



**UNIVERSITY OF LEEDS**

This is a repository copy of *High-Dimensional Clustering of 4000 Irritable Bowel Syndrome Patients Reveals Seven Distinct Disease Subsets*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/194599/>

Version: Accepted Version

---

**Article:**

Byale, A, Lennon, RJ, Byale, S et al. (7 more authors) (2024) High-Dimensional Clustering of 4000 Irritable Bowel Syndrome Patients Reveals Seven Distinct Disease Subsets. *Clinical Gastroenterology and Hepatology*, 22 (1). E12. pp. 173-184. ISSN: 1542-3565

<https://doi.org/10.1016/j.cgh.2022.09.019>

---

© 2022 by the AGA Institute. This is an author produced version of a paper published in *Clinical Gastroenterology and Hepatology*. Uploaded in accordance with the publisher's self-archiving policy. This manuscript version is made available under the Creative Commons CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

## **High-dimensional clustering of 4000 Irritable Bowel Syndrome Patients reveals Seven Distinct Disease Subsets**

Anjali Byale, MD<sup>1</sup>, Ryan J. Lennon, MS<sup>2</sup>, Siddharth Byale, BE<sup>3</sup>, Margaret Breen-Lyles<sup>1</sup>, Adam L. Edwinson, PhD<sup>1</sup>, Ruchi Gupta, MS<sup>2</sup>, Brian E. Lacy, MD<sup>4</sup>, Janet E. Olson, PhD<sup>2</sup>, Lesley A. Houghton, PhD<sup>4,5</sup>, Madhusudan Grover, MD<sup>1</sup>

<sup>1</sup>Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN, USA

<sup>2</sup>Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA

<sup>3</sup>Viterbi School of Engineering, University of Southern California, Los Angeles, CA, USA

<sup>4</sup>Division of Gastroenterology and Hepatology, Mayo Clinic, Jacksonville, FL, USA

<sup>5</sup>Division of Gastroenterology and Surgical Sciences, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, U.K.

**Abbreviations:** IBS, irritable bowel syndrome; GI, gastrointestinal; IBS-M, mixed IBS; IBS-D, diarrhea-predominant IBS; IBS-C, constipation-predominant IBS; IBS-U, unsubtyped IBS; OR, odds ratio; CI, confidence interval; MDCP, Multi-Dimensional Clinical Profile; IBS-SSS, IBS-symptom severity scale; ARM, anorectal manometry; LCA, Latent Class Analysis; BIC, Bayesian Information Criterion; QoL, quality of life

### **Corresponding author:**

Madhusudan Grover, M.D.

Associate Professor

Department of Medicine and Physiology

Enteric Neuroscience Program

Mayo Clinic

200 First St SW

Rochester, MN 55905, USA

E-mail: [grover.madhusudan@mayo.edu](mailto:grover.madhusudan@mayo.edu)

Twitter: @mgrover\_gi

Phone: 507-284-2478; Fax: 507-284-0266

**Word count (including table/figure legends and references):** 4,000

**Writing Assistance:** None

**Acknowledgment:** The authors acknowledge Ms. Amanda Bedard for administrative assistance. The graphical abstract was generated using BioRender.com.

**Conflicts of interest:** The authors disclose no conflicts.

**Funding:** This work was supported by NIH DK 127998 to the corresponding author.

**Author contributions:**

**AB:** Conceptualization, Data curation, Investigation, Methodology, Software, Visualization, Writing

**RJL:** Data curation, Investigation, Methodology, Software, Supervision, Writing

**SB:** Investigation, Methodology, Software, Writing

**MBL:** Data curation, Investigation, Writing

**ALE:** Resources, Visualization, Writing

**RG:** Data curation, Investigation, Methodology, Software, Writing

**BEL:** Conceptualization, Visualization, Writing

**JEO:** Conceptualization, Data curation, Funding acquisition, Methodology, Resources, Software, Writing

**LAH:** Conceptualization, Visualization, Writing

**MG:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Visualization, Writing

**Abstract (260 words)**

**Background:** Irritable bowel syndrome (IBS) is a pain disorder classified by bowel habits, disregarding other factors that may influence the clinical course. The aim of this study was to determine if IBS patients can be clustered based on clinical, dietary, lifestyle, and psychosocial factors.

