Supplementary Information for:

First experience in clinical application of hyperspectral endoscopy for evaluation of colonic polyps

Jonghee Yoon[1,2], James Joseph[1,2], Dale J. Waterhouse[1,2], Charlie Borzy[4], Kyla Siemens[4], Sarah Diamond[3], V. Liana Tsikitis[4,*] and Sarah E. Bohndiek[1,2,*]

[1] Department of Physics, University of Cambridge, JJ Thomson Avenue, Cambridge CB3 0HE, United Kingdom

[2] Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge CB2 0RE, United Kingdom

[3] Department of Medicine, Oregon Health and Science University, Portland, Oregon, 97239, United States

[4] Department of Surgery, Oregon Health and Science University, Portland, Oregon 97239, United States

JY is now at Department of Physics, Ajou University, Suwon, 16499, South Korea.

JJ is now at School of Science and Engineering, Fulton Building, University of Dundee, Dundee DD1 4HN, United Kingdom.

*Corresponding authors, e-mail: tsikitis@ohsu.edu, seb53@cam.ac.uk

Supplementary Methods

Photobiological safety evaluations

In order to address photobiological safety of HySE, we compared the spectral irradiance and illumination angle of HySE to that of an FDA-approved clinical colonoscope (Olympus Colonoscope CF-H290). For the spectral irradiance calculation, spectral profiles and optical power were measured. Spectral profiles of a supercontinuum source coupled to HySE and illumination light of the colonoscope were measured using a spectrometer (AvaSpec–ULS2048L, Avantes). For the optical power measurement in the worst-case scenario, we used a calibrated thermal power meter (A-01-D12-USB, Laserpoint) and put the distal end of the HySE or clinical colonoscope directly against the surface of the power meter. Then, spectral irradiance was calculated by the following equation:

$$\text{Spectral irradiance}(\lambda) = (S(\lambda)\times P)/(A\times D)$$

, where $S(\lambda)$ is the signal (a.u.) at a wavelength of λ measured by the spectrometer, P is an optical power (mW) of the illumination measured by the calibrated thermal power meter, A is the total spectral area under the curve (a.u.$\times$nm), and D is the area (cm^2) of the calibrated thermal power meter. The spectral irradiance of the compact supercontinuum source was within that of the clinically approved colonoscope (Fig. 1d).

For the measurement of the illumination angle, we measured diameter of the illumination light at a known working distance (Fig. 1e). The illumination angle was determined by the following equation:

$$\alpha = \tan^{-1}(r/WD)\times 180/pi$$

, where α is a half illumination angle, WD is working distance, and r is a radius of the illuminating area. The measured illumination angles of HySE and the colonoscope were 53.2° – 65.0° and > 150°, respectively, which means that HySE illumination was again within that of the clinical colonoscope.

Photothermal safety was also tested to check whether HySE could damage on epithelial tissue. The surface temperature of the polyscope was measured for 1 hour every 10 minutes using a non-contact thermometer (RS-8662, RS Components). The measured surface temperature of the polyscope coupled to the supercontinuum source and room temperature were 22.09 ± 0.26 and 22.1 ± 0.17°C respectively, which demonstrates that HySE will not introduce any thermal risks related to high surface temperature during the endoscopy procedure (Fig. 1f).

As the supercontinuum laser is a nanosecond laser source (pulse duration: ~ 1 ns, pulse duration: ~10kHz), the single-pulse energy was calculated to confirm laser safety under clinical conditions. The single-pulse energy was estimated to be 8.56×10^{-3} mJ/cm^2, which is much lower than maximum permissible exposures (0.02 J/cm^2) recommended by the American National Standards Institute and International Electotechnical Commission.

Data analysis

Matlab R2020a and imageJ were used for hyperspectral image processing. GraphPad Prism was used for statistical analysis. In total, 92 measurements obtained from eight patients were analysed, and each measurement acquired 100 spectral images consisting of 200 spectral signals. Therefore, in total $200 \times 100 \times 92$ spectra were analysed.

