Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population

C.Y. Ponsioen\textsuperscript{a}, S.M.E. Vrouwenraets\textsuperscript{a}, W. Prawirodirdjo\textsuperscript{b}, R. Rajaram\textsuperscript{a} E.A.J. Rauws\textsuperscript{b}, C.J.J. Mulder\textsuperscript{b}, J.B. Reitsma\textsuperscript{a}, S.H. Heisterkamp\textsuperscript{b}, G.N.J. Tytgat\textsuperscript{b}

\textsuperscript{a}Department of Gastroenterology & Hepatology, Academic Medical Center, Amsterdam, The Netherlands
\textsuperscript{b}Department of Internal Medicine, Rijnstate Hospital, Arnhem, The Netherlands
\textsuperscript{c}Department of Epidemiology & Biostatistics, Academic Medical Center, Amsterdam, The Netherlands

Submitted for publication
ABSTRACT

Background-The median survival of primary sclerosing cholangitis (PSC) has been estimated to be 12 years. Cholangiography is the golden standard for diagnosis, but is hardly used in estimating prognosis.

Aims-To assess the natural history of Dutch PSC patients and to evaluate the prognostic value of a cholangiographic classification system.

Patients-174 patients with established PSC from a university hospital and three teaching hospitals seen from 1970 through 1999.

Methods-Charts were reviewed for validity and time of diagnosis, concurrent inflammatory bowel disease, interventions, liver transplantation, and occurrence of cholangiocarcinoma, and death. Follow-up data were obtained from the charts and from the attending clinician or family physician. Median follow-up was 76 months (range 1-300). The earliest available cholangiography was scored as to the severity of sclerosis, using a radiological classification system developed in our institution. Survival curves were computed by the Kaplan-Meier method. Cholangiographic staging was used to construct a prognostic model, applying Cox proportional-hazards analysis.

Results-The estimated median survival from time of diagnosis to death from liver disease or liver transplantation was 18 years. Cholangiocarcinoma was found in 18 (10 %) of patients. Fourteen patients (8 %) underwent liver transplantation. Cholangiographic scoring was inversely correlated with survival. A combination of intrahepatic and extrahepatic scoring, together with the age at ERCP proved strongly predictive for survival.

Conclusions-The observed survival is considerably better than reported in earlier series from Sweden, the UK, and the USA. Classification and staging of cholangiographic abnormalities has prognostic value.

Keywords: primary sclerosing cholangitis; cholangiocarcinoma; natural history; survival; liver transplantation; prognosis
INTRODUCTION

Primary sclerosing cholangitis is a chronic progressive cholestatic disease, characterised by patchy inflammation of the biliary tree, resulting in obliterative fibrosis. The etiology is unknown and to date there is no therapy that can effectively halt disease progression. Many patients are nowadays treated with ursodeoxycholic acid, which in a considerable proportion of patients improves biochemical cholestatic parameters, and to a lesser extent alleviates cholestatic symptoms such as pruritus and fatigue. However, there are two randomised placebo controlled studies published so far, which evaluated medium long-term ursodeoxycholic acid treatment. Neither showed clear benefit on clinical parameters, histology, and delay in time until treatment failure (1, 2). Whether higher doses could yield more favourable outcomes, remains to be seen. Dominant strictures in the major draining bile ducts are treated by endoscopic or percutaneous balloon dilatation or temporary stent placement. Several groups have reported favourable results of dilatation in terms of remission of symptoms and improvement of cholestatic biochemical parameters (3-9). Whether it also delays disease progression is not known.

Regarding the natural history of PSC three larger independent studies containing more than 100 cases have been published so far (10-12). They are concordant in the estimated median survival, which is about 12 years. All three studies report an incidence of cholangiocarcinoma (CCA) of 6-8%. When counselling patients, we usually rely on these 1970-1980s based series.

Two studies have been published that assessed the value of cholangiography in predicting prognosis. Both concluded that high-grade intrahepatic strictures were indicative of a poor prognosis (13, 14). However, the cholangiographic abnormalities encountered in PSC are not widely used to assess disease stage, probably because there is no uniformly accepted classification system. In 1984, Li-Yeng et al. were the first to propose a classification of the cholangiographic findings in PSC, based on a series of 19 patients (15). Several years later, Majoie et al. from our institution proposed a modification of this system of radiographic findings (16).

