ARGES-Ulcer$_{Surface\text{-}Score}$: A continuous AI-derived SES-CD subcomponent for Crohn's Disease severity assessment and clinical endpoint recapitulation

AUTHORS: **Lukas Hirsch**[2], Krishna Chaitanya[1], Fabio Gunderson[2], Pooya Mobadersany[2], Chaitanya Parmar[2], Lindsey Surace[2], Shadi Yarandi[2], Louis R Ghanem[2], Yuliya Halchenko[2], Nicholas Skomrock[2], Tommaso Mansi[2], Gabriela Oana Cula[2], Kristopher Standish[2], Pablo F Damasceno[1]. AFFILIATIONS: [1] Janssen Research & Development, LLC, Switzerland, [2] Janssen Research & Development, LLC, USA, [3] University of Michigan Medicine, Ann Arbor, MI, United States.

DDW2026 Digestive Disease Week® MAY 2-5, 2026 | CHICAGO, IL EXHIBIT DATES: MAY 3-5, 2026

Background

- The Simple Endoscopic Score for Crohn's Disease (SES-CD) is a standard tool for assessing endoscopic disease severity in Crohn's disease clinical trials.
- Ulcerated surface is a categorical subscore from SES-CD that estimates the percent mucosal ulceration across intestinal segments.
- We previously developed an AI pipeline to automatically segment ulcerative lesions in each frame of ileocolonoscopy videos.
- Translating frame-level ulcer segmentation into an accurate clinical metric is required to recapitulate the SES-CD ulcerated surface subscore.
- AI-based automation of this subscore could improve consistency, decreasevariability, and speed up disease assessment in clinical trials.

Objective

To develop and validate an AI-based pipeline (ARGES-Ulcer) that automatically computes the SES-CD ulcerated surface subscore (ARGES-Ulcer$_{Surface\text{-}Score}$) from ileocolonoscopy videos.

Methods

- A deep neural network was trained on human ulcer segmentations to automatically segment ulcers in ileocolonoscopies (see details in poster **Tu1497** and presentation **#88**).
- AI-derived ulcer segmentation as well as pooled statistical features (mean, max, standard deviation, ulcer/non-ulcer frame ratio) from SEAVUE [3] Phase 3b (NCT03464136) were used as model input to a shallow AI model trained to assess the SES-CD ulcerated surface subscore (Fig. 1).
- Frame-level ulcer percentages were computed for five anatomical regions (ileum, ascending, transverse, sigmoid/descending colon, and rectum) by measuring the ratio between pixels belonging to ulcers and all pixels belonging to a colon region;
- Training optimized Mean Absolute Error (MAE) relative to the human-assigned ulcerated surface subscore.
- Validation was performed on an independent GALAXI [4] Phase 3 (NCT03466411) hold-out dataset to assess score recapitulation, clinical correlations, and disease state differentiation.

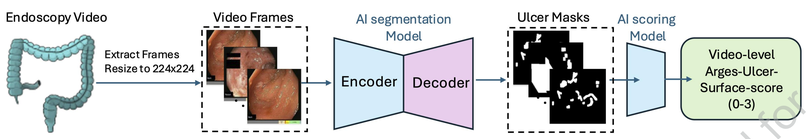

Figure 1: Automated estimation of SES-CD ulcerated surface from ileocolonoscopy video. Frames are extracted from an ileocolonoscopy video and processed by ARGES-Ulcer: an AI ulcer segmentation model to generate pixel-level ulcer masks for each frame. Ulcer mask areas across all frames are then fed into a shallow model that recapitulates the SES-CD ulcerated surface subscore (Arges-Ulcer$_{Surface\text{-}Score}$) at the video level.

https://www.congresshub.com/Immunology/DDW2026/Ustekinumab/Hirsch

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

Results

a) Changes in Arges-Ulcer$_{Surface\text{-}Score}$ show strong, significant associations with most changes in endoscopic, clinical, biomarker, and histologic disease activity.

Change in Clinical Measures	Change in ARGES-Ulcer $_{surface\text{-}score}$	
Disease Severity Scores	**Correlation Coefficient**	**p-value**
SES-CD and subscores	0.59	<0.001
Ulcerated surface	0.58	<0.001
Affected surface area	0.57	<0.001
Presence and Size of Ulcers	0.59	<0.001
Narrowing	0.07	0.048
CDAI score	0.35	<0.001
Biomarkers		
Fecal calprotectin (FCP)	0.43	<0.001
C-reactive protein (CRP)	0.33	<0.001
Histological Score		
Total GHAS Score	0.55	<0.001

Table 1: Association between change in ARGES-Ulcer$_{Surface\text{-}Score}$ and change in disease severity, biomarker and histological scores on the prospective GALAXI CD cohort. Reported values are Spearman correlation coefficients and two-tailed p-values.

b) Lower Arges-Ulcer-surface-score is associated with clinical remission at Week 48.

