GASTROENTEROLOGY 1993;105:1865–1876

Updating Prognosis in Primary Biliary Cirrhosis Using a Time-Dependent Cox Regression Model

ERIK CHRISTENSEN,* DOUGLAS G. ALTMAN,‡ JAMES NEUBERGER,§ BIANCA L. DE STAVOLA,‡ NIELS TYGSTRUP,¶ ROGER WILLIAMS,# and the PBC1 and PBC2 TRIAL GROUPS

*Department of Medical Gastroenterology, Hvidovre University Hospital; *Department of Medicine B, Bispebjerg University Hospital, Copenhagen, Denmark; ¶Medical Department A, Rigshospitalet, Copenhagen, Denmark; ‡Imperial Cancer Research Fund, London, England; #Liver Unit, King's College Hospital, London, England; and §Liver Unit, Queen Elizabeth Hospital, Birmingham, England

Background: **The precision of current prognostic models in primary biliary cirrhosis (PBC) is rather low, partly because they are based on data from just one time during the course of the disease. The aim of this study was to design a new, more precise prognostic model by incorporating follow-up data in the development of the model.** ***Methods:*** **We have performed Cox regression analyses with time-dependent variables in 237 PBC patients followed up regularly for up to 11 years. The validity of the obtained models was tested by comparing predicted and observed survival in 147 independent PBC patients followed for up to 6 years.** ***Results:*** **In the obtained model the following time-dependent variables independently indicated a poor prognosis: high bilirubin, low albumin, ascites, gastrointestinal bleeding, and old age. When including histological variables, cirrhosis, central cholestasis, and low immunoglobulin (Ig)M also indicated a poor prognosis. The survival predicted by the models agreed well with the survival observed in the independent PBC patients. The time-dependent models predicted better than our previously published time-fixed model.** ***Conclusions:*** **Using the time-dependent Cox models, one can estimate a more precise probability of surviving the next 1, 3, or 6 months for any given patient at any time during the course of the disease. This may improve monitoring of PBC patients.**

In primary biliary cirrhosis (PBC), several prognostic models have been developed using Cox regression analysis.[1–5] They are all so-called time-fixed models developed from the relation between survival and patient characteristics at just one fixed time during the course of the disease (usually the time of diagnosis or the time of inclusion in a controlled clinical trial). Although application of the time-fixed Cox models in new patients has shown reasonable agreement between predicted and observed survival for groups of patients having a good, intermediate, or poor prognosis,[2–3] the prognostic estimates for individual patients are not very precise.[2–3,5–6] Furthermore, because the time-fixed models are not based on follow-up data, it is not appropriate to use such models on follow-up data to obtain an updated prognosis, although they may be misused for that purpose.

We have previously published a time-dependent prognostic model[7] for patients affected by cirrhosis of mixed etiology (alcoholic, posthepatitic, and cryptogenic). The time-dependent model uses the follow-up data to estimate the effect of the evolution of the variables over time. Therefore, it can provide updated short-term prognostic estimates during the course of the disease. This can improve follow-up monitoring of patients by translating a changed condition into a changed prognosis and provide a better basis for decisions about changes in therapy including liver transplantation, which is particularly relevant for patients with PBC because current medical therapy[8–10] cannot stop progression of the disease or the occurrence of complications. To improve our understanding of the development of the disease and thus the timing of transplantation,[6,11] better prognostic models adapted to the follow-up situation become important.

This paper presents a time-dependent Cox model developed from the data of patients with PBC in a large controlled clinical trial of azathioprine vs. placebo[2] and tested on independent patients from another similar controlled clinical trial of d-penicillamine vs. placebo.[12]

Patients and Methods

Model Data

The data of 248 patients with PBC included in a previously published randomized clinical trial (PBC1 Trial) of azathioprine (1 mg/kg body wt) vs. placebo (respectively 127 and 121 patients) have been used in this study.[2] The

Abbreviations used in this paper: PI(t), time-dependent prognostic index; PBC, primary biliary cirrhosis.

0016-5085/93/$3.00

patients were assessed clinically and biochemically at the time of admission to the trial and every 6 months thereafter.[2] The total number of sets of clinical and biochemical recordings was 1590. For the key variables, the percentage of missing data was as follows: bilirubin, 5.0%; albumin, 7.4%; age, 0%; immunoglobulin (Ig) M, 18%; ascites, 5.6%; and gastrointestinal bleeding, 5.5%. Liver biopsy was performed before entry into the trial and thereafter at yearly intervals.[2] The total number of sets of biopsy data was 706. For the key variables, the percentage of missing data was as follows: cirrhosis, 4.2%; central cholestasis, 4.2%. The patients were followed for up to 11 years (median 2.9 years for those who died and 4.6 years for those censored). During that period 121 died (57 during azathioprine treatment and 64 during placebo treatment). The main cause of death was hepatic (mainly liver failure and gastrointestinal bleeding) in 98 patients (azathioprine 46, placebo 52). Two patients had a liver transplant, and both died from the procedure. Their deaths are included as endpoints and regarded as hepatic deaths. The above numbers differ slightly from those previously published[2] because some late incoming data have now been included. The cumulative survival curves have been published previously.[2]

Test Data

The independent data for testing the predictive power of the developed time-dependent Cox model were obtained from 189 patients with PBC followed up for up to 6 years (median 2.2 years for those who died and 3.6 years for those censored) and included a randomized clinical trial (PBC2 trial) of D-penicillamine, 1200 mg daily, vs. placebo in 98 and 91 patients, respectively.[12] This trial showed no significant effect of d-penicillamine.[12] The PBC2 trial was conducted in the same manner as the PBC1 trial with the same follow-up scheme. During the follow-up, 67 patients died (32 in the d-penicillamine group and 35 in the placebo group). The cumulative survival curves have been published previously.[12]

Statistical Analysis

The association between patient data (admission and follow-up data) and death risk (hazard) was analyzed using the Cox model for time-dependent variables[13]:

$$PI(t) = b_1z_1(t) + \cdots + b_qz_q(t) \quad \textbf{(Equation 1)}.$$

The model states that at a given time t the prognostic index ($PI(t)$) of a patient with the q variables $z_1(t)$ to $z_q(t)$ at that time is a function of these variables weighted by the corresponding regression coefficients b_1 to b_q. The time-dependent model uses the current value of each variable. Because each variable can vary over time, the death risk estimated by the model varies accordingly. If, for example, a patient develops gastrointestinal bleeding, the risk is likely to increase; if the bleeding can be effectively treated, the risk is likely to decrease again. The analysis requires that values for all variables are defined for each patient over his complete follow-up time. We have assumed that the values observed on a patient at a given time remain unchanged until the next observation because this corresponds to the clinical situation.

Probability of Surviving a Given Time

The PI(t) can be estimated repeatedly and used to assess prognosis during the course of the disease in a given patient. Although the future development in the variables of a patient is not known in advance, one can nevertheless use their current value at time t to compute a heuristic probability of surviving the following time interval h as

$$P(t, t + h) = \exp(-\lambda_0 \cdot h \cdot \exp[PI(t)]) \quad \textbf{(Equation 2)}$$

provided that h is short relatively to the course of the disease and λ_0, the estimated baseline hazard rate, is constant[7,13–14] (see appendix for details). It is reasonable to assume that prognostic variables do not change greatly over a 6-month period. Therefore we have used equation 2 for h up to 6 months, but not beyond, to assess the properties of the time-dependent model.

