Machine learning-based phenogroups and prediction model in patients with functional gastrointestinal disorders to reveal distinct disease subsets associated with gas production

Lingling Zhu[1,2], Shuo Xu[3], Huaizhu Guo[1,2], Siqi Lu[1,2], Jiaqi Gao[1,2], Nan Hu[1,2], Chen Chen[1,2], Zuojing Liu[1,2], Xiaolin Ji[1,2], Kun Wang[1,2], Liping Duan[1,2]

[1]Department of Gastroenterology, Peking University Third Hospital, Beijing 100191, China;
[2]Beijing Key Laboratory for Helicobacter Pylori Infection and Upper Gastrointestinal Diseases, Beijing 100191, China;
[3]Beijing Aerospace Wanyuan Science Technology Co., Ltd., China Academy of Launch Vehicle Technology, Fengtai, Beijing 100176, China

ABSTRACT

Background and Objectives: Symptom-based subtyping for functional gastrointestinal disorders (FGIDs) has limited value in identifying underlying mechanisms and guiding therapeutic strategies. Small intestinal dysbiosis is implicated in the development of FGIDs. We tested if machine learning (ML) algorithms utilizing both gastrointestinal (GI) symptom characteristics and lactulose breath tests could provide distinct clusters. **Materials and Methods:** This was a prospective cohort study. We performed lactulose hydrogen methane breath tests and hydrogen sulfide breath tests in 508 patients with GI symptoms. An unsupervised ML algorithm was used to categorize subjects by integrating GI symptoms and breath gas characteristics. Generalized Estimating Equation (GEE) models were used to examine the longitudinal associations between cluster patterns and breath gas time profiles. An ML-based prediction model for identifying excessive gas production in FGIDs patients was developed and internal validation was performed. **Results:** FGIDs were confirmed in 300 patients. K-means clustering identified 4 distinct clusters. Cluster 2, 3, and 4 showed enrichments for abdominal distention and diarrhea with a high proportion of excessive gas production, whereas Cluster 1 was characterized by moderate lower abdominal discomforts with the most psychological complaints and the lowest proportion of excessive gas production. GEE models showed that breath gas concentrations varied among different clusters over time. We further sought to develop an ML-based prediction model to determine excessive gas production. The model exhibited good predictive capabilities. **Conclusion:** ML-based phenogroups and prediction model approaches could provide distinct FGIDs subsets and efficiently determine FGIDs subsets with greater gas production, thereby facilitating clinical decision-making and guiding treatment.

Key words: functional gastrointestinal disorders, small intestinal dysbiosis, gas production, machine learning

Address for Correspondence:
Liping Duan, Department of Gastroenterology, Peking University Third Hospital, No. 49 North Garden Road, Haidian District 100191, Beijing, China. E-mail: duanlp@bjmu.edu.cn

Access this article online

Website: www.intern-med.com

DOI: 10.2478/jtim-2024-0009

INTRODUCTION

Functional gastrointestinal disorders (FGIDs) are highly prevalent and remain the most common cause of gastrointestinal (GI) complaints.[1,2] However, FGIDs often overlap, and the limited link between the current classifications and putative pathophysiology remains poorly translated into efficacious treatment options.[3]

Machine learning (ML) methods have been employed to classify individuals with FGIDs to discover novel subcategories.[4] Supervised ML utilizes iterative processes to "learn" from a well-labeled training set for high diagnostic accuracy.[5] Whereas unsupervised methods classify patients into

clusters based on various attributes to examine how these attributes may be connected to treatment responses.[6]

Several studies have applied unsupervised ML approaches utilizing GI and extra-GI symptoms to discover new phenogroups.[4,7,8] However, these studies have primarily utilized a single set of clinical symptoms. Incorporating clinical manifestations and biomarkers that delineate the underlying mechanisms is imperative for distinguishing FGIDs from other conditions with similar or overlapping symptoms and guiding treatments.

Imbalances in the small intestine microbiota are closely linked to FGIDs.[9–12] Small intestinal bacterial overgrowth is the most widely recognized type of dysbiosis involving the small intestine, resulting in symptoms caused by carbohydrate fermentation, usually confirmed through a positive breath test.[13] Association studies between breath gas characteristics and symptoms following sugar provocation tests in FGIDs have shown that acute GI symptoms following sugar ingestion are associated with breath gas levels.[14,15] However, these examinations were conducted based on acute provocation tests, and abnormal gas production resulting from small intestinal dysbiosis as a persistent pathogenic element in FGIDs requires further investigation.

Consequently, mechanisms underlying the association among GI symptoms, gas production, and small intestinal dysbiosis remain unclear. We hypothesized that ML algorithms utilizing both GI characteristics and lactulose breath tests could provide more well-separated clusters, and that identifying excessive gas production in FGIDs patients may allow a more targeted antibacterial or diet-based approach to treatment.

MATERIALS AND METHODS

Patients and data sources

This was a prospective cohort study. This study was approved by the ethics committee of Peking University Third Hospital (2021-132-01) and was conducted in accordance with the ethical guidelines outlined in the Declaration of Helsinki. Informed consent was obtained from the participants on an opt-out basis given the noninvasive nature of the study.

All patients with GI symptoms referred to the Outpatient Department of Gastroenterology at Peking University Third Hospital from December 2021 to November 2023 were eligible for inclusion in this study. FGIDs were established according to Rome IV criteria. The exclusion criteria were described in Supplementary Method 1. All medical history-taking was conducted by skilled gastroenterologists (Supplementary Method 2). Standardized questionnaires were used to classify GI symptoms into FGIDs according to the Rome IV criteria, and contained questionnaires assessing the severity of anxiety, depression, and sleep quality (Supplementary Table 1). The intensity of GI symptoms was assessed using a 4-point Likert scale ranging from 0 (none) to 3 (severe).

Breath tests

All patients underwent standardized lactulose hydrogen methane breath tests (LHMBT) and hydrogen sulfide breath tests performed by the same technician (Supplementary Method 3). Excessive gas production was defined as a positive LHMBT according to the North American Consensus, with an increase of hydrogen (H_2) ≥20 parts per million (ppm) or methane (CH_4) levels reaching ≥10 ppm within 90 min.[16]

Unsupervised machine learning

Forty-three baseline variables, including clinical characteristics, H_2 peak, excessive H_2 production, CH_4 peak, excessive CH_4 production, hydrogen sulfide (H_2S) peak, time to H_2S peak, and the area under the curve (AUC) of H_2S, were identified and utilized as input for the unsupervised ML algorithm (Supplementary Method 4). The dataset was supplemented using the multiple imputation method to fill in missing data representing less than 40% of the total data. Twenty complete datasets were imputed. The variation analyses of the 20 imputation datasets were provided (Supplementary Table 2).

Individuals were grouped and different categories of FGIDs were identified based on their phenotypic characteristics using the K-means algorithm. Each imputed dataset was individually processed using the prescribed steps for the clustering algorithm.[17] The clustering algorithm was executed with a range of predefined group numbers, starting from 2 and increasing to 9. The elbow method was used to determine the optimal number of clusters (Supplementary Figure 1). Principal components analysis (PCA) was used to visualize and investigate the clustering of our data (Supplementary Figure 2).