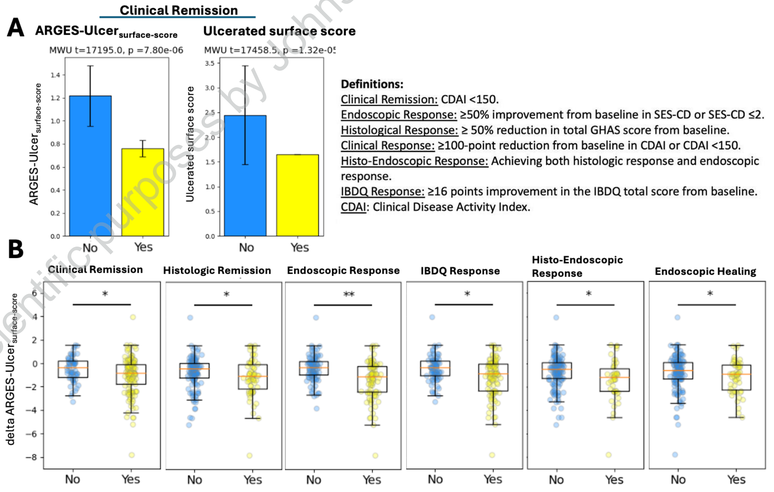

Figure 2: Arges-Ulcer-$_{Surface\text{-}Score}$ and its change can differentiate disease status in CD trials based on several clinical endpoints. A: Both Arges-Ulcer-$_{Surface\text{-}Score}$ and the standard SES-CD ulcerated surface score successfully distinguish patients based on their clinical remission status at week 48. **B:** Decrease in Arges-Ulcer-surface-score from baseline to week 48, for patients achieving a positive (yellow) or negative (blue) outcome in each respective clinical endpoint. Statistical significance was assessed with the Mann-Whitney U-test (* for p<0.05, ** for p<1e-3).

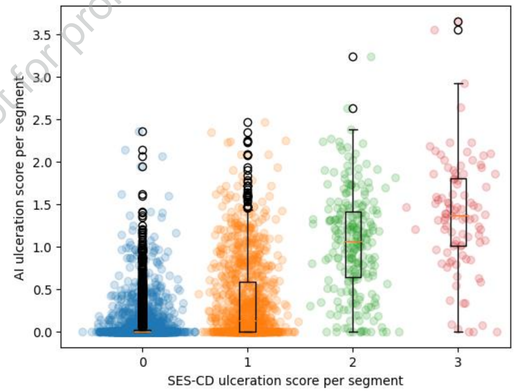

Figure 3: Predicted ulceration scores from ARGES-Ulcer$_{Surface\text{-}Score}$ in the GALAXI test set vs the ground truth ulcerated surface SES-CD scores (mean absolute error = 0.39, spearman correlation = 0.55, p<1e-20)

Discussion

- ARGES-Ulcer$_{Surface\text{-}Score}$ closely recapitulated the SES-CD ulcerated surface subscore with low mean absolute error (Fig 3).
- Strong correlations were observed between ARGES-Ulcer$_{Surface\text{-}Score}$ and both SES-CD components and inflammatory biomarkers (Table 1).
- Lower scores were associated with endoscopic remission and favorable disease status (Fig. 2 & Table 2).
- Longitudinal changes in the score aligned with multiple clinical, endoscopic, and quality-of-life response measures.

Conclusion

- ARGES-Ulcer$_{Surface\text{-}Score}$ provides an automated, objective approximation of the SES-CD ulcerated surface subscore.
- This approach may enhance reproducibility and resolution of disease activity assessment in Crohn's Disease clinical trials.

References

[1] Cai, Lingrui, et al. "Adapting Segment Anything Model for Ulcer Segmentation in Inflammatory Bowel Disease." 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2024.
[2] Cai, Lingrui, et al. "Artificial Intelligence for Quantifying Endoscopic Mucosal Ulceration in Crohn's Disease." Clinical Gastroenterology and Hepatology (2025).
[3] Sands, Bruce E et al. "Ustekinumab versus adalimumab for induction and maintenance therapy in biologic-naive patients with moderately to severely active Crohn's disease: a multicentre, randomised, double-blind, parallel-group, phase 3b trial." Lancet (London, England) vol. 399,10342 (2022): 2200-2211. doi:10.1016/S0140-6736(22)00688-2.
[4] Panaccione, Remo, et al. "Efficacy and safety of intravenous induction and subcutaneous maintenance therapy with guselkumab for patients with Crohn's disease (GALAXI-2 and GALAXI-3): 48-week results from two phase 3, randomised, placebo and active comparator-controlled, double-blind, triple-dummy trials." The Lancet 406.10501 (2025): 358-375.

Acknowledgements & Disclosures

We would like to acknowledge the clinical and statistical programming colleagues (Tadeusz Lewandowski, Jie Wang, Tamsin Sargood, Hongyan Zhang, and Vincent Guo) for assistance with data preparation.

This work was supported by Janssen Research & Development, LLC.

All development of AI models and analyses presented in this poster has been conducted internally within Janssen. L.Hirsch, K. Chaitanya, F. Gunderson,, P. Mobadersany, C. Parmar, Y. Halchenko, N. Skomrock, S. Yarandi, L. Surace, L.R. Ghanem, T. Mansi, G.O. Cula, K. Standish, P. Damasceno are employees of Johnson & Johnson and own company stock/stock options.

CONTACT: Pablo Damasceno (pdamasc1@its.jnj.com)

Johnson&Johnson