Health-related quality of life questionnaires used in primary biliary cholangitis: a systematic review

Xin Ai[1], Xian Yang[1], Jia-min Xu[1], Wen-xia Yang[1], and Ying-Mei Tang[2]

[1]The Second Affiliated Hospital of Kunming Medical University
[2]The Second Affiliated Hospital of Kunming Medical University,Yunnan Research Center for Liver Diseases

November 30, 2020

Abstract

Objective The purpose of this systematic review was to assess the suitability of HRQOL questionnaires in patients with primary biliary cholangitis. Methods Five electronic databases were searched. The validity of translated questionnaires, floor and ceiling effects, internal consistency and test-retest reliability were investigated. Results Forty-four studies were included, of which fifteen HRQOL questionnaires were identified. The instruments used most frequently were the PBC-40 (n = 22), followed by the SF-36 (n = 19), PBC-27(n=4), CLDQ (n = 3) and NIDDK-QA(n=2), the remaining instruments were uesd only once. Tweenty-six studies used a translated HRQOL questionnaire and only six reported or referenced a validation of the translated questionnaire. Conclusions PBC-specific HRQOL questionnaires used in primary biliary cholangitis have generally good psychometric properties. But lots of studies directly applied the HRQOL tools without verifying the HRQOL tools validity and reliability in PBC patients. Thus, it is better for clinicians and researchers to test the measurement properties of HRQOL questionnaires before use it.

1

Key words: review, primary biliary cholangitis, health-related quality of life

*Corresponding author: Ying-mei Tang

mail address: tangyingmei_med@kmmu.edu.cn

Review criteria:

Inclusion criteria

1. Original studies of adult patients ([?]18 years old) with primary biliary cholangitis, studies reporting generic or disease-specific HRQOL questionnaires as outcome measure.

2. The study had to be a randomized controlled trial, cohort study, case control study, clinical trial, or validation study of HRQL instruments.

Message for the clinic:

PBC-specific HRQOL questionnaires used in primary biliary cholangitis have generally good psychometric properties. It is better for clinicians and researchers to test the measurement properties of HRQOL questionnaires before use it.

1.Introduction

Primary biliary cholangitis(PBC) is an autoimmune disease, which characterized by interlobular cholangitis with progressive destruction, cholestasis, fibrosis and may develop into end-stage liver disease eventually[1, 2]. High-titer specific anti-mitochondrial antibodies and anti-nuclear antibodies GP120 and SP100 appeared in serum of PBC patients, and the level of IgM in peripheral blood also increased correspondingly[3]. Fatigue and pruritus are the most common symptoms of primary biliary cirrhosis, which affects patients' quality of life seriously. Ursodeoxycholic acid(UDCA) is the first valid drug in the treatment of PBC, but 30% of patients with PBC do not have a biochemical response to UDCA[4, 5].The World Health Organization (WHO) defines quality of life as individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns[6]. And health-related quality of life (HRQOL) is defined as the influence of health and disease on quality of life[7]. Health-related quality of life (HRQOL) measures in chronic disease management is necessary as it assesses the overall impact on health from the patient's perspective. A range of different questionnaires are used to evaluate HRQOL in PBC, including disease-specific and generic questionnaires, but no guidance of HRQOL questionnaires properties for clinicians and researchers to choose. The aim of this review was to evaluate the suitability of questionnaires used to assess HRQOL in primary biliary cholangitis.

2.Materials and methods

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses

(PRISMA) statement guidelines for systematic reviews[8] to ensure a complete reporting in the development of the study protocol. This systematic review was performed in January 2020.

2.1 Data sources

We performed electronic searches in the following five databases: PubMed, the Cochrane Library, Web of Science, Wan Fang Database, and Chinese National Knowledge Infrastructure. The search was run on 17 January 2020.

2.2 Search strategy

The search strategy included a combination of the following relevant keywords: "health-related quality of life", "quality of life", "primary biliary cholangitis" and "primary biliary cirrhosis". Electronic database searches were limited to English-language publications.

