



## Quality of life in patients with primary biliary cholangitis: A cross-geographical comparison



Lorenzo Montali<sup>a</sup>, Andrea Gragnano<sup>a</sup>, Massimo Miglioretti<sup>a</sup>, Alessandra Frigerio<sup>a</sup>, Luca Vecchio<sup>a</sup>, Alessio Gerussi<sup>b,c</sup>, Laura Cristofori<sup>b,c</sup>, Vincenzo Ronca<sup>b,c</sup>, Daphne D'Amato<sup>b,c</sup>, Sarah Elizabeth O'Donnell<sup>b,c</sup>, Clara Mancuso<sup>b,c</sup>, Martina Lucà<sup>b,c</sup>, Minami Yagi<sup>d</sup>, Anna Reig<sup>e</sup>, Laura Jopson<sup>f</sup>, Sesé Pilar<sup>e</sup>, Dave Jones<sup>f</sup>, Albert Pares<sup>e</sup>, George Mells<sup>g</sup>, Atsushi Tanaka<sup>d</sup>, Marco Carbone<sup>b,c</sup>, Pietro Invernizzi<sup>b,c,\*</sup>

<sup>a</sup> Department of Psychology, University of Milan Bicocca, Milan, Italy

<sup>b</sup> Division of Gastroenterology, Department of Medicine and Surgery, University of Milan Bicocca, Milan, Italy

<sup>c</sup> European Reference Network on Hepatological Diseases (ERN RARE-LIVER), San Gerardo Hospital, Monza, Italy

<sup>d</sup> Department of Medicine, Teikyo University School of Medicine, Tokyo, Japan

<sup>e</sup> Liver Unit, Hospital Clínic, University of Barcelona, IDIBAPS, CIBERehd, Barcelona, Spain

<sup>f</sup> Institute of Cellular Medicine, Newcastle University, Newcastle Upon Tyne, UK

<sup>g</sup> Department of Medical Genetics, University of Cambridge, Cambridge, UK

### ARTICLE INFO

**Keywords:**  
Autoimmunity  
Liver  
Cholestasis  
Fatigue

### ABSTRACT

**Background & aims:** Several symptoms impair the quality of life (QoL) of patients with primary biliary cholangitis (PBC). They are reported to vary significantly in different countries. Aim of our study was to explore whether there is a geographical clustering that accounts for symptoms in PBC.

**Methods:** Data was analysed from four cohorts of PBC patients from the UK, Spain, Japan and Italy using the PBC-27 scale.

**Results:** Overall, 569 patients from four cohorts were identified, including 515 females (90.5%) with a mean age of 61 years. The analysis provided evidence for strict factorial invariance of the scale, a robust indicator of its validity for cross-cultural research. The mean of the fatigue domain of British patients was significantly greater than that of the Japanese ( $p < 0.05$ ), Italian ( $p < 0.05$ ), and Spanish patients ( $p < 0.001$ ). The mean of the cognitive domain after 54 years of age, was significantly greater in the British patients than in the Japanese ( $p < 0.05$ ) and Spanish patients ( $p < 0.01$ ). However, after 69 years of age, there were not significant differences between countries. The mean of the emotion domain after 54 years of age, was greater in the British than in the Spanish ( $p < 0.01$ ) and Italian patients ( $p < 0.01$ ).

**Conclusions:** Differences in the four countries concerning fatigue, cognitive and emotional dysfunction were found. The association of latitude and symptoms might provide new insights into the role of sun exposure, genetics and/or cultural component into disease phenotype in PBC.

### 1. Introduction

Several non-specific symptoms frequently impair the quality of life of Primary Biliary Cholangitis (PBC) patients [1,2]. Fatigue is one of the major determinants of life quality deterioration [3–6] albeit there is no clear relationship between fatigue severity and the illness progression [7]. Other common symptoms are pruritus, itch, cognitive impairment,

social and emotional dysfunctions [2,8,9].

The burden of impaired quality of life in different countries is not clear. A UK study on a large national cohort of PBC patients, showed that only 25% of patients experienced no symptoms and perceived good quality of life. In contrast, larger proportion experienced symptoms in multiple domains, resulting in severe deterioration of their quality of life [1]. Early studies in Italy showed that global mental status is not altered

\* Corresponding author. Division of Gastroenterology and Center for Autoimmune Liver Diseases, Department of Medicine and Surgery, University of Milano - Bicocca, Via Cadore 48, 20900, Monza, (MB), Italy.

E-mail address: [pietro.invernizzi@unimib.it](mailto:pietro.invernizzi@unimib.it) (P. Invernizzi).

<https://doi.org/10.1016/j.jtauto.2021.100081>

Received 30 December 2020; Accepted 30 December 2020

2589-9090/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

in PBC, but early changes in orientation and in personal memory are present in cirrhotic stage [10].

Some studies pointed out at geographical and cultural variations in symptom relevance and impact for PBC patients. For example, according to a recent study on a Polish sample [11] gender had no impact on HRQoL, contrasting evidences from the UK where men had significantly lower levels of fatigue than women. Another study conducted on a US cohort of 1032 patients [12], indicated that the QoL of patients with PBC in the United States was generally well preserved.

To assess PBC patients' quality of life, Jacoby and colleagues developed a specific questionnaire, called PBC-40 [13]. PBC-40 is a patient-derived instrument that investigates the impact of the illness in six domains: fatigue, pruritus, cognitive, social and other symptoms. The patients rate 40 items on a five-point scale, with a higher score indicating a worse HRQoL. PBC-40 was initially validated on a British sample and it was subsequently used in extensive studies and validated in different languages, proving to be valid and highly effective at quantifying PBC symptoms [1,4,7,11,14,15]. Montali and colleagues derived a short version of the PBC-40 scale, called PBC-27, in a study on Italian and Japanese patients [16]. The items of the PBC-27 are the same as those of the PBC-40, even if reduced in number, and the domains correspond to those of the original scale, with the addition of a domain called dryness.

