Stratification and model-based analysis of patients with Irritable Bowel Syndrome using advanced biostatistics and medical data mining techniques

Annikka Virginia Polster

Department of Internal Medicine and Clinical Nutrition
Institute of Medicine
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2018

Cover illustration: Annikka Polster

Stratification and model-based analysis of patients with Irritable Bowel Syndrome using advanced biostatistics and medical data mining techniques

annikka.polster@gu.se
annikka.polster@gmx.de

ISBN 978-91-7833-028-7 (PRINT)
ISBN 978-91-7833-027-0 (PDF, e-pub)

Printed in Gothenburg, Sweden 2018
Printed by BrandFactory

To my family, whose endless support made this work possible.

And to the sofa on which I wrote the whole thing.

Stratification and model-based analysis of patients with Irritable Bowel Syndrome using advanced biostatistics and medical data mining techniques

Annikka Virginia Polster

Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

ABSTRACT

Irritable Bowel Syndrome (IBS) is characterized by symptoms that are dominated by abdominal pain and abnormal bowel habits, as defined by the Rome criteria. The complexity of the disorder is exemplified by the heterogeneity of symptom profiles and the number of putative pathophysiological mechanisms. Currently it is unclear whether IBS is a multifactorial disorder or rather a summary diagnosis for several distinct disease entities displaying similar symptoms. This thesis aims to identify subgroups of clinical relevance by developing and demonstrating symptom- and mechanism-based stratification approaches, as well as an integrative analysis pipeline aiming to link different pathophysiological mechanisms.

In a clinical sample of IBS patients, as well as in subjects fulfilling IBS in a population-based sample, symptom-based stratification yielded reproducible subgroups, characterized by combinations of gastrointestinal, extra-intestinal somatic and psychological symptoms. In the population-based sample this subgrouping was associated with differences in healthcare utilization. Mechanism-based stratification, focusing on the function of the autonomic nervous system (ANS), demonstrated altered ANS function in IBS patients compared to healthy controls, and identified a subgroup of IBS patients with aberrant overall ANS function, which was associated with more severe diarrhea. This thesis also introduces a stepwise multilevel integrative analysis pipeline using network theory, which presents associations of host-gene expression with mucosa-adherent gut microbiota as well as key IBS symptoms, revealing distinct IBS-specific associations.

In conclusion, IBS patients show reproducible subgroups with specific profiles of a comprehensive set of IBS related symptoms and differences in healthcare needs based on these subgroups. Further, multivariate comparisons between IBS patients and healthy controls aid in identifying individuals for which specific complex pathophysiological mechanisms may be of relevance, as demonstrated by identifying a subset of IBS patients with aberrant overall ANS function. This stratification approach could be applied to other pathophysiological mechanisms. Our stepwise multilevel integrative analysis pipeline showed differences in variable associations at the gut mucosal level between IBS patients and healthy controls, and is therefore a model for further, comprehensive analysis of the complex pathophysiology of IBS

Keywords: Irritable bowel syndrome (IBS), subgroup analysis, Gaussian mixture models, heart rate variability, integrative analysis, network medicine

ISBN: 978-91-7833-028-7 (Print); ISBN: 978-91-7833-027-0 (PDF)

SAMMANFATTNING PÅ SVENSKA

Irritable Bowel Syndrome (IBS) är en vanlig sjukdom som karaktäriseras av buksmärta i kombination med diarré och/eller förstoppning. IBS är en komplex sjukdom där symtombilden varierar mellan olika patienter, och där ett stort antal underliggande orsaker till symtomen lyfts fram. För närvarande är det oklart om IBS är en separat sjukdom med många olika orsaker, eller snarare en paraplydiagnos för flera olika specifika sjukdomar som uppvisar likartade symtom.

Denna avhandling syftar till att identifiera kliniskt relevanta undergrupper bland IBS-patienter baserat på skillnader och likheter i symtom och sjukdomsmekanismer, och att använda olika avancerade statistiska metoder för att belysa samband mellan olika sjukdomsmekanismer vid IBS.

I en grupp IBS-patienter som sökt vård vid vår mag-tarmmottagning och bland personer med IBS som identifierats i en populationsbaserad internet-studie, kunde vi påvisa reproducerbara grupper av IBS-patienter baserat på förekomst av olika symtom. Dessa grupper karaktäriserades av olika kombinationer av mag-tarmsymtom, psykologiska symtom och kroppsliga symtom från andra delar av kroppen än magen och tarmen. I den populationsbaserade studien så var denna gruppering av patienter associerad med skillnader i hur mycket sjukvård individerna utnyttjade.

I ett annat arbete påvisade vi skillnader i funktion i det autonoma nervsystemet (ANS) (den del av vårt nervsystem som står utanför viljans kontroll) mellan IBS-patienter och friska kontrollpersoner. Dessutom identifierades med hjälp av mer avancerade statistiska metoder en grupp av IBS-patienter med avvikande ANS-funktion, och detta var kopplat till mer uttalade diarré-besvär. Användning av den här typen av avancerade statistiska analyser kan underlätta identifieringen av grupper av IBS-patienter med olika dominerande sjukdomsmekanismer, så som här skett genom identifiering av en grupp IBS-patienter med avvikande ANS-funktion.

I den här avhandlingen introducerar vi också ett statistiskt tillvägagångsätt för att studera hur olika sjukdomsmekanismer interagerar och hur detta kan associeras till uppkomst av symtom. Detta exemplifieras genom att påvisa skillnader i interaktioner mellan tarmbakterier och faktorer i tarmens slemhinna mellan IBS-patienter och friska kontrollpersoner, och hur dessa interaktioner sedan är kopplade till symtom vid IBS. Detta nya tillvägagångssätt kan fungera som en modell för ytterligare omfattande analyser av komplexa orsakssamband hos patienter med IBS.

Sammanfattningsvis kan man med avancerade statistiska analysmetoder identifiera distinkta grupper av IBS-patienter baserat på symtomprofil och underliggande sjukdomsmekanismer. Dessa metoder kan användas vidare för att studera komplexa interaktioner vid IBS för att bättre förstå denna patientgrupp och dess underliggande orsaker.

Nyckelord: Irritable bowel syndrome (IBS), subgruppsanalys, Gaussian mixture models, hjärtfrekvensvariabilitet, integrativ analys, medicinsk nätverksanalys

ISBN: 978-91-7833-027-0 (digital copy), ISBN: 978-91-7833-028-7 (print)

List of papers

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. **Polster A**, Van Oudenhove L, Jones M, Öhman L, Törnblom H, Simrén M. *Mixture model analysis identifies IBS subgroups characterised by specific profiles of gastrointestinal, extraintestinal somatic and psychological symptoms.* Aliment Pharmacol Ther. 2017;00:1–11. https://doi.org/10.1111/apt.14207

II. **Polster A**, Palsson OS, Törnblom H, Öhman L, Sperber AD, Whitehead WE, Simrén M. *Population-based IBS subgroups characterized by specific profiles of GI and non-GI symptoms identified through Mixture model analysis report differences in healthcare utilization.* Submitted

III. **Polster A**, Friberg P, Gunterberg V, Öhman L, Le Nevé B, Törnblom H, Cvijovic M, Simrén M. *Heart rate variability characteristics of patients with irritable bowel syndrome and associations with symptoms*. Neurogastroenterol Motil. 2018;e13320. https://doi.org/10.1111/nmo.13320

IV. **Polster A**, Öhman L, Tap J, Derrien M, Le Nevé B, Sundin J, Törnblom H, Cvijovic M, Simrén M. *A network model reveals distinct microbiota-host interactions in irritable bowel syndrome.* In manuscript

V. Supplementary publications

Black, C. J., and A. C. Ford. *Subgroups in irritable bowel syndrome—more than just diarrhoea and constipation?* Alimentary pharmacology & therapeutics 46.7 (2017): 697-697.

Polster, A, et al. *Subgroups in irritable bowel syndrome—more than just diarrhoea and constipation? Authors' reply.* Alimentary pharmacology & therapeutics 46.7 (2017): 698-699.

Contents

Abbreviations

ANS	Autonomic Nervous System
BIC	Bayesian Information Criterion
CHG	Chromogranin
CLD	Claudin
DOUX	Dual-Oxidase
FFAR	Free Fatty Acid Receptor
GI	Gastrointestinal
GSRS	Gastrointestinal Symptom Rating Scale
HAD	Hospital Anxiety and Depression scale
HC	Healthy controls
HRV	Heart Rate Variability
IBS	Irritable Bowel Syndrome
IBS-C	IBS with constipation
IBS-D	IBS with diarrhea
IBS-M	IBS with mixed loose and hard stools
IBS-SSS	IBS-Severity Scoring System
IBS-U	Unsubtyped IBS
IL	Interleukin
MUC	Mucin
NOX	NAD(P)H oxidase
OCLN	Occludin
OPLS-DA	Orthogonal Partial Least Squares Discriminant Analysis

PAR2	Protease activated receptor 2
PCA	Principal Component Analysis
PHQ	Patient Health Questionnaire
PNS	Parasympathetic Nervous System
SCG	Secretogranin
SLC	Solute Carrier
SNS	Sympathetic Nervous System
TJP	Tight Junction Protein
TLR	Toll-like receptor
TNF	Tumor necrosis factor
TPH 1	Tryptophan hydroxylase 1

1. Purpose of this thesis

Complex disorders are one of the major challenges of modern day medicine. They are prevalent, and are characterized by very heterogeneous patient cohorts as well as multifaceted pathophysiology. Ongoing efforts aim to tackle this complexity by stratifying patients into more homogenous subgroups and by developing methods to better understand the relationship between the multiple physiological and pathophysiological mechanisms within individuals and groups of patients.

Irritable Bowel Syndrome (IBS) is such a complex disorder: It is frequent, affecting around 10% of the worldwide population, and defined by medically unexplained recurrent abdominal pain and changes in bowel habits. Patients commonly display several fluctuating gastrointestinal (GI) symptoms and often additional unexplained non-GI symptoms. A number of pathophysiological mechanisms are considered to be relevant for IBS, but not all of these are equally relevant for the individual patient.

Current subgrouping of IBS patients only takes into account bowel habits, disregarding the full symptom spectrum of these patients, and has not shown to be stable over time or to sufficiently relate to either pathophysiology or treatment response of the individual patients.

This thesis hypothesizes that the clinical phenotype of IBS is a summary diagnosis of several distinct subgroups with different underlying pathophysiological mechanisms (Figure 1).

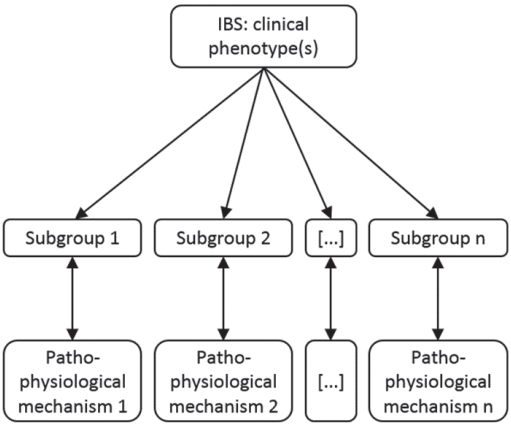
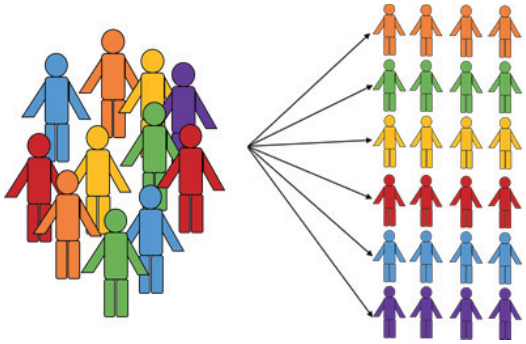


Figure 1: Hypothesis underlying this thesis: IBS is a summary diagnosis consisting of different subgroups with distinct pathophysiological mechanisms.

The purpose of this thesis is to explore data analysis methods and develop analytical strategies suitable to stratify IBS patients into novel subgroups (Aim 1, Figure 2), and to understand the relationship between multiple pathophysiological mechanisms and symptoms by developing suitable strategies for performing integrative analyses (Aim 2, Figure 2).

Aim 1: Stratification of heterogeneous IBS cohorts into novel subgroups



Aim 2: Developing integrative analysis strategies suitable to explore the relationship between multiple pathophysiological factors in IBS

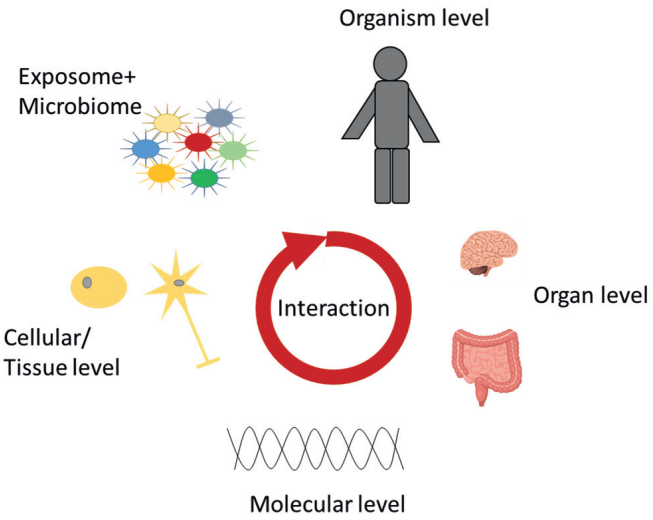


Figure 2: Aims of this thesis

2. Background

2.1. Irritable Bowel Syndrome: Definition and Diagnosis

Functional gastrointestinal disorders (FGIDs) describe a number of disorders characterized by chronic GI and non-GI symptoms. The symptoms are medically unexplained, meaning that no physiological or biochemical abnormalities that are detectable in clinical routine are associated with symptom occurrence. Therefore, these disorders are defined by symptom-based diagnostic criteria, as currently no or few biomarkers are available to facilitate definition and diagnosis of these disorders[1, 2]. One of the most common FGIDs is Irritable Bowel Syndrome (IBS)[3], describing patients who suffer predominantly from symptoms of the lower GI tract.

As of writing, IBS is defined by the Rome criteria[4], with the vast majority of recent studies (including the manuscripts of this thesis) being based on their third iteration, the Rome III criteria[5]. In 2016 an updated version, the Rome IV criteria[6], was published, which describes FGIDs as disorders of gut-brain interaction and presents slightly stricter diagnostic criteria for IBS. The detailed criteria of both Rome III and Rome IV, which are used to define and diagnose IBS in clinical practice and research, are shown in Table 1.

1.1. History of IBS

The Rome IV criteria are the newest attempt of defining a disorder which has undergone various changes of names and diagnostic criteria over the years. A look at this terminology gives insight into the pathophysiological assumptions made to explain this disorder, with the most relevant examples here being the terms 'Mucous colitis', 'Membranous enteritis', 'Spastic colon', 'Nervous Stomach' as well as 'Intestinal Neurosis'. With increasing understanding of the mechanisms of this disorder, none of these terms have shown to sufficiently or unambiguously describe it. Patients do not show signs of active mucosal inflammation[6, 7], nor is the syndrome limited solely to the large intestine[8, 9], making terms such as 'colitis' and 'enteritis' misleading. While increased muscular activity of the GI tract has indeed been identified in some patients[10], other patients instead exhibit a decrease in

Table 1: Rome III and Rome IV diagnostic criteria

Rome III diagnostic criteria* for IBS[5]
Recurrent abdominal pain or discomfort** at least 3 days per month in the last 3 months associated with 2 or more of the following: 1. Improvement with defecation 2. Onset associated with a change in frequency of stool 3. Onset associated with a change in form (appearance) of stool *Criteria fulfilled for the last 3 months with symptom onset at least 6 months prior to diagnosis **Discomfort means an uncomfortable sensation not described as pain. In pathophysiology research and clinical trials, a pain/discomfort frequency of at least 2 days a week during screening evaluation is recommended for subject eligibility.
Rome IV diagnostic criteria* for IBS[6]**
Recurrent abdominal pain on average at least 1 day per week in the last 3 months, associated with 2 or more of the following criteria: 1. Related to defecation 2. Associated with a change in frequency of stool 3. Associated with a change in form (appearance) of stool ***Criteria fulfilled for the last 3 months with symptom onset at least 6 months before diagnosis

muscular activity/motility[11], therefore the term 'spastic colon' also does not capture the full nature of the problem. The last two examples refer to a more psychosomatic explanatory model, assuming that anxiety or stress triggers the symptoms. While some patients indeed report such an association[12, 13], others relate the onset of their symptoms for example to GI infections or intake of certain foods[14], suggesting that these factors cannot be neglected in understanding this disorder, rendering also terms such as 'Nervous Stomach' or 'Intestinal Neurosis' insufficient for describing the disorder.

