

Computer vision endoscopy scoring for ulcerative colitis disease severity (ARGES-CMES): A comparison between adult and paediatric clinical trials

N. Alves¹, K. Chaitanya², P. Mobadersany³, C. Parmar³, A. Griffiths⁴, R.K. Russell⁵, D. Turner⁶, J. Hyams⁷, L. Ghanem³, C. Richmond³, E. Van Limbergen⁸, P. Ufberg³, R. Strauss³, N. Skomrock³, G.O. Cula³, K. Standish³, P. F. Damasceno³.

[1] Janssen R&D, LLC- a Johnson & Johnson Company, The Netherlands; [2] Janssen R&D, LLC- a Johnson & Johnson Company, Switzerland; [3] Janssen R&D, LLC- a Johnson & Johnson Company, USA. [4] Hospital for Sick Children, University of Toronto, Canada; [5] Royal Hospital for Children & Young People, United Kingdom; [6] Shaare Zedek Medical Center, Israel. [7] Connecticut Children's Medical Center, USA; [8] Janssen R&D, LLC- a Johnson & Johnson Company, Belgium.

Background and Objective

A standardized, objective, and finely granular endoscopic scoring method for UC that is validated across adult and paediatric clinical trial populations is presently lacking.

- Paediatric IBD trials are often small, open-label, difficult to recruit, and challenging to interpret.
- Endoscopic endpoints are considered the most-objective signs of improvement. However, the current Mayo endoscopic scoring (MES) system is coarse (relying on a 4-point scale), subjective (relying on manually performed central reads), and limited in sensitivity.
- Novel computer-vision derived endpoints can improve upon MES but lack validation on paediatric populations to support extrapolation and regulatory decision-making.
- We compared **ARGES-CMES**, a previously developed novel computer-vision endpoint with improved sensitivity, to MES for endoscopy scoring across adult and paediatric clinical trials.
- The goal of this study was to generate robust evidence to support reliable extrapolation and enhance detection of treatment effects.

Conclusions

ARGES-CMES demonstrates similar temporal improvements and consistent dose-response patterns in adults and children:

1. Similar improvements in endoscopic disease severity (ARGES-CMES) between adult (UNIFI) and paediatric (UNIFI-JR) populations following induction and maintenance phase.
2. Using ARGES-CMES, relative improvement between screening and end of induction follow similar trend in UNIFI and UNIFI-JR.
3. This model provides compelling evidence supporting extrapolation of novel endoscopic endpoints from adult to paediatric clinical trial populations.
4. Greater precision in estimates of endoscopic endpoints with ARGES-CMES may facilitate smaller trial sample size.

These findings provide robust evidence to adopt a standardized, automated, and more sensitive endoscopic assessment in paediatric UC trials, and substantiate the responsible extrapolation of adult trial outcomes to paediatric populations.

Methods

The **ARGES-CMES** endpoint was developed exclusively using adult UC clinical trial populations to improve the granularity of the categorical MES by providing a continuous score of disease severity across the colon (Figure 1) [1,2].

The ARGES-UC framework [3,4], previously trained on colonoscopy videos from adult UC trials, was used without any parameter tuning to compute ARGES-CMES across two phase 3 clinical trials assessing ustekinumab (UST) as treatment of UC in adult (UNIFI, NCT02407236) and paediatric (UNIFI-JR, NCT04630028) patients.

To allow for comparisons between the placebo-controlled UNIFI and the open-label UNIFI-JR trials, UNIFI participants receiving placebo were excluded from the analysis.

ARGES-CMES scores were compared at screening, week 8 (end of induction) and week 52 (end of maintenance). No adjustments for differences in baseline characteristics were made. Two-sided permutation tests were used to evaluate differences in the median CMES scores between the two trials at each time point.