Measurement of QoL in PBC can become a routine in clinical practice to favour a more patient sensitive approach, and researchers can more easily compare their results in a cross-cultural perspective. However, a prerequisite for cross-cultural comparisons is that the instrument activates similar perceptive, cognitive and interpretative processes in all the groups, so that the factorial structure of a scale is invariant across different nations [17]. The analysis of the cross-cultural invariance of the scale may be particularly interesting in the domain of quality of life, in which culture may play a relevant role in shaping perceptions and attitudes.

Within this perspective, our research had two main aims. First to assess the cultural invariance of the PBC-27 scale. The second aim was to compare the scale results in a cross-cultural perspective, to deepen our knowledge concerning differences in QoL for PBC patients.

## 2. Materials and methods

### 2.1. Study design

For this study we developed a consortium of four research groups, already studying PBC patients' quality of life in the UK, Spain, Japan and Italy. Each group provided datasets from its own Country, all consisting of patients with a diagnosis of PBC [18–21].

### 2.2. Study population

We defined PBC according to EASL guidelines [22]. For the cross-national study we used data from the *UK-PBC Research Cohort*, part of the UK-PBC project [18]; data from the *Liver Unit, Digestive Diseases Institute, Hospital Clínic, IDIBAPS, CIBEREhd, Barcelona, Spain* [19], data from the *Japan PBC Study Group* [20] and data from the *Italian PBC Study Group* [21].

The databases were harmonized to ensure that liver biochemistry, disease classification and risk factor definitions were comparable. The study was done in accordance with the Declaration of Helsinki and the principles of good clinical practice. In all cohorts, participants provided written informed consent. The protocol was approved by the local Ethical Committee.

### 2.3. Questionnaire

The PBC-27 consists of 27 items taken from the original PBC-40 scale, which are divided in 7 domains: Symptoms (3 items), Dryness (2 items), Itch (3 items), Fatigue (8 items), Cognitive (5 items), Emotional (3 items)

and social (3 items). PBC-40 was not used since its theoretical factor structure does not fit in the Italian and Japanese populations [16]. Domains including Symptoms, Itch, Dryness, Fatigue and Cognition refer to the last four weeks, with a 1 to 5-point scale, with 1 corresponding to "Never" and 5 to "Always". The other domains do not refer to a specific time, and patients are asked to indicate their preference on a 1 to 5-point scale, with 1 representing "Not at all" and 5 "Very much".

The PBC 27 has been initially validated in a Japanese and an Italian population [16], and then it proved to be valid in English-speaking patients [23]. More recently, the scale has been used in a study on a large cohort of Polish PBC patients, confirming that it performs well in assessing the extent of HRQoL impairment [11].

### 2.4. Statistical analysis

To validate the factorial structure of the PBC-27 we performed one confirmatory factor analyses (CFA) with maximum likelihood estimation from the covariance matrix of each Country following the model presented in Fig. 1. We considered the model fit to be acceptable when the Root Mean Square Error of Approximation (RMSEA)  $\leq .08$ , non-normed fit index (NNFI)  $\geq 0.95$ , confirmatory fit index (CFI)  $\geq 0.95$ , and index root-mean-square residuals (SRMR)  $\leq .07$  [24,25], all factor loading are statistically significant, and the composite reliability of every latent variable is higher than .60.

We tested the measurement invariance of the PBC-27 up to the level of strong factorial invariance for the four considered countries. This test involved the sequential examination of differently constrained multiple-group CFA models (MG-CFA), from the least constrained to the most constrained (configural, metric, and strong factorial invariance) [26]. We evaluated the deterioration of the model's fit with the difference in the CFI ( $\Delta$ CFI) between the less and the more constrained model. A  $\Delta$ CFI smaller than or equal to 0.01 indicates that the null hypothesis of invariance should not be rejected [27].

To test if PBC affects patients' quality of life differently across the countries, we performed ANCOVAs for each PBC-27 factor controlling for age and the clinical indicators of disease severity (Albumin, Alkaline phosphatase, Alanine aminotransferase and Bilirubin). In Ancova the F-test of significance is used to test each main and interaction effect, for the case of a single interval dependent and multiple ( $>2$ ) groups formed by a categorical independent. F is between groups variance divided by within-groups variance. If the computed p-value is small, then significant relationships exist. After ANCOVAs we performed post-hoc comparisons with p values adjusted using the Hochberg method. We reported Cohen's d for the significant differences. We tested ANCOVA's assumptions, and when severely violated, we performed alternative non-parametric tests (Kruskal-Wallis Rank Sum Test followed by Conover-Iman post-hoc test with p values adjusted using the Benjamini-Hochberg method). We used full information maximum-likelihood and multiple-imputation methods to deal with missing data. Statistical analysis was performed using R [28] and LISREL 8.80 [29].

## 3. Results

### 3.1. Characteristics of the sample

Overall, 569 patients from the four cohorts were identified, including 515 females (90.5%) with a mean age of 61 years (range 30–86). Table 1 presents the main clinical and laboratory characteristics of the patients. We found 3.5% of missing data overall. The variables with more missing value were the albumin and bilirubin with 13% (75) and 11% (64) of missing data, respectively. Missing values were considered as missing at random (MAR) after exploration of data.