The term Irritable Bowel Syndrome was introduced 1950 in the Rocky Mountain Medical Journal by Philip W Brown[15]. It is a more generic term than the previous, merely describing a bouquet of symptoms that are associated with a hypersensitive small and large intestine without detailing the assumed pathophysiology underlying these symptoms.

As mentioned above, not only the terminology describing this disorder has changed over the years, also the criteria defining the disorder and distinguishing it from similar disorders has been adapted to the growing knowledge about it. The first diagnostic criteria, termed

Manning criteria[16], were created in 1978, and succeeded 1989 by the first version of the Rome criteria[17], which have recently been updated to their fourth version[6], as described above.

1.2. Rome Subtypes

Given the lack of biomarkers[1] and rather broad diagnostic criteria, IBS presents as a very heterogeneous disorder. It is currently subdivided into four Rome subtypes, which are based on the individual patients' predominant bowel habits. The focus of these current subtypes lies solely on the seemingly paradoxical GI symptoms constipation and diarrhea, and can be defined either by estimating the predominant bowel type using the Rome diagnostic questionnaire[5, 6], or by registering the types in a diary using the Bristol Stool Form Scale [18]. If 25% or more the bowel movements are hard or lumpy stools, the patient is considered as having constipation-predominant IBS (IBS-C), if 25% or more are loose or mushy, the disorder is termed as diarrhea-predominant IBS (IBS-D). If both forms occur to 25% or more, it is termed mixed IBS (IBS-M), and if both occur less often, unsubtyped IBS (IBS-U)[5, 6].

Since the current definitions of IBS focuses purely on abdominal pain/discomfort and abnormal bowel habits[5, 6], using these symptoms in the definition of subtypes seems logical. Nevertheless, the current subtypes have three major disadvantages:

I) Patients often switch subgroups over time, as commonly bowel habits fluctuate both regarding form and frequency as well as regarding severity[19-22]

II) The subtypes have not shown to be sufficiently associated to pathophysiological mechanisms or treatment response[3, 23-25]

III) IBS patients show a variety of additional GI and non-GI symptoms, which are not taken into account[26, 27]. More on these additional symptoms is covered in the 'Symptoms and comorbidities' chapter.

1.3. Epidemiology

1.3.1. Prevalence

IBS is assumed to be the most common GI disorder encountered in primary care settings[28] and likely also the most common disorder presented to specialized gastroenterologists[29]. The prevalence of IBS in most Western countries and China is estimated to be between 1.1% - 35.5% [30], whereas for many African, South American and Asian countries such estimates are unavailable, in part due to practical difficulties in assessing the prevalence in these countries, but also in part likely due to an elevated prevalence of infectious diseases, "drowning" the IBS diagnoses[31]. While IBS is present in children as well as adults of all ages, the prevalence is negatively correlated to age[32, 33].

There are also attempts to estimate the prevalence of the Rome subtypes, which has proven difficult, given the fluctuation of predominant bowel habits in the individual patients[19-22] and the resulting switching of subtype over the course of time. It is assumed that, at a given time point, approximately one third of the IBS population can be subtyped as IBS-C, and likewise one third as IBS-D, with IBS-M and –U sharing the last third to varying proportions[34].

1.3.2. Gender distribution

A recurring phenomenon observed in IBS patients is the approximately two-fold higher proportion of females[32], and several reasons for this phenomenon are being discussed. It has been shown that some females experience symptom exacerbation in association to the menstrual cycle[35], suggesting hormonal factors to be of relevance here. Indeed it has been shown that estrogen can interact with gut serotonin receptors, thus influencing GI motility[35], and it has been shown that females often exhibit slower transit time[11, 36] as well as more visceral hypersensitivity than males[37], but these observations have not sufficiently explained gender differences. Psychosocial factors such as psychological wellbeing and quality of life, as well as symptom reporting, coping abilities and gender role have also been assessed and compared between the genders[38-42], but without sufficiently explaining this phenomenon.

1.4. Quality of life and economic impact

When measuring the quality of life (QOL) in patients, it is necessary to distinguish between general (i.e. the overall health perception) and disease-specific (i.e. the disturbance of wellbeing specifically caused by the disease) QOL. With regards to general QOL, IBS patients show a significant reduction regardless of culture, and have even shown a greater reduction in perceived general QOL than patients with gastroesophageal reflux disease, diabetes mellitus or severe chronic kidney disease[43]. It is noteworthy though to mention here that a reduction in general QOL is more prominent in IBS patients when compared to individuals fulfilling IBS diagnostic criteria who do not seek health care[44].

With regards to this thesis it is of high importance to mention that disturbed QOL in IBS patients was not solely associated to GI symptoms, but also to psychiatric comorbidities, such as nervousness and hopelessness, as well as extraintestinal somatic symptoms, especially tiredness, reduced energy, sleeping difficulties and reduced libido[45, 46].

The impact IBS has on the patients' lives is not limited to this reduction in QOL, but also impacts the social life and manifests in high healthcare utilization and work absenteeism, resulting in high costs for the patients as well as for the healthcare system and society[47].

1.5. Risk factors

The differences in IBS prevalence between genders, as described above, make it the most explored risk factor for IBS[48], but more factors are associated with an increased risk of IBS: Family aggregation, likely due to genetic factors or social learning, has been observed[49-51], but also individual factors such as psychological factors, stressful life events and abuse history, young age, abdominal surgery or endometriosis have been reported[23, 32, 52-55].

The occurrence of IBS after a healed gastrointestinal infection, termed post-infectious IBS, has gained special attention as the probably strongest risk factor for IBS, and is by some even considered an independent disease entity. It is assumed to occur in around 10% of cases, and may be higher (up to 30%) after epidemic infections[56-59], although this may be due to reporting or recall biases. The occurrence of post-infectious IBS is associated to the severity

of the GI infection (which may again relate to recall bias), and is more likely in younger age and females, and related to several psychosocial factors[59].

1.6. Symptoms and comorbidities

Besides the defining symptoms of IBS, abdominal pain and constipation or diarrhea, patients display a variety of additional GI and non-GI symptoms. The IBS-specific Gastrointestinal Rating Scale, a commonly used questionnaire developed to register key IBS symptoms, considers bloating and abdominal distention, urgency, incomplete bowel emptying, loose or hard stools, passing gas as well as early satiety and postprandial fullness to be the most relevant GI symptoms of IBS[60]. In addition to that, it is known that a subset of patients also suffer from other upper GI symptoms[61], and a substantial number of patients suffer from more than one FGID[62]. Despite the fact that IBS patients foremost seek help for their gastrointestinal problems, a large proportion also exhibits extra-intestinal somatic symptoms such as back or joint pain, headaches, sleep impairment and tiredness, palpitations, problems during sexual intercourse and menstrual distress[26, 63-67].

There are several comorbidities that are frequently present in IBS patients, especially psychological disorders such as anxiety and depression[12, 68, 69] . Non-GI functional syndromes such as fibromyalgia and chronic fatigue syndrome[66], but also migraine, eating disorders, prostatic pain syndrome, and the urologic chronic pelvic pain syndrome are present in many patients[67], and show a higher prevalence in the IBS population than can be expected based on the prevalence of the respective disorders in the general population[23].

These symptoms and comorbidities are of high relevance for the overall symptom burden of the patients, as especially the presence of extra-intestinal somatic symptoms has been shown to be equally relevant in predicting the overall burden of illness and quality of life of the affected patients as are the GI symptoms[70].

1.7. Potential pathophysiological mechanisms

As mentioned initially, the exact pathophysiological mechanisms underlying the different symptoms of IBS are undetermined. Nevertheless, various mechanisms have been identified that seem to be of relevance for at least a part of the patients[3, 23]. At the time of writing, it is unclear whether IBS is a multifactorial disorder (i.e. several pathophysiological mechanisms interact in the same individual and in sum lead to the disorder), or whether it is a summary diagnosis for different etiologies emitting similar symptoms. In this thesis, besides symptom-based stratification, disturbances of the gut-brain axis as mediated by the autonomic nervous system, as well as altered host-microbiota interaction have been focused on, and will therefore be covered in greater depth here.

1.7.1. Gut-Brain Axis

The term Gut-Brain Axis describes the multifaceted bidirectional communication between the GI tract and the central nervous system. In a broad definition it includes the central and autonomic nervous system (which is comprised of the enteric, sympathetic and parasympathetic nervous systems), as well as the hypothalamic-pituitary-adrenal (HPA) axis and the neuroendocrine and neuroimmune system. Often it is extended to also include the gut ecosystem, then termed microbiome-gut-brain-axis[71, 72]. As mentioned initially, at the time of writing FGIDs are considered disorders of gut-brain interaction, which is why a strong focus of IBS research has been put on the functioning of the autonomic nervous system as a proxy measure for overall gut-brain communication, and on the role of gut sensitivity and the role of gut microbiota, as well as structural alterations and aberrant activity of the central nervous system[3, 23].

With visceral sensitivity and altered motility being important observations in IBS patients, the question with regards to gut-brain communication is where these alterations originate, whether in the central processing of interoceptive information and modulation of intestinal function (motility, secretion, permeability etc.), in the afferent and/or efferent transport or generation of information via the autonomic nervous system, or in the peripheral reception and execution.

With regards to central functioning, noteworthy findings have been made related to registration and affect-modulation of sensory information, as well as functional and structural alterations of the brain as reviewed in Enck et al[23]. Likewise, peripheral mechanisms such as enteroendocrine and neuroimmune function[73], as well as changes in microbial composition[74] have shown noteworthy findings.

Altered autonomic function has been observed in IBS patients in several studies[75], but it is a mechanism that is difficult to study. Commonly focus is set on the activity of the parasympathetic and sympathetic nervous system (PNS and SNS) function in IBS, assuming that here the transport of information takes place, and altered communication may occur. One way of estimating this activity is the measurement of heart rate variability, which has been performed in various studies. Given the different study protocols, these studies are difficult to compare, but many have shown aberrations in estimated autonomic function in IBS patients, mostly reflecting reduced PNS or elevated SNS activity[76-82].

The concept of the gut-brain axis currently gives a useful framework aiding to categorize the factors involved here, but to this date there is not comprehensive understanding on how all these factors relate to each other, and which are truly of relevance for the individual patient.

1.7.2. Microbiota

The microbial ecosystem colonizing the gut lumen has experienced a hugely increased research interest in the last years due to the development of new, culture-independent molecular techniques (next-generation sequencing), giving a much better estimation of the identity and quantity of the species present. It is comprised of bacterial and fungal species as well as archaea, protozoa, viruses and bacteriophages, which interact with each other, with food and other gut content, as well as with the host. Current publications have identified over 1000 bacterial species that can colonize the gut lumen[83], and estimate a minimum of 160 of those colonizing an individual's gut[84]. These numbers highlight the high inter-individual variability, which is one major challenge when researching the role of the microbial flora in health and disease. The second major challenge is the difficulty of culturing gut microbial species, limiting our understanding of the physiological function and the mechanisms of host-microbiota interaction.

In IBS the role of the microbial ecosystem seems to be of great importance in symptom generation, and presents a promising therapeutic target[85]. The majority of research on gut microbiota has been conducted on fecal samples, mainly due to the ease of access. Here a growing body of evidence suggests a shift in the microbial composition (termed dysbiosis)[86], but studies show inconsistent results regarding the direction of this shift[87-93]. An association between microbial profiles and brain structures has been reported both in healthy women[94] as well as IBS patients[95], and several studies have associated microbial profiles and dysbiosis to IBS symptoms or symptom severity[91, 96, 97].

It is noteworthy to keep in mind that more or less all of these findings are based on relative bacterial abundance, and not the absolute quantity, which may well be neglectable with regards to fecal samples, but may be of relevance when studying microbiota in tissue biopsies. Fecal samples contain microbial cells of the various species present in the gut lumen, whereas biopsies show mucosa-adherent microbiota of the intestinal tissue, and although the microbial composition in these two compartments differ substantially, the inter-individual variability is higher than the intra-individual variability. Thus, within an individual, fecal microbiota composition is a proxy for mucosa-adherent microbiota composition[96]. It is currently unclear whether the microbial cells themselves or rather their metabolites are of most relevance for host-microbiota interaction, therefore both the fecal microbial composition and the metabolites produced by these species, as well as the identity and quantity of mucosa-adherent microbiota is of high interest for studying this interaction.

1.7.3. Epithelial barrier and mucosal crosstalk

The gut epithelium is a large surface area that constantly is in touch with microbial and fecal material present in the gut lumen. It acts as a semipermeable membrane and represents a barrier between the gut tissue and luminal content. This barrier function is further enhanced by the mucus system of the gastrointestinal tract, which is of high importance for the luminal protection of the gut mucosa[98, 99]. The integrity of the epithelium is maintained by various nervous, enteroendocrine and immune cells[73], and seems to be modified by the gut microbiota[100, 101]. Increased intestinal permeability, referring to a reduced integrity of the epithelial cell composite and the tight junctions linking cells of the epithelial lining, seems to

be frequently found in IBS patients[102-104]. This may be due to reduced expression of tight junction proteins in IBS patients[105], as well as bacteria- and proteasome-mediated degradation[106]. These morphological and functional alterations seem to occur in conjunction with low-grade inflammation[106], modified by neurons of the enteric nervous system[104, 107], and may involve genetic and epigenetic factors as well as effects of dysbiosis and food intolerances[102, 108]. At the time of writing the details of interaction between microbiota (fecal or mucosa-adherent) and the host epithelial barrier are insufficiently understood.

1.7.4. Further mechanisms

Further mechanisms that seem to be involved in IBS pathophysiology, at least for a subset of patients, but which are not focus of this thesis, include altered central processing, visceral hypersensitivity, and altered intestinal motility[3, 23]. Genetic factors playing a role in IBS are highlighted by family disease amassment and further supported by genome-wide association studies[51, 109]. The role of food intake and diet-composition in symptom generation seems to be of high relevance for many patients[14, 110, 111], and the importance of biopsychosocial factors for symptom exacerbation, disease management and quality of life has been shown in many studies[12].

1.8. Stratification of complex disorders

Not only IBS, but many common diseases, especially the ones with chronic disease progression, are characterized by a complex etiology and limited treatment success. This makes them one of the key challenges of modern health care, as they are cause for both the suffering of many individuals as well as cause for high healthcare costs. In many of these diseases, subgroups have been described which may actually be distinct disease entities. Noteworthy examples are asthma[112], diabetes[113] and cancer[114], examples which highlight the importance of such stratification for the understanding of the pathophysiological mechanisms and treatment decisions relevant for the individual patient. Some of these stratification approaches defined disease subtypes by treatment response or –non-response, others by differences in molecular pathophysiology[115]. Despite these improvements in the

understanding of complex disorders, these stratification approaches have so far not sufficiently addressed the whole complexity of the respective pathophysiology, and often neglect parts of the multifactorial disease etiology. In addition to this, since most of these stratification attempts are based on –omics data, they are not suitable for current everyday clinical settings. Taken together, this highlights the need for simple and cost-effective stratification approaches for various complex disorders, which can easily be implemented into current clinical practice.

1.9. Network medicine

Network medicine is a fairly new field of biomedical research focusing on the utilization of graph theory to understand relationships between variables of interest[116]. Prominent examples are biological networks based on protein-protein interactions[117] and disease networks[118], which estimate the similarity of diseases or relationships between diseases and underlying biological factors, such as genotype or metabolome[116].

The term 'network medicine' was coined by Albert-Laszlo Barabasi in 2007 in an article titled *Network Medicine – From Obesity to the "Diseasome"*[119], which was published in The New England Journal of Medicine. Since then various applications of this analysis approach have been published, which all follow the same principle, that complex systems can be comprehensibly represented by network plots of linked variables (with the variables termed nodes and the connections between them termed edges), and can this way be better understood. The key consideration is the question how to define and, if necessary, quantify relationships between the nodes of a network. In many fields the relationship between two nodes is defined by physical interaction (for example experimentally determined protein-protein interaction[117]), shared genes or shared metabolic pathways, shared phenotypes, or, especially in psychological research, by social relationships or correlations.

2. Specific aims of the manuscripts included in this thesis

I) Perform symptom-based stratification of a clinical cohort of IBS patients into novel subgroups taking into account the individual severity of a comprehensive set of IBS-related symptoms.

II) Determine the reproducibility of these subgroups and symptom associations in a population-based cohort fulfilling IBS diagnostic criteria, and examine the relationship of these symptom profiles to healthcare utilization

III) Conduct mechanism-based stratification focusing on individual ANS characteristics to identify subgroups of IBS patients with distinct ANS function differentiating them from HC, and explore associations between ANS status and symptoms.