2.3 Inclusion and exclusion criteria

Inclusion criteria

1. Original studies of adult patients ([?]18 years old) with primary biliary cholangitis, studies reporting generic or disease-specific HRQOL questionnaires as outcome measure.

2. The study had to be a randomized controlled trial, cohort study, case control study, clinical trial, or validation study of HRQL instruments.

3.Articles published between January 1999 and January 2020 (The publication period of 20 years considered for this systematic review is defined based on the reference list from Osborne et al.)[9].

Exclusion criteria

1. Studies reporting quality of life for primary sclerosing cholangitis, autoimmune hepatitis.

2. Case reports, conference abstracts, case series, and dissertations were excluded.

3.Using fatigue scales or pruritus scales to assess patients' quality of life.

4.Number of included cases<=10.

2.4 Data extraction

Two reviewers (X.Y. and J.-M.X.) assessed the full-text publications and extracted data using standardized electronic forms independently. Disagreement between the reviewers about eligibility was resolved through discussion. Data extracted from the selected studies including: author, year of publication, study design, aim of the study, study characteristics, intervention, HRQOL tools, results. For translated questionnaires, we evaluated floor and ceiling effects, missing data, internal consistency and test-retest reliability. Finally, we performed a quality analysis of the included studies following the Hawker's tool, which provides clear description of ratings, that is, "good," "fair," "poor" and "very poor," and has been designed to assess quality of studies covering a variety of research paradigms[10].

3.Results

3.1 Study selection

The search retrieved 1175 publications and 435 full-text records were obtained after duplicates removed. Studies were published between 1999.1 and 2020.1. 391 studies were excluded with reasons. Finally, 44 studies met the inclusion criteria. A PRISMA flow chart illustrates the studies selection process and reasons for exclusion (Figure 1).

3.2 Overview of included studies

A total of 15 HRQOL questionnaires were used in the included 44 studies, descriptions of the used questionnaires are presented in table 1. The instruments used most frequently were the PBC-40 (n = 21), followed by the SF-36 (n = 19), PBC-27(n=4), CLDQ (n = 3) and NIDDK-QA(n=2), the remaining instruments were uesd only once (Supplementary table 1~3). These instruments were initially designed for various purposes and referred to different concepts. The PBC-40 was the first questionnaire for use in patients with PBC specifically. The SF-36 is a standardized measure of health status, commonly used as a HRQOL questionnaire. The PBC-27 and the PBC-10 were brief vision of PBC-40. The CLDQ, CLDQ-I and LDSI2.0 were liver disease-specific questionnaires. The other questionnaires were developed to measure generic quality of life.

Validity of translated questionnaires

Tweenty-six studies used a translated HRQOL questionnaire and only six reported or referenced a validation of the translated questionnaire[11-16].

Floor/ceiling effects and missing data

Floor and ceiling effects were reported for only two HRQOL questionnaires (PBC-40 and PBC-10).There were no significant ceiling effects but a moderate floor effect was found for the itch domain in PBC-40. The PBC-10 did not have a significant ceiling effect but flooring effects were observed in all 10 items similarly[15, 17].

Internal consistency

Cronbach's α coefficients for the HRQOL questionnaires ranged from adequate to excellent. The PBC-40 had the highest internal consistency (Cronbach's α coefficient 0.72–0.95)[17]. PBC-27 Cronbach's α coefficients ranged from 0.45 to 0.93[13], NIDDK-QA from 0.87 to 0.94 and for PBC-10 it was 0.90[15, 16]. The internal consistency of other questionnaires was not reported.