**Methods:** Between 2013-20, Mayo Clinic Biobank surveyed and received 40,291 responses to a questionnaire incorporating Rome III criteria. Factors associated with IBS were determined and, latent class analysis, a model-based clustering, was performed on IBS cases.

**Results:** We identified 4,021 IBS patients (mean 64 years; 75% female) and 12,063 controls. Using 26 variables separating cases from controls, the optimal clustering revealed seven latent clusters. These were characterized by perceived health impairment (moderate or severe), psychoneurological factors, and bowel dysfunction (diarrhea or constipation predominance). Health impairment clusters demonstrated more pain, with “severe” cluster also having more psychiatric comorbidities. The next three clusters had unique enrichment of psychiatric, neurological or both comorbidities. The bowel dysfunction clusters demonstrated less abdominal pain, with diarrhea cluster most likely to report pain improvement with defecation. Constipation cluster had the highest exercise score, consumption of fruits, vegetables, and alcohol. The distribution of clusters remained similar when Rome IV criteria were applied. Physiologic tests were available on a limited subset (6%), and there were no significant differences between clusters.

**Conclusions:** In this cohort of older IBS patients, seven distinct clusters were identified demonstrating varying degrees of GI symptoms, comorbidities, dietary, and lifestyle factors. Further research is required to assess whether these unique clusters could be used to direct clinical trials and individualize patient management.

**Keywords:** severity; psychosocial; disorders of gut-brain interaction; pain; comorbidities

## **Introduction**

Irritable bowel syndrome (IBS) is a disorder of gut-brain interaction characterized by abdominal pain related to defecation and changes in the consistency and/or frequency of bowel movements. IBS has a high global prevalence (7-21%); and peripheral (increased intestinal permeability, visceral hypersensitivity, dysmotility) and central (altered brain signaling, disrupted coping) mechanisms play a role in its pathophysiology.<sup>1</sup> While co-morbidities, diet, and lifestyle influence the mechanisms and clinical course, these factors are overlooked in the current classification paradigm.<sup>2</sup> The Rome Foundation created the concept of Multi-Dimensional Clinical Profile (MDCP) to help capture the clinical spectrum of IBS but it has not gained broad clinical or research utility.<sup>3</sup>

The IBS-symptom severity scale (IBS-SSS), the most widely used scoring system for severity, captures abdominal pain, distension, bowel dysfunction, and quality of life/global well-being.<sup>4</sup> A study found that the IBS-SSS is not unidimensional and subgroups of IBS patients were variably influenced by components of IBS-SSS.<sup>5</sup> In addition to psychiatric conditions, musculoskeletal and neurological comorbidities are common in IBS.<sup>6,7</sup> Additionally, there is an expanding understanding for the role of diet in the pathophysiology of IBS.<sup>8</sup> A recent study showed that 27% of IBS patients were following restrictive diets, which was associated with greater IBS symptom severity.<sup>9</sup> Lastly, a few studies have inversely associated physical activity with IBS symptom severity.<sup>10,11</sup> There is a paucity of large studies that incorporate complex dimensions of diet, exercise, and comorbidities noted in IBS patients. We hypothesized that IBS patients can be segregated into subsets that go beyond the conventional categories driven by predominant bowel habits.

The aims of this study were to determine (a) clinical, dietary, lifestyle, and psychosocial factors associated with IBS, (b) whether latent-class analysis can predict clusters within IBS, based on these factors, and (c) if sensory-motor testing provides mechanistic insight of patients characterized by the clusters.