Derivation of Time-Dependent Cox Model

Variables included in the previously published time-fixed Cox regression model for the PBC1 data (bilirubin, age, albumin, cirrhosis, central cholestasis, and therapy [azathioprine or placebo]) were considered, together with other variables that, in univariate analyses, had been shown to have the following prognostic information: jaundice, pruritus, gastrointestinal hemorrhage, cholestyramine treatment, diuretic treatment, pigmentation, ascites, alkaline phosphatase, alanine aminotransferase, cholesterol, IgA, IgG, IgM, bile ducts present, granulomas, bile duct destruction, ductules, necrosis, fibrosis, and peripheral cholestasis. Univariate time-dependent Cox analyses were performed for each of these variables.

All variables were then taken as candidates in multiple time-dependent Cox regression analyses to derive predictive models. Stepwise selection was used with $P < 0.05$ being the significance level for a variable to be included in the model. The same set of variables was selected whether forward selection or backward elimination procedures[15] were used. Interactions between treatment (azathioprine or placebo) and other variables were investigated in the same manner as described previously.[16–17] The scoring of the variables was adapted to fulfill model assumptions.[15]

We fitted two models to the data, one including clinical and biochemical variables (A) and the other also including biopsy variables (B). In addition, we performed the analyses considering all deaths as events (I) or considering only deaths from hepatic cause (liver failure, gastrointestinal bleeding, or both and hepatoma) as events while deaths from other causes were censored (II). The latter set of analyses was performed to make the model more clinically rele-

vant to the transplantation situation, because liver transplantation can only be justified to prevent death from liver disease. Thus four models corresponding to all combinations of the above criteria (A-I, A-II, B-I and B-II) were fitted.

Validation of Model

The goodness of fit of the time-dependent Cox model to the data was assessed by a number of methods described in the Appendix. The model's predictive power was tested in 147 independent PBC patients (with complete data) who were included in the randomized clinical trial of D-penicillamine vs. placebo (PBC2)[12] (see Appendix).

Furthermore, the heuristic predictions derived from the time-dependent model were compared with those of our previously developed time-fixed model[2] (see Appendix).

Results

Univariate Time-Dependent Analyses

Each of the following time-varying variables had significant association (i.e., $P < 0.05$) with shorter survival in univariate time-dependent analysis: high current age, ascites, high bilirubin, low albumin, jaundice, treatment with diuretics, gastrointestinal bleeding, pigmentation, high alanine aminotransferase, high IgA, central cholestasis, cirrhosis, peripheral cholestasis, no (noncirrhotic) fibrosis, no granulomas, no ductules.

Model Based on Clinical and Biochemical Variables

The final time-dependent Cox model based on clinical and biochemical variables only is shown in Table 1. It is based on 237 patients on whom entry values of the variables in the model were available, and it shows that high serum bilirubin, ascites, low serum albumin, old age, and gastrointestinal bleeding were each associated with a poor prognosis. The same variables were selected when only the hepatic deaths (A-II) were taken as end-points. The regression coefficients for bilirubin, ascites, and gastrointestinal bleeding were somewhat bigger than in the model with all deaths as end-points (A-I), although the differences were rather small. Closest to significance for inclusion into either model was alanine aminotransferase ($P = 0.05$ and $P = 0.06$, respectively).

Model Based on Clinical, Biochemical, and Histological Variables

The final time-dependent Cox model selected from the full set of variables (including histology) is shown in Table 2. The estimates were based on 191 patients on whom entry values of all variables in the model were available. The variables selected included all those in the previous model with the addition of low IgM, cirrhosis, and central cholestasis. When only hepatic deaths were considered as end-points, the same variables were again selected, although the regression coefficients for bilirubin, ascites, and gastrointestinal bleeding were slightly bigger. The coefficients for the other variables were nearly the same in the two models. No other variable was close to inclusion in these models.

Effect of Azathioprine

In the time-dependent models (Tables 1 and 2) the time-fixed variable treatment (azathioprine or placebo) was not significant. Separate graphs of the data for the patients in the two treatment groups suggested a noticeable influence of therapy on bilirubin and albumin (Figures 1 and 2). For bilirubin, the initial drop (regression toward the mean) was greater and the sub-

Table 1. Final Time-Dependent Cox Regression Model Including Only Clinical and Biochemical Variables in 237 Patients With PBC

		Model A-I (all deaths)[a]			Model A-II (hepatic deaths)[b]		
Variable	Scoring	Regression coefficient	SE	P	Regression coefficient	SE	P
Serum bilirubin	$\log_{10}$ (value in μmol/L) − 1.53	2.53	0.26	<0.001	3.02	0.31	<0.001
Ascites	Absent: 0 Present: 1	1.39	0.21	<0.001	1.43	0.24	<0.001
Serum albumin	Value in g/L − 34.3	−0.085	0.019	<0.001	−0.077	0.021	<0.001
Age	Years − 55	0.040	0.011	<0.001	0.043	0.012	<0.001
Gastrointestinal bleeding	Absent: 0 Present: 1	0.65	0.21	0.001	0.74	0.24	0.002

[a]No. of deaths = 116.
[b]No. of deaths = 94.

Table 2. Final Time-Dependent Cox Regression Model Including Clinical, Biochemical, and Liver Biopsy Variables in 191 Patients With PBC

		Model B-I (all deaths)[a]			Model B-II (hepatic deaths)[b]		
Variable	Scoring	Regression coefficient	SE	P	Regression coefficient	SE	P
Serum bilirubin	$\log_{10}$ (value in μmol/L) − 1.53	2.26	0.32	<0.001	2.79	0.39	<0.001
Central cholestasis	Absent: 0 Present: 1	1.22	0.25	<0.001	1.29	0.29	<0.001
Ascites	Absent: 0 Present: 1	1.18	0.25	<0.001	1.30	0.29	<0.001
Cirrhosis	Absent: 0 Present: 1	0.87	0.23	<0.001	0.88	0.26	<0.001
Gastrointestinal bleeding	Absent: 0 Present: 1	0.95	0.25	<0.001	1.10	0.29	<0.001
Serum albumin	Value in g/L − 34.3	−0.070	0.022	0.001	−0.062	0.024	0.01
Serum IgM	$\log_{10}$ (value in g/L) − 0.47	−0.90	0.33	0.006	−0.85	0.39	0.03
Age	Years − 55	0.027	0.012	0.02	0.026	0.014	0.06

[a]No. of deaths = 96.
[b]No. of deaths = 77.

sequent rate of increase less marked for the azathioprine group (Figure 1*B*) than for placebo treated patients (Figure 1*A*) when overlapping groups with complete data in each time span were considered.[17–18] For albumin, the levels in the placebo group were, despite some short-term fluctuation, relatively constant over time (Figure 2*A*) whereas marked initial increases followed by slight long-term decreases were seen during azathioprine treatment (Figure 2*B*).

Separate models designed to study interactions between therapy and all the variables included in the model showed that the favorable prognostic influence of a high albumin tended to be more marked in the azathioprine group, whereas the harmful prognostic influence of old age tended to be more marked in the control group. However, neither of these interactions was statistically significant. Therefore, therapy was not included in the final model.

Time-Dependent Prognostic Index

The information in Tables 1 and 2 can be used to calculate time-dependent prognostic indices for patients whose data are updated during their follow-up.