Test-retest reliability

Test-retest reliability data were available for PBC-40, NIDDK-QA and PBC-10. The test-retest reliability of NIDDK-QA was measured by Pearson Correlation, it was ranged from 0.82 to 0.99. Test-retest reliability was calculated using intraclass correlation coefficients (ICCs) of PBC-40 and PBC-10. The PBC-40 ICCs ranged from 0.83 to 0.96 and for PBC-10 it was 0.945.
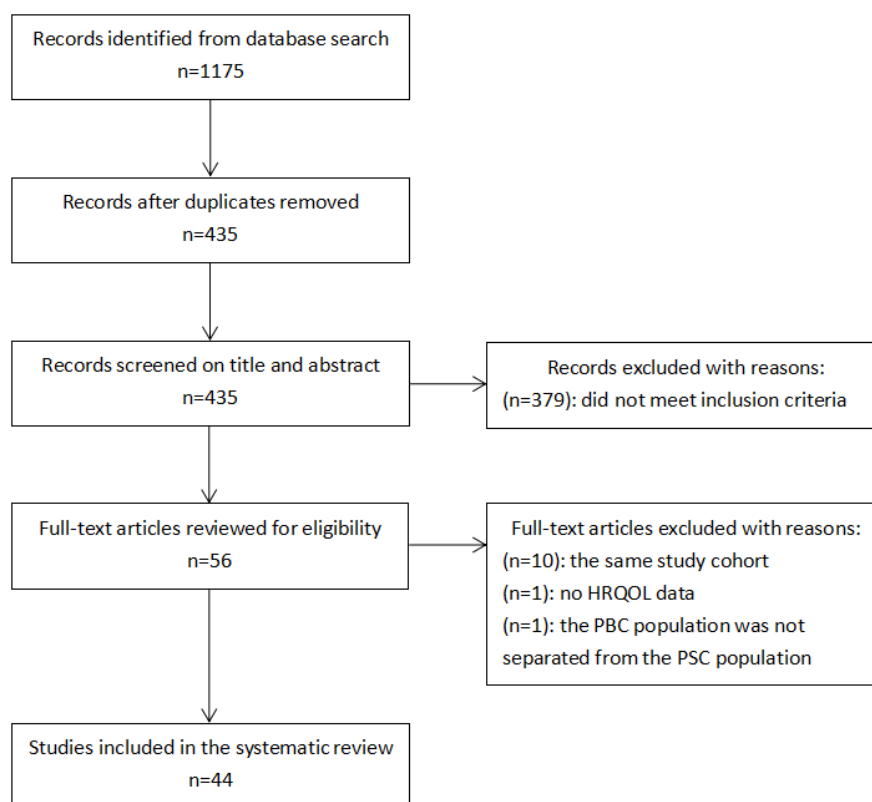

Figure 1 PRISMA flow diagram of the study selection process.

Table 1 Generic quality of life questionnaires and disease-specific quality of life questionnaires

Questionnaires	Scale (item)	Domains description	Scoring/administration
Disease-specific quality of life questionnaires			

Questionnaires	Scale (item)	Domains description	Scoring/administration
PBC-40[17]	6(40)	Fatigue, emotional, social, cognitive, itch and symptoms	5 point scale,1(not at all) to 5(very much) Total and subscale score Higher score indicates poorer health
PBC-27[12]	7(27)	Other Symptoms, dryness, itch, fatigue, cognitive, emotional and social	5 point scale,1(not at all) to 5(very much) Total and subscale score Higher score indicates poorer health
PBC-10[15]	6(10)	Fatigue, emotional, social, cognitive, itch and symptoms	5 point scale,1(not at all) to 5(very much) Total and subscale score Higher score indicates poorer health
CLDQ[18]	6(29)	Abdominal symptoms, fatigue, systemic symptoms, activity , emotional function, worry	7 point scale,1(not at all) to 7(very much) Total and subscale score Higher score indicates poorer health
CLDQ-I[19]	6(28)	Abdominal symptoms, fatigue, systemic symptoms, activity , emotional function, worry	6 point scale,1(not at all) to 7(very much) Total and subscale score Higher score indicates poorer health
LDSI 2.0[20]	18	Itch, joint pain, pain in the right upper abdomen, sleepiness during the day, worry about family situation, decreased appetite, depression, fear of complications and jaundice, nine other items measure the hindrance of these symptoms to daily activities	5 point scale,1(not at all) to 5(to a high extent) Subscale score Higher score indicates poorer health
Generic quality of life questionnaires			
SF-36[21]	8(36)	Physical functioning, physical role functioning, limitation emotional role limitation, bodily pain, mental health, social functioning, vitality and general health perception	100 point scale,0(not at all) to 100(very much) Total and subscale score Higher score indicates greater health
SF-12[22]	2(12)	Physical component and mental summary	100 point scale,0(not at all) to 100(very much) Total and subscale score Higher score indicates greater health