The aims of the present study were to assess the natural history of Dutch PSC patients, and to evaluate the predictive value of our cholangiographic classification system.
PATIENTS AND METHODS

Patients

The records from all 181 patients, who were filed with a diagnosis of primary sclerosing cholangitis between 1970-1999 from one tertiary referral centre and three teaching hospitals, were retrieved.

The diagnosis was reviewed together with cholangiograms and histology results, whenever available. Seven patients were excluded, because they lacked sufficient evidence for the diagnosis of primary sclerosing cholangitis. Three of these patients had a Klatskin tumour, and four had complicated gallstone disease. From the remaining 174 patients the following data were recorded: age at diagnosis, sex, concurrent IBD, cholecystectomy, ABO blood type, endoscopic retrograde cholangiopancreatography (ERCP) findings, age at index ERCP (AGEERCP), disease duration until index ERCP, liver biopsies, total follow-up (FU) time, follow-up from index ERCP, clinical status at end of follow-up, occurrence of cholangiocarcinoma (CCA), and orthotopic liver transplantation (OLT). Follow-up data were extracted from the charts, and obtained from the attending clinicians, the primary care physician, the patient, or from one of the liver transplantation centres in the Netherlands.

Natural history assessment

Cumulative survival from time of diagnosis to death from liver disease or OLT was estimated by Kaplan-Meier analysis. The same method was applied to assess the survival of patients with a cholangiocarcinoma from time of diagnosis of this complication to death.

Cholangiographic scoring

All available cholangiograms were retrieved and scored by two investigators applying the scoring system proposed by Majoie et al., see table 1. We introduced one extra type 0 score for both the extrahepatic and intrahepatic compartments, to provide for the occurrence of no apparent visible abnormalities. Survival curves were computed to assess whether increasing intra- and extrahepatic scorings were correlated with poorer prognosis.

Prognostic model computation

A variable based on the cholangiographic scoring system was delineated using Kaplan-Meier and univariate Cox proportional-hazards analysis. Subsequently, using
backward regression, 'age at index ERCP', 'duration of disease until index ERCP', sex, concurrent IBD, and cholecystectomy were introduced in the regression analysis as possible covariates. The assumption of proportionality for a Cox regression model fit was checked as described previously (17). For statistical calculations the SPSS package version 8.0 (SPSS inc., Chicago, IL, USA) and S-plus version 4.5 (MathSoft, Seattle, WA, USA) were used.

Table 1.
Classification of cholangiographic findings in PSC

<table>
<thead>
<tr>
<th>type of duct involvement/classification</th>
<th>cholangiographic abnormalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrahepatic (IHD)</td>
<td>no visible abnormalities</td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Multiple strictures; normal calibre of bile ducts or minimal dilatation</td>
</tr>
<tr>
<td>II</td>
<td>Multiple strictures, saccular dilatations, decreased arborization</td>
</tr>
<tr>
<td>III</td>
<td>Only central branches filled despite adequate filling pressure; severe pruning</td>
</tr>
<tr>
<td>Extrahepatic (EHD)</td>
<td>no visible abnormalities</td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>slight irregularities of duct contour; no stricture</td>
</tr>
<tr>
<td>II</td>
<td>segmental stricture</td>
</tr>
<tr>
<td>III</td>
<td>stricture of almost entire length of duct</td>
</tr>
<tr>
<td>IV</td>
<td>extremely irregular margin; diverticulumlile outpouchings</td>
</tr>
</tbody>
</table>

Reproduced after modification with permission from (16).