If, for example, at a given time a patient has the following variables: bilirubin 45 μmol/L, ascites, albumin 30 g/L, age 48 years, and gastrointestinal bleeding, PI(t) based on model A-I (Table 1) is calculated as the sum of regression coefficients each multi-

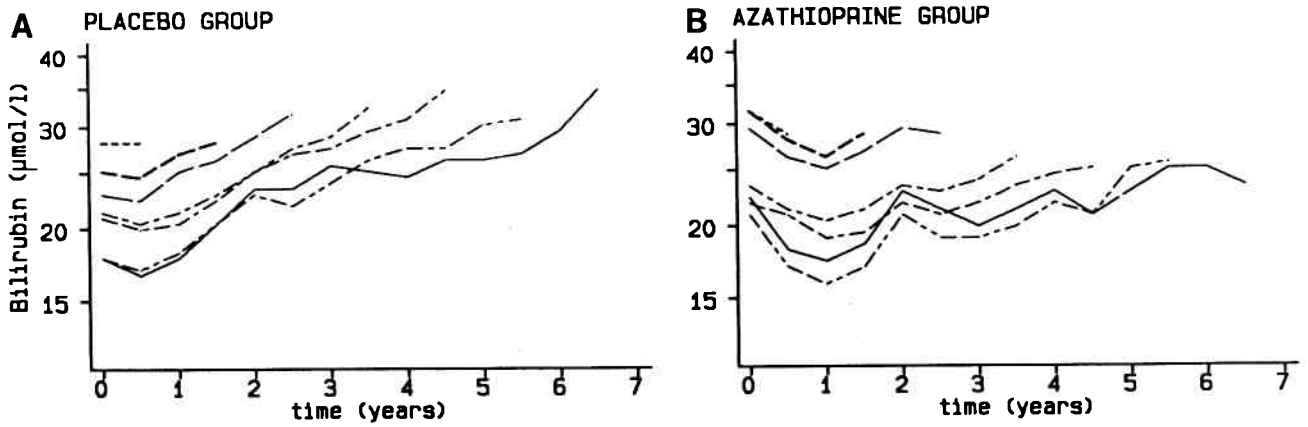

Figure 1. Course of serum bilirubin in overlapping groups of patients with complete data in each period of observation. Geometric means in μmol/L are presented separately for placebo- (*A*) and azathioprine-treated (*B*) groups. The number of patients contributing to the curves are as follows: (*A*) 6 months, 98; 18 months, 75; 30 months, 67; 42 months, 51; 54 months, 42; 66 months, 28; and 78 months, 17; (B) 6 months, 102; 18 months, 95; 30 months, 79; 42 months, 64; 54 months, 56; 66 months, 44; 78 months, 28. Six months, -----; 18 months, – – –; 30 months, — — —; 42 months, -----; 54 months, – – – –; 66 months, – – – – –; 78 months, ——.

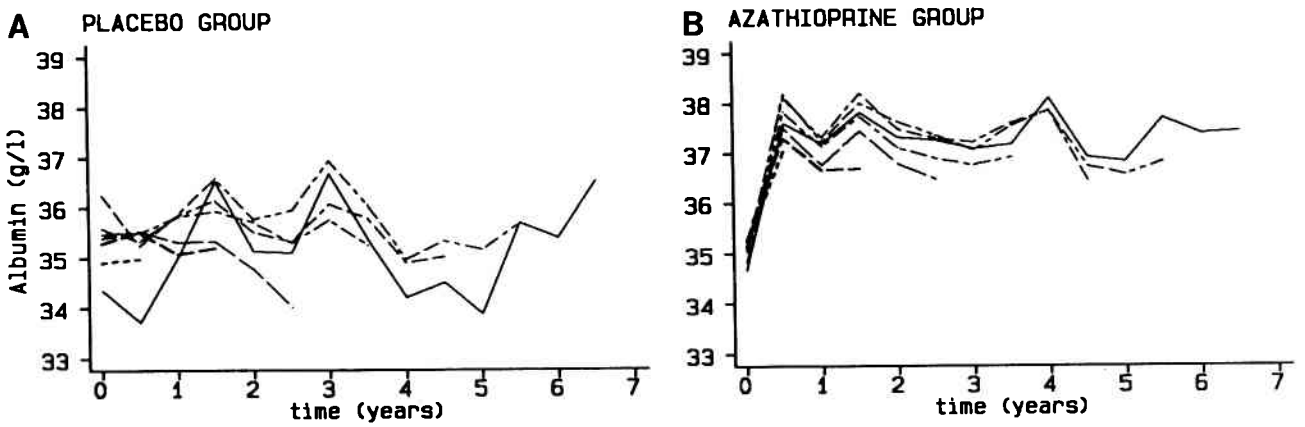

Figure 2. Course of serum albumin in overlapping groups of patients with complete data in each period of observation. Arithmetic means in grams per liter are presented separately for placebo (*A*) and azathioprine (*B*) groups. The number of patients contributing to the curves are as follows: (*A*) 6 months, 97; 18 months, 74; 30 months, 66; 42 months, 50; 54 months, 41; 66 months, 27; 78 months, 16; (*B*) 6 months, 100; 18 months, 93; 30 months, 77; 42 months, 64; 54 months, 54; 66 months, 43; 78 months, 28. Six months, -----; 18 months, – – –; 30 months, — — —; 42 months, –––––; 54 months, – ––– –; 66 months, – ––– –; 78 months, ——.

plied by the scoring of the corresponding variable in this way: PI(t) = 2.53 × ($\log_{10}$[45] − 1.53) + 1.39 × 1 − 0.085 × (30 − 34.3) + 0.040 × (48 − 55) + 0.65 × 1 = 2.53 × 0.123 + 1.39 × 1 − 0.085 × (−4.3) + 0.040 × (−7) + 0.65 × 1 = 2.44.

Transformation of PI(t) to Survival Probability

From a given value of PI(t), the survival probability can be obtained using equation 2 in Patients and Methods. The λ_0 value to be used is given in the Appendix (Figure 4) for each of the four models. For example, for model A-I, for which $\lambda_0 = 0.035$ years^{-1}, the probability of surviving the next 6 months ($\lambda = 0.5$ years) given the above PI(t) of 2.44 can be estimated as exp(−0.035 × 0.5 × exp[2.44]) = exp(−0.035 × 0.5 × 11.47) = 0.82 or 82%.

Figure 3 shows the estimated probability of surviving the next 1, 3, and 6 months as a function of the time-dependent prognostic index PI(t) for each of the 4 models presented. For the example presented above (with PI(t) = 2.44), the probability of surviving the following 1, 3, and 6 months can be read as 97%, 91%, and 82%, respectively (Figure 3 [A-I]).

Table 3 shows more examples of prognostic indices for different patients and the corresponding 6-month survival probabilities predicted for each of the four presented models. There seems to be a reasonable agreement between the 6-month survival probability estimated by the four models, although slightly higher survival probabilities are obtained from models using only hepatic deaths as events as would be expected.

Calculation of PI(t) can be simplified using the pocket charts shown in the Appendix.

Confidence Limits of Survival Probabilities

Approximate confidence limits of estimated survival probabilities[2,19] can be calculated assuming unchanged variables and death risk for the period concerned (e.g., over 6 months). For example, for patient 3 in Table 3 having an estimated 6-month survival probability of 81.8% (model A-I), the 95% confidence interval would be 71.2%–92.5%. For patient 2 in the same table, who has an estimated 6-month survival probability of 98.7%, the corresponding 95% confidence interval would be 98.0%–99.4%. As for the time-fixed Cox model,[2] the confidence interval is widest for survival probabilities in the middle range (around 50%).

Validation of Model

As described in the Appendix, the goodness of fit of the time-dependent Cox model to the data was confirmed by a number of methods. In independent PBC patients, the observed survival was found to agree well with the survival predicted by the model (see Appendix).

