A VIRTUAL CHROMOENDOSCOPY ARTIFICIAL INTELLIGENCE SYSTEM TO DETECT ENDOSCOPIC AND HISTOLOGIC REMISSION IN ULCERATIVE COLITIS

M. Iacucci , R. Cannatelli , T.L. Parigi , A. Buda , N. Labarile , O.M. Nardone , G.E. Tontini , A. Rimondi , A. Bazarova , P. Bhandari , R. Bisschops , G. De Hertogh , R. del Amor , J.G Ferraz , M. Goetz , X. Gui , B. Hayee , R. Kiesslich , M. Lazarev , V. Naranjo , R. Panaccione , A. Parra-Blanco , L. Pastorelli , T. Rath , E.S Røyset , M. Vieth , V. Villanacci , D. Zardo , S. Ghosh , E. Grisan

**Aims** We aimed to develop an artificial intelligence (AI) system to assess endoscopic remission (ER) and histologic remission (HR) of ulcerative colitis (UC) in both white light (WL, using Ulcerative Colitis Endoscopic Index of Severity [UCEIS]) and virtual chromoendoscopy (VCE, using Paddington International Virtual ChromoendoScopy ScOre [PICaSSO]).

**Methods** A convolutional neural network (CNN) was developed based on 559 endoscopy videos, from 302 UC patients prospectively included in the PICaSSO multicentre study. The videos were divided in training (254), validation (62), and testing (243), and comprised 67280 frames in total. The CNN was trained to predict both ER (defined as UCEIS≤1 in WL and as PICaSSO≤3 in VCE) and HR (defined as Robarts Histological Index ≤ 3 with no neutrophils in lamina propria or epithelium) in video clips.

**A picture containing plastic

Description automatically generated**

**Results** In the validation cohort, our system predicted ER in WL videos with 82% sensitivity, 94% specificity and an area under the ROC curve (AUROC) of 0.92. In VCE, sensitivity was 74%, specificity 95%, and AUROC 0.95. In the testing cohort, the diagnostic performance remained similar. The diagnostic performance for the prediction of HR in the validation set had sensitivity, specificity, and accuracy of 92%, 83%, and 85%, respectively, using VCE; and 83%, 87%, and 86% respectively, with WL. In the testing set, these metrics declined modestly while remaining good. Of note, the algorithm’s prediction of histology was similar with VCE and WL videos.

**Conclusions** Our AI system accurately recognizes ER in videos and predicts HR equally well.