RESULTS

General characteristics
Among the 174 patients 105 (60 %) were male. Their mean age at time of diagnosis was 40.4 ± 14.8 year. One hundred fourteen patients (66 %) were known with concurrent IBD ( ulcerative colitis n=83, Crohn’s disease n=28, indeterminate colitis n=3 ). Thirty-nine patients (22 %) underwent cholecystectomy, which not infrequently revealed the diagnosis of PSC. The ABO blood type distribution closely matched the normal distribution in the Dutch population (18). The general characteristics are summarised in table 2.
Table 2.
General characteristics of PSC cohort (n=174)

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>105</td>
<td>60</td>
</tr>
<tr>
<td>Inflammatory bowel disease</td>
<td>114</td>
<td>66</td>
</tr>
<tr>
<td>UC</td>
<td>83</td>
<td>48</td>
</tr>
<tr>
<td>CD</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>Indeterminate</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Cholecystectomy</td>
<td>9</td>
<td>22</td>
</tr>
<tr>
<td>ABO type n = 114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>53</td>
<td>46.5</td>
</tr>
<tr>
<td>A</td>
<td>47</td>
<td>41.2</td>
</tr>
<tr>
<td>B</td>
<td>11</td>
<td>9.6</td>
</tr>
<tr>
<td>AB</td>
<td>3</td>
<td>2.6</td>
</tr>
</tbody>
</table>

Natural history assessment

Follow-up data were complete within 12 months from data entry closure, which was in July 1999, in 161 cases. The median follow-up time was 76 months, range 1-300 months. During the follow-up there were 32 deaths (18.4 %), six of which were non-liver disease related. An event was defined as 1) death from liver disease or 2) OLT. The estimated median survival from time of diagnosis until death from liver disease (n = 26) or OLT (n = 14) was 18 years (95 % CI: 15-21 years). The Kaplan-Meier curve is shown in figure 1. The 14 transplanted patients (8%) underwent OLT after a median disease duration of 95 months (range 2-221 months). At a median of 44 months (range 5-117 months) after OLT, all were alive. Cholangiocarcinoma occurred in 18 patients (10 %), which accounted for 14 (54 %) of the liver disease related deaths. Biliary obstruction caused by a secondary cholangiocarcinoma was the initial mode of presentation in 6 cases. The other 12 patients developed CCA a median time of 27 months (range 9-258 months) after the diagnosis of PSC was made. One patient was lost to follow-up after a diagnosis of CCA was made. The estimated median survival in the remaining 17 patients was 9 months (95 % CI: 7-11 months). The survival curve is depicted in figure 2.
Survival analysis

Fig 1. Kaplan-Meier plot + 95% CI (----) of cumulative survival of the entire PSC cohort (n = 174).

Estimated median survival until death from liver disease or OLT was 18 years, 95% CI: 15-21 years.
Chapter VIII

Fig 2. Kaplan-Meier plot of cumulative survival of patients (n=17) after a diagnosis of cholangiocarcinoma (CCA) was made. Estimated median survival = 9 months, 95% CI: 7-11 months.

Cholangiographic scoring
From 133 patients ERCPs were available, either initial (n=78) or within 2-257 months after diagnosis. Mean age at the time of ERCP was 42.6 ± 14.0 years. There were 83 males (62%) in this subgroup. Figures 3 and 4 show two typical examples of the cholangiographic spectrum of PSC.

The cumulative survival from time of index ERCP for the various intrahepatic- and extrahepatic types is displayed in figures 5 and 6, respectively. As can be seen in figure 6, the differences in survival between EHD stage 0 and I, and EHD stage III and IV were negligible and not significant (p=0.91 and p=0.58 respectively). From a morphological standpoint, these differences can also be very subtle. Therefore, stage 0 and I were combined to stage I', and stage III and IV were grouped into stage III'.
Fig. 3  Cholangiography showing type I intrahepatic, and type 0 extrahepatic
Fig 4. Cholangiography showing type II intrahepatic, and type III extrahepatic
Survival analysis

Fig 5. Survival curves for intrahepatic types 0-III.
Type 0: n=9; type I: n=66; type II: n=40; type III: n=18

Fig 6. Survival curves for extrahepatic types 0-IV
Type 0: n=13; type I: n=22; type II: n=70; type III: n=17; type IV: n=10
Chapter VIII

Computing a prognostic model

The assumption was made that both intra- and extrahepatic abnormalities would reflect disease severity. Hence, both intra- and extrahepatic scoring would be expected to play a role in determining prognosis. Extensive univariate Cox proportional-hazards analysis yielded a single parameter that had the strongest association with survival. This categorical parameter (SUMIHDEHD') was composed of the sum of the intrahepatic scoring (IHD) and the modified extrahepatic scoring (EHD'). The range of possible values for SUMIHDEHD' was 1-6. Since the difference of the survival curves for the values 3 and 4 was almost nil (p = 0.91), these values were taken together. The resulting categorical parameter SUMIHDEHD" contained 5 possible values (1-5), and had a highly significant overall log likelihood (p = 0.0008). Table 3 lists the resultant SUMIHDEHD" score as a function of IHD and EHD.