Supervised machine learning

Variables associated with excessive gas production were included as candidate variables based on a literature review, expert opinions, and the results of univariate analyses. Multicollinearity between variables was analyzed using Spearman correlation analysis and collinearity diagnostics (Supplementary Method 5). Least absolute shrinkage and selection operator (Lasso) regression was employed on all 20 imputed datasets to assess the importance of potential variables. To create a robust and dependable model, we chose the top seven characteristics for constructing the

Table 1: Demographic and clinical characteristics of patients with functional gastrointestinal disorders grouped by cluster

Characteristics	Overall (*n* = 300)	Cluster 1 (*n* = 59)	Cluster 2 (*n* = 139)	Cluster 3 (*n* = 64)	Cluster 4 (*n* = 38)	*P* value
Sex, female (%)	153 (51.0)	31 (52.5)	56 (40.3)	41 (64.1)	25 (65.8)	0.003
Age (years, median [IQR])	29.0 (25.0, 37.0)	29.0 (25.0, 38.0)	31.0 (25.0, 39.0)	25.5 (23.0, 33.5)	29.0 (25.3, 32.8)	0.011
Age						0.135
≤ 30 years (%)	171 (57.0)	32 (54.2)	68 (48.9)	46 (71.9)	25 (65.8)	
31–40 years (%)	83 (27.7)	18 (30.5)	45 (32.4)	10 (15.6)	10 (26.3)	
41–50 years (%)	28 (9.3)	4 (6.8)	16 (11.5)	6 (9.4)	2 (5.3)	
≥ 51 years (%)	18 (6.0)	5 (8.5)	10 (7.2)	2 (3.1)	1 (2.6)	
BMI (kg/m^2, median [IQR])	22.0 (20.0, 24.8)	21.5 (20.0, 24.8)	23.0 (20.6, 25.5)	21.2 (19.1, 23.1)	21.6 (20.2, 22.9)	0.010
BMI grade						0.212
<18.5 kg/m^2 (%)	32 (10.7)	6 (10.2)	14 (10.1)	9 (14.1)	3 (7.9)	
18.5–23.9 kg/m^2 (%)	177 (59.0)	36 (61.0)	71 (51.1)	43 (67.2)	27 (71.1)	
24.0–27.9 kg/m^2 (%)	74 (24.7)	14 (23.7)	45 (32.4)	9 (14.1)	6 (15.8)	
≥28 kg/m^2 (%)	17 (5.7)	3 (5.1)	9 (6.5)	3 (4.7)	2 (5.3)	
Duration (months, median [IQR])	44.0 (18.5, 87.0)	44.5 (21.8, 95.5)	53.5 (20.0, 116.5)	28.0 (15.0, 55.0)	52.0 (20.0, 99.0)	0.007
Smoking history (%)	24 (8.1)	5 (8.6)	14 (10.3)	3 (4.8)	2 (5.3)	0.523
Drinking history (%)	43 (14.6)	10 (17.2)	22 (16.2)	6 (9.5)	5 (13.2)	0.580
History of allergy (%)	48 (17.6)	14 (29.2)	19 (15.0)	8 (12.7)	7 (20.0)	0.101
Allergic diseases (%)	116 (42.5)	21 (43.8)	53 (41.7)	25 (39.7)	17 (48.6)	0.851
Heart disease (%)	4 (1.4)	1 (1.7)	2 (1.5)	1 (1.6)	0 (0.0)	0.893
Respiratory disease (%)	6 (2.0)	0 (0.0)	4 (2.9)	2 (3.1)	0 (0.0)	0.402
Endocrine disorders (%)	14 (4.7)	4 (6.9)	4 (2.9)	2 (3.1)	4 (10.5)	0.186
Hematologic disorders (%)	2 (0.7)	1 (1.7)	0 (0.0)	1 (1.6)	0 (0.0)	0.407
Autoimmune disease (%)	3 (1.0)	1 (1.7)	0 (0.0)	2 (3.1)	0 (0.0)	0.176
Gastroenteritis (%)	111 (46.4)	31 (58.5)	50 (45.1)	18 (38.3)	12 (42.9)	0.205
Infection history of *H. pylori* (%)	55 (21.8)	22 (39.3)	21 (20.0)	7 (11.5)	5 (16.7)	0.002
PPI prescription (%)	83 (38.4)	23 (46.0)	31 (32.3)	22 (46.8)	7 (30.4)	0.191
Antibiotic prescription (%)	113 (58.3)	30 (61.2)	57 (58.8)	15 (57.7)	11 (50.0)	0.848
Infection history of COVID-19 (%)	115 (63.9)	40 (93.0)	40 (45.5)	18 (85.7)	17 (60.7)	<0.001
Oral ulcers (%)	44 (21.9)	22 (41.5)	15 (22.7)	3 (5.2)	4 (16.7)	<0.001
Type of BSF (%)						<0.001
Constipation	18 (6.0)	3 (5.1)	8 (5.8)	5 (7.8)	2 (5.3)	
Diarrhea	134 (44.7)	21 (35.6)	77 (55.4)	16 (25.0)	20 (52.6)	
Mixed	3 (1.0)	3 (5.1)	0 (0.0)	0 (0.0)	0 (0.0)	
Normal	145 (48.3)	32 (54.2)	54 (38.9)	43 (67.2)	16 (42.1)	
IBS (%)	128 (42.7)	22 (37.3)	86 (61.9)	2 (3.1)	18 (47.4)	<0.001
IBS subtype (%)						<0.001
IBS-C	3 (1.0)	0 (0.0)	2 (1.4)	0 (0.0)	1 (2.6)	
IBS-D	73 (24.3)	11 (18.6)	53 (38.1)	0 (0.0)	9 (23.7)	
IBS-M	1 (0.3)	1 (1.7)	0 (0.0)	0 (0.0)	0 (0.0)	
IBS-U	51 (17.0)	10 (17.0)	31 (22.3)	2 (3.1)	8 (21.1)	
FD (%)	135 (45.0)	42 (71.2)	15 (10.8)	64 (100.0)	14 (36.8)	<0.001
EPS (%)	62 (20.7)	19 (32.2)	3 (2.2)	31 (48.4)	9 (23.7)	<0.001
PDS (%)	122 (40.7)	42 (71.2)	15 (10.8)	52 (81.3)	13 (34.2)	<0.001
The overlap of EPS and PDS (%)	49 (16.3)	19 (32.2)	3 (2.2)	19 (29.7)	8 (21.1)	<0.001
Other FGIDs (%)						0.222
U-FBD	8 (2.7)	0 (0.0)	4 (2.9)	2 (3.1)	2 (5.3)	
FC	12 (4.0)	1 (1.7)	7 (5.0)	4 (6.3)	0 (0.0)	

(*To be continued*)

(*Continued*)

