**Supplementary Material**

**Microbiota and Radiation-induced gastrointestinal Side-effects (MARS)**

**S.1 – Supplementary materials and methods**

**1.a – Inclusion and exclusion criteria**

All subjects were male, ≥18 years and able to give informed consent. Exclusion criteria were therapy with systemic antibiotics ≤3 months before enrolment (BE), or cytotoxic or immunosuppressive therapies ≤6 months BE, or consumption of ≥108 CFU/day of commercial probiotics ≤12 months BE, or any treatment for gastrointestinal cancer prior to enrolment.

**1.b – Prostate cancer treatment**

In the early and late cohorts, all patients received intensity-modulated radiotherapy to the prostate, seminal vesicles and pelvic lymph nodes with conventionally-fractionated (CFRT) or hypofractionated (HFRT) schedules according to a previously published protocol.1 Patients treated with CFRT received 70-74Gy in 35-37 fractions to the prostate and seminal vesicles and 50-60Gy to the pelvic lymph nodes. Patients treated with HFRT received 60Gy to the prostate and seminal vesicles and 47Gy to the pelvic lymph nodes.

In the colonoscopy cohort, cases received radiotherapy with either intensity-modulated or 3D-conformal techniques to the prostate and seminal vesicles, with or without pelvic lymph nodes irradiation. Radiotherapy fractions were given daily, 5 days/week. Control patients did not receive radiotherapy.

**1.c – Patient physiology and diet**

Patients completed a 7-day food diary starting 2 days after colonoscopy. A pragmatic approach of counting portions of cereals, vegetables, meat and fish, was used to characterise diet. Alcohol intake (g/week) was also recorded.

**1.d – Sample size**

The study we report was an exploratory study and, as such, all patients attending relevant clinics were invited to participate during a 2-year period during which the study recruitment ran. As this was an exploratory study and no previous published data was available in humans, we aimed to recruit as many patients as possible in one of the main oncology practices in the UK, with indicative rather than definitive sample sizes. Estimates of sample size were based on studies carried out in small samples of patients or in studies of the microbiota in inflammatory bowel disease, as no data for radiation enteropathy was available at the time of study design. We powered our study for 95% confidence level and 80% power, where differences in proportions of micro-organisms of the microbiota were expected to range between 24-76% (average 50%). It was therefore aimed that each cohort should recruit:

- Early cohort (cohort 2 in the study protocol, n=32): minimum of 26 patients;

- Late cohort (cohort 1 in the study protocol, n=87): minimum of 50 patients;

- Colonoscopy cohort (cohort 3 in the study cohort, n=15): minimum of 16 patients.

In the colonoscopy cohort, due to difficulties in recruiting healthy controls for the study, 9 "case" (irradiated) patients (out of a minimum of 8) and 6 "control" subjects were recruited and analysis was therefore carried out.

**1.e – Sampling procedures and processing**

**1.e.1 – Sampling of stool**

Stool from a single bowel movement was obtained within 12h prior to the visit by each subject by using the supplied materials kit and stored on dry ice up to 4h before further processing. The 12h maximum time was considered safe to avoid degradation of DNA, RNA and the composition of the microbial community as previously described.2 Samples were weighed, homogenized, and aliquoted in three sterile 1.5mL microcentrifuge plastic tubes (≥1.5mL of stool per aliquot). 100μL of phenol per 500 mg of stools were added to avoid DNA degradation during the freezing period, which was vortexed for 1 minute and immediately frozen at -80°C until DNA extraction.2

**1.e.2 – Sampling of intestinal mucosa (colonoscopy cohort only)**

Biopsies for 16S-rRNA gene-based phylotyping were stored in 5% v/v glycerol in 2mL sterile microcentrifuge tubes with screw caps, snap-frozen in liquid nitrogen in the endoscopy room and frozen at -80oC within 60 minutes until DNA extraction. Biopsies for cytokine analysis were stored in 2mL sterile microcentrifuge tubes with screw caps, snap-frozen in liquid nitrogen in the endoscopy room and frozen at -80oC within 60 minutes until analysis. Biopsies for pathology were fixed in 10% neutral buffered formalin in a 20mL container and sent to the pathology laboratory for analysis.

**1.e.3 – DNA extraction, data acquisition and processing**

Genomic DNA was extracted from faecal (250mg) and gut biopsy (whole biopsy) samples using the Qiagen Stool Kit (Qiagen, Crawley, UK) according to manufacturer instructions with an additional bead beating step for homogenisation of sample and lysis of bacterial cells (0.1g 0.1 mm sterile glass beads, FastPrep bead-beater (Q-BIOgene), setting six (6 metres per second) for 20 seconds, repeated a further two times with 5 minutes on ice between cycles), and quantified using the Invitrogen Qubit platform.

Library preparation and Illumina (MiSeq) sequencing of the V1-2 regions of the 16S-rRNA gene were performed at RTLGenomics (Lubbock, Texas, USA). Samples were amplified for sequencing in a two-step process. The forward primer was constructed with the (5’-3’) Illumina i5 sequencing primer (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG) and the 28F-GAGTTTGATCNTGGCTCAG primer. The reverse primer was constructed with the (5’-3’) Illumina i7 sequencing primer (GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG) and the 388R-TGCTGCCTCCCGTAGGAGT primer. Amplifications were performed in 25μL reactions with Qiagen HotStar Taq master mix (Qiagen Inc, Valencia, California), 1μL of each 5uM primer, and 1μL of template. Reactions were performed on ABI Veriti thermocyclers (Applied Biosytems, Carlsbad, California) under the following thermal profile: 95°C for 5 min, then 35 cycles of 94°C for 30 sec, 54°C for 40 sec, 72°C for 1 min, followed by one cycle of 72°C for 10 min and 4°C hold.

Products from first stage amplification were added to a second PCR to qualitatively determine concentrations. Primers for the second PCR were designed based on the Illumina Nextera PCR primers as follows:

- Forward - AATGATACGGCGACCACCGAGATCTACAC[i5index]TCGTCGGCAGCGTC and

- Reverse - CAAGCAGAAGACGGCATACGAGAT[i7index]GTCTCGTGGGCTCGG.

The second stage amplification was run the same as the first stage except for 10 cycles.

Amplification products were visualized with eGels (Life Technologies, Grand Island, New York). Products were then pooled equimolar and each pool was size selected in two rounds using Agencourt AMPure XP (BeckmanCoulter, Indianapolis, Indiana) in a 0.75 ratio for both rounds. Size selected pools were then quantified using the Quibit 2.0 fluorometer (Life Technologies) and loaded on an Illumina MiSeq (Illumina, Inc. San Diego, California) 2x300 flow cell at 10pM.

Filtered sequences were classified using the RDP classifier and relative proportions of phyla and families determined. Community analysis of the data was undertaken using MOTHUR.

**1.e.4 – Mucosal sample processing for cytokine detection**

A lysis buffer was produced, consisting of 20μL/mL of phosphatase inhibitor cocktail 2 and 20μL/mL of phosphatase inhibitor cocktail 3 (both purchased from Sigma-Aldrich) and 10μL/mL of protease inhibitor cocktail (Sigma-Aldrich). The lysis buffer and the unthawed samples were kept on ice. 250μL of buffer and the mucosa sample were loaded in 2mL tubes with screw caps and metal beads. These were loaded into a Precellys24® homogenizer and bead beating occurred in 2 cycles of 20s at 6000 rpm. This step was repeated if the sample was observed to not be fully lysed. Samples were put on ice and tubes were spun at 4°C at 18,894g for 10mins. The supernatant was transferred to a microcentrifuge tube and protein yield was quantified using a Direct Detect® infrared spectrometer. Samples were frozen overnight and after thawing, cytokine detection was carried out with the MSD® V-PLEX Human Cytokine 30-Plex Kit according to manufacturer instructions. Sensitivity (lower limit of detection) data is detailed in the product inserts of the kit, available at www.mesoscale.com.

**1.e.5 – Bioinformatic processing of 16S-rRNA gene data**

Sequences generated from Illumina (MiSeq) sequencing were analysed with MOTHUR for identification of operational taxonomic units (OTU), taxonomic assignment, community comparison, and data cleaning by adapting its standard operational procedure.3 Experimental sequences were firstly processed using the make.contigs and screen.seqs scripts with minimum length set at the 2.5% tile of the data (345bp), maximum length at the 97.5% tile (365bp), and maximum ambiguity at 0. This filtered data to minimize effects of poor sequence quality and sequencing errors by removing sequences with more than one ambiguous base call, and retaining only sequences that were between 345 and 365bp. Only sequences with the forward primer motif were included to ensure that the V1-V2 region was available for taxonomic assignment. Unique sequences were obtained with the unique.seqs script and aligned to Silva reference files.4 The screen.seqs script was again used to remove poorly aligned sequences starting at or before the 25% tile (position 1044) and ending at or after the 75% tile (position 23444). Any non-bacterial taxon was removed with the remove.lineage command.

**S.2 – Supplementary text**

**2.a – Symptoms and diet in the early cohort**

2.a.1 – Clinician-reported outcomes

Overall, clinician-reported gastrointestinal toxicity followed similar patterns to the study reported in chapter 2 (supplementary figure 2.a-1). At baseline, 15.6% and 18.7% of patients with CRO-late and composite score symptoms of grade 1 or over (G1+) respectively, with diarrhoea and tenesmus mainly reported. It is acknowledged that this scale was designed to assess late symptoms after radiotherapy, however, the decision of measuring these prior to radiotherapy allows for assessment of “toxicity-analogous” symptom rates before the start of radiotherapy.

Acute gastrointestinal symptoms peaked at 4/5 weeks and abated at 12 weeks, with G1+/G2+ rates of toxicity of 84.4%/34.4% and 71.88/12.5% at these timepoints respectively. Late gastrointestinal symptoms decreased with time, with G1+/G2+ rates of 53.1%/6.3% and 34.4%/3.1% maximum CRO, and 59.4%/9.4% and 37.5%/3.1% maximum composite toxicity at the 6 and 12 month timepoints respectively. Diarrhoea, proctitis and tenesmus were the main reported problems in the late enteropathy setting. Only one instance of grade 3 toxicity was observed, with one patient experiencing bleeding at 4/5 weeks which required pads and was managed by urgent endoscopic treatment.