IV) Develop an integrative analysis pipeline suitable for identifying relevant microbiota-host interactions and their association to IBS symptoms.

3. Basic materials and methods

This chapter contains sections providing information about the cohorts, physiological measures and questionnaires used. Figure 3 presents an overview over the main aims as well as materials and methods used in the respective manuscripts.

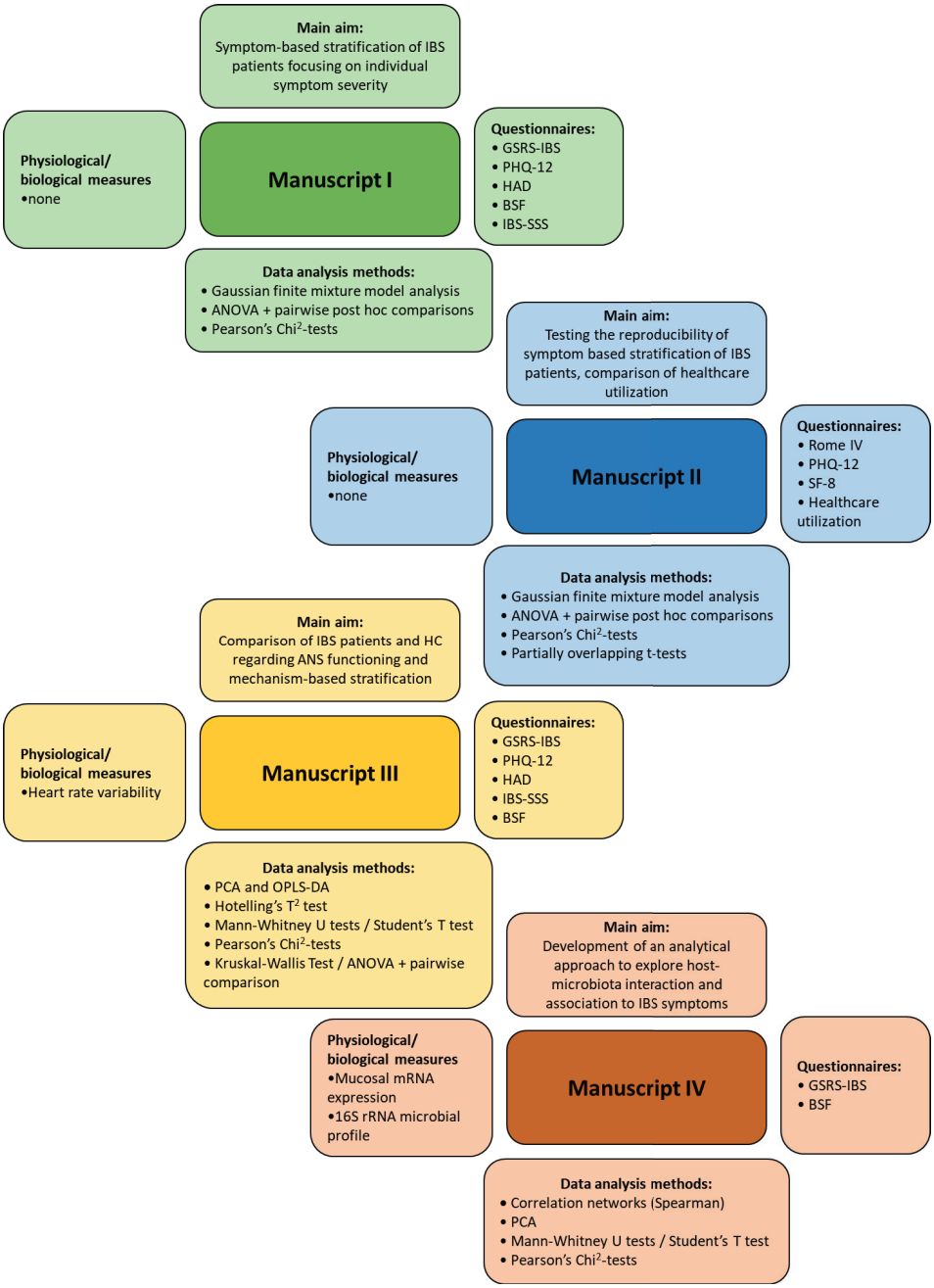


Figure 3: Overview of the main aims, materials and methods of manuscripts I-IV

3.1. Cohorts

This section gives additional information regarding the cohorts of manuscripts I-IV. All studies were approved by either the Swedish Regional Ethical Review Board at the University of Gothenburg (Manuscript I, III, IV: Nr: 48902 and 73109) or the Institutional Review Board of the University of North Carolina (Manuscript II, Nr: 15-2313).

3.1.1. Clinical cohorts

The study subjects of manuscripts I, III and IV were part of extensive clinical and pathophysiological phenotyping studies conducted at our combined clinical and research unit at Sahlgrenska University hospital in Gothenburg, Sweden. IBS patients were recruited at the outpatient clinic from primary care- or self-referrals, whereas HC were volunteers with no history of GI disorders, chronic diseases or current bowel symptoms. Not all participants conducted all tests, therefore the cohorts of manuscripts I, III and IV are not identical. Please see Table 2 for details regarding demographics of the cohorts of the respective manuscripts.

3.1.2. Population-based cohort

The study subjects of manuscript II were participants fulfilling IBS diagnostic criteria based on Rome III[5] and IV[6] of an internet-based health survey conducted by adults of the general population from the US, Canada, and UK. Due to the almost ubiquitous availability of internet access in these countries, this internet-based survey methodology can be utilized to reach a relevant proportion of the general population, and built-in quality checks assure high-quality data without missing values. The demographic characteristics of the subjects are detailed in Table 2.

Table 2: Demographics of the cohorts of manuscripts I-IV

	Manuscript I	Manuscript II		Manuscript III		Manuscript IV	
	IBS patients	IBS based on Rome III	IBS based on Rome IV	IBS patients	Healthy controls	IBS patients	Healthy controls
Number	172	637	341	158	39	42	20
Age (range)	34 (18-60)	46 (18-87)	46 (18-77)	35 (19-64)	29 (19-49)	34 (18-58)	27 (23-41)
Female/male	119/53	420/217	217/124	113/45	22/17	24/18	12/8
IBS-C	23%	17%	28%	24%	-	31%	-
IBS-D	37%	21%	35%	34%	-	36%	-
IBS-M	16%	60%	33%	13%	-	IBS-NonCnonD: 34%	-
IBS-U	24%	3%	5%	25%	-		-

3.2. Questionnaires

Several validated questionnaires were utilized in the respective manuscripts to measure GI and non-GI symptom severity or frequency, and the reasoning for the choice of questionnaire in the respective manuscripts is discussed in the following text.

3.2.1. GI-symptoms

To measure the severity/frequency of GI symptoms we have utilized either the IBS-specific version of the Gastrointestinal symptom rating scale (GSRS-IBS)[60], the IBS severity scoring system (IBS-SSS)[120], or the Rome IV diagnostic questionnaire[6]. Additionally, we have used a Bristol Stool Form (BSF) diary[18], which registers stool form and frequency over a given period of time.

The GSRS-IBS consists of 13 questions measuring different typical IBS symptoms on a scale from 1 to 7, with higher scores representing more severe symptoms[60]. These 13 questions can be used as single items or summarized to five subscores representing key IBS symptoms (abdominal pain, constipation, diarrhea, bloating and satiety). The advantage of the GSRS-IBS is the quite comprehensive measurement of relevant IBS symptoms, making it interesting for detailed analyses. The main disadvantage of the GSRS-IBS is the Likert-scale type data collected, limiting the level of detail measured, but since this data is commonly interpreted as ordinal data, we have deemed it the best choice to represent detailed GI symptom severity for the research questions in manuscripts I, III and IV. In manuscript I we have utilized single

item scores to perform the mixture-model analysis as described in the data analysis chapter, whereas in manuscripts III and IV we have used the 5 sum-scores (with a further detailed analysis in manuscript III), as here we were interested in more generalized symptom associations.

The IBS-SSS is specifically designed to measure the disease burden of IBS on a scale from 0-500 using visual analogue scales (VAS)[120]. Since it puts a strong focus on abdominal pain and does not have validated subscales for different symptoms, we have only used it to provide information about the overall severity of IBS in the descriptive part of manuscripts I and III.

In manuscript II we utilized several questions from the Rome IV diagnostic questionnaire[6] to measure GI symptoms. These questions were selected to match the GSRS-IBS questions as closely as possible, since the latter was not available in the cohort of manuscript II. Differences of the two datasets are discussed below.

The BSF[18] was utilized to subtype IBS patients based on the Rome III criteria[5] in manuscripts I, III and IV. In manuscript I it was furthermore utilized to complement the GSRS-IBS in providing a measure of bowel habit-related symptom severity.

3.2.2. Non-GI symptoms

To measure extraintestinal somatic symptoms we have used the non-GI questions of the IBS patient health questionnaire-15 (PHQ-15)[63], commonly referred to as PHQ-12[64], whereas psychological symptoms were measured using either the hospital anxiety and depression scale (HAD)[121] in the Swedish clinical cohorts or selected questions from the SF-8 health survey[122] in the population-based cohorts.

The PHQ-12 measures symptoms on a scale from 0 = “not bothered at all” to 2 = “bothered a lot”, which are summed up to create a total score[64]. We have used the single items in the Mixture model analysis of manuscripts I and II and the sum score to describe the cohorts of manuscript III.

The HAD contains 14 items, 7 of which are summarized to measure anxiety and 7 to measure depression, each on a scale from 0-21, with higher scores reflecting increasing severity[121]. Commonly a cut-off of 11 points is used to define clinically relevant anxiety or depression. We

have used these sum-scores to quantify both anxiety and depression in manuscripts I and IV, while also using the above mentioned cutoff for clinically relevant scores in manuscript III.

For manuscript II we selected three questions from the SF-8[122] to match as closely as possible the anxiety and depression sum scores of the HAD, with the main difference being that the individual SF-8 items provide a smaller scale range than the overall HAD scores (from 1 (not at all) to 5 (extremely) vs. 0-21)[122] and that this questionnaire asks for the severity of emotional problems without making a distinction between anxiety or depression, which needs to be taken into account when comparing the results of manuscripts I and II.

3.2.3. Differences in symptom registration in manuscript I and II

While the questions measuring the severity of GI and non-GI symptoms in manuscript II were chosen to replicate those used in manuscript I as closely as possible, not all symptoms were equally represented in the two datasets. Table 3 gives an overview over similarities and differences in the two datasets. The most noteworthy difference is that in the population-based cohort a question specifically asking for diarrhea (frequent stools) is missing.

4.3. Physiological measures and laboratory analyses

This section discusses the physiological measures used to assess ANS function in manuscript III as well as the laboratory analysis of mucosal biopsies used in manuscript IV.

4.3.1. Heart rate variability

Given the anatomical location of the parasympathetic and sympathetic nervous system (PNS and SNS), it is challenging to measure the respective activity non-invasively. Measuring an individual's heart rate variability (HRV) is therefore a commonly used method to indirectly measure and estimate the activity and/or reactivity of an individual's ANS function[123]. HRV is a physiological phenomenon describing variations in the intervals between single heart beats, and can also be referred to using the terms 'cycle length variability', 'RR-variability' or 'NN-variability' (with N referring to normal heartbeats). It is based on heartbeats measured by

Table 3: Overview of symptoms measured in the cohorts of manuscripts I and II.

	Clinical cohort	Population-based cohort
Abdominal pain (severity)	X	X
Abdominal pain (frequency)		X
Abdominal pain associated to meal		X
Abdominal pain relieved by defecation	X	
Bloating	X	X
Passing gas	X	
Constipation (infrequent stools)	X	X
Hard stools	X	X
Abdominal distention	X	
Incomplete bowel emptying	X	X
Early satiety	X	
Fullness	X	
Diarrhea (frequent stools)	X	
Loose stools	X	X
Urgency	X	X
Stool form	X	
Stool frequency	X	
Straining		X
Back pain	X	X
Joint pain	X	X
Menstrual cramps or other problems with the period	X	
Intercourse related pain or problems	X	X
Headaches	X	X
Chest pain	X	X
Dizziness	X	X
Palpitations	X	X
Shortness of breath	X	X
Trouble sleeping	X	X
Tiredness	X	X
Anxiety	X	
Depression	X	
Overall emotional/physical health		X
Anxiety/depression/stress		X
Overall emotional/physical health		X

Symptom measured in both cohorts
Symptom measured only in the clinical cohort
Symptom measured only in the population-based cohort

electrocardiogram, and takes only consecutive normal heartbeats into account, excluding any extra-systoles or similar events[123].

The following limitations are inherent to this method: HRV is an indirect estimation of ANS activity, and may therefore not fully reflect the overall ANS activity. Also, many variables reflect multiple aspects of the ANS and not single entities (see Table 1 in manuscript III), making it challenging to fully interpret the results. In our study the data was derived from a 24-h ECG, which also has some limitations to consider, mainly potentially insufficient connection of the electrodes over the course of the day, leading to noisy registration. Also it is important to acknowledge that the 24h of the study do not necessarily reflect the IBS patient's ANS status over time.

Nevertheless, we have carefully designed our study to overcome these limitations where possible. Despite HRV being an indirect measure of overall ANS activity/reactivity, several studies have confirmed its suitability as a proxy measure[123], also with regards to GI-related ANS activity[124, 125]. Additionally, since all study participants underwent the exact same protocol, a comparison of HRV measures between IBS patients and HC is useful. We have chosen multivariate approaches taking all registered variables into account at once, and thus creating a comprehensive representation of the individual's ANS activity, reducing the importance of underlying mechanisms of the single variables for result interpretation, and instead acknowledging the importance of these different factors for the overall analysis. The ECG data was analyzed by the same (blinded) investigator, ensuring that only correctly registered normal heartbeats were utilized for HRV calculation, and segments with artefacts were excluded from the analysis.

Given the time-intensive ECG analysis and HRV calculation as well as the inconvenience of wearing an ECG for an extended period of time, limiting the registration to 24h is a commonly used timeframe, but regarding future studies a follow-up HRV measurement after a given amount of days or weeks might be worth considering.

Taken together we have to the best of our possibilities acknowledged and overcome the major limitations of HRV registration and are confident that the described results are of high quality.

4.3.2. Mucosal biopsies

In manuscript IV, mucosal biopsies collected during an unprepared sigmoidoscopy were used to measure both host mRNA gene expression of selected genes, as well as abundance of mucosa-adherent microbiota. The sigmoid colon was chosen due to ease of access compared to other parts of the colon, and was additionally considered favorable since the vast majority of measurements of visceral sensitivity in the study protocols (a central aspect in the study protocols, although not included in this thesis) were conducted in the recto-sigmoid region, therefore making it easier to relate potential findings to this. Since we were aiming to measure the condition and composition of microbiota as similarly as possible to in-vivo conditions, bowel preparations normally used for diagnostic endoscopies were skipped to preserve these conditions.

4.3.3. Host gene expression

For the pilot study conducted in manuscript IV the mRNA expression level of several host genes relevant for mucosal integrity were included. The mRNA expression analysis was performed using quantitative reverse transcription PCR, and normalized to multiple housekeeping genes to improve standardization and to achieve high-quality inter-plate validation. The analyzed genes were selected based on information as to which genes were likely to be of interest for IBS pathophysiology derived from literature searches, and were aimed to be a selection of genes relevant for microbial recognition, local immune-activity, mucosal permeability and enteroendocrine activity[73, 105, 126-128]. The chosen gene-repertoire is therefore suitable for the conducted pilot analysis aiming to estimate the usefulness of the analysis pipeline presented in manuscript IV. Nevertheless, it is important to consider that this preselection is the strongest limitation to the results of the integrative analysis, meaning that a follow-up study requires a broader repertoire of mucosal gene measures, for example using multiple arrays or whole transcriptome shotgun sequencing, to create a more comprehensive dataset of the potential mediators of host-microbiota interaction.

4.3.4. Gut microbiota assessment

As mentioned in the introduction, due to strong difficulties in culturing gut microbial species, the relatively new next-generation sequencing techniques have become the standard method for the identification of gut microbial species. Currently the most common technique is based on 16S-rRNA based counts of operational taxonomic units (OTUs), which is used to estimate the abundance of various gut microbial species. This has proven to be a useful method to identify the respective species in a given sample, but despite of its several advantages, this technique has some limitations to consider. OTU counts provide only information regarding the relative amount of species in a sample, and the sequencing step itself is affected by systematic variability, such as differences in sequencing depth between the given samples. The data therefore requires an additional normalization step preceding any data analysis, which is commonly done by either calculating the relative abundance of species in a given sample, or through rarefication. Both of these methods have disadvantages regarding suitability for statistical analyses, and have been widely criticized[129-131].