### 3.2. Factorial structure

The CFAs were performed for each country with the structure

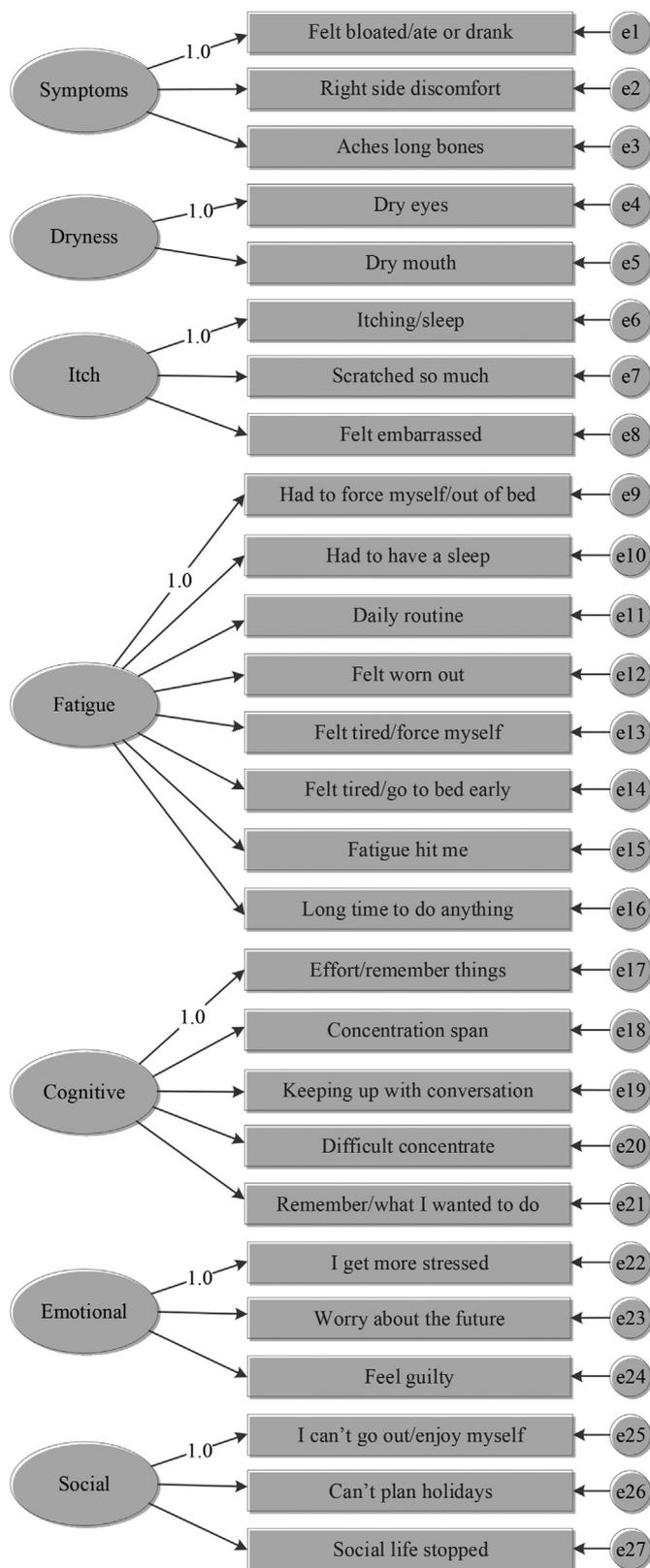


Fig. 1. Graphical counterpart of the equation of the model used in the CFAs for PBC-27 in the present study. The latent factors are free to co-vary.

presented in Fig. 1. The fit indexes showed that measurement model fit the data well for all the subgroups (the range of the fit indexes were:  $\chi^2_{(df)} = 367.2_{(303)}-479.94_{(303)}$ , RMSEA = 0.03–0.077, NNFI = 0.97–.99, SRMR = 0.052–0.67). Table 2 presents the within-group completely

standardized CFA solutions for each sample. All factor loading in each sample were statistically significant.

### 3.3. Measurement invariance

All the models showed acceptable fit indexes (Configural invariance  $\chi^2_{(df)} = 2442.553_{(1212)}$ , RMSEA = 0.067, NNFI = 0.98, SRMR = 0.053, CFI = 0.981; Metric invariance  $\chi^2_{(df)} = 2511.28_{(1272)}$ , RMSEA = 0.065, NNFI = 0.98, SRMR = 0.075, CFI = 0.981; Strong factorial invariance  $\chi^2_{(df)} = 2759.1_{(1332)}$ , RMSEA = 0.069, NNFI = 0.98, SRMR = 0.074, CFI = 0.978) and  $\Delta$ CFIs lower than .01. These results indicated negligible deterioration in the fit between adjacent nested models supporting strict factorial invariance for each factor of the questionnaire.

### 3.4. Differences of domains across countries

#### 3.4.1. Fatigue

Fatigue was most affected by the patient's country, after controlling for age and the clinical indicators,  $F_{(3, 404803)} = 12.12, p < .001$ . The significant covariates were age with a  $F_{(1, 16379)} = 6.97, p < .01$  and Albumin with  $F_{(1, 984)} = 10.71, p < .01$ . Post hoc test with Hochberg method revealed that fatigue of the British sample was the highest and significantly greater than that of the Japanese ( $t = 3.03, p < .05, d = 0.23$ ), Italian ( $t = 2.83, p < .05, d = 0.25$ ), and Spanish sample ( $t = 5.79, p < .001, d = 0.50$ ). Fatigue of the Spanish sample was the lowest and significantly lower than that of the Japanese ( $t = 2.53, p < .001, d = 0.32$ ), Italian ( $t = 3.58, p < .05, d = 0.26$ ), and, as stated before, the British sample (Table 3).

#### 3.4.2. Dryness

The variable Country has a significant effect on Dryness after controlling for age and the clinical indicators,  $F_{(3, 13334)} = 3.92, p < .01$ . The only significant covariate was Albumin  $F_{(1, 1743)} = 4.90, p < .05$ . Post hoc test with Hochberg method revealed that the only significant difference was between Spanish and Japanese sample ( $t = 3.17, p < .05, d = 0.28$ ) with the highest and lowest level of dryness, respectively (Table 3).

