

# Psychometric evaluation of the Immune Thrombocytopenic Purpura Patient Assessment Questionnaire Item-10 (Physical Fatigue) and Idiopathic Thrombocytopenic Purpura Bleeding Scale using LUNA 3 Phase 3 trial data

Imene Gouia<sup>1</sup>, Mattea Orsini<sup>2</sup>, Joshua Maher<sup>3</sup>, George Skingley<sup>4</sup>, Asha Lehane<sup>4</sup>, Helena Bradley<sup>4</sup>

<sup>1</sup>Health Economics and Value Assessment, Sanofi, Gentilly, France; <sup>2</sup>Aixial, Sèvres, France; <sup>3</sup>Clinical Outcome Assessment Group, Sanofi, Reading, United Kingdom; <sup>4</sup>Adelphi Values, Bollington, Cheshire, United Kingdom

Poster No. MSR174



Copies of this poster obtained through Quick Response (QR) Code are for personal use only

## OBJECTIVE



To evaluate the psychometric properties of Immune Thrombocytopenic Purpura Patient Assessment Questionnaire (ITP-PAQ) Item-10 (physical fatigue) and Idiopathic Thrombocytopenic Purpura Bleeding Scale (IBLS) total score in the context of LUNA 3 phase 3 trial.

## BACKGROUND

- Immune thrombocytopenic purpura (ITP) is a rare autoimmune disorder characterized by immune-mediated platelet destruction and impaired platelet production, resulting in platelet count below 100,000/ $\mu$ L.<sup>1</sup>
- The efficacy and safety of rilzabrutinib, a Bruton tyrosine kinase inhibitor, was evaluated in patients with ITP in the LUNA 3 phase 3 trial (NCT04562766).<sup>2</sup>
- The LUNA 3 phase 3 trial included the ITP-PAQ Item-10 to assess physical fatigue and the IBLS to support key secondary endpoints.<sup>2</sup>