2.a.2 – Patient-reported outcomes

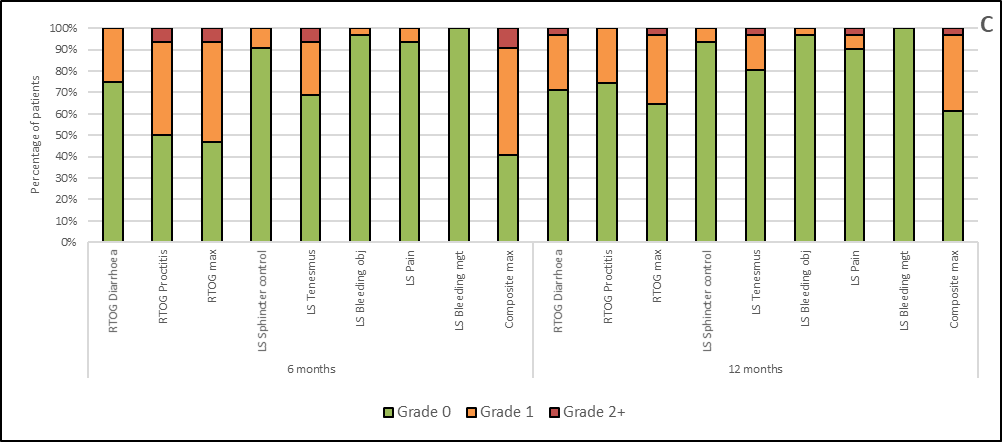
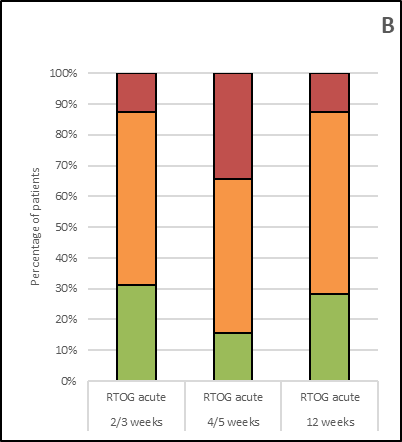
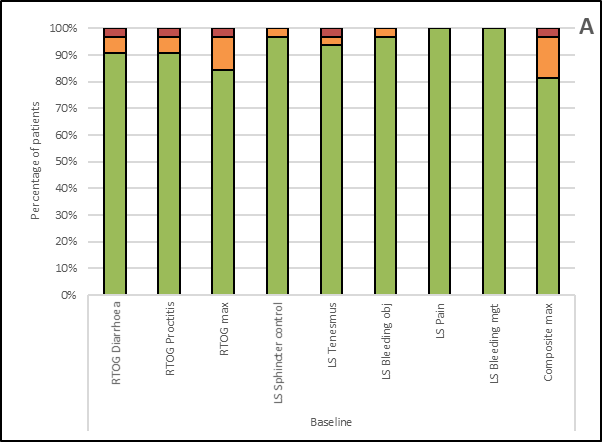
Patient-reported outcomes followed a similar pattern to CRO (supplementary figure 2.a.2). At baseline, PRO median (IQR) score was 68 (65-70), decreasing to 62 (57-66) and 58.5 (48.8-66) at 2/3 and 4/5 weeks respectively. PRO scores recovered in the immediate aftermath of treatment to near-baseline levels. It is noted that at 12 months there was a high rate of patients not returning their PRO questionnaires (supplementary table 2.a.2).

2.a.3 – Diet

Comparison between dietary habits was made by grouping patients according to PRO and CRO-based stratification. There were no vegetarians in the cohort. When analysing dietary differences in groups defined with PRO (supplementary figure 2.a.3-1), significant differences between groups were detected for meat (p=0.05), fish (p=0.02), vegetables (p=0.05) and cereals (p=0.03), while a trend was detected for differences in alcohol consumption (p=0.1). On post-hoc analysis, we detected that, when compared to the group with no symptoms, the group with symptoms at both timepoints had higher consumption of fish (mean (SD) portions per week: 3 (2) vs 1 (1); p=0.02), and a small trend was detected for differences in alcohol intake (g/week: 228 (277) vs 0 (0); p=0.11), which is largely explained by a single episode of very high intake by one patient in the symptomatic group prior to treatment initiation, as illustrated by the ample standard deviation. Also, when compared to the group with no symptoms, the group with symptoms at one timepoint had a higher consumption of meat (10 (2) vs 7 (3); p=0.03), fish (4 (2) vs 1 (1); p=0.007) and vegetables (27 (8) vs 18 (6); p=0.04), with trends detected for cereal (19 (5) vs 25 (5); p=0.1) and alcohol (g/week: 57 (66) vs 0 (0); p=0.09). Significant differences when comparing symptomatic groups were found for cereals only (29 (8) vs 19 (5); p=0.02).

When analysing groups defined by CRO-based toxicity scores (supplementary figure 2.a.3-2), significant differences were found for meat consumption (p=0.006). On post-hoc analysis, significant differences in meat consumption were detected between the groups with no symptoms and symptoms at one timepoint (9 (2) vs 5 (2); p=0.02) only.

**Supplementary figure 2.a.1:** Clinician-reported toxicity endpoints in cohort 2 of the MARS study*.*



A: Symptoms at baseline. B: Acute symptoms. C: Late symptoms.

*LS=LENT-SOM. Obj.=objective. Mgt=management. Max=maximum. Composite max=maximum of composite score*

*including all CRO (RTOG and LENT-SOM) endpoints. Baseline scores were measured with late toxicity scales (see text).*

*The acute setting is contemplated only in the RTOG scale. Higher scores represent more symptoms.*

**Supplementary figure 2.a.2:** Patient-reported toxicity endpoints in cohort 2 of the MARS study*. Mean PRO score and standard deviation (error bar). Higher scores reflect better function (less symptoms).*

**Supplementary table 2.a.2:** Rates of patients not sending back self-reported toxicity questionnaires*.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Timepoint | Baseline | 2/3 weeks | 4/5 weeks | 12 weeks | 6 months | 12 months |
| Patients missing PRO data – n (%) | 5 (16%) | 4 (13%) | 4 (13%) | 12 (38%) | 10 (31%) | 14 (44%) |

**Supplementary figure 2.a.3-1:** Differences in (A) mean portions per week of meat (p=0.05), fish (p=0.02), vegetables (p=0.05) and cereals (p=0.03); and (B) mean grams of alcohol per week (p=0.1) between **PRO-stratified** symptom groups in the early cohort. *Significance testing reported here was made between all groups using the Kruskal-Wallis test.*

*The group with no symptoms had a mean alcohol intake of 0g/week with SD=0.*

**Supplementary figure 2.a.3-2:** Differences in (A) mean portions per week of meat (p=0.006), fish (p=0.73), vegetables (p=0.13) and cereals (p=0.28); and (B) mean grams of alcohol per week (p=0.28) between **CRO-stratified** symptom groups in the early cohort. *Significance testing reported here was made between all groups using the Kruskal-Wallis test.*

2.a.4 – Correlation between CRO and PRO scores

To study the correlation between CRO and PRO scores, all observations at all timepoints were included in a single dataframe. CRO and PRO were weakly correlated, with Spearman’s rank rho coefficient of -0.47 (p<0.0001). The median PRO score (IQR) was 67 (64.5-70), 62.5 (57.3-66), and 58 (52.7-62.3) for CRO grades 0, 1 and 2 respectively, reflecting a clinically significant increase in patient-reported symptoms per clinician-reported toxicity grade.

**2.b – Symptoms and diet in the late cohort**

2.b.1 – Clinician-reported outcomes

Toxicity prevalence data at the time of sampling was obtained as for the early cohort (termed “actual”). However, cumulative peak toxicity (termed “historic”), obtained as per the IMRT for Prostate Cancer trial protocol, was also used. This data was obtained from the long-term late toxicity database of the IMRT-PCa trial, which included outcomes from 6 months after radiotherapy onwards, reflecting the late enteropathy timeline.

As expected, actual toxicity followed similar prevalence patterns as late toxicity timepoints in the IMRT-PCa trial (supplementary figure 2.b.1-1).1 Twenty patients (23%) had CRO toxicity grade 2 or above (grade 2+), mostly explained by proctopathy (19%), although 9.2% reported grade 2+ diarrhoea. Grade 2+ LENT-SOM tenesmus and bleeding were reported by 14% and 3.4%, making tenesmus the most frequent symptom of proctopathy at the time of sampling. Composite maximum prevalence scores were similar to CRO maximum, with grade 2+ toxicity in 25% of patients.

Historic toxicity rates were obviously higher, with 15%/20%, 39%/17% and 32%/29% of patients experiencing grade 1/2+ diarrhoea, proctopathy and maximum toxicity respectively (supplementary figure 2.b.1-2). It is noteworthy that, unlike for actual toxicity endpoints, CRO rectal ulcer, rectal-anal stricture and bowel obstruction were included for computing the maximum figure, with three distinct patients experiencing bowel obstruction at one timepoint during follow-up graded 3, 2 and 1 respectively, but no other such toxicities observed. Correlation between peak historical and MARS-prevalence toxicity existed but was weak, with a Spearman rho coefficient of 0.56 (p<0.0001).

2.b.2 – Patient-reported outcomes

Actual patient-reported symptoms were generally low. Median PRO score was 67 (IQR=62-70) out of a maximum of 70, which represents best intestinal function (figure 2.b.2-1).

Historic patient-reported outcomes were available scored with the UCLA-PCI scale, as per the IMRT for Prostate Cancer study protocol.1 Specifically, two items of the scale were used for comparison: bowel distress and bowel problem, representing overall bowel function.5 Using peak scores obviously provided a different picture. Moderate to severe bowel distress had been reported by 29 patients (34%, supplementary figure 2.b.2-2A), while moderate to big bowel problem had been reported by 28 patients (33%, supplementary figure 2.b.2-2B) since radiotherapy. As with clinician-reported outcomes, correlation between peak historical and actual patient-reported symptoms existed but was very weak, with a Spearman’s rho coefficient of 0.37 (p=0.001).

2.b.3 – Diet

Comparison between dietary habits was made by grouping patients according to CRO maximum toxicity at the time of sampling only, as significant differences in microbiota endpoints were detected using this scale. There were no vegetarians in the cohort. There were no significant or biologically relevant differences between patients grouped by CRO in the cohort for any of the analysed diet parameters, including meat, fish, vegetable, cereal and alcohol intake per week (supplementary figure 2.b.3).

**Supplementary figure 2.b.1-1:** Clinician-reported actual toxicity in the late cohort*.*

*LS=LENT-SOM. Obj.=objective. Mgt=management. Max=maximum. Composite max=maximum*

*of composite score including all CRO endpoints. Higher scores represent more symptoms.*

**Supplementary figure 2.b.1-2:** Clinician-reported historical toxicity (since 6 months post-radiotherapy)

In the late cohort*. Higher scores represent more symptoms.*

**Supplementary figure 2.b.2-1:** Patient-reported actual toxicity in the late cohort*.*

*Mean PRO score and standard deviation (error bar). Higher scores reflect better function (less symptoms).*

**Supplementary figure 2.b.2-2:** Patient-reported historical bowel distress (A) and bowel problem (B) in the early cohort*. Higher scores represent more symptoms.*

Supplementary figure 2.b.3: Differences in (A) mean portions per week of meat (p=0.28), fish (p=0.48), vegetables (p=0.42) and cereals (p=0.35); and (B) mean grams of alcohol per week (p=0.83) according to CRO maximum.

**2.c – Symptoms and diet in the colonoscopy cohort**

2.c.1 – Clinician-reported outcomes

Case subjects had significantly higher proctitis (p=0.02) and maximum toxicity (p=0.006), as well as trends for higher tenesmus (p=0.051), objective bleeding (p=0.051) and maximum toxicity (p=0.06) when compared to controls. There was also a trend for higher diarrhoea (p=0.08). Results are summarized in supplementary figure 2.c.1. Supplementary table 2.c.1 summarizes symptoms of case patients as reported by the specialist gastroenterology clinical team.

2.c.2 – Patient-reported outcomes

Case subjects had a lower PRO scores than controls, with median scores of 61.5 (IQR: 54.5-64.5) and 63.5 (62.25-65.5) for cases and controls respectively (supplementary figure 2.c.2). This difference did not reach a difference of ≥6 defined as clinically meaningful difference and was not statistically significant (p=0.3).6

2.c.3 – Diet

Except for one control subject, there were no other vegetarians. There were no significant differences in diet between the control and case populations (supplementary figure 2.c.3). However, alcohol consumption appeared higher in case subjects, although this was not significant (p=0.62).

Supplementary figure 2.c.1: Mean clinician-reported bowel symptoms. *LS = LENT-SOM*.