Functional diarrhea	68 (22.7)	14 (23.7)	25 (18.0)	17 (26.6)	12 (31.6)	
Functional bloating	16 (5.3)	5 (8.5)	10 (7.2)	0 (0.0)	1 (2.6)	
Belching (%)	72 (24.0)	35 (59.3)	17 (12.2)	8 (12.5)	12 (31.6)	<0.001
Overlapping of FGID syndromes	115 (38.3)	47 (79.7)	23 (16.6)	29 (45.3)	16 (42.1)	<0.001
The count of overlap syndromes						<0.001
2	92 (30.7)	35 (59.3)	21 (15.1)	25 (39.1)	11 (29.0)	
3	23 (7.7)	12 (20.3)	2 (1.4)	4 (6.3)	5 (13.2)	
Nausea (median [IQR])	0.0 (0.0, 1.0)	0.0 (0.0, 1.0)	0.0 (0.0, 0.0)	1.0 (0.0, 1.0)	0.0 (0.0, 1.0)	<0.001
Upper abdominal pain (median [IQR])	0.0 (0.0, 1.0)	1.0 (0.0, 1.0)	0.0 (0.0, 0.0)	1.0 (1.0, 2.0)	0.0 (0.0, 1.0)	<0.001
Stomach fullness (median [IQR])	1.0 (0.0, 2.0)	2.0 (1.0, 3.0)	0.0 (0.0, 0.0)	1.0 (1.0, 2.0)	0.0 (0.0, 1.0)	<0.001
Upper abdominal discomfort (median [IQR])	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	1.0 (0.0, 2.0)	0.0 (0.0, 0.0)	<0.001
Bloating (median [IQR])	0.0 (0.0, 1.0)	1.0 (1.0, 2.0)	0.0 (0.0, 0.0)	1.0 (1.0, 2.0)	0.0 (0.0, 1.0)	<0.001
Vomiting (median [IQR])	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.815
Early satiety (median [IQR])	0.0 (0.0, 1.0)	0.0 (0.0, 1.0)	0.0 (0.0, 0.0)	1.0 (0.0, 1.0)	0.0 (0.0, 0.0)	<0.001
Fullness after meals (median [IQR])	1.00 (0.0, 2.0)	2.0 (1.0, 2.0)	0.0 (0.0, 1.0)	2.0 (1.0, 2.0)	0.0 (0.0, 1.8)	<0.001
Abdominal pain (median [IQR])	1.0 (0.0, 1.0)	1.0 (0.0, 2.0)	1.0 (0.0, 1.0)	0.0 (0.0, 0.0)	0.5 (0.0, 1.0)	<0.001
Heartburn (median [IQR])	0.0 (0.0, 1.0)	0.0 (0.0, 1.0)	0.0 (0.0, 0.0)	0.0 (0.0, 1.0)	0.0 (0.0, 1.0)	0.004
Reflux (median [IQR])	0.0 (0.0, 1.0)	1.0 (0.0, 1.0)	0.0 (0.0, 1.0)	0.0 (0.0, 1.0)	0.0 (0.0, 1.0)	0.005
Borborygmus (median [IQR])	1.0 (0.0, 1.0)	1.0 (1.0, 2.0)	1.0 (0.0, 1.0)	0.0 (0.0, 0.0)	1.0 (1.0, 1.8)	<0.001
Eructation (median [IQR])	1.0 (0.0, 2.0)	2.0 (1.0, 2.5)	1.0 (0.0, 1.0)	1.0 (0.0, 1.0)	1.0 (1.0, 2.0)	<0.001
Abdominal distention (median [IQR])	1.0 (0.0, 2.0)	2.0 (1.5, 3.0)	1.0 (0.0, 2.0)	0.0 (0.0, 0.0)	1.0 (0.0, 2.0)	<0.001
Increased flatus (median [IQR])	1.0 (0.0, 2.0)	2.0 (1.0, 2.0)	1.0 (1.0, 2.0)	0.0 (0.0, 0.0)	2.0 (1.0, 2.0)	<0.001
Increased defecation (median [IQR])	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	1.0 (1.0, 2.0)	0.0 (0.0, 0.0)	1.0 (1.0, 2.0)	<0.001
Reduced defecation (median [IQR])	0.0 (0.0, 1.0)	1.0 (0.0, 2.0)	0.0 (0.0, 0.5)	0.0 (0.0, 0.0)	0.0 (0.0, 1.0)	<0.001
Loose stools (median [IQR])	1.0 (0.0, 2.0)	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)	0.0 (0.0, 1.0)	1.0 (0.0, 2.0)	<0.001
Hard stools (median [IQR])	0.0 (0.0, 1.0)	1.0 (0.0, 1.0)	0.0 (0.0, 1.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.8)	<0.001
Urgent need for defecation (median [IQR])	1.0 (0.0, 2.0)	1.0 (0.0, 1.0)	1.0 (0.0, 2.0)	0.0 (0.0, 0.0)	1.0 (0.0, 2.0)	<0.001
Feeling of incomplete evacuation (median [IQR])	1.0 (0.0, 1.3)	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)	0.0 (0.0, 0.0)	1.0 (0.0, 1.0)	<0.001
Anxiety (%)	79 (38.5)	33 (62.3)	23 (33.8)	18 (28.6)	5 (23.8)	<0.001
HAMA score (median [IQR])	6.0 (3.0, 8.0)	8.0 (5.0, 12.0)	5.0 (3.0, 8.0)	5.0 (2.5, 7.0)	4.0 (2.0, 6.0)	0.001
HAMA grade (%)						0.006
Mild	65 (31.7)	24 (45.3)	21 (30.9)	16 (25.4)	4 (19.1)	
Moderate	10 (4.9)	7 (13.2)	2 (2.9)	1 (1.6)	0 (0.0)	
Severe	1 (0.5)	1 (1.9)	0 (0.0)	0 (0.0)	0 (0.0)	
Very severe	3 (1.5)	1 (1.9)	0 (0.0)	1 (1.6)	1 (4.8)	
Depression (%)	52 (25.4)	25 (47.2)	9 (13.2)	13 (20.6)	5 (23.8)	<0.001
HAMD score (median [IQR])	4.0 (2.0, 7.0)	6.0 (3.0, 10.0)	3.0 (2.0, 5.0)	4.0 (1.5, 6.0)	3.0 (2.0, 5.0)	0.001
HAMD grade (%)						0.003
Mild	50 (24.4)	24 (45.3)	9 (13.2)	12 (19.1)	5 (23.8)	
Moderate	2 (1.0)	1 (1.9)	0 (0.0)	1 (1.6)	0 (0.0)	
Decreased sleep quality (%)	137 (62.8)	47 (81.0)	44 (57.9)	33 (52.4)	13 (61.9)	0.007
PSQI (median [IQR])	6.00 (5.0, 8.0)	8.0 (6.0, 11.0)	6.0 (4.0, 8.0)	6.0 (5.0, 8.0)	7.0 (5.0, 7.0)	<0.001

(*To be continued*)

(*Continued*)

PSQI grade (%)						0.001
Mild	110 (50.5)	30 (51.7)	39 (51.3)	28 (44.4)	13 (61.9)	
Moderate	23 (10.6)	14 (24.1)	5 (6.6)	4 (6.4)	0 (0.0)	
Severe	4 (1.8)	3 (5.2)	0 (0.0)	1 (1.6)	0 (0.0)	
Positive LHMBT (%)	183 (61.0)	21 (35.6)	91 (65.5)	33 (51.6)	38 (100.0)	<0.001
Positive HBT (%)	183 (61.0)	21 (35.6)	91 (65.5)	33 (51.6)	38 (100.0)	<0.001
Positive MBT (%)	39 (13.0)	1 (1.7)	0 (0.0)	0 (0.0)	38 (100.0)	<0.001

[#]Data are presented as median (IQR) for continuous variables and number (frequency, %) for categorical variables. FGIDs: functional gastrointestinal disorders, IQR: interquartile ranges, BMI: Body Mass Index, PPI: proton pump inhibitor, COVID-19: coronavirus disease 2019, BSF: Bristol stool form, IBS: irritable bowel syndrome, IBS-C: IBS with predominant constipation, IBS-D: IBS with predominant diarrhea, IBS-M: IBS with mixed bowel habits, IBS-U: unclassified IBS, FD: functional dyspepsia, EPS: epigastric pain syndrome, PDS: postprandial distress syndrome, U-FBD: unspecified functional bowel disorders, FC: functional constipation, HAMA: Hamilton Anxiety Scale, HAMD: Hamilton Depression Scale, PSQI: Pittsburgh sleep quality index, LHMBT: lactulose hydrogen methane breath tests, HBT: hydrogen methane breath tests, MBT: methane breath tests.