## **Methods**

### **Study design and population**

**Data collection**

The Biobank recruited adult patients scheduled for appointments between 2009-2015.<sup>12, 13</sup> This study utilized responses from the first follow-up survey administered via mail 4 years post-enrollment. A total of 40,291 individuals (of 54,584 eligible) responded to the survey (74% response rate). Non-responders were mailed a follow-up packet at one and two months after first mailing. The questionnaire recorded demographic information; perceptions of general, mental, and social health; diagnoses of any mental, neurologic, rheumatologic, gastrointestinal (GI), or gynecologic disorders; dietary and lifestyle practices; and Rome III IBS. The data were collected under Biobank IRB #08-007049 and study IRB #21-008816. All participants provided informed consent.

**Identification of IBS cases and controls**

The Rome III criteria was used to define IBS cases and those with diagnosed or self-reported disorders that may mimic IBS symptoms were excluded. These are detailed in the **supplementary methods**. Gender and age-matched ( $\pm 5$  years) controls (3 per case) without any of the Rome III criteria were identified.

**Factors associated with IBS**

Variables were chosen to explore specific lifestyle/dietary behaviors (e.g., smoking, alcohol, high-fat diet, fruits, vegetables, regular and diet soft-drink consumption, and weekly leisure activity score (WLAS)) and co-morbid conditions (e.g., migraine and endometriosis). Furthermore, the six Rome III variables (abdominal pain frequency, abdominal pain that decreased with defecation, frequent bowel movements, loose stools, less frequent bowel movements, hard stools) were also included.

**Physiologic tests**

Electronic medical records were searched to abstract GI transit studies and/or anorectal manometry (ARM) with rectal sensitivity testing on the IBS subjects.

**Statistical analysis**

We performed latent class analysis (LCA) using the poLCA package<sup>14</sup> in RStudio (version 4.1.2 (2021-11-01) R Foundation for Statistical Computing, Vienna, Austria) to identify clusters of subjects within the IBS cohort.<sup>15</sup> LCA is a mixture model that assumes a population can be divided into mutually exclusive and exhaustive latent classes based on probability distributions of one or more underlying variables. Multiple solutions are tested for any given number of clusters and statistically evaluated to determine the best fit of the model and the optimal number of clusters<sup>16</sup>. Conditional logistic regression was used to determine factors significantly associated with the IBS status. These and other clinically relevant variables were included in the LCA. All observations were included in the LCA regardless of missing values in manifest variables. The proportion of individuals with missing data for each variable is presented in **Supplementary Table 1**. In order to analyze the data structure with the least number of classes, LCA was conducted from two (most parsimonious) to 11 clusters. Because the estimated models were non-nested, the Bayesian Information Criterion (BIC) model selection strategy was adopted, with the optimal model being the one with lowest BIC.

Subjects were grouped into "clusters" using the vector assignment of predicted class memberships. After determining the optimal number of latent clusters,  $z$ -scores for each variable were calculated, which represents the distance from the cohort mean for that variable. These were calculated by adjusting each variable's cluster mean to the cohort mean and standard deviation. Comparison of clusters with respect to characteristics and variables was done using analysis of variance followed by Tukey post-hoc testing.  $\chi^2$  was used for comparing the categorical variables.

## Results

### *Study cohort*

Of 40,291 respondents to the Biobank questionnaire, 5,005 subjects met Rome III criteria for IBS. After applying exclusion criteria, 4,021 IBS subjects (9.9% prevalence among survey responders) were eligible for analysis and were compared with 12,063 matched controls. The mean age of cases was 63.6 years (18–86 years), and 3,011 (75%) were female. The cohort consisted of 1,280 patients with IBS-D (32%), 1,021 with IBS-C (25%), 1,597 with IBS-M (40%) and 81 with IBS-U (2%). 1% IBS subjects could not be subtyped due to incomplete data.