#### 3.4.3. Cognitive, emotion, other symptoms

Results concerning these three domains are presented together because for all of them in the ANCOVA the homogeneity of regression slopes assumption was violated for the covariate age. Hence, we decided to include the interaction of age with Country in our model. The interaction term was always significant, as well as the effect of countries at mean level of age.

To make country comparison and understand the effect of the interaction between Country and age, we performed post hoc test with Hochberg method at two levels of age, the 25th and 75th percentile of our sample corresponding to an age of 54 and 69 respectively (Table 4).

For the Cognitive domain, at 54 years of age, the adjusted mean of the British sample was the highest and significantly greater than that of the Japanese ( $t = 3.03, p < .05, d = 0.29$ ) and Spanish sample ( $t = 5.79, p < .01, d = 0.26$ ). At 69 years of age, there were not significant differences between countries. As the age grows, the adjusted level of Cognition decreases in British patients whereas it increases in Japanese and Italian patients (Fig. 2).

For the Emotion domain, at 54 years of age, the adjusted mean of the British sample was the highest and significantly greater than that of the Spanish ( $t = 3.72, p < .01, d = 0.32$ ) and Italian sample ( $t = 3.36, p < .01, d = 0.29$ ). At 69 years of age, the adjusted mean of the Spanish sample was the lowest and significantly lower than that of the British ( $t = 3.49, p < .01, d = 0.30$ ) and Japanese sample ( $t = 4.68, p < .001, d = 0.40$ ). As the age grows, Emotion decreases in British and Spanish patients, whereas it remains stable in Japanese Italian patients (Fig. 2).

For the Other symptoms domain, at 54 years of age, the adjusted mean of the Japanese sample was the lowest and significantly lower than that of the Italian ( $t = 5.27, p < .001, d = 0.46$ ), British ( $t = 5.04, p <$

**Table 1**  
Characteristics of the study population.

	English	Spanish	Japanese	Italian	All subjects (569)
Women	85	171	143	116	515
Age	65 (37–84)	59 (33–86)	61 (30–83)	62 (39–84)	61 (30–86)
Albumin median (IQR)	4.3 (0.5)	4.2 (0.3)	4.1 (0.4)	4.3 (0.5)	4.2 (0.4)
Alkaline phosphatase					
Alanine aminotransferase	37 (34)	34 (30)	25 (19)	34 (23.25)	31.5 (28)
Bilirubin	0.5 (0.4)	0.6 (0.4)	0.7 (0.4)	0.6 (0.3)	0.6 (0.365)

**Table 2**  
Composite Reliability and standardized factor loadings of the PBC-27.

	Italian	Japanese	Spanish	English
<i>Factor 1: symptoms - CR</i>	0.65	0.72	0.71	0.84
Felt bloated/ate or drank	0.44	0.74	0.63	0.72
Right side discomfort	0.62	0.57	0.73	0.84
Aches long bones	0.77	0.73	0.66	0.82
<i>Factor 2: dryness - CR</i>	0.63	0.75	0.77	0.85
Dry eyes	0.58	0.71	0.71	0.88
Dry mouth	0.77	0.83	0.86	0.85
<i>Factor 3: itch - CR</i>	0.81	0.92	0.89	0.95
Itching/sleep	0.68	0.85	0.84	0.94
Scratched so much	0.84	0.89	0.84	0.92
Felt embarrassed	0.8	0.92	0.89	0.94
<i>Factor 4: fatigue - CR</i>	0.92	0.93	0.94	0.96
Had to force myself/out of bed	0.68	0.71	0.64	0.83
Had to have a sleep	0.6	0.63	0.63	0.69
Daily routine	0.86	0.90	0.87	0.89
Felt worn out	0.85	0.90	0.90	0.94
Felt tired/force myself	0.88	0.90	0.93	0.94
Felt tired/go to bed early	0.73	0.75	0.84	0.85
Fatigue hit me	0.8	0.83	0.79	0.87
Long time to do anything	0.74	0.73	0.90	0.89
<i>Factor 5: cognitive - CR</i>	0.87	0.90	0.92	0.95
Effort/remember things	0.77	0.72	0.78	0.86
Concentration span	0.79	0.88	0.88	0.90
Keeping un with conversation	0.73	0.79	0.76	0.90
Difficult concentrate	0.77	0.89	0.91	0.91
Remember/what I wanted to do	0.75	0.76	0.82	0.86
<i>Factor 6: emotional - CR</i>	0.77	0.75	0.73	0.81
I get more stressed	0.85	0.81	0.76	0.86
Worry about the future	0.78	0.72	0.79	0.82
Feel guilty	0.53	0.58	0.49	0.61
<i>Factor 7: social - CR</i>	0.83	0.89	0.88	0.88
I can't go out/enjoy myself	0.81	0.82	0.84	0.83
Can't plan holidays	0.73	0.92	0.79	0.80
Social life stopped	0.81	0.82	0.89	0.89

Abbreviation: CR = Composite Reliability

.001,  $d = 0.43$ ), and Spanish sample ( $t = 5.54, p < .001, d = 0.48$ ). The same but less marked significant differences reply at 69 years of age. The Japanese sample was the lowest and significantly lower than that of the Italian ( $t = 4.56, p < .001, d = 0.42$ ), British ( $t = 2.65, p < .05, d = 0.24$ ), and Spanish sample ( $t = 2.84, p < .05, d = 0.26$ ). As the age grows, the adjusted level of Other symptoms decreases in British and Spanish patients, whereas it tends to increase in Italian and especially Japanese patients (Fig. 2).