Questionnaires	Scale (item)	Domains description	Scoring/administration
15D(generic)[23]	15	Breathing, mental function, speech (communication), vision, mobility, usual activities, vitality, hearing, eating, elimination, sleeping, distress, discomfort and symptoms, sexual activity, and depression	5 point scale,1(not at all) to 5(very much) Total score Higher score indicates poorer health
NIDDK LTD-QOL[24]	5(21)	Measures of disease, psychological status, personal function, social and role function, and general health perception	5 point scale,0(not at all) to 4(extremely) Total and subscale score Higher score indicates poorer health
NIDDK-QA[16]	4(47)	Liver disease symptoms, physical functioning, health satisfaction and overall well-being	5 point scale,0(not at all) to 4(extreme) Total and subscale score Higher score indicates greater health
HUI-Mark-2[25]	7(15)	Sensation, mobility, emotion, cognition, selfcare, pain and fertility	2 point scale,0(not at all) to 1(very much) Total and subscale score Higher score indicates greater health
HUI-Mark-3[26]	8(15)	Vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain	2 point scale,0(not at all) to 1(very much) Total and subscale score Higher score indicates greater health
NHP[27]	6(38)	Energy, sleep, pain, emotional reactions, social isolation, and physical mobility	2 point scale,0(no) to 1(yes) Total and subscale score Higher score indicates poorer health
NHANES III (https://wwwn. 90 cdc.gov/nchs/nhanes/nhanes3/default.aspx)	4(83)	Articular symptoms, activity limitations (in both household and job-related activities), hobbies and leisure activities, and social life	Total and subscale score Higher score indicates poorer health
PROMIS-29(generic)[28]	7(29)	Physical function, depression, anxiety, fatigue, sleep disturbance, ability to participate in social roles and activities (social roles), pain interference, and one pain intensity item	5 point scale,1(not at all) to 5(very much) Total and subscale score Higher score indicates poorer health

Table 2 Summary of psychometric properties of the identified HRQOL questionnaires

Questionnaires	Validity	Reliability	Responsiveness and floor/ceiling effects	Minimal important difference
PBC-40	**Construct validity** Moderate to high correlations with the SF-36 and ESI-55 scales, except for "Physical functioning" domain of the SF-36 and "Itch" domain of the PBC-40[17]	**Internal consistency** Cronbach's alpha >0.7 for all scales[13] Cronbach's alpha >0.7 for all scales[14] Cronbach's alpha >0.7 for all scales[17] **Test-retest** ICC ranged 0.83-0.96[17]	**Floor and ceiling effects** There were no ceiling effects but there was a noticeable floor effect in the itch domain (36.7%)[15]	
PBC-27	**Convergent validity** minor correlations between the PBC-27 social factor and the SF-36 social functioning, but moderate to high correlations in majority of scales[12]	**Internal consistency** Cronbach's alpha >0.7 for all scales[12] Cronbach's alpha >0.7 for all scales, except for domains "Dryness", "Symptoms" and "Fatigue" [13]		
PBC-10	**Content validity** Generally good, developed by hepatology experts in PBC[15]	**Internal consistency** Internal consistency was excellent with Cronbach's alpha of 0.936. **Test-retest** ICC=0.945 for all scales[15]	**Floor and ceiling effects** There were no ceiling effects but a floor effect was obvious[15] responsiveness ratio of 0.38, effect size of 0.39 and standardised response mean of 1.25. **Responsiveness** Small to moderate responsiveness[15]	The minimal import- ant difference was 3.539[15]
NIDDK-QA	**Concurrent validity** Small to high correlations with the SF-36 scales, P <0 .01[16]	**Internal consistency** Cronbach's alpha ranged from 0.87 to 0.94 **Test-retest** Pearson coefficients: 0.82-0.99, P <0 .01[16]		

4.Discussion

This is the first systematic review of the suitability of HRQOL questionnaires used to assess primary biliary cholangitis, with the goal to provide evidence for clinicians and researchers concerning its use within clinical practice and research. We explored the suitability of these instruments through their psychometric properties and their content. There are 15 HRQOL questionnaires currently available.