*Factors associated with IBS status*

The IBS patients had greater perceived health impairment, greater psychiatric diagnoses, and comorbidities like fibromyalgia, endometriosis, migraines, and other neurological disorders. They were more likely to be current smokers, consume high-fat food, and soft-drinks but less likely to consume fruits, vegetable, and alcohol (**Supplementary Table 2**). The top ten variables with strongest association with IBS status (by smallest  $P$ -value) were identified. Any pain in the last seven days (OR=1.41) and perceived impairment overall health (OR=2.42) were found to be most significantly associated ( $P<0.001$  for both). These were followed by perceived impairment in social health, mental health, reported diagnosis of fibromyalgia, any rheumatologic disorder, any mental health disorder, anxiety, depression, and any neurologic disorder ( $P<0.001$  for all). The three IBS subtypes were compared (**Supplementary Table 3**).

*IBS clustering using latent class analysis*

In addition to the top ten significant variables identified, 16 additional variables were selected for their relevance to the clinical course of IBS. The variables that were included in model along with their ORs are listed in **Table 1**. The LCA included a total of 4,015 IBS subjects because six were lost to an update of their medical record number. The best LCA solution was achieved with seven clusters (BIC=210204.1) (**Supplementary Figure 1**). The seven clusters were distinguished by GI symptoms, lifestyle behaviors, and prevalence of extra-intestinal somatic and psychological comorbidities. **Figure 1** shows radar plots demonstrating the z-scores of each variable for each cluster.

Clusters with greater perceived impairment in overall health

**Cluster 1 (Moderate impairment in health):** This cluster comprised 17% (n=689) of the IBS cohort and had the lowest proportion of females (62%). The cluster most closely approximated the z-score means for the overall IBS cohort (**Table 2, Figure 1A**). These patients had moderate perceived impairment in overall and social health. They had slightly above average scores for any pain in the last seven days (mean 4.1, SD 2.1,  $P<0.05$  compared to last five clusters). This



cluster had much lower psychiatric comorbidities, migraines, and neurological disorders than the average for the IBS cohort. Vegetable, fruit, and alcohol consumption was lower while consumption of high-fat foods was higher. Lastly, WLAS of this cluster was below the IBS cohort mean.

**Cluster 2 (Severe impairment in health):** This cluster comprised 16% (n=665) of the IBS cohort and was characterized by severe perceived impairment in overall, mental, and social health. These patients also reported highest scores for any pain in last seven days (mean 5.8, SD 2.0,  $P<0.05$  compared to other six clusters) (**Figure 1A**). Moreover, anxiety, depression, any mental health disorder, endometriosis, fibromyalgia, and any rheumatologic disorder were more prevalent. This cluster had above average scores for regular soft drink consumption and highest smoking prevalence. Abdominal pain was also highest (mean 4.6, SD 1.1) and least likely to improve with defecation (mean 1.6, SD 1.1),  $P<0.05$  compared to other clusters for both. Vegetable, fruit, alcohol consumption, and WLAS were the lowest for this cluster.

#### Clusters with greater prevalence of psycho-neurological diagnoses

**Cluster 3 (Increased psychoneurological burden):** This cluster comprised 10% (n=428) of the IBS cohort and was defined by above-average scores for psychiatric diagnoses: anxiety, depression, or any other mental health disorder as well as migraine and any neurologic disorder (**Figure 1B**). This cluster contained a significantly higher proportion of women (91%,  $P<0.05$  compared to other six clusters) and was the youngest of all the clusters (58 years). The remaining variables were close to the IBS cohort means.

**Cluster 4 (Increased psychiatric burden):** This cluster was the largest by proportion (20%, n=793) and was defined by above-average scores for psychiatric diagnoses: anxiety, depression, or any other mental health disorder but below average scores for migraine and any neurologic disorder (**Figure 1B**).

**Cluster 5 (Increased neurological burden):** This cluster comprised 11% (n=451) of the cohort and had the second highest proportion of women (86%). This cluster had higher prevalence of migraine and any neurologic disorder but, below-average scores for psychiatric diagnoses (anxiety, depression, or any other mental health disorder) (**Figure 1B**).