**Supplementary table 2.c.1:** Symptoms and diagnoses made by the gastroenterology team for case patients.

|  |  |  |
| --- | --- | --- |
| **Trial ID** | **Symptoms reported by clinical team prior to colonoscopy** | **Diagnoses made by clinical team for each patient** |
| 3001 | Flatulence, faecal incontinence, steatorrhoea. | Mild bile acid malabsorption, and radiation proctopathy (telangectasia) grade 1. |
| 3002 | Erratic bowel function with intermittent rectal mucus discharge, faecal incontinence, tenesmus, urgency, borborygmi. | Radiation proctopathy (telangectasia) grade 1, no other diagnoses. |
| 3003 | Stool frequency 6 times/day, urgency, tenesmus, faecal incontinence, excessive flatulence, borborygmi. | Small intestinal bacterial overgrowth, excessive fibre intake, and radiation proctopathy (telangectasia) grade 1. |
| 3004 | Urgency, faecal incontinence, rectal bleeding, nocturnal defecation, bowel frequency (3-6 times/day). | Pancreatic insufficiency, excessive fibre, and radiation proctopathy (telangectasia) grade 2. |
| 3006 | Peri-anal pain and pruritus, tenesmus, mucus discharge, rectal bleeding. | Bile acid malabsorption + radiation proctopathy (telangectasia) grade 2, and peri-anal radiation-induced ulcers. |
| 3007 | Abdominal cramps, upper and lower abdominal pain, perianal discomfort, abdominal bloating, excessive flatulence, borborygmi, tenesmus, intermittent steatorrhoea, perianal pruritus, nocturnal defecation. | Small intestinal bacterial overgrowth, and radiation proctopathy (telangectasia) grade 1. |
| 3008 | Urgency, mucus discharge, bowel frequency (4-6 times/day). | Excessive fibre intake, diverticular disease, and radiation proctopathy (telangectasia) grade 1. |
| 3010 | Urgency, tenesmus, abdominal cramps, occasional mucus discharge, occasional flatulence. | Radiation proctopathy (telangectasia) grade 2, no other diagnoses. |
| 3015 | Abdominal cramps, abdominal pain, perianal pain, excessive flatulence, borborygmi, tenesmus, bowel frequency 4 times/day | Excessive fibre intake and radiation proctopathy (telangectasia) grade 1. |

Supplementary figure 2.c.2: Mean patient-reported bowel symptoms.

*Mean PRO score and standard deviation (error bar). Higher scores represent best intestinal function.*

Supplementary figure 2.c.3: : Mean diet scores (A) and alcohol consumption (B). wk = week. No significant differences were observed for meat (p=0.27), fish (p=0.71), vegetable (p=0.94), cereal (p=0.20), or alcohol (p=0.62).

**2.d – Comparisons of microbiota features in the colonoscopy cohort**

2.d.1 – α-diversity

The Chao microbial α-diversity index was compared between intestinal mucosa and stool samples obtained from control and case subjects and is summarized in table 2.a.1 and figure 2.a.1. Microbial communities in stool samples appeared less diverse than communities in rectal biopsy samples, with median (IQR) Chao diversity of 54.1 (51.3-65.2) and 69.8 (67.6-73.8) in stools compared to 77 (59.3-117.1) and 76.4 (66-102.1) in rectal mucosa of cases and controls, respectively. This difference was not apparent when comparing the diversity of distal sigmoid microbial communities (median: 52.3, IQR: 44.6-75.2) with stool diversity. However, when testing the significance of these observations, only the difference between stools and biopsies from the anterior rectal wall samples in cases showed a trend for significance (p=0.07).

2.d.2 – Comparison of β-diversity between stool and mucosal samples

β-diversity distance matrices were generated using UniFrac distances between mucosal and stool samples in cases and controls. UniFrac is a distance metric based upon the unique fraction of branch length in a phylogenetic tree built from two sets of taxa.404 Presence of significant effects of sample type on β-diversity metrics were tested using permutational multivariate analysis of variance (PERMANOVA) through the function adonis() in the “vegan” R package.405 PERMANOVA tests group-level differences through an ordination method and provides an assessment of grouping factors in an iterative, pairwise subject-to-subject procedure.404 Non-metric multidimensional scaling (NMDS) plots were also generated using the vegan and MASS R packages.406 NMDS allows for collapsing information from multiple dimensions (in this case, bacterial communities) into a two-dimensional format, allowing for visual representations of distance between communities by sample types

There was no evidence of a significant impact of sample type (sigmoid or rectum mucosa and stool samples) on UniFrac distance matrices in either control or case samples (table 2.d.2). This indicates that stool and mucosa-associated bacterial communities overlapped significantly, as summarized in NMDS plots (figure 2.d.2). Of note, a trend was found for distance between case sigmoid and case stool samples, suggesting that faecal samples are more representative of rectal than sigmoid colon communities. However, R2 and pseudo-F values were low (R2=0.14, pseudo-F=2.55), and the NMDS plot does not suggest that this difference is relevant in terms of overall microbial communities.

Table 2.d.2: PERMANOVA results based on UniFrac distance matrices by sample type.

|  |  |  |  |
| --- | --- | --- | --- |
| Pairs | Pseudo-F | R2 | p-value |
| Rectum-Case vs Stool-Case | 0.23 | 0.01 | 0.88 |
| Sigmoid-Case vs Stool-Case | 2.55 | 0.14 | 0.09 |
| Rectum-Control vs Stool-Control | 0.31 | 0.03 | 0.80 |

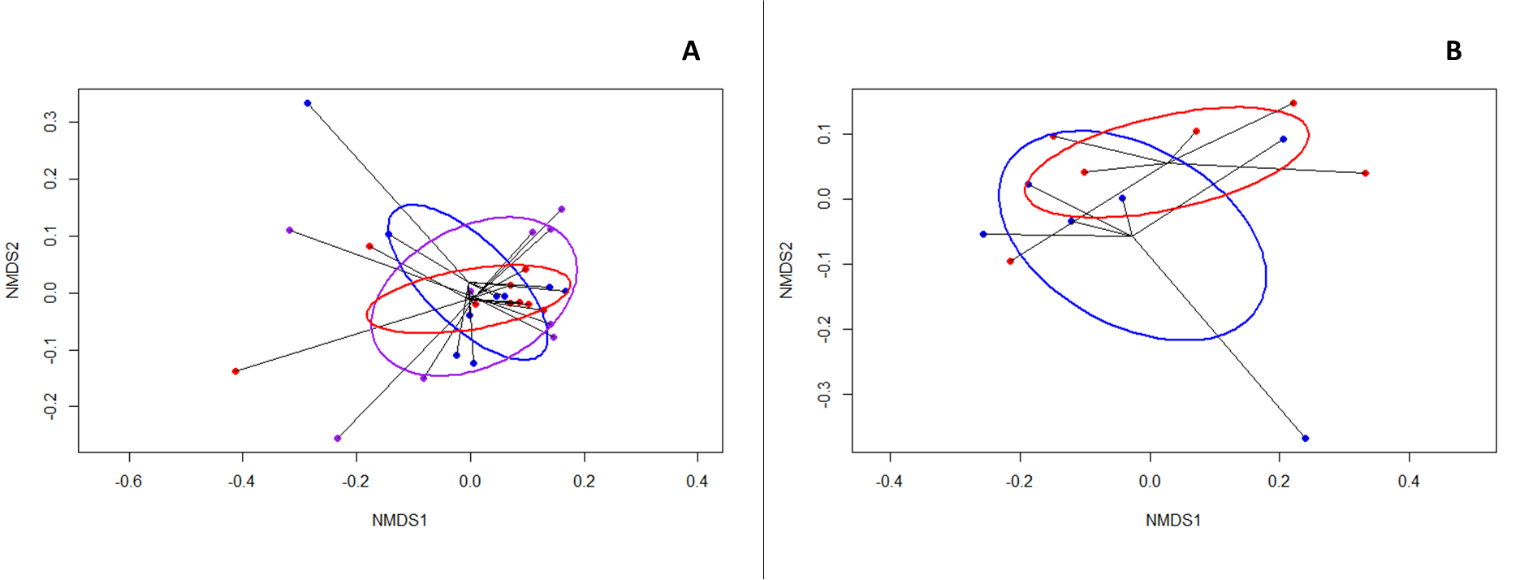


Figure 2.d.2: NMDS plots of UniFrac distances between samples. A: Cases (red=stool samples; blue=rectal biopsy samples; purple=distal sigmoid samples). B: Controls (red=stool samples; blue=rectal biopsy samples).

2.d.3 – Stool samples

Stool samples of cases and controls were highly correlated (R2=0.88 and R2=0.86 for phyla and genera respectively). No observed differences at the phylum or genus level were statistically significant. Bacteroidetes appeared higher in case subjects (difference in mean proportions (DiM) = 17.2%, p>0.05) and Firmicutes were conversely lower (DiM = 10.7%, p>0.05). When analysing genera, the proportion of *Faecalibacterium* appeared higher in controls (DiM=9.61%, p>0.05) and *Bacteroides* higher in cases (DiM=13.79%, p>0.05).

2.d.4 – Mucosal samples – rectal biopsy (controls) vs rectal biopsy (cases)

Rectal biopsy samples were highly correlated (phylum: R2=1; genus: R2=0.96). No significant differences were observed between cases and controls at the phylum level (figure 13). At the genus level, *Faecalibacterium* (DiM=6.21%, p>0.05) appeared higher in control samples although differences were non-statistically significant.

2.d.5 – Mucosal samples – rectal biopsy (controls) vs rectal biopsy (cases)

Rectal biopsy control samples were highly correlated with distal sigmoid mucosa samples in cases (phylum: R2=0.992; genus: R2=0.952). No differences were observed between cases and controls at the phylum level. At the genus level, *Faecalibacterium* (DiM=7.14%, p>0.05) appeared higher in control samples although non-significantly.

2.d.6 – Mucosal samples – rectal biopsy (cases) vs distal sigmoid (cases)

No evident differences were found when comparing microbiota features of the rectal biopsy and the distal sigmoid in cases. Samples were highly correlated at the phylum (R2=0.993) and genus (R2=0.982) levels.

**2.e – Comparison of the microbiota according to androgen deprivation therapy (ADT) and testosterone recovery status in the late cohort**

In the late cohort, 10 (11%) patients were under active androgen deprivation therapy (ADT) and 40 (46%) had unrecovered testosterone levels. To test if these aspects had any impact on microbiota profiles, patients were divided in groups according to each characteristic and differences in the microbiota were tested. Table 2.e.1 summarizes overlapping features between these two groups of patients.

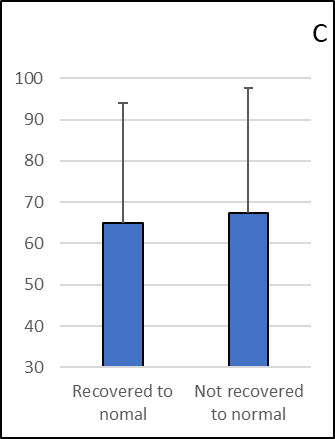
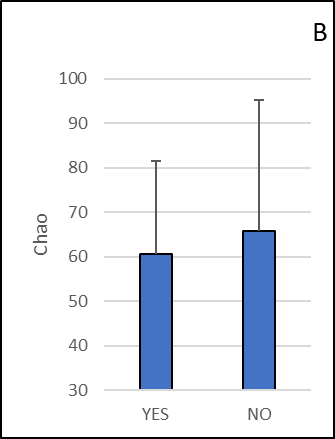
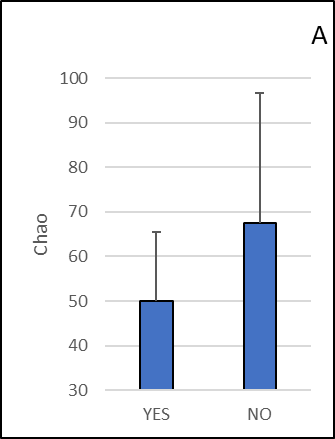
**Supplementary table 2.e.1:** Number of patients on ADT and/or with testosterone levels not recovered to normal.

|  |  |  |
| --- | --- | --- |
|  | On hormone therapy (ADT) | Testosterone not recovered |
| On hormone therapy (ADT) | *10\** | 10 |
| Testosterone not recovered | 10 | *40\** |

*\*Numbers in italic/light grey cells indicate the total number of patients per group. Results are n in all cases.*

2.e.1 – Diversity

Differences in the diversity of patients stratified by hormone therapy status and testosterone recovery status were not significantly different. Patients receiving ADT had median Chao of 51.6 (44.5-77.0) compared to 54.6 (47.3-79.3) for patients not under this treatment (p=0.63, figure 2.e.1-A). Patients with recovered levels of testosterone had a median Chao 55.7 (45.1-75.8) compared to 54.3 (48.7-87.5) in patients with testosterone levels <6nmol/L (p=0.69, figure 2.e.1-B).



