Psychometric validation of Generalized Pustular Psoriasis (GPP) Physician Global Assessment (GPPGA) and Generalized Pustular Psoriasis Area and Severity Index (GPPASI) as clinician-reported outcomes in GPP

Birgit Gradl[1], Na Hu[2], Christian Thoma[3], Thomas Zimmermann[1], Ismail Budhiarso[4], Milena Anatchkova[4], Anne Skalicky[4]

[1]Boehringer Ingelheim International GmbH, Ingelheim, Germany; [2]Boehringer Ingelheim (China) Investment Corporation Limited, Shanghai, People's Republic of China; [3]Boehringer Ingelheim International GmbH, Biberach, Germany; [4] Evidera, Bethesda, MD, USA

GPPGA and GPPASI are reliable, valid and responsive measures that can detect meaningful changes in GPP severity. The findings of this study support their use as endpoints in GPP clinical trials

PURPOSE

To evaluate the reliability, validity and responder definitions of GPPGA and GPPASI using data from the Effisayil™ 1 study, and confirm that these measures are suitable for the assessment of GPP disease severity in clinical trials

INTRODUCTION

- GPP is a rare autoinflammatory skin disease characterized by sterile, neutrophilic pustules often accompanied by systemic inflammation, with mortality rates ranging from 2%–16%[1–8]
- GPPGA and GPPASI are novel clinician-reported measures of GPP-specific severity, adapted from the established PGA and PASI with input from both dermatologists and patients with GPP[9,10]
- As the GPPGA and GPPASI were used in the definitions of the primary and secondary endpoints of Effisayil™ 1, a trial of the anti-IL-36R monoclonal antibody spesolimab in GPP, it is therefore necessary to evaluate their psychometric properties, consistent with US FDA guidance[9]

CONCLUSIONS

- Overall, psychometric analyses of the GPPGA and GPPASI indicate that these measures are valid, reliable and are responsive endpoints to assess meaningful change in GPP severity
- Our findings support the use of GPPGA and GPPASI in Effisayil™ 1, and also as measures of clinical efficacy in future studies in patients with GPP

METHODS

- All analyses were calculated using data from Week 1
- Item correlations, internal consistency and CFA were not conducted for the GPPASI as items are not inter-related

Analysis step	Objective/question	Methods applied
Confirmatory factor analysis	Does the data support the structure of the measure?	Two CFAs for GPPGA (constrained & unconstrained model)
Item-to-item and item-to-total correlations	Assess relationships among items and between items and total score within a measure	Pearson correlation for GPPGA items and total score
Internal consistency reliability	Assess degree of agreement between items – is the measure consistent?	Cronbach's alpha coefficient
Convergent validity	Do the measures correlate with similar measures?	Correlations between GPPGA, GPPGA pustulation subscore, GPPASI with DLQI total score, items 1 and 2, EQ-VAS, EQ-5D pain/discomfort item, CGI-I
Test-retest reliability	Are the measures reproducible over time in stable patients?	ICC values for GPPGA at Days 1-8. Stability defined as no change in 1) JDA GPP Severity Index Part A or 2) JDA Part B
Known-groups validity	Do the measures distinguish between distinct groups (e. g. based on disease severity)?	Compare GPPGA and GPPASI scores across groups defined by anchors (JDA Part A; JDA Part B; DLQI total score; EQ-5D pain item; EQ-VAS)
Responsiveness/ability to detect change	Are the measures sensitive to change in health status?	Correlations between change in GPPGA and change in anchors (DLQI item 1; EQ-5D pain/discomfort; CGI-I)
Responder definition	What magnitude of change is considered important to the patient?	Determine the change in GPPGA, GPPGA pustulation subscore, GPPASI in sub-populations with subjective patient-reported meaningful change in anchors (DLQI item 1; EQ-5D pain item; EQ-VAS) or clinician-reported change (CGI-I)

RESULTS

CFA and inter-item correlations

GPPGA	CFA (constrained model)	CFA (unconstrained model)
Erythema	0.708	0.708
Pustules	0.893	0.896
Scaling/crusting	0.893	0.889
1 factor		
Chi-Square, p-value (df)	0.002, p=0.9655 (1)	0.000, p=0.000 (0)
CFI	1.00	1.000
RMSEA	0.000	0.000
90% CI for RMSEA	0.000–0.000	0.000–0.000
SRMR	0.001	0.000

Item correlations and CFA were not conducted for the GPPASI as items are not inter-related.
Acceptable ranges: Cronbach's α: ≥0.70; CFI: ≥0.9; RMSEA: <0.8; SRMR: <0.1.