<table>
<thead>
<tr>
<th>IHD</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>EHD</td>
<td>I</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

In order to assess the influence of other variables such as ‘age at index ERCP’ (AGEERCP), ‘duration of disease until index ERCP’, gender, concurrent IBD, and cholecystectomy, backward regression was done including these variables in the model. Only AGEERCP moderately improved the significance of the overall likelihood (p = 0.0003). However, independent contribution to the predictive power of the model was only borderline significant (p = 0.06).

The assumption of proportionality for a Cox regression model fit was tested and found valid (p = 0.84).

Figure 7 shows the predicted survival curves for the three categories of SUMIHDEHD" present in the cohort, along with the Kaplan-Meier plots of the
actual observed survival in these categories. The curves were truncated at 15 years to avoid misinterpretation of the increasingly unreliable right-hand side. The prognostic index (PI) based on the survival function was as follows:

$$PI = 1.13X_3 + 1.98X_4 + 0.024Y$$

Where

- $X_3 = 0$ and $X_4 = 0$ if $SUMIHDEHD'' = 2$
- $X_3 = 1$ and $X_4 = 0$ if $SUMIHDEHD'' = 3$
- $X_3 = 0$ and $X_4 = 1$ if $SUMIHDEHD'' = 4$
- $Y = AGEERCP$ in years

(SUMIHDEHD'' score 1 and 5 not present in the cohort)

Fig7. Actual observed survival and predicted survival by the Cox model for the various categories of SUMIHDEHD'' at mean of covariate AGEERCP

Higher scores were associated with considerably poorer prognosis. The relative risk at death from liver disease or OLT of SUMIHDEHD'' score 3 vs. score 2 was 3.1 (95% CI 0.7-13.4). The relative risk of SUMIHDEHD'' score 4 vs. score 2 was 7.3 (95% CI 1.5-34.7). Each 10-year increment of AGEERCP was associated with a relative risk of 1.3.
After calculation of the PI, the estimated median survival can be determined from the graph in figure 8. For example, the patient, whose cholangiogram is shown in figure 4, was at the time of the ERCP 48 years old. Her SUMIHDHEHD" score = $X$. Thus, her PI = 3.13. From figure 8 it can be read that her estimated median survival would be 80 months.

**DISCUSSION**

The results of this cohort study, which is the second largest (together with the Mayo clinic series) reported, containing the longest follow-up so far, show that the median survival appears to be considerably better than reported previously. There are three published, individual studies involving over 100 patients (10-12). They report cohorts mainly from the 1970s and 1980s from the USA, the UK, and Sweden. All three report a median survival of approximately 12 years. In contrast, we found a 50 % better median survival of 18 years. The reasons for this discrepancy with the earlier
series may be multifactorial. Selection bias may play a role. However, the general patient characteristics were rather similar in all four series. In the study from the Mayo Clinic, the date of referral was taken as time of diagnosis, which may have given rise to a considerable underestimation of total disease duration. Moreover, death from any cause was regarded as an endpoint. However, when the six non-liver related death causes in our series were included as endpoint in the survival analysis, the estimated median survival was 17.5 years, still much longer than in the Mayo cohort. The King’s college and Sweden studies used the same endpoints as in the present study. The Swedish study defined base-line time as the date of diagnosis, as we did. The British study defined base-line as the time of the first symptom consistent with -in retrospect- the diagnosis of PSC. Compared with our series, as well as with the Swedish cohort, this could give rise to a substantial increase in median survival time, inasmuch as parameters such as elevated liver function tests may precede the diagnosis of PSC by many years. Yet, the median survival in our series was 50 % longer than the median survival in the series from King’s college. Alternatively, the time frame studied may be of influence. The series from the USA, the UK, and Sweden contained cohorts mainly from the 1970s-1980s, while our cohort also spanned 1990-1999. Substantial therapeutic advances aimed at slowing down disease progression have not been made in the last decade, but perhaps the diagnosis is made earlier nowadays.

