Introduction Mismatch Repair (MMR) deficient colorectal cancers are an important cohort to study as these patients have a high neo-epitope load, suggesting they may respond well to immunotherapy. We previously reported deficient MMR due to loss of MLH1 expression occurs in >20% of all inflammatory bowel disease-associated colorectal cancers (IBD-CRC). Manually identifying MLH1 deficient tumours from immunostained samples is time-consuming and labour-intensive. Therefore, we aimed to train QuPath, a digital image analysis program, to accurately identify MLH1 deficient IBD-CRCs from a tissue microarray (TMA) containing normal colon and IBD-CRC.

Methods A TMA (n=162 cores) containing normal colon and IBD-CRC was immunostained for MLH1, digitalised and imported into QuPath. A representative sample (14 sections of 500μm x 500μm) was used to train a QuPath algorithm to identify tissue histology (normal epithelium, tumour, immune infiltrate or stroma) and MLH1 expression pattern. The algorithm was applied to the whole TMA. Accuracy level for histology and MLH1 classification was set at >75% of cells per category; cores where neither category met this threshold were flagged for review and these cores were histopathologically reviewed qualitatively. Data were exported to IBM®SPSS® (V25.0) for statistical analysis.

Results Tissue histology and MLH1 expression was correctly identified in 87/113 (77%) cases, misclassified in 1/113 (0.9%) case and flagged for manual review in 25/113 (22.1%) cases. Reasons cores were flagged for review included: small quantity of tissue, atypical morphology, excessive immune infiltrate, normal epithelium and patchy immunostaining. Of cores not flagged for review (n=88), the algorithm was highly sensitive (100%) and specific (98.6%) for identifying MLH1-deficient IBD-CRCs, [1-0.967 (95%CI 0.902-1.032)] (p<0.001).

Conclusions MLH1 deficient IBD-CRC can be accurately identified using our QuPath algorithm, with an acceptable proportion of equivocal samples highlighted for manual evaluation. This process could be efficiently automated in conventional NHS IBD-CRC surveillance and treatment programmes to examine all colonic tissue and tumour specimens for MLH1 expression, thus identifying patients who may respond well to immunotherapy. This approach now needs to be validated using endoscopic biopsies.