Psychometric Validation of the Sunlight Exposure Diary and Erythropoietic Protoporphyria Impact Questionnaire (EPIQ) Using Data from the AURORA Clinical Study in Erythropoietic Protoporphyria (EPP)

Susan D. Mathias, MPH[1], Michelle P. Turner, MS[1], Megan O'Grady, PhD[2], Chelsea Norregaard, PhD, MPH[2], Hilary H. Colwell, MPH[1], William Savage, MD, PhD[2], Hetanshi Naik, PhD[3], Melanie Chin, PhD[2]

[1]Health Outcomes Solutions, Palm Beach Gardens, FL, USA, [2]Disc Medicine, Watertown, MA, USA, [3]Stanford University School of Medicine, Palo Alto, CA, USA.

MSR61

INTRODUCTION

Erythropoietic protoporphyria (EPP) and X-linked protoporphyria (XLP) are rare, life-long diseases that can negatively affect one's health-related quality of life (HRQoL) because of symptoms including pain from phototoxic reactions.

There is a need for reliable and valid patient-reported outcome measures (PROs) that address all aspects of EPP/XLP, including symptoms and impacts. The psychometric properties of 2 PROs in individuals with EPP, the Sunlight Exposure Diary and EPP Impact Questionnaire (EPIQ), were evaluated.

METHODS

Data Collection

- Data from participants enrolled in AURORA, a randomized, double-blind, placebo-controlled study of bitopertin in EPP, with a baseline and at least 1 follow-up assessment on the EPIQ, were included.

Overview of Measures

- The Sunlight Exposure Diary[1] and the EPP Impact Questionnaire (EPIQ)[2] are recently developed PROs for capturing the experiences of those with EPP/XLP.
- **Sunlight Exposure Diary** recorded daily sunlight exposure and the severity (mild, moderate, severe) of symptoms (tingling, burning, itching, stinging, sensitivity to hot or cold, sensitivity to touch, pain, feelings of warmth, swelling, redness/discoloration, burst blood vessels, and blisters).
- **EPIQ** assessed the impact of EPP/XLP, including the occurrence of any phototoxic reactions, duration of sun exposure before a reaction, time to improvement and complete resolution, impact on daily activities, and overall quality of life compared to someone without EPP/XLP. It also included 3 Patient Global Impression of Severity (PGI-S) items and 5 Patient Global Impression of Change (PGI-C) items.
  - Individual EPIQ items were rescaled to 0 to 100, where higher scores indicate better outcomes.
- Participants also completed Patient-Reported Outcomes Measurement Information System (PROMIS)-57 v2.1, PROMIS Short Form v2.0-Social Isolation, and PROMIS-Neuropathic Pain Quality v2.0 scales.

Analyses Conducted

- Exploratory factor analysis (EFA). To evaluate the scale structure of the Sunlight Exposure Diary and the EPIQ
- Internal consistency reliability. Using Cronbach's alpha coefficient, with values of ≥0.70 considered acceptable
- Test-retest reliability. Using intraclass correlation coefficient (ICC), with values of ≥0.75 indicating good reliability
- Construct validity. Comparing hypothesized scale and item correlations with PROMIS measures
- Known-groups validity. Categorizing participants based on:
  - 2-week symptom average (dichotomized at median of 100) and baseline light tolerance (<30 minutes vs ≥30 minutes)
  - Baseline light tolerance categories of ≤10 minutes, >10 to <30 minutes, and ≥30 minutes to assess the potential for improved differentiation
- Responsiveness (using standardized effect sizes [SES] and standardized response means [SRM]), and meaningful change (minimal detectable change [MDC] and minimal important change [MIC] using distribution- and anchor-based approaches)
- Data from all treatment groups were combined

RESULTS

Data from 65 participants were included (Table 1).