Overall, it can be observed a consistent trend in the three domains: as the age grows, the perception of symptoms' severity decreases in the

**Table 3**  
Adjusted mean of Fatigue/Dryness and group comparison with Hochberg method.

Fatigue				Dryness			
Group	Adjusted mean	SD	Grouping	Group	Adjusted mean	SD	Grouping
British	2.76	0.102	A	Spanish	3.10	0.096	A
Japanese	2.43	0.079	B	British	2.72	0.135	A, B
Italian	2.39	0.088	B	Italian	2.79	0.114	A, B
Spanish	2.04	0.073	C	Japanese	2.65	0.103	B

Note: Means that do not share a letter are significantly different.

British and Spanish sample, while it increases in the Japanese and Italian sample.

### 3.4.4. Itch and Social

The ANCOVAs for the Itch and Social domains showed severe violations of the assumptions. We therefore performed Kruskal-Wallis Rank Sum Tests without covariates. There was a significant difference between countries in the level of Itch ( $H_{(df)} = 28.33_{(3)}, p < .001$ ) and Social ( $H_{(df)} = 65.16_{(3)}, p < .001$ ). Fig. 3 shows the distribution of the level of Itch and Social with the results of the post hoc comparisons. British and Japanese patients reported comparable levels of itching problems, significantly higher than Spanish and Italian patients ( $p < .05$ ). Spanish patients reported the highest impact on the Social domain ( $p < .001$ ) and British patients the lowest ( $p < .005$ ), with Japanese and Italian patients reporting an intermediate level compared to the other two groups.

## 4. Discussion

This study validates the PBC-27 as an accurate tool for measuring QoL in patients with PBC across different geographical areas; furthermore,

**Table 4**  
Adjusted mean of Cognitive, Emotional and Other Symptoms and group comparison with Hochberg method at two different ages.

	age 54			age 69		
	Adjusted mean at	SD	Grouping	Adjusted mean at	SD	Grouping
Cognitive						
British	2.45	0.145	A	2.10	0.105	A
Italian	2.03	0.105	A, B	2.19	0.98	A
Spanish	1.96	0.075	B	1.96	0.089	A
Japanese	1.86	0.088	B	2.15	0.095	A
Emotional						
British	2.98	0.158	A	2.43	0.115	A
Japanese	2.56	0.096	A, B	2.57	0.102	A
Italian	2.33	0.116	B	2.29	0.107	A, B
Spanish	2.33	0.082	B	1.91	0.098	B
Symptom						
British	2.68	0.147	A	2.39	0.107	A
Italian	2.54	0.109	A	2.65	0.100	A
Spanish	2.46	0.077	A	2.38	0.091	A
Japanese	1.80	0.090	B	2.01	0.097	B

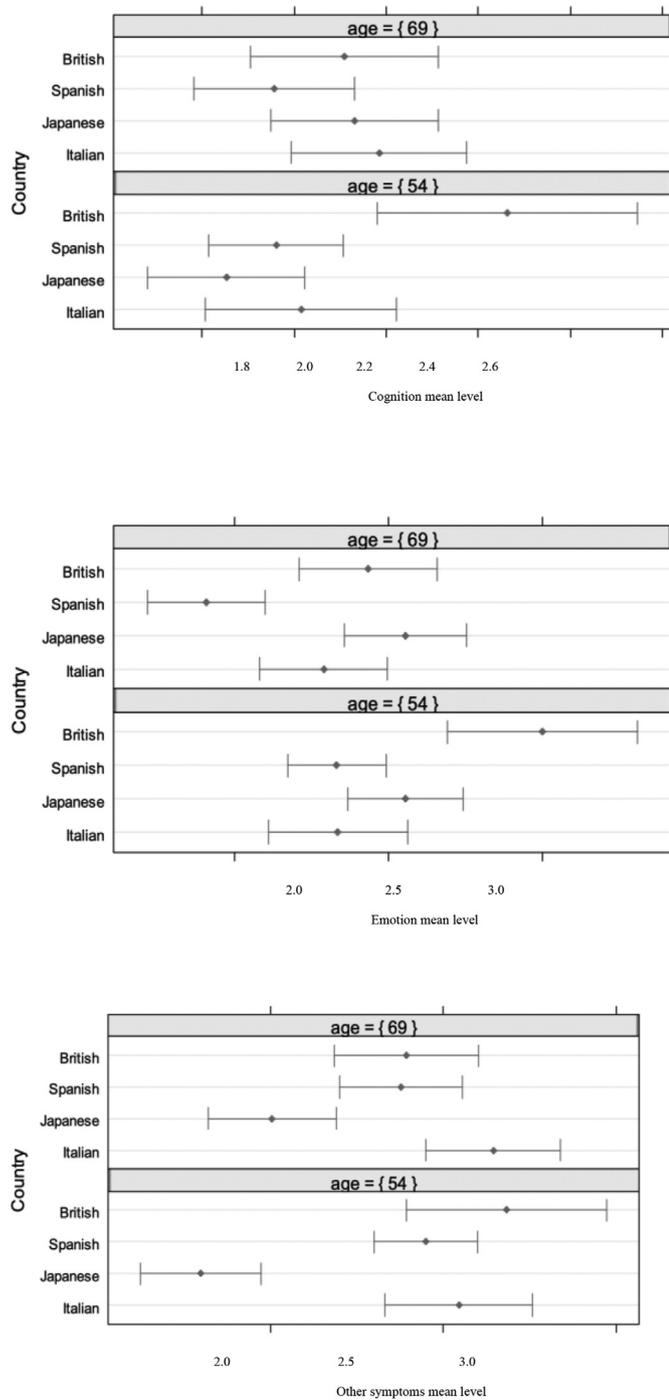


Fig. 2. Estimated means for Cognition (above), Emotion (in the middle) and Other symptoms (below) in the four countries at 69 and 54 years of age adjusted for albumin, alkaline phosphatase, alanine aminotransferase and bilirubin.