When the predictions of the time-dependent model were compared with those of our previously developed time-fixed model,[2] we found that the time-dependent predictions were markedly better (see Appendix).

Discussion

Even though the course of PBC is one of progression, short-term fluctuations occur. Following

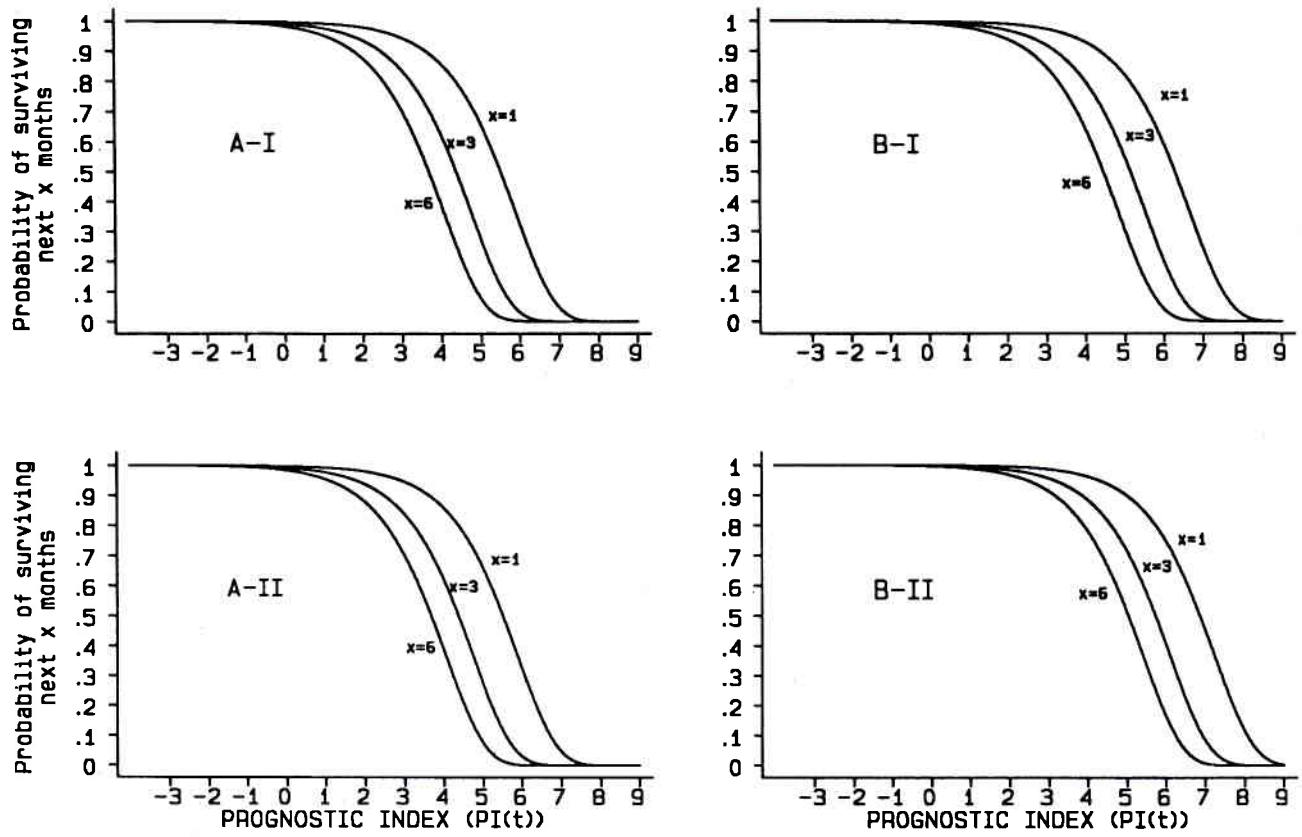

Figure 3. Probability of surviving 1, 3, and 6 months as a function of the PI(t) for models excluding biopsy variables (A-I [all deaths treated as events], A-II [only hepatic deaths treated as events]) and models including biopsy variables (B-I [all deaths treated as events], B-II [only hepatic deaths treated as events]).

diagnosis, which tends to be made during an exacerbation for which the patients seek medical assistance, slight improvement (regression toward the mean[18]) may occur, even in placebo treated patients, as shown in Figures 1*A* and 2*A*. This means that later "steady state" values (e.g., after some months) may better reflect the degree of permanent liver damage than the initial values. However, the previously published time-fixed models[1–5] are all based on the relation between the initial (more abnormal) status and the subsequent survival. Therefore, if time-fixed models are applied to follow-up data, they may tend to overestimate survival. (In the development of the model, the observed survival is being related to a "too abnormal" status. Therefore, from a less abnormal follow-up status the model will predict a better survival than that observed.)

The present analysis shows (Table 4) that our previously published time-fixed model[2] tends to overestimate survival (or underestimate death risk) when applied to follow-up data. A similar overestimation of survival has been observed for the time-fixed Mayo model when used on follow-up data.[20] This implies that time-fixed models are not well suited for updating of prognosis from follow-up data; they are not designed for that purpose.

Although some developments have taken place in medical therapy,[8–10] progress has been very limited and no known medical therapy can stop either progression of the disease or the occurrence of complications and the need for transplantation. Therefore, we think that our data contain information of current value.

Previously, we reported that azathioprine has a small but significantly beneficial effect on survival and incapacitation of patients with PBC.[2] The more detailed analysis performed in this study shows that the beneficial effect of azathioprine is apparent also on the course of bilirubin and albumin, especially just after the start of treatment. It seems that azathioprine treatment was not significant in the time-dependent models because the treatment effect is carried by the time-dependent prognostic variables in the model, especially bilirubin, so that therapy becomes redundant in the model.

Because follow-up data contribute in a time-dependent Cox model, it may be expected to describe more

Table 3. Example of Calculation of Time-Dependent Prognostic Indices for Three Patients With Good, Medium, and Poor Prognosis

Variable	Patient 1 (good prognosis; 10th percentile of PI)	Patient 2 (medium prognosis; median PI)	Patient 3 (poorer prognosis; 90th percentile of PI)
Bilirubin ($\mu mol/L$)	11	35	45
Central cholestasis	No	No	No
Ascites	No	No	Yes
Cirrhosis	No	Yes	Yes
Gastrointestonal bleeding	No	No	Yes
Albumin (g/L)	45	40	30
IgM (g/L)	3.0	3.8	4.0
Age (yr)	63	59	48
Models excluding biopsy variables			
$PI_{A\text{-}I}$	−1.83	−0.29	2.44
$P_{A\text{-}I}$ (6 mo)	99.7%	98.7%	81.8%
$PI_{A\text{-}II}$	−1.96	−0.22	2.57
$P_{A\text{-}II}$ (6 mo)	99.9%	99.2%	87.2%
Models including biopsy variables			
$PI_{B\text{-}I}$	−1.64	0.51	3.27
$P_{B\text{-}I}$ (6 mo)	99.8%	98.7%	82.0%
$PI_{B\text{-}II}$	−1.82	0.58	3.60
$P_{B\text{-}II}$ (6 mo)	99.9%	99.2%	84.8%

precisely the effect of the prognostic variables, provided the model assumptions can be considered fulfilled.

We present two sets of models, one based on clinical and biochemical variables and the other also including histological variables. Because liver biopsies are not often performed during follow-up, the former models are more relevant in the clinical follow-up situation than the latter.