**Supplementary figure 2.e.1:** Chao richness by recurrent prostate cancer (A, p=0.03), current hormone therapy (B, p=0.63), testosterone recovery (C, p=0.69). The table below the figure indicates the number of patients assessed by group.

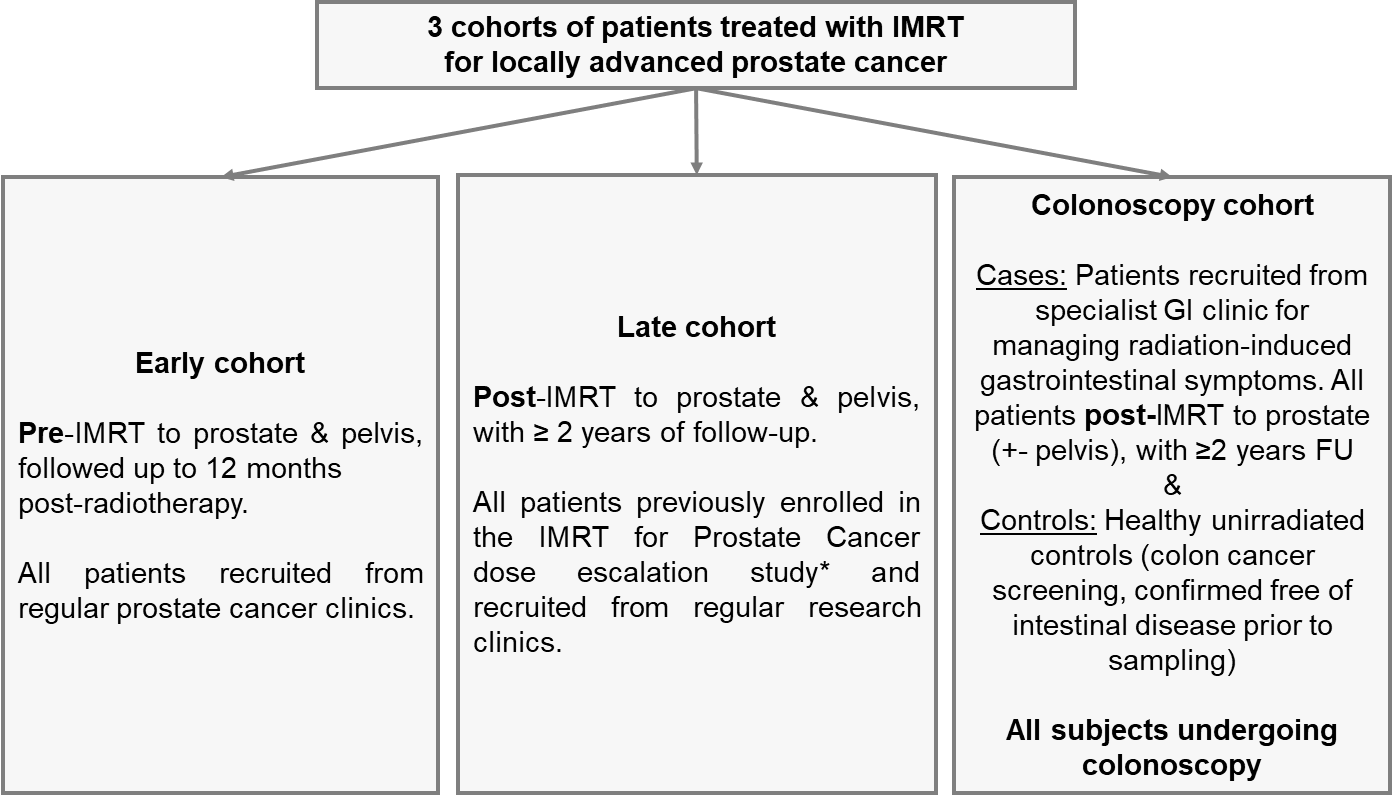
2.e.2 – Microbial features

No differences were detected at the phylum level between patients with and without recurrent tumours. At the genus level, *Odoribacter* was higher in patients with recurrent tumours, with a trend for significance before correction (p\*=0.07, p>0.1). No other features appeared different between groups.

No differences were detected at the phylum level according to ADT status. At the genus level, *Oscillibacter* was lower in patients receiving ADT (p\*=0.03, p>0.1). No other differential features were apparent.

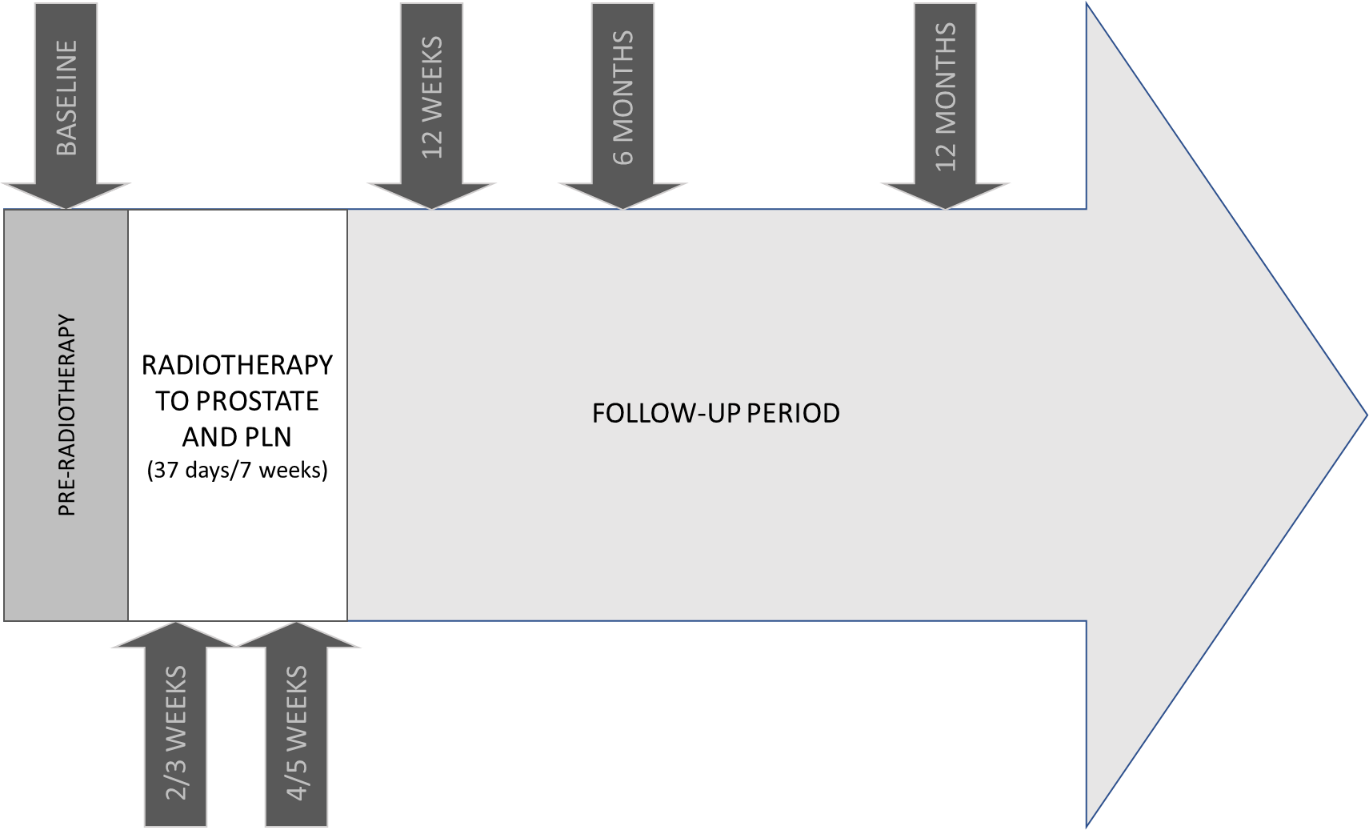
At the phylum level, Firmicutes were higher in patients with normal testosterone levels (p\*=0.04, p>0.1). At the genus level, *Faecalibacterium* was lower in patients with testosterone values lower than the normal range (p\*=0.03, p>0.1, figure 17D). No other differential features were apparent between groups.

**S.3 – Supplementary figures**



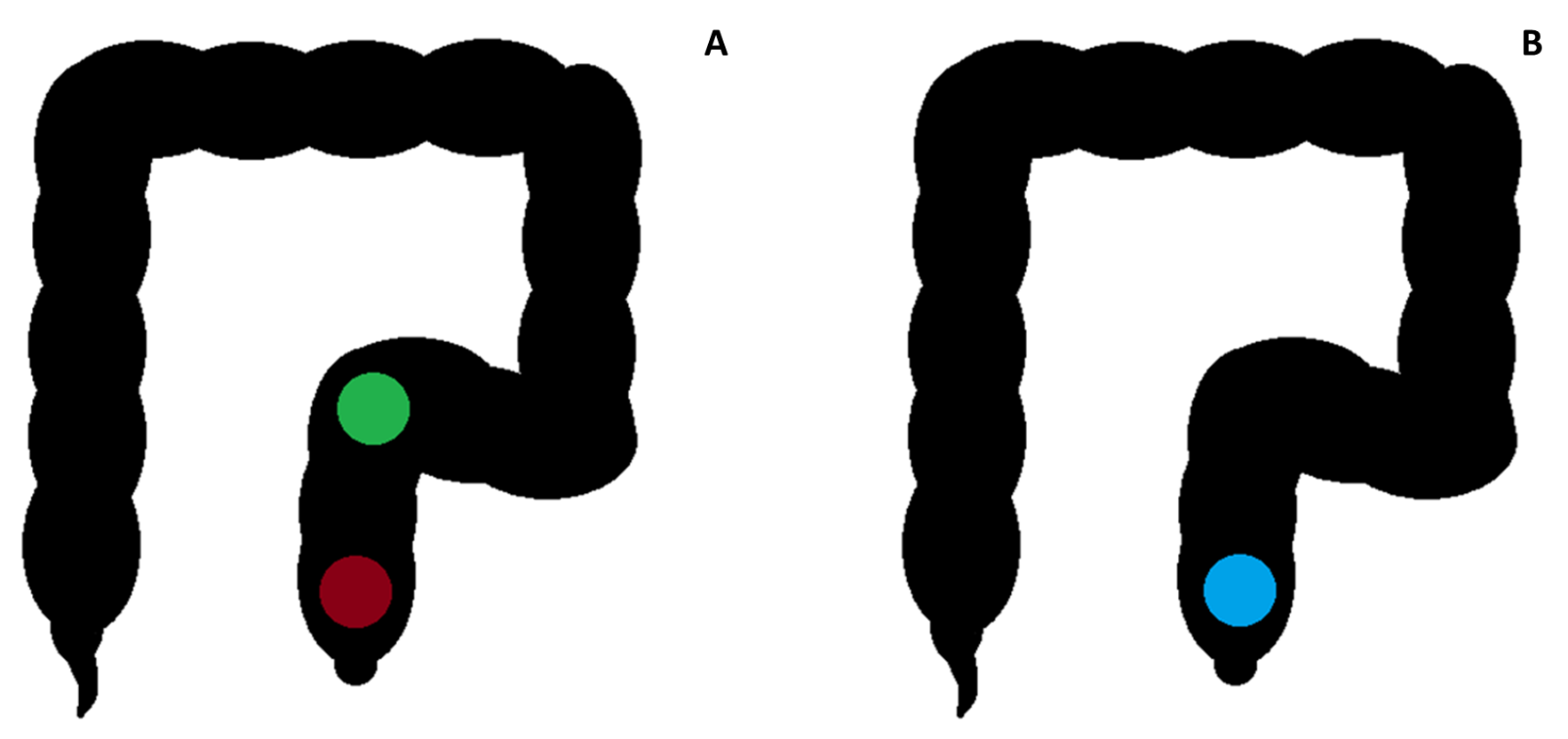
**Figure SUPP-1:** Design of the MARS study.