We have therefore utilized a newer method primarily developed for whole transcriptome shotgun sequencing data, which is based on the ratio of counts and the sample geometric mean and available in the R-package DESeq2[132]. It is important to mention that the method has been shown to be useful also for 16S rRNA data[129]. This method normalizes next-generation sequencing data by estimating a "pseudo-reference" based on the data available, which is then used to normalize the reads to this reference[132]. The method was designed to be robust against outliers, and shown to perform well under most conditions[130], making it suitable for the analyses performed in manuscript IV.

5. Data analysis approaches

This chapter contains detailed discussions of the statistical methodology used in the respective manuscripts, focusing on the reasoning behind the applied analytical strategies.

5.3. Symptom-based stratification: Considerations and methodology

In manuscripts I and II we have aimed to stratify patients into clinically relevant subgroups based on a comprehensive set of IBS-related symptoms.

Stratification is the process of sorting data into distinct groups or clusters. It is probably the most common aim of data mining projects, and is applied in many fields such as computer sciences, image analysis and data compression, but also commonly used in bioinformatics and biomedical data analysis. Given the broad variety of datatypes and –structure as well as tasks required in these different fields, various stratification techniques have been developed. The most important of these are the more classic statistical clustering algorithms, as well as the rather novel machine learning techniques. In order to choose the most adequate stratification method several considerations need to be taken into account:

1) The type of data (categorical, ordinal, continuous…)
2) Data distribution
3) Sample size
4) Research question
5) Previous knowledge: are supervised analyses possible?

Type of data: For choosing the appropriate analysis method, the following was considered in the manuscripts I and II. In medical research the presence and severity of symptoms is commonly measured using validated questionnaires, and the questionnaires used in these two manuscripts use Likert scales to assess symptom severity. The data derived from these questionnaires therefore must be considered as either categorical or ordinal, with the latter being the most common interpretation of this data.

Data distribution: In the clinical patient cohorts used in this thesis, a common observation was a skewed distribution of the data the symptom measures. This may be interpreted as a hint towards the presence of different subgroups in this cohort, but can also be interpreted as fluctuations of symptom severity throughout an individuals' disease progression. In both manuscripts I and II the data was therefore logarithmically transformed for the multivariate analyses to ensure that the assumptions for the applied tests were fulfilled.

Sample size: This was the most important factor in choosing the appropriate stratification technique. The data used here was collected as part of a very extensive phenotyping study, which limited the amount of patients included in the time available for data collection. While data mining algorithms usually are optimized for very large datasets (number of observations >> number of variables), the dataset used in these manuscripts was characterized by high dimensionality and a rather low number of individuals. This was especially relevant in manuscript I, in which the analytical approach was developed.

Research question and previous knowledge: As mentioned, the aim of these manuscripts was to stratify patient cohorts into novel subgroups. That means that the ideal number of subgroups was unknown, and no previous knowledge was available which could have been used as a suitable template for stratification procedures as conducted in many common machine learning procedures such as neural networks, decision trees or support vector machines.

In conclusion, the ideal analysis strategy was required to be able to work with ordinal data, be suitable for small-but-high-dimensional datasets, and conduct unsupervised stratification into an optimized number of subgroups. This suggests unsupervised clustering algorithms, especially connectivity-based such as hierarchical clustering[133], centroid-based such as k-means clustering[134], distribution-based clustering such as Gaussian mixture models as well as density-based clustering such as Density-based spatial clustering of applications with noise (DBSCAN)[135].

The first two, connectivity- and centroid-based clustering, are methods that perform a hard split of the dataset into groups, which may give spurious results in smaller datasets. They also require the user to choose or predefine the number of clusters, making them not optimally suitable for the research question at hand. Most k-means algorithms also assume equal-sized

clusters, an assumption not in line with our research question. Density-based clustering only links observations that satisfy a specified density criterion and classifies all others as outliers, which is also not in line with our dataset and research question.

Gaussian mixture models, our method of choice, has several benefits. It defines the number of clusters and group membership using an expectation–maximization (EM) algorithm[136, 137], a likelihood estimation based on a Bayesian information criterion (BIC)[138] in which clusters are at first randomly assigned and then iterated until the optimal fit is reached. This holds the risk for overfitting, which can be minimized by limiting model complexity. This likelihood calculation reduces the risk for spurious findings, making it the most suitable approach for the dataset at hand, as well as for the task of finding an unknown number of novel subgroups.

The subgroups resulting from this stratification analysis were then characterized by the severity of their symptoms as visualized in Figure 4. To make this easier to compare, the respective symptom severity was described relative to the cohort average, resulting in a description of the individual symptoms of the group's symptom profiles as '*above-average'* or '*below-average'* symptom severity.

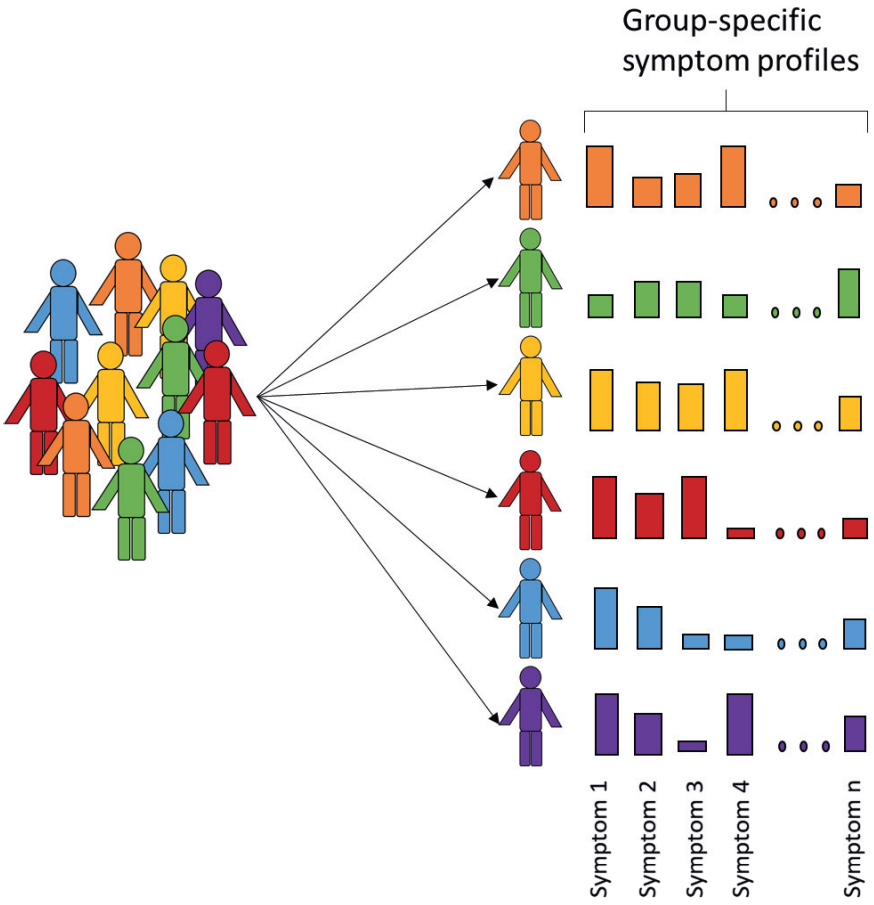


Figure 4: Gaussian mixture modelling was used to stratify the cohort into several subgroups characterized by specific profiles of differing symptom severity

5.4. Mechanism-based stratification: Considerations and methodology

In manuscript III we aimed to generate a representation of the overall ANS status (as derived from measures of heart rate variability (HRV)) in individual study participants in order to understand the similarities and differences between individual patients, as well as patients and HC. The rationale for this approach is explained in detail in manuscript III.

Given the complexity of the ANS, several distinct HRV measurements are commonly generated, reflecting different aspects of ANS function[123]. This complexity can presumably be better understood using multivariate analytical methods rather than univariate comparisons, as this way a simultaneous analysis of all HRV measures is achieved. Such a simultaneous analysis is likely better suited to estimate the overall ANS functioning of the individuals, which can then be compared between the respective individuals.

The data available for manuscript III consisted of continuous, non-parametric data with a rather low degree of noise. To answer the research question of this study, we decided to apply an unsupervised exploratory multivariate analysis, the principal component analysis (PCA), as well as a supervised multivariate analysis, the orthogonal partial least squares discriminant analyses (OPLS-DA). The PCA is a dimensionality reduction method used to summarize the underlying data structure of a multidimensional set of variables to a low-dimensional summative model, which is useful to explore a dataset and extract information that would otherwise not be human-readable in the raw data. This makes the PCA an ideal tool to create such a summative model of the individual ANS function (termed global HRV in the manuscript) of our study subjects by simultaneously using all relevant HRV variables and plotting this information in the two-dimensional space, thus transforming this information into an interpretable format. Since the exploratory PCA as well as a comparison of the multivariate means of the dataset hinted towards differences between IBS patients and HC regarding their global HRV, we further conducted an OPLS-DA in a secondary analysis to better understand the factors underlying these differences. An OPLS-DA is, simply explained, similar to a PCA, with an additional regression step adding a vector which contains information about which group an individual observation belongs to, in our case whether they are IBS patients or HC. It is a supervised analysis, used to highlight the variance in the data which is most discriminating between the groups. This enabled us to estimate whether differences between IBS patients and HC were rather due to specific variables, or to some IBS patients' global HRV differing from that of HC, which would enable us to stratify the IBS patient group based on their global HRV.

5.5. The stepwise integrative analysis pipeline

In manuscript IV we aimed to develop an analytical approach enabling us to link measurement of different potential pathophysiological mechanisms as well as the IBS patients' clinical phenotype represented by the severity of key IBS symptoms. One way of approaching this task is to view it as a multilayer problem, resulting in two questions to be answered: i) How do the variables within each layer relate to each other, and ii) how do the variables between each layer relate to each other?

For the presented manuscript we have chosen to limit the complexity of this task to three layers and perform a pilot study to test the feasibility of our approach. These three layers were gut microbiota, mucosal mRNA gene expression and key IBS symptoms. This leads to the more specific research question as to how host and microbiota interact, if there are differences between IBS patients and HC, and which of these variables relate to symptoms (Figure 5). Conceptually, this results to the following dual-layer visualization: i) Mucosal layer: The variables of the gut microbiota and host gene expression layers are summarized into one combined layer, the mucosal layer, representing the interaction between microbiota and host cells taking part, and ii) Symptom layer, which represents key IBS symptoms.

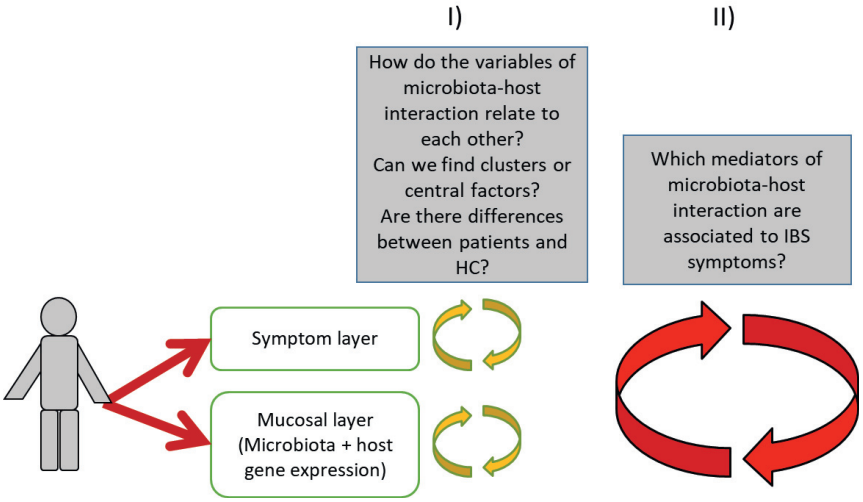


Figure 5: The specific focus of manuscript IV

To understand the relationship between the different variables, we have chosen to apply network analyses, a method developed to visualize relationships between variables and highlight patterns that are over-represented relative to the overall network structure, such as clusters of variables. In case of biological or biomedical data, such patterns could hint towards a functional connectedness of the respective variables and thus highlight potential key mechanisms, making this a very suitable analysis approach for the given research question.

Key considerations regarding the conduction of network analyses are i) how to quantify the grade of connection between two variables, and ii) how to limit the level of detail in order to not obscure important findings, with both considerations being dependent on the data available.

We have used the strength of correlation between the respective variables as given by a Spearman's rank correlation to quantify the grade of connection between the respective variables in the pilot study presented in manuscript IV. Correlation is a good approximation of the grade of connection in this dataset, as in biological systems connected mediators commonly either up- or downregulate each other, a phenomenon quite well represented by the degree of correlation.

To limit the level of detail we have developed an approach as visualized in Figure 6. The gut microbiome as well as a host cell transcriptome or proteome usually contain several hundred to several thousand mediators, which, if taken into account in raw form, would produce an overwhelmingly large network. The approach suggested here therefore is conceptualized as a stepwise analysis pipeline (Figure 6), which starts off with an initial summarizing or clustering step. Since at this time point very little is known about the physiological functioning of most species in the gut microbiome, we did not have the option of applying clustering strategies resulting in functionally related groups, but instead chose to summarize the microbiota based on phylogeny, thus summarizing the available genera into their respective phyla. When it comes to the host cell mediators more is known about their biological function, meaning that these could be stratified into functional groups before the integrative analysis step. In the pilot study presented in manuscript IV, we only had data available for a few of these mediators, which is why in this manuscript we have abstained from summarizing these.

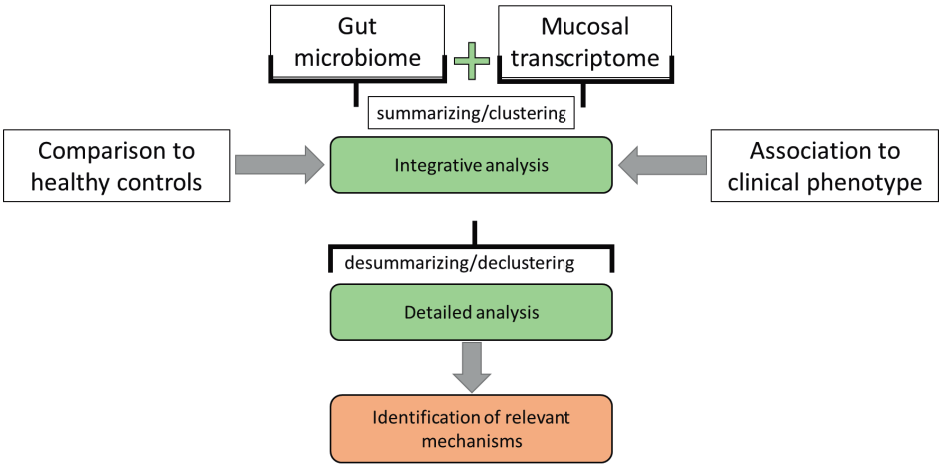


Figure 6: Analysis pipeline as proposed in manuscript IV

The main step in this analysis strategy, the integrative analysis, is comparatively simple in its execution. A network based on a correlation matrix of both summarized microbiome data as well as host cell mediators is conducted in both IBS patients and HC, and compared. The resulting visualization of relationships between mediators of the mucosal layer in IBS patients is additionally associated to key IBS symptoms, and visualized as a dual-layer network graph as demonstrated in manuscript IV. The aim of this step is to identify "patterns of interest", i.e. relationship patterns that differ between IBS patients and HC, which hint towards different mechanisms in microbiota-host interaction, as well as to understand which patterns are associated to clinical symptoms. Both steps together ideally result in a better understanding of which mediators may potentially be of relevance in IBS pathophysiology.

Since the initial summarizing/clustering step has the drawback of some loss of information, it is succeeded by the final step, a detailed analysis. The idea here is to separately analyze the previously identified patterns of interest by using unsummarized variables, but for this step only those that are part of the summary variables. This way a detailed understanding of the relationship patterns is achieved, without the drawback of getting lost in an overload of details.

6. Results and Discussion

This chapter highlights the most important results obtained in the four manuscripts and aims to put them into context to the previously stated purpose of this thesis. While a detailed discussion of the respective results has been performed in the corresponding manuscripts, the focus here lies in discussing the results in relation to the purpose of this thesis.

6.3. Symptom-based patient stratification

Manuscript I focused on symptom-based stratification of IBS patients into novel subgroups, while manuscript II explored the reproducibility of these subgroups in a population-based cohort and compared the groups regarding healthcare utilization. Figure 7 summarizes the main results of manuscripts I and II.