Most of the variables included in the models correspond to those included in the published time-fixed models.[1–5] Serum bilirubin and age have been found to be important prognostic in all the time-fixed models.[1–5] Serum bilirubin is also by far the most important prognostic variable in the time-dependent models. In addition, the following included time-dependent variables have also been included in time-fixed models: albumin,[2–3,5] ascites,[4–5] gastrointestinal bleeding,[4] cirrhosis,[1–2,5] and central cholestasis.[2,4] The only new prognostic variable identified in the time-dependent analysis was IgM, low values being associated with a poor prognosis. The biological significance of this is uncertain, but it may reflect a more severe immune disturbance and proneness to infection in some patients.[21] Even though IgM was not significant in the univariate analysis, it was highly significant in the multivariate analysis. Nevertheless, we cannot rule out completely that it is a false-positive finding. Unlike in our time-fixed model,[2] a linear scoring of age was adequate in the time-dependent models.

The prognostic precision of time-fixed models is not very high.[2–3,5–6] Besides not making use of the follow-up information, a possible cause for this prognostic imprecision may be that all deaths irrespective of the cause have been considered endpoints in those analyses. This may not be reasonable because deaths from nonhepatic causes may not be associated with the severity of PBC. To study the effect of the nonhepatic deaths, time-dependent models were developed with nonhepatic deaths being either included as endpoints (models A-I and B-I) or censored (models A-II and B-II). The effect of censoring the nonhepatic deaths turned out to be surprisingly small because the differences between the models with or without hepatic deaths as end-points were quite small (Tables 1 and 2).

The final models were validated by a number of methods as described in the Appendix. Despite the fact that the assumption of variables being constant between recordings may not be entirely correct (in fact, we found a slight increase in the average PI(t) over time, implying that the risk may be slightly underestimated by the model) the fit of the model to the data was satisfactory in all cases. Furthermore, as expected, the time-dependent model gave more accurate short-term prediction than the previously published time-fixed model.[2]

The time-dependent prognostic indices based on the models are easily calculated (and they can be obtained even more simply using the pocket charts in the Appendix). The analysis showed that a given value of PI(t) has the same prognostic significance whether early or late in the course of the disease. Thus the varying risk can be described just by the varying levels of the prognostic variables. A prognostic index can be

easily transformed to estimates of surviving the following 1, 3, or 6 months. We chose 6 months as the limit for prediction because the covariates are likely to change beyond that time, which will affect prognosis.

Because the time-dependent models have been based on a fuller use of the data, they should be preferred to the time-fixed models when the aim is to monitor a PBC patient during follow-up. Here they may contribute to the decision of if and when to perform a liver transplantation,[22] which involves comparison of pretransplant prognosis with and without transplantation.[6,11,23–27] As a preliminary practical guideline for using the time-dependent model, we recommend that transplantation be considered when the estimated 6-month survival probability drops to 80% or less. This is in accordance with recent recommendations based on the Mayo model.[11] However, to define more precisely the optimal use of the time-dependent model for timing of liver transplantation, further evaluation in prospective studies is necessary.

Even though patients selected for liver transplantation may remain on a waiting list for a long time, they will be followed closely for monitoring the course of the disease. If prognosis becomes worse, their position in the waiting list may be advanced. Thus, to keep the waiting list up-to-date, it should be sorted at all times according to the urgency of the need. In this context, the time-dependent model can also provide a valuable objective contribution.

Appendix

Statistical Methods

In this study, we used the Cox model for time-dependent variables.[13] This model has the same form as the time-fixed Cox model but allows the variables to change in time. The prognostic index PI(t), which is a function of the patient's variables at a given time t being weighted by the corresponding regression coefficients, is also equal to log ($\lambda[t]/\lambda_0[t]$) where $\lambda(t)$ is the patient's death risk or hazard a time t given his variables at that time and $\lambda_0(t)$ is the so-called baseline or underlying hazard function for the time-dependent model.[7] ($\lambda_0(t)$ is defined as the hazard of a hypothetical subject with variable scorings of zero at time t.[15]) In the analysis, continuous variables were "centred" by subtracting the mean. This had the effect of making the underlying hazard correspond to that of an average patient (with variables equal to the mean for continuous variables and zero for binary variables).

For all four models, Figure 4 shows the cumulative baseline hazard function $\Lambda_0(t)$, being the integral of the baseline hazard function $\lambda_0(t)$ over the time interval from zero to t. Because this function turned out to be linear in all four models apart from the last 2 years, where the uncertainty of the estimated curves is great because of the few patients at risk at that time, the underlying hazard function may be considered constant (independent of time t) for the first 6 years. This means that, for a given value of PI(t), the prognostic information for a subsequent short time period is the same whether early or late in the course of the disease, at least within the first 6 years. The underlying hazards were estimated by simple linear regression over that period for each specification of the time-dependent model. The values obtained (Figure 4) are not directly comparable because they refer to different specifications of the prognostic index PI(t).

Statistical Programs

The statistical analyses were performed using (1) the BMDP 2L program,[28] with the support of a purposely written Fortran subroutine and (2) the time-varying option for the Cox regression analysis in the statistical package STATA.[29] The two procedures led to identical results. The underlying hazard function was obtained from a modified version of the program by Kalbfleisch and Prentice.[30]

Assessment of Model Goodness of Fit

There are no standard procedures to assess the goodness of fit of a time-dependent Cox model, but several steps can be taken to evaluate its adequacy in fitting the original data and its predictive power on independent observations. We have used four approaches, which follow.

Examination of residuals. In linear regression analyses, residuals, defined as the discrepancy between the observed and fitted values, are often examined to assess the goodness-of-fit of the model and the presence of extreme values. Equivalent values have been defined for the Cox model. In this paper, the unstandardized generalized residuals of Barlow and Prentice,[31] which are defined for each subject and each variable in a model, were examined to identify individuals having an undue influence in the analysis. For each specification (A-I and B-I) of the model, we examined the residuals by plotting them against the ranks of the subjects' observation times.

We found that the unstandardized residuals of Barlow and Prentice[31] computed for each variable in the models with or without histological variables identified a few outliers in the data. Their deletion did not have a substantial effect on the coefficients.

Comparison of observed and expected survival. Observed survival frequencies were computed over separate 6-month intervals for up to 6 years of follow-up (i.e., a total of 12 intervals) for nine categories of the two time-dependent prognostic indices derived from the models with and without histological variables for all deaths treated as events (A-I and B-I). These frequencies were calculated at the beginning of each interval as the number of interval survivors divided by the total number of subjects alive at the beginning of the interval. Subjects censored during the interval were included in the computation.

The corresponding heuristic 6-month survival probabili-

Table 4. Observed and Expected Number of Events From the Previously Published Time-Fixed Model[2] and Two Time-Dependent Models With All Deaths Treated as Events

Model	No. of subjects	Categories of bilirubin								
		Low			Medium			High		
		Obs	E_F	E_T	Obs	E_F	E_T	Obs	E_F	E_T
Excluding histologic variables (A-I)	237	17	8.3	10.9	38	22.7	36.5	36	25.2	33.9
Including histologic variables (B-I)	191	15	7.2	8.0	35	21.6	31.0	30	19.9	24.5

E_F, Expected number of events in our (previously published[2]) time-fixed model; E_T, Expected number of events in the time-dependent model; Obs, Observed number of events.

ties were estimated as described in the text for each specification (A-I and B-I) of the time-dependent model.[7]

Observed survival rates and the expected survival probabilities from Equation 2 in Patients and Methods were computed over the separate six-month intervals and then pooled and plotted as a function of PI(t).