*IMRT: Intensity-Modulated Radiotherapy. FU = follow-up. \*: see reference 7 in main text.*



**Figure SUPP-2:** Follow-up in the late cohort.

*PLN: Pelvic-lymph nodes.*

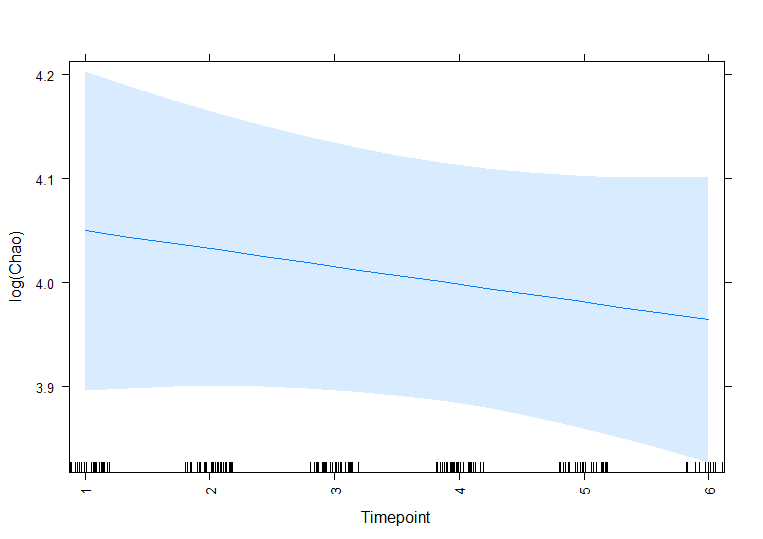


**Figure SUPP-3:** Location of sampling for patients with enteropathy (A) and control subjects (B).

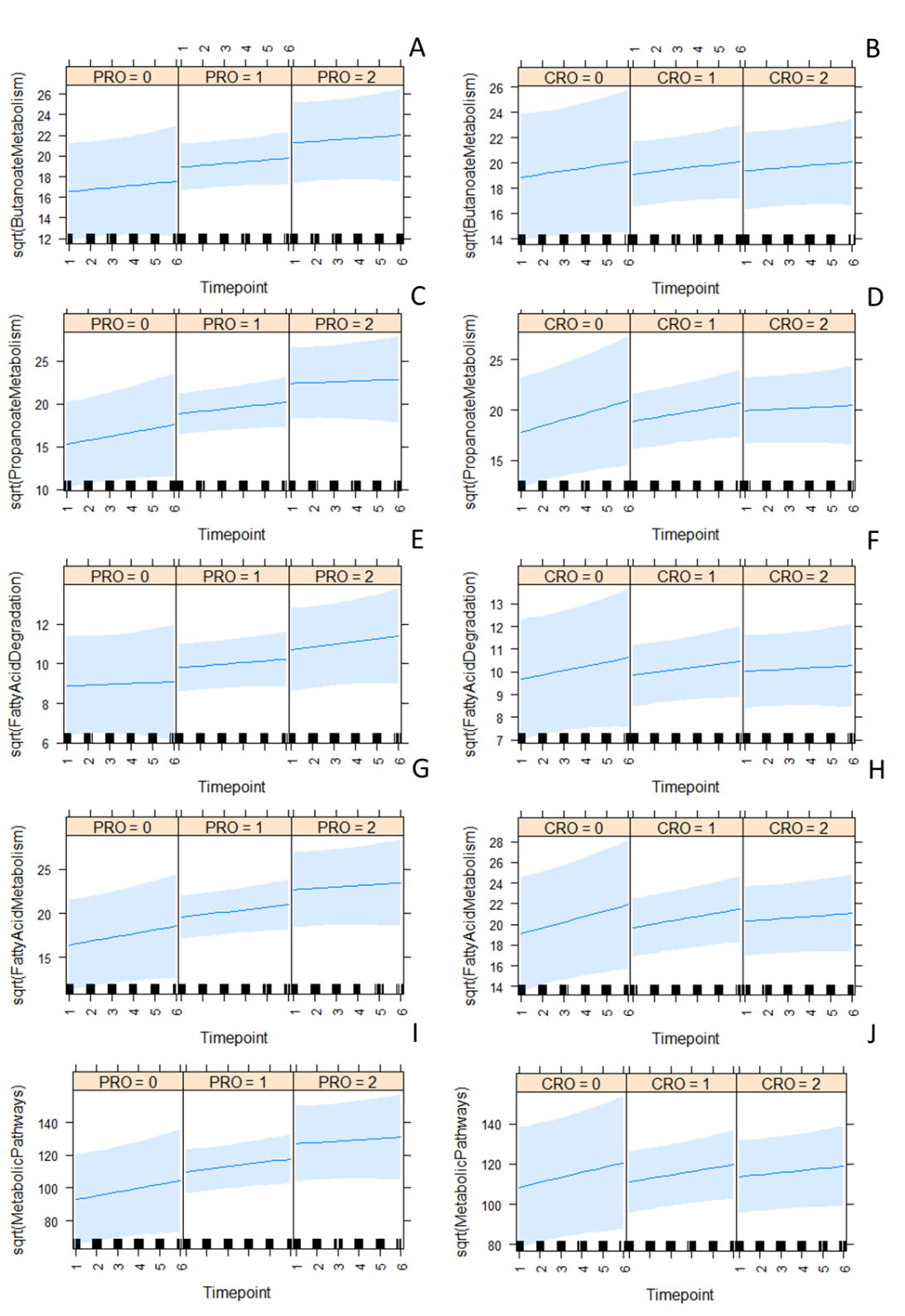
*In patients with radiation enteropathy, 3 samples were taken from an area with (red) and another 3 samples were taken from an area without (green) enteropathy. In control subjects, only samples from the anterior rectum (blue) were taken.*

**Figure SUPP-4:** Mean Chao richness by patient-reported toxicity group (p=0.97).

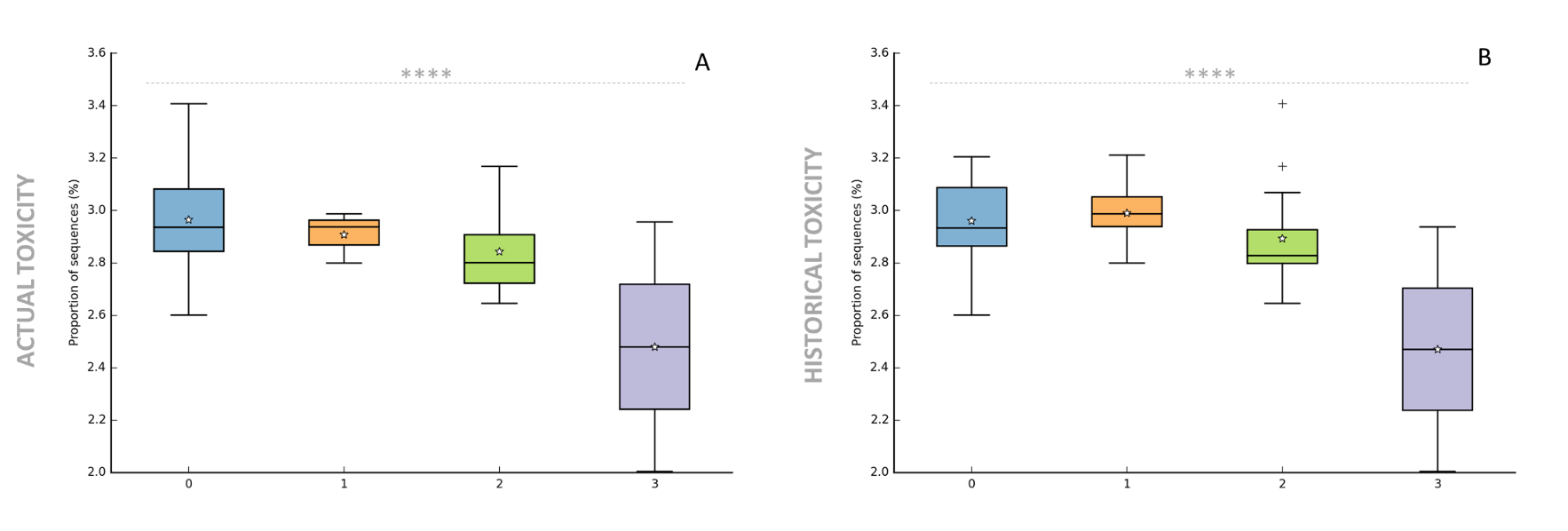
*The table below the figure indicates the number of patients assessed per symptom group.*



**Figure SUPP-5:** Dynamics of Dynamics of Chao diversity over time in the acute cohort not stratified by GI symptoms.

****

**Figure SUPP-6:** Dynamics of abundance of KEGG microbial pathways of butyrate\* metabolism (A/B), propionate\* metabolism(C/D), fatty acid degradation (E/F) fatty acid metabolism (G/H), and metabolic pathways (I/J) over time in PRO (left) and CRO (right) stratified groups. The effect of PRO symptom group was significant for biosynthesis of unsaturated fatty acids (PRO: p=0.04; CRO: p=0.08). *Groups: 0 = no symptoms, 1 = non-persistent symptoms, and 2 = persistent symptoms. Timepoints: 1=baseline, 2=2/3 weeks, 3=4/5 weeks, 4=12 weeks, 5=6 months, and 6=12 months after radiotherapy initiation. A square root transformation was used due to a positive skew of the data, which was confirmed to provide superior goodness of fit when compared to a log transformation in all models. \*: Butyrate/butanoate and propionate/propanoate are interchangeable expressions. For graphical purposes we chose to use the names of pathways as defined in the KEGG database.*

****

**Figure SUPP-7:** Fatty acid metabolism abundance with actual (A) and historical (B) by CRO diarrhoea grade in the late cohort. Higher grades reflect more serious symptoms. \*\*\*\*: p≤0.001. The x axis shows grade of diarrhoea.

**S.4 – Supplementary tables**

Table SUPP-1: Bowel subset of gastrointestinal symptom score validated for radiation enteropathy used for patient-reported bowel symptom assessment (10 items).7

|  |  |
| --- | --- |
| Item | Problem addressed |
| Question 1 | Bowel frequency |
| Question 5 | Loose stools |
| Question 9 | Anal pain |
| Question 13 | Abdominal or anal pain |
| Question 17 | Flatulence |
| Question 20 | Abdominal bloating |
| Question 22 | Rectal bleeding |
| Question 24 | Bowel urgency/tenesmus |
| Question 26 | Accidental soiling |
| Question 29 | Bowel problem |

Please refer to supplementary reference 4 for the full questionnaire.