First, this review found that a high degree of heterogeneity in HRQOL used instruments in the primary biliary cholangitis studies: fifteen different types of instruments were used in forty-four included articles. Among these forty-four included articles, only three articles were randomized and among non-randomized controlled trials, only eight assessed HRQOL as a primary endpoint. This finding shows that the relevance of HRQOL is not fully recognized in primary biliary cholangitis studies. If we consider observational studies, few studies were longitudinal; the others were cross-sectional with no possibility of exploring a change over time.

Many studies used translated HRQOL questionnaires. While there were good examples of translated validity in PBC-40, there were many studies where this procedure was either not conducted or reported. It is essential that HRQOL questionnaires are translated and validated using well recognized and standardized procedures to ensure that they are appropriately adapted to accommodate cultural differences. Since one's functional status, such as language barriers, quality of education, literacy level will have a direct effect on HRQOL scores. Strenthening investigators' ability to interpret results and formulating clinically relevant conclusions when using HRQOL questionnaire will ameliorate variability of scores. Inductive method was performed to assess psychometric properties of instruments in this review, minimal important difference, responsiveness and floor/ceiling effects, test-rest reliability, internal consistency (Cronbach's alpha) and various forms of construct validity (convergent, content and concurrent) were assessed. Except PBC-10, no other questionnaire provides complete psychometric validation in patients with PBC. Another common problem of most studies was that no sample size calculation was provided, which offers stronger implications on validity.

Our review highlights the questionnaire options available, reminding researchers should take into account psychometric properties and content when choose which HRQOL is the most appropriate to use in primary biliary cholangitis. The PBC-40 was the most widely studied questionnaire, with most extensive psychometric evaluation, which is the first disease-specific HRQOL assessment tool for PBC patients. Cronbach's α scores in all six domains were above 0.7, three of the domains (fatigue, cognitive, and itch) met the recommended minimum of 0.7 for individual patient comparisons. Test-retest reliability scores were above 0.8[17]. PBC-40 is a good choice for a research study if extensive experience of use of an HRQOL questionnaire is important. Its disadvantage is that it is too lengthy and time consuming to use in the normal clinical practice. The next most commonly used tool was the SF-36. Of generic quality of life questionnaires, the SF-36 performed best in terms of internal consistency and test-retest reliability. Though SF-36 provides a generic HRQOL assessment tool, it is less sensitive to impairments in HRQOL caused by pruritus and did not explore the quality of sleep and quality of wake domains[29, 30]. Briefer tools with simplified scoring for use in the clinical and research settings are needed urgently. The PBC-27 is one such example, which was shown to be equally effective in detecting the impact of PBC on HRQOL. However, the PBC-27 has potential language problems, since its original version was developed for use in Italian and Japanese patients with primary biliary cirrhosis[12]. Evaluations of PBC-27 in Poland were performed after translating it into Polish[31]. Thus, the PBC-27 requires testing of its measurement stability for the different language versions. Another short HRQL questionnaire, PBC-10, had good internal consistency (Cronbach's α 0.905) and test-retest reliability. PBC-10 demonstrated no ceiling effects but a floor effect was noted, which requires further study[15]. The CLDQ, a liver disease-specific HRQOL questionnaire, was developed to measure longitudinal change in individuals with chronic liver disease, although it didn't cover pruritus and fatigue domains in PBC patients well[32, 33]. The CLDQ-I(CLDQ Italy vision) and LDSI 2.0 are liver disease-specific HRQOL questionnaires. However, the LDSI 2.0 lacks comprehensiveness as a QOL instrument in terms of emotional well-being/emotional functioning and, if used, should be used in conjunction with other measures of well-

8

being. The SF-12, 15D, NIDDK LTD-QOL, NIDDK-QA, HUI-Mark-2, HUI-Mark-3, NHP, NHANES III and PROMIS-29 are generic HRQOL tools. The psychometric properties of the above tools has not been verified in PBC patients(except for NIDDK-QA),thus it is not recommended use them in PBC patients.