Standard endpoint: MES Novel endpoint: ARGES-CMES

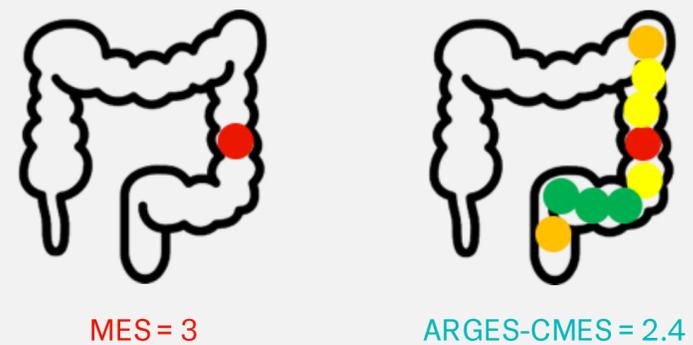


Figure 1. The ARGES-CMES endpoint measures disease activity across endoscopy video frames in the descending colon, sigmoid, and rectum, producing a comprehensive, continuous severity score. In contrast, the traditional MES score only considers the worst disease severity in the colon using a categorical scale between 0 and 3.

Results

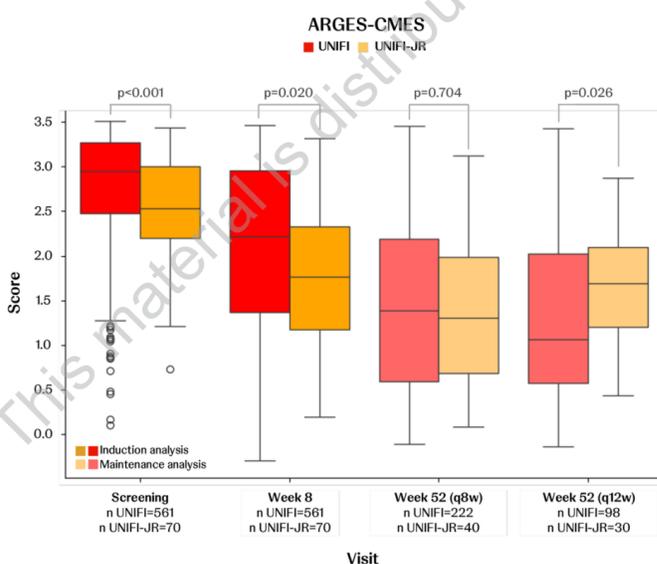


Figure 2. ARGES-CMES score distribution across visits for UNIFI and UNIFI-JR.

Table. Mean and standard deviation of ARGES-CMES score across trials and visits, as well as the average change in score between the screening visit and week 8.

Visit	UNIFI		UNIFI-JR	
	N	CMES (mean ± std)	N	CMES (mean ± std)
Screening	561	2.78 ± 0.62	70	2.55 ± 0.57
Week 8	561	2.10 ± 0.97	70	1.72 ± 0.77
Week 52 q8w	222	1.46 ± 0.98	40	1.36 ± 0.85
Week 52 q12w	98	1.30 ± 0.91	30	1.62 ± 0.68
Change Screening -Week 8	561	0.68 ± 0.88	70	0.83 ± 0.96

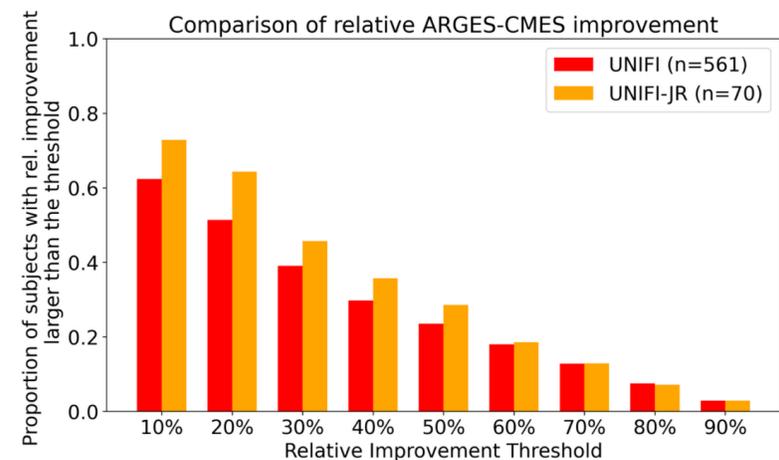


Figure 3 ARGES-CMES relative improvement between screening and week 8 for UNIFI and UNIFI-JR



This QR code is intended to provide scientific information for individual reference and the information should not be altered or reproduced in any way.