### Clusters with predominant bowel dysfunction

**Cluster 6 (Diarrhea predominance):** Twelve percent (n=493) of the IBS cohort was classified into this cluster, characterized by a higher proportion of patients with loose and frequent bowel movements (**Figure 1C**). Abdominal pain was lower in this cluster than the IBS cohort mean (mean 3.9, SD 0.9 vs 4.2, 1.1,  $P<0.05$ ). However, abdominal pain improvement with defecation was more likely in this cluster (mean 2.3, SD 1.2,  $P<0.05$  compared to the prior five clusters). Seventy five percent (n=370) of individuals in this cluster met Rome III criteria of IBS-D. The cluster had above-average WLAS mean whereas, the scores for co-morbidities like endometriosis, migraine, any neurologic disorder, and fibromyalgia were very low in this cluster (**Table 2**).

**Cluster 7 (Constipation predominance):** The seventh cluster (12% of the cohort, n=487) was characterized by a higher proportion of patients with hard and a less frequent bowel movements (**Figure 1C**). Similar to the diarrhea cluster, abdominal pain was lower than the IBS cohort mean in this cluster (mean 3.8, SD 1.0 vs 4.2, 1.1,  $P<0.05$ ). Also, this cluster perceived a significantly lower impairment of overall, social, or mental health and a lower prevalence of neuropsychiatric diagnoses. Only 56% (n=274) of individuals in this cohort met Rome III criteria of IBS-C. These patients had the highest WLAS in relation to other clusters (mean 2.6, SD 0.7,  $P<0.05$  compared to other six clusters). The cluster had high scores for fruit, vegetable, and alcohol consumption. Similar to the diarrhea predominance cluster, the scores for co-morbidities were very low.

**Figure 2** provides a visual path to compare the cluster-specific symptom profiles and co-morbid conditions.

### *Distribution of IBS symptoms in the seven clusters*

We compared the Rome III IBS symptoms among the seven clusters (**Figure 3**). In clusters with moderate impairment and severe impairment in health, 37% and 38%, respectively, reported abdominal pain 1 day/week, while for the remaining five clusters, the most common response was 2-3 days/month. For the abdominal pain improvement with defecation variable, the first five clusters had 35-47% individuals responding as “sometimes” making it the most common response. However, for the diarrhea and constipation predominant clusters, ~33% individuals

responded as “most of the time”, making it the most common response. In diarrhea predominant cluster, frequent and loose stools variables was “most of the time” in 31% and 35%, respectively. For constipation predominant cluster, the most common response for infrequent and hard stools variables was “sometimes” (40% and 48%, respectively).

#### *Subset analysis of sensorimotor profiles of the 7 clusters*

Within the IBS cohort, 246 underwent ARM (6%), 217 gastric emptying (5.4%), 108 small bowel transit (2.7%), and 91 colonic transit studies (2.3%). The severe perceived impairment cluster was much more likely to undergo testing compared to the other 6 clusters (11% vs 3-6%). The ARM or the GI transit variables showed no significant difference between the clusters ( $P>0.05$  for all) (**Table 3**).

#### *Putative classification using Rome IV*

One of the key changes in the Rome IV criteria is threshold for frequency of abdominal pain (1 day/week vs 2-3 days/month with Rome III). Sixty percent ( $n=2,402$ ) would have met Rome IV pain criteria along with the bowel criteria. The mean age for this cohort was 63 years and 75% were female. Furthermore, the distribution of IBS subtypes was comparable to the Rome III cohort (32% were IBS-D; 24% IBS-C; 41% IBS-M and 3% IBS-U). Lastly, the distribution of Rome IV characterized patients in the seven identified clusters also remained broadly the same as the presented Rome III (cluster 1:19%, 2:20%, 3:10%, 4:19%, 5:11%, 6:11% and 7:9%).