QoL measures concerning fatigue, cognitive and emotional dysfunction significantly differed in the four countries analysed.

Evidence for strict factorial invariance of the PBC 27 scale in the four cohorts of patients with PBC is provided. This implies that an unbiased comparison of factor means can be performed, i.e. the scale items reflect equivalent underlying constructs and the response scales are used likewise across the different cultures. This conclusion was a robust indicator of the validity of the PBC-27 for cross-cultural research, and its importance may be fully appreciated considering the paucity of studies that have examined the factorial invariance of HRQoL measures [30].

Symptoms of PBC significantly changed across different geographical

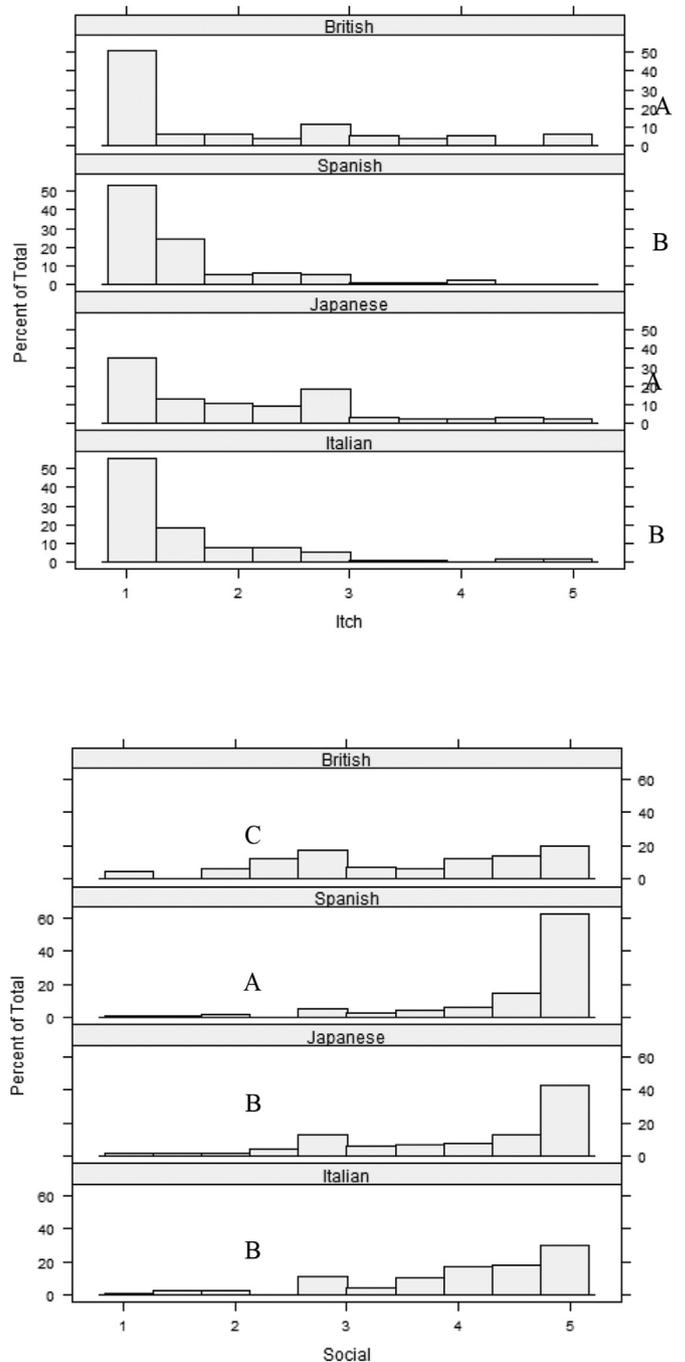


Fig. 3. Distribution of the level of Itch (left) and Social (right) with the results of the post hoc comparisons. Note: groups that do not share a letter are significantly different.

areas. As per fatigue, which is considered the most common and debilitating symptom in PBC, our results indicate that the British study participants had a significantly higher perception of fatigue than the other groups, whereas the Spanish had the lowest. Regarding dryness, a significant difference was observed between the Spanish cohort, which perceived the highest level of this symptom, and the Japanese one, which perceived the lowest one. At the country level, other differences emerged in the Social and Itch domains: Spanish patients perceived a significantly more relevant impact of the illness on their social activities compared to the other groups, whereas in the Itch domain, Japanese and British patients were the most significantly affected. Itching is a symptom strictly related to PBC severity. Our data shows a pattern in the impact of itch and

fatigue in the four groups compared in this study. Indeed, the order of the countries with patients most affected by Itch and Fatigue is the same (i.e., British, Japanese, Italian, Spanish). This pattern indicates that the impact of a subjective symptom such as fatigue assessed in this study is corroborated by its relationship with the more objective symptom of Itch. However, caution is required because we observed a marked floor and ceiling effect (non-normal distribution of answers with most of the patients choosing the same lowest, "floor", or highest, "ceiling", answer category) for Itch and Social, respectively. Non-normality prevented us from controlling the group effect for any covariate, limiting the value of our data.

From a transcultural perspective, another conclusion of our study is that PBC patients' age has a different impact on symptoms' perception in different countries. Indeed, at 54 years of age, i.e., the median age at diagnosis, the British cohort perceived a significantly more severe cognitive impairment than the Spanish and the Japanese sample, but this difference disappeared at 69 years of age. A similar pattern was identified in the Emotional domain: at 54 years of age, the British sample perceived a significantly more severe emotional impairment than the Spanish and the Italian sample. However, in this domain a difference at 69 years of age was also observed: the Spanish sample perceived significantly less emotional impairment than the British and Japanese. Finally, in the Other symptoms domain, our results indicate a significant difference both at 54 and 69 years of age: these symptoms have the lowest impact on the quality of life of the Japanese sample compared to the other three groups. Overall, these results indicate that the oldest patients' quality of life tends to be more similar in the four groups, whereas a more marked difference is present in the youngest group of patients. Our results also show a consistent trend that differentiates the four countries in relation to the effect of age on Cognitive, Emotional and Other symptoms. As the age grows, the perception of symptoms' severity decreases in the British and Spanish sample, while it increases in the Japanese and Italian sample.