## METHODS

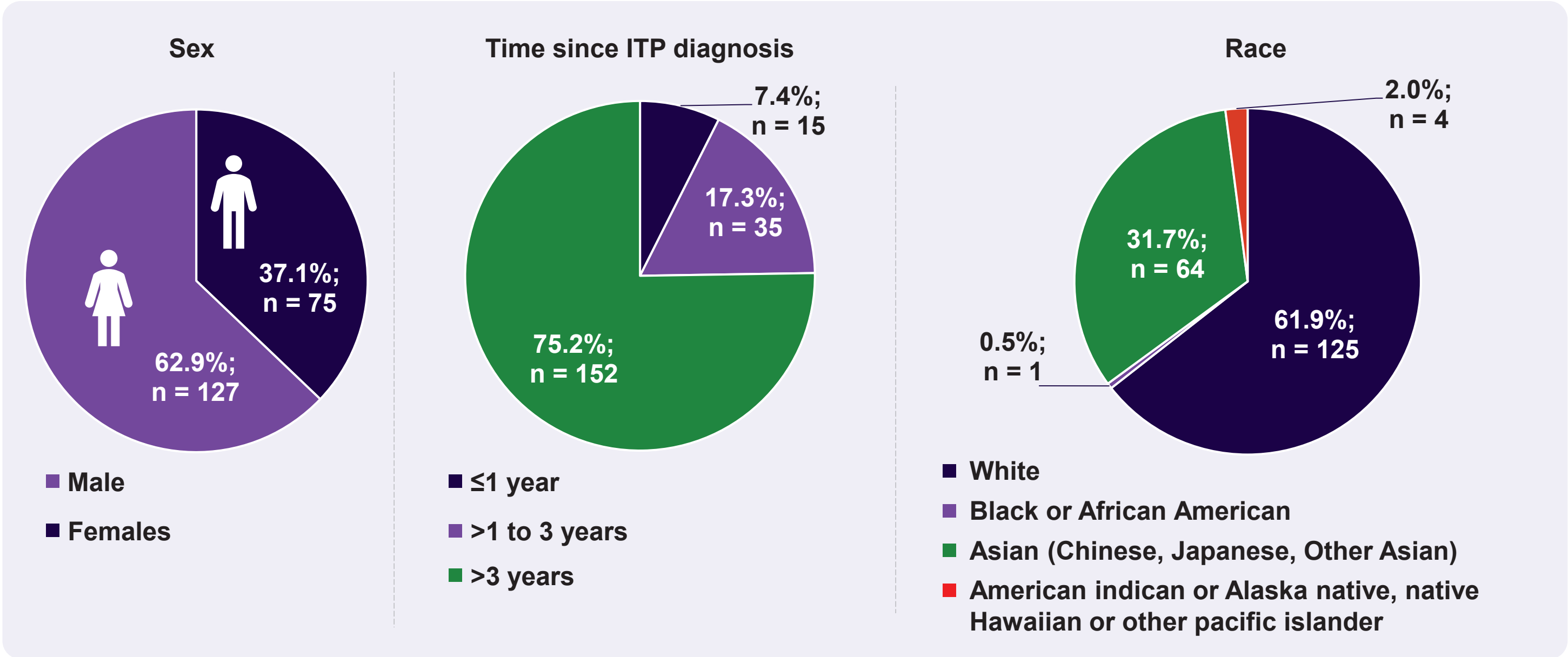
- Blinded data from the LUNA 3 phase 3 trial collected across timepoints (baseline to Week 25) were analysed for the intent-to-treat population.
- The psychometric evaluation assessed the distributional properties of scores and items, test-retest reliability, construct validity (convergent correlations and known-groups comparisons), and the ability to detect change.
- The study also aimed to derive between-group meaningful score difference (MSD) thresholds used to define clinically meaningful change.<sup>3</sup> Between-group MSD thresholds were established using anchor-based methods and supported by distribution-based analysis.
- Between group mean change scores and mixed-models for repeated measures (MMRMs) were used to estimate MSDs, a correlation weighted average was then calculated with 95% confidence intervals which guided the threshold range.

## RESULTS

### Study population

- The analysis sample comprised 202 patients (mean age: 46.9 years [range: 18–80]; female patients: n = 127; 62.9%; **Figure 1**).
- The mean platelet count at baseline was 15.0  $\times$  10<sup>3</sup>/ $\mu$ L.
- The patients were predominantly White (n = 125; 61.9%) and non-Hispanic/Latino (n = 156; 77.2%).

Figure 1. Baseline characteristics of patients



### Reliability

- Test-retest reliability was assessed between screening and baseline, and Week 5 and Week 9 using Patient Global Impression of Severity (PGIS), PGIS-Fatigue (PGIS-F) and Patient Global Impression of Change scales (PGIC).
- ITP-PAQ Item-10 exhibited good test-retest reliability ( $\kappa$  = 0.69–0.75) and IBLS total score exhibited moderate test-retest reliability (intra-class correlation coefficient = 0.65–0.75).

### Convergent validity

- Convergent correlations (assessed at Week 9) of ITP-PAQ Item-10 with PGIS and relevant EQ-5D items were moderate-to-strong ( $>|0.40|$ ); all hypothesized thresholds were met or exceeded (**Table 1**).
- The correlations between IBLS total score and PGIS and EQ-5D Visual Analogue Scale scores were smaller ( $r$  = 0.36 and 0.24, respectively), but met the hypothesized threshold, demonstrating convergent validity.

