**Introduction**:

Machine learning algorithms (MLA) trained on capsule endoscopy (CE) images have shown high accuracy in detecting signs of villous atrophy in patients with coeliac disease. However, using MLA to quantify coeliac disease extent and the proportion of small bowel affected has not been studied.

**Methods**:

We analysed CE videos of adult patients (≥18 years) with biopsy-proven coeliac disease from a large prospectively maintained database. Demographic, clinical and biochemical data were collected for all patients. The extent of coeliac disease and the proportion of small bowel affected were calculated for each video by an expert human reader and by a trained MLA. The consistency of agreement between the expert human and the MLA was estimated by intraclass correlation coefficients (ICCs) using a two-way mixed-effects model.

**Results**:

CE videos of 50 patients with coeliac disease (median age 56.5 years [IQR 36 – 65 years], 64% female) were included. The median gastric passage time and small bowel transit time were 18 minutes (IQR 9 – 65 minutes) and 273 minutes (IQR 213 – 319 minutes), respectively. Macroscopic features of coeliac disease, including scalloping, fissuring, mosaicism, nodularity, erosions and ulcers, were identified beyond the first tertile in 42% of patients (extensive disease). The overall agreement between the expert human and MLA for the proportion of small bowel affected was good (ICC 0.78, 95% CI 0.62-0.87). Patients with extensive small bowel disease were significantly more likely to have nutritional deficiencies compared with those with proximal disease (p=0.002).

**Conclusion**:

Trained MLA can accurately quantify the extent of coeliac disease and the proportion of small bowel involvement on CE, showing a good agreement with expert human assessment. This machine learning quantification of coeliac disease extent allows closer monitoring for nutritional deficiencies associated with extensive disease and provides an objective and standardised endpoint on response to therapy in clinical trials. Larger training and validation datasets are required to optimise the performance of the MLA.