Wide-area hyperspectral image reconstruction

The overall hyperspectral image reconstruction process was described previously[1] and was tested for this study to evaluate whether it was sufficient to operate under clinical conditions, which provides low contrast images and random image distortions during measurement. Briefly, geometric transformation matrices (GMs) were estimated by extracting and comparing scale-invariant features from the wide-field colour images. Then, the wide-area hyperspectral image was reconstructed using the spectral images and estimated GMs. The identification of regions of interest containing normal mucosa (8 patients), polyp (5 patients), and post-resection tissue (2 patients), were confirmed by clinicians using a combination of the colonoscopy video and the wide-field colour images; locations were then mapped onto the reconstructed hyperspectral images.

Reflection and absorption calculation

In order to calculate reflectance, measured spectra were normalised according to:

$$\mathrm{Reflectance}(\mathrm{x},\mathrm{y},\lambda) = \frac{\mathrm{I}_{\mathrm{sample}}(x,y,\lambda) - I_{dark}(x,y,\lambda)}{I_{white}(x,y,\lambda) - I_{dark}(x,y,\lambda)}$$

, where $I_{sample}(x,y,\lambda)$ is the measured intensity at the wavelength of λ at the point x and y in the image, and I_{dark} is the dark counts measured from the EMCCD in the absence of illumination, and I_{white} is the intensity measured from a standard white reflectance target (Spectralon Diffuse Reflectance Standards, LabSphere) using the white light source from the standard-of-care colonoscope. I_{white} and I_{dark} were measured in every experiment to avoid variations introduced depending on experimental conditions.

Absorbance was calculated by taking the logarithm of reflectance via the following

equation:

$$\mathrm{Absorbance}(\mathrm{x},\mathrm{y},\lambda) = -\log 10\left(\frac{\mathrm{I}_{\mathrm{sample}}(x,y,\lambda) - I_{dark}(x,y,\lambda)}{I_{white}(x,y,\lambda) - I_{dark}(x,y,\lambda)}\right)$$

In hyperspectral image analysis, reflectance and absorbance were selectively exploited for better data interpretation. As the measured intensity changes depending on the distance between the illumination fibre and the target, it is difficult to accurately compute the reflectance and absorbance under *in vivo* conditions. In our previous report[2], we found that inaccurate intensity measurement alters the scale and offset of reflectance and absorbance, respectively. Therefore, reflectance was exploited for quantitative analysis, and absorbance was shown when the graph shape is important.

Coefficient of variation

The coefficient of variation (CV) was calculated to evaluate intra- / inter-patient variability in hyperspectral imaging quality. Hyperspectral images from normal mucosa (5 patients) and polyps (2 patients), which were clearly diagnosed by the clinicians, were selected for the accurate calculation of the variability. In a patient, 10 – 30 spectral images which has 10 – 20 spectral data points were tested for calculating CV. Intra- / inter-patient CVs were calculated via following equations[3]:

$$Intra - patient\ CV(\lambda) = \sum_{m=1}^{N} \frac{standard\ deviation\left(Reflectance_m(\lambda)\right)}{mean(Reflectance_m(\lambda))} \times 100$$

$$Inter - patient\ CV(\lambda) = \sum_{p=1}^{N} \frac{standard\ deviation\ (mean\left(Reflectance_p(\lambda)\right)}{mean(mean(Reflectance_p(\lambda))} \times 100$$

, where *m* and *p* indicate a spectral image number in a patient and patient number, respectively.

Spectral angle mapping

For quantitative comparison of spectra acquired from different tissue types, spectral angle mapping (SAM) was exploited[4]. The spectral angle, α, was calculated by the following equation:

$$\alpha = \cos^{-1}\left(\frac{\sum_{\lambda=1}^{n} t_\lambda r_\lambda}{\left(\sum_{\lambda=1}^{n} t_\lambda^2\right)^{0.5}\left(\sum_{\lambda=1}^{n} r_\lambda^2\right)^{0.5}}\right)$$

, where t_λ and r_λ are values of target and reference spectral signals at the wavelength of λ, respectively. *n* indicates the total number of spectral samples. The average spectra of normal tissue, polyps and post-resection tissue were used as the reference spectral signals in order to compare differences among these tissue types. Therefore, three spectral angles were calculated from one spectral signal via exploiting three reference spectral signals (Supplementary Fig 1a,b). To enhance the visualisation of SAM results, a synthetic RGB SAM image was created by assigning green, blue, and red colours to the three SAM values obtained via the reference spectral signals of normal mucosa, polyp, and post-resection tissue, respectively. However, SAM analysis shows small values if the target and reference signals are similar; thus, 1-SAM values were assigned as RGB values of a synthetic RGB SAM image to make normal mucosa, polyps, and post-resection tissue as green, blue, and red colours. (Supplementary Fig 1c-e).