Table 1. Convergent correlations between study measures at Week 9 (N = 195)

	EQ-5D-5L			EQ-5D VAS	PGIS	PGIS-Fatigue
	Mobility	Self-care	Usual activities			
IBLS total score				–0.24	0.36	
ITP-PAQ Item-10	–0.57	–0.49	–0.58	0.48	–0.61	–0.73

### Known-group comparison

- Known-group comparisons at Week 9 confirmed the ability of both scores to differentiate between distinct PGIS severity levels (**Table 2**).
- A clear monotonic decrease in ITP-PAQ Item-10 scores and increase in IBLS total scores was observed from none to moderate, with a plateau between moderate and severe/very severe.
- ITP-PAQ Item-10 scores showed significant differences across PGIS and PGIS-Fatigue severity groups ( $p < 0.001$ ).
- IBLS total scores demonstrated that bleeding symptoms worsened significantly with increasing PGIS severity ( $p < 0.001$ ).

## CONCLUSIONS



The findings support the reliability, validity and ability of ITP-PAQ Item-10 and IBLS total score to detect improvement.



The data substantiate their use as key secondary endpoints in clinical trials in ITP.



The between-group MSD thresholds derived support the interpretation of a clinically meaningful improvement in symptoms.

Table 2. Known-groups comparisons for the ITP-PAQ Item 10 and IBLS Total score at Week 9.

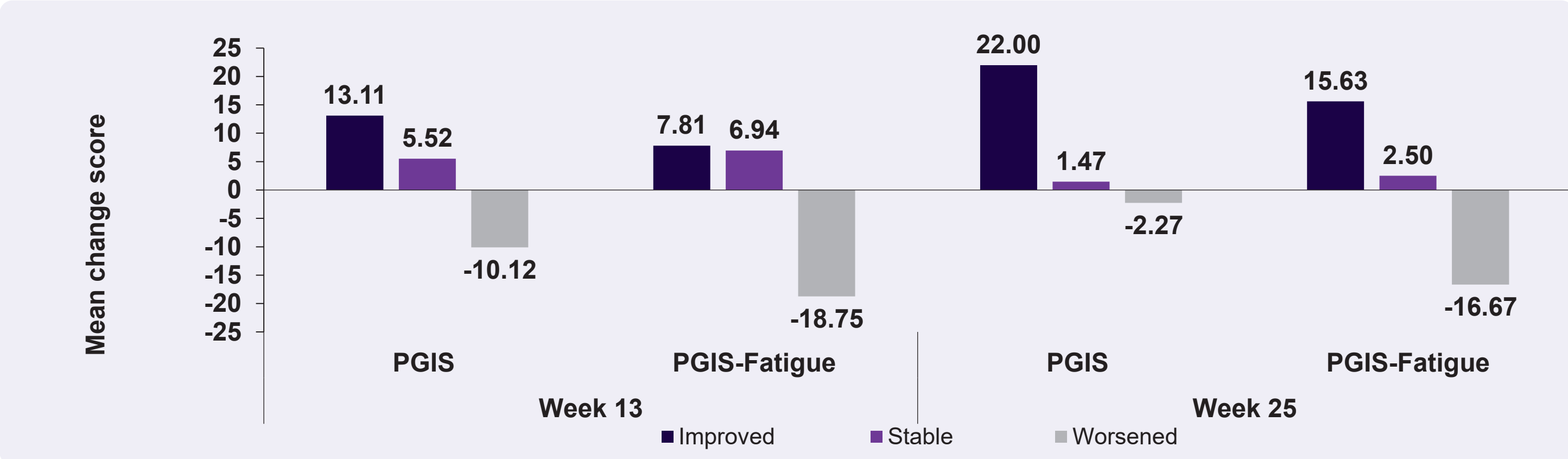
	N	Mean (SD)	Between-group effect size	Pairwise p-value	One-way ANOVA p-value
ITP-PAQ Item-10					
PGIS					
None	55	77.73 (21.88)			<0.001
Mild	67	57.09 (26.04)	−0.85	<0.001	
Moderate	51	35.78 (24.62)	−1.81	<0.001	
Severe/Very severe	11	36.36 (30.34)	−1.77	0.946	
PGIS-Fatigue					
None	18	88.89 (15.39)			<0.001
Mild	24	61.46 (22.09)	−1.41	<0.001	
Moderate/Severe/Very severe	15	41.67 (24.40)	−2.37	0.013	
Clinical anchor – Platelet count					
≥50k/μL	44	56.82 (29.22)			0.148
30–49k/μL	45	62.22 (26.98)	0.19	0.734	
20–29k/μL	24	62.50 (28.55)	0.20	0.968	
<20k/μL	69	50.72 (30.91)	−0.20	0.314	
IBLS total score					
PGIS					
None	55	0.07 (0.11)			<0.001
Mild	68	0.14 (0.16)	0.50	0.014	
Moderate	51	0.25 (0.24)	0.96	0.011	
Severe/Very severe	11	0.24 (0.26)	1.16	0.901	
Clinical anchor – Platelet count					
≥50k/μL	45	0.04 (0.10)			<0.001
30–49k/μL	44	0.12 (0.14)	0.64	0.010	
20–29k/μL	25	0.15 (0.14)	0.91	0.436	<0.001
<20k/μL	74	0.26 (0.25)	1.062	0.068	

