Validation of the worst itch numerical rating scale (WI-NRS) and related patient-reported outcomes (PROs) to assess severity and impact of cholestatic pruritus in primary biliary cholangitis (PBC): An observational study

PCR123

Digital poster

SCAN ME

WI-NRS, Pruritus-related Sleep Interference NRS and Fatigue NRS items as well as the PBC-40 (7-day recall version) are valid and reliable PRO measures that are appropriate for the assessment of symptom severity and impact of pruritus in patients with PBC and cholestatic pruritus

Brooke M. Currie[1], Donald M. Bushnell[2], Mona L. Martin[2], Hayley Karn[3], Marlyn J. Mayo[4], David E. Jones[5], Megan M. McLaughlin[1], Robyn von Maltzahn[6]

[1]GSK, Collegeville, PA, USA; [2]Evidera | PPD, Wilmington, NC, USA; [3]Evidera | PPD, London, UK; [4]University of Texas Southwestern Medical Center, Dallas, TX, USA; [5]Newcastle University, Newcastle, UK; [6]GSK, London, UK

Background

- Pruritus, fatigue and sleep disturbance are common symptoms of PBC assessed by PROs in clinical trials of PBC[1,2]
- NRS are appropriate for assessing unidimensional PRO concepts and are easily interpreted; while NRS have been validated for assessment of symptom severity in other diseases,[3–6] their validity and reliability have not been evaluated for use in a PBC population with cholestatic pruritus
- The PBC-40, a disease-specific questionnaire designed to assess health-related QoL in patients with PBC, is used extensively in clinical research in PBC[7,8]
  - The PBC-40 was originally developed and validated using a 4-week recall period in a PBC population both with and without pruritus[9,10]
- To utilise PRO instruments in clinical trials, rigorous psychometric validation is required by regulatory authorities in the relevant patient population, particularly to confirm their ability to quantify change in the context of therapeutic intervention[11]
- Herein, we present findings from an observational study in patients with PBC and pruritus, designed to evaluate psychometric properties of PRO measures to assess pruritus, pruritus-related sleep interference, fatigue, and health-related QoL

Conclusions

Results from this study confirm the reliability and validity of the WI-NRS, Sleep Interference NRS and Fatigue NRS as well as the PBC-40 (modified with a 7-day recall period)

This analysis strongly supports the use of these PROs as assessments for patient-centric trial endpoints, including in the Phase 3 GLISTEN trial (NCT04950127) of linerixibat for pruritus treatment in patients with PBC

Use of these PROs in an interventional study involving participants with PBC and pruritus will enable the assessment of any potential treatment benefit of the investigational product

Aims

To evaluate the psychometric properties of the following PRO measures for assessing severity and impact of pruritus in patients with PBC and cholestatic pruritus:

NRS items:
- Worst Itch Numerical Rating Scale (WI-NRS)
- Pruritus-related Sleep Interference NRS (hereafter referred to as 'Sleep Interference NRS')
- Fatigue NRS

PBC-40 (7-day recall)[12,13]

Methods

- This web-based survey recruited individuals with self-reported PBC and prior/current pruritus from market research panels in the US, UK and Canada

Table 1: **PRO assessments in the observational study**

PROs	Time points of assessment	Response options	Scores
WI-NRS	Twice daily (morning and evening) Day 1 to Day 8	0–10 scale (0=no itching; 10=worst imaginable itching)	Worst daily itch score: worst of the two daily itch scores Weekly itch score: average of worst daily itch scores over a 7-day period
Sleep Interference NRS	Once daily (morning) Day 1 to Day 8	0–10 scale (0=did not interfere; 10=completely interfered)	Weekly sleep score: average of daily sleep scores over a 7-day period
Fatigue NRS	Once daily (evening) Day 1 to Day 8	0–10 scale (0=no fatigue; 10=worst possible fatigue)	Weekly fatigue score: average of daily fatigue scores over a 7-day period
PBC-40 (7-day recall)	Day 1 and Day 8	5-level response scale (e.g., never to always or not at all to very much)	Scores for individual items range from 1 (or 0) to 5, with higher scores indicating worse health-related QoL Score ranges for the six domains (Symptoms, Itch, Fatigue, Cognitive, Social, Emotional) vary depending on the domain