Figure 5 shows the pooled observed 6 months survival frequency as a function of the mean PI(t) values for 9 PI(t) intervals covering the full range, calculated for all the patients at 6-month intervals for the first 6 years of observation. For comparison, the 6 months heuristic survival probabilities expected from the model are also given. The results are given for both the model excluding histology (A-I) and the model including histology (B-I). Because the model was fitted on the same data, one would expect a good correspondence between observed and expected 6-month survival probabilities. However, the closeness of the values suggests an adequate fit of the models to the data.

Comparison of observed and expected number of events. Following the same approach described in the comparison of observed and expected survival, we computed the expected numbers of deaths for each 6-month interval of the first 6 years of follow-up on the basis of the models A-I and B-I. We then compared these with the observed values. To

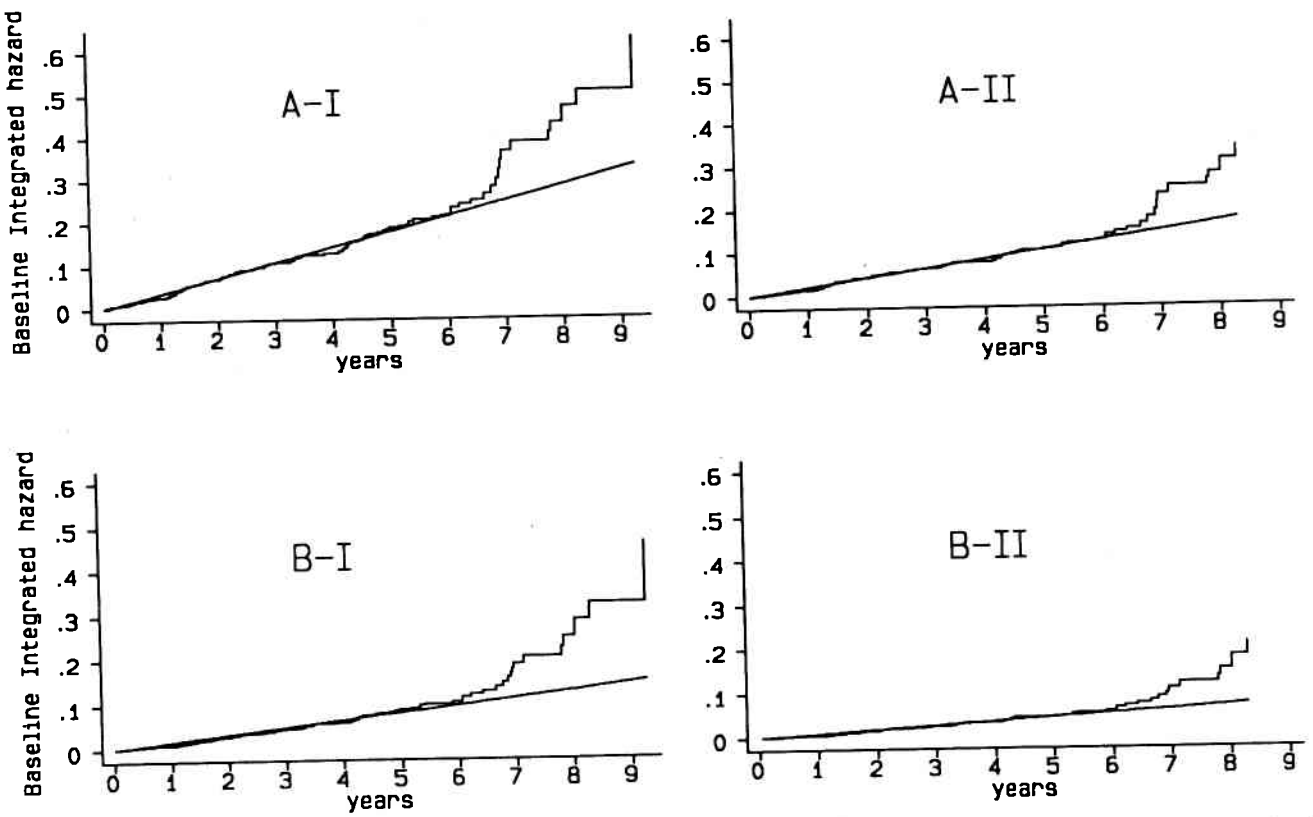

Figure 4. Estimated integrated baseline hazard functions for the four presented time-dependent models. The underlying hazard for the first 6 years were estimated by linear regression to give the following values, models excluding biopsy variables: A-I (all deaths treated as events), 0.035 years^{-1} (SE, 0.0002); A-II (only hepatic deaths treated as events), 0.021 years^{-1} (SE, 0.0001). Models including biopsy variables: B-I (all deaths treated as events), 0.016 years^{-1} (SE, 0.0002); B-II (only hepatic deaths treated as events), 0.009 years^{-1} (SE, 0.0001).

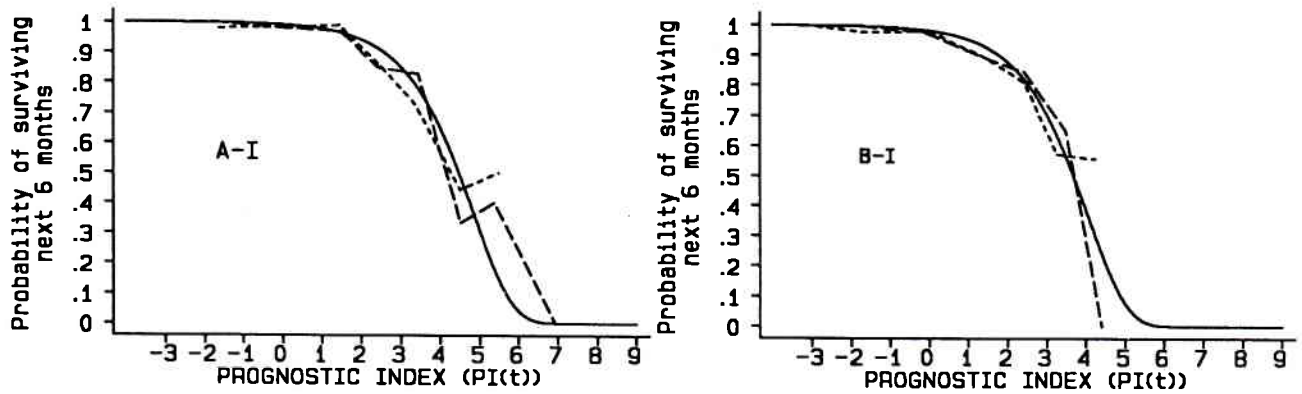

Figure 5. Comparison of observed survival frequency and expected survival probability for different values of the PI(t) (in PBC1 patients) of model A-I (excluding biopsy variables and all deaths treated as events) and B-I (including biopsy variables and all deaths treated as events). ——, expected survival probability; -----, observed survival frequency for placebo group; – – –, observed survival frequency for azathioprine group.

assess the predictive power of the model in different subsets of the patients, we carried out this comparison separately for subjects whose bilirubin at the beginning of the interval was <36.3 μmol/L, between 36.3 and 120.2 μmol/L, and >120.2 μmol/L (the 60th and 90th percentile of the observed bilirubin values, respectively). We used categories of bilirubin to evaluate the fit because bilirubin is known to be highly prognostic.[1–5]

Table 4 compares the observed total number of events in the three bilirubin categories with those predicted by the two time-dependent models (A-I and B-I). The observed values differ because the model with histological variables was fitted on a smaller set of patients. Both models appear to fit the data adequately for medium and high bilirubin values but underestimate the number of events in the lowest category. The consistent slight excess of observed over expected numbers of events appears to result from a very slight increase over time (0.02 per month) in the average prognostic index, in contrast to the assumption that the covariates remain unchanged.