Table 2: Gas results for patients with functional gastrointestinal disorders according to clusters during lactulose breath testing

Variables[a]	Cluster 1	Cluster 2	Cluster 3	Cluster 4	*P* value[b]
H_2-peak, PPM	26.0 (16.0, 50.0)	41.0 (24.0, 60.0)	35.5 (20.5, 54.0)	89.5 (77.0, 107.0)	< 0.001***
H_2-time to peak, min	90.0 (45.0, 90.0)	90.0 (75.0, 90.0)	90.0 (45.0, 90.0)	90.0 (75.0, 90.0)	0.405
H_2-AUC	21.4 (15.4, 39.4)	30.8 (17.0, 43.4)	26.8 (18.2, 39.0)	63.6 (51.8, 78.8)	< 0.001***
CH_4-peak, PPM	1.0 (1.0, 3.0)	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)	19.0 (14.0, 27.3)	< 0.001***
CH_4-time-to-peak, min	0.0 (0.0, 45.0)	0.0 (0.0, 30.0)	0.0 (0.0, 0.0)	90.0 (75.0, 90.0)	< 0.001***
CH_4-AUC	1.5 (1.5, 2.8)	1.5 (1.5, 1.8)	1.5 (1.5, 1.8)	9.5 (5.6, 15.1)	< 0.001***
H_2S-peak, PPB	87.0 (58.0, 135.0)	94.0 (56.0, 121.0)	65.5 (34.3, 114.0)	148.0 (120.3, 177.3)	< 0.001***
H_2S-time-to-peak, min	90.0 (60.0, 90.0)	90.0 (60.0, 90.0)	75.0 (33.3, 90.0)	90.0 (75.0, 90.0)	0.021*
H_2S-AUC	90.5 (54.6, 138.6)	94.0 (57.5, 127.4)	64.5 (36.5, 105.8)	128.2 (79.7, 165.2)	< 0.001***

[a]Data are presented as median (IQR) for continuous variables. [b]Statistical significance after Bonferroni adjustment for multiple comparisons. *$P < 0.05$, ***$P < 0.001$. H_2: hydrogen, PPM: parts per million, AUC: area under the curve, CH_4: methane, H_2S: hydrogen sulfide, PPB: parts per billion, IQR: interquartile ranges.

model based on a majority vote.

Internal validation was performed using a bootstrap method. Six classical ML models, namely logistic regression (LR), linear discriminant analysis (LDA), k-nearest neighbor (KNN), naive Bayes (NB), support vector machine (SVM), and random forest (RF), were developed to predict excessive gas production. The performance of the model was evaluated by examining the receiver operating characteristic (ROC) and calibration curves. Furthermore, we used decision curve analysis (DCA) to visually evaluate the net benefits of each model. The SHapley Additive exPlanations (SHAP) method was employed to visualize the importance of characteristics selected in the model. A dynamic nomogram was constructed for clinical application.

Statistical analysis

Categorical variables were presented as percentages and were analyzed using the Chi-squared test. Variables displaying a non-normal distribution were presented as medians along with their interquartile ranges and analyzed using the Mann-Whitney U-test. Multiple comparisons of breath gas features were analyzed using the Kruskal-Wallis test with Bonferroni post hoc correction, and repeated measurements of gas concentrations were analyzed using generalized estimating equation (GEE) models in R 4.3.0 (R Foundation for Statistical Computing, Vienna, Austria). The ML algorithm was executed using Python 3.11.5 (Python Software Foundation, Wilmington, USA) and Project Jupyter 6.5.4 (Anaconda, Inc., Austin, USA). Statistical significance was determined for all two-sided *P* values < 0.05.

RESULTS

Patient baseline characteristics categorized by phenogroups

FGIDs were confirmed in 300 of the 508 patients. A flowchart of the patients involved in this study was

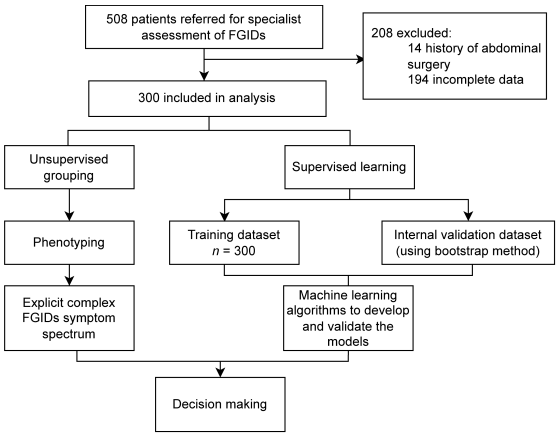

Figure 1: Flow diagram of patient cohort. FGIDs: functional gastrointestinal disorders.

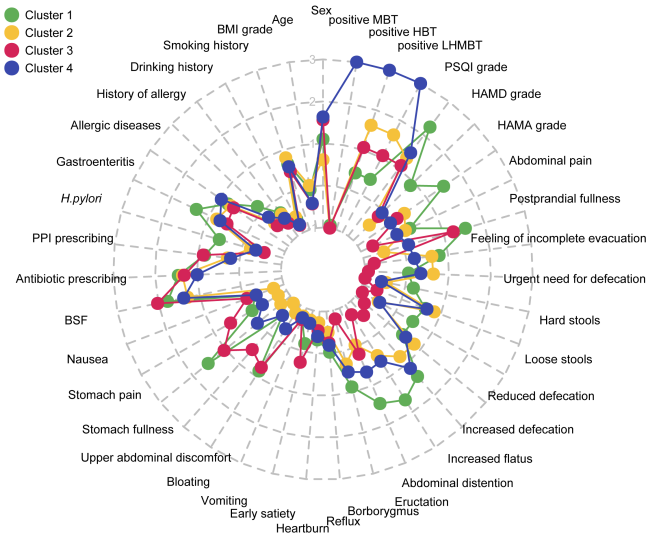

Figure 2: Description of clinical characteristics among clusters using a single spider plot superimposing the signature of each cluster. The values depicted were normalized to the entire population. Green, yellow, red, and deep purple indicate Clusters 1, 2, 3, and 4, respectively. MBT: methane breath tests, HBT: hydrogen methane breath tests, LHMBT: lactulose hydrogen methane breath tests, PSQI: Pittsburgh sleep quality index, HAMD: Hamilton Depression Scale, HAMA: Hamilton Anxiety Scale, BSF: Bristol stool form, PPI: proton pump inhibitor, BMI: body mass index.

shown in Figure 1. Using the Rome IV criteria, functional dyspepsia (FD) was the most common FGID (45.0%), followed by irritable bowel syndrome (IBS; 42.7%, $n = 128$; Table 1, Supplementary Figure 3 and 4A). Overlapping of FGIDs were seen in 115 (38.3%) patients (Table 1).

Excessive gas production was observed in 183 patients (61.0%, Table 1), with IBS representing the largest proportion (42.6%, $n = 78$), followed by FD (39.9%, $n = 73$; Supplementary Table 3 and Supplementary Figure 4B).

The K-means algorithm was used to integrate clinical symptoms and breath gas characteristics for phenogrouping. The most statistically significant clustering solution divided the entire patient population into four clusters using the elbow method based on a majority vote.

Cluster 1: Lower abdominal discomfort with the highest proportion of psychological burden

The cluster comprised 19.7% ($n = 59$) of the total participants and was characterized by higher-than-average scores for psychiatric scales (Table 1, all $P < 0.05$). Those patients presented more severe complaints of lower abdominal discomfort ($P < 0.001$) compared with the other three clusters, and many of them overlapped with other GI symptoms. The cluster had a higher proportion of extraintestinal symptoms such as oral ulcers. This cluster had the lowest proportion of excessive gas production ($P < 0.001$).