There are some limitations in our systematic review. First, PRISMA is just one methodology that can be used to synthesis or evaluate outcome measures and other methods might be equally valid or provide different perspectives. Secondly, we did not search the gray literature (inability to identify ongoing or unpublished studies), which produces selection bias. And our review was limited to English-language articles , some studies may not have been included due to this.

5.Conclusion

PBC-specific HRQOL questionnaires used in primary biliary cholangitis have generally good psychometric properties. But lots of studies directly applied the HRQOL tools without verifying the HRQOL tools validity and reliability in PBC patients. Investigators should select questionnaires for their study based on the ease of administration and the questionnaire's correlation with the primary health domain under investigation. Thus, more robust psychometric design studies to test the measurement properties of HRQOL questionnaires as the primary outcome of investigation are highly needed.

Conflict of interest

The authors have no conflict of interest to declare.

Contributions

X.A. contributed significantly to the conception and design of the study, data extraction, critical appraisal, interpretation of data and drafting of the manuscript. X.Y. and J.-M.X. were involved in literature search, critical appraisal and interpretation of data and drafting. W.-X.Y. was involved in critical appraisal and drafting. Y.-M.T. was involved in the drafting and review of the manuscript. All authors have given their final approval on the manuscript to be published.

Patient consent for publication

Not required.

Copyright permissions

All figures and tables were original.

References

[1] E.J. Carey, A.H. Ali, K.D. Lindor. Primary biliary cirrhosis. Lancet, 2015, 386: 1565-1575.

[2] F. de Liso, C. Matinato, M. Ronchi, R. Maiavacca. The diagnostic accuracy of biomarkers for diagnosis of primary biliary cholangitis (PBC) in anti-mitochondrial antibody (AMA)-negative PBC patients: a review of literature. Clin Chem Lab Med, 2017, 56: 25-31.

[3] J.K. Dyson, G.M. Hirschfield, D.H. Adams, U. Beuers, D.A. Mann, K.D. Lindor, D.E. Jones. Novel therapeutic targets in primary biliary cirrhosis. Nat Rev Gastroenterol Hepatol, 2015, 12: 147-158.

[4] J. Mattner. Impact of Microbes on the Pathogenesis of Primary Biliary Cirrhosis (PBC) and Primary Sclerosing Cholangitis (PSC). Int J Mol Sci, 2016, 17.

[5] Y. Zhang, S. Li, L. He, F. Wang, K. Chen, J. Li, T. Liu, Y. Zheng, J. Wang, W. Lu, Y. Zhou, Q. Yin, Y. Xia, Y. Zhou, J. Lu, C. Guo. Combination therapy of fenofibrate and ursodeoxycholic acid in patients with primary biliary cirrhosis who respond incompletely to UDCA monotherapy: a meta-analysis. Drug Des Devel Ther, 2015, 9: 2757-2766.

[6] Study protocol for the World Health Organization project to develop a Quality of Life assessment instrument (WHOQOL). Qual Life Res, 1993, 2: 153-159.

[7] Measuring Health-Related Quality of Life. 1993, 118: 622-629.

[8] A. Liberati, D.G. Altman, J. Tetzlaff, C. Mulrow, P.C. Gotzsche, J.P. Ioannidis, M. Clarke, P.J. Devereaux, J. Kleijnen, D. Moher. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol, 2009, 62: e1-34.

[9] T.R. Osborne, C. Ramsenthaler, R.J. Siegert, P.M. Edmonds, S.A. Schey, I.J. Higginson. What issues matter most to people with multiple myeloma and how well are we measuring them? A systematic review of quality of life tools. Eur J Haematol, 2012, 89: 437-457.

[10] S. Hawker, S. Payne, C. Kerr, M. Hardey, J. Powell. Appraising the evidence: reviewing disparate data systematically. Qual Health Res, 2002, 12: 1284-1299.