Higher ITP-PAQ scores indicate a better quality of life. Higher IBLS scores indicate worse symptoms. The between-group effect size was calculated using Hedge's  $g$  for each group against the reference group (best health state). Hedge's  $g$  was calculated as the difference in the means divided by the pooled SD. The pairwise comparisons were tested using t-tests with p-values adjusted for multiple comparisons using Holm's correction method. The statistical significance ( $p < 0.05$ ) of differences in scores between groups was calculated using the F-test of one-way ANOVA.

### Sensitivity to change

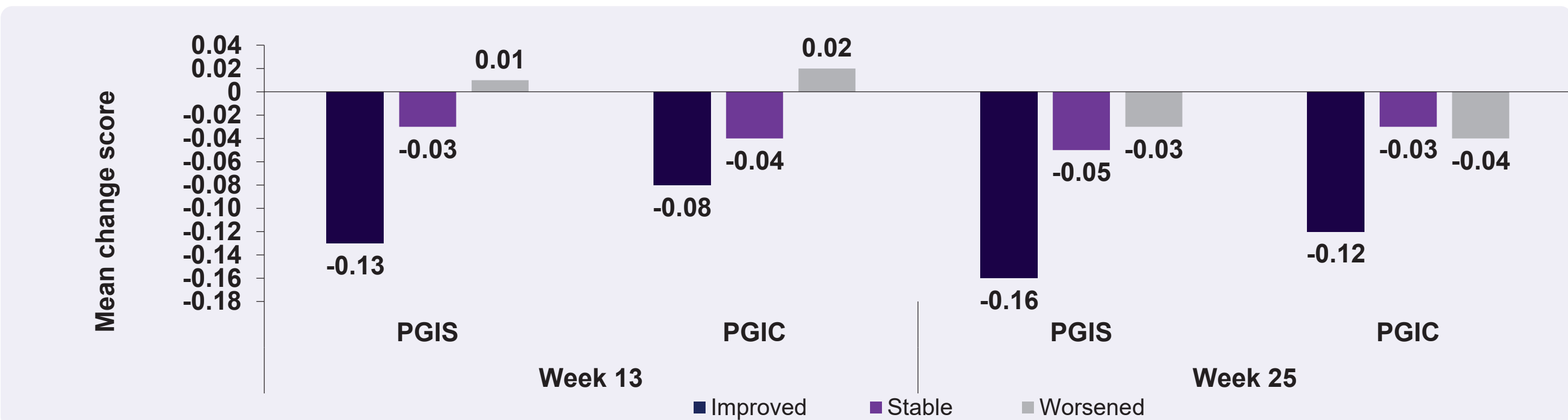
- Larger within-group effect sizes for means of 'improved' compared to 'stable' participants supported the ability of ITP-PAQ Item-10 (**Figure 2A**) and IBLS total score (**Figure 2B**) to detect improvement.