Table SUPP-2: Comorbidity and physiology data recorded.

|  |  |  |  |
| --- | --- | --- | --- |
| Constitutional factors | Habits | Comorbidities | Medications |
| Age  Body mass index | Smoking habits  Drinking habits  Nutrition habits (7-day food diary) | Previous abdominal surgery  Arterial hypertension  Collagen vascular disease  Diabetes mellitus  Dyslipidemia  Any other metabolic diseases  Irritable bowel syndrome  Any systemic immune-allergic disease  Any other chronic disease of the gastrointestinal tract  Any other acute or chronic disease. | Laxatives  Corticosteroids used within 6 months of enrolment |

Table SUPP-3: Histopathology score to score intestinal mucosa samples (colonoscopy cohort only).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item | Score = 0 | Score = 1 | Score = 2 | Score = 3 |
| 1 - Thickening of Serosa | No changes | Slight thickening of serosa; hyperplasia of peritoneal mesothelium | Marked thickening of serosa | Extreme thickening and fibrosis of serosa |
| 2 - Mucosal Ulcerations | No changes | Small superficial ulcerations | Ulcerations involving more than half of the intestinal circumference | N/A |
| 3 - Epithelial atypia | No changes | Abnormally oriented crypts | Irregular crypt regeneration with atypical epithelial cells | Adenocarcinoma |
| 4 - Vascular Sclerosis | No changes | Slight thickening and hyalinization of vessel wall | Vessel wall double normal thickness: hyalinization and stenosis | Extreme sclerosis with marked stenosis or complete occlusion; fibrinoid necrosis |
| 5 - Intestinal wall fibrosis | No changes | Submucosa double normal thickness; broadened and hyalinized collagen fibres | Submucosa three to four times normal thickness; abnormal collagen fibres | Massive fibrosis including muscularis |
| 6 - Lymph Congestion | No changes | Dilated lymph vessels or cystic collections of lymph | N/A | N/A |
| 7 - Ileitis Cystica Profunda | No changes | Submucosal glandular inclusions | submucosal cysts with polypoid elevation of the mucosa | Large cysts extending into the muscularis |

**Table SUPP-4:** Definition of symptom groups according to PRO in the late cohort.

|  |  |  |
| --- | --- | --- |
| **PRO score** | **Symptom group** | |
| ≤ first quartile (62≤) | | Severe GI symptoms |
| First quartile to median (62< & ≤67) | | Moderate GI symptoms |
| Median to third quartile (67< & <70) | | Mild GI symptoms |
| ≥ third quartile (=70) | | No GI symptoms |

*The best possible PRO score is 70. GI = gastrointestinal*

Table SUPP-5: Number of patients per longitudinal symptom group in the early cohort.

|  |  |
| --- | --- |
| Symptom group | n (%) PRO/CRO [n concordant\*] |
| No symptoms PRO/CRO | 4 (13%)/4 (13%) [2] |
| Non-persistent symptoms PRO/CRO | 19 (59%)/12 (37%) [11] |
| Persistent symptoms PRO/CRO | 9 (28%)/16 (50%)[8] |

*\* : Concordant classification is defined as patients being classified in the same group with both PRO and CRO.*

**Table SUPP-6:** Comorbidity comparisons between symptom groups in the early cohort.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| PRO | | | | |
| *↓: Item / 🡪: symptom group* | ***No symptoms***  ***n=4***  ***(Mean (SD) or %)*** | ***Non-persistent symptoms n=20***  ***(Mean (SD) or %)*** | ***Persistent symptoms***  ***n=8***  ***(Mean (SD) or %)*** | ***p-value*** |
| Age | 64(12.8) | 67.5(6.7) | 65.125(3.8) | 0.59 |
| Presenting PSA | 30.5(15.9) | 43.82(45.5) | 31.3(22.9) | 0.97 |
| BMI | 34.1 (7.2) | 27.2 (3.9) | 28.2 (3.4) | 0.14 |
| Smoking status (NS/ES/S) | 50%/50%/0% | 60%/30%/10% | 63%/37%/0% | 0.92 |
| Pack/year | 14.5(17.5) | 12.1(22.8) | 15.25(22.9) | 0.93 |
| Previous abdominal surgery | 75% | 55% | 63% | 0.75 |
| Diabetes | 50% | 10% | 38% | 0.11 |
| Statins | 50% | 35% | 13% | 0.36 |
| Other metabolic disease | 0% | 0% | 0% | NA |
| IBS | 0% | 0% | 0% | NA |
| Antihypertensives | 50% | 45% | 25% | 0.58 |
| Metformin | 25% | 10% | 13% | 0.72 |
| Aspirin | 0% | 30% | 13% | 0.33 |
| Warfarin | 0% | 0% | 0% | NA |
| CRO | | | | |
| *↓: Item / 🡪: symptom group* | ***No symptoms***  ***n=4***  ***(Mean (SD) or %)*** | ***Non-persistent symptoms***  ***n=12***  ***(Mean (SD) or %)*** | ***Persistent symptoms***  ***n=16***  ***(Mean (SD) or %)*** | ***p-value*** |
| Age | 72.5(3.7) | 66.7(7.9) | 64.8 (3.8) | 0.10 |
| Presenting PSA | 23.1(21.1) | 54.9(44.6) | 31.1(22.9) | 0.41 |
| BMI | 30.4 (4.8) | 27.3 (5.2) | 28.5 (3.4) | 0.46 |
| Smoking status (NS/ES/S) | 50%/50%/0% | 58%/32%/8% | 63%/31%/6% | 0.94 |
| Pack/year | 21.5(25.7) | 9.7 (21.1) | 13.8(22.9) | 0.74 |
| Previous abdominal surgery | 100% | 58% | 50% | 0.20 |
| Diabetes | 25% | 17% | 25% | 0.86 |
| Statins | 50% | 33% | 25% | 0.63 |
| Other metabolic disease | 0% | 0% | 0% | NA |
| IBS | 0% | 0% | 0% | NA |
| Antihypertensives | 50% | 50% | 31% | 0.57 |
| Metformin | 0% | 17% | 13% | 0.69 |
| Aspirin | 25% | 25% | 19% | 0.92 |
| Warfarin | 0% | 0% | 0% | NA |

**Table SUPP-7A:** Comorbidity comparisons between PRO-stratified symptom groups in the late cohort.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| PRO (actual) | | | | | | |
| *↓: Item / 🡪 : symptom group* | ***No symptoms***  ***n=24***  ***(Mean (SD) or %)*** | ***Mild symptoms***  ***n=10***  ***(Mean (SD) or %)*** | ***Moderate symptoms***  ***n=20***  ***(Mean (SD) or %)*** | ***Severe symptoms***  ***n=20***  ***(Mean (SD) or %)*** | ***-*** | ***p-value*** |
| Age | 74 (7) | 75 (6) | 73(7) | 73(6) | - | 0.84 |
| Presenting PSA | 22 (16) | 29 (19) | 42 (38) | 22 (30) | - | 0.05 |
| Recurrent PCa | 8% | 20% | 15% | 10% | - | 0.77 |
| On ADT | 8% | 10% | 10% | 10% | - | 1 |
| Testosterone recovery to normal | 58% | 50% | 60% | 45% | - | 0.76 |
| BMI | 26 (6) | 27 (3) | 26 (5) | 28 (4) | - | 0.73 |
| Smoking status (NS/ES/S) | 54%/38%/8% | 40%/60%/0% | 60%/40%/0% | 35%/35%/30% | - | 0.13 |
| Pack/year | 9 (16) | 28 (35) | 8(16) | 23(27) | - | 0.08 |
| Previous abdominal surgery | 46% | 70% | 30% | 50% | - | 0.22 |
| Diabetes | 13% | 20% | 20% | 15% | - | 0.90 |
| Statins | 46% | 90% | 50% | 55% | - | 0.12 |
| Other metabolic disease | 8% | 0% | 0% | 5% | - | 0.49 |
| IBS | 0% | 0% | 5% | 0% | - | 0.44 |
| Antihypertensives | 46% | 80% | 60% | 50% | - | 0.30 |
| Metformin | 8% | 20% | 20% | 10% | - | 0.61 |
| Aspirin | 25% | 40% | 35% | 30% | - | 0.82 |
| Warfarin | 4% | 10% | 5% | 10% | - | 0.84 |
| PRO (historic) | | | | | | |
| *↓: Item / 🡪 : symptom group* | ***No bowel problem***  ***n=16***  ***(Mean (SD) or %)*** | ***Very small bowel problem***  ***n=22***  ***(Mean (SD) or %)*** | ***Small bowel problem***  ***n=18***  ***(Mean (SD) or %)*** | ***Moderate bowel problem***  ***n=21***  ***(Mean (SD) or %)*** | ***Big bowel problem***  ***n=7***  ***(Mean (SD) or %)*** | ***p-value*** |
| Age | 73 (6) | 72 (7) | 76 (6) | 73 (7) | 72 (7) | 0.29 |
| Presenting PSA | 34 (35) | 35 (37) | 22 (14) | 26 (21.7) | 20 (10) | 0.86 |
| Recurrent PCa | 13% | 14% | 22% | 0% | 29% | 0.21 |
| On ADT | 6% | 18% | 6% | 5% | 29% | 0.27 |
| Testosterone recovery to normal | 31% | 59% | 56% | 71% | 29% | 0.10 |
| BMI | 28 (4) | 24 (6) | 28 (4) | 28 (4) | 27 (4) | 0.27 |
| Smoking status (NS/ES/S) | 44%/44%/12% | 46%/40%/14% | 61%/33%/6% | 24%/52%/24% | 43%/43%/14% | 0.18 |
| Pack/year | 21 (31) | 14 (23) | 10 (15) | 13 (17) | 22 (30) | 0.53 |
| Previous abdominal surgery | 50% | 36% | 44% | 48% | 57% | 0.86 |
| Diabetes | 13% | 14% | 17% | 24% | 14% | 0.89 |
| Statins | 50% | 45% | 67% | 48% | 43% | 0.68 |
| Other metabolic disease | 6% | 5% | 0% | 0% | 14% | 0.39 |
| IBS | 0% | 0% | 6% | 10% | 0% | 0.41 |
| Antihypertensives | 50% | 45% | 72% | 62% | 43% | 0.42 |
| Metformin | 6% | 14% | 17% | 14% | 14% | 0.93 |
| Aspirin | 13% | 32% | 33% | 48% | 29% | 0.27 |
| Warfarin | 0% | 9% | 6% | 5% | 14% | 0.68 |