[11] G. Cerri, C.A. Cocchi, M. Montagna, M. Zuin, M. Podda, P. Cavallari, C. Selmi. Patients with primary biliary cirrhosis do not show post-exercise depression of cortical excitability. Clin Neurophysiol, 2010, 121: 1321-1328.

[12] L. Montali, A. Tanaka, P. Riva, H. Takahashi, C. Cocchi, Y. Ueno, M. Miglioretti, H. Takikawa, L. Vecchio, A. Frigerio, I. Bianchi, R. Jorgensen, K.D. Lindor, M. Podda, P. Invernizzi, P.B.C.S.G. Italian-Japanese. A short version of a HRQoL questionnaire for Italian and Japanese patients with Primary Biliary Cirrhosis. Dig Liver Dis, 2010, 42: 718-723.

[13] S. Zec, D. Popovic, V. Matovic, V. Nikolic, K. Bojovic, J. Jovic, L. Markovic Denic, T. Milosavljevic, T. Alempijevic. Translation and validation of the Serbian primary biliary cholangitis-40 questionnaire. PLoS One, 2017, 12: e0175697.

[14] G.F. Mells, G. Pells, J.L. Newton, A.J. Bathgate, A.K. Burroughs, M.A. Heneghan, J.M. Neuberger, D.B. Day, S.J. Ducker, R.N. Sandford, G.J. Alexander, D.E. Jones, U.-P. Consortium. Impact of primary biliary cirrhosis on perceived quality of life: the UK-PBC national study. Hepatology, 2013, 58: 273-283.

[15] L. Alrubaiy, G. Mells, S. Flack, H. Bosomworth, H. Hutchings, J. Williams, D. Jones, U.-P.R. Consortium. PBC-10: a short quality of life measure for clinical screening in primary biliary cholangitis. Aliment Pharmacol Ther, 2019, 50: 1223-1231.

[16] W.R. Kim, K.D. Lindor, M. Malinchoc, J.L. Petz, R. Jorgensen, E.R. Dickson. Reliability and validity of the NIDDK-QA instrument in the assessment of quality of life in ambulatory patients with cholestatic liver disease. Hepatology, 2000, 32: 924-929.

[17] A. Jacoby, A. Rannard, D. Buck, N. Bhala, J.L. Newton, O.F. James, D.E. Jones. Development, validation, and evaluation of the PBC-40, a disease specific health related quality of life measure for primary biliary cirrhosis. Gut, 2005, 54: 1622-1629.

[18] Z.M. Younossi, G. Guyatt, M. Kiwi, N. Boparai, D. King. Development of a disease specific questionnaire to measure health related quality of life in patients with chronic liver disease. Gut, 1999, 45: 295-300.

[19] P. Rucci, G. Taliani, L. Cirrincione, A. Alberti, D. Bartolozzi, N. Caporaso, M. Colombo, R. Coppola, M. Chiaramonte, A. Craxi, I. De Sio, A.R. Floreani, G.B. Gaeta, M. Persico, G. Secchi, I. Versace, A. Mele. Validity and reliability of the Italian version of the Chronic Liver Disease Questionnaire (CLDQ-I) for the assessment of health-related quality of life. Dig Liver Dis, 2005, 37: 850-860.

[20] S.M. van der Plas, B.E. Hansen, J.B. de Boer, T. Stijnen, J. Passchier, R.A. de Man, S.W. Schalm. The Liver Disease Symptom Index 2.0; validation of a disease-specific questionnaire. Qual Life Res, 2004, 13: 1469-1481.

[21] J. Ware, K. Snoww, K. Ma, G. Bg. SF36 Health Survey: Manual and Interpretation Guide. Lincoln, RI: Quality Metric, Inc, 1993, 1993, 30.

[22] J. Ware, Jr., M. Kosinski, S.D. Keller. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care, 1996, 34: 220-233.

[23] H. Sintonen. The 15D instrument of health-related quality of life: properties and applications. Ann Med, 2001, 33: 328-336.

[24] S.H. Belle, M.K. Porayko, J.H. Hoofnagle, J.R. Lake, R.K. Zetterman. Changes in quality of life after liver transplantation among adults. National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Liver Transplantation Database (LTD). Liver Transpl Surg, 1997, 3: 93-104.