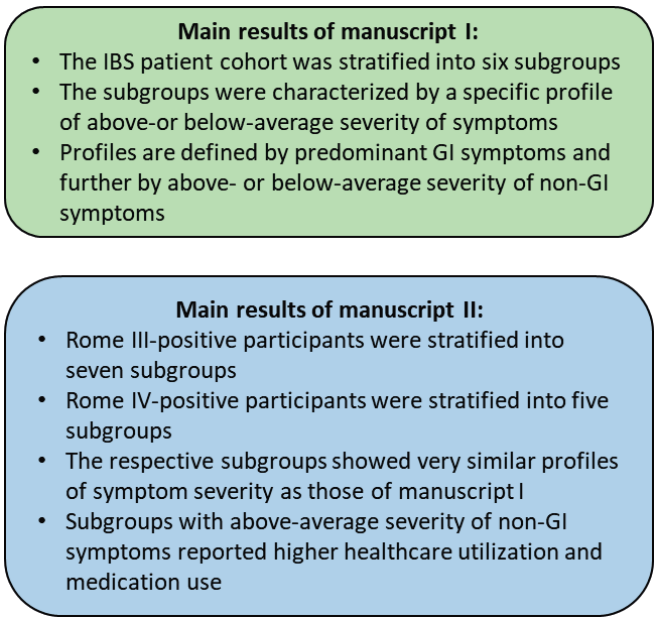


Figure 7: Main results of manuscript I and II

The symptom-based patient stratification conducted in manuscript I yielded a six-subgroup solution, whereas in manuscript II a seven-subgroup solution showed the best statistical fit when analyzing Rome III-positive participants, and a five-subgroup solution when analyzing Rome IV-positive participants.

Of the groups in the clinical cohort, defined by the Rome III criteria for IBS[5], and the Rome III-positive population-based cohort, four groups were characterized by either predominantly constipation-related (two groups) or diarrhea-related symptoms together with above-average severity of pain (two groups), with the further distinction made by above-average or below-average severity of non-GI symptoms. Furthermore, two groups were characterized by an unspecific profile of GI and non-GI symptoms, one group showing high severity of most of these symptoms, and one group showing a profile of overall mild severity of all symptoms. The Rome III-positive population-based cohort showed an additional seventh subgroup which was characterized by above-average severity of psychological symptoms and overall mild severity for the remaining symptoms (Table 4).

In the population-based cohort subgroups which showed elevated severity of non-GI symptoms reported higher healthcare utilization and medication usage than the subgroups with below-average non-GI symptoms.

Table 4: Overview of the corresponding groups identified in manuscripts I and II

Clinical cohort	Rome III-positive population-based cohort	Rome IV-positive population-based cohort
Constipation-low comorbidities	Constipation-low comorbidities	Constipation-predominant
Constipation-high comorbidities	Constipation-high comorbidities	
Diarrhea-low comorbidities	Diarrhea-low comorbidities	Diarrhea-predominant
Diarrhea-high comorbidities	Diarrhea-high comorbidities	
Overall mild symptoms	Overall mild symptoms	Overall mild symptoms
Mixed-GI-high comorbidities	Mixed-GI-high comorbidities	Mixed-moderate psychological symptoms
		Mixed-high psychological symptoms
-	Psychological symptoms	-

The group-specific symptom profiles of the six corresponding groups derived from these two Rome III cohorts overlapped to a large extent, as visualized in Figure 8. The only exception

was the mixed GI-high comorbidities group, which showed higher relative extraintestinal somatic symptoms in the population-based cohort.

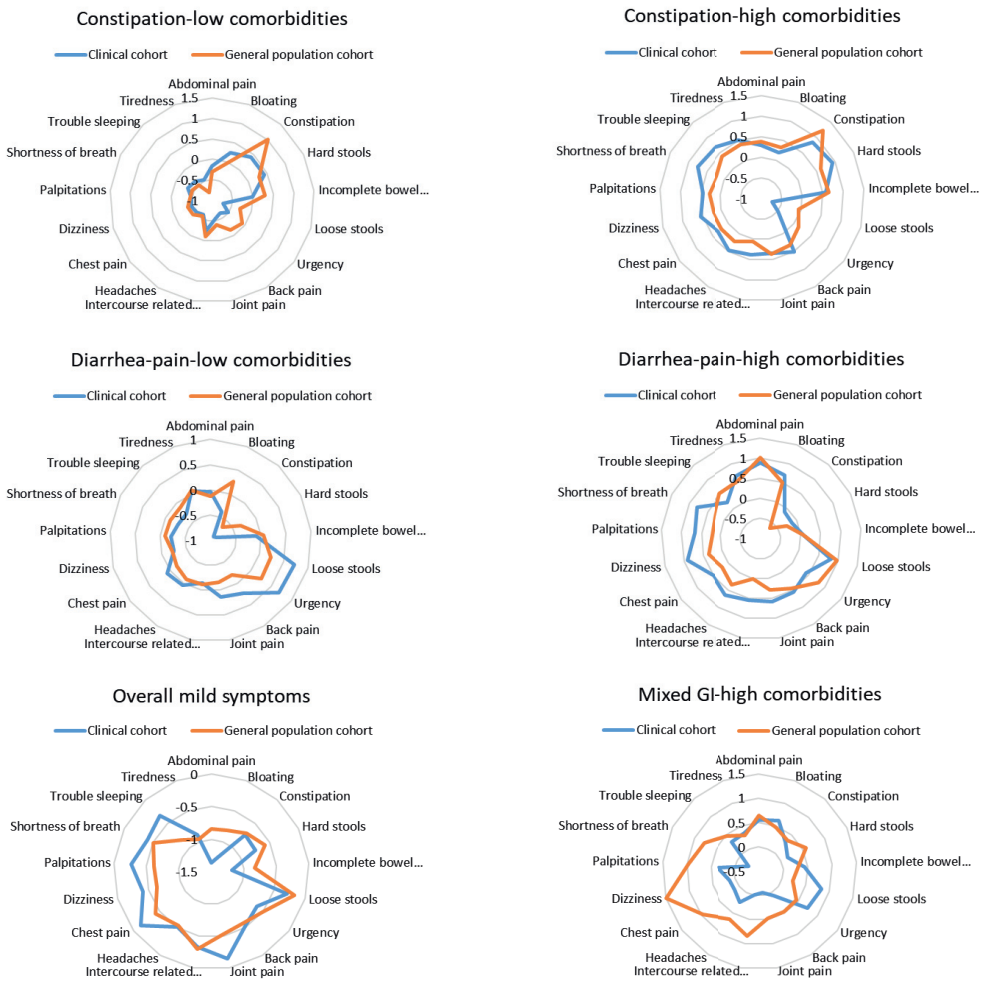


Figure 8: Comparison of group specific symptom profiles of the corresponding groups in the clinical Rome-III positive cohort (blue) and the Rome III-positive population-based cohort (orange). Symptom measures are calculated as Z-scores (respective group average set to zero and standard deviation set to one).

This high degree of correspondence between the subgroups identified in these independent Rome III-positive cohorts is a strong argument for the validity of our subgrouping solution. It

shows that the symptom associations shown here are mostly constant, and that the relevance of non-GI symptoms differs for individual IBS patients. This approach may therefore be useful to develop an effective way of implementing a more comprehensive clinical assessment into clinical practice.

The five subgroups identified in the Rome IV-positive population-based cohort showed a similar tendency but less distinct symptom profiles (Table 4). One group was characterized by predominantly constipation-related symptoms and average severity of non-GI symptoms, and one by predominantly diarrhea-related symptoms in association to above average severity of pain (Figure 9). The remaining three groups were characterized by an unspecific mix of GI and non-GI symptoms of varying severity, especially of the psychological symptoms (Figure 10).

When comparing the constipation- and diarrhea- predominant subgroups identified in the Rome IV population-based cohort to those in the Rome III population-based cohort (Figure 9), the Rome IV-subgroups showed a profile that was in between the two respective Rome III-subgroups, especially with regards to non-GI comorbidities, where some had symptom severities above and some below average.

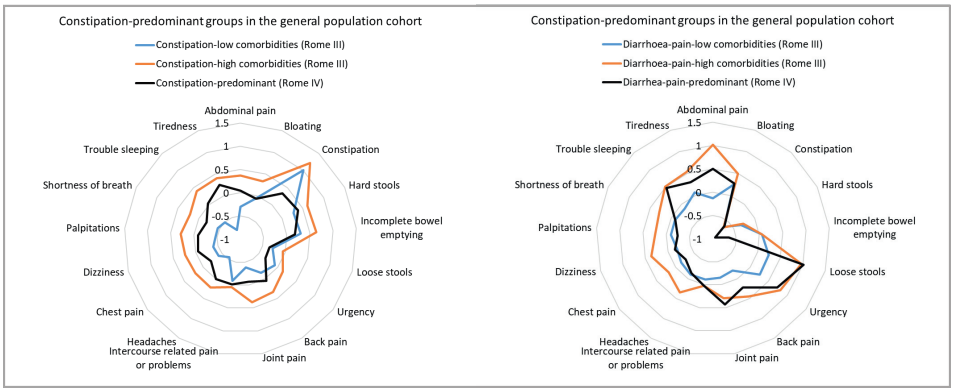


Figure 9: Comparison of constipation- and diarrhea-predominant subgroups in the Rome III-positive (orange and blue) and Rome IV-positive (black) population-based cohort. Symptom measures are calculated as Z-scores (respective group average set to zero and standard deviation set to one).

This may, as discussed in detail in manuscript II, be explained by an overall higher severity of symptoms, including the non-GI symptoms, and furthermore supports the relevance of

implementing a comprehensive symptom registration into clinical routine, as suggested previously[26, 27].

The three mixed-symptom groups of the Rome IV cohort are compared in Figure 10. The overall mild symptoms subgroup differs clearly from the other two groups, which show very similar symptom profiles with a moderate difference in the severity of psychological problems and a large difference regarding intercourse related pain.

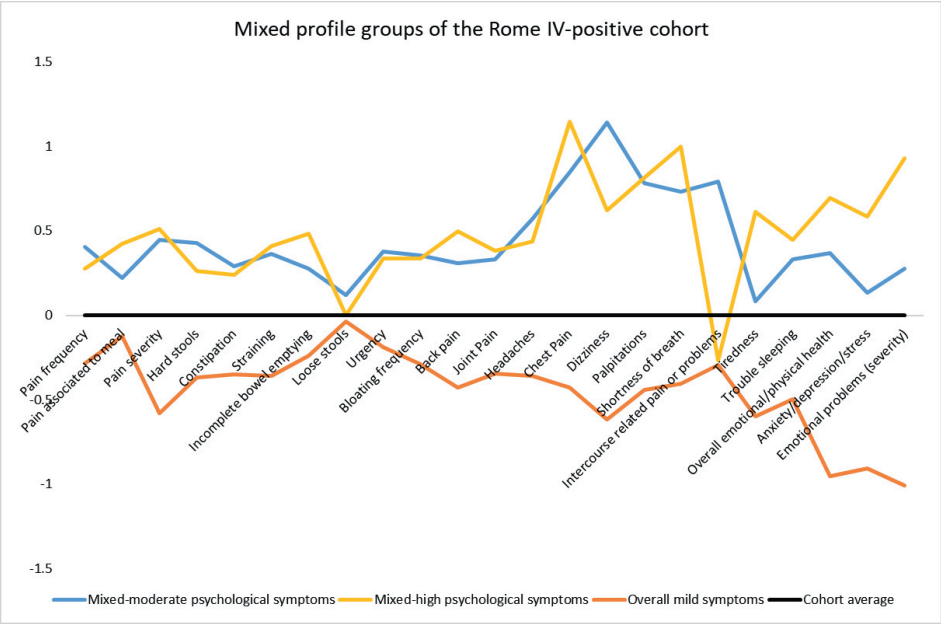


Figure 10: Comparison of the three mixed profile groups identified in the Rome IV-positive cohort. Symptom measures are calculated as Z-scores (respective group average set to zero and standard deviation set to one).

A subgroup of IBS patients reporting overall mild symptoms was present in all of the analyzed cohorts, also in the clinical sample. It is unclear whether this group represents a status of remission, or just represent those on the lower end of symptom severity, or if these are subjects that tend to underreport symptoms when using questionnaires.

Since the subgroups in these cohorts are based on statistical likelihood, it is also interesting to compare the goodness of fit (meaning how well the model matches the data, shown in Figure 11), an aspect not covered in the manuscripts. The clinical cohort had the highest

Bayesian information criterion (BIC, used to mathematically estimate the goodness of fit), followed by the Rome IV-positive population-based cohort. The Rome III-positive population-based cohort had a noticeably less good fit. This observation may give us additional information on the structure and especially the heterogeneity of the three cohorts.

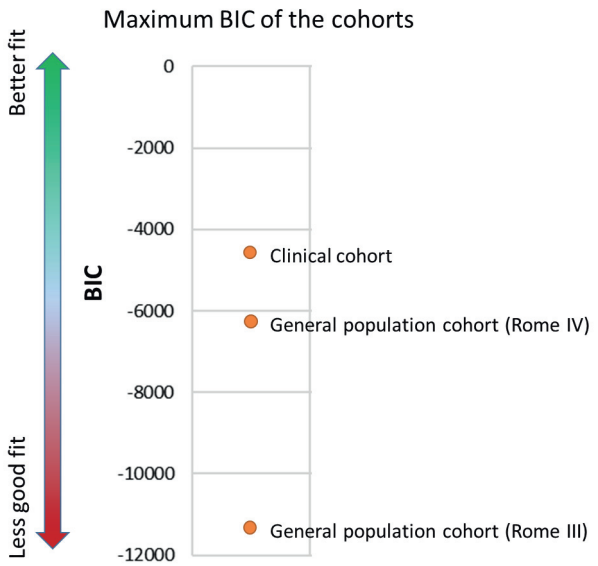


Figure 11: Comparison of the statistical fit of the mixture model solution in the three cohorts. BIC=Bayesian information criterion.

While all of the groups are heterogeneous, as also indicated by the negative BIC values (-4529.2 to -11335.2), the clinical cohort had the best statistical fit regarding the subgrouping solution. This may be expected, as the clinical cohort consists of well characterized IBS patients, and each IBS diagnosis was confirmed by an experienced gastroenterologist. When it comes to the population-based cohort it is noteworthy that the Rome IV-positive cohort showed a far better statistical fit than the Rome III-positive cohort, which, together with the differences in the resulting subgroups, indicates that the updated and stricter diagnostic criteria may have substantially altered the structure and heterogeneity of IBS cohorts in general.

Taken together, the mixture model analysis of a comprehensive set of symptoms has shown to reproducibly identify subgroups with specific symptom profiles in Rome III-defined IBS patients. These symptom profiles related to the extent of healthcare needs of the respective IBS patients, and may therefore be useful in clinical practice to aid the clinician in making individual treatment choices.

How well these subgroups relate to pathophysiology and treatment response needs to be evaluated in future studies. Recurring symptom associations such as the association of above average severity of diarrhea and pain, which was found in all cohorts, may point towards common denominators and are therefore important to take into account when searching for pathophysiological mechanisms.

The newly introduced Rome IV diagnostic criteria seem to have altered the composition of the IBS cohort, with future studies necessary to elucidate in detail the impact these changes have on the cohort and clinical practice.

Regarding the purpose of this thesis, a symptom-based stratification has shown to be useful to better understand the clinical presentation of these very heterogeneous cohorts, and highlight recurring symptom associations. By focusing on a comprehensive set of simply measurable symptoms, this approach can be transferred to clinical practice rather easily. It shows that the relevance of specific symptoms varies between individuals, but is similar between many individuals, thus enabling the definition of several subgroups with specific profiles. This approach can also be applied to similar diseases or research questions, and thus improve our understanding of the clinical presentation of complex disorders.

6.4. Mechanism-based patient stratification

Manuscript III compared IBS patients and HC regarding their ANS function as represented by the individual heart rate variability. Figure 12 summarizes the main results of manuscript III.

Main results of manuscript III:

- Univariate comparisons between all IBS patients and HC showed significant differences during daytime and orthostatic stress regarding ANS function
- A comparison of the multivariate mean showed significant differences between all IBS patients and HC
- A subset of IBS patients showed an altered ANS function, which distinguished them from HC as well as the remaining IBS patients

Figure 12: Main results of manuscript III

An overall univariate comparison of IBS patients and HC showed significant differences for HRV variables obtained during daytime and during orthostatic stress, but not during nighttime or in supine position. This is noteworthy, as the ANS is likely more "challenged" during daytime, as well as during orthostatic stress. Similarly, an overall comparison of both groups comparing all variables simultaneously showed significant differences. Since both of these comparisons were conducted on ranks or averages, these comparisons do not answer the question whether these findings are due to an altered HRV function in all patients or only in a subset, which is why we have further conducted an OPLS-DA to perform a mechanism-based stratification of the IBS patient cohort.