**Table SUPP-7B:** Comorbidity comparisons between CRO-stratified symptom groups in the late cohort.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| CRO (actual) | | | | | |
| *↓: Item / 🡪: symptom group* | ***Grade 0***  ***n=53***  ***(Mean (SD) or %)*** | ***Grade 1***  ***n=14***  ***(Mean (SD) or %)*** | ***Grade 2***  ***n=18***  ***(Mean (SD) or %)*** | ***Grade 3***  ***n=2***  ***(Mean (SD) or %)*** | ***p-value*** | |
| Age | 73 (7) | 76 (7) | 71 (5) | 75(11) | 0.29 | |
| Presenting PSA | 28 (26) | 20 (14) | 37 (38) | 13 (3) | 0.37 | |
| Recurrent PCa | 9% | 21% | 17% | 0% | 0.57 | |
| On ADT | 11% | 7% | 17% | 0% | 0.81 | |
| Testosterone recovery to normal | 58% | 43% | 50% | 50% | 0.74 | |
| BMI | 27 (5) | 28 (4) | 27 (3) | 29 (8) | 0.83 | |
| Smoking status (NS/ES/S) | 43%/45%/11% | 36%/57%/7% | 44%/28%/28% | 50%/50%/0% | 0.92 | |
| Pack/year | 12 (22) | 21 (28) | 17 (18) | 2(2) | 0.47 | |
| Previous abdominal surgery | 43% | 36% | 56% | 100% | 0.30 | |
| Diabetes | 11% | 36% | 22% | 0% | 0.15 | |
| Statins | 51% | 57% | 56% | 0% | 0.49 | |
| Other metabolic disease | 6% | 0% | 0% | 0% | 0.58 | |
| IBS | 0% | 0% | 11% | 50% | **0.0004** | |
| Antihypertensives | 57% | 50% | 61% | 50% | 0.94 | |
| Metformin | 8% | 36% | 17% | 0% | **0.05** | |
| Aspirin | 26% | 57% | 28% | 0% | 0.12 | |
| Warfarin | 11% | 0% | 0% | 0% | 0.25 | |
| CRO (historic) | | | | | |
| *↓: Item / 🡪: symptom group* | ***Grade 0***  ***n=33***  ***(Mean (SD) or %)*** | ***Grade 1***  ***n=28***  ***(Mean (SD) or %)*** | ***Grade 2***  ***n=20***  ***(Mean (SD) or %)*** | ***Grade 3***  ***n=5***  ***(Mean (SD) or %)*** | ***p-value*** | |
| Age | 74 (7) | 73 (7) | 73 (7) | 73 (7) | 0.92 | |
| Presenting PSA | 28 (30) | 26 (16) | 28 (29) | 44 (56) | 0.92 | |
| Recurrent PCa | 9% | 11% | 20% | 20% | 0.65 | |
| On ADT | 12% | 7% | 15% | 0% | 0.70 | |
| Testosterone recovery to normal | 58% | 46% | 65% | 40% | 0.54 | |
| BMI | 28 (4) | 26 (7) | 26 (3) | 27 (6) | 0.41 | |
| Smoking status (NS/ES/S) | 39%/49%/12% | 43%/50%/7% | 50%/25%/25% | 40%/40%/20% | 0.97 | |
| Pack/year | 13 (23) | 17 (24) | 12 (16) | 23 (35) | 0.86 | |
| Previous abdominal surgery | 52% | 36% | 45% | 60% | 0.58 | |
| Diabetes | 18% | 18% | 20% | 0% | 0.77 | |
| Statins | 45% | 54% | 55% | 60% | 0.86 | |
| Other metabolic disease | 6% | 0% | 0% | 20% | 0.10 | |
| IBS | 3% | 0% | 5% | 20% | 0.16 | |
| Antihypertensives | 52% | 61% | 55% | 80% | 0.65 | |
| Metformin | 12% | 18% | 15% | 0% | 0.74 | |
| Aspirin | 27% | 39% | 35% | 0% | 0.33 | |
| Warfarin | 6% | 4% | 10% | 20% | 0.55 | |

**Table SUPP-8:** Comorbidity comparisons between CRO-stratified symptom groups in the colonoscopy cohort.

|  |  |  |  |
| --- | --- | --- | --- |
| *↓: Item / 🡪: symptom group* | *Cases*  *n=9*  *(Mean (SD) or %)* | *Controls*  *n=6*  *(Mean (SD) or %)* | *p-value* |
| Age | 62 (13) | 73 (4) | 0.04 |
| BMI | 24 (2) | 27 (4) | 0.10 |
| Smoking status (NS/ES/S) | 17%/83%/0% | 44%/44%/11% | 0.50 |
| Pack/year | 13 (11) | 19 (22) | 0.95 |
| Previous abdominal surgery | 50% | 67% | 0.53 |
| Diabetes | 0% | 0% | NA |
| Statins | 33% | 44% | 0.68 |
| Other metabolic disease | 0% | 0% | NA |
| IBS | 33% | 22% | 0.65 |
| Antihypertensives | 17% | 78% | **0.02** |
| Metformin | 0% | 0% | NA |
| Aspirin | 0% | 33% | 0.13 |
| Warfarin | 0% | 11% | 0.41 |

**Table SUPP-9:** Results of linear mixed model regression analysis for exploring

associations between Chao richness and symptom group over time*.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Effect | Estimate | Standard error | Degrees of freedom | t-value | p-value |
| PRO-stratified groups | | | | | |
| Timepoint | -0.08 | 0.04 | 97.10 | -2.21 | **0.03** |
| Symptom group\* | -0.20 | 0.15 | 33.03 | -1.31 | 0.20 |
| Timepoint by symptom group | 0.06 | 0.03 | 97.21 | 1.99 | **0.05** |
| CRO-stratified groups | | | | | |
| Timepoint | -0.02 | 0.04 | 88.84 | -0.54 | 0.59 |
| Symptom group\* | -0.05 | 0.13 | 30.91 | -0.41 | 0.69 |
| Timepoint by symptom group | 0.003 | 0.03 | 84.96 | 0.13 | 0.90 |

*\*: Symptom group was coded as 0 = no symptoms at either 4/5 weeks or 6 months; 1 = symptoms at either 4/5 weeks or 6 months; 2 = symptoms at 4/5 weeks and 6 months. Therefore a positive effect indicates rising counts of the feature with rising symptoms.*

Table SUPP-10: Exploration of differences in microbial features between patients with

and without persistent gastrointestinal side-effects in the early cohort.

| Timepoint | Microbial feature | Mean proportion of microbial feature  (No symptoms/Symptoms at one timepoint/Persistent symptoms) | SD of microbial feature  (No symptoms/Symptoms at one timepoint/Persistent symptoms) | Eta-squared | p\*-value  (p-value) |
| --- | --- | --- | --- | --- | --- |
| PRO-stratified groups | | | | | |
| Baseline | *Clostridium IV* | 0.00%/0.14%/1.37% | 0.00%/0.39%/1.66% | 0.27 | 0.003 (>0.1) |
| *Roseburia* | 0.07%/2.06%/1.98% | 0.13%/3.02%/1.44% | 0.06 | 0.38 (>0.1) |
| 2/3 weeks | *Clostridium IV* | 0.00%/0.07%/0.32% | 0.00%/0.15%/0.44% | 0.18 | 0.06 (>0.1) |
| *Clostridium XIVa* | 0.14%/0.72%/2.10% | 0.14%/1.07%/2.98% | 0.13 | 0.13 (>0.1) |
| 4/5 weeks | Proteobacteria | 0.72%/1.97%/1.44% | 0.95%/1.83%/0.80% | 0.07 | 0.34 (>0.1) |
|  | *Sutterella* | 0.00%/0.59%/1.05% | 0.00%/0.95%/0.97% | 0.11 | 0.20 (>0.1) |
| 12 weeks | Proteobacteria | 0.65%/1.49%/2.52% | 0.96%/1.59%/3.34% | 0.07 | 0.37 (>0.1) |
| *Clostridium IV* | 0.00%/0.06%/0.79% | 0.005/0.21%/1.11% | 0.23 | 0.03 (>0.1) |
| *Odoribacter* | 0.00%/0.13%/0.29% | 0.00%/0.26%/0.25% | 0.13 | 0.15 (>0.1) |
| *Sutterella* | 0.00%/0.42%/0.76% | 0.00%/0.61%/0.64% | 0.14 | 0.14 (>0.1) |
| 6 months | Proteobacteria | 0.87%/1.43%/1.44% | 0.20%/1.42%/0.61% | 0.03 | 0.69 (>0.1) |
| *Roseburia* | 1.37%/0.64%/1.91% | 0.96%/0.62%/1.85% | 0.20 | 0.04 (>0.01) |
| CRO-stratified groups | | | | | |
| Baseline | NA | NA | NA | - | - |
| 2/3 weeks | NA | NA | NA | - | - |
| 4/5 weeks | *Sutterella* | 0.07%/0.72%/0.70% | 0.13%/1.05%/0.93% | 0.05 | 0.48 (>0.1) |
| *Phascolarctobacterium* | 0.00%/0.50%/1.15% | 0.00%/1.07%/2.19% | 0.06 | 0.42 (>0.1) |
| 12 weeks | *Sutterella* | 0.07%/0.50%/0.52% | 0.13%/0.66%/0.63% | 0.06 | 0.45 (>0.01) |
| 6 months | *Odoribacter* | 0.07%/0.19%/0.35% | 0.13%/0.21%/0.30% | 0.13 | 0.13 (>0.01) |
| *Roseburia* | 0.65%/0.77%/1.40% | 0.31%/0.80%/1.56% | 0.07 | 0.35 (>0.01) |
| *Sutterella* | 0.22%/0.72%/0.56% | 0.37%/0.82%/0.76% | 0.04 | 0.55 (>0.01) |

*Only biologically plausible relationships are reported (see main text).SD=standard deviation. Significance testing was done with ANOVA (comparison of proportions of microbial feature between symptom groups).* *Direct comparisons at 12 months were not undertaken to avoid biases, as only 10 patients were sampled at that timepoint.*

**Table SUPP-11:** Results of linear mixed model regression analysis for exploring associations

between microbial features and longitudinal symptom group over time in the early cohort*.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Microbial feature | Effect | Estimate | Standard error | Degrees of freedom | t-value | p-value |
| PRO-stratified groups | | | | | | |
| Clostridium IV | Timepoint | 0.03 | 0.03 | 116.77 | 1.04 | 0.30 |
| Symptom group\* | 0.40 | 0.14 | 33.86 | 2.90 | **0.007** |
| Timepoint by symptom group | -0.04 | 0.03 | 116.60 | -1.62 | 0.11 |
| Roseburia | Timepoint | 0.02 | 0.07 | 28.70 | 0.27 | 0.80 |
| Symptom group\* | 0.37 | 0.21 | 28.23 | 1.79 | **0.08** |
|  | Timepoint by symptom group | -0.07 | 0.06 | 29.86 | -1.22 | 0.23 |
| Phascolarctobacterium | Timepoint | 0.00 | 0.05 | 25.31 | 0.08 | 0.94 |
| Symptom group\* | -0.03 | 0.18 | 31.21 | -0.18 | 0.86 |
| Timepoint by symptom group | 0.03 | 0.04 | 25.90 | 0.85 | 0.40 |
| Sutterella | Timepoint | -0.01 | 0.05 | 26.61 | -0.30 | 0.78 |
| Symptom group\* | 0.19 | 0.18 | 29.07 | 1.03 | 0.31 |
| Timepoint by symptom group | 0.02 | 0.04 | 27.35 | 0.64 | 0.53 |
| CRO-stratified groups | | | | | | |
| Clostridium IV | Timepoint | -0.00 | 0.03 | 94.30 | -0.13 | 0.90 |
| Symptom group\* | 0.17 | 0.13 | 31.01 | 1.32 | 0.20 |
| Timepoint by symptom group | -0.01 | 0.02 | 90.29 | -0.31 | 0.76 |
| Roseburia | Timepoint | -0.07 | 0.08 | 27.98 | -0.93 | 0.36 |
| Symptom group\* | 0.42 | 0.18 | 26.44 | 0.23 | 0.82 |
| Timepoint by symptom group | 0.01 | 0.05 | 24.76 | 0.18 | 0.86 |
| Phascolarctobacterium | Timepoint | 0.05 | 0.05 | 25.11 | 1.05 | 0.31 |
| Symptom group\* | 0.26 | 0.14 | 30.22 | 1.77 | **0.09** |
| Timepoint by symptom group | -0.01 | 0.03 | 23.11 | -0.28 | 0.78 |
| Sutterella | Timepoint | -0.03 | 0.05 | 25.57 | -0.77 | 0.45 |
| Symptom group\* | -0.07 | 0.16 | 28.29 | -0.46 | 0.65 |
| Timepoint by symptom group | 0.03 | 0.03 | 23.01 | 1.16 | 0.26 |