Machine learning

To investigate the potential for discrimination of the different tissue types from which spectra were collected, we employed a k-nearest neighbour (k-NN, k = 4) classification model. The k-NN algorithm predicts the class of newly observed data by choosing the most frequent class labels of k-nearest neighbour data points in the feature space[5,6]. We exploited SAM values obtained by calculating three reference signals of normal mucosa, polyps, and post-resection tissue, which clearly shows a distinct distribution of spectral signals of each tissue type. The established classifier was validated via k-fold cross-validation using data that was not utilized for training. The training, cross-validation, and prediction were performed using Matlab function: *fitcknn*, *crossval*, and *predict*, respectively.

References

1 Yoon, J. et al. A clinically translatable hyperspectral endoscopy (HySE) system for imaging the gastrointestinal tract. Nature communications 10, 1-13 (2019).

2 Yoon, J., Grigoroiu, A. & Bohndiek, S. E. A background correction method to compensate illumination variation in hyperspectral imaging. Plos one 15, e0229502 (2020).

3 Dumas, M.-E. et al. Assessment of analytical reproducibility of 1H NMR spectroscopy based metabonomics for large-scale epidemiological research: the INTERMAP Study. Analytical chemistry 78, 2199-2208 (2006).

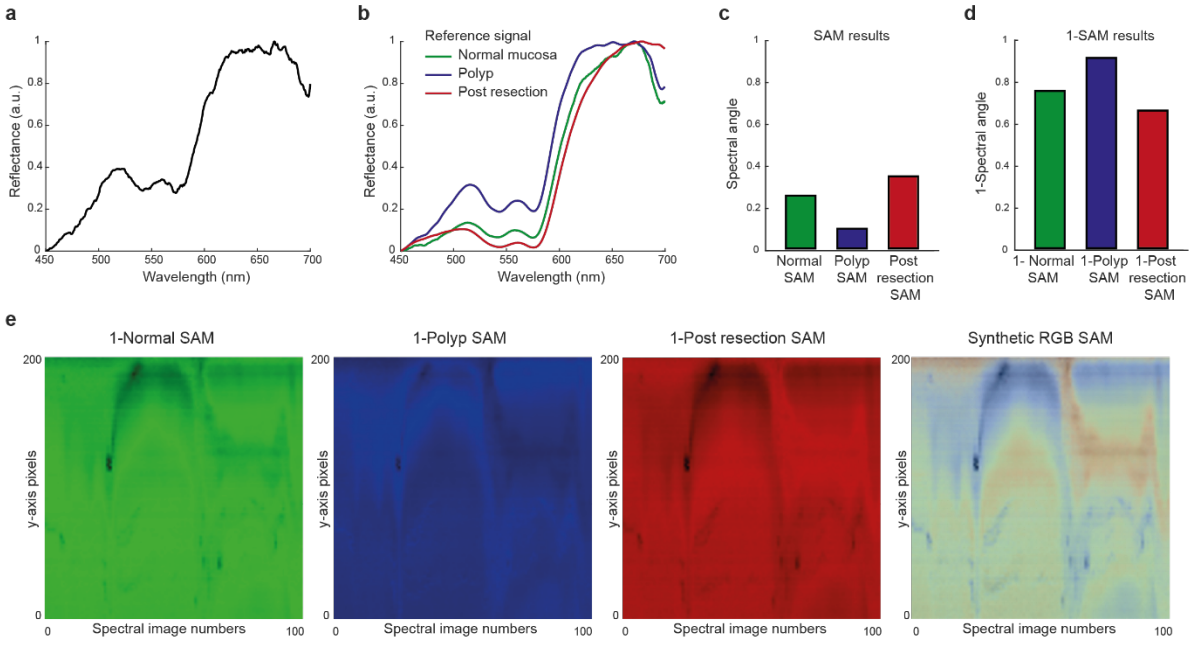
4 Park, B., Windham, W., Lawrence, K. & Smith, D. Contaminant classification of poultry hyperspectral imagery using a spectral angle mapper algorithm. Biosystems

Engineering 96, 323-333 (2007).