- Other select PROs, including the PROMIS-43, patient global impression of severity (PGI-S) and 5D-Itch scale were collected on Days 1 and 8 to support the psychometric evaluation of the NRS items and the PBC-40 (7-day recall)
- The following psychometric properties were assessed: reliability, including test–retest reliability and internal consistency reliability for PBC-40 (7-day recall); construct validity, including convergent and known-groups validity; and confirmatory factor analysis to evaluate the dimensionality and existing structure of the PBC-40 (7-day recall)

Results

Figure 1: **Participant demographics and clinical characteristics (N=141)**

≥35 years of age, n (%)	Female, n (%)	Race, n (%)	Region, n (%)
131 (93)	131 (93)	White 122 (87)	US: 75 (53) UK: 56 (40) Canada: 10 (7)

WI-NRS at screening, mean (SD)	Diagnosis of PBC ≥12 months, n (%)	Experienced pruritus for >1 year, n (%)
4.7 (2.8)	141 (100)	112 (79)

- At screening, 69% were currently experiencing pruritus and 50% characterised their overall itching severity as moderate or greater (WI-NRS ≥4)

Test–retest reliability

- Results showed acceptable test–retest reliability for the PBC-40 (7-day recall) domains (ICCs: 0.85–0.90)
- Test–retest reliability was also acceptable for the WI-NRS (ICC: 0.78)

Table 2: **Internal consistency reliability of PBC-40 (7-day recall) was excellent and acceptable to excellent for individual domains at Day 1**

PBC-40 (7-day recall) domains	Cronbach's alpha (N=141)
Overall alpha	0.95
Fatigue	0.95
Emotional	0.78
Social	0.87
Cognitive	0.95
Symptoms	0.73
Itch	0.87

Cronbach's alpha was calculated for raw variables

Table 3: **Confirmatory factor analysis of the PBC-40 (7-day recall) at Day 1 confirmed the PBC-40 (7-day recall) domain structure with acceptable factor coefficients and adequate fit statistics**

PBC-40 (7-day recall) domains	Items	Factor coefficient range	RMSEA (90% CI)	CFI
Fatigue	11	0.72–0.94	0.059 (0.051–0.066)	0.966
Emotional	3	0.56–0.87		
Social	10	0.43–0.91		
Cognitive	6	0.87–0.92		
Symptoms	7	0.39–0.76		
Itch	3	0.82–0.92		

Table 4: **Convergent validity assessed at Day 1 using Spearman's correlations between each of the NRS items, PBC-40 (7-day recall) and predicted convergent concept, was confirmed for the NRS items and the PBC-40 (7-day recall) Itch, Fatigue, Cognitive, Emotional and Social domains**

	r†
WI-NRS‡	
PBC-40 – Itch domain score	0.79****
5-D Itch Scale – Total score	0.71****
Sleep Interference NRS	
PBC-40 – Sleep item 8 score	0.79****
Fatigue NRS	
PBC-40 – Fatigue domain score	0.68****
PBC-40 (7-day recall)§	
Itch domain	
5-D Itch Scale – Total score	0.70****
WI-NRS‡	0.79****
Fatigue domain	
Fatigue NRS	0.68****
Cognitive domain	
Concentration item¶	0.64****
Remember item††	0.69****
Emotional domain	
PROMIS-43 – Depression/Anxiety domain score	0.61****
Social domain	
PROMIS-43 – Social roles/Activities domain score	|0.80|****

*†r denotes Spearman's rank-order correlations; asterisks indicate the level of statistical significance, all ****p<0.0001. ‡Worst Daily Itch score. §The 'Symptoms' domain was deliberately excluded from evaluation due to the absence of appropriate measures for assessment. ¶How difficult was it for you to concentrate today? (Not at all, A little, Somewhat, Quite, Extremely). ††How difficult was it for you to remember things today? (Not at all, A little, Somewhat, Quite, Extremely)*

Table 5: **Known-groups validity was confirmed for the NRS items based on groups defined by patient global impression of severity (PGI-S) at Day 1 (all p<0.001)**