Table 1. Demographic and Clinical Characteristics

Characteristic	Total (N=65)
Sex, n (%)	
Female	32 (49.2%)
Male	33 (50.8%)
Age in years, mean (SD)	44.8 (13.0)
Whole-blood metal-free PPIX levels, µg/L	
Mean (SD)	9335.5 (5682.1)
Median (Interquartile Range)	7790 (5250, 11,230)
Two-week interval symptom average	
Mean (SD)	93.1 (11.9)
Median	100
Missing, n (%)	4 (6.2%)
<100, n (%)	29 (44.6%)
=100, n (%)	32 (49.2%)
Geographic region, n (%)	
US Midwest or Northeast	43 (66.2%)
US South or West Region	22 (33.8%)
Baseline Light Tolerance, n (%)	
<30 minutes	22 (33.8%)
≥30 minutes	43 (66.2%)

EFA

- Supported domains for Duration of Full Phototoxic Reaction (DFR) [2 items], Overall Change (OC) [5 items], and Overall Severity and Impact (OSI) [5 items].
- Psychometric properties of a single item, Daily Daylight Tolerance (DDT), were evaluated over 2-week (TWI) and 1-month (MON) intervals.

Reliability

- Most scales demonstrated strong internal consistency and test-retest reliability
  - DDT-TWI had acceptable test-retest reliability (ICC = 0.89).
  - DFR had acceptable internal consistency (Cronbach's alpha = 0.86).
  - OC had acceptable reliability (Cronbach's alpha = 0.97 and ICC = 0.97). However, it was used to define the stable cohort used for ICC, so that value is likely artificially inflated.
  - OSI had acceptable reliability (Cronbach's alpha = 0.88), but some variability in item-total correlations; ICC was lower (0.62) than desired.

Construct Validity

- OSI was correlated as expected with PROMIS scales (some were higher, some were lower), except with the PROMIS Physical Function.
- DFR and OC scales correlated as expected (some were higher, some were lower), although with some variability.
- DDT-TWI did not correlate with PROMIS as expected, which may be a result of using 2-week period (instead of a longer period), causing increased variability.

Known-Groups Validity (Figures 1 and 2)

- Mean OSI scores differed between symptom severity and time to prodrome groups.
- DDT-TWI only differentiated by time to prodrome.
- DFR and OC did not significantly differ by known groups.
- Exploratory known-groups analysis using the 3-category time-to-prodrome classification showed results similar to the 2-category analysis (data not shown).

Ceiling Effects

- High ceiling effects (eg, blisters, burst blood vessels) were noted starting at Day 1.

Figure 1. Forest Plot of Known-Groups Validity: Mean Difference in Average Symptom Score Categories

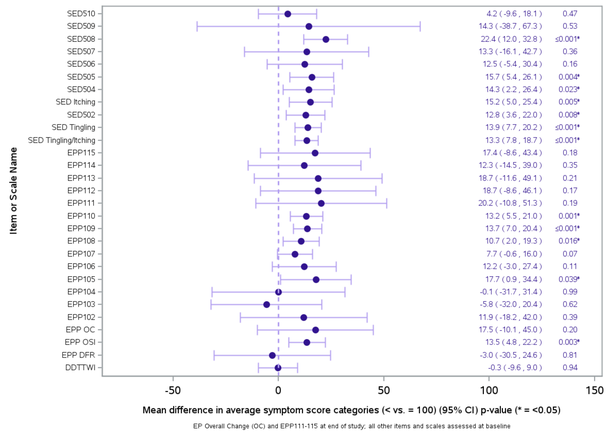
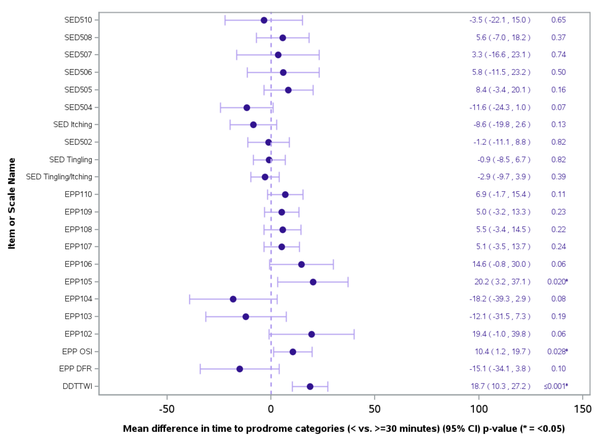