The prevalence of cholangiocarcinoma in the above-mentioned studies was 6-8 %. In our series the prevalence of secondary cholangiocarcinoma was somewhat worse at 10 %. This may be as a result of the longer follow-up in our series. The actual prevalence may even be somewhat higher, since not all patients who died were autopsied.

Several prognostic models have been proposed to predict the course of PSC and to tailor the timing for OLT (19, 20). Independent variables for these models are age, histological staging and serum bilirubin. However, hepatic histologic biopsy findings can be non-specific, with a considerable degree of sampling variability, because the disease may not be evenly spread throughout the liver, and the pathognomonic onion-skin lesion can easily be missed (21-23). Thus, for diagnosis liver biopsy is usually not done, and histological staging may be quite variable. Likewise, serum bilirubin level at presentation may be a rather fluctuating indicator of disease stage. For instance, when a patient presents with suppurative cholangitis caused by a bile-plugged distal dominant stricture, but otherwise patent biliary tree, the elevated serum bilirubin may soon revert to a normal level, when the obstruction is relieved.
The mainstay of diagnosis is cholangiography, so it seems obvious to develop predictors based on cholangiographic abnormalities. Two studies have been performed that assessed the prognostic value of cholangiography (13, 14). Both concluded that high-grade intrahepatic strictures indicated poor prognosis. However, grading of intrahepatic strictures in these studies was rather detailed and may be difficult to reproduce, because the degree of narrowing of strictures may vary considerably, depending on the amount of contrast injected in the biliary tree. The modified classification of cholangiographic findings of Majoie et al. uses a more qualitative appreciation of the abnormalities in the biliary tree in PSC patients (16). So far, this classification has not been evaluated clinically in terms of predicting prognosis. We recently evaluated this radiographic scoring system and found it well-applicable in the vast majority of PSC patients, providing it was supplemented by two type 0 categories, (unpublished data). Survival analysis for each individual category revealed that the difference between extrahepatic type 0 and I, as well as the difference between extrahepatic type III and IV, were negligible. On morphological grounds, the difference between stage 0 and I can indeed be very minor. Stage III and IV both imply serious involvement of almost the entire extrahepatic duct. Hence, it seemed prudent to combine stage 0 and I to category I', and stage III and IV to category III'. On the assumption that both the intrahepatic scoring and the extrahepatic scoring would have influence on disease severity and -consequently- survival, a variable was delineated representing both scorings. The resulting variable SUMIHDEHD' showed that there was a clear inverse correlation of combined morphologic scoring with survival. Moreover, a prognostic index could be computed on the basis of SUMIHDEHD'. The predicted survival plots for the various categories of SUMIHDEHD' closely matched the actual observed cumulative survival curves. However, the definitive value of our model can only be assessed after cross-validation.

PSC is nowadays regarded as an immune dysbalance disease, occurring in genetically susceptible hosts. Important candidate genes that may play a role in predisposing an individual to chronic inflammation are those of the human leukocyte antigen DR locus, situated at the short arm of chromosome 6. One of the simplest and oldest means to search for common genetic background features is determination of the ABO blood type. Information on ABO blood type was available in 70 % of patients in our series. There was no evidence of a prevailing particular ABO phenotype.

In conclusion: the survival of Dutch PSC patients is considerably better than reported
previously. The Amsterdam radiographic classification system reflects disease stage, and has the potential to serve as a predictor in determining prognosis. This may prove useful in patient counselling, stratifying patients in therapeutic trials, and patient selection and timing for liver transplantation.

ACKNOWLEDGEMENTS

The authors are indebted to Dr. A.W.M. van Milligen de Wit Red Cross Hospital, The Hague, and Dr. W. Bruins Slot, Spaarne Hospital, Haarlem, for providing patient’s data.

REFERENCES