This mechanism-based patient stratification yielded two subgroups. One group showed a global HRV profile which was in the same range as those of the HC ('*healthy-like',* Figure 13), and the other, smaller group of IBS patients showed a global HRV profile that differed clearly from that of the HC, as well as from the other IBS patients ('*non-healthy-like',* Figure 13). This subgroup with non-healthy-like global HRV differed significantly from HC on all single HRV measures (Figure 14), whereas no statistically significant difference was found between HC and IBS patients with healthy-like global HRV. This may suggest that the differences seen in the initial comparisons are rather due to the IBS-patients with non-healthy like global HRV, and not the whole IBS patient cohort.

When comparing key IBS symptoms between IBS patients with healthy-like and those with non-healthy-like global HRV, the IBS patients with non-healthy-like global HRV suffered from more severe diarrheal symptoms than those with a healthy-like global HRV.

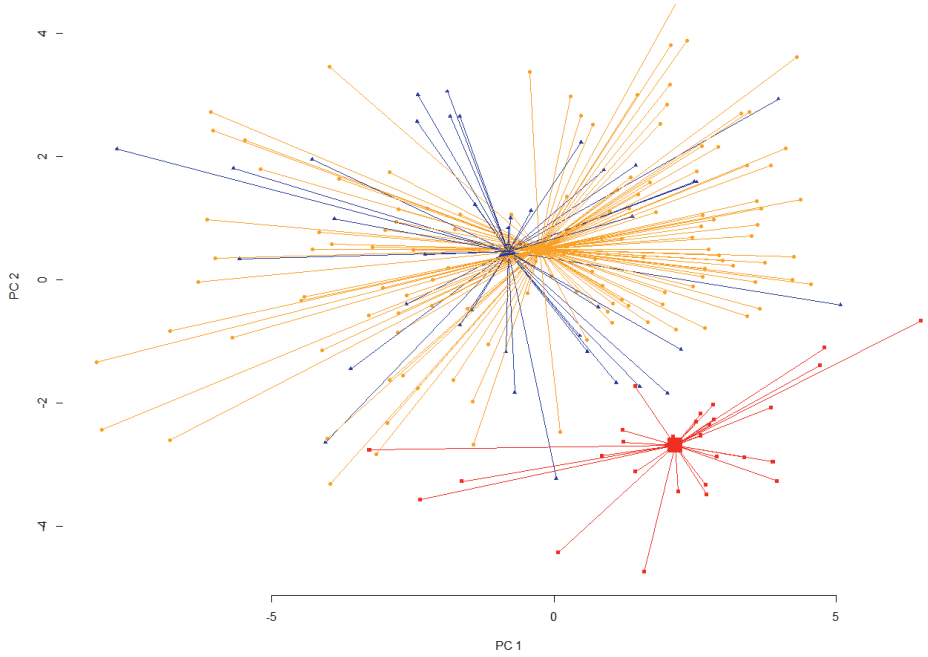


Figure 13: PCA plot of the global HRV profile of HC (blue), patients with healthy-like global HRV (orange) and non-healthy-like global HRV (red). The centroids mark the respective multivariate mean of the three groups.

This mechanism-based stratification approach has therefore shown to be useful to identify a subgroup of IBS patients characterized by aberrant measurements in all of the variables used for this analysis.

The principle of this stratification approach, using a well-selected group of HC to define what is a normal range and what is not, has been used for example for many laboratory measurements such as markers for liver or kidney function or inflammatory markers, which are commonly used in everyday clinical routine. We have here extended this principle by comparing a combination of variables simultaneously, which may be an advantageous approach when investigating complex mechanisms such as ANS function, and can surely be applied for other complex mechanisms such as brain function or hormonal status, given a careful selection of relevant variables. In complex syndromes such as IBS, assuming different

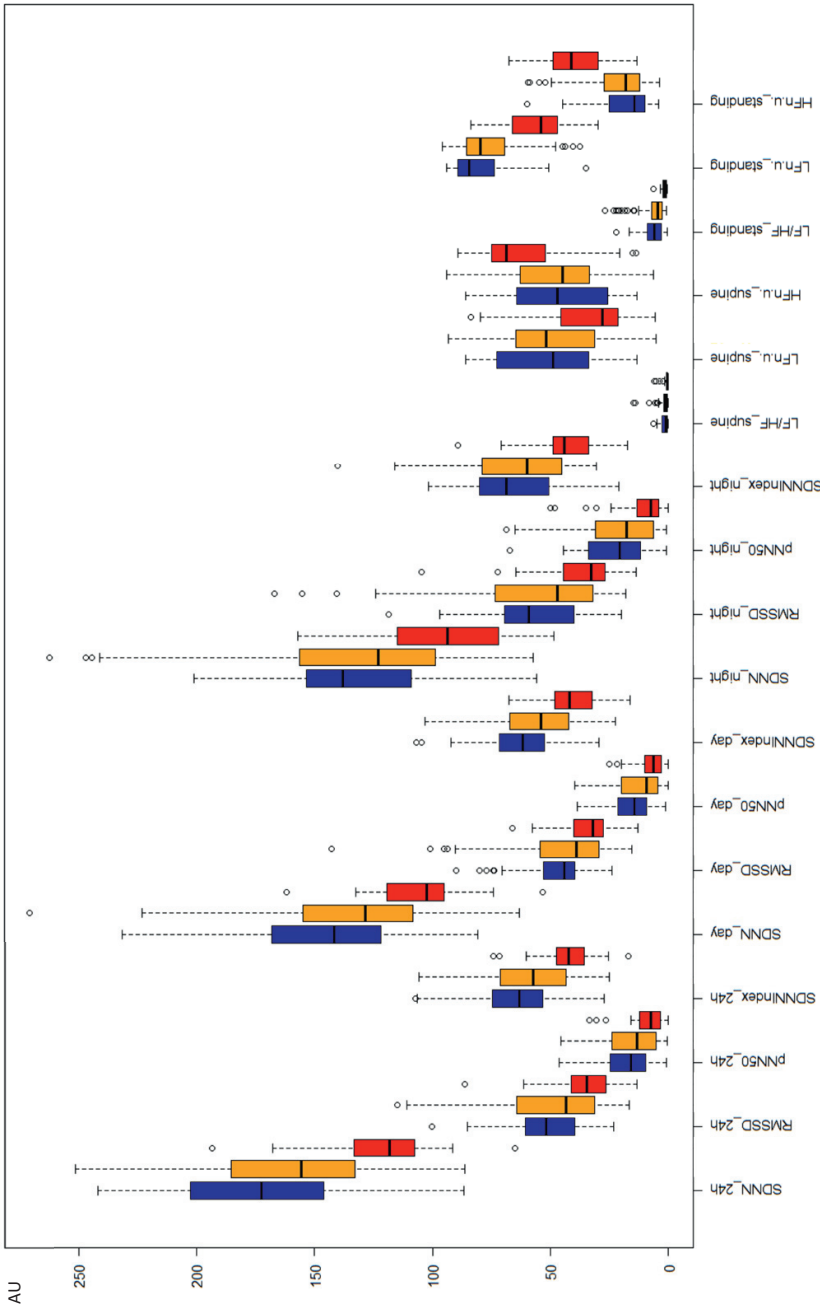


Figure 14: Comparison of single HRV measures between HC (blue), IBS patients with healthy-like global HRV (orange) and IBS patients with non-healthy-like HRV (red). The between group comparisons were significant with p<0.01 for all variables. SDNN: standard deviation normal-to-normal; RMSSD: root mean square successive differences; pNN50: proportion normal-to-normal50 (differing by more than 50 milliseconds); LF: low frequency; HF: high frequency; n.u.: normalized units

distinct disease entities, it may be interesting to conduct such an analysis on several mechanisms of interest, to see whether IBS patients with aberrant profiles are aberrant for several mechanisms, or if this way several groups with a respective aberrant global status could be identified.

Two limitations may be most important regarding the methodology of this analysis. It could be hypothesized that a larger sample of HC might display a higher variance of HRV measures, but there are to the best of our knowledge no publications where the same method of HRV measurement resulted in noteworthy differences to the here presented variable range. Nevertheless, this is an important factor to consider in general when conducting such comparisons. The second limitation is the cross-sectional nature of this analysis. Repeating such an analysis after a given amount of time would enhance the reliability of these findings and improve our understanding of the relevance of an aberrant global HRV in IBS. Nevertheless, while this limitation is relevant for this specific analysis, it does not limit the usefulness of the stratification approach, and is therefore mostly important to consider when designing future studies where such a stratification analysis is planned.

Taken together, a mechanism-based stratification utilizing a group of HC as reference for defining a physiological or aberrant HRV function has shown useful to identify a subset of IBS patients significantly differing from HC. This analysis approach can easily be transferred to other mechanisms, and may thus be utilized for identifying IBS patients for which a certain mechanism may be of relevance.

6.5. Integrative analysis of host-microbiota interaction

Manuscript IV presented a stepwise integrative analysis pipeline and tested the usefulness of this pipeline on a pilot-dataset of selected variables of host-microbiota interaction. Figure 15 summarizes the main results of manuscript IV.

Main results of manuscript IV:

- In the network plot of both IBS patients and HC a cluster of mucosal genes involved in immune and enteroendocrine function was observed and termed *'immune-enteroendocrine cluster'*
- Patients uniquely showed an additional cluster of genes involved in mucosal permeability, termed *'permeability cluster'*
- The clusters observed in IBS patients were associated to specific microbiota, whereas in HC no statistically significant association was found
- IBS symptoms were associated to a part of the permeability cluster as well as to further genes and microbiota

Figure 15: Main results of manuscript IV

The results of manuscript IV comprised of two parts. The development of a pipeline for conducting a stepwise integrative analytical approach, as well as the conduction of a pilot analysis to explore the usefulness of this pipeline. The pipeline itself is discussed in detail in the methods chapter 5.3. Therefore, the focus here will be on the pilot analysis and based on this the usefulness of the analysis approach will be discussed. Figure 16 conceptually displays how the different steps of the pipeline were conducted in the pilot study. The 188 genera screened for in the sequence analysis of the gut microbiome were summarized to the respective phyla based on phylogenetic relatedness, which were then correlated to the mRNA expression of the mucosal genes available in the data for this pilot study in order to conduct the integrative analysis and produce overview network plots. The resulting network structures were then compared between IBS patients and HC. Both IBS patients and HC showed a cluster of positively correlated mucosal genes, which was comprised of the toll-like receptors (TLR) 2, 6 and 9, the chromogranins A and B as well as secretogranin 3, and in HC additionally of TLR4, and which was termed '*immune-enteroendocrine cluster'*. In IBS patients this cluster was associated to Bacteria of the Firmicutes and Fusobacteria phylum, whereas no statistically significant association to any specific phyla could be detected in HC. IBS patients uniquely featured an additional cluster, termed '*permeability cluster'*, which was comprised of occludin and tight junction protein 1, TLR 4 and protease activated receptor 2, and which was associated to bacteria of the Chlamydia and Lentisphaerae phylum.

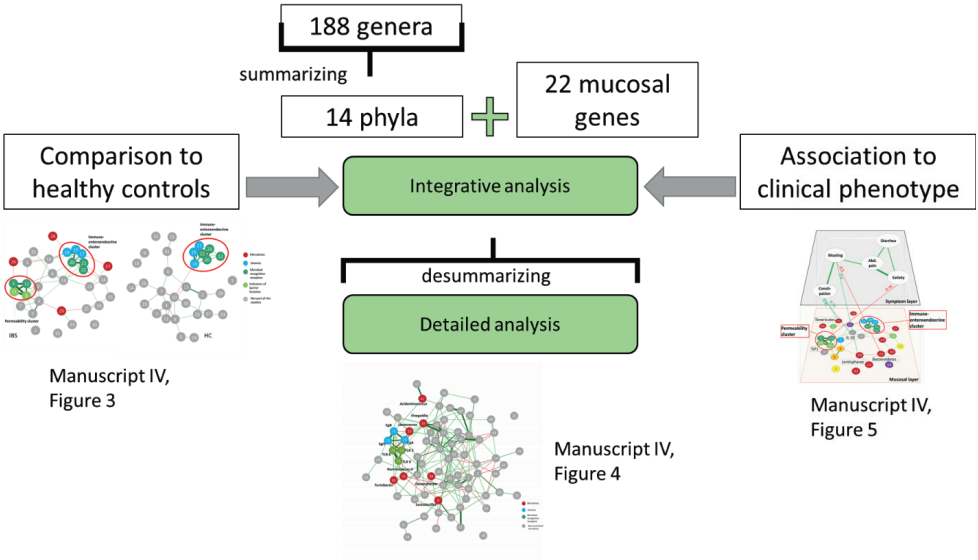


Figure 16: Conceptual visualization of the pilot study conducted to assess the usefulness of the stepwise integrative analysis approach.

This comparison of network structures between IBS patients and HC showed, despite utilizing a limited dataset, noteworthy differences in variable association between IBS patients and HC, thus pinpointing towards variables of high interest for further, more detailed studies. It can be assumed that a more extensive dataset could show even more noteworthy differences between IBS patients and HC. It can therefore be concluded that this part of the analysis pipeline is indeed useful for approaching the research question at hand.

Associating the overview network plots of IBS patients to the clinical phenotype, in this case five key IBS symptoms, as visualized in Figure 5 of manuscript IV, showed that there was a statistically significant association between symptoms and the permeability cluster uniquely found in IBS patients, but no significant association to the immune-enteroendocrine cluster found in both groups. Additionally, the symptoms were associated with bacteria of the Tenericutes, Lentisphaerae and Bacteroidetes phyla as well as to the interleukin IL-10. This shows that an association to IBS symptoms gives additional information to an association of microbiota and mucosal gene expression, making also this step of the analysis pipeline useful for identifying which variables may be of relevance for IBS and therefore potentially interesting targets for more detailed analysis. The last step, a detailed analysis on the genus level, using the variables of the immune-enteroendocrine cluster and bacteria of the

Firmicutes phylum, showed that only a few genera, namely *Lactococcus, Lactobacillus, Finegoldia, Ruminococcus II, Flavonifractor, Turicibacter* and *Acidaminococcus,* were significantly associated to the mucosal targets. This step of the pipeline therefore is useful to get more detailed information about which variables are most relevant for understanding the associations identified. In further studies, these genera and mucosal targets could then selectively be analyzed, for example by experimental tests suitable for elucidating functional interactions on a molecular level.

In conclusion, the here presented pilot study has confirmed the suitability of our pipeline for a stepwise integrative analysis for exploring host-microbiota interaction in IBS and HC. Applying these steps on more comprehensive datasets can likely aid in selecting which variables are of most interest for experimental investigations improving our functional understanding of this interaction. With increasing knowledge, the edges, which are currently based on correlation analyses, could then be defined by experimentally confirmed interaction, further increasing the validity of this analysis. Furthermore, this approach could be extended to an analysis integrating several layers of pathophysiological mechanisms and symptoms.

7. Conclusion and future perspectives

This thesis demonstrates the identification of reproducible symptom-based subgroups with specific profiles of a comprehensive set of IBS-related symptoms and differences in healthcare needs. These subgroups need to be tested in future studies to evaluate the long-term stability of group membership, and, most importantly, whether they relate to differences in pathophysiological mechanisms and/or are suitable predictors for treatment response, as discussed previously[139, 140]. Especially recurring symptom associations such as above-average severity of diarrhea and pain are interesting in this regard, as they may point towards common denominators. In order to answer these questions, future studies could encompass a follow-up symptom registration after a given time period, or, even better, a longitudinal symptom assessment over a certain time period, which could be performed for example by utilizing app-based questionnaires or similar methods suitable for simplifying symptom registration and analysis.

Mechanism-based stratification by multivariate comparisons between IBS patients and healthy controls has demonstrated its usefulness for identifying individuals for which specific complex pathophysiological mechanism may be of relevance, as exemplified by identifying a subset of IBS patients with aberrant overall ANS reactivity. This stratification approach may, in future studies, be applied to other pathophysiological mechanisms, and thus pinpoint which mechanisms are of relevance for which IBS patients, and from this bottom-up approach potentially identify distinct IBS-endotypes. Also here longitudinal studies would prove useful in understanding the relevance of the respective mechanisms in symptom generation and exacerbation, and reduce spurious findings.