Cluster 2: Diarrhea predominance with a high prevalence of excessive gas production

This cluster consisted of 46.3% ($n = 139$) of the cohort (Table 1, Figure 2). The duration of this cluster was longer ($P = 0.007$). Most reported type 6/7 in the Bristol stool form (BSF). Many patients have reported experiencing abdominal pain and symptoms associated with gas production, such as borborygmus and increased flatus. This cluster had an above-average proportion of excessive gas production.

Cluster 3: Upper abdominal discomfort with a high prevalence of excessive gas production

This cluster comprised 21.3% ($n = 64$) of the total cohort (Table 1, Figure 2). The patients in this cluster were significantly younger, had a lower body mass index (BMI), and had a relatively short disease duration (all $P < 0.05$). This cluster was distinguished by a higher percentage of individuals experiencing upper abdominal discomfort (Supplementary Figure 5). This cluster had a relatively high proportion of excessive gas production, and most met the diagnosis of FD.

Cluster 4: Diarrhea predominance combined with postprandial symptoms presenting the highest prevalence of excessive gas production

This cluster comprised 12.7% ($n = 38$) of the cohort. Many of these patients had a history of allergic diseases. Similar to Cluster 2, patients in this cluster presented with loose stools and moderate lower GI symptoms. Many of them had postprandial symptoms. The cluster had the highest proportion of excessive gas production (100%, $P < 0.001$). The lowest scores were noted for the anxiety and depression scales ($P = 0.001$).

Gas time profiles after lactulose ingestion

The gas time profiles of the different clusters were shown

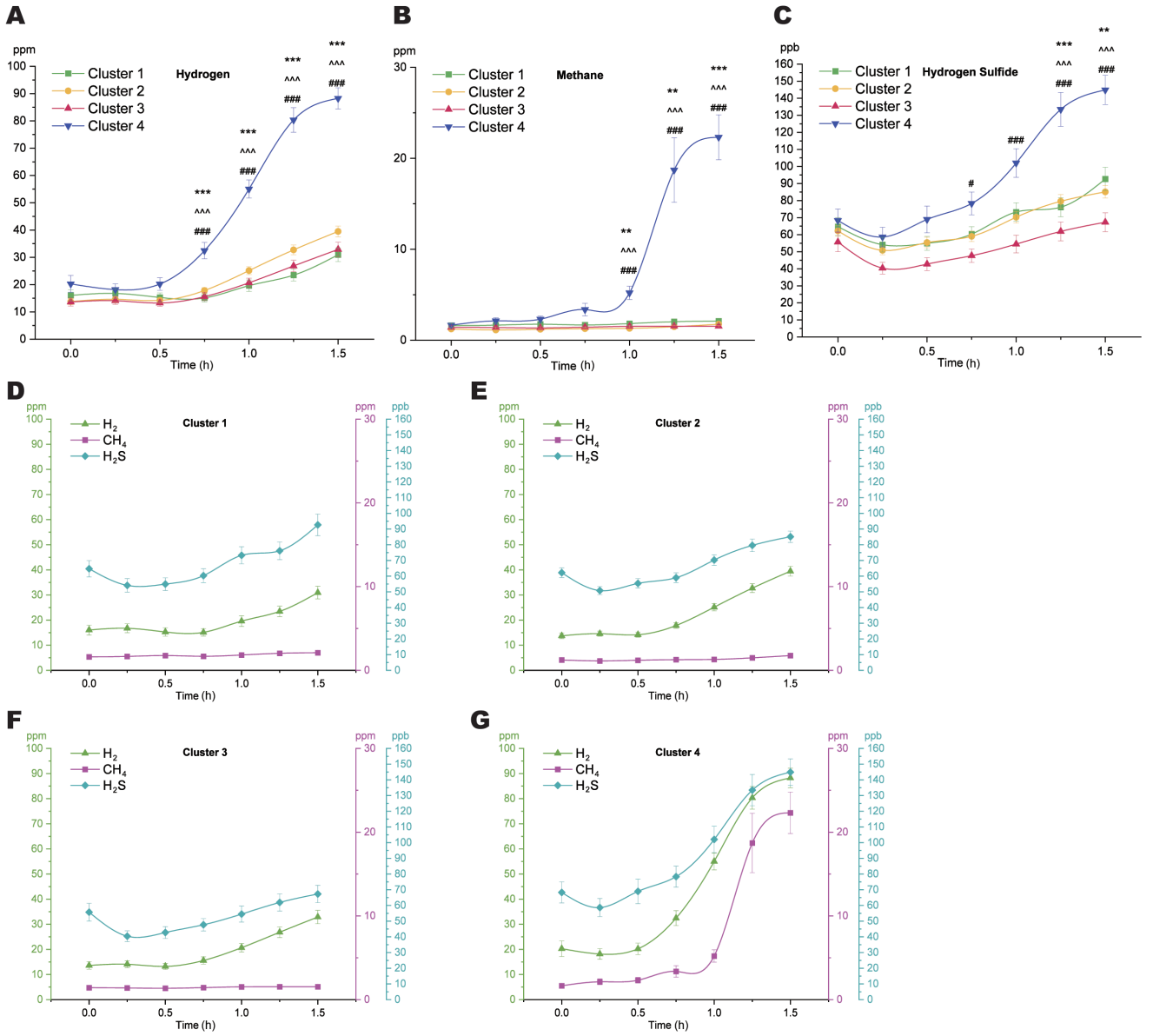

Figure 3: Gas time profiles in patients with FGIDs (n = 300). (A) Breath hydrogen concentration-time profiles grouped by clusters following lactulose ingestion. (B) Breath methane concentration-time profiles grouped by clusters following lactulose ingestion. (C) Breath hydrogen sulfide concentration-time profiles grouped by clusters following lactulose ingestion. (D) Gas time profiles in Cluster 1. (E) Gas time profiles in Cluster 2. (F) Gas time profiles in different Cluster 3. (G) Gas time profiles in Cluster 4. Data were mean ± SD. The differences of gas production between these clusters were examined by GEEs with time, group (Cluster 1, Cluster 2, Cluster 3 and Cluster 4), and group-by-time interaction as the covariates. Pairwise comparisons were performed with Bonferroni corrections. **Cluster 1 *vs*. Cluster 4, $P < 0.01$; ***Cluster 1 *vs*. Cluster 4, $P < 0.001$; ^Cluster 2 *vs*. Cluster 4, $P < 0.05$; ^^Cluster 2 *vs*. Cluster 4, $P < 0.01$; ^^^Cluster 2 *vs*. Cluster 4, $P < 0.001$; #Cluster 3 *vs*. Cluster 4, $P < 0.05$; ###Cluster 3 *vs*. Cluster 4, $P < 0.001$. FGIDs: functional gastrointestinal disorders, ppm: parts per million, ppb: parts per billion, H_2: hydrogen, CH_4: methane, H_2S: hydrogen sulfide, SD: standard deviation.

in Figure 3. Temporal trends and variations in breath gas following lactulose ingestion in different clusters of patients with FGIDs were analyzed using a generalized estimation equation (Supplementary Table 4 and 5). The levels of exhaled H_2, CH_4, and H_2S showed significant variations following ingestion of lactulose (all $P < 0.001$, Supplementary Table 4).

The increasing trends of exhaled H_2 concentrations were similar among the four clusters (Figure 3A). However, significant differences were observed in both the H_2 peak and H_2 AUC among the four clusters (all $P < 0.001$, Table 2). Compared to Cluster 1 and Cluster 3, Cluster 2 and Cluster 4 had higher H_2 peaks and AUCs, especially in Cluster 4 (all $P < 0.001$, Table 2). Further analysis using GEEs showed that H_2 concentrations were significantly elevated at multiple time points in Cluster 4 (all $P < 0.001$, Supplementary Table 5).