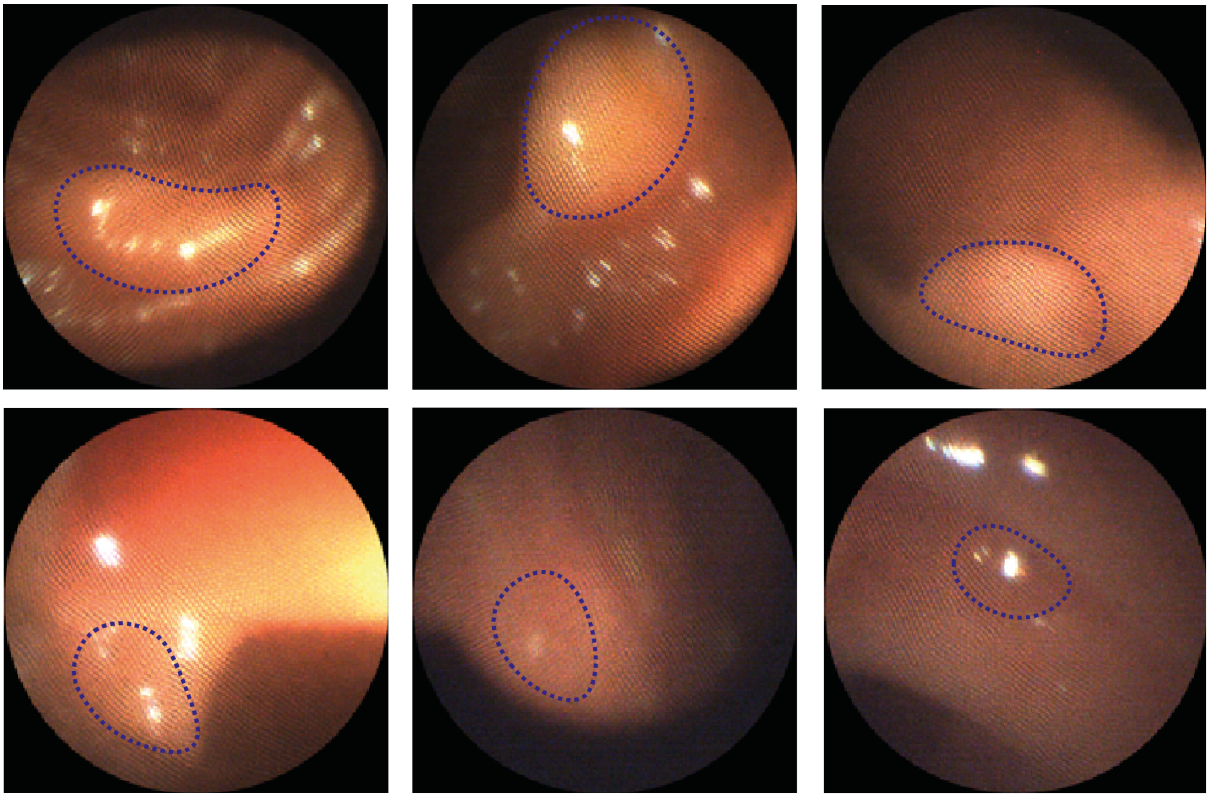
5 Lu, J. et al. MicroRNA expression profiles classify human cancers. nature 435, 834-838 (2005).

6 Chen, X., Zhou, X. & Wong, S. T. Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. IEEE Transactions on Biomedical Engineering 53, 762-766 (2006).