Testing of model prediction in independent patients. The developed time-dependent Cox model was further assessed by testing its predictive power in 147 independent PBC patients (with complete data) who were included in the randomized clinical trial of D-penicillamine vs. placebo (PBC2 trial).[12] The survival predicted by the time-dependent Cox model was compared with the observed survival in the same way as described above.

In Figure 6 the results are presented in the same way as in Figure 5. For these independent patients there is also a good agreement between observed and expected 6-month survival probabilities, supporting the validity of the model.

Comparison of fit between time-fixed and time-dependent models. The comparison of observed and expected number of events was also performed for the time-fixed model presented by Christensen et al.[2] The updated values available at the beginning of each interval were used for the prediction to mimic the use clinicians might make of such a model.

Table 4 also shows the values computed when the previously published time-fixed model[2] was used to predict the number of events in the 237 and 191 patients used to fit the

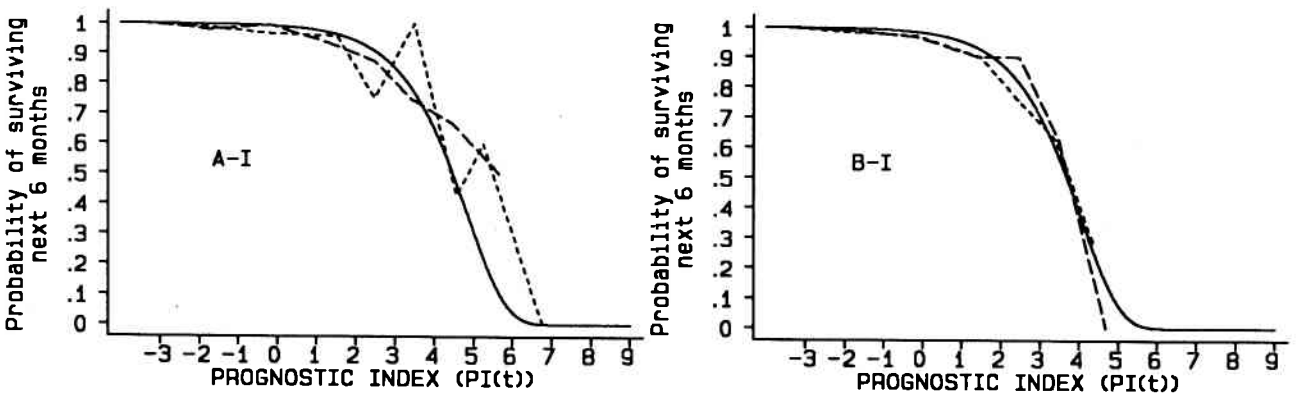

Figure 6. Comparison of observed survival frequency and expected survival probability for different values of the time-dependent prognostic index (in independent PBC2 patients) of model A-I (excluding biopsy variables and all deaths treated as events) and B-I (including biopsy variables and all deaths treated as events). ——, expected survival probability; ----, observed survival frequency for placebo group; – – –, observed survival frequency for D-penicillamine group.

Table 5. Pocket Chart for Easy Calculation of the PI(t) Excluding Liver Biopsy Variables

Variable		Points to add: All deaths (model A-I)	Points to add: Hepatic deaths (model A-II)
Bilirubin			
mg/100 mL	µmol/L		
0.6	10	−6	−9
1.2	20	1	0
2.0	34	7	7
2.9	50	11	12
4.1	70	15	17
5.8	100	19	21
8.8	150	23	27
11.7	200	27	30
17.5	300	31	36
23.4	400	34	39
29.2	500	37	42
Ascites			
Absent		0	0
Present		14	14
Albumin			
µmol/L	g/L		
304	20	−6	−7
456	30	−14	−15
517	34	−18	−18
608	40	−23	−22
760	50	−31	−30
912	60	−40	−38
Age (*yr*)			
30		1	0
40		5	5
50		9	9
55		11	11
60		13	13
70		17	17
80		21	22
Gastrointestinal bleeding			
Absent		0	0
Present		7	7
Sum of added points (S) = PI(t) = S/10 =			

Table 6. Pocket Chart for Easy Calculation of the PI(t) Including Liver Biopsy Variables

Variable		Points to add: All deaths (model B-I)	Points to add: Hepatic deaths (model B-II)
Bilirubin			
mg/100 mL	µmol/L		
0.6	10	−6	−9
1.2	20	1	0
2.0	34	6	6
2.9	50	10	11
4.1	70	13	15
5.8	100	17	19
8.8	150	21	24
11.7	200	23	28
17.5	300	27	32
23.4	400	30	36
29.2	500	32	39
Central cholestasis			
Absent		0	0
Present		12	13
Ascites			
Absent		0	0
Present		12	13
Cirrhosis			
Absent		0	0
Present		9	9
Gastrointestinal bleeding			
Absent		0	0
Present		10	11
Albumin			
µmol/L	g/L		
304	20	0	−1
456	30	−7	−7
517	34	−10	−10
608	40	−14	−14
760	50	−21	−20
912	60	−28	−26
IgM			
µmol/L	g/L		
1.1	1	0	0
3.2	3	−4	−4
5.3	5	−6	−6
10.5	10	−9	−9
15.8	15	−10	−10
21.0	20	−11	−11
Age (*yr*)			
30		1	2
40		4	4
50		7	7
55		8	8
60		9	9
70		12	12
80		15	14
Sum of added points (S) = PI(t) = S/10 =			

time-dependent models. In both cases, the predictions are closer to the observed values when using the time-dependent model.

Pocket charts to obtain time-dependent prognostic index PI(t). The time-dependent prognostic indices can be obtained in a simple way at the bed side using the pocket charts presented in Tables 5 and 6 in which the regression terms have been replaced by simple numbers. Only one number for each variable should be used. Interpolation should be performed if a patient has values between those in the tables.

If, for example, at a given time a patient has the variables bilirubin (45 µmol/L), ascites, albumin (30 g/L), age (48 years), and gastrointestinal bleeding, the PI(t) based on the

model (A-I) without histological variables and all deaths treated as end points (Table 3) are calculated to one decimal as (10 [for bilirubin] + 14 [for ascites] − 14 [for albumin] + 8 [for age] + 7 [for gastrointestinal bleeding])/10 = 2.5, which is only slightly larger than the precise value of 2.44 in Table 3.