## **Discussion**

Amongst 4,015 well-characterized IBS patients, seven unique latent clusters were identified. Two clusters (constituting 34% of patients) had greater perceived impairment in health, moderate in one and severe in the other. These patients had greater pain but not significant bowel dysfunction. Three clusters (42% of patients) were predominated by neuropsychological complaints but not enriched for pain, bowel dysfunction or perceived impairment in health. Last two clusters (24% of patients) showed enrichment for bowel dysfunction with lower than average perceived health impairment and other comorbidities. These findings provide novel

insight into the IBS symptom experience beyond the subtyping based on bowel habits or IBS-SSS.

There is a knowledge gap in the precise understanding of GI symptom variability, perceived health impairment, and how these factors relate to specific comorbidities, diet, and exercise patterns in IBS patients. Our current paradigms for clinical trial design capture a limited spectrum of disease experience. The clusters identified demonstrate features that help understand disease heterogeneity and optimize clinical care for IBS. The first two clusters have greatest abdominal pain, highest perceived health impairment, and least bowel dysfunction; however, a diverse spectrum for the remaining variables. The severe impairment cluster had significantly more psychiatric comorbidities, and pain beyond abdominal pain (migraines, fibromyalgia and “any pain”). It also reported the least improvement in pain with defecation. Thus, it likely represents the tertiary care group of patients with chronic and disabling pain with an overlap of psychiatric comorbidities. The next three clusters pivoted towards variable representation of five diagnoses (anxiety, depression, any mental disorder, migraines, and any neurological disorder), while having similar bowel symptomatology. Despite the high prevalence of psychiatric comorbidities, the perceived impairment of health and pain was less than the IBS cohort mean. The last two clusters were clearly dominated by bowel habits (diarrhea or constipation). Interestingly, the perceived health impairment, psychiatric, and other comorbidities were significantly lower in these two clusters. Both clusters had higher alcohol consumption and exercise frequency than the rest, particularly the constipation cluster. The constipation cluster had the highest fruit and vegetable consumption of all clusters and one of the lowest consumptions of high-fat foods and soda. It is possible that patients in this cluster modified their diet to help alleviate the constipation symptoms. These bowel dysfunction clusters likely reflect the primary/secondary care IBS patients with less pain, lower comorbidities, and greater engagement with dietary changes and exercise.

Some previous studies have investigated methods of subgrouping IBS patients beyond the stool pattern.<sup>17-19</sup> Using 172 tertiary care IBS patients Polster et al., identified six subgroups based on GI, psychiatric, somatic, and cardiovascular symptoms.<sup>17</sup> Another study by the same group using psychiatric and GI symptoms, identified seven clusters for Rome III (n=637), and five for Rome IV IBS (n=341).<sup>18</sup> They found that subgroups with a high prevalence of extra-intestinal and

psychological symptoms used healthcare and medications frequently. This was also observed in our second cluster, which had the largest proportion of patients undergoing physiological testing. In a recent larger study by Black et al., self-identified IBS subjects from the community also fitted into a seven-cluster model.<sup>19</sup> These were diarrhea/urgency or constipation/bloating each with low and high psychological burden, less GI symptoms with low or high psychological burden and more GI symptoms along with higher psychological comorbidity. These studies have limitations of smaller cohorts of patients, lack of assessment of important medical comorbidities (neurological, gynecological, rheumatological), diet, and lifestyle variables which are known to affect the clinical course of IBS. For example, our study reveals two clusters uniquely differentiated by having migraines and neurological diagnoses, one with and one without added psychiatric diagnoses. The patients in these clusters are best suited for neurological work-up and potentially benefit from treatments suited for abdominal migraines.<sup>20, 21</sup> Our study also reveals that contrary to general belief, only 46% of IBS patients have a high burden of psychiatric diagnoses and only 16% had significant abdominal and overlapping pain. These groups may need targeted interventions with neuromodulators or psychotherapy. The bowel dysfunction predominant clusters (24% patients) have less abdominal pain and more likely to have pain improvement with defecation compared to the other five clusters, which questions the paradigm suggested by the FDA in which a target drug must meet the pain and bowel dysfunction criteria. The variability in dietary responses between the clusters was less than the other variables; however, the severe health impairment cluster endorsed most smoking, use of non-diet soft drinks and the least alcohol. It is plausible that high sugar consumption may play a role in exacerbating their symptoms and they may be smoking to alleviate their symptoms. In contrast, the bowel dysfunction clusters have the most alcohol and vegetable consumption, particularly in the constipation cluster. These findings suggest that integration of GI care with neurological, psychiatric, dietary and lifestyle care in specific subsets of IBS patients can plausibly result in better patient outcomes.