Both symptom-based and mechanism-based stratification have pros and cons. Symptom-based subgroups are easier to translate into clinical practice, for example the groups identified in this thesis highlight the relevance of non-GI symptoms for a group of IBS patients. They can quite easily be used to identify which IBS patients might need additional treatment for these symptoms, and for which GI-directed therapy may be sufficient. On the other side, given the rather unspecific nature of many of these symptoms, the groups may not necessarily relate to a distinct underlying mechanism. Here mechanism-based stratification may be more useful. Subgroups with aberrant profiles can be identified, and, if proven reproducible, can give more profound information regarding mechanistic characteristics of these IBS patients,

which may be possible to translate into improved treatment. On the other hand, many of these measures are not easily integrated into clinical practice as there are very costly and time intensive, limiting the direct clinical implications and usefulness for these patients. These two approaches may therefore be utilized in parallel in future studies, to improve both clinical practice and mechanistic understanding.

The mechanism-based stratification approach could also be utilized to generate an "IBS-score". Patients that are phenotyped regarding several mechanisms could be compared to HC regarding all of these mechanisms separately, and be classified based on whether or not they are aberrant on one or several of these mechanisms. Also if the same patients show aberrant profiles for several mechanisms, this could be seen as a hint towards a common denominator and further enhance our understanding of the complex pathophysiology of IBS.

The stepwise integrative analysis pipeline has proven suitable to identify differences in variable associations at the gut mucosal level between IBS patients and healthy controls. Further studies can utilize this approach to comprehensively explore the host-microbiota interaction on the mucosal level by creating more extensive datasets. A truly quantitative measurement of mucosa-adherent microbiota species would further enhance the results of such an analysis. The stepwise integrative analysis pipeline can also be further extended to integrate additional levels representing other relevant pathophysiological mechanisms (Figure 17), thus creating a more comprehensive model of IBS pathophysiology.

Taken together, this thesis has confirmed the high importance of stratification efforts in complex disorders such as IBS. Given the rather broad, purely symptom-based IBS diagnostic criteria, it is quite possible that several distinct etiological mechanisms can lead to an IBS diagnosis. If these are treated as one single cohort, relevant findings may be obscured and treatment development averted. Both symptom- and mechanism-based stratifications can therefore be utilized to improve future studies and can be combined with integrative analysis approaches to further improve our understanding of the mechanisms underlying human physiology and IBS pathophysiology.

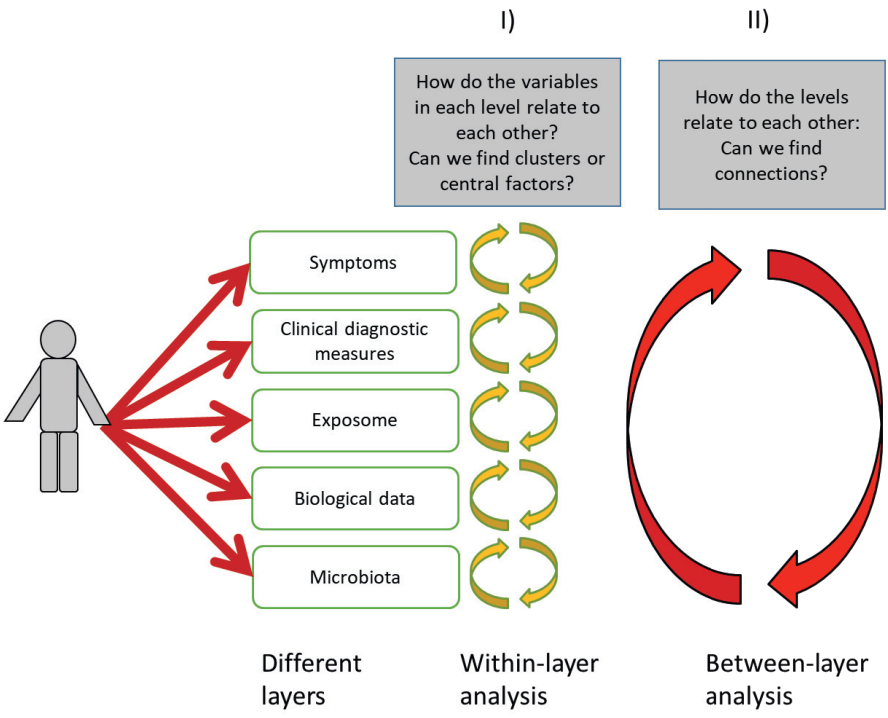


Figure 17: Extension of the stepwise integrative analysis pipeline to associate multiple layers.

8. Acknowledgements

Supervisors

To my three supervisors a tremendous Thank you for your support and trust over the last years.

Magnus, thanks for being the perfect boss, for shepherding me through all of the research experiences over the last years, for giving me the freedom to discover my passion, the trust to execute my ideas, and the support to turn these into actual science.

Lena, I already told you, you are actually really cool;) Thanks also to you for supporting me through this development and giving me lots of expert insight. You are a tremendously clever person with great humor.

Hans, you have the amazing ability to formulate critique in a way that makes people feel better about it. I definitely need to learn that;) Thanks for your support, and for all the shared experiences.

To all three, I hope we can continue working together, I have tons of more ideas what to do with all that data;)

Supporters

Marija, thank you for becoming such a valuable mentor and taking care of me. I'm very glad I met you.

Co-Authors

Thanks to all my co-authors: Magnus Simrén, Lena Öhman, Hans Törnblom, Lukas van Oudenhove, Mike Jones, Oli Palsson, Ami Sperber, Bill Whitehead, Peter Friberg, Veronica Gunterberg, Boris Le Nevé, Marija Cvijovic, Julien Tap, Muriel Derrien. Thanks for the great collaborations, none of this would have been possible without you.

Special thanks to Bill, Oli and Ami for letting me use the amazing 3Country database.

Collegues

Mag-Tarm Lab: Thanks to all fellow researchers and nurses for helping me through the culture-shock and supporting me through all this time. It was a pleasure travelling with you.

Mountain-people: Thanks to all of you for the talks and ideas and especially the laughs. This also includes all mountain-dwarves, just to make sure.

Flascovic-group: Good times, most definitely. Hopefully more to come. Pusspuss.

Special thanks to Irina and Dağsu for translating the summary.

9. References

1. Sood R, Gracie DJ, Law GR, Ford AC. Systematic review with meta-analysis: the accuracy of diagnosing irritable bowel syndrome with symptoms, biomarkers and/or psychological markers. *Aliment Pharmacol Ther* 2015;**42**(5):491-503.
2. Sood R, Law GR, Ford AC. Diagnosis of IBS: symptoms, symptom-based criteria, biomarkers or 'psychomarkers'? *Nat Rev Gastroenterol Hepatol* 2014;**11**(11):683-91.
3. Drossman DA. Functional Gastrointestinal Disorders: History, Pathophysiology, Clinical Features and Rome IV. *Gastroenterology* 2016;**150**(6):1262-1297.
4. Ford AC, Bercik P, Morgan DG, Bolino C, Pintos-Sanchez MI, Moayyedi P. Validation of the Rome III criteria for the diagnosis of irritable bowel syndrome in secondary care. *Gastroenterology* 2013;**145**(6):1262-70.e1.
5. Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC. Functional bowel disorders. *Gastroenterology* 2006;**130**(5):1480-91.
6. Lacy BE, Mearin F, Chang L, et al. Bowel Disorders. *Gastroenterology* 2016;**150**(6):1393–1407.
7. Fass R, Longstreth GF, Pimentel M, et al. Evidence- and consensus-based practice guidelines for the diagnosis of irritable bowel syndrome. *Arch Intern Med* 2001;**161**(17):2081-8.
8. Kellow JE, Phillips SF. Altered small bowel motility in irritable bowel syndrome is correlated with symptoms. *Gastroenterology* 1987;**92**(6):1885-93.
9. Pimentel M, Chow EJ, Lin HC. Eradication of small intestinal bacterial overgrowth reduces symptoms of irritable bowel syndrome. *Am J Gastroenterol* 2000;**95**(12):3503-6.
10. Kanazawa M, Palsson OS, Thiwan SI, et al. Contributions of pain sensitivity and colonic motility to IBS symptom severity and predominant bowel habits. *Am J Gastroenterol* 2008;**103**(10):2550-61.
11. Tornblom H, Van Oudenhove L, Sadik R, Abrahamsson H, Tack J, Simren M. Colonic transit time and IBS symptoms: what's the link? *Am J Gastroenterol* 2012;**107**(5):754-60.
12. Van Oudenhove L, Crowell MD, Drossman DA, et al. Biopsychosocial Aspects of Functional Gastrointestinal Disorders. *Gastroenterology* 2016;**150**(6):1355-1367.
13. Drossman DA, Creed FH, Olden KW, Svedlund J, Toner BB, Whitehead WE. Psychosocial aspects of the functional gastrointestinal disorders. *Gut* 1999;**45 Suppl 2**:II25-30.
14. Bohn L, Storsrud S, Tornblom H, Bengtsson U, Simren M. Self-reported food-related gastrointestinal symptoms in IBS are common and associated with more severe symptoms and reduced quality of life. *Am J Gastroenterol* 2013;**108**(5):634-41.
15. Brown PW. The irritable bowel syndrome. *Rocky Mountain medical journal* 1950;**47**(5):343-6.
16. Manning AP, Thompson WG, Heaton KW, Morris AF. Towards positive diagnosis of the irritable bowel. *Br Med J* 1978;**2**(6138):653-4.
17. Thompson W, Dotevall G, Drossman D, al e. Irritable bowel syndrome: Guidelines for the diagnosis. *Gastroenterol Int* 1989(2):92–95.
18. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. *Scand J Gastroenterol* 1997;**32**(9):920-4.

19. Drossman DA, Morris CB, Hu Y, et al. A prospective assessment of bowel habit in irritable bowel syndrome in women: defining an alternator. *Gastroenterology* 2005;**128**(3):580-9.
20. Garrigues V, Mearin F, Badia X, et al. Change over time of bowel habit in irritable bowel syndrome: a prospective, observational, 1-year follow-up study (RITMO study). *Aliment Pharmacol Ther* 2007;**25**(3):323-32.
21. Tillisch K, Labus JS, Naliboff BD, et al. Characterization of the alternating bowel habit subtype in patients with irritable bowel syndrome. *Am J Gastroenterol* 2005;**100**(4):896-904.
22. Olafsdottir LB, Gudjonsson H, Jonsdottir HH, Thjodleifsson B. Stability of the irritable bowel syndrome and subgroups as measured by three diagnostic criteria - a 10-year follow-up study. *Aliment Pharmacol Ther* 2010;**32**(5):670-80.
23. Enck P, Aziz Q, Barbara G, et al. Irritable bowel syndrome. *Nat Rev Dis Primers* 2016;**2**:16014.
24. Van Oudenhove L, Tornblom H, Storsrud S, Tack J, Simren M. Depression and Somatization Are Associated With Increased Postprandial Symptoms in Patients With Irritable Bowel Syndrome. *Gastroenterology* 2016;**150**(4):866-74.
25. Simren M, Tornblom H, Palsson OS, et al. Visceral hypersensitivity is associated with GI symptom severity in functional GI disorders: consistent findings from five different patient cohorts. *Gut* 2018;**67**(2):255-262.
26. Simrén M, Törnblom H, Palsson OS, Whitehead WE. Management of the multiple symptoms of irritable bowel syndrome. *The Lancet Gastroenterology & Hepatology* 2017;**2**(2):112-122.
27. Drossman DASE. Rome IV Multidimensional Clinical Profile for Functional Gastrointestinal Disorders: MDCP (Second Edition). *Rome Foundation* 2016.
28. Thompson WG, Heaton KW, Smyth GT, Smyth C. Irritable bowel syndrome in general practice: prevalence, characteristics, and referral. *Gut* 2000;**46**(1):78-82.
29. Harvey RF, Salih SY, Read AE. Organic and functional disorders in 2000 gastroenterology outpatients. *Lancet* 1983;**1**(8325):632-4.
30. Sperber AD, Dumitrascu D, Fukudo S, et al. The global prevalence of IBS in adults remains elusive due to the heterogeneity of studies: a Rome Foundation working team literature review. *Gut* 2017;**66**(6):1075-1082.
31. Gwee KA. Irritable bowel syndrome in developing countries--a disorder of civilization or colonization? *Neurogastroenterol Motil* 2005;**17**(3):317-24.
32. Lovell RM, Ford AC. Global prevalence of and risk factors for irritable bowel syndrome: a meta-analysis. *Clin Gastroenterol Hepatol* 2012;**10**(7):712-721 e4.
33. Goodwin L, White PD, Hotopf M, Stansfeld SA, Clark C. Life course study of the etiology of self-reported irritable bowel syndrome in the 1958 British birth cohort. *Psychosom Med* 2013;**75**(2):202-10.
34. Hungin AP, Whorwell PJ, Tack J, Mearin F. The prevalence, patterns and impact of irritable bowel syndrome: an international survey of 40,000 subjects. *Aliment Pharmacol Ther* 2003;**17**(5):643-50.
35. Houghton LA, Brown H, Atkinson W, et al. 5-hydroxytryptamine signalling in irritable bowel syndrome with diarrhoea: effects of gender and menstrual status. *Aliment Pharmacol Ther* 2009;**30**(9):919-29.
36. Frissora CL, Koch KL. The role of gender and biological sex in irritable bowel syndrome. *Current gastroenterology reports* 2005;**7**(4):257-63.

37. Chang L, Mayer EA, Labus JS, et al. Effect of sex on perception of rectosigmoid stimuli in irritable bowel syndrome. *American journal of physiology. Regulatory, integrative and comparative physiology* 2006;**291**(2):R277-84.
38. Toner BB, Akman D. Gender role and irritable bowel syndrome: literature review and hypothesis. *Am J Gastroenterol* 2000;**95**(1):11-6.
39. Bjorkman I, Jakobsson Ung E, Ringstrom G, Tornblom H, Simren M. More similarities than differences between men and women with irritable bowel syndrome. *Neurogastroenterol Motil* 2015;**27**(6):796-804.
40. Cain KC, Jarrett ME, Burr RL, Rosen S, Hertig VL, Heitkemper MM. Gender differences in gastrointestinal, psychological, and somatic symptoms in irritable bowel syndrome. *Dig Dis Sci* 2009;**54**(7):1542-9.
41. Tang YR, Yang WW, Wang YL, Lin L. Sex differences in the symptoms and psychological factors that influence quality of life in patients with irritable bowel syndrome. *Eur J Gastroenterol Hepatol* 2012;**24**(6):702-7.
42. Simren M, Abrahamsson H, Svedlund J, Bjornsson ES. Quality of life in patients with irritable bowel syndrome seen in referral centers versus primary care: the impact of gender and predominant bowel pattern. *Scand J Gastroenterol* 2001;**36**(5):545-52.
43. Gralnek IM, Hays RD, Kilbourne A, Naliboff B, Mayer EA. The impact of irritable bowel syndrome on health-related quality of life. *Gastroenterology* 2000;**119**(3):654-60.
44. Halder SL, Locke GR, 3rd, Talley NJ, Fett SL, Zinsmeister AR, Melton LJ, 3rd. Impact of functional gastrointestinal disorders on health-related quality of life: a population-based case-control study. *Aliment Pharmacol Ther* 2004;**19**(2):233-42.
45. Creed F, Ratcliffe J, Fernandez L, et al. Health-related quality of life and health care costs in severe, refractory irritable bowel syndrome. *Ann Intern Med* 2001;**134**(9 Pt 2):860-8.
46. Spiegel BM, Gralnek IM, Bolus R, et al. Clinical determinants of health-related quality of life in patients with irritable bowel syndrome. *Arch Intern Med* 2004;**164**(16):1773-80.
47. Canavan C, West J, Card T. Review article: the economic impact of the irritable bowel syndrome. *Aliment Pharmacol Ther* 2014;**40**(9):1023-34.
48. Lovell RM, Ford AC. Effect of gender on prevalence of irritable bowel syndrome in the community: systematic review and meta-analysis. *Am J Gastroenterol* 2012;**107**(7):991-1000.
49. van Tilburg MA, Levy RL, Walker LS, et al. Psychosocial mechanisms for the transmission of somatic symptoms from parents to children. *World J Gastroenterol* 2015;**21**(18):5532-41.
50. Bengtson MB, Aamodt G, Vatn MH, Harris JR. Co-occurrence of IBS and symptoms of anxiety or depression, among Norwegian twins, is influenced by both heredity and intrauterine growth. *BMC Gastroenterol* 2015;**15**:9.
51. Saito YA, Petersen GM, Larson JJ, et al. Familial aggregation of irritable bowel syndrome: a family case-control study. *Am J Gastroenterol* 2010;**105**(4):833-41.
52. Drossman DA, Leserman J, Nachman G, et al. Sexual and physical abuse in women with functional or organic gastrointestinal disorders. *Ann Intern Med* 1990;**113**(11):828-33.
53. Koloski NA, Jones M, Kalantar J, Weltman M, Zaguirre J, Talley NJ. The brain--gut pathway in functional gastrointestinal disorders is bidirectional: a 12-year prospective population-based study. *Gut* 2012;**61**(9):1284-90.
54. Sperber AD, Morris CB, Greemberg L, et al. Development of abdominal pain and IBS following gynecological surgery: a prospective, controlled study. *Gastroenterology* 2008;**134**(1):75-84.