CFA demonstrated unidimensionality of the GPPGA total score at Week 1 (RMSEA<0.08)

Item-to-item and item-to-total correlations were statistically significant (r=0.61–0.90; data not shown)

The GPPGA total score showed good internal consistency (Cronbach's α=0.81; data not shown)

Convergent validity

PRO/ClinRO variables	Correlations†		
	GPPGA total score	GPPGA pustulation subscore	GPPASI total score
CGI global improvement	0.45*	0.48**	0.24
DLQI total score	0.36*	0.33*	0.14
DLQI item 1: How itchy, sore, painful, or stinging has your skin been?	0.49**	0.45**	0.37*
DLQI item 2: How embarrassed or self-conscious have you been because of your skin?	0.39*	0.30*	0.25
EQ-5D pain/discomfort	0.54***	0.47**	0.46**
EQ-VAS score	-0.47**	-0.47**	-0.40*

†Spearman's rank order correlation. Correlation interpretation: less than 0.3 = weak, between 0.3 and 0.7 = moderate, between 0.7 and 0.9 = strong, and above 0.9 = very strong. Significance levels for correlations p-values are: *p<0.05, **p<0.001, ***p<0.0001.

The GPPGA total score and GPPGA pustulation subscore showed good evidence of convergent validity, with moderate-to-strong correlations with selected anchors

The GPPGA total score, pustulation score and GPPASI total score demonstrated good test-rest reliability (ICC=0.70, 0.91, and 0.95, respectively; data not shown)‡

‡using JDA GPP part A assessment of skin symptoms to define the stable population, from Day 3 to Day 4; ICC≥0.7 is acceptable for establishing test-retest reliability

Responsiveness

Change score from baseline to Week 1	CGI-I change score, 3-category						Overall F-test	
	Worsened/no change/minimally improved		Much improved		Very much improved			
	N	LS mean (SE)	N	LS mean (SE)	N	LS mean (SE(	F-test	p-value*
GPPGA total score	18	-0.53 (0.22)	12	-1.08 (0.27)	19	-1.50 (0.21)	3.42	0.0250

Change score from baseline to Week 1	DLQI item 1 change score, 2-category				Overall F-test	
	Worsened/no change		Improved			
	N	LS mean (SE)	N	LS mean (SE)	F-test	p-value*
GPPGA total score	28	-0.57 (0.16)	23	-1.56 (0.18)	8.21	0.0009

Change score from baseline to Week 1	EQ-5D pain/discomfort change score, 2-category				Overall F-test	
	Worsened/no change		Improved			
	N	LS mean (SE)	N	LS mean (SE)	F-test	p-value*
GPPGA total score	16	-0.36 (0.23)	36	-1.34 (0.15)	6.49	0.0032

*ANCOVA adjusted by baseline/Day 1 score and anchor change score. Pairwise comparisons are calculated only if at least five patients are in each group.

The GPPGA total score was able to differentiate between select known groups measuring different levels of symptom or disease severity

Responsiveness of GPPGA detected change in anchor severity categories from baseline (CGI–I, p<0.05; DLQI item 1, p<0.001; EQ-5D pain/discomfort, p<0.01)

Responder definitions

LS mean score change from baseline to Week 1	Anchor			
	DLQI item 1†	EQ-5D pain/discomfort‡	EQ-VAS§	CGI-I¶
GPPGA total score	-1.56	-1.34	-1.45	-1.36
GPPGA pustulation subscore	-2.24	-2.11	-2.30	-2.17
GPPASI total score	-12.64	-11.88	-12.65	-10.82

†group=small/moderate/large improvement. ‡group=minimally/moderate/much/very much improved. §group=minimal/large improvement. ¶group=minimally/much/very much improved

Using anchor-based analyses, responder definitions for the GPPGA total score, pustulation subscore, and GPPASI total score were reductions of approximately 1.4, 2.2, and 12.0 points, respectively*

*Responder definitions were calculated as an average of the estimates for DLQI Item 1 , EQ-5D pain/discomfort, EQ-VAS, and CGI-I, rounded to 1 decimal place.