[25] TORRANCE, GEORGE, W., FEENY, DAVID, H., FURLONG, WILLIAM, J., B.J.M. Care. Multiattribute Utility Function for a Comprehensive Health Status Classification System. 1996.

[26] D. Feeny, W. Furlong, G.W. Torrance, C.H. Goldsmith, Z. Zhu, S. DePauw, M. Denton, M. Boyle. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care, 2002, 40: 113-128.

[27] S.M. Hunt, S.P. McKenna, J. McEwen, J. Williams, E. Papp. The Nottingham Health Profile: subjective health status and medical consultations. Soc Sci Med A, 1981, 15: 221-229.

[28] B.B. Reeve, R.D. Hays, J.B. Bjorner, K.F. Cook, P.K. Crane, J.A. Teresi, D. Thissen, D.A. Revicki, D.J. Weiss, R.K. Hambleton, H. Liu, R. Gershon, S.P. Reise, J.S. Lai, D. Cella, P.C. Group. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Med Care, 2007, 45: S22-31.

[29] M. Yagi, A. Tanaka, T. Namisaki, A. Takahashi, M. Abe, A. Honda, Y. Matsuzaki, H. Ohira, H. Yoshiji, H. Takikawa, P.B.C.S.G. Japan. Is patient-reported outcome improved by nalfurafine hydrochloride in patients with primary biliary cholangitis and refractory pruritus? A post-marketing, single-arm, prospective study. J Gastroenterol, 2018, 53: 1151-1158.

[30] M. Turco, N. Cazzagon, I. Franceschet, C. Formentin, G. Frighetto, F. Giordani, N. Cellini, G. Mazzotta, R. Costa, B. Middleton, D.J. Skene, A. Floreani, S. Montagnese. Morning Bright Light Treatment for Sleep-Wake Disturbances in Primary Biliary Cholangitis: A Pilot Study. Front Physiol, 2018, 9: 1530.

[31] J. Raszeja-Wyszomirska, E. Wunsch, M. Krawczyk, E.I. Rigopoulou, K. Kostrzewa, G.L. Norman, D.P. Bogdanos, P. Milkiewicz. Assessment of health related quality of life in polish patients with primary biliary cirrhosis. Clin Res Hepatol Gastroenterol, 2016, 40: 471-479.

[32] S. Bondini, J. Kallman, A. Dan, Z. Younoszai, L. Ramsey, F. Nader, Z.M. Younossi. Health-related quality of life in patients with chronic hepatitis B. Liver Int, 2007, 27: 1119-1125.

[33] P. Charatcharoenwitthaya, J.A. Talwalkar, P. Angulo, A.A. Gossard, J.C. Keach, J.L. Petz, R.A. Jorgensen, K.D. Lindor. Moexipril for treatment of primary biliary cirrhosis in patients with an incomplete response to ursodeoxycholic acid. Dig Dis Sci, 2010, 55: 476-483.

Hosted file

Tables.pdf available at https://authorea.com/users/380099/articles/496208-health-related-quality-of-life-questionnaires-used-in-primary-biliary-cholangitis-a-systematic-review

```
┌─────────────────────────────────────┐
│ Records identified from database search │
│              n=1175                    │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│   Records after duplicates removed     │
│              n=435                     │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐        ┌──────────────────────────────────────┐
│ Records screened on title and abstract │──────▶│ Records excluded with reasons:          │
│              n=435                     │        │ (n=379): did not meet inclusion criteria │
└─────────────────────────────────────┘        └──────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐        ┌──────────────────────────────────────┐
│ Full-text articles reviewed for eligibility │─▶│ Full-text articles excluded with reasons: │
│              n=56                      │        │ (n=10): the same study cohort           │
└─────────────────────────────────────┘        │ (n=1): no HRQOL data                    │
                  │                              │ (n=1): the PBC population was not        │
                  ▼                              │ separated from the PSC population        │
┌─────────────────────────────────────┐        └──────────────────────────────────────┘
│ Studies included in the systematic review │
│              n=44                      │
└─────────────────────────────────────┘
```

12