The temporal trends of exhaled CH_4 concentration were

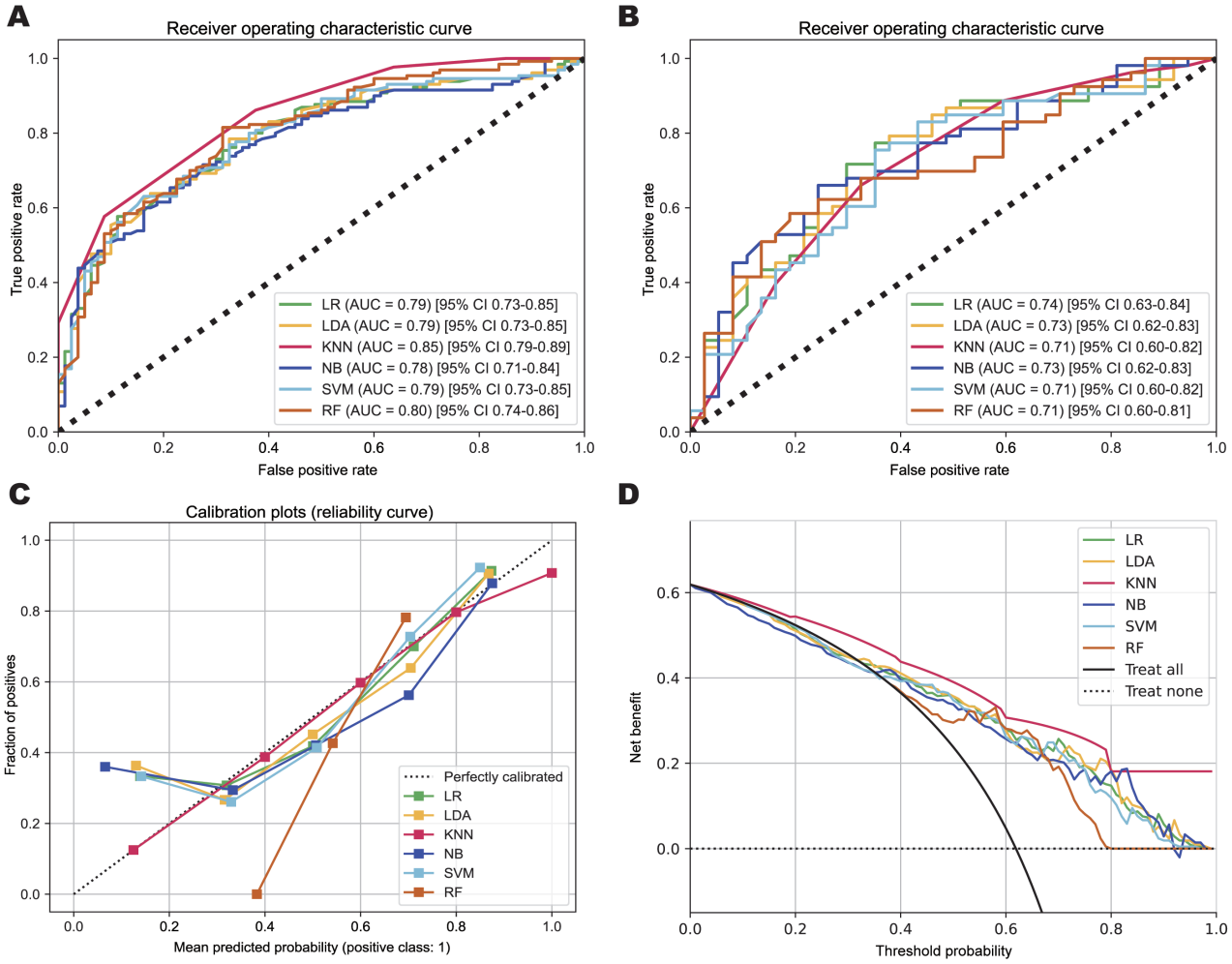

Figure 4: Performance evaluation of ML models for predicting excessive gas production in FGIDs. (A) ROC curves of different models adopted for predicting excessive gas production in the FGIDs cohort training dataset. (B) ROC curves of different models adopted for predicting excessive gas production in the FGIDs cohort internal validation dataset. (C) Calibration plot of different models for predicting excessive gas production in patients with FGIDs. The model-predicted probability of excessive gas production was plotted on the x-axis and the actual proportion of excessive gas production was plotted on the y-axis. An ideal calibration plot is indicated by a 45-diagonal line. (D) Decision curve analysis of the model. The x-axis indicates the threshold probability of the excessive gas production. The threshold probability is the level of certainty above which a patient or physician would choose to intervene. The y-axis indicates the net benefit, which is calculated as true positive rate − (false positive rate × weighting factor). ML: machine learning, FGIDs: functional gastrointestinal disorders, ROC: receiver operating characteristic curve, LR: logistic regression, LDA: linear discriminant analysis, KNN: k-nearest neighbor, NB: naive Bayes, SVM: support vector machine, RF: Random Forest, AUC: area under the curve, CI: confidence interval.

relatively flat among Cluster 1, 2, and 3, whereas CH_4 concentration showed a significant elevation at 75–90-min time points in Cluster 4 ($P < 0.001$, Figure 3B, Table2). Similar results were obtained using GEEs (Supplementary Table 5).

Re-increasing trends in H_2S were observed in all four clusters after a short period of decrease (Figure 3C). The change in H_2S concentration over time was distinct among the four clusters. Both multiple comparisons among clusters and the results of GEEs found that the H_2S concentration was significantly higher in Cluster 4 *vs.* the other clusters at multiple time points ($P < 0.05$, Supplementary Table 5). Notably, the total concentration of H_2S was not high in Cluster 3, but H_2S peaked earlier ($P = 0.021$, Table 2).

A supervised machine learning approach to predict excessive gas production

The FGIDs cohort was split into two sets: training with 300 patients and internal validation using bootstrap method. In the training dataset, 183 patients (61%) were identified as having excessive gas production, as defined by a positive LHMBT. Detailed clinical characteristics were provided in Supplementary Table 3.

Features selection

Based on literature reviews, expert opinions, and univariate

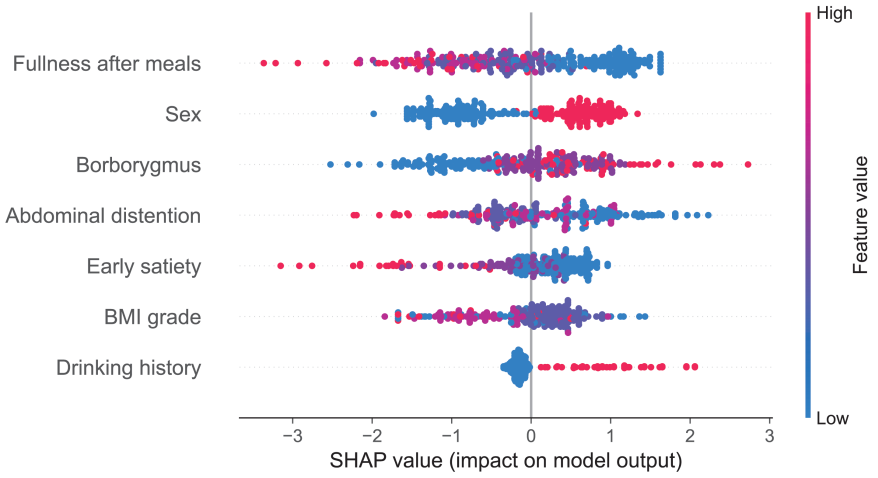

Figure 5: SHAP summary plot for the predictors contributing to the ML model prediction of excessive gas production. The organization of variables in the plot showcased their significance, with color coding showing the size and direction of impacts. A higher likelihood of having excessive gas production was represented by SHAP values exhibiting positive magnitudes. SHAP: SHapley Additive explanation, ML: machine learning.

comparisons between positive and negative LHMBT results in the training dataset, 36 features related to excessive gas production were identified. After adjusting for multicollinearity between variables, LASSO regression was utilized to pinpoint seven features for model construction: sex, BMI grade, drinking history, early satiety, fullness after meals, borborygmus, and abdominal distension. The predictors were utilized in the following modeling procedures (Supplementary Figure 6 and Supplementary Table 6).