55. Wu CY, Chang WP, Chang YH, Li CP, Chuang CM. The risk of irritable bowel syndrome in patients with endometriosis during a 5-year follow-up: a nationwide population-based cohort study. *Int J Colorectal Dis* 2015;**30**(7):907-12.
56. Schwille-Kiuntke J, Frick JS, Zanger P, Enck P. Post-infectious irritable bowel syndrome--a review of the literature. *Z Gastroenterol* 2011;**49**(8):997-1003.
57. Thabane M, Kottachchi DT, Marshall JK. Systematic review and meta-analysis: The incidence and prognosis of post-infectious irritable bowel syndrome. *Aliment Pharmacol Ther* 2007;**26**(4):535-44.
58. Schwille-Kiuntke J, Mazurak N, Enck P. Systematic review with meta-analysis: post-infectious irritable bowel syndrome after travellers' diarrhoea. *Aliment Pharmacol Ther* 2015;**41**(11):1029-37.
59. Donnachie E, Schneider A, Mehring M, Enck P. Incidence of irritable bowel syndrome and chronic fatigue following GI infection: a population-level study using routinely collected claims data. *Gut* 2017.
60. Wiklund IK, Fullerton S, Hawkey CJ, et al. An irritable bowel syndrome-specific symptom questionnaire: development and validation. *Scand J Gastroenterol* 2003;**38**(9):947-54.
61. Locke GR, 3rd, Zinsmeister AR, Fett SL, Melton LJ, 3rd, Talley NJ. Overlap of gastrointestinal symptom complexes in a US community. *Neurogastroenterol Motil* 2005;**17**(1):29-34.
62. Aziz I, Palsson OS, Tornblom H, Sperber AD, Whitehead WE, Simren 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**(1):86-96.
63. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* 2002;**64**(2):258-66.
64. Spiller RC, Humes DJ, Campbell E, et al. The Patient Health Questionnaire 12 Somatic Symptom scale as a predictor of symptom severity and consulting behaviour in patients with irritable bowel syndrome and symptomatic diverticular disease. *Aliment Pharmacol Ther* 2010;**32**(6):811-20.
65. Riedl A, Schmidtmann M, Stengel A, et al. Somatic comorbidities of irritable bowel syndrome: a systematic analysis. *J Psychosom Res* 2008;**64**(6):573-82.
66. Whitehead WE, Palsson O, Jones KR. Systematic review of the comorbidity of irritable bowel syndrome with other disorders: what are the causes and implications? *Gastroenterology* 2002;**122**(4):1140-56.
67. Whitehead WE, Palsson OS, Levy RR, Feld AD, Turner M, Von Korff M. Comorbidity in irritable bowel syndrome. *Am J Gastroenterol* 2007;**102**(12):2767-76.
68. Addolorato G, Marsigli L, Capristo E, Caputo F, Dall'Aglio C, Baudanza P. Anxiety and depression: a common feature of health care seeking patients with irritable bowel syndrome and food allergy. *Hepatogastroenterology* 1998;**45**(23):1559-64.
69. Creed F. The relationship between psychosocial parameters and outcome in irritable bowel syndrome. *Am J Med* 1999;**107**(5A):74S-80S.
70. Vu J, Kushnir V, Cassell B, Gyawali CP, Sayuk GS. The impact of psychiatric and extraintestinal comorbidity on quality of life and bowel symptom burden in functional GI disorders. *Neurogastroenterol Motil* 2014;**26**(9):1323-32.
71. Grenham S, Clarke G, Cryan JF, Dinan TG. Brain-gut-microbe communication in health and disease. *Frontiers in physiology* 2011;**2**:94.

72. Bercik P, Collins SM, Verdu EF. Microbes and the gut-brain axis. *Neurogastroenterol Motil* 2012;**24**(5):405-13.
73. Ohman L, Tornblom H, Simren M. Crosstalk at the mucosal border: importance of the gut microenvironment in IBS. *Nat Rev Gastroenterol Hepatol* 2015;**12**(1):36-49.
74. Rajilic-Stojanovic M, Jonkers DM, Salonen A, et al. Intestinal microbiota and diet in IBS: causes, consequences, or epiphenomena? *Am J Gastroenterol* 2015;**110**(2):278-87.
75. Mazurak N, Enck P, Muth E, Teufel M, Zipfel S. Heart rate variability as a measure of cardiac autonomic function in anorexia nervosa: a review of the literature. *Eur Eat Disord Rev* 2011;**19**(2):87-99.
76. Adeyemi EO, Desai KD, Towsey M, Ghista D. Characterization of autonomic dysfunction in patients with irritable bowel syndrome by means of heart rate variability studies. *Am J Gastroenterol* 1999;**94**(3):816-23.
77. Heitkemper M, Burr RL, Jarrett M, Hertig V, Lustyk MK, Bond EF. Evidence for autonomic nervous system imbalance in women with irritable bowel syndrome. *Dig Dis Sci* 1998;**43**(9):2093-8.
78. Heitkemper M, Jarrett M, Cain KC, et al. Autonomic nervous system function in women with irritable bowel syndrome. *Dig Dis Sci* 2001;**46**(6):1276-84.
79. Elsenbruch S, Lovallo WR, Orr WC. Psychological and physiological responses to postprandial mental stress in women with the irritable bowel syndrome. *Psychosom Med* 2001;**63**(5):805-13.
80. Jarrett ME, Burr RL, Cain KC, Rothermel JD, Landis CA, Heitkemper MM. Autonomic nervous system function during sleep among women with irritable bowel syndrome. *Dig Dis Sci* 2008;**53**(3):694-703.
81. Mazurak N, Seredyuk N, Sauer H, Teufel M, Enck P. Heart rate variability in the irritable bowel syndrome: a review of the literature. *Neurogastroenterol Motil* 2012;**24**(3):206-16.
82. Camilleri M. Physiological underpinnings of irritable bowel syndrome: neurohormonal mechanisms. *J Physiol* 2014;**592**(14):2967-80.
83. Rajilic-Stojanovic M, de Vos WM. The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS microbiology reviews* 2014;**38**(5):996-1047.
84. Qin J, Li R, Raes J, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 2010;**464**(7285):59-65.
85. Rodino-Janeiro BK, Vicario M, Alonso-Cotoner C, Pascua-Garcia R, Santos J. A Review of Microbiota and Irritable Bowel Syndrome: Future in Therapies. *Advances in therapy* 2018;**35**(3):289-310.
86. Enck P, Mazurak N. Dysbiosis in Functional Bowel Disorders. *Annals of nutrition & metabolism* 2018;**72**(4):296-306.
87. Simren M, Barbara G, Flint HJ, et al. Intestinal microbiota in functional bowel disorders: a Rome foundation report. *Gut* 2013;**62**(1):159-76.
88. Matto J, Maunuksela L, Kajander K, et al. Composition and temporal stability of gastrointestinal microbiota in irritable bowel syndrome--a longitudinal study in IBS and control subjects. *FEMS immunology and medical microbiology* 2005;**43**(2):213-22.
89. Malinen E, Rinttila T, Kajander K, et al. Analysis of the fecal microbiota of irritable bowel syndrome patients and healthy controls with real-time PCR. *Am J Gastroenterol* 2005;**100**(2):373-82.
90. Rajilic-Stojanovic M, Biagi E, Heilig HG, et al. Global and deep molecular analysis of microbiota signatures in fecal samples from patients with irritable bowel syndrome. *Gastroenterology* 2011;**141**(5):1792-801.

91. Jeffery IB, O'Toole PW, Ohman L, et al. An irritable bowel syndrome subtype defined by species-specific alterations in faecal microbiota. *Gut* 2012;**61**(7):997-1006.
92. Carroll IM, Ringel-Kulka T, Siddle JP, Ringel Y. Alterations in composition and diversity of the intestinal microbiota in patients with diarrhea-predominant irritable bowel syndrome. *Neurogastroenterol Motil* 2012;**24**(6):521-30, e248.
93. Jalanka-Tuovinen J, Salojarvi J, Salonen A, et al. Faecal microbiota composition and host-microbe cross-talk following gastroenteritis and in postinfectious irritable bowel syndrome. *Gut* 2014;**63**(11):1737-45.
94. Tillisch K, Mayer EA, Gupta A, et al. Brain Structure and Response to Emotional Stimuli as Related to Gut Microbial Profiles in Healthy Women. *Psychosom Med* 2017;**79**(8):905-913.
95. Labus JS, Hollister EB, Jacobs J, et al. Differences in gut microbial composition correlate with regional brain volumes in irritable bowel syndrome. *Microbiome* 2017;**5**(1):49.
96. Tap J, Derrien M, Tornblom H, et al. Identification of an Intestinal Microbiota Signature Associated With Severity of Irritable Bowel Syndrome. *Gastroenterology* 2017;**152**(1):111-123 e8.
97. Jalanka-Tuovinen J, Salonen A, Nikkila J, et al. Intestinal microbiota in healthy adults: temporal analysis reveals individual and common core and relation to intestinal symptoms. *PLoS One* 2011;**6**(7):e23035.
98. Johansson ME, Sjovall H, Hansson GC. The gastrointestinal mucus system in health and disease. *Nat Rev Gastroenterol Hepatol* 2013;**10**(6):352-61.
99. McGuckin MA, Linden SK, Sutton P, Florin TH. Mucin dynamics and enteric pathogens. *Nature reviews. Microbiology* 2011;**9**(4):265-78.
100. Sommer F, Backhed F. The gut microbiota--masters of host development and physiology. *Nature reviews. Microbiology* 2013;**11**(4):227-38.
101. Peterson LW, Artis D. Intestinal epithelial cells: regulators of barrier function and immune homeostasis. *Nature reviews. Immunology* 2014;**14**(3):141-53.
102. Bischoff SC, Barbara G, Buurman W, et al. Intestinal permeability--a new target for disease prevention and therapy. *BMC Gastroenterol* 2014;**14**:189.
103. Martinez C, Lobo B, Pigrau M, et al. Diarrhoea-predominant irritable bowel syndrome: an organic disorder with structural abnormalities in the jejunal epithelial barrier. *Gut* 2013;**62**(8):1160-8.
104. Piche T, Barbara G, Aubert P, et al. Impaired intestinal barrier integrity in the colon of patients with irritable bowel syndrome: involvement of soluble mediators. *Gut* 2009;**58**(2):196-201.
105. Bertiaux-Vandaele N, Youmba SB, Belmonte L, et al. The expression and the cellular distribution of the tight junction proteins are altered in irritable bowel syndrome patients with differences according to the disease subtype. *Am J Gastroenterol* 2011;**106**(12):2165-73.
106. Coeffier M, Gloro R, Boukhettala N, et al. Increased proteasome-mediated degradation of occludin in irritable bowel syndrome. *Am J Gastroenterol* 2010;**105**(5):1181-8.
107. Buhner S, Li Q, Vignali S, et al. Activation of human enteric neurons by supernatants of colonic biopsy specimens from patients with irritable bowel syndrome. *Gastroenterology* 2009;**137**(4):1425-34.
108. Fritscher-Ravens A, Schuppan D, Ellrichmann M, et al. Confocal endomicroscopy shows food-associated changes in the intestinal mucosa of patients with irritable bowel syndrome. *Gastroenterology* 2014;**147**(5):1012-20.e4.

109. Bonfiglio F, Henstrom M, Nag A, et al. A GWAS meta-analysis from 5 population-based cohorts implicates ion channel genes in the pathogenesis of irritable bowel syndrome. *Neurogastroenterol Motil* 2018:e13358.
110. Bohn L, Storsrud S, Simren M. Nutrient intake in patients with irritable bowel syndrome compared with the general population. *Neurogastroenterol Motil* 2013;**25**(1):23-30.e1.
111. Posserud I, Strid H, Storsrud S, et al. Symptom pattern following a meal challenge test in patients with irritable bowel syndrome and healthy controls. *United European Gastroenterol J* 2013;**1**(5):358-67.
112. Bartminski G, Crossley M, Turcanu V. Novel biomarkers for asthma stratification and personalized therapy. *Expert Rev Mol Diagn* 2015;**15**(3):415-30.
113. Liu M, Liberzon A, Kong SW, et al. Network-based analysis of affected biological processes in type 2 diabetes models. *PLoS Genet* 2007;**3**(6):e96.
114. Cui J, Antoniou AC, Dite GS, et al. After BRCA1 and BRCA2-what next? Multifactorial segregation analyses of three-generation, population-based Australian families affected by female breast cancer. *American journal of human genetics* 2001;**68**(2):420-31.
115. Benson M. Clinical implications of omics and systems medicine: focus on predictive and individualized treatment. *J Intern Med* 2016;**279**(3):229-40.
116. Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* 2011;**12**(1):56-68.
117. De Las Rivas J, Fontanillo C. Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS Comput Biol* 2010;**6**(6):e1000807.
118. Goh KI, Choi IG. Exploring the human diseasome: the human disease network. *Brief Funct Genomics* 2012;**11**(6):533-42.
119. Barabasi AL. Network medicine--from obesity to the "diseasome". *N Engl J Med* 2007;**357**(4):404-7.
120. 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**(2):395-402.
121. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;**67**(6):361-70.
122. Ware JE, Kosinski M, Dewey JE, Gandek B. How to score and interpret single-item health status measures: a manual for users of the SF-8 health survey. *Lincoln, RI: QualityMetric Incorporated* 2001;**15**(10):5.
123. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Eur Heart J* 1996;**17**(3):354-81.
124. Buysschaert M, Donckier J, Dive A, Ketelslegers JM, Lambert AE. Gastric acid and pancreatic polypeptide responses to sham feeding are impaired in diabetic subjects with autonomic neuropathy. *Diabetes* 1985;**34**(11):1181-5.
125. Emmanuel AV, Kamm MA. Laser Doppler flowmetry as a measure of extrinsic colonic innervation in functional bowel disease. *Gut* 2000;**46**(2):212-7.
126. Ohman L, Stridsberg M, Isaksson S, Jerlstad P, Simren M. Altered levels of fecal chromogranins and secretogranins in IBS: relevance for pathophysiology and symptoms? *Am J Gastroenterol* 2012;**107**(3):440-7.

127. McKernan DP, Gaszner G, Quigley EM, Cryan JF, Dinan TG. Altered peripheral toll-like receptor responses in the irritable bowel syndrome. *Aliment Pharmacol Ther* 2011;**33**(9):1045-52.
128. Bashashati M, Rezaei N, Shafieyoun A, et al. Cytokine imbalance in irritable bowel syndrome: a systematic review and meta-analysis. *Neurogastroenterol Motil* 2014;**26**(7):1036-48.
129. McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* 2014;**10**(4):e1003531.
130. Pereira MB, Wallroth M, Jonsson V, Kristiansson E. Comparison of normalization methods for the analysis of metagenomic gene abundance data. *BMC Genomics* 2018;**19**(1):274.
131. Weiss S, Xu ZZ, Peddada S, et al. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 2017;**5**(1):27.
132. Love MI, Huber, W., Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 2014;**15**(12):550.
133. Ward Jr JH. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* 1963;**58**(301):236-244.
134. Lloyd S. Least squares quantization in PCM. *IEEE transactions on information theory* 1982;**28**(2):129-137.
135. Ester M, Kriegel H-P, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Kdd; 1996; 1996. p. 226-231.
136. Fraley C, Raftery AE. Model-based clustering, discriminant analysis, and density estimation. *J Am Stat Assoc* 2002(97):611–631.
137. Fraley C, Raftery AE. Enhanced model-based clustering, density estimation, and discriminant analysis software: MCLUST. *J Classif* 2003(20):263–286.
138. Schwarz G. Estimating the dimension of a model. *Ann Stat* 1978(6):461–464.
139. Black CJ, Ford AC. Editorial: subgroups in irritable bowel syndrome-more than just diarrhoea and constipation? *Aliment Pharmacol Ther* 2017;**46**(7):697.
140. Polster A, Van Oudenhove L, Jones M, Ohman L, Tornblom H, Simren M. Editorial: subgroups in irritable bowel syndrome-more than just diarrhoea and constipation? Authors' reply. *Aliment Pharmacol Ther* 2017;**46**(7):698-699.