Figure 2. Forest Plot of Known-Groups Validity: Mean Difference in Time-to-Prodrome Categories



Responsiveness

- DDT-TWI showed low responsiveness (small SES [0.20] and SRM [0.16], alongside a negative responsiveness statistic range [-0.91 to -0.55]).
- DFR demonstrated moderate responsiveness (SES of -0.11 [low] and SRM of 0.92 [high]), although data to compute the responsiveness statistic are not available.
- OSI scale exhibited high responsiveness (strong SES of 1.23 and SRM of 0.92), supported by a positive, albeit low, responsiveness statistic range (0.06 to 0.14).

Meaningful Change

- MIC range for OSI was 21-24 points.
- MIC range for DDT-TWI was -9.2 to -0.2 (expected to be +), while MIC range from DDT-MON was -0.9 to 7.1 (Note: MIC range could not be calculated for DFR and OC).
- Exploratory meaningful change analyses for the DDT-TWI measures based on combined anchor categories revealed a range of 6.7 to 7.2 for the TWI, whereas the MON interval showed a higher range of 13.3 to 13.5.

A summary of the psychometric properties is provided in Table 2.

Table 2. Summary of Psychometric Properties

Primary Finding	Measure	Reliability	Validity	Responsiveness	Minimal Detectable Change
Strong psychometric properties	Overall Severity and Impact	Acceptable	Excellent	High	6.4
Acceptable psychometric properties	Duration of Full Reaction	Acceptable	Fair	Moderate	5.3
	Overall Change	Potential item redundancy	Moderate	Moderate	Unable to calculate
Weak psychometric properties	Daily Daylight Tolerance Over 2-Week Intervals	Acceptable	Moderate	Low	6.1
	Daily Daylight Tolerance Over 1-Month Intervals	Not assessed	Not assessed	Not assessed	Not assessed

Strengths and Limitations

- Analyses confirm the robustness of the EPIQ and Sunlight Exposure Diary, particularly the OSI scale, which demonstrated strong psychometric properties in terms of reliability, validity, and responsiveness. The DFR and OC scales also exhibited acceptable psychometric properties.
- These PROs address the need for EPP/XLP-specific measures that assess the duration, severity, and impact of early warning symptoms and full phototoxic reactions, and capture impacts of EPP and XLP on well-being and HRQoL.
- Some limitations included small sample sizes in some analyses, including responsiveness and meaningful change assessments, where sparse data limited the calculation of key MDC and MIC statistics for some scales.
- Limited variability in the known-groups analyses restricted the ability to detect meaningful differences, further affecting the robustness of the findings.

Areas to Explore

- Confirmation of the identified factor structure with a different sample.
- Re-evaluation of test-retest reliability using a shorter time interval between "test" and "retest."
- Exploring a 1-month interval for daylight tolerance instead of 2 weeks to strengthen validity testing. This approach will increase sample sizes to improve robustness of responsiveness and meaningful change analyses and enhance the variability of known groups.

Conclusion

- **Overall, the Sunlight Exposure Diary and EPIQ were found to be reliable and valid, supporting their use in future EPP clinical studies.**

References

1. Mathias SD, et al. Development and content validation of a sunlight exposure diary in patients with erythropoietic protoporphyria. J Patient Rep Outcomes. 2023;7(1):119.
2. Mathias SD, et al. Development and content validation of novel patient-reported outcome measures to assess disease severity and change in patients with erythropoietic protoporphyria: the EPP Impact Questionnaire (EPIQ). Patient Relat Outcome Meas. 2024;15:17-30.

Contact

Will Savage, MD, PhD | Chief Medical Officer, Disc Medicine
wsavage@discmedicine.com

Presented at ISPOR, May 13-16, 2025, Montreal, QC, Canada