Performance of the ML models in predicting excessive gas production

We conducted an analysis of the predictive performance of various ML models, as illustrated in Figure 4. The models exhibited varied capabilities for predicting excessive gas production, as defined by the positive LHMBT (Figure 4A). Among the six classification algorithms, KNN demonstrated superior performance, achieving an AUC of 0.85 for the cohort, demonstrating good discrimination ability (Figure 4A). The ML models exhibited robust performance in discriminating excessive gas production within the internal validation cohort, as evidenced by an AUC ranging from 0.71 to 0.74 (Figure 4B).

Furthermore, the calibration plots verified the agreement between the predicted excessive gas production from the ML models and actual observations (Figure 4C). Additionally, DCA revealed that by setting a probability threshold for excessive gas production occurrence within the range of 38.1% and 79.9%, utilizing the model yielded significantly superior net benefits compared to both strategies of "treating none" and "treating all" (Figure 4D). These findings highlighted substantial clinical applicability of our prediction model.

Interpreting models using the SHAP approach

We conducted a comprehensive investigation of the fundamental workings of our ML model using the SHAP method (Figure 5). The SHAP summary plot revealed that female individuals with a history of drinking, low BMI, moderate abdominal distention and postprandial symptoms, and exaggerated bowel sounds were more likely to be identified as having excessive gas production.

Finally, we successfully developed a user-friendly and easily accessible web-based nomogram for predicting excessive gas production (https://jxl-xs.shinyapps.io/DynNomapp/).

DISCUSSION

In this study, we demonstrated that the application of ML enabled novel integration of clinical parameters and gas production features to phenotype patients with FGIDs. These findings offer novel insights into FGID symptom experience that goes beyond the classification based on the Rome criteria.

We divided the entire patient population into four clusters using an unsupervised ML method. Cluster 2, 3, and 4 showed enrichments for abdominal distention and diarrhea, which were accompanied by a significant prevalence of small intestinal bacterial dysbiosis. Symptoms may be relieved by microbiota-targeted therapy, such as rifaximin, aimed at improving the imbalance in small intestinal flora.[18] Additionally, reducing the intake of fermented carbohydrates through a low fermentable oligo-, di-, mono-saccharides and polyols (FODMAP) diet may also provide relief from abdominal distention.[18] Cluster 1 was characterized by moderate lower abdominal discomfort

with the most psychological complaints and the lowest proportion of excessive gas production, which may benefit from targeted interventions with neuromodulators or psychotherapy. We further used a generalized estimation equation to analyze the gas concentration trends over time and variations between clusters. Compared to the other clusters, Cluster 4 showed higher H_2 and CH_4 concentrations at multiple time points and had a longer duration of excessive gas production. It is believed that the moment when gas levels peak aligns with when the main substance interacts with the majority of fermenting microorganisms.[14] This strongly suggests that small intestinal microbial overloads in Cluster 4 may lead to higher gut permeability and mild inflammation, leading to symptoms like diarrhea and bloating.

Furthermore, we investigated how breath H_2S levels were altered in the different clusters. H_2S is produced by both endogenous and exogenous pathways.[19] At low concentrations, endogenous basal H_2S production has beneficial effects such as reducing oxidative stress-related tissue injury.[20] Exogenous H_2S is usually produced by gut microbes,[19] which may drive the production of H_2S to high concentrations and exert a pro-inflammatory role.[21,22] In our study, we observed that H_2S levels were markedly elevated in Cluster 4, which was characterized by diarrhea and postprandial GI symptoms, suggesting that H_2S could be a biomarker for small intestinal dysbiosis in IBS.[23,24] Interestingly, we also observed that Cluster 3, which met the diagnosis of FD based on traditional classifications, reached the H_2S-peak earlier. In previous studies, H_2S-generating enzymes expressed in the enteric nervous system produced endogenous H_2S involved in gastric accommodation and induced a stimulatory effect on duodenal motility in animal models.[25,26] Excessive exogenous H_2S from sulfate-reducing bacteria could play an inhibitory role in regulating gastric and duodenal motility.

Because breath gas production was significantly associated with FGID symptoms and different clusters, we established a specific model to accurately predict excessive gas production. Sex, BMI grade, drinking history, early satiety, fullness after meals, borborygmus, and abdominal distension were identified as predictors in the models. The identified symptoms were good predictors of excessive gas production, aligning with the fact that moderate GI dysfunction with gas-related symptoms such as borborygmus is a cardinal symptom of small intestinal dysbiosis. Sex was identified as an influential factor in determining model output and is potentially associated with variations in estrogen levels that affect gastric acid secretion and motility in animals.[27] BMI was also an influential factor. This is justifiable because dysbiosis in the small intestine leads to poor carbohydrate absorption owing to microbial competition with the host for essential nutrients.[28] The consumption of alcohol in moderation significantly increases the likelihood of developing small intestinal dysbiosis through direct toxic damage to the intestinal epithelium.[29]

Although ML has garnered significant interest for clustering and accurate prediction, its "black-box" nature often yields outcomes that are difficult to interpret. Therefore, we chose unsupervised ML to offer a more significant explanation and differentiation of patient categories associated with small intestinal dysbiosis. Subsequently, we developed a predictive model for excessive gas production using supervised ML. We employed the SHAP method to elucidate the decision-making process of our prediction model, enabling clinicians to gain comprehensive insights into the roles of different variables and establish informed decisions for their patients. Finally, a feasible nomogram for predicting excessive gas production was developed, which can be externally validated.

Our study had some limitations. Long-term follow-up of the effectiveness of personalized treatment in different clusters of FGIDs patients is needed to further understand and optimize management strategies for these conditions. Lastly, the prediction model should be validated externally to test its generalizability.

CONCLUSION

We demonstrated that the application of unsupervised ML can yield comprehensive, comprehensible, and medically significant categorization of diverse groups of patients with FGIDs. All four clusters revealed considerable differences in clinical features and gas time profiles, suggesting that evaluation of gas production could be valuable in guiding therapeutic decisions. Given the strong association between gas production and FGID symptoms, we developed and optimized a predictive model that exhibited favorable performance for excessive gas production. Our new clustering and predictive models using ML hold promise for effectively exploring subsets of FGIDs associated with gas production, thus facilitating clinical decision-making and providing improved treatment guidance.

Acknowledgement

We express our gratitude to Xin Liu, Songfei Li, Cheng Zhang, Qiuhong Li, Xin Guan, Chao Ma, Fan He, Hui Zhang, Huanhuan Shi, and Mingfei Zhu (Department of Gastroenterology, Peking University Third Hospital) for their help in data collection for this research.

Author Contributions

Lingling Zhu: Investigation, Methodology, Software,

Validation, Formal analysis, Visualization, Writing—Original draft preparation, and Writing—Reviewing and Editing. Shuo Xu: Methodology, Software, Validation, Formal analysis, and Visualization. Huaizhu Guo: Investigation, Methodology. Siqi Lu: Investigation, Methodology. Jiaqi Gao: Investigation, Methodology. Nan Hu: Investigation, Methodology. Chen Chen: Investigation, Methodology. Zuojing Liu: Investigation, Methodology. Xiaolin Ji: Investigation, Methodology. Kun Wang: Conceptualization, Supervision; Liping Duan: Conceptualization, Supervision, and Funding acquisition. All authors approved the final version of the manuscript.