Our study has some limitations. Compared to the geographic populations, Biobank participants tend to be better educated, have lower BMI and were more likely to self-report their race as white.<sup>12, 13</sup> Our IBS cohort is likely to represent the general population of IBS patients because subjects were not chosen for any specific disease or condition or based on GI complaints. The

population enrolled was older than the typical large IBS cohorts and IBS was not confirmed by a clinician. The cross-sectional nature of the study precludes determination if the associations studied influenced GI symptoms or vice versa. Another limitation is the questionnaire did not capture complete dietary information, which results in limited interpretation of dietary factors. However, it captured exercise data well, allowing us to calculate the validated WLAS.<sup>22</sup> Lastly, the absence of IBS-SSS data limits contrasting these clusters against established score for symptom severity.

In conclusion, we have demonstrated that a combination of GI symptoms, medical, and psychiatric comorbidities, dietary, and lifestyle factors allows clustering of IBS patients into distinct subgroups. These findings provide important insights into the heterogeneity of IBS and imply that significant bowel dysfunction, the primary input for clinical care and research, may only be present in a subset and may not cause remarkable impairment in QoL. Secondly, these findings demonstrate the presence of highly specific subgroups, which have comorbidities like migraines or fibromyalgia, raising questions about overlapping mechanisms that may underlie these disorders in those clusters. Finally, the findings question the paradigm of expecting improved QoL for IBS patients by simply targeting bowel dysfunction or abdominal pain. Future prospective studies should validate these findings in larger cohorts and investigate the possibility of directing clinical trials towards the unique IBS subsets characterized by this clustering. Ultimately, these findings emphasize the importance of recognizing clinical profiles that go beyond GI symptoms in clinical decision-making and personalizing treatment for IBS patients.<sup>23</sup>

## Figure legends

**Figure 1. Radar plots showing seven latent clusters (A)** Clusters with moderate or severe impairment in perceived health **(B)** Clusters with increased psychoneurological, psychological or neurological burden **(C)** Clusters with diarrhea or constipation predominance. Dark circles represent adjusted IBS cohort mean.

**Figure 2. Heatmap contrasting the z-scores for various variables in the seven latent clusters.** The clusters represent variable influence of perceived impairment in health, psychiatric diagnoses, bowel symptoms, dietary and lifestyle behaviors, and comorbidities. Z-scores -0.5 to 1.5.

**Figure 3. Distribution of responses on the Rome III questions for the seven latent clusters (A)** Abdominal pain frequency **(B)** Improvement of abdominal pain with defecation **(C)** Frequent bowel movements **(D)** Loose stools **(E)** Less frequent bowel movements **(F)** Hard stools

## Table legends

**Table 1.** Selected variables used in latent class analysis

**Table 2.** Comparison of clusters with respect to characteristics and variables input using analysis of variance (ANOVA) followed up by pairwise comparisons of all clusters using post hoc t tests and Tukey correction for multiple testing