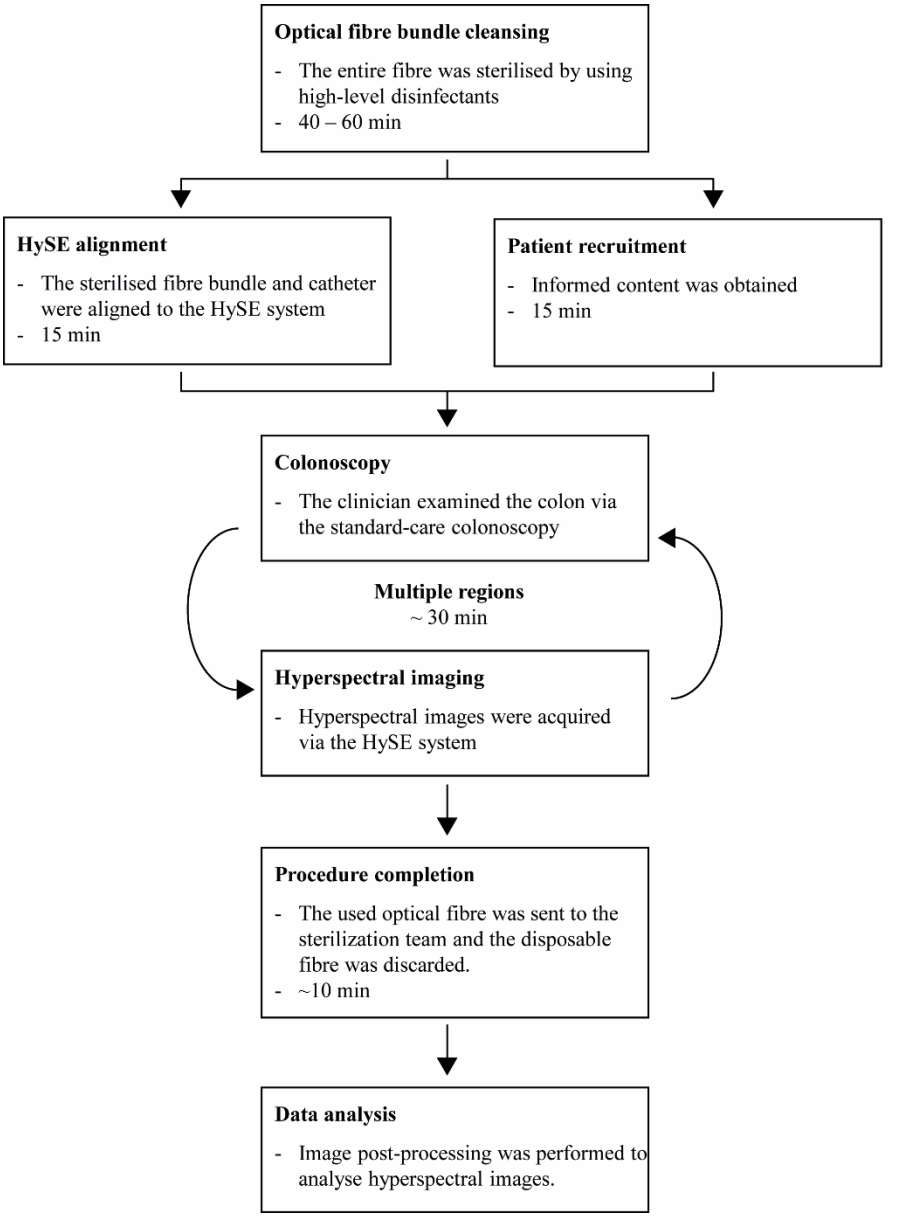
Supplementary Figures



Supplementary Fig. 1. Creation of a synthetic RGB SAM image. **a** A min-max normalized reflectance signal obtained from polyp. **b** Three reference spectral signals for SAM calculation. **c** Spectral angles calculated based on three references. **d** Bar graph showing 1-SAM results. **e** Representative images of green (normal mucosa), blue (polyp) and red (post-resection tissue) channels and synthetic RGB image to aid visualisation of the spatial data.

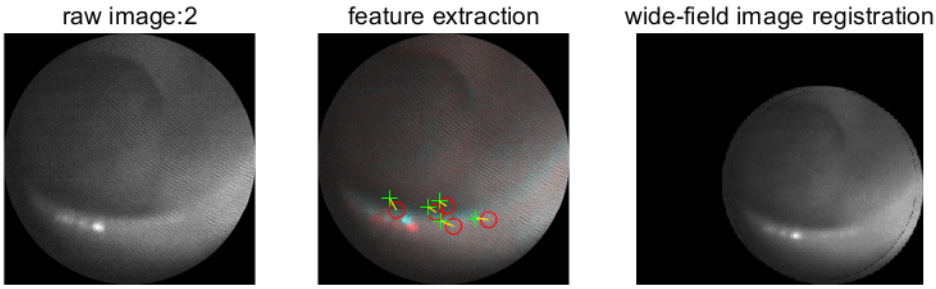


Supplementary Fig. 2. Wide-field white-light images of polyps. Blue dashed lines indicate polyps diagnosed by the clinicians.