Ethical Approval

The study was approved by the ethics committee of Peking University Third Hospital (2021-132-01). This study was conducted in accordance with the Declaration of Helsinki.

Informed Consent

Informed consent was obtained from the participants on an opt-out basis given the noninvasive nature of the study.

Source of Funding

This work was supported by the National Key Research and Development Program of China (No. 2019YFA0905600) and the Clinical Cohort Construction Program of Peking University Third Hospital (No. BYSYDL2023002).

Conflict of Interest

The authors declare no competing interest.

Data Availability Statement

The data used to support the findings of this study are available from the corresponding author upon request.

REFERENCES

1. Black CJ, Drossman DA, Talley NJ, Ruddy J, Ford AC. Functional gastrointestinal disorders: advances in understanding and management. Lancet 2020;396:1664–1674.
2. Aziz I, Palsson OS, Törnblom H, Sperber AD, Whitehead WE, Simrén M. The prevalence and impact of overlapping Rome IV-diagnosed functional gastrointestinal disorders on somatization, quality of life, and healthcare utilization: a cross-sectional general population study in three countries. Am J Gastroenterol 2018;113:86–96.
3. Vanheel H, Carbone F, Valvekens L, Simren M, Tornblom H, Vanuytsel T, *et al*. Pathophysiological abnormalities in functional dyspepsia subgroups according to the Rome III criteria. Am J Gastroenterol 2017;112:132–140.
4. Mousavi E, Keshteli AH, Sehhati M, Vaez A, Adibi P. Re-investigation of functional gastrointestinal disorders utilizing a machine learning approach. BMC Med Inform Decis Mak 2023;23:167.
5. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, *et al*. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402–2410.
6. Cikes M, Sanchez-Martinez S, Claggett B, Duchateau N, Piella G, Butakoff C, *et al*. Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy. Eur J Heart Fail 2019;21:74–85.
7. Byale A, Lennon RJ, Byale S, Breen-Lyles M, Edwinson AL, Gupta R, *et al*. High-dimensional clustering of 4000 irritable bowel syndrome patients reveals seven distinct disease subsets. Clin Gastroenterol Hepatol 2024;22:173–184.e12.
8. Siah KTH, Gong X, Yang XJ, Whitehead WE, Chen M, Hou X, *et al*. Rome Foundation-Asian working team report: Asian functional gastrointestinal disorder symptom clusters. Gut 2018;67:1071–1077.
9. Wauters L, Li H, Talley NJ. Editorial: disruption of the microbiota-gut-brain axis in functional dyspepsia and gastroparesis: mechanisms and clinical implications. Front Neurosci 2022;16:941810.
10. Gurusamy SR, Shah A, Talley NJ, Koloski N, Jones MP, Walker MM, *et al*. Small intestinal bacterial overgrowth in functional dyspepsia: a systematic review and meta-analysis. Am J Gastroenterol 2021;116:935–942.
11. Shah A, Talley NJ, Jones M, Kendall BJ, Koloski N, Walker MM, *et al*. Small intestinal bacterial overgrowth in irritable bowel syndrome: a systematic review and meta-analysis of case-control studies. Am J Gastroenterol 2020;115:190–201.
12. Chuah KH, Wong MS, Tan PO, Lim SZ, Beh KH, Chong SCS, *et al*. Small intestinal bacterial overgrowth in various functional gastrointestinal disorders: a case-control study. Dig Dis Sci 2022;67:3881–3889.
13. Lacy BE, Cangemi D, Vazquez-Roque M. Management of chronic abdominal distension and bloating. Clin Gastroenterol Hepatol 2021;19:219–231.e1.
14. Wilder-Smith CH, Olesen SS, Materna A, Drewes AM. Fermentable sugar ingestion, gas production, and gastrointestinal and central nervous system symptoms in patients with functional disorders. Gastroenterology 2018;155:1034–1044.e6.
15. Wilder-Smith CH, Drewes AM, Materna A, Olesen SS. Extragastrointestinal symptoms and sensory responses during breath tests distinguish patients with functional gastrointestinal disorders. Clin Transl Gastroenterol 2020;11:e00192.
16. Rezaie A, Buresi M, Lembo A, Lin H, McCallum R, Rao S, *et al*. Hydrogen and methane-based breath testing in gastrointestinal disorders: the North American consensus. Am J Gastroenterol 2017;112:775–784.
17. Basagaña X, Barrera-Gómez J, Benet M, Antó JM, Garcia-Aymerich J. A framework for multiple imputation in cluster analysis. Am J Epidemiol 2013;177:718–725.
18. Pimentel M, Saad RJ, Long MD, Rao SSC. ACG clinical guideline: small intestinal bacterial overgrowth. Am J Gastroenterol 2020;115:165–178.
19. Szabo C. Gasotransmitters in cancer: from pathophysiology to experimental therapy. Nat Rev Drug Discov 2016;15:185–203.
20. Wallace JL, Wang R. Hydrogen sulfide-based therapeutics: exploiting a unique but ubiquitous gasotransmitter. Nat Rev Drug Discov 2015;14:329–345.
21. Wolf PG, Cowley ES, Breister A, Matatov S, Lucio L, Polak P, *et al*. Diversity and distribution of sulfur metabolic genes in the human gut microbiome and their association with colorectal cancer. Microbiome 2022;10:64.
22. Wallace JL, Motta JP, Buret AG. Hydrogen sulfide: an agent of stability at the microbiome-mucosa interface. Am J Physiol Gastrointest Liver Physiol 2018;314:G143–G149.
23. Banik GD, De A, Som S, Jana S, Daschakraborty SB, Chaudhuri S, *et al*. Hydrogen sulphide in exhaled breath: a potential biomarker for small intestinal bacterial overgrowth in IBS. J Breath Res 2016;10:026010.
24. Villanueva-Millan MJ, Leite G, Wang J, Morales W, Parodi G, Pimentel ML, *et al*. Methanogens and hydrogen sulfide producing bacteria guide

distinct gut microbe profiles and irritable bowel syndrome subtypes. Am J Gastroenterol 2022;117:2055–2066.
25. Lu W, Li J, Gong L, Xu X, Han T, Ye Y, *et al*. H2 S modulates duodenal motility in male rats via activating TRPV1 and K(ATP) channels. Br J Pharmacol 2014;171:1534–1550.
26. Xiao A, Wang H, Lu X, Zhu J, Huang D, Xu T, *et al*. H2S, a novel gasotransmitter, involves in gastric accommodation. Sci Rep 2015;5:16086.
27. Chen B, Kim JJW, Zhang Y, Du L, Dai N. Prevalence and predictors of small intestinal bacterial overgrowth in irritable bowel syndrome: a systematic review and meta-analysis. J Gastroenterol 2018;53:807–818.
28. Wielgosz-Grochowska JP, Domanski N, Drywień ME. Influence of body composition and specific anthropometric parameters on SIBO type. Nutrients 2023;15:4035.
29. Gabbard SL, Lacy BE, Levine GM, Crowell MD. The impact of alcohol consumption and cholecystectomy on small intestinal bacterial overgrowth. Dig Dis Sci 2014;59:638–644.

How to cite this article: Zhu L, Xu S, Guo H, Lu S, Gao J, Hu N, *et al*. Machine learning-based phenogroups and prediction model in patients with functional gastrointestinal disorders to reveal distinct disease subsets associated with gas production. J Transl Intern Med 2024; 12: 355-366.