GPPASI % improvement threshold

Mean change from Week 1/Day 8	GPPASI percent improvement category			Overall p-value		Pairwise comparison‡
	GPPASI <50%	GPPASI 50 to <75%	GPPASI ≥75%	F-test	p-value†	
	n, mean (SE)	n, mean (SE)	n, mean (SE)			
EQ-VAS score	29, 4.41 (3.55)	18, 42.78 (6.17)	5, 34.00 (5.57)	23.12	<.0001	Impr. <50% vs. Impr. 50–<75%: <.0001*** Impr. <50% vs. Impr. ≥75%: 0.0200* Impr. 50–<75% vs. Impr. ≥75%: 0.9633
DLQI total score	29, −1.28 (1.18)	17, −6.65 (1.65)	5, −7.20 (1.77)	3.40	0.0252	Impr. <50% vs. Impr. 50–<75%: 0.0659 Impr. <50% vs. Impr. ≥75%: 0.1417 Impr. 50–<75% vs. Impr. ≥75%: 0.9167

†From ANCOVA using GPPASI percent change categories and baseline value of COA/PRO as independent variables and change from baseline in COA/PRO as dependent variable. ‡p–values of pairwise comparisons, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. Pairwise comparisons are calculated only if at least five patients are in each group.

Anchor-based analyses support the GPPASI 50% as a meaningful threshold for improvement

References

1. Navarini AA, et al. *J Eur Acad Dermatol Venereol* 2017;31:1792–1799.
2. Fujita H, et al. *J Dermatol* 2018;45:1235–1270.
3. Augey F, et al. *Eur J Dermatol* 2006;16:669–673.
4. Jin H, et al. *J Dermatol* 2015;42:674–678
5. Miyachi H, et al. *J Am Acad Dermatol* 2021;86:1266–1274
6. Zelickson BD, et al. *Arch Dermatol* 1991;127:1339–1345
7. Choon SE, et al. *Int J Dermatol* 2014; 53:676–684
8. Choon SE, et al. *Am J of Clin Dermatol* 2022;23:21–29
9. Baker H, et al. *Br J Dermatol* 1972;80:771–793
10. Bachelez H, et al. *N Engl J Med* 2021;385:2431–2440.
11. Choon SE, et al. *BMJ Open* 2021;11:e043666.

Abbreviations

CFA, confirmatory factor analysis; CFI, comparative fit index; CGI-I, clinical global impression – improvement; CI, confidence interval; ClinRO, clinician-reported outcome; COA, clinical outcome assessment; df, degrees of freedom; DLQI, dermatology life quality index; EQ-VAS, EuroQol-Visual Analog Scale; FDA, food and drug administration; GPP, generalized pustular psoriasis; GPPASI, generalized pustular psoriasis area and severity index; GPPGA, generalized pustular psoriasis physician global assessment; ICC, intraclass correlation coefficient; JDA, Japanese Dermatological Association; LS, least squares; PASI, psoriasis area severity index; PGA, physician global assessment; PRO, patient-reported outcome; RMSEA, root mean square error of approximation; SE, standard error; SRMR, standardised root mean square residual; US, United States

Disclosures & Acknowledgements

The study was supported and funded by Boehringer Ingelheim. BG, NH, CT, and TZ are employees of Boehringer Ingelheim. IB, MA and AS are employees of Evidera, which was contracted by Boehringer Ingelheim for the purposes of this study. The authors met criteria for authorship as recommended by the International Committee of Medical Journal Editors (ICMJE). The authors did not receive payment related to the development of the poster. Boehringer Ingelheim was given the opportunity to review the poster for medical and scientific accuracy as well as intellectual property considerations. Isabella Goldsbrough, PhD, of OPEN Health Communications (London, UK) provided writing, editorial and formatting support, which was contracted and funded by Boehringer Ingelheim.

Scan QR code for an interactive, electronic device-friendly copy of the poster
https://bit.ly/3wnVcKe

Click the icon to access an interactive microsite for this Smart poster