Figure 2. A) Sensitivity to change of ITP-PAQ Item-10 based on the mean change from baseline to Week 13 and Week 25 by PGIS, and PGIS-Fatigue



	Week 13						Week 25					
	PGIS			PGIS-Fatigue			PGIS			PGIS-Fatigue		
Number of patients	61	77	42	16	18	12	25	34	11	8	10	3
Within-groups ES	0.45	0.19	–0.32	0.37	0.24	–0.70	0.76	0.05	–0.08	0.60	0.10	–0.58
Between-groups ES	-	–0.31	–0.75	-	–0.04	–1.16	-	–0.84	–0.15	-	–0.71	–0.82
One-way ANOVA p-value	<0.001			0.007			0.003			0.107		

B) Sensitivity to change of IBLS total score based on the mean change from baseline to Week 13 and Week 25 by PGIC, and PGIS.



	Week 13						Week 25					
	PGIS			PGIC			PGIS			PGIC		
Number of patients	61	77	40	81	83	17	25	34	11	45	20	5
Within-groups ES	–0.52	–0.16	0.07	–0.45	–0.20	0.08	–0.68	–0.30	–0.20	–0.60	–0.14	–0.18
Between-groups ES	-	0.57	0.22	-	0.20	0.32	-	0.57	0.14	-	0.48	–0.05
One-way ANOVA p-value	<0.001			0.121			0.059			0.173		

Higher ITP-PAQ scores indicate a better quality of life. Higher IBLS scores indicate worse symptoms. The between-group ES was calculated using Hedge's  $g$  for each group against the reference group (best health state). Hedge's  $g$  was calculated as the difference in the means divided by the pooled SD. The within-group ES was calculated as the mean change score divided by the SD of the score at the earlier of the two time points. The statistical significance of differences in scores between groups was calculated using the F-test of one-way ANOVA.

### Between-group MSD threshold

- The between-group MSD thresholds for ITP-PAQ Item-10 and IBLS total score ranged from 8 to 18 points and 0.05 to 0.10 points, respectively.

## REFERENCES

- Cooper N, Ghanima W. *N Engl J Med*. 2019;381(10):945–955.
- Kuter DJ, Ghanima W, Cooper N, et al. *Blood*. 2025;145(24):2914–2926.
- Food and Drug Administration (FDA). Patient-Focused Drug Development: Incorporating Clinical Outcome Assessments Into Endpoints for Regulatory Decision-Making (PFDD Guidance 4 [draft]). Available at: <https://www.fda.gov/media/166830/download>. Published 2023. Accessed 28 April 2025.

## ABBREVIATIONS

ANOVA, analysis of variance; ES, effect size; IBLS, Immune Thrombocytopenia Purpura Bleeding Scale; ITP-PAQ, Immune Thrombocytopenic Purpura Patient Assessment Questionnaire; N, number of patients; PGIS, Patient Global Impression of Severity; PGIC, Patient Global Impression of Change; PGIS, Patient Global Impression of Severity; SD, standard deviation; VAS, Visual Analogue Scale.

## CONFLICTS OF INTEREST

Imene Gouia is an employee of Sanofi may hold stock and/or stock options in Sanofi. Mattea Orsini was an employee of Aixial, and a consultant for Sanofi when this study was conducted. Joshua Maher was an employee of Sanofi when this study was conducted and may hold stock and/or stock options in Sanofi. George Skingley, Asha Lehane and Helena Bradley are employee of Adelphi Values, Patient Centered Outcomes.

## ACKNOWLEDGEMENTS AND FUNDING

Deepshikha Roy and Deepti Sharda from Sanofi provided medical writing and editorial assistance during the development of this poster. The authors thank Karuna Naik from Sanofi for her contributions to planning, review, and coordination of this poster. This study was funded by Sanofi.