	PGI-S					
PRO (mean score [SD])	Absent (n=19)	Mild (n=70)	Moderate (n=40)	Severe (n=12)	F Value and significance†	Post hoc comparison‡
WI-NRS§	0.4 (0.8)	3.3 (1.8)	5.8 (1.4)	8.0 (1.1)	81.6***	1***, 2***, 3***, 4***, 5***, 6***
Sleep Interference NRS¶	0.0 (0.0)	1.9 (2.0)	5.0 (2.1)	7.8 (1.2)	70.8***	1**, 2***, 3***, 4***, 5***, 6***
Fatigue NRS††	3.3 (2.4)	5.5 (2.5)	6.5 (1.7)	6.8 (2.6)	8.9***	1**, 2***, 3**

*The PGI-S has a 'very severe' category; however, no one selected that category. P-values are: **<0.01; ***<0.001.*
†An ANOVA. ‡One-way ANOVA using Scheffe's Method. 1=Absent vs mild; 2=Absent vs moderate; 3=Absent vs severe; 4=Mild vs moderate; 5=Mild vs severe; 6=Moderate vs severe. §Worst Daily Itch score. ¶Daily Sleep score. ††Daily Fatigue score

- Similarly, known-groups validity was confirmed for the PBC-40 (7-day recall)
  - At Day 1, PBC-40 (7-day recall) domains discriminated between patients in the expected direction according to groups defined by PGI-S, with the Itch, Fatigue, Social and Emotional domains being highly discriminant (p<0.001)
  - When examining groups by the General Health status item of the PBC-40 (7-day recall), five of the six domains were able to differentiate between levels of health (p<0.001)

Abbreviations

5-D, 5-dimension; ANOVA, analysis of variance; CFI, comparative fit index; CI, confidence interval; ICC, intraclass correlation coefficient; NRS, numerical rating scale; PBC, primary biliary cholangitis; PBC-40, primary biliary cholangitis – 40 items; PGI-S, patient global impression of severity; PRO, patient-reported outcome; QoL, quality of life; RMSEA, root mean square error of approximation; SD, standard deviation; WI-NRS, worst itch numerical rating scale

References

1. Levy C, et al. *Clin Gastro Hep.* 2022;21(7):1902–12
2. De Vries E, et al. *Gastroenterology.* 2021;160(3):734–43
3. Rams A, et al. *Adv Ther.* 2024;41(4):1512–25
4. Kimball AB, et al. *Br J Dermatol.* 2016;175(1):157–62
5. Gladman D, et al. *RMD Open.* 2020 Jan;6(1):e000928
6. Ständer S et al. *Dermatol Ther (Heidelb).* 2023;13(7):1587–1602
7. von Maltzahn R, et al. *J Patient Rep Outcomes.* 2024;8(1):60
8. Jones D, et al. *Hepatol Commun.* 2023;7(3):e0057
9. Jacoby A, et al. *Gut.* 2005;54(11):1622–9
10. Mayo MJ, et al. *Dig Dis Sci.* 2023;68(3):995–1005
11. Gnanasakthy A, et al. *Value Health.* 2017;20(3):420–9
12. Martin ML, et al. *J Patient Rep Outcomes* 2019;3(1):2
13. Gilchrist K, Vallow S. *Value Health* 2015;18:A30

Acknowledgements

Editorial support (in the form of writing assistance, including preparation of the draft poster under the direction and guidance of the authors, collating and incorporating authors' comments for each draft, assembling tables and figures, grammatical editing, and referencing) was provided by Angeliki Karamitri, PhD, of Fishawack Indicia Ltd, part of Avalere Health, and was funded by GSK

Disclosures

This study was funded by GSK (study 212144).
BMC, **MMM** and **RvM** are employed by GSK and hold financial equities in GSK. **DMB**, **MLM** and **HK** are employees of Evidera, who was funded by GSK to carry out this study. **MJM** has received research support from CymaBay, GSK, Intercept, Ipsen, Mallinckrodt and Mirum, and has consulted for CymaBay, GSK and Mallinckrodt. **DEJ** has received grant funding or fees from Abbott, Falk, GSK, Intercept, Novartis and Pfizer