Concerning clinical variables, our results indicate that only lower albumin seems to impact on our patients' quality of life. Indeed, albumin was the only significant covariate related to Fatigue, Dryness and Other symptoms. This may be related to the fact that lower albumin levels are a surrogate biochemical marker of advanced cirrhosis. On the contrary, alkaline phosphatase, alanine aminotransferase and bilirubin, biochemical markers of disease activity, had no significant effect on any PBC-27 factors. This is consistent with the known lack of relationship between fatigue severity and the illness progression [7]. Subclinical hepatic encephalopathy could also play a relevant role fatigue in PBC patients [31].

Whether sun exposure might play a role to explain the lower fatigue at lower latitudes, it is worth mentioning that there is evidence going in the opposite direction. Indeed, in a previous study by Björnsson et al., fatigue levels in patients from Sweden and the UK were comparable to those from Spain and Italy found in the current study [32]. While comparisons are difficult, this discrepancy should reinforce the need for future studies validating these findings possibly with the aim of putting together both genetic and environmental factors as covariate factors.

In conclusion, this is the first study that explores cross-sectionally the QoL of PBC patients concerning fatigue, cognitive and emotional dysfunction in different geographical areas. The association of latitude and symptoms might provide new insights into the role of sun exposure, genetics and/or cultural component into disease phenotype in PBC.

The authors declare no conflicts of interest.

## Financial Support

We acknowledge that this research was partially supported by the Italian Ministry of University and Research (MIUR) - Department of Excellence project PREMIA (PRECision Medicine Approach: bringing biomarker research to clinic).

## Data availability statement

The data that support the findings of this study are available on request from the corresponding author.

## Author contributions

Designed and supervised the project: ML, CMar, IP.

Acquisition of data: all authors.

Data analysis: ML, GA, MM, FA, LV.

Interpretation of data and drafting of the manuscript: ML, GA, CMar, IP.

Revision of the manuscript for important intellectual content: all authors.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

Alessio Gerussi, Laura Cristoferi, Vincenzo Ronca, Dave Jones, Albert Pares, Marco Carbone and Pietro Invernizzi are members of the European Reference Network on Hepatological Diseases (ERN RARE LIVER). The authors thank AMAF Monza ONLUS and AIRCS for the unrestricted research funding.

## References

- [1] G.F. Mells, G. Pells, J.L. Newton, A.J. Bathgate, A.K. Burroughs, M.A. Heneghan, et al., Impact of primary biliary cirrhosis on perceived quality of life: the UK-PBC national study, *Hepatology* 58 (2013) 273–283, <https://doi.org/10.1002/hep.26365>.
- [2] J.L. Newton, K.G. Hollingsworth, R. Taylor, A.M. El-Sharkawy, Z.U. Khan, R. Pearce, et al., Cognitive impairment in primary biliary cirrhosis: symptom impact and potential etiology, *Hepatology* (2008), <https://doi.org/10.1002/hep.22371>.
- [3] E. Björnsson, E. Kalaitzakis, M. Neuhauser, F. Enders, H. Maetzel, R.W. Chapman, et al., Fatigue measurements in patients with primary biliary cirrhosis and the risk of mortality during follow-up, *Liver Int.* (2010), <https://doi.org/10.1111/j.1478-3231.2009.02160.x>.
- [4] J.L. Newton, G.J. Gibson, M. Tomlinson, K. Wilton, D. Jones, Fatigue in primary biliary cirrhosis is associated with excessive daytime somnolence, *Hepatology* (2006), <https://doi.org/10.1002/hep.21230>.
- [5] J. Goldblatt, P.J.S. Taylor, T. Lipman, M.I. Prince, A. Baragiotta, M.F. Bassendine, et al., The true impact of fatigue in primary biliary cirrhosis: a population study, *Gastroenterology* (2002), <https://doi.org/10.1053/gast.2002.32993>.
- [6] K. Cauch-Dudek, S. Abbey, D.E. Stewart, E.J. Heathcote, Fatigue in primary biliary cirrhosis, *Gut* (1998), <https://doi.org/10.1136/gut.43.5.705>.
- [7] N. Al-Harthy, T. Kumagi, C. Coltescu, G.M. Hirschfield, The specificity of fatigue in primary biliary cirrhosis: evaluation of a large clinic practice, *Hepatology* (2010), <https://doi.org/10.1002/hep.23683>.
- [8] R.E. Poupon, Y. Chrétien, O. Chazouillères, R. Poupon, J. Chwalow, Quality of life in patients with primary biliary cirrhosis, *Hepatology* (2004), <https://doi.org/10.1002/hep.20276>.
- [9] E.A. Jones, N.V. Bergasa, The pathogenesis and treatment of pruritus and fatigue in patients with PBC, *Eur. J. Gastroenterol. Hepatol.* (1999), <https://doi.org/10.1097/00042737-199906000-00007>.
- [10] A. Floreani, M. Marchiori, S. Bonato, M. Zucchetto, R. Naccarato, M. Chiaramonte, Cognitive assessment in primary biliary cirrhosis: a case-control study, *Am. J. Gastroenterol.* (1995).
- [11] J. Raszęja-Wyszomirska, E. Wunsch, M. Krawczyk, E.I. Rigopoulou, K. Kostrzewa, G.L. Norman, et al., Assessment of health related quality of life in Polish patients with primary biliary cirrhosis, *Clin Res Hepatol Gastroenterol* (2016), <https://doi.org/10.1016/j.clinre.2015.10.006>.
- [12] C. Selmi, M.E. Gershwin, K.D. Lindor, H.J. Worman, E.B. Gold, M. Watnik, et al., Quality of life and everyday activities in patients with primary biliary cirrhosis, *Hepatology* (2007), <https://doi.org/10.1002/hep.21953>.
- [13] A. Jacoby, A. Rannard, D. Buck, N. Bhala, J.L. Newton, O.F.W. James, et al., Development, validation, and evaluation of the PBC-40, a disease specific health related quality of life measure for primary biliary cirrhosis, *Gut* 54 (2005) 1622–1629, <https://doi.org/10.1136/gut.2005.065862>.
- [14] S. Zec, D. Popovic, V. Matovic, V. Nikolic, K. Bojovic, J. Jovic, et al., Translation and validation of the Serbian primary biliary cholangitis-40 questionnaire, *PloS One* (2017), <https://doi.org/10.1371/journal.pone.0175697>.