References

1. Roll J, Boyer JL, Barry D, Klatskin G. The prognostic importance of clinical and histologic features in asymptomatic and symptomatic primary biliary cirrhosis. N Engl J Med 1983;308:1–7.
2. Christensen E, Neuberger J, Crowe J, Altman DG, Popper H, Portmann B, Doniach D, Ranek L, Tygstrup N, Williams R. Beneficial effect of azathioprine and prediction of prognosis in primary biliary cirrhosis. Final results of an international trial. Gastroenterology 1985;89:1084–1091.
3. Dickson ER, Grambsch PM, Fleming TR, Fisher LD, Langworthy A. Prognosis in primary biliary cirrhosis: Model for decision making. Hepatology 1989;10:1–7.
4. Goudie BM, Burt AD, Macfarlane GJ, Boyle P, Gillis CR, MacSween RNM, Watkinson G. Risk factors and prognosis in primary biliary cirrhosis. Am J Gastroent 1989;84:713–716.
5. Biagini MR, Guardascione M, Raskino C, McIntyre N, Surrenti C, Burroughs AK. Poor prognostication for survival of individual PBC patients with Cox models (abstr). J Hepatol 1990;11(Suppl 2):S7.
6. Christensen E. Prognostication in primary biliary cirrhosis: Relevance to the individual patient. Hepatology 1989;10:111–113.
7. Christensen E, Schlichting P, Andersen PK, Fauerholdt L, Schou G, Pedersen BV, Juhl E, Poulsen H, Tygstrup N, Copenhagen Study Group for Liver Diseases. Updating prognosis and therapeutic effect evaluation in cirrhosis with Cox's multiple regression model for time-dependent variables. Scand J Gastroenterol 1986;21:163–174.
8. Wiesner RH, Grambsch PM, Lindor KD, et al. Clinical and statistical analysis of new and evolving therapies for primary biliary cirrhosis. Hepatology 1988;8:668–676.
9. Slitzky BE, Ouellette GS, Boyer JL. Approaches to the treatment of primary biliary cirrhosis—A status report. Gastroenterol Int 1990;3:134–139.
10. Lombard M, Portmann B, Neuberger J, Williams R, Tygstrup N, Ranek L, Ring-Larsen H, Rodes J, Navasa M, Trepo C, Pape G, Schou G, Badsberg JH, Andersen PK. Cyclosporin A treatment in primary biliary cirrhosis: Results of a long-term placebo controlled trial. Gastroenterology 1993;104:519–526.
11. Wiesner RH, Porayko MK, Dickson ER, Gores GJ, LaRusso NF, Hay JE, Wahlstrom HE, Krom RAF. Selection and timing of liver transplantation in primary biliary cirrhosis and primary sclerosing cholangitis. Hepatology 1992;16:1290–1299.
12. Neuberger J, Christensen E, Portmann B, Caballeria J, Rodes J, Ranek L, Tygstrup N, Williams R. Double-blind controlled trial of d-penicillamine in patients with primary biliary cirrhosis. Gut 1985;26:114–119.
13. Andersen PK, Borgan Ø, Gill RD, Keiding N. Statistical models based on counting processes. New York. Springer, 1993.
14. Altman DG, De Stavola BL. Practical problems in fitting a proportional hazards model to data with updated measurements of the covariates. Statistics in Medicine (in press).
15. Christensen E. Multivariate survival analysis using Cox's regression model. Hepatology 1987;7:1346–1358.
16. Christensen E, Schlichting P, Andersen PK, Fauerholt L, Juhl E, Poulsen H, Tygstrup N, Copenhagen Study Group for Liver Diseases. A therapeutic index that predicts the individual effects of prednisone in patients with cirrhosis. Gastroenterology 1985;88:156–165.
17. Christensen E. Individual therapy-dependent prognosis based on data from controlled clinical trials in chronic liver disease (thesis). Dan Med Bull 1988;35:167–182.
18. Christensen E, Schlichting P, Fauerholdt L, Juhl E, Poulsen H, Tygstrup N, Copenhagen Study Group for Liver Diseases. Changes of laboratory variables with time in cirrhosis. Prognostic and therapeutic significance. Hepatology 1985;5:843–853.
19. Altman DG, Andersen PK. A note on the uncertainty of a survival probability estimated from Cox's regression model. Biometrika 1986;73:722–724.
20. Klion FM, Fabry TL, Palmer M, Schaffner F. Prediction of survival of patients with primary biliary cirrhosis. Examination of the Mayo Clinic model on a group of patients with known endpoint. Gastroenterology 1992;102:310–313.
21. Morreale M, Tsirigotis M, Hughes MD, Brumfitt W, McIntyre N, Burroughs AK. Significant bacteriuria has prognostic significance in primary biliary cirrhosis. J Hepatol 1989;9:149–158.
22. Esquivel CO, Van Thiel DH, Demetris AJ, Bernardos A, Iwatsuki S, Markus B, Gordon RD, Marsh JW, Makawka L, Tzakis AG, Todo S, Gavaler JS, Starzl TE. Transplantation for primary biliary cirrhosis. Gastroenterology 1988;94:1207–1216.
23. Neuberger J, Altman DG, Christensen E, Tygstrup N, Williams R. Use of a prognostic index in evaluation of liver transplantation for primary biliary cirrhosis. Transplantation 1986;41:713–716.
24. Markus BH, Dickson ER, Grambsch PM, Fleming TR, Mazaferro V, Klintholm GB, Wiesner RH, Van Thiel DH, Starzl TE. Efficacy of liver transplantation in patients with primary biliary cirrhosis. N Engl J Med 1989;320:1709–1713.
25. Bonsel GJ, Klompmaker IJ, Van't Veer F, Babbema JDF, Slooff MJH. Use of prognostic models for assessment of value of liver transplantation in primary biliary cirrhosis. Lancet 1990;335:493–497.
26. Neuberger J, Altman DG, Polson R, Buckels J, Rolles K, Elias E, Calne R, McMaster P, Williams R. Prognosis after liver transplantation for primary biliary cirrhosis. Transplantation 1989;48:444–447.
27. Keiding S, Ericzon B-G, Eriksson S, Flatmark A, Hockerstedt K, Isoniemi H, Karlberg I, Keiding N, Olsson R, Samela K, Schrumpf E, Soderman C. Survival after liver transplantation of patients with primary biliary cirrhosis in the Nordic countries. Scand J Gastroenterol 1990;25:11–18.
28. Dixon WJ. BMDP statistical software. Berkeley, California: University of California, 1988.
29. STATA Reference manual, release 2, Los Angeles Computing Resource Center, 1989.
30. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. New York: Wiley, 1980.
31. Barlow WE, Prentice RL. Residuals for relative risk regression. Biometrika 1988;75:65–74.

Received June 25, 1992. Accepted July 20, 1993.

Address requests for reprints to: Erik Christensen, M.D., D.M.Sc., Department of Medicine B, Bispebjerg University Hospital, Bispebjerg Bakke 23, DK-2400 Copenhagen N, Denmark.

The PBC1 Trial Group consisted of Jean-Piere Benhamou, Clichy, France; Erik Christensen, Hvidovre, Denmark; John Crowe, London, England; Deborah Doniach, London, England; Jan de Groote, Leuven, Belgium; Kurt Iversen, Copenhagen, Denmark; Erik Juhl, Hvidovre, Denmark; Steven Mistilis, Sidney, Australia; James Neuberger, London, England; Hans Popper, New York, New York; Bernard Portmann, London, England; Leo Ranek, Copenhagen, Denmark; Helmer Ring-Larsen, Hvidovre, Denmark; Juan Rodes, Barcelona, Spain; Fenton Schaffner, New York, New York; Niels Tygstrup, Copenhagen, Denmark; Geoffrey Watkinson, Glasgow, Scotland; Roger Williams, London, England. The PBC2 Trial Group consisted of Juan Caballeria, Barcelona, Spain; Erik Christensen, Hvidovre, Denmark; Erik Juhl, Hvidovre, Denmark; James Neuberger, London, England; Hans Popper, New York, New York; Bernard Portmann, London, England; Leo Ranek, Copenhagen, Denmark; Helmer Ring-Larsen, Hvidovre, Denmark; Juan Rodes, Barcelona, Spain; Niels Tygstrup, Copenhagen, Denmark; Roger Williams, London; England.