	0.68	1.76
8	1.35	0.88	26	0.93	2.16
9, 10, 11	0.70	2.60	27	0.98	1.47
12	2.38	1.35	28, 29	1.28	1.11
13	0.79	1.34	43	1.43	1.35
14	2.86	1.29	45	2.73	1.17

*Lymph node number.

REFERENCES

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, *2nd Inter. Symp. on Information Theory* (eds. B. N. Petrov and F. Csaki), 267–281, Akademiai Kiado, Budapest. (Reproduced in *Breakthroughs in Statistics*, Volume 1 (eds. S. Kotz and N. L. Johnson), Springer, New York (1992).)

Ashby, M., Neuhaus, J. M., Hauck, W. W., Bacchetti, P., Heilbron, D. C., Jewell, N. P., Segal, M. R. and Fusaro, R. E. (1992). An annotated bibliography of methods for analysing correlated categorical data, *Statistics in Medicine*, **11**, 67–99.

Davidon, W. C. (1968). Variance algorithm for minimization, *Comput. J.*, **10**, 406–410.

Ishiguro, M. and Akaike, H. (1989). DALL: Davidon's algorithm for log likelihood maximization —A FORTRAN subroutine for statistical model builders, *Comput. Sci. Monographs*, No. 25, The Institute of Statistical Mathematics, Tokyo.

Johnson, N. L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions-1*, 250–266, Wiley, New York.

Matsubara, T. (1992). Pattern of lymphatic spreading in cancer of the thoracic esophagus; Analysis in cases undergoing cervical dissection, *Journal of Japan Surgical Society*, **93**(4), 377–387 (in Japanese).

Nelson, W. (1982). *Applied Life Data Analysis*, 36–39, Wiley, New York.

Sakamoto, Y., Ishiguro, M. and Kitagawa, G. (1986). *Akaike Information Criterion Statistics*, D. Reidel, Dordrecht, Holland.