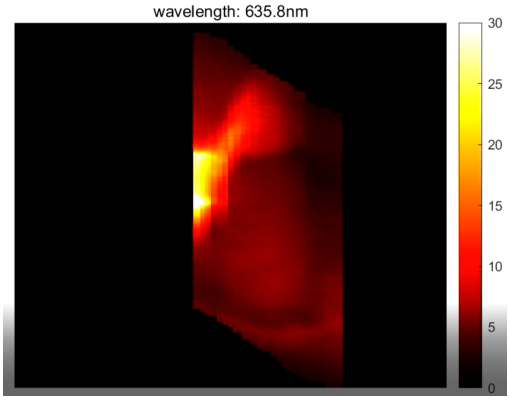


Supplementary Fig. 3. A flow chart of hyperspectral endoscopy procedure.

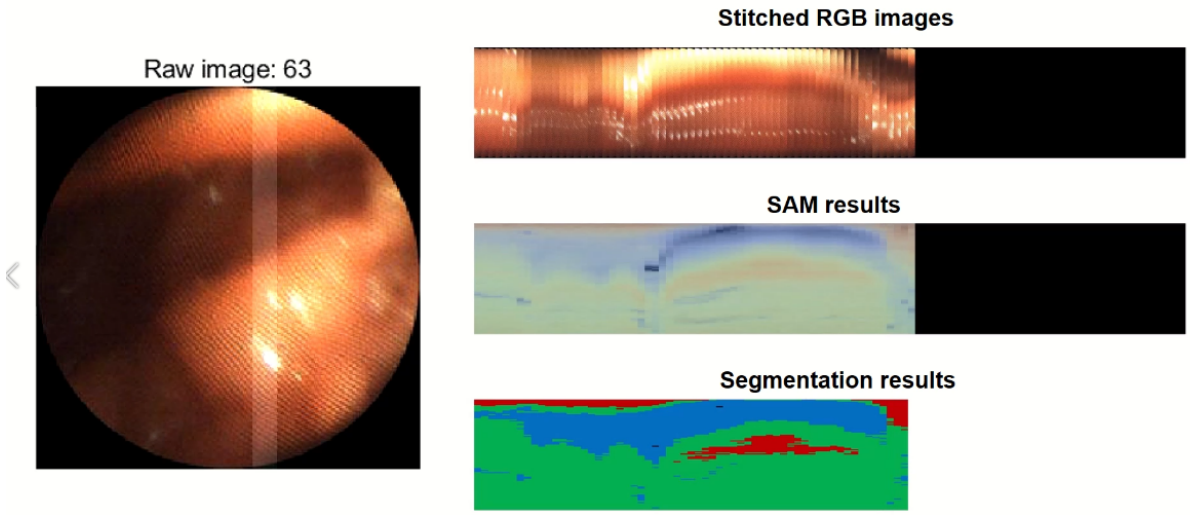
Supplementary Movies (single frame shown for reference)



Supplementary Movie 1. Wide-field image registration process by extracting features and calculating geometric transformation matrices.



Supplementary Movie 2. Slice images of the reconstructed wide-area hyperspectral image.



175

176 **Supplementary Movie 3.** A representative movie that shows wide-field image scanning and
177 quantitative analysis of corresponding spectral images based on SAM and kNN.

178

179

Supplementary Table

Supplementary Table 1. Summary of recruited patients

Patient number	Date	Optic Fibre	Notes
P1	26.Aug.2019	Optic Fibre A	System optimization
P2	27.Aug.2019	Optic Fibre A	System optimization
P3	15.Jan.2020	Optic Fibre A	Polyp
P4	15.Jan.2020	Optic Fibre B	Polyp
P5	17.Jan.2020	Optic Fibre A	Fail to data acquisition
P6	17.Jan.2020	Optic Fibre B	Normal
P7	21.Jan.2020		Didn’t perform endoscopy
P8	22.Jan.2020	Optic Fibre A	Polyp
P9	22.Jan.2020	Optic Fibre B	Normal
P10	22.Jan.2020		Didn’t perform endoscopy
P11	24.Jan.2020		Didn’t perform endoscopy
P12	27.Jan.2020	Optic Fibre A	Polyp
P13	27.Jan.2020	Optic Fibre B	Polyp