- [15] P. Blackburn, M. Freeston, C.R. Baker, D.E.J. Jones, J.L. Newton, The role of psychological factors in the fatigue of primary biliary cirrhosis, *Liver Int.* (2007), <https://doi.org/10.1111/j.1478-3231.2007.01500.x>.
- [16] L. Montali, A. Tanaka, P. Riva, H. Takahashi, C. Cocchi, Y. Ueno, et al., A short version of a HRQoL questionnaire for Italian and Japanese patients with Primary Biliary Cirrhosis, *Dig. Liver Dis.* (2010), <https://doi.org/10.1016/j.dld.2010.01.004>.
- [17] J. He, F. van de Vijver, Bias and equivalence in cross-cultural research, *Online Readings Psychol Cult* (2012), <https://doi.org/10.9707/2307-0919.1111>.
- [18] M. Carbone, S.J. Sharp, S. Flack, D. Paximadas, K. Spiess, C. Adgey, et al., The UK-PBC risk scores: derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cholangitis, *Hepatology* 63 (2016) 930–950, <https://doi.org/10.1002/hep.28017>.
- [19] A. Parés, L. Caballería, J. Rodés, Excellent long-term survival in patients with primary biliary cirrhosis and biochemical response to ursodeoxycholic acid, *Gastroenterology* 130 (2006) 715–720, <https://doi.org/10.1053/j.gastro.2005.12.029>.
- [20] A. Honda, A. Tanaka, T. Kaneko, A. Komori, M. Abe, M. Inao, et al., Bezafibrate improves GLOBE and UK-PBC scores and long-term outcomes in patients with primary biliary cholangitis, *Hepatology* 70 (2019) 2035–2046, <https://doi.org/10.1002/hep.30552>.
- [21] M. Carbone, A. Nardi, S. Flack, G. Carpino, N. Varvaropoulou, C. Gavrilá, et al., Pretreatment prediction of response to ursodeoxycholic acid in primary biliary cholangitis: development and validation of the UDCA Response Score, *Lancet Gastroenterol Hepatol* 1253 (2018) 1–9, [https://doi.org/10.1016/S2468-1253\(18\)30163-8](https://doi.org/10.1016/S2468-1253(18)30163-8).
- [22] EASL Clinical Practice Guidelines, The diagnosis and management of patients with primary biliary cholangitis, *J. Hepatol.* 67 (2017) 145–172, <https://doi.org/10.1016/J.JHEP.2017.03.022>.
- [23] D. Jones, I. Patanwala, J. Newton, Comparison of the PBC-40 and PBC-27 tools for quality of life assessment in English speaking primary biliary cirrhosis patients, *Gut* (2011), <https://doi.org/10.1136/gut.2011.239301.516>.
- [24] R.P. Bagozzi, Y. Yi, Specification, evaluation, and interpretation of structural equation models, *J. Acad. Market. Sci.* (2012), <https://doi.org/10.1007/s11747-011-0278-x>.
- [25] K. Lai, S.B. Green, The problem with having two watches: assessment of fit when RMSEA and CFI disagree, *Multivariate Behav. Res.* (2016), <https://doi.org/10.1080/00273171.2015.1134306>.
- [26] R.J. Vandenberg, C.E. Lance, A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research, *Organ. Res. Methods* (2000), <https://doi.org/10.1177/109442810031002>.
- [27] G.W. Cheung, R.B. Rensvold, Evaluating goodness-of-fit indexes for testing measurement invariance, *Struct. Equ. Model.* (2002), [https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5).
- [28] R.C.R. Team, *A Language and Environment for Statistical Computing*, R Found Stat Comput, 2016.
- [29] K.G. Joreskog, D. Sorbom, *Lisrel 8.8: interactive LISREL: technical support*, Mooresville, Sci Softw (2006).
- [30] D.A. Newman, C.A. Limbers, J.W. Varni, Factorial invariance of child self-report across English and Spanish language groups in a hispanic population utilizing the pedsQL™ 4.0 generic core scales, *Eur. J. Psychol. Assess.* (2010), <https://doi.org/10.1027/1015-5759/a000026>.
- [31] E. Kalaitzakis, E. Björnsson, Fatigue and liver transplantation in patients with primary biliary cirrhosis, *J. Hepatol.* 60 (2014) 1326–1328, <https://doi.org/10.1016/j.jhep.2014.01.023>.
- [32] E. Björnsson, M. Simren, R. Olsson, R.W. Chapman, Fatigue is not a specific symptom in patients with primary biliary cirrhosis, *Eur. J. Gastroenterol. Hepatol.* 17 (2005) 351–357, <https://doi.org/10.1097/00042737-200503000-00015>.

**Graphical Abstract: sunlight gradient exposition hypothesis for fatigue in PBC**