*\*: Symptom group was coded as 0 =* ***no*** *symptoms at either 4/5 weeks or 6 months; 1 = symptoms at* ***either*** *4/5 weeks* ***or*** *6 months; 2 = symptoms at 4/5 weeks* ***and*** *6 months. Therefore a positive effect indicates rising counts of the feature with rising symptoms.*

**Table SUPP-12:** Differences in proportions of *Roseburia* between patients with and without RE in the late cohort.\*

| Microbial feature | Mean  (Grade 0/1/2/3) | SD  (Grade 0/1/2/3) | Eta-squared | p\*-value  (p-value) |
| --- | --- | --- | --- | --- |
| Actual toxicity\* | | | | |
| CRO maximum toxicity | | | | |
| *Roseburia* | 0.78%/1.11%/1.51%/9.65% | 1.08/2.13/2.18/8.79 | 0.30 | <0.0001 (<0.0001) |
| CRO diarrhoea | | | | |
| *Roseburia* | 0.83%/0.67%/3.17%/9.65% | 1.30%/0.76%/3.07%/8.79% | 0.34 | <0.0001 (<0.0001) |
| CRO proctitis | | | | |
| NA (see main text) | - | - | - | - |
| HISTORICAL toxicity | | | | |
| CRO maximum toxicity | | | | |
| *Roseburia* | 0.55%/0.92%/1.67%/4.90% | 0.80%/1.55%/2.19%/7.00% | 0.17 | 0.001 (0.06) |
| CRO diarrhoea | | | | |
| *Roseburia* | 0.59%/1.15%/2.36%/9.22% | 0.77%/2.06%/2.45%/9.22% | 0.33 | <0.001 (<0.001) |
| CRO proctitis | | | | |
| *Roseburia* | 0.66%/1.30%/2.67%/0.38% | 1.11%/2.00%/5.01%/0.36% | 0.07 | 0.09 (>0.1) |
| PRO | | | | |
| *Roseburia* | 0.29%/0.71%/1.19%/1.75%/3.10% | 0.38%/0.90%/1.94%/2.53%/5.85% | 0.09 | 0.11 (>0.1) |

*\*: No significant or relevant differences with actual PRO-based symptom grouping were found.*

*p\* represents the p-value prior to correction for multiple testing.*

**Table SUPP-13:** Results of linear mixed model regression analysis for exploring associations

between metagenomic SCFA-related metabolic pathways and longitudinal symptom group over time in the early cohort*.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Pathway and KEGG identifier | Effect | Estimate | Standard error | Degrees of freedom | t-value | p-value |
| PRO-stratified groups | | | | | | |
| Butanoate metabolism  ko00650 | Timepoint | 0.20 | 0.34 | 109.41 | 0.58 | 0.57 |
| Symptom group\* | 2.38 | 1.92 | 30.83 | 1.24 | 0.22 |
| Timepoint by symptom group | -0.02 | 0.27 | 110.53 | -0.09 | 0.93 |
| Propanoate metabolism  ko00640 | Timepoint | 0.46 | 0.36 | 83.06 | 1.26 | 0.21 |
| Symptom group\* | 3.74 | 2.00 | 30.95 | 1.87 | 0.07 |
|  | Timepoint by symptom group | -0.18 | 0.29 | 84.40 | -0.63 | 0.53 |
| Fatty acid degradation  ko00071 | Timepoint | 0.04 | 0.20 | 110.35 | 0.20 | 0.84 |
| Symptom group\* | 0.87 | 1.04 | 31.13 | 0.84 | 0.41 |
| Timepoint by symptom group | 0.05 | 0.16 | 111.51 | 0.28 | 0.78 |
| Fatty acid metabolism  ko01212 | Timepoint | 0.43 | 0.35 | 101.27 | 1.22 | 0.23 |
| Symptom group\* | 3.28 | 2.07 | 30.76 | 1.58 | 0.12 |
| Timepoint by symptom group | -0.14 | 0.28 | 102.50 | -0.50 | 0.62 |
| Metabolic pathways  ko01100 | Timepoint | 2.31 | 1.94 | 109.99 | 1.19 | 0.24 |
| Symptom group\* | 17.80 | 11.31 | 30.85 | 1.57 | 0.13 |
| Timepoint by symptom group | -0.76 | 1.53 | 111.08 | -0.50 | 0.62 |
| CRO-stratified groups | | | | | | |
| Butanoate metabolism  ko00650 | Symptom group\* | 0.25 | 0.35 | 109.99 | 0.72 | 0.48 |
| Timepoint | 0.32 | 1.69 | 30.08 | 0.19 | 0.85 |
| Timepoint by symptom group | -0.06 | 0.22 | 106.36 | -0.26 | 0.80 |
| Propanoate metabolism  ko00640 | Timepoint | 0.62 | 0.37 | 88.44 | 1.68 | 0.10 |
| Symptom group\* | 1.32 | 1.80 | 30.17 | 0.73 | 0.47 |
| Timepoint by symptom group | -0.26 | 0.23 | 83.63 | -1.10 | 0.27 |
| Fatty acid degradation  ko00071 | Timepoint | 0.19 | 0.21 | 107.17 | 0.90 | 0.37 |
| Symptom group\* | 0.24 | 0.90 | 30.27 | 0.26 | 0.80 |
| Timepoint by symptom group | -0.07 | 0.13 | 103.19 | -0.51 | 0.61 |
| Fatty acid metabolism  ko01212 | Timepoint | 0.57 | 0.36 | 106.03 | 1.58 | 0.12 |
| Symptom group\* | 0.83 | 1.84 | 30.08 | 0.45 | 0.66 |
| Timepoint by symptom group | -0.21 | 0.23 | 102.11 | -0.92 | 0.36 |
| Metabolic pathways  ko01100 | Timepoint | 2.43 | 1.99 | 114.96 | 1.23 | 0.22 |
| Symptom group\* (RTOG class) | 3.23 | 10.06 | 30.11 | 0.32 | 0.75 |
| Timepoint by symptom group | -0.68 | 1.25 | 111.87 | -0.54 | 0.59 |

*\*: Symptom group was coded as 0 =* ***no*** *symptoms at either 4/5 weeks or 6 months; 1 = symptoms at* ***either*** *4/5 weeks* ***or*** *6 months; 2 = symptoms at 4/5 weeks* ***and*** *6 months. Therefore a positive effect indicates rising counts of the feature with rising symptoms. KEKK = Kyoto Encyclopaedia of Genes and Genomes.*

**Table SUPP-14:** Differences in proportions of microbial metabolic pathway abundance between

patients with and without late clinician-reported diarrhoea in the late cohort.

| Pathway and KEGG identifier | Mean  (Grade 0/1/2/3) | SD  (Grade 0/1/2/3) | Eta-squared | p-value |
| --- | --- | --- | --- | --- |
| Actual toxicity – CRO diarrhoea | | | | |
| Butanoate metabolism  ko00650 | 2.70/3.19/2.56/3.09 | 0.41/0.67/0.21/0.76 | 0.07 | 0.12 |
| Propanoate metabolism  ko00640 | 2.84/2.84/2.67/3.08 | 0.50/0.50/0.33/0.36 | 0.01 | 0.78 |
| Fatty acid degradation  ko00071 | 0.78/0.93/0.66/0.82 | 0.19/0.24/0.12/0.63 | 0.04 | 0.33 |
| Fatty acid metabolism  ko01212 | 2.96/2.91/2.84/2.48 | 0.15/0.08/0.17/0.48 | 0.19 | 0.0006 |
| Metabolic pathways  ko01100 | 90.72/90.13/91.27/90.53 | 0.97/1.28/0.31/2.23 | 0.03 | 0.43 |
| HISTORICAL toxicity – cro diarrhoea | | | | |
| Butanoate metabolism  ko00650 | 2.72/2.82/2.61/2.60 | 0.42/0.60/0.33/0.27 | 0.02 | 0.66 |
| Propanoate metabolism  ko00640 | 2.84/2.99/2.72/2.50 | 0.53/0.46/0.35/0.21 | 0.03 | 0.4 |
| Fatty acid degradation  ko00071 | 0.79/0.82/0.72/0.52 | 0.20/0.28/0.15/0.33 | 0.05 | 0.22 |
| Fatty acid metabolism  ko01212 | 2.96/2.99/2.89/2.47 | 0.14/0.12/0.19/0.47 | 0.19 | 0.0007 |
| Metabolic pathways  ko01100 | 90.69/90.38/91.05/91.91 | 1.02/1.19/0.59/0.86 | 0.07 | 0.13 |

Table SUPP-15: Differences in histology scores (colonoscopy cohort).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item | Thickening of Serosa | Mucosal Ulcerations | Epithelial atypia | Vascular Sclerosis | Intestinal wall fibrosis | Lymph Congestion | Ileitis cystica profunda | Total score |
| Score in cases  Distal sigmoid  (mean/SD) | 0 (0) | 0 (0) | 0.22 (0.44) | 0.22 (0.44) | 0.33 (0.50) | 0 (0) | 0 (0) | 0.78 (0.83) |
| Score in cases  Rectal biopsy  (mean/SD) | 0 (0) | 0 (0) | 0.22 (0.44) | 0.44 (0.53) | 0.44 (0.53) | 0 (0) | 0 (0) | 1.11 (1.17) |
| Score in controls  Rectal biopsy  (mean/SD) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| p-value | 0.22 | 0.22 | 1.00 | 0.55 | 0.71 | 0.22 | 0.22 | 0.48 |

*P-values were obtained with the Kruskal-Wallis test.*

Table SUPP-16: Functions of cytokines significantly different between rectal biopsy samples of cases and controls.

|  |  |  |
| --- | --- | --- |
| Cytokine | Function | Observation in THIS STUDY |
| **Eotaxin** | Chemotactant for eosinophils and basophils.5 | Elevated in cases |
| **IL-7** | Regulatory factor for intestinal lymphocytes and intestinal epithelial cells. Contributes for intestinal epithelial cell homeostasis (IL-7 promotes hyperplasia). Microbiota antigens promote IL7 gene expression in the intestine and antibiotic-induced microbiota depletion associates with reduced colon wall thickness.6 | Decreased in cases |
| **IL-12/IL-23p40** | IL-23 is a heterodimeric cytokine (IL-23p19 and IL-12p40).7  Mediates local inflammatory response against infection through the innate immune system and is essential for the development of IBD.8,9 The IL-23/Th17 axis is associated to protective immunity and repair at mucosal surfaces.10 | Decreased in cases |
| **IL-15** | Upregulates T cell and NK cell activation and proliferation through the IL-2 receptor, for which it competes with IL-2.11  Is essential in the microbiota signal-detecting NOD2 pathway for maintenance of intestinal intra-epithelial lymphocytes (which reduce risk of colitis).12 Also protects from colitis through inhibition of apoptosis of intestinal epithelial cells in a caspase 3-dependent mechanism.13 | Decreased in cases |
| **IL-16** | Chemoattractant which can be produced by cells including the lymphoid lineage and epithelial cells.14 Recruits and reversibly activates CD4+ cells (lymphocytes, dendritic cells, monocytes, eosinophils).15 Induces the expression of the IL-2 receptor and epidermal growth factor (which promotes intestinal barrier function).16,17 | Decreased in cases |