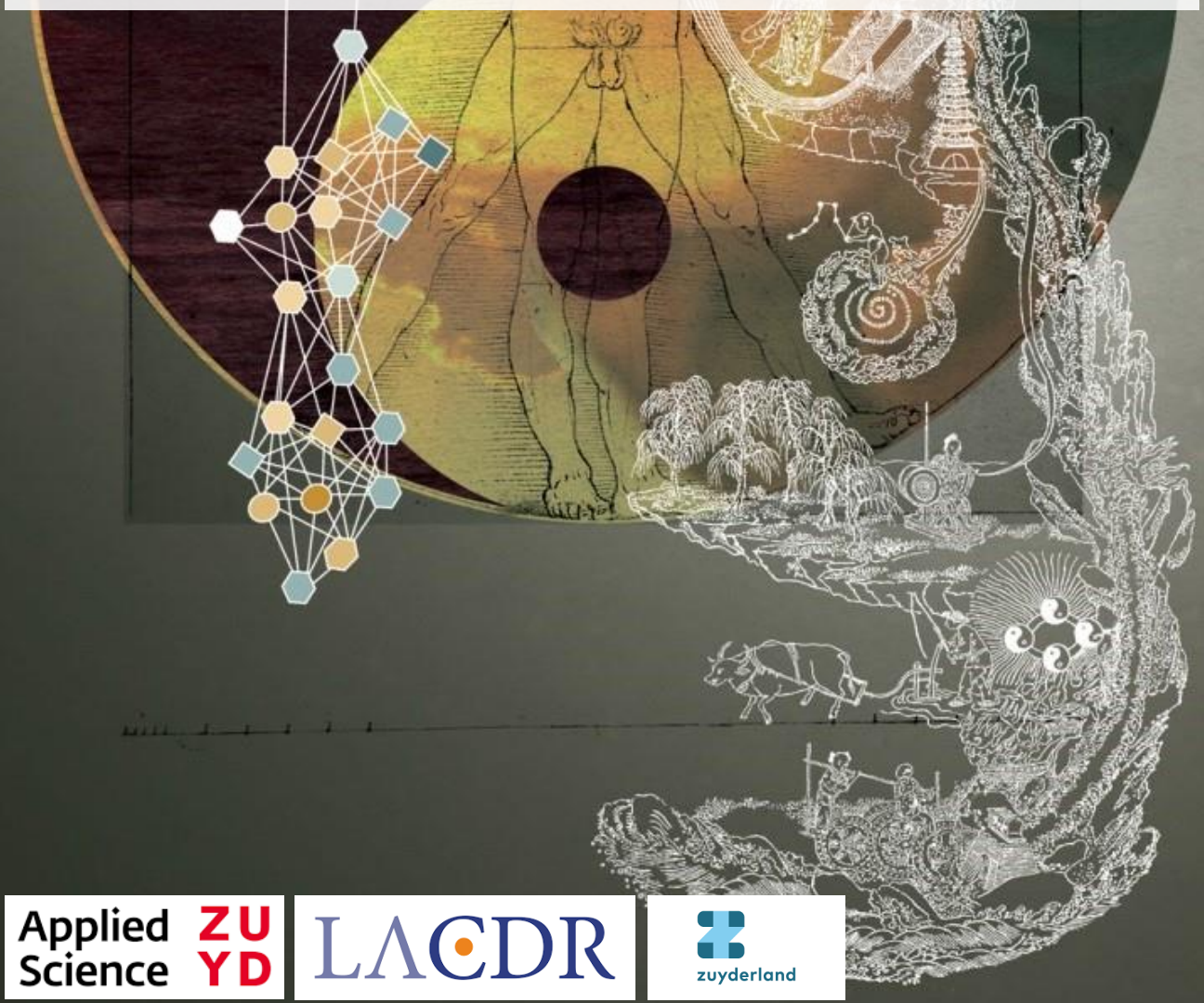
Louis Bolk Instituut

Funded by
EKHAGASTIFTELSEN

Systems biology of COPD subgroups based on Western and Chinese diagnosis

Final report project 2019-79

Herman van Wietmarschen



Applied Science ZUYD

LACDR

zuyderland

Systems biology of COPD subgroups based on Western and Chinese diagnosis - Final report project 2019-79

Herman van Wietmarschen

Project partners:
Applied Science Zuyd, Heerlen, the Netherlands
Analytical Biosciences, LACDR, Leiden, the Netherlands
Zuyderland Medical Center, Heerlen, the Netherlands

Keywords: COPD, metabolomics, Chinese diagnosis, subtyping,

71 pages

Publication number: 2024-6068-VG

This publication is available at
https://louisbolk.nl/en/publications

www.louisbolk.nl
info@louisbolk.nl
T +31(0) 343 523 860
Kosterijland 3-5
3981 AJ Bunnik,
the Netherlands
𝕏 @LouisBolk

Preface

With great pleasure we present the final report of the project titled: 'Systems biology of COPD subgroups based on Western and Chinese diagnosis'. The project was funded by the Ekgahastiftelsen (project number 2019-79). We thank all project partners and especially all COPD patients participating in this study.

Contents

Samenvatting

Achtergrond. COPD is een heterogene progressieve longziekte en is wereldwijd een van de belangrijkste doodsoorzaken. De behandeling van COPD zou baat kunnen hebben bij een integratie van traditionele medische perspectieven uit andere delen van de wereld. In deze studie worden COPD-subgroepen gebaseerd op conventionele westerse geneeskunde vergeleken met subgroepen uit de Traditionele Chinese Geneeskunde.

Patiënten & Methoden. 56 COPD-patiënten werden geworven op 5 verschillende locaties in de regio Heerlen/Sittard in Nederland. Bloedmonsters werden verzameld, een TCM symptoom vragenlijst werd ingevuld, een TCM diagnose werd gesteld en andere COPD gerelateerde vragenlijsten werden ingevuld. De bloedmonsters werden gebruikt voor een metabolomics analyse met een nieuw signaallipiden platform.

Resultaten. In de data werden specifieke clusters van TCM-symptomen gevonden, die overeenkwamen met Stagnatiesymptomen en Deficiëntie zonder Stagnatiesymptomen. Er werden verschillende metabolietprofielen gevonden die overeenkwamen met westerse COPD-subgroepen (KNGF en GOLD) en met TCM-subtypen.

Conclusies. De studie biedt een biologisch perspectief op zowel COPD-subgroepen als TCM-subgroepen. De TCM-diagnose wordt ondersteund door specifieke patronen van metabolieten. Dit suggereert dat een toekomstige behandeling van COPD-patiënten geoptimaliseerd zou kunnen worden door het integreren van gepersonaliseerde leefstijladviezen op basis van specifieke TCM-subgroepen.

Abstract

Background. COPD is a heterogeneous progressive lung disease and one of the leading causes of death worldwide. COPD treatment could benefit from an integration of traditional medical perspectives that originated in different parts of the world. In this study COPD subgroups based on conventional Western medicine are compared with Traditional Chinese Medicine subgroups.

Patients & Methods. 56 COPD patients were recruited in 5 different locations in the Heerlen/Sittard region of the Netherlands. Blood samples were collected, a TCM symptom questionnaire was completed, a TCM expert diagnosis was taken, and other COPD related questionnaires were completed. The blood samples were used for a metabolomics profiling with a novel signalling lipids platform.

Results. Specific clusters of TCM symptoms were found in the data, corresponding to Stagnation symptoms and Deficiency without Stagnation symptoms. Different metabolite profiles were found to correspond with Western COPD subgroups (KNGF and GOLD) and with TCM subtypes.

Conclusions. The study offers a biological perspective on COPD subgroups as well as on TCM subgroups. TCM diagnosis is supported by specific metabolite clusters. This suggests that a future treatment of COPD patients could be optimized by integrating personalized lifestyle advise based on specific TCM symptom clusters.

Abbreviations

13-HOTE	13-Hydroxyoctadeca-9,11,15-trienoic acid
16-HDoHE/DCA	16-hydroxy Docosahexaenoic Acid
6MWT	Six-minute walk test
ACV	Analysis of Covariance
ANCOVA	Analysis of Covariance
BMI	Body Mass Index
CAT	COPD Assessment Test
CATPCA	Categorical Principal Component Analysis
CCR3	C-C chemokine receptor type 3
CCQ	Clinical COPD Questionnaire (van der Molen 2003)
COPD	Chronic obstructive pulmonary disease
COX	Cyclooxygenase
CYP	Cytochrome P450
DALY	Disability-adjusted life-year
DCA	Docosanoic acid
EDTA	Ethylenediaminetetraacetic acid
EPA	Eicosapentaenoic Acid
ERK1/2	Extracellular signal-regulated protein kinases 1 and 2
FEV_1	Forced Expiratory Volume in one second
FXR	Farnesoid X receptor
GOLD	Global Initiative for Chronic Obstructive Lung Disease
GPRs	G-protein coupled receptors
GUDCA	Glycoursodeoxycholic acid
HEPE	Hydroxyicosapentaenoic acid
IL-17	Interleukin-17
KNGF	The Royal Dutch Society for Physical Therapy
LOX	Lipoxygenase
LPA	Lysophosphatidic acid
LPE	Lysophosphatidylethanolamine
LPS	Lipopolysaccharide
MAPK	Mitogen-activated protein kinase
MART	Physical Activity Questionnaire (Marshall 2005)
METC	Medisch Ethische Toetsings Commissie [Medical Ethical Research Council]
MISS	Meten Is Super Simpel [measuring is super simpel] (Ummels 2020)
mMRC	Modified Medical Research Council dyspnoea scale (Mahler 1988)
MUST	Malnutrition Universal Screening Tool (Elia 2003)
MS/MS	Mass spectrometry/mass spectrometry
MRC	Medical Research Council dyspnoea scale
NF-κB	Nuclear factor kappa B
NHG	Nederlands Huisartsen Genootschap [Dutch College of General Practitioners]

ORM	Ordinal regression model
PGF3a	Prostaglandin F3a
RhoA	Ras homolog gene family member A
ROCK	Rho-associated coiled-coil-containing kinase
S1P	Sphingosine 1-phosphate
SARC-F	Sarcopenia questionnaire (Woo 2014)
TCM	Traditional Chinese Medicine
TUDCA	Taurine conjugate of ursodeoxycholic acid
UHPLC	Ultra High Performance Liquid Chromatography
WHO	World Health Organisation

1 Introduction

1.1 Problem definition

Respiratory diseases are among the leading causes of death worldwide, which is a public health problem that is currently challenging our society. Respiratory infections, lung cancer and chronic obstructive pulmonary disease (COPD), together account for 9.5 million yearly deaths worldwide, one-sixth of the global total (Gibson 2013). Within this top four respiratory diseases, COPD accounts for 35% of the mortality and 21% of the disability- adjusted life-years (DALYs) lost worldwide (2008-2012, Gibson 2013). The global prevalence is currently 251 million cases of COPD, which is expected to further increase in the future (WHO 2017). The most common risk factor for COPD is active or passive smoking (Rastron 2014), but since 2-4% of patients with COPD are non-smokers, other factors such as work-related exposure to gases or dusts and α_1-antitrypsin deficiency have also been associated with increased risk for COPD (Bang 2015, Mannino 2007).

COPD is a progressive life-threatening lung disease that causes breathlessness on exertion, chronic cough and predisposes to exacerbations and serious illness (WHO 2017; Gibson 2013). Apart from the airway obstruction, which defines the medical diagnosis of COPD, serious systemic consequences, also called extra-pulmonary manifestations of COPD, like deconditioning, exercise intolerance, skeletal muscle dysfunction and metabolic impact (e.g. cachexia) can arise as the disease progresses (Gibson 2013; Decramer 2008; Schols 2009). There is also great variation in this patient group regarding co-existing diseases i.e. comorbidity. At least one comorbidity exists in 50-98% of patients with COPD (Ferrer 1997; van Manen 2001; Chatila 2008; Vanfleteren 2013); and in 16-46% of patients even three to four comorbid conditions can be detected (van Manen 2001; Chatila 2008; Hillas 2015). Because of these pulmonary and various extra-pulmonary manifestations and comorbidity, COPD is regarded as complex and heterogeneous disease (Agusti 2010). Due to this heterogeneity, several studies tried to identify homogeneous patient subgroups in the past years (Han 2010, Beekman 2016, Garcia-Aymerich 2017). There are indications that each subgroup has its own disease course in time and responds differently to treatment (Han 2010). However, there is no consensus yet about which variations constitute clear subgroups for which optimized treatment can be provided. Thus, identification and validation of such clinically relevant COPD subgroups is needed. This would enable evidence based personalized treatment which in turn might lead to improved outcomes (Beasley 2009, Burgel 2010, Fingleton 2011).

1.2 Current management of patients with COPD

The increasing knowledge about the relevance of various COPD subgroups to enable personalized care, has not yet been incorporated into the current understanding and management of patients with COPD (Agusti 2011). Existing guidelines recommend different pharmacological interventions depending on increasing severity of the phenotype (GOLD 2019,

Snoeck-Stroband 2015), but no specific clinical interventions regarding physical activity and exercise are mentioned based on different phenotypes.

The American Thoracic Society and the European Respiratory Society states that physical exercise and education regarding dealing with the disease are the cornerstone of pulmonary rehabilitation, with the ultimate goal of patients being able to self-manage their condition effectively (Blackstock 2018; Spruit 2013). However, the difficulty of implementing relevant results from clinical studies into daily practice emerges from the use of homogeneous populations in studies in contrast to the heterogeneous population that is seen in daily practice (Schellevis 2007). Consequently, guidelines for healthcare professionals rely on these homogeneous oriented scientific studies. This means that healthcare professionals are not informed on evidence based care for their heterogeneous patients. Literature, and thus guidelines, generally provide more generic 'one size fits all' treatment advises for patients with COPD. However, this approach does not match with individual differences in patient characteristics, preferences and needs and the principle of personalized care. Therefore, not the disease but the individual patient needs to be the starting point in diagnosis and treatment and clinicians need tools to select the right care, in the right place, at the right time for individual patients.

1.3 Vision

It becomes increasingly clear that a 'one size fits all' healthcare system is not going to work for many people with chronic conditions. A personalized care approach is needed that takes into account the history of the patient, preferences, lifestyle, and willingness to change lifestyle, social

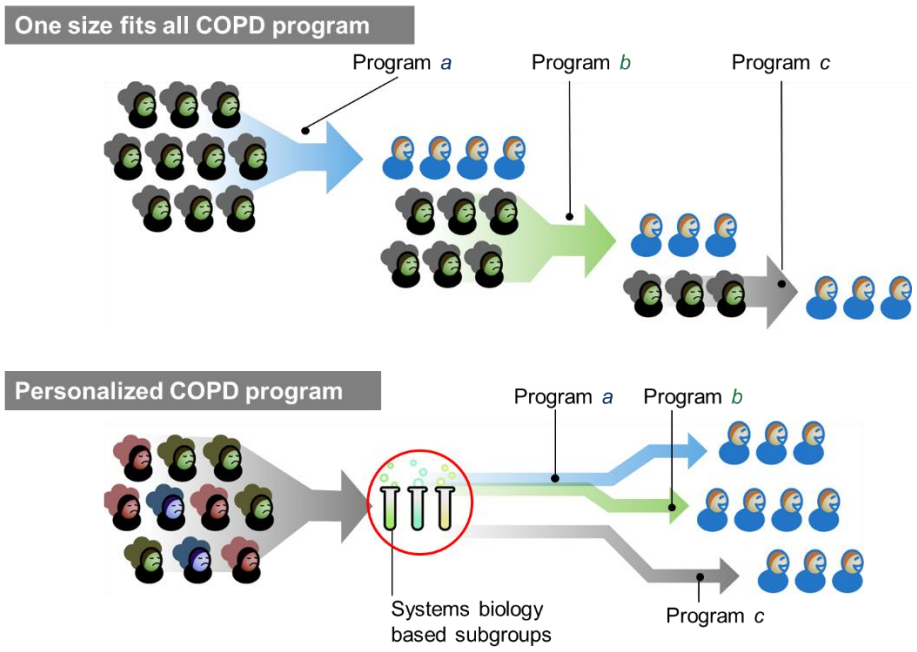


Figure 1 From 'one size fits all' towards a personalized COPD program

environment, and what is meaningful in the patient's life. However, it is a challenge to develop individualized treatment programs based on current scientific and clinical evidence that is obtained from large controlled randomized trials. A first step towards a more personalized approach is to identify subgroups within patient populations, validate these subgroups and develop personalized treatment programs based on these subgroups (see Figure 1). Several studies have reported a variety of subgroups of patients with COPD based on clinical parameters (Pinto 2015, Han 2010, Zhou 2017). However, the biological mechanisms or pathophysiology underlying these subgroups are still unknown.
Traditional Chinese medicine (TCM) represents a holistic medical system that already makes use of different subgroups to treat patients in a much more individualized manner than Western medicine currently does, focussing on the self-healing capacity of the body (Schroën 2014). The identification of relevant COPD subgroups in Western medicine could be supported by the TCM knowledge about COPD subgroups (Maciocia 2004). Our vision is to combine and compare TCM and Western COPD subgroups to better understand the biological mechanisms behind these subgroups. We then select relevant subgroups and develop a multidisciplinary treatment plan including physiotherapy and lifestyle advise targeted to those subgroups (van der Greef 2010).

Besides studying the TCM symptoms and clinical parameters related to COPD subgroups, more detailed insights into the biology behind COPD subgroups can improve understanding of the biological mechanisms related to the subgroups. This can be provided by a modern systems biology approach (Kitano 2002, Agusti 2011). Systems biology is a relatively new field in biology that uses data analysis techniques to evaluate differences in patterns of large numbers of variables between subgroups obtained from bio fluids such as blood or urine in combination with clinical parameters, physical test results and symptoms (van der Greef 2007). The focus of systems biology is on relationships between variables and patterns of relationships.

Our vision for this project is to use a systems biology approach to discover biological processes underlying subgroups of patients with COPD in order to better understand these subgroups and provide a biological validation of these subgroups. The interpretation of the COPD subgroups is the result of an integrated multivariate analysis of (1) metabolic data, (2) Western clinical data (body functions and activities) and Chinese symptom data. Validation of these potential classification parameters of patients with COPD itself is needed to eventually guide clinical practice.

The result of the project is a characterization and selection of the most promising COPD subgroups. These subgroups can then be implemented in multidisciplinary programs including amongst others physiotherapy, exercise, nutrition, and coaching for COPD. The results will also be used to inform clinical practitioners to optimize treatment, inform the development of COPD treatment guidelines and the scientific community.

1.4 Preliminary data

1.4.1 Efforts to discover relevant COPD subgroups

In de past years several subtypes of patients with chronic diseases, like COPD, are described in clinical guidelines and used in clinical practice. A common ground of these clinical patients subtypes according to the western (para)medical healthcare system is the classification of disease severity as a starting point. The severity of airway obstruction, expressed in spirometric GOLD-stages I-IV by the Global Initiative for Chronic Obstructive Lung Disease (GOLD), is mainly used to classify patients with COPD, based on the Forced Expiratory Volume in one second (FEV_1) (GOLD 2019). However, to gain a holistic understanding regarding the impact of COPD on an individual level, solely assessing the patient's spirometric capacity is insufficient. The severity of airway obstruction (FEV_1) is poorly related with the severity of breathlessness, exercise limitation and health status impairment (GOLD 2019; Spruit 2010; Agusti 2010; O'Donnell 2015; Pinto-plata 2004). Therefore, a combined assessment of GOLD-stage (FEV_1), symptoms (modified Medical Research Council dyspnoea scale (mMRC) or COPD Assessment Test (CAT)) and exacerbation frequency into four patient groups (A-D) was recently introduced by the Global Initiative for Chronic Obstructive Lung Disease (GOLD 2011-2019): (A) low risk, less symptoms, (B) low risk, more symptoms, (C) high risk, less symptoms and (D) high risk, more symptoms (GOLD 2019). However, this subgroup classification remains doubtful, since the distinction between those subgroups is arbitrary and does not allow for personalised care based on valid phenotypes (Smid 2017).

National guidelines also evolved in this period with own classifications of patients with COPD, such as the Dutch College of General Practitioners (NHG) guidelines 'Adult asthma' and 'COPD'(Snoeck-Stroband 2015). It divides patients into three groups with respect to the burden of disease (low, moderate and high burden of disease) based on airway obstruction severity (FEV_1), symptoms and disabilities (Medical Research Council dyspnoea scale (MRC) or Clinical COPD Questionnaire (CCQ), nutrition status and exacerbation frequency (Snoeck-Stroband 2015). Currently, the Dutch clinical practice guidelines for physical therapy in patients with COPD is being revised by the Royal Dutch Society for Physical Therapy (KNGF). These guidelines also use a range of clinical parameters to classify patients according to patients' needs and goals. Based on the best available scientific evidence, clinical expertise of COPD related healthcare providers and patient values and preferences six different patient groups are suggested in relation to disease stability (exacerbations), symptoms, physical capacity and level of physical activity. This classification does not take into account the severity of airway obstruction as a clinical classification parameter (Vreeken 2019). Despite these different approaches to classify patients with COPD into certain subgroups using different clinical parameters, there is no consensus yet of which parameters would be used best in order to classify patients with COPD.

1.4.2 **Systems biology approach to validate subgroups based on Western and Chinese diagnosis**

In 2008 the Sino-Dutch center for Preventive and Personalized Medicine (https://www.sinodutchcentre.nl/) was founded with the vision to study Chinese medicine diagnosis patterns using a systems biology approach (Schroën 2014). One of the authors of this research proposal (Herman van Wietmarschen) conducted his PhD research at this center. In 2009 the first paper was published in which distinct gene expression and metabolite patterns were presented for two subgroups of rheumatoid arthritis patients based on Chinese diagnosis (van Wietmarschen 2009, Van Wietmarschen 2012). Further research showed that certain elementary concepts in Chinese diagnosis (eg. the Ba Gang concepts internal, external, cold and heat) could be matched with metabolite and clinical parameters (van Wietmarschen 2011).
A similar approach revealed distinct urine metabolite patterns for subgroups of subjects with metabolic syndrome based on Chinese diagnosis (Qi and Yin Deficiency with Stagnation, or with Dampness) (Wei 2012). Based on these results a personalized treatment program for people with overweight was developed at a rehabilitation center de Vogellanden (Zwolle, the Netherlands). Participants were first subjected to a Chinese medicine diagnosis and designated to one out of five subgroups based on that diagnosis. For each of these subgroups a lifestyle program including diet, exercise, and coaching was developed including insights and advice from Chinese medicine. The participants of the first few groups in the program were very enthusiastic, the data collected in the program are not published yet. However, we feel that the addition of Chinese medicine knowledge can greatly contribute to the development of more personalized health promotion programs.

2 Study aim and objectives

The goal of this research is to identify biological mechanisms and patterns underlying COPD subgroups based on Western clinical variables, TCM symptom patterns and metabolic data. The aim is to come to a practical number of clinically relevant subgroups.

The objectives of the study are:

1. Collecting clinical data according to a Western and Chinese medicine approach in patients with COPD.
2. Finding patterns in multiple physical parameters, clinical parameters and symptoms related to COPD.
3. Comparing these patterns of variables with Western and Chinese COPD leading to a selection of relevant subgroups.
4. Elucidate the biological mechanisms underlying the established COPD subgroups using a systems biology approach, i.e. validation by means of metabolic data.
5. Use the results to inform clinical practitioners on possible validated subgroups. These subgroups may lead to the development of more tailored COPD healthcare in the future, in relation to the preferences and goals of the patient.
6. Use the results to inform the development of COPD treatment (para)medical guidelines.
7. Disseminate the results to a wider (scientific) community.

2.1 Hypothesis

The hypothesis of this study is that clinically relevant COPD subgroups can be discovered through an integration of Western and Chinese medicine views on COPD. This is done by applying a systems biology approach to a combination of Western clinical data, physical data, Chinese symptom data and metabolomics data.

2.2 Relevance

The current project proposal ultimately aims at providing a more personalized treatment for patients with COPD. COPD is the 4th leading cause of mortality worldwide and has an enormous impact on patients' quality of life. COPD is a complex heterogeneous disease with respect to clinical presentation, physiology, response to therapy, decline in lung function, inflammation process and survival. It is known that a large heterogeneity exists amongst patients, resulting in different responses to standard treatment. Still no consensus is formed about clinically relevant COPD subgroups that could be used to optimize treatment programs. Current literature provides health care providers with a 'one-size-fits-all' approach and mainly consist of medication, health promotion, exercise and functional training, diet and psychosocial support that might lead to suboptimal treatment effects. A more subgroup based advice regarding health promotion, exercise and functional training, diet and psychosocial support can be helpful in the future.

Therefore, the identification of clinically relevant and validated COPD subgroups can contribute to the development of an evidence based personalized health approach for COPD.

2.3 Innovation

The innovative aspect in this project is to use knowledge from traditional Chinese medicine about COPD subgroups and integrate it with Western clinical and physical data, which will result in clinically relevant COPD subgroups. This knowledge might be used in the future to create a more personalized treatment for patients with COPD. Hereby, patterns of variables are integrated with clinical parameters that have physical impact on a patient level using a systems biology approach. A novel signalling lipids metabolomics analysis of blood samples will be used to provide a biological validation as well as an understanding of biological mechanisms underlying the identified COPD subgroups.

In the field of systems biology multivariate data analysis techniques have been developed that allow the integration of various sources of data, such as symptoms, metabolite concentrations and physical parameters. Furthermore, since biology and life is inherently is non-linear, techniques such as non-linear principal component analysis have been developed to deal with non-linear relationships between variables. Finally, network analysis techniques can elucidate patterns of relationships which are not revealed by standard univariate statistical analyses.

All these systems biology techniques improve the chances of finding relevant COPD subgroups (see also Figure 2). The impact of potential new and validated COPD subgroups on a patient level is reflected by more than one component from the International Classification of Functioning, Disability and Health (ICF): body functions (e.g. physiological measure FEV_1, physical and mental symptoms) and activities (e.g. physical functioning, physical activity in everyday life) (WHO 2001).

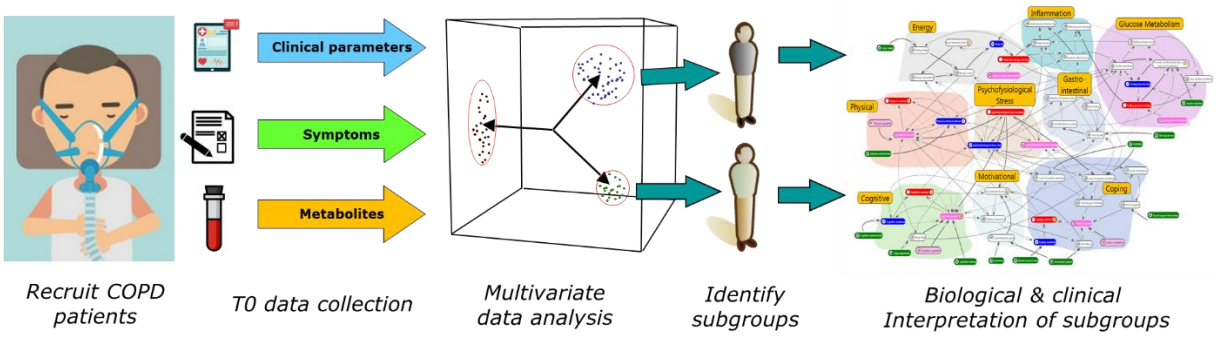


Figure 2 Systems biology workflow

This strategy agrees with the vision of Ekhaga to support integrative medicine, focus on health promotion strategies and integration of physical impact and body functions. Furthermore, the proposed systems biology approach emphasises the study of biological, physiological and

behavioural mechanisms underlying COPD subgroups, which is one of the priorities of Ekhaga. Finally, the results will be used to inform future revisions of national guidelines and international GOLD guidelines.

3 Materials and methods

3.1 Design

In this cross-sectional study physiological and clinical parameters, TCM symptoms and blood samples will be collected from patients suffering from COPD at one time point.

3.2 Study population

Patients with chronic obstructive pulmonary disease (COPD) were recruited by means of a combination of probability and non-probability sampling to achieve a wide range in patient characteristics (e.g. age, gender, GOLD-stage). First, a convenience sample of primary and/or secondary clinics where physiotherapists are specifically treating patients with COPD were contacted through existing networks in the south of the Netherlands. Potential participants were recruited within these institutes. In addition, a voluntary sample of patients were recruited outside these institutes through affiliated patient organisations. To ensure heterogeneity and rich data collection in the study population and to cover all relevant patient subgroups, broad selection criteria were used: eligible adult patients (> 18 years) with the medical diagnosis of COPD and a known GOLD-stage of I-IV will be included.

Based on earlier studies using systems biology approach in other populations we expect the number of subjects needed to achieve reliable subgroups in this study is at least 20 per subgroup (van Wietmarschen 2009, van Wietmarschen 2012). Based on earlier studies (Peters 2017, Garcia-Aymerich 2017, Burgel 2010, Snoeck-Stroband 2015) we expect approximately three different subgroups, resulting in a sample size of 60 patients. In order to account for a heterogeneous study population in relation to the relevant strata and taking into account a non-response rate of 20% we aimed to address 75 participants.

All subjects provided written informed consent prior to inclusion into the study. Ethical approval for this study was obtained from METC Brabant (NL74360.028.20).

3.3 Study procedure

Patient recruitment and handling was carried out by Zuyd University of Applied Sciences. Potential participants were contacted and screened for study inclusion by a research assistant by phone. Participants received written and verbal information about the aim of the study and were required to give written informed consent prior to inclusion. After informed consent, participants were asked to provide the following personal data: age, gender and GOLD-stage (or permission to contact physician by the researchers to collect this fact). Participants received an online Chinese medicine symptoms questionnaire, the modified Medical Research Council dyspnoea scale (mMRC) (Mahler 1988; Cazzola 2015), Clinical COPD Questionnaire (CCQ including subgroups

symptoms, mental health and physical functioning) (van der Molen 2003), a short exacerbation questionnaire (exacerbation history in the past year and related use of oral corticosteroids/prednisolone or hospitalisations) (Burge 2003; Caramori 2009), physical activity questionnaire (Marshall 2005) a short nutrition status questionnaire (BMI, non-voluntarily weight loss over the past month and past six months)(Elia 2003), and a sarcopenia questionnaire (SARC-F)(Woo 2014). Furthermore, participants were asked to perform on the physical six-minute walk test (6MWT) (ATS 2002) to indicate physical capacity, in order to collect additional information on body functions and physical impact related to their COPD.

Finally, a fasting EDTA blood plasma sample was collected from each of the patients for a systems biology analysis. The blood samples were stored at -80 degrees Celsius until the entire batch of samples could be analysed by the metabolomics facility. The plasma samples were measured at the Metabolomics Facility in Leiden (Leiden University). A novel validated UHPLC-MS/MS targeted signalling lipids platform was used to measure 150 metabolites (Yang 2024, de Leeuw 2017). This platform includes free fatty acids (omega-3, omega-6, omega-9), oxylipins (isoprostanes, prostaglandins and other oxidized lipids), lysophospholipids, sphingosine lipids, endocannabinoids and bile acids covering a wide variety of biological pathways and processes.

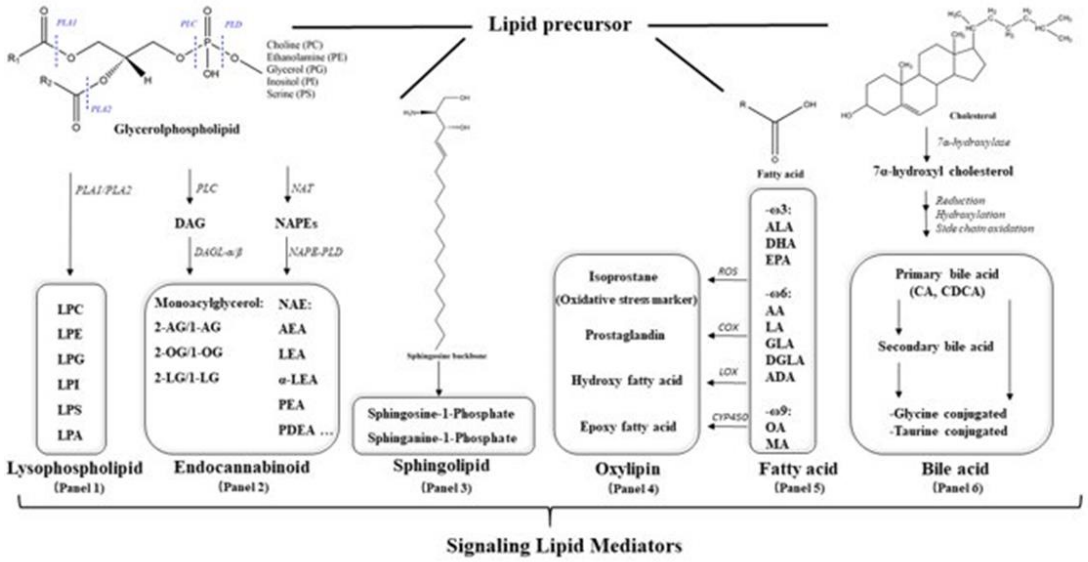


Figure 3 Overview of the signaling lipids platform (de Leeuw 2017)

3.4 TCM symptom questionnaire

A TCM symptom questionnaire was designed by TCM experts and scientists according to TCM theory and practical experience including 28 symptoms related to 4 expected subgroups of COPD patients. Each of the symptoms was asked as a question on a 7 point Likert scale from not or never present to very often present or very severe, in a random order without being labelled to a subtype. The questions per subgroup were the following (see also Table 1):

1. Lung Qi Deficiency
    - Do you experience shortness of breath?
    - Do you cough?
    - Are you short of breath with light exercise?
    - Are you easily tired or fatigued?
    - Do you dislike windy weather?
    - Are you sensitive to cold?
    - Do you easily perspire?

2. Lung Qi and Spleen Qi Deficiency
    - Do you recognize a bloated feeling in your stomach area?
    - Do you experience abdominal distention?
    - Do you experience loss of appetite?
    - Do you have loose stools?

3. Lung Qi and Spleen Qi Deficiency + Kidney Yang Deficiency
    - Do you experience shortness of breath when lying down?
    - Do you have white slime in your mouth?
    - Do you have a swollen face?
    - Do you experience dizziness?
    - Do you have tinnitus?
    - Do you experience tension in your chest area?
    - Do you experience palpitations?
    - Do you experience pain or weak knees and lower back?
    - Do you urinate frequently?
    - Do you have to urinate at night?
    - Do you experience loss of urine at night?

4. Lung Qi and Spleen Qi Deficiency + Kidney Yang and Yin Deficiency
    - Do you cough with spittle?
    - Do you experience a dry mouth and throat?
    - Do you have clammy hands and feet?
    - Do you experience night sweating?
    - Are you sensitive to catching a cold?

In addition to the TCM symptom questionnaire and the diagnosis by the TCM expert, an algorithm was constructed in Excel to calculate a subtype score from the symptom data. The algorithm attributes extra points to a subtype for combinations of symptoms present. The algorithm provides percentage scores for the 4 subtypes for each subject.

The mixed subtype scores that were sometimes provided by the TCM expert and the algorithm were recoded into a single TCM subtype yes/no score. The TCM expert scores were categorized into the 4 subtype groups. For the algorithm subtype 3 and 4 were merged into a single category because of too few entrees (<7) in the separate categories.

3.5 Clinical and TCM symptom data processing

In order to run a non-linear principal component (CATPCA) analysis the data needs to be processed in several steps: data screening, data recording, and discretization.

The data screening involved removing variables with very little to no variation, which were smoking (only 5 people smoked out of the 56), smoking history (all but one had a smoking history), malnutrition (only 7 subjects with non-zero score), number of days with tracker (all 7 days except 5), number of hospitalisations (only 6 subjects with hospitalisations).

Then some variables were recoded. Blood pressure was recoded into two variables, systolic blood pressure and diastolic blood pressure. Walking aids was recoded into categories none, walker, walking stick, other. The data using the activity tracker was recoded into how many days the tracker was used. A new variable was created for the average active minutes per day. Subsequently, in order to ensure stability of the categorical variables in the analyses each categorical variable should be recoded to have at least 7 entrees per category. Categories were merged to ensure this. Numerical variables were discretized into 19 categories.

Based on the transformation plots optimal transformations were chosen for the variables. All TCM symptom variables receive Ordinal analysis level. The following variables receive a Spline ordinal with 2 degrees and 2 inknots analysis level: D4, CCQ11, CCQ12, CCQ13, CCQT, D9, SARCT, MIST1, MIST2, D7, D8, MWT16, MWT17, MWTT. EXC1, EXC2, MRCT, MART also receive Ordinal analysis level, and variable D3 is Nominal. An overview of all the variables in the data analysis is presented in table 1. The Screeplot indicated 3 relevant principal components.

Table 1 Overview of the variables

Data name	Data label	Questionnaire	Data type	Analysis vars	Labelling vars
NR	Patient number		Nominal		x
AN	Aanalysis number		Numeric		x
BN	Batch number		Nominal		x
MWT16	Systolic blood pressure	Measurement	Numeric	x	
MWT17	Diastolic blood pressure	Measurement	Numeric	x	
MWTT	MWTT	6 Minute walk test	Numeric	x	
D3	Gender	Demographics	Nominal	x	
D4	Age	Demographics	Numeric	x	

Data name	Data label	Questionnaire	Data type	Analysis vars	Labelling vars
GLDR	GOLD stage	GOLD stage	Ordinal		x
EXC1	Number exacerbations	Exacerbation Q	Ordinal	x	
EXC2	Number antibiotics courses	Exacerbation Q	Ordinal	x	
CCQ11	Total score symptoms	Clinical COPD Q	Numeric	x	
CCQ12	Total score emotions	Clinical COPD Q	Numeric	x	
CCQ13	Total score condition	Clinical COPD Q	Numeric	x	
CCQT	CCQ total	Clinical COPD Q	Numeric	x	
TCM1	Shortness of breath	TCM symptom	Ordinal	x	
TCM2	Shrt breath excertion	TCM symptom	Ordinal	x	
TCM3	Shrt breath lying down	TCM symptom	Ordinal	x	
TCM4	Cough	TCM symptom	Ordinal	x	
TCM5	Perspire	TCM symptom	Ordinal	x	
TCM6	Moisture retention in face	TCM symptom	Ordinal	x	
TCM7	Dizziness	TCM symptom	Ordinal	x	
TCM8	Tinnitis	TCM symptom	Ordinal	x	
TCM9	Cough with phlegm	TCM symptom	Ordinal	x	
TCM10	White mucus formation	TCM symptom	Ordinal	x	
TCM11	Dry mouth/throat	TCM symptom	Ordinal	x	
TCM12	Clammy hands/feet	TCM symptom	Ordinal	x	
TCM13	Nightsweat	TCM symptom	Ordinal	x	
TCM14	Easily tired	TCM symptom	Ordinal	x	
TCM15	Pressure on the chest	TCM symptom	Ordinal	x	
TCM16	Palpitations	TCM symptom	Ordinal	x	
TCM17	Sensitive to windy weather	TCM symptom	Ordinal	x	
TCM18	Sensitive to cold	TCM symptom	Ordinal	x	
TCM19	Sensitive for catching colds	TCM symptom	Ordinal	x	
TCM20	Full feeling in stomach	TCM symptom	Ordinal	x	
TCM21	Bloating in the stomach	TCM symptom	Ordinal	x	
TCM22	Decreased appetite	TCM symptom	Ordinal	x	
TCM23	Loose stools	TCM symptom	Ordinal	x	
TCM24	Pain	TCM symptom	Ordinal	x	
TCM25	Weak knees/lower back	TCM symptom	Ordinal	x	
TCM26	Frequently urinating	TCM symptom	Ordinal	x	
TCM27	Urinating at night	TCM symptom	Ordinal	x	
TCM28	Unwanted urine loss at night	TCM symptom	Ordinal	x	
SUB7	TCM tanje all rec 2		Ordinal		x
SUB8	TCM algo rec		Nominal		x

Data name	Data label	Questionnaire	Data type	Analysis vars	Labelling vars
MART	MART	Physical activity Q	Ordinal	x	
D7	Length	Demographics	Numeric	x	
D8	Weight	Demographics	Numeric	x	
D9	BMI	Demographics	Numeric	x	
SARCT	SARC-F	Sarcopenia Q	Ordinal	x	
MRCT	MRCT	Modified Medical Research Council dyspnoea Q	Ordinal	x	
MIST1	Average steps	MISS activity	Numeric	x	
MIST2	Average minutes per day	MISS activity	Numeric	x	
KNGF3	KNGF rec	KNGF profile	Ordinal		x

3.6 Metabolomics data processing

Sciex OS (AB SCIEX, Version 2.1.6) was used to integrate metabolite peaks from raw LC-MS/MS data. The relative concentration was calculated from the area of the targets divided by the area of the internal standard assigned to the respective peak. A quality control sample was prepared by pooling all study samples and integrated into the analysis. Targets with RSD of the quality control samples above 30% or background signal over 40% were excluded from statistical analyses. A total of 139 metabolites passed the criteria, which are listed in Appendix 1. The sums of metabolites within the same lipid subclass and ratios of specific targets to their precursors were also calculated to provide detailed information for metabolic profiling (Appendix 2). Finally, the 166 metabolomics variables for the data analysis consisted of 139 detectable lipids, 5 lipid subclass sums and 22 lipid ratios.

3.7 Data analysis

3.7.1 Comparing Western and Chinese COPD subgroups

Non-linear Principal Component Analysis (CATPCA) will be used to detect the largest variation in the Western clinical and physical parameters and Chinese medicine symptoms of the patients (Linting 2012, van Wietmarschen 2011). The sources of variation in the data will be compared with existing Western and Chinese subgroups of COPD. In order to perform the CATPCA analysis the data will first be screened, recoded and discretized. Missing data were imputed with the mode of the respective optimally scaled variables. Optimal transformations were checked with the transformation plots. The number of relevant principal components was determined using a Scree plot.

A forced classification approach is then used to discover the Western clinical parameters and Chinese medicine symptoms most relevant to the subgroups of patients with COPD

(Nishisato1984). The KNGF profile variable will be used with a weight of 20 to find relationships in the data with the KNGF profile groups. Furthermore, the TCM expert subtype variable will be used with a weight of 10 to find relationships in the data with the TCM subtypes.

In addition to the CATPCA analysis a network view will be constructed from the correlations between the variables using Cytoscape 3.10.2 (Zitnik 2024, de Souza 2020). Such a view provides an intuitive visual way to look for clusters in the data and how clinical variables are connected to the TCM symptoms. A selection of the most relevant COPD subgroups will then be made based on the patterns of clinical, physical and symptoms variables if possible.

3.7.2 **Biological interpretation of COPD subgroups with metabolomics**

Prior to the analysis, the relative concentrations of all signaling lipids were log2-transformed and scaled to achieve a normal distribution and comparability among the targets. The missing values were replaced by one-tenth of the minimum value observed from the respective metabolite. Furthermore, to eliminate the interference from possible confounders for metabolites, gender, age and Body Mass Index (BMI) were inspected with an independent T-test and Pearson correlation analysis. Afterward, the existence of gender-associated metabolites was checked, and age and BMI were taken into the mathematical models as covariates during the following analyses.

An ordinal regression model was constructed to investigate the relationship between metabolites and COPD subgroups. Only targets with $p \leq 0.05$ were selected for visualization using effects plots (Fox 2009) which indicate how the probability distribution varies across different types of subgroups with changes in metabolite abundance. Then, Analysis of Covariance (ANCOVA) was performed to elucidate the variation of metabolites in different types of subgroups of COPD patients. The abundance of metabolites was expressed by estimated marginal (EM) means after correction by age and BMI instead of detected abundance after normalization (Searle 1980). The exclusion criteria for lipid targets were set at p-value above 0.05.

In addition, the correlations between clinical characteristics and metabolites, and categories were investigated with partial correlation analysis adjusting for age and BMI. Spearman correlations were used in order to include the categorical variables into the analysis. Correlation coefficients above $|0.35|$ were chosen for analyzing the correlations between metabolites and clinical characteristics. All the statistical analyses of the metabolomics data were performed using R (version 4.3.2).

4 Results and discussion

A total of 56 patients were recruited at 5 different locations in the Heerlen/Sittard region of the Netherlands for the study. The distribution of patients over the locations was 9 (17%), 5 (9%), 3 (5%), 4 (7%), 35 (63%). All patients completed the clinical questionnaires and the TCM symptom questionnaire. Many patients scored profile 6 on the KNFG stage which would indicate lungrevalidation in a hospital, which is a category apart from the other 5 KNGF stages. In order to make more distinction between the KNGF profile 6 patients, these patients were recategorized into KNGF stages 2-5 based on their physical capacity. 38 of the 56 patients were seen by a TCM expert in order to provide a TCM diagnosis. Blood samples were collected and stored from 56 patients. However, 4 samples were incorrectly labelled and couldn't be traced back to the original patient. After running the metabolomics platform a total of 49 patients provided data suitable for analysis. An overview of the available data is presented in Figure 4 below. The baseline data for the two sets of participants is presented in Table 2.

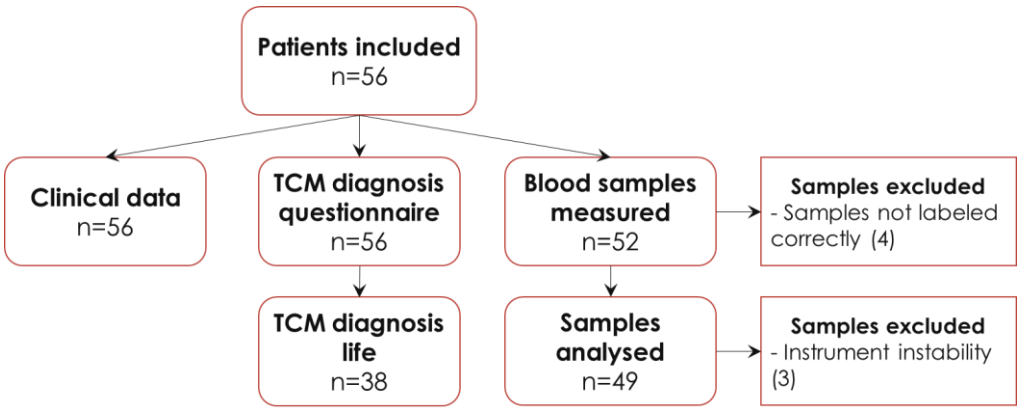


Figure 4 Overview of the available data

Table 2 Baseline data of the participants available for analysis of the TCM symptoms and clinical variables (n=56) and available for the metabolomics analysis (n=49)

	TCM & Clinical	Metabolomics
Demographics		
Age		
N	56	49
Mean (SD)	68.9 (8.19)	68.3 (8.3)
Gender n (%)		
Male	19 (34%)	18 (37%)
Female	37 (66%)	31 (63%)
Weight (Mean, SD)	75.9 (15.5)	76.2 (16.2)
Height (Mean, SD)	1.65 (0.08)	1.65 (0.08)
BMI (Mean, SD)	27.9 (5.33)	27.9 (5.6)
Smoking n (%)	5 (9%)	4 (8%)

	TCM & Clinical	Metabolomics
Smoking history n (%)	55 (98%)	48 (98%)
Systolic blood pressure (Mean, SD)	139 (18.6)	139 (19.5)
Diastolic blood pressure (Mean, SD)	18.5 (11.1)	75.2 (11.6)
COPD related measurements		
GOLD stage		
GOLD 1	0 (0%)	0 (0%)
GOLD 2	18 (33%)	17 (35%)
GOLD 3	28 (52%)	22 (45%)
GOLD 4	8 (15%)	8 (16%)
KNGF profile		
KNGF 1	5 (9%)	5 (10%)
KNGF 2	20 (36%)	18 (37%)
KNGF 3	2 (4%)	0 (0%)
KNGF 4	18 (32%)	18 (37%)
KNGF 5	10 (18%)	7 (14%)
Exacerbation history		
Lung attacks n (%)	23 (42%)	21 (43%)
Antibiotic treatment n (%)	21 (38%)	20 (41%)
Hospitalisation n (%)	6 (11%)	5 (10%)
MRCT		
Stage 0	4 (7%)	4 (8%)
Stage 1	13 (23%)	11 (22%)
Stage 2	23 (41%)	20 (41%)
Stage 3	12 (21%)	10 (20%)
Stage 4	4 (7%)	4 (8%)
CCQ (Mean, SD)	2.46 (1.14)	2.4 (1.2)
Activity		
MART ≥ 4 n (%)	25 (45%)	24 (49%)
6MWT (Mean, SD)	65% (20%)	66 (21%)
Activity tracker		
Average steps (Mean, SD)	8466 (4982)	8877 (5132)
Average active minutes (Mean, SD)	72.8 (50)	77 (51)
SARC-F > 4 n (%)	20 (36%)	18 (37%)
Nutrition		
MUST malnutrition risk		
Low risk n (%)	49 (88%)	43 (88%)
Medium risk n (%)	4 (7%)	3 (6%)
High risk n (%)	3 (5%)	3 (6%)

	TCM & Clinical	Metabolomics
TCM subtypes		
Subtype expert		
Lung Qi Def	4 (7%)	5 (10%)
Lung & Spleen Qi Def	13 (23%)	11 (22%)
Lung & Spleen Qi Def + Kidney Yang Def	20 (36%)	17 (35%)
Lung & Spleen Qi Def + Kidney Yang Def + Kidney Yin Def	19 (34%)	16 (33%)
Subtype algorithm		
Lung Qi Def	22 (39%)	19 (39%)
Lung & Spleen Qi Def	28 (50%)	14 (29%)
Lung & Spleen Qi Def + Kidney Yin & Yang Def	2 (4%)	16 (33%)

4.1 Comparing Western and Chinese COPD subgroups

In this section the results of an exploratory analysis of the clinical variables and the TCM symptoms is presented in 7 subsections. The first three subsections concern an unsupervised CATPCA analysis. In the first subsection all variables are analysed in one model. The second subsection focusses on the clinical variables only, while the third subsection focusses on the TCM symptoms only. In the fourth and fifth section a forced classification approach is used, in order to force a classification and find the variables most related to these classifications. In subsection 4 the KNGF profile variables is used as a classification variable, while in subsection 5 the TCM expert subgrouping is used as a classification variable. In subsection 6 a network analysis of the correlations between the variables is described. In subsection 7 the conclusions of this section is described.

4.1.1 Unsupervised CATPCA analysis all variables

Figure 5 shows the score plots of the CATPCA analysis including all variables, the left panel shows dimension 1 and 2 and the right panel dimension 1 and 3. The scores are labelled by batch number, representing the location the patients were recruited from. Each dot represents one subject, the same subjects are plotted in the left and right panel. The figures show no clear groupings in the distribution of the subjects, indicating that there are no clear groups of patients with different characteristics from others. Furthermore, the subjects from the different locations are distributed throughout the plots indicating that there is no effect of the location on the data. Apparently there are no clear differences between the patients groups treated at the various locations. No outliers are detected in the data.

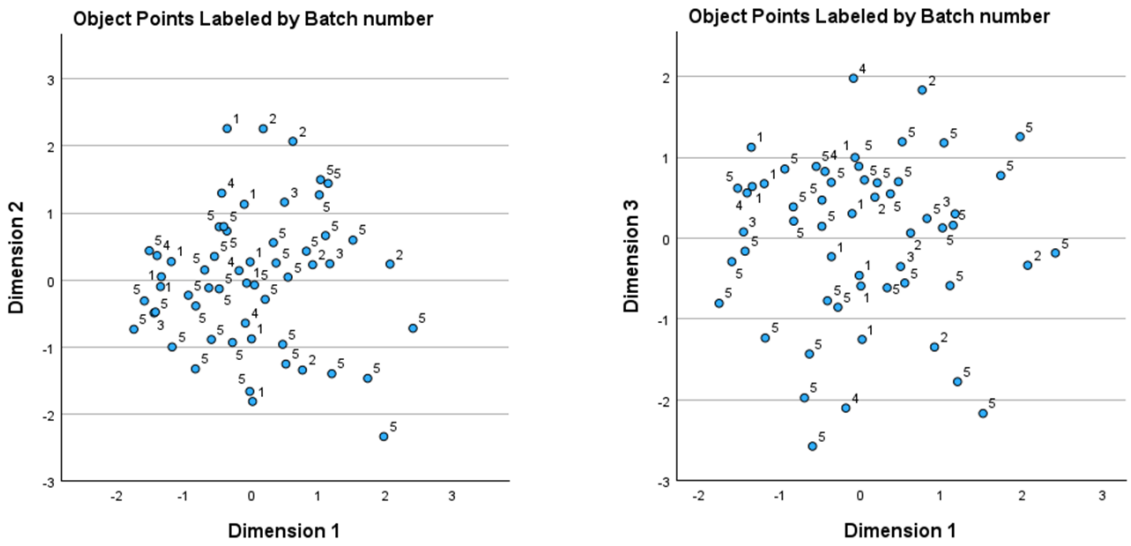


Figure 5 Score plots labelled by recruitment location

Figure 6, 7 and 8 show the loading plots of the CATPCA analysis with all variables for three combinations of dimensions. In the direction of the first dimension all the TCM symptoms have a large loading, indicating a large variation in TCM symptom severity. Also the COPD symptoms measured with the CCQ, the dyspnoea score (mRCT), and the sarcopenia score (SARC-F) is large in the first dimension. This suggests that the first dimension in the analysis can be interpreted as a general symptom severity. Subjects scoring higher on dimension 1 have a higher symptom severity, subjects scoring negative on dimension1 have a lower symptom severity.

In dimension 2 the variables MWT, average minute per day and average steps have a large loading, and are separate from the large group of symptoms on the right part of the figure. These variables are related to activity. The KNGF profile score is determined based on the 6 minute walk test, which provides the MWT score and the average minutes and steps scores, and also the CCQ questionnaire. The CCQ score variables are in the lower right part of the loading plot, which is partly consistent with KNGF. Subjects with more symptoms (higher CCQ scores) in general show less activity (low on MWT, average minutes and average steps), however there are also subjects with high activity and high CCQ scores (upper right part of the plot).

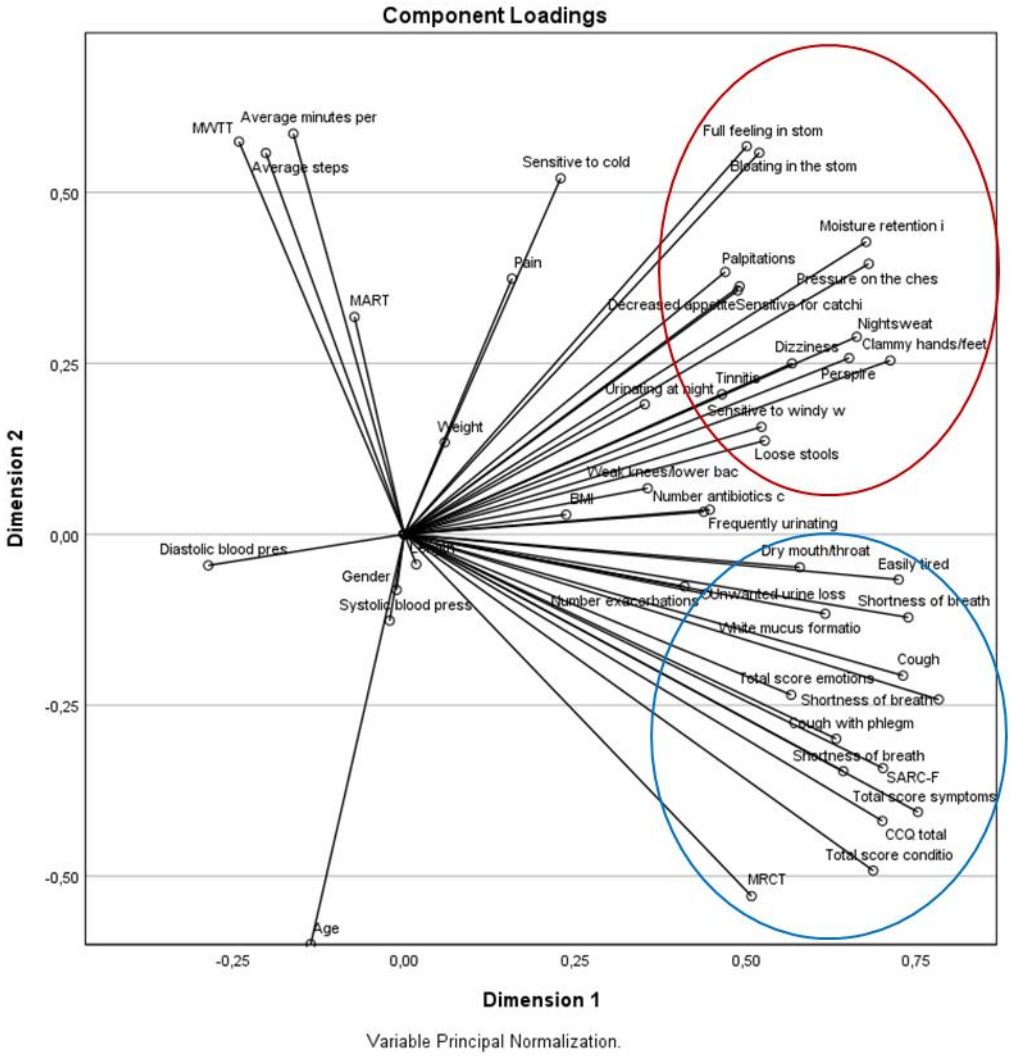


Figure 6 Loading plot dimension 1 versus 2. Two groups of TCM symptoms are circled.

We also asked the TCM expert of the project team to interpret the distribution of the TCM symptoms in Figure 6. The upper right group of symptoms include Stagnation symptoms (eg. palpitations, pressure on the chest). Stagnation of the liver can result in a rebelling stomach, symptoms of which can also be seen in the upper right part (full feeling in stomach, bloating in the stomach). Below this group, more in the center right part, Kidney symptoms appear (loose stools, weak knees/lower back, tinnitus, night sweating, dizziness). Then in the lower right part mainly Deficiency symptoms are clustered (eg. shortness of breath, tired, cough). In this lower part the disease seems to be more severe and deeper into the system, according to TCM theory. This is often seen in somewhat older people.

In the third dimension (Figure 7) the variables weight, BMI, gender and length have large loadings as well as number of antibiotics courses used and number of exacerbations. The third dimension therefore seems to be a more constitutional dimension (weight, gender, length). The number of exacerbations seems to be related with weight and gender.

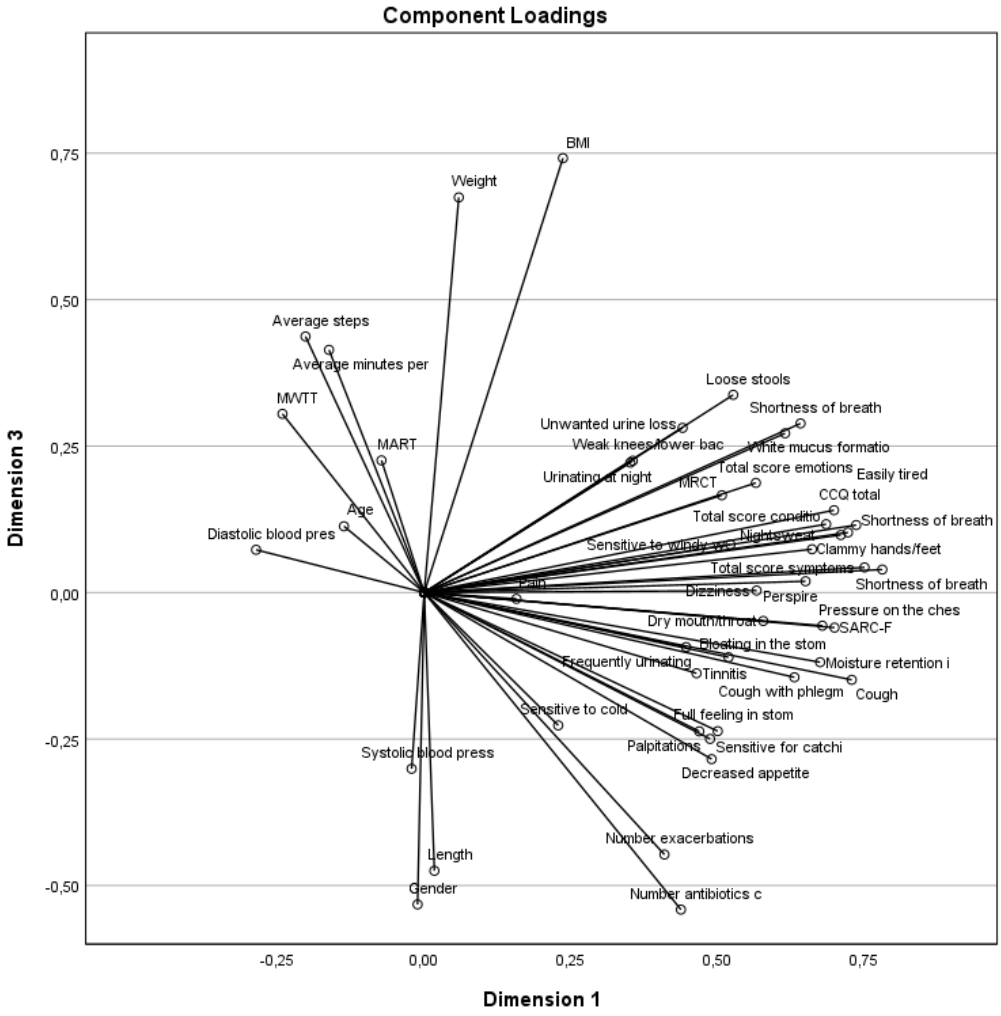


Figure 7 Loading plot dimension 1 versus 3

When the second and third dimension are plotted against each other (see Figure 8), many of the TCM symptoms have a much smaller loading than in the first dimension indicating that symptom severity in general is more related to dimension 1. However, there is one group of symptoms in the lower right part that stands out with large loadings: bloating feeling in the stomach, full feeling in the stomach, sensitive to cold, decreased appetite, and palpitations. These symptoms are related to Stagnation. Age is opposite, indicating that this symptom group occurs mainly in younger patients. This group of symptoms is also opposite to the CCQ scores, indicating a relationship with patients with less strong COPD symptoms and maybe an earlier phase of the disease. In this figure it is also clearly visible that the variables weight, BMI, gender, length, number of antibiotics and number of exacerbations represent the third dimension.

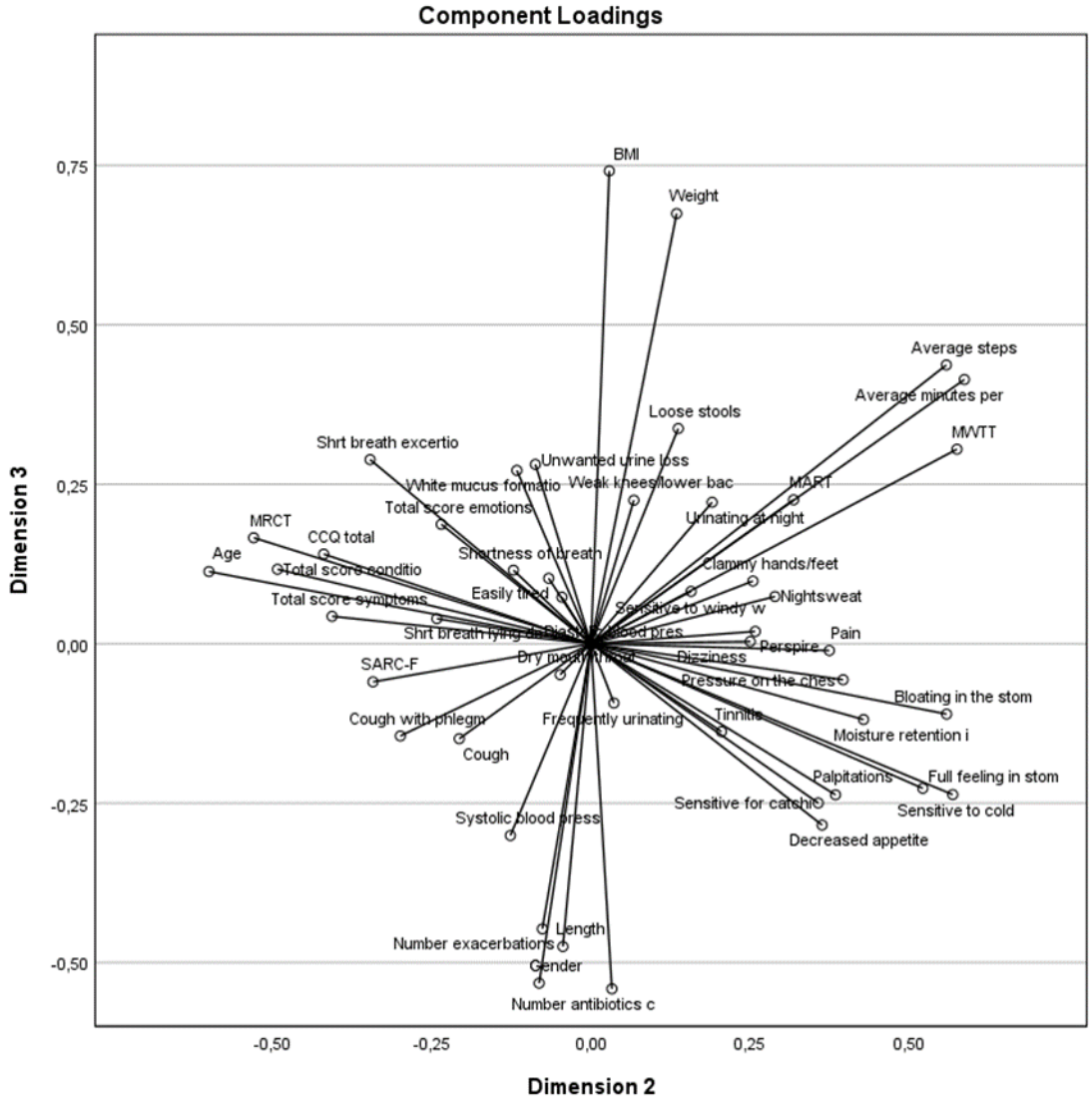


Figure 8 Loading plot dimension 2 versus 3

In the next analysis step the subjects in the score plots were labelled according to their GOLD stage (Figure 9), KNGF profile (Figure 10), TCM subtype according to the expert (Figure 11) and TCM subtype according to the algorithm (Figure 12). The idea is to check whether the subjects cluster together in the score plots based on these labels, and if that is the case, which variables then correspond to these clusters. In the following figures the outer boundaries of the groups of subjects are marked with a colored line to visualize the overlap between groups.

The score plots labelled by GOLD stage (Figure 9) show a large overlap between the three GOLD stages in all three dimensions. In dimension 2 GOLD stage 4 has a smaller variation than the GOLD stages 2 and 3. In the right panel is can be observed that the GOLD stage scores are lower in the upper part and higher in the lower part, suggesting a weak relationship with GOLD stage severity.

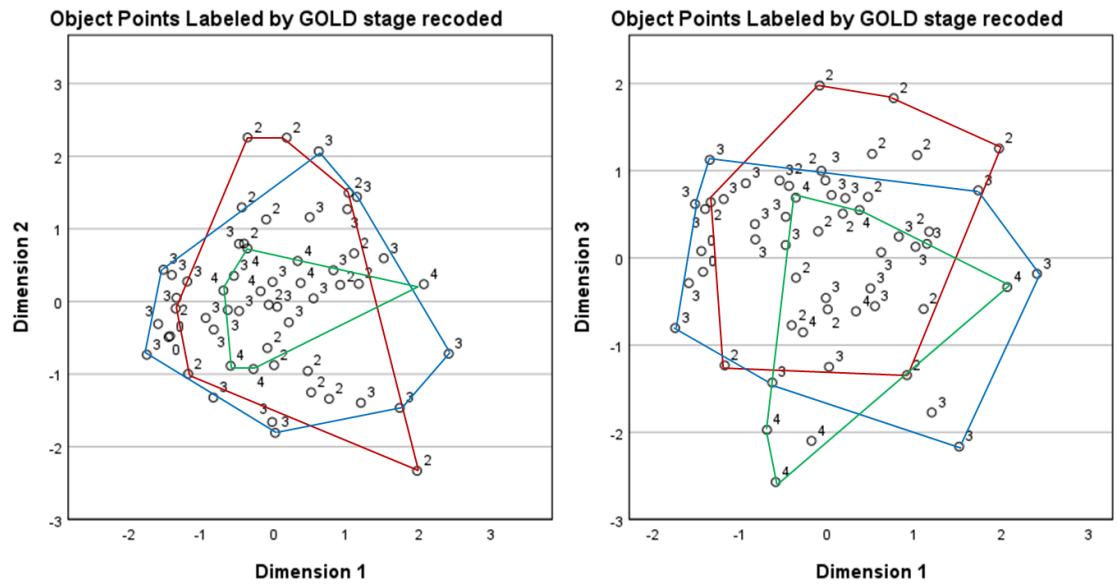


Figure 9 Score plots labelled by GOLD stage. Outer boundaries of each GOLD stage are marked by colored lines (red = 2, blue = 3, green = 4).

The score plots labelled by KNGF profile (Figure 10) also show a large overlap between the KNGF stages in all three dimensions. There is a tendency for a gradient from KNGF 1 to 5 from left to bottom right in both panels. This is conform the loading plot which shows more severity in the lower right part, higher CCQ scores, and lower MWT and activity. This also corresponds to the TCM interpretation that the lower right corner corresponds to more severe Kidney related deficiency symptoms. KNGF group 1 seems to cluster more tightly together, indicating less variation in this group than in the other KNGF groups. However, the KNGF groups are not clearly separate.

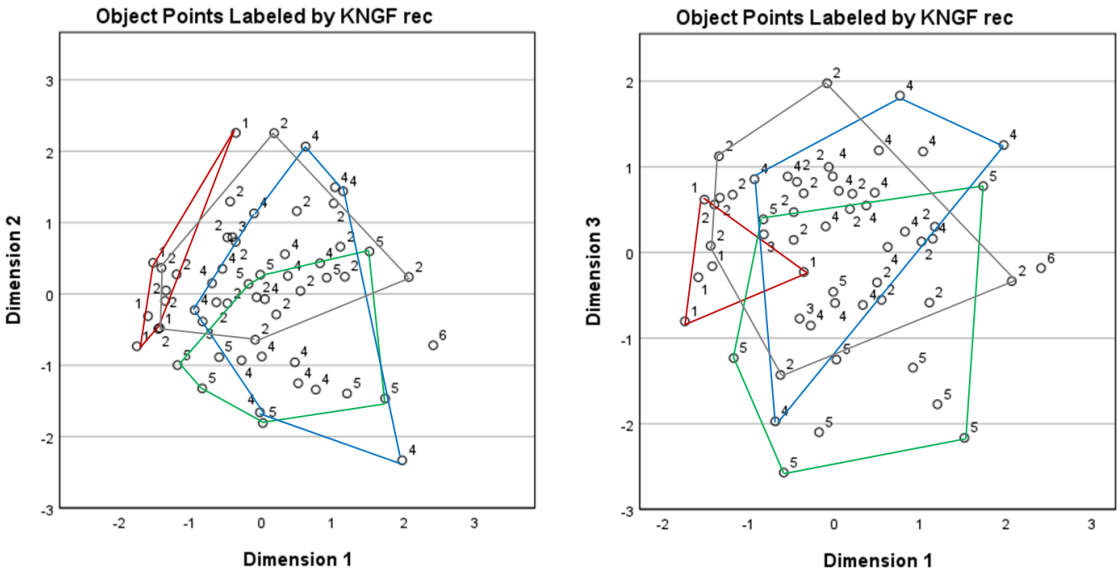


Figure 10 Score plots labelled by KNGF profile. Outer boundaries of each KNGF stage are marked by colored lines (red = 1, grey = 2, blue = 4, green = 5).

The score plots labelled by TCM expert subtype (Figure 11) also show a large overlap. Subtype 1 is closer together as a group, indicating less variation in this group. Subtype 1 is mostly negative in dimension and around 0 in dimension 2, indicating patients with less severe symptoms. The groups 2, 3 and 4 almost completely overlap.

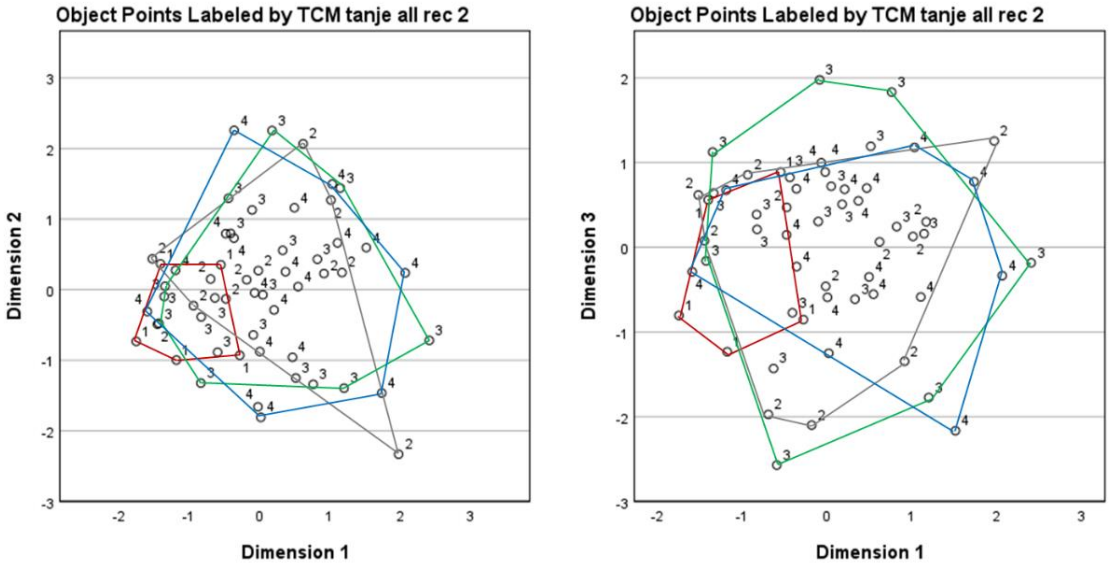


Figure 11 Score plots labelled by TCM expert subtype. Outer boundaries of each Subtype are marked by colored lines (red = 1, grey = 2, green = 3, blue = 4).

The score plots labelled by the TCM algorithm subtype (Figure 12) show some separation in the first dimension, indicating that symptom severity (loadings in first dimension) is the main contributor to the algorithm based TCM subtype, which is also according to theory. There is no difference in the second or third dimension.

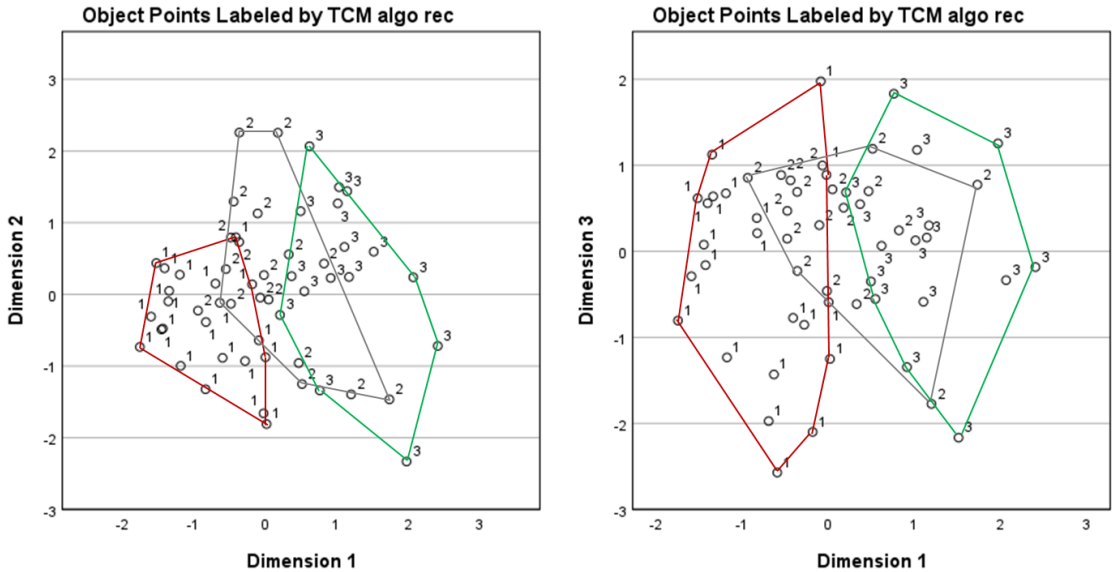


Figure 12 Score plots labelled by TCM algorithm subtype. Outer boundaries of each Subtype are marked by colored lines (red = 1, grey = 2, green = 3).

The above analyses don't show any clear clustering of subjects in the score plots. KNGF severity is related to symptom severity in general. No relationships between symptoms and GOLD stage could be found. The TCM algorithm corresponds to TCM symptom severity. Looking at the distribution of the TCM symptom loadings in Figure 6, there is a clear distinction between a group of symptoms in the upper right part with Stagnation symptoms, the lower right part with Deficiency symptoms, and in between the Kidney symptoms. This indicates that the second dimension could be interpreted as an extra dimension of COPD disease activity and development which is related to the TCM concepts Stagnation and Deficiency, which is not covered by KNGF profiles or GOLD stages.

4.1.2 **Unsupervised CATPCA analysis clinical variables only**

Figure 13 shows the loading plots of the clinical variables. The picture is not very different from the analysis of all variables together described in the previous section. The MWT and activity variables are opposite of the CCQ symptom scores in the first dimension. This represents the largest source of variation in the data. The third dimension again is gender, length, and age. Now weight number of exacerbations and number of antibiotics courses is more related to the second dimension and the first dimension. MART measured with the physical activity questionnaire stands out in the second dimension.

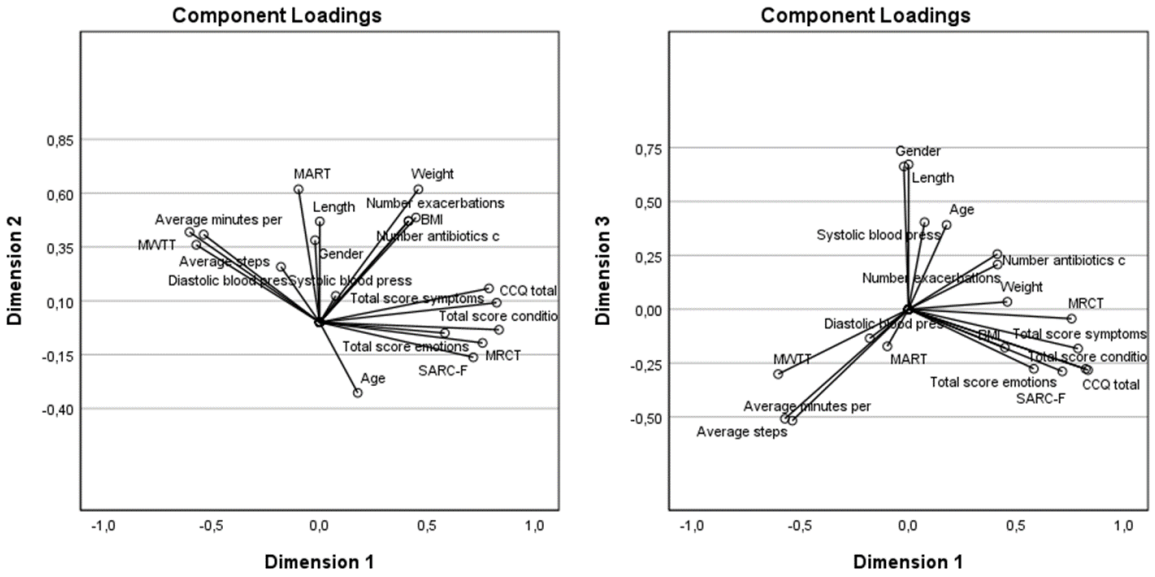


Figure 13 Loading plots of the clinical variables

The score plots of this analysis don't show any clear groupings in GOLD stage, KNGF profiles, and TCM subtypes expert and algorithm (see Figure 14, other data not shown). The left panel of Figure 14 shows that subjects with subtype 1 are closer together and centered in the lower left corner, indicating less severe symptoms and more activity.

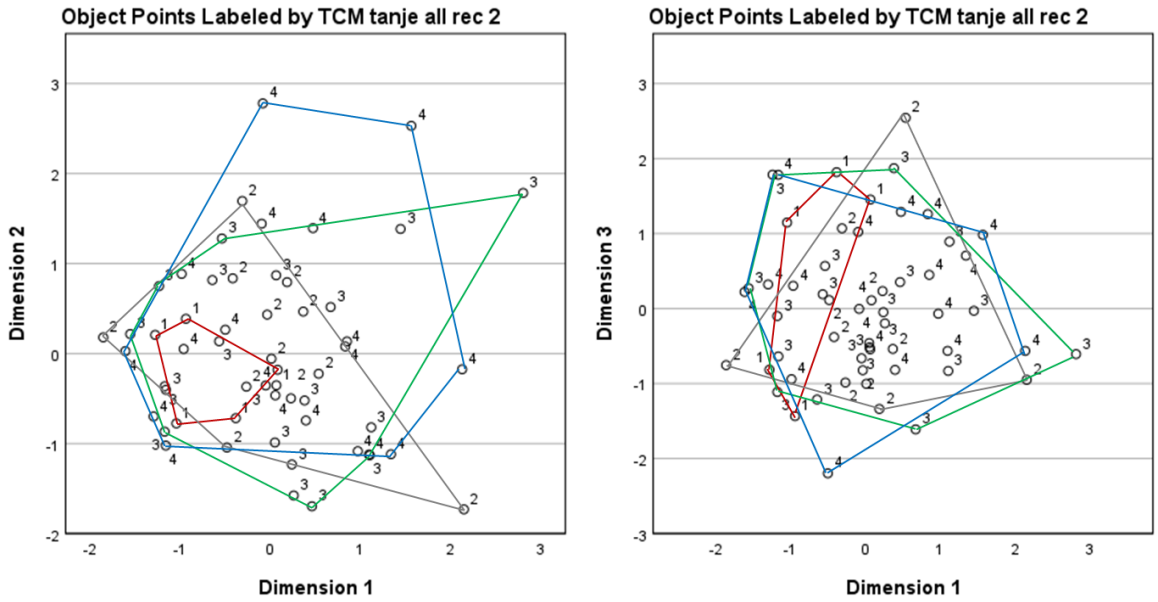


Figure 14 Score plots of the clinical variables by TCM expert subtype. Outer boundaries of each Subtype are marked by colored lines (red = 1, grey = 2, green = 3, blue = 4).

4.1.3 **Unsupervised CATPCA analysis TCM symptoms only**

Figure 15 shows the loading plots of the TCM symptoms, without the clinical variables. The loadings are colored according to the relationship of the TCM symptom with a TCM subgroups according to theory (1 = red: Lung Qi Deficiency, 2 = grey: Lung and Spleen Qi Deficiency, 3 = green: Lung and Spleen Qi Deficiency and Kidney Yang Deficiency, 4 = blue: Lung and Spleen Qi Deficiency and Kidney Yang and Yin Deficiency). The first dimension is clearly related to symptom severity. In the second dimension the symptoms cover a wide range of values. The red variables mostly have a negative loading in the second dimension, the grey ones a positive loading, the blue ones are not related to the second dimension, and the green ones are spread out.

The distribution of the TCM symptoms is very similar to the distribution of Figure 6, suggesting that adding the clinical variables to the analysis doesn't have a big effect on the TCM symptom distribution. What can be observed in this figure is that the original subtyping of COPD patients into 4 TCM subtypes is not visible in clearly separate clusters of symptom loadings. However, when looking with new eyes at the symptoms, the upper part can be related to Stagnation symptoms and the lower part to Deficiency symptoms without Stagnation. In the center part are the Kidney related symptoms.

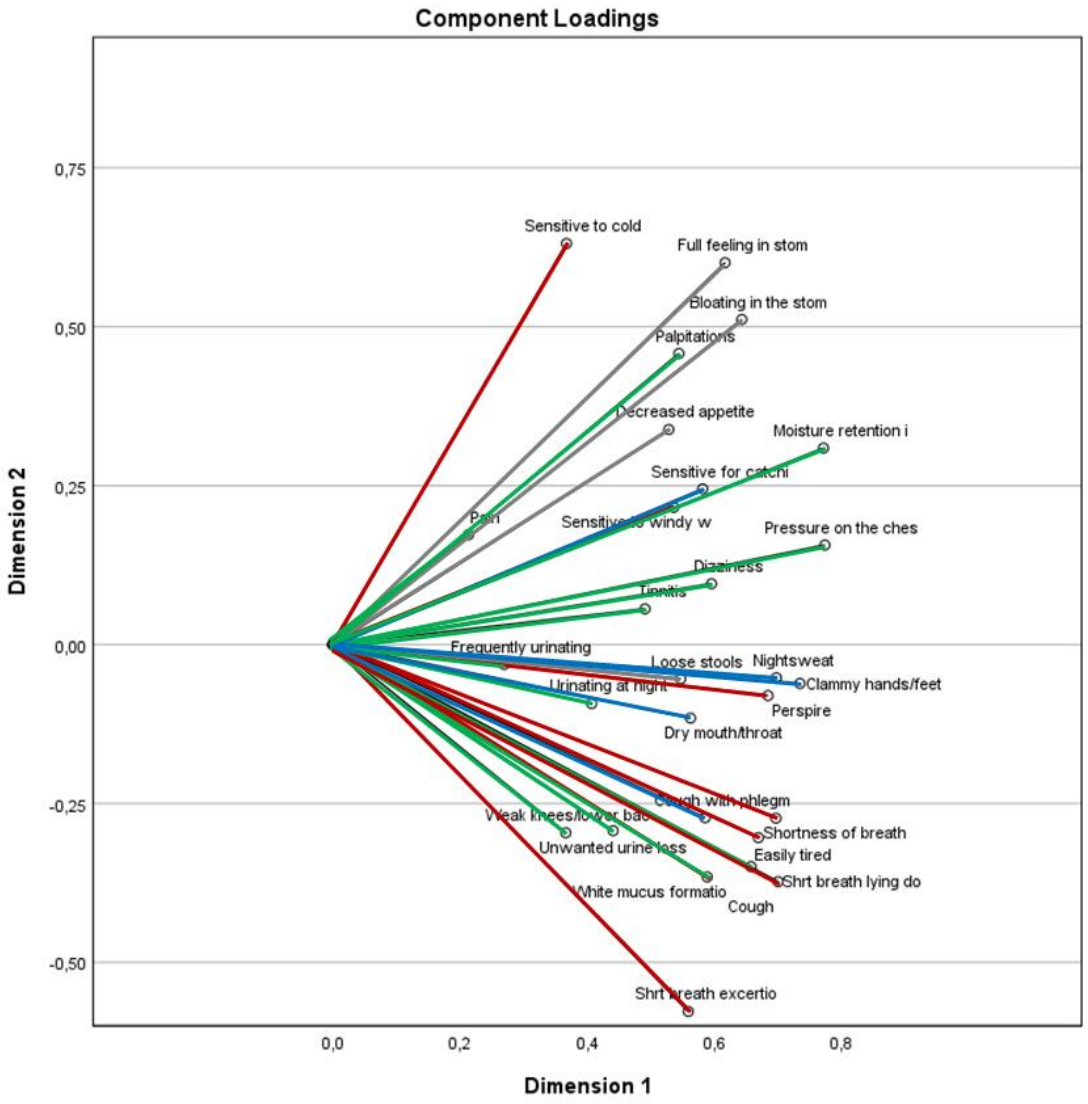


Figure 15 Loadings of the TCM symptoms (dimension 1 versus 2), colored by TCM expert subtype (1 = red, 2 = grey, 3 = green, 4 = blue).

In Figure 16 the loadings on the first and third dimension are visualized. There is a large variation in the third dimension. The grey variables (Lung and Spleen Qi Deficiency) are more in the center here, indicating these are not related to the third dimension. Some of the blue variables (Lung and Spleen Qi Deficiency and Kidney Yang and Yin Deficiency) are negative in dimension 3. The green variables are spread out, but the positive variables in dimension 3 are green (Lung and Spleen Qi Deficiency and Kidney Yang Deficiency).

When the TCM expert looked at the TCM symptom distribution with fresh eyes, the lower right part shows more Kidney Yin Deficiency symptoms. The upper right part shows more Yang Deficiency symptoms (frequent urination, urination at night, pain, unwanted urine loss, tired). In the center right there are more general Yin Deficiency symptoms and drying of the fluids.

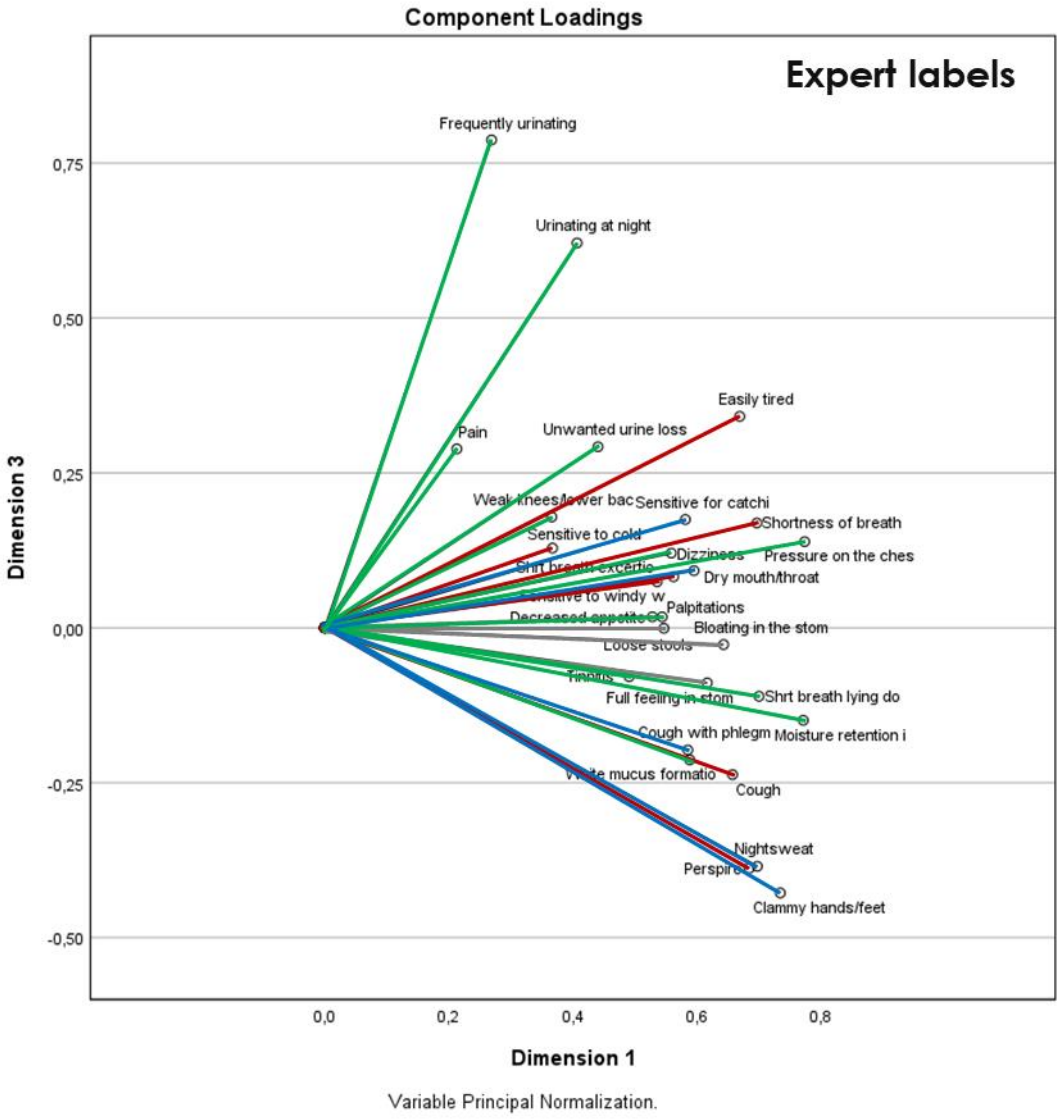


Figure 16

Loadings of the TCM symptoms (dimension 1 versus 3), colored by TCM expert subtype (1 = red, 2 = grey, 3 = green, 4 = blue).

In Figure 17 the loading plot is shown of the second versus the third dimension. This figure confirms the above two loading plots. The red and grey variables are mostly related to the second dimension, the red ones negative and the grey ones positive. The blue variables are more related to the third dimension. And some green variables are large in the third dimension. In this Figure a similar TCM symptom distribution can be observed as in Figures 15 and 16.

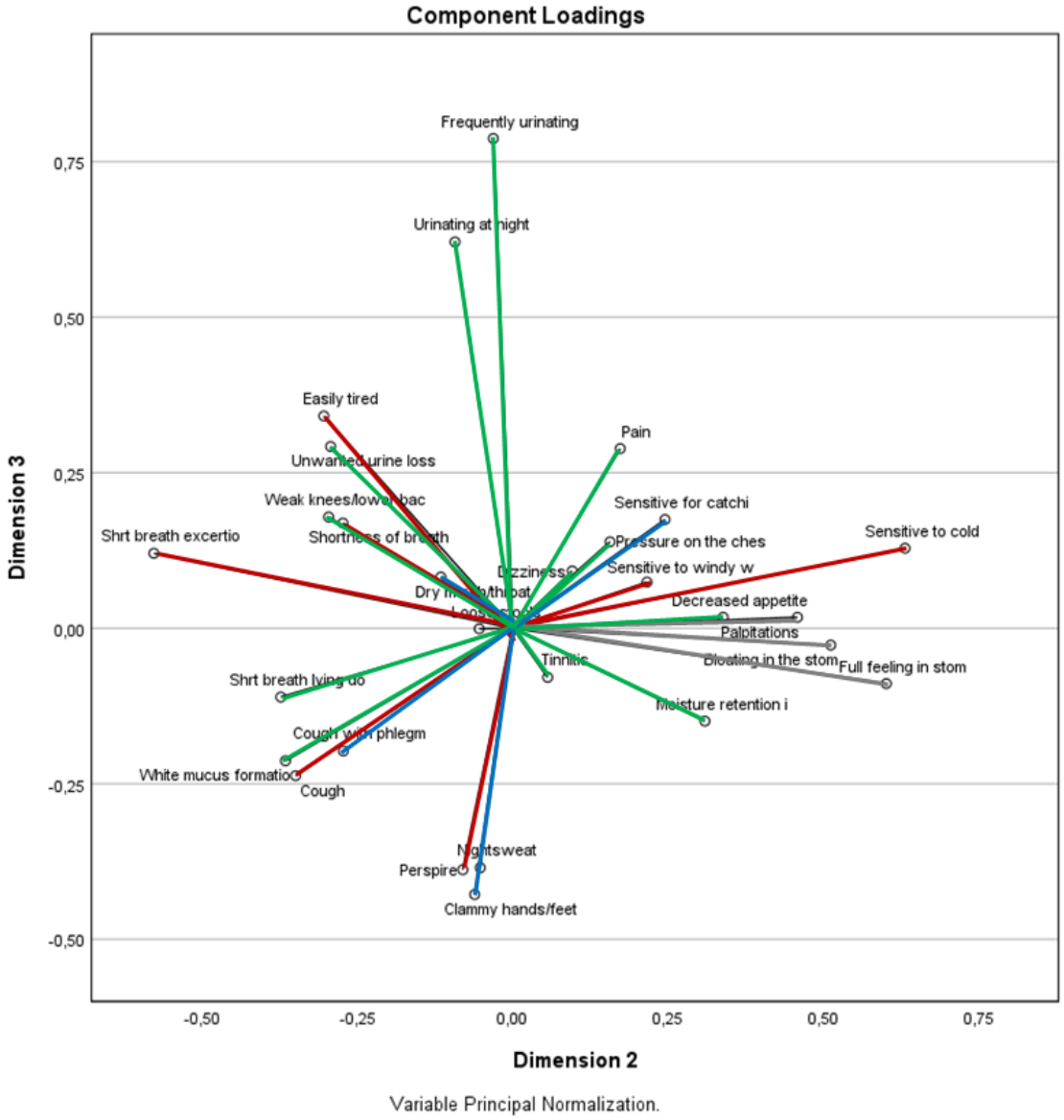


Figure 17 Loadings of the TCM symptoms (dimension 2 versus 3), colored by TCM expert subtype (1 = red, 2 = grey, 3 = green, 4 = blue).

The score plots for the TCM symptoms are very similar to the score plots in which all variables were included (data not shown). No clear groupings of subjects are seen. Only TCM expert subgroup 1 is again more tight together and scores low in dimension 1 and 2, similar to the findings described above. Similar to the analysis with all variables, again a distinction can be made in groups of COPD patients with Stagnation symptoms and patients with mainly Deficiency symptoms without Stagnation.

4.1.4 **Forced classification CATPCA analysis KNGF profiles**

In order to find out which variables are most related to the KNGF profiles a forced classification approach using CATPCA was conducted. This is an explorative method allowing weighing of

variables which then have a larger or smaller effect on the model. In our analysis the KNGF profile variable received a weight of 20. The score plots shown in Figure 18 clearly show that this method separates the subjects into KNGF profile in de first dimension. KNGF profile 1 is separate from 2 which are separate from 3, 4 and 5. KNGF profiles 3, 4 and 5 are so similar that they remain together as one group, they also don't separate in dimension 2 or 3.

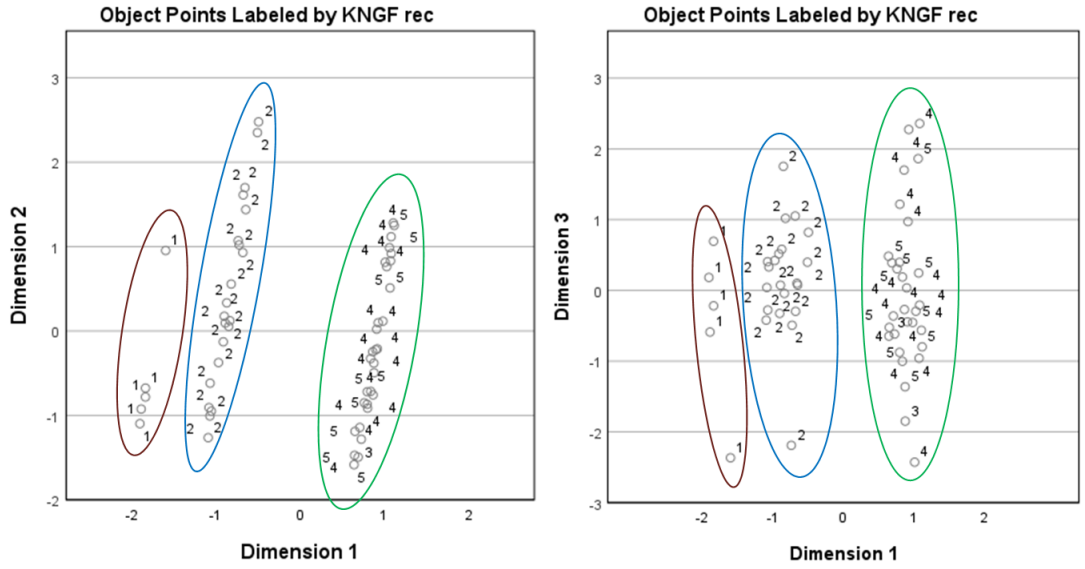


Figure 18 Score plots forced classification with KNGF profile weight 20. Groups of subjects are marked by colored circles per KNGF stage (red = 1, grey = 2, green = 4 & 5).

In Figure 19 the loading plot of the first and second dimension is shown. In this figure is becomes clear that KNGF profile almost entirely represents the first dimension. Some symptoms also point towards the first dimension, for instance easily tired, shortness of breath and dry mouth and throat. Overall the TCM symptoms seems to represent the second dimension, and thus are orthogonal to the KNGF profiles. This indicates the the TCM symptom information represents other sources of variation in the patient group, which is little related to the KNGF profiles.

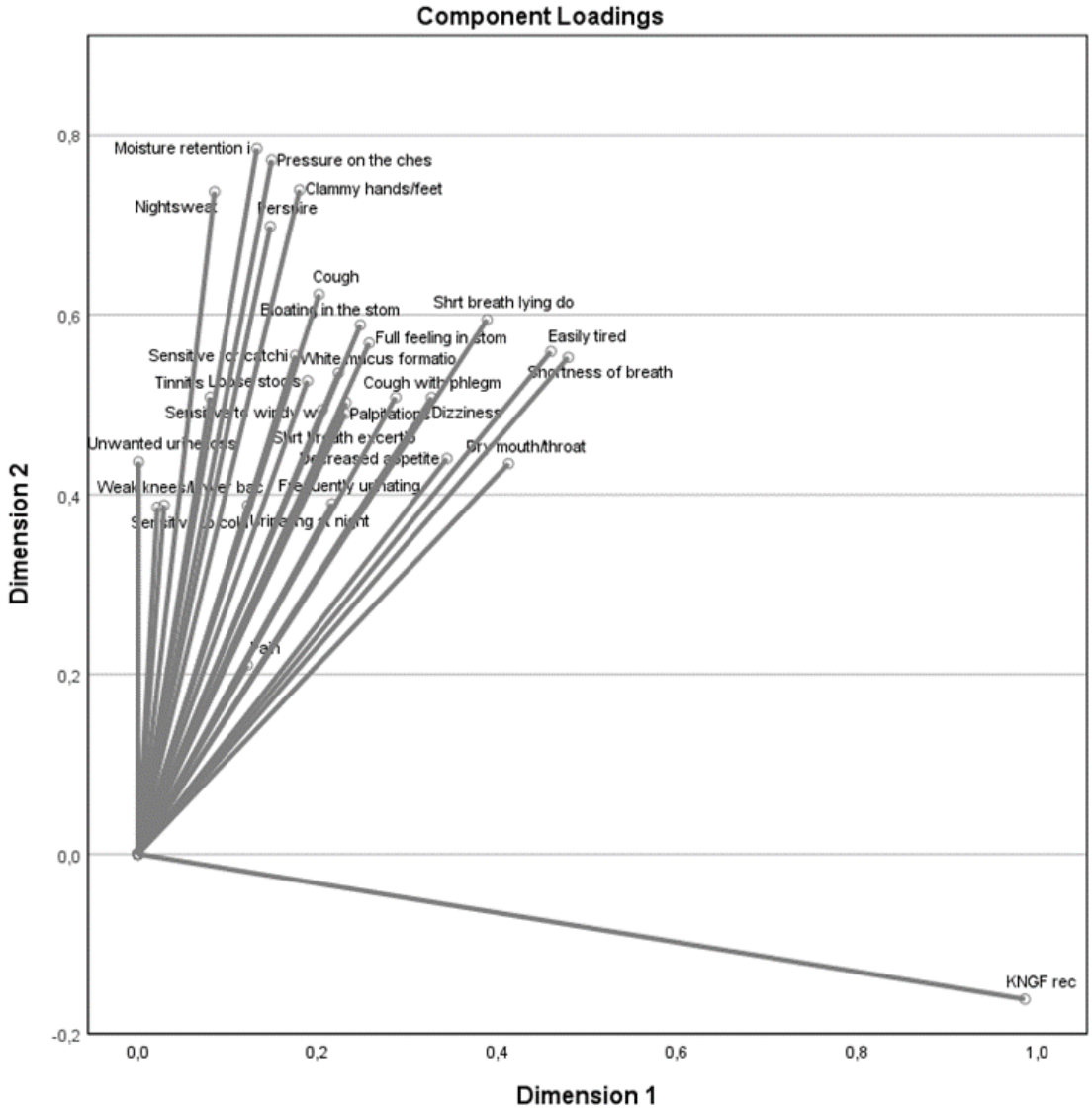


Figure 19 Loading plot forced classification KNGF profile weight 20, first and second dimension.

In Figure 20 the loading plot of the first and third dimension are shown. In this figure it can be observed that the TCM symptoms are both pointing up and down in the third dimension, indicating that there are distinct subsets of symptoms related to patients scoring high in dimension 3 and those scoring low in dimension 3. In this figure the separation between Stagnation (lower part) and Deficiency without Stagnation (upper part) symptoms can be observed, similar to the previous analyses.

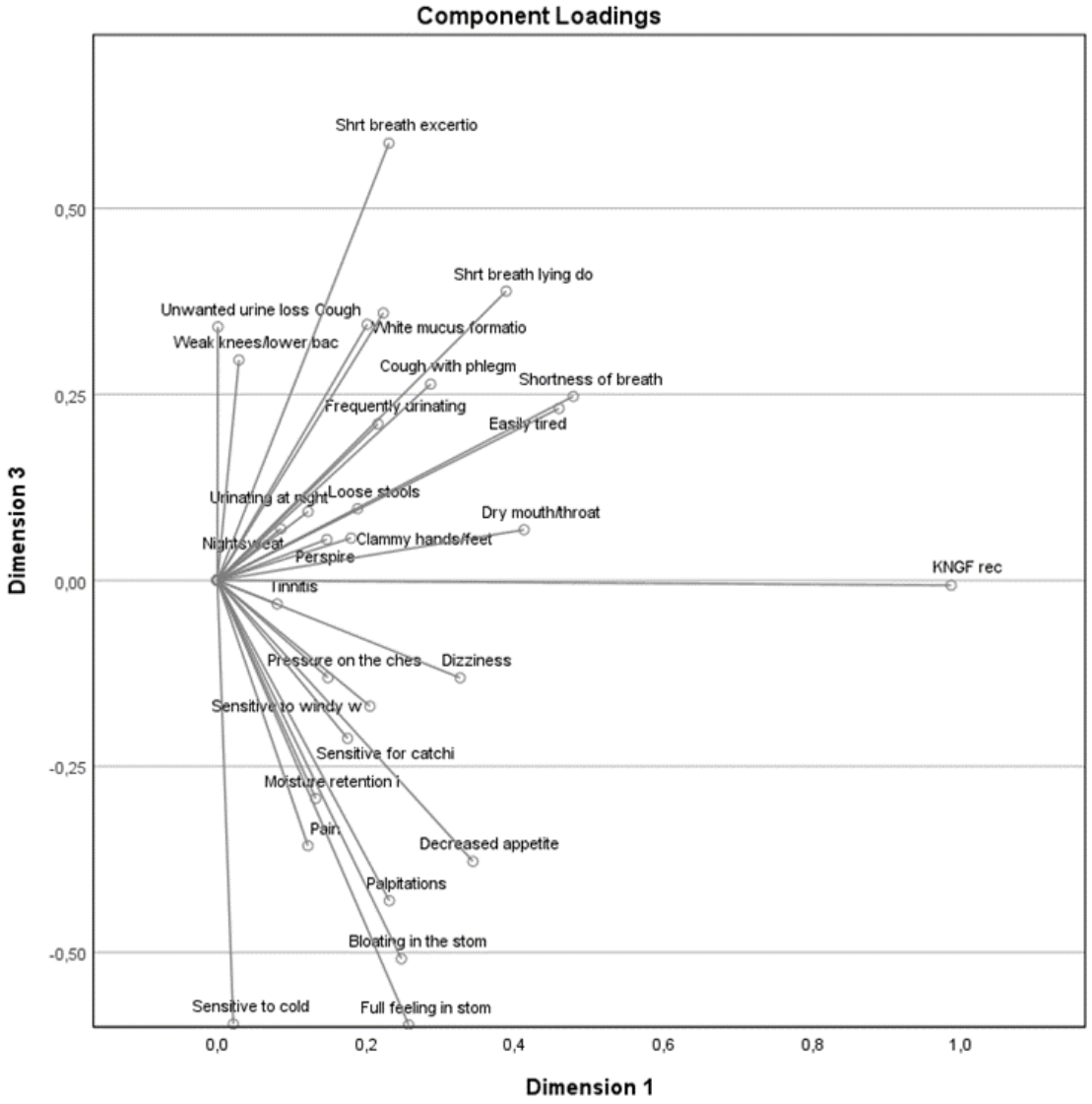


Figure 20 Loading plot forced classification KNGF profile weight 20, first and third dimension.

4.1.5 **Forced classification CATPCA analysis TCM expert subtypes**

Then a similar forced classification approach is conducted to find relationships between the TCM expert classification and the clinical variables. The TCM expert classification is given a weight of 10. In Figure 21 the score plots show a separation of the TCM expert subtypes into groups. One is separate from 2 and 3, and separate from 4. TCM subtypes 2 and 3 are together in one group, indicating that these subtypes are more similar.

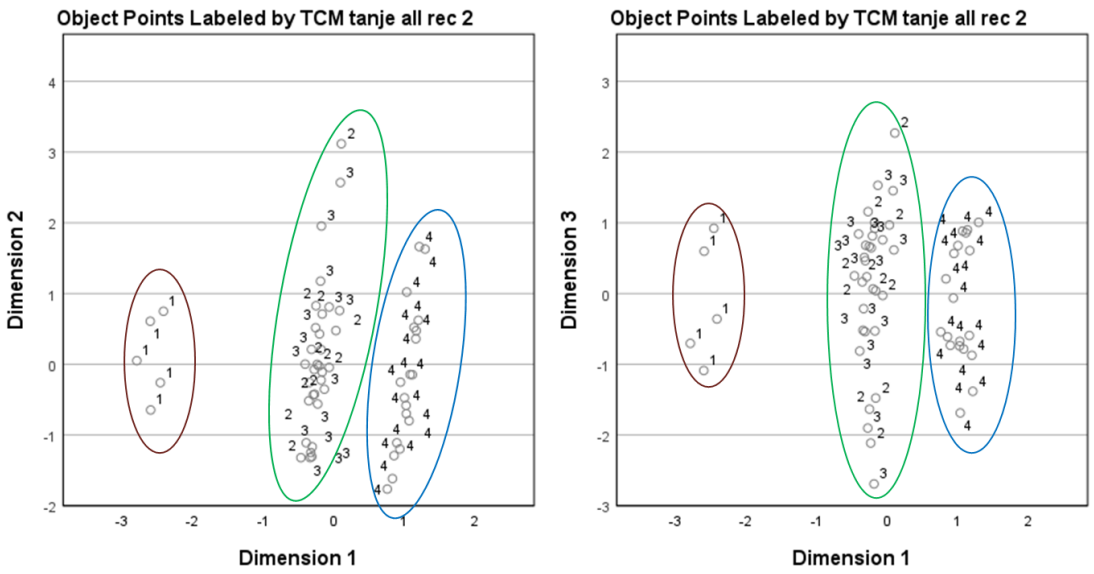


Figure 21 Score plots forced classification TCM expert subtype weight 10. Groups of subjects are marked by colored circles per TCM expert subgroup (red = 1, green = 2 & 3, blue = 4).

The loading plot of the first and second dimension is shown in Figure 22. The clinical variables most related to the TCM expert subtypes (dimension 1) are blood pressure and BMI, sarcopenia (SARC-F) and then after that number of antibiotics courses and number of exacerbations.