**Table 3.** Gastrointestinal transit and anorectal manometry (ARM) assessments in the seven IBS clusters

**References**

1. Mearin F, Lacy BE, Chang L, et al. Bowel Disorders. *Gastroenterology* 2016.
2. Lackner JM, Gudleski GD, Dimuro J, et al. Psychosocial predictors of self-reported fatigue in patients with moderate to severe irritable bowel syndrome. *Behav Res Ther* 2013;51:323-31.
3. Schmulson MJ, Drossman DA. What Is New in Rome IV. *J Neurogastroenterol Motil* 2017;23:151-163.
4. Francis CY, Morris J, Whorwell PJ. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progress. *Aliment Pharmacol Ther* 1997;11:395-402.
5. Lackner JM, Jaccard J, Baum C. Multidomain patient-reported outcomes of irritable bowel syndrome: exploring person-centered perspectives to better understand symptom severity scores. *Value Health* 2013;16:97-103.
6. Shiha MG, Aziz I. Review article: Physical and psychological comorbidities associated with irritable bowel syndrome. *Aliment Pharmacol Ther* 2021;54 Suppl 1:S12-S23.
7. Grover M, Kolla BP, Pamarthy R, et al. Psychological, physical, and sleep comorbidities and functional impairment in irritable bowel syndrome: Results from a national survey of U.S. adults. *PLoS One* 2021;16:e0245323.
8. Chey WD, Hashash JG, Manning L, et al. AGA Clinical Practice Update on the Role of Diet in Irritable Bowel Syndrome: Expert Review. *Gastroenterology* 2022.
9. Lenhart A, Dong T, Joshi S, et al. Effect of Exclusion Diets on Symptom Severity and the Gut Microbiota in Patients With Irritable Bowel Syndrome. *Clin Gastroenterol Hepatol* 2022;20:e465-e483.
10. Hajishafiee M, Keshteli AH, Saneei P, et al. Healthy lifestyle score and irritable bowel syndrome: A cross-sectional study in adults. *Neurogastroenterol Motil* 2020;32:e13793.
11. Sadeghian M, Sadeghi O, Hassanzadeh Keshteli A, et al. Physical activity in relation to irritable bowel syndrome among Iranian adults. *PLoS One* 2018;13:e0205806.
12. Olson JE, Ryu E, Johnson KJ, et al. The Mayo Clinic Biobank: a building block for individualized medicine. *Mayo Clin Proc* 2013;88:952-62.



13. Olson JE, Ryu E, Hathcock MA, et al. Characteristics and utilisation of the Mayo Clinic Biobank, a clinic-based prospective collection in the USA: cohort profile. *BMJ Open* 2019;9:e032707.
14. Linzer DA, Lewis JB. *poLCA: An R Package for Polytomous Variable Latent Class Analysis*. *Journal of Statistical Software* 2011;42:1 - 29.
15. Team R. *RStudio: Integrated Development Environment for R*. RStudio, Boston, MA, 2021.
16. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. *Prev Sci* 2013;14:157-68.
17. Polster A, Van Oudenhove L, Jones M, et al. Mixture model analysis identifies irritable bowel syndrome subgroups characterised by specific profiles of gastrointestinal, extraintestinal somatic and psychological symptoms. *Aliment Pharmacol Ther* 2017;46:529-539.
18. Polster AV, Palsson OS, Törnblom H, et al. Subgroups of IBS patients are characterized by specific, reproducible profiles of GI and non-GI symptoms and report differences in healthcare utilization: A population-based study. *Neurogastroenterol Motil* 2019;31:e13483.
19. Black CJ, Yiannakou Y, Guthrie EA, et al. A Novel Method to Classify and Subgroup Patients With IBS Based on Gastrointestinal Symptoms and Psychological Profiles. *Am J Gastroenterol* 2021;116:372-381.
20. Welander NZ, Olivo G, Pisanu C, et al. Migraine and gastrointestinal disorders in middle and old age: A UK Biobank study. *Brain Behav* 2021;11:e2291.
21. Fila M, Chojnacki J, Pawłowska E, et al. Kynurenine Pathway of Tryptophan Metabolism in Migraine and Functional Gastrointestinal Disorders. *Int J Mol Sci* 2021;22.
22. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. *Can J Appl Sport Sci* 1985;10:141-6.
23. Simrén M, Törnblom H, Palsson OS, et al. Management of the multiple symptoms of irritable bowel syndrome. *Lancet Gastroenterol Hepatol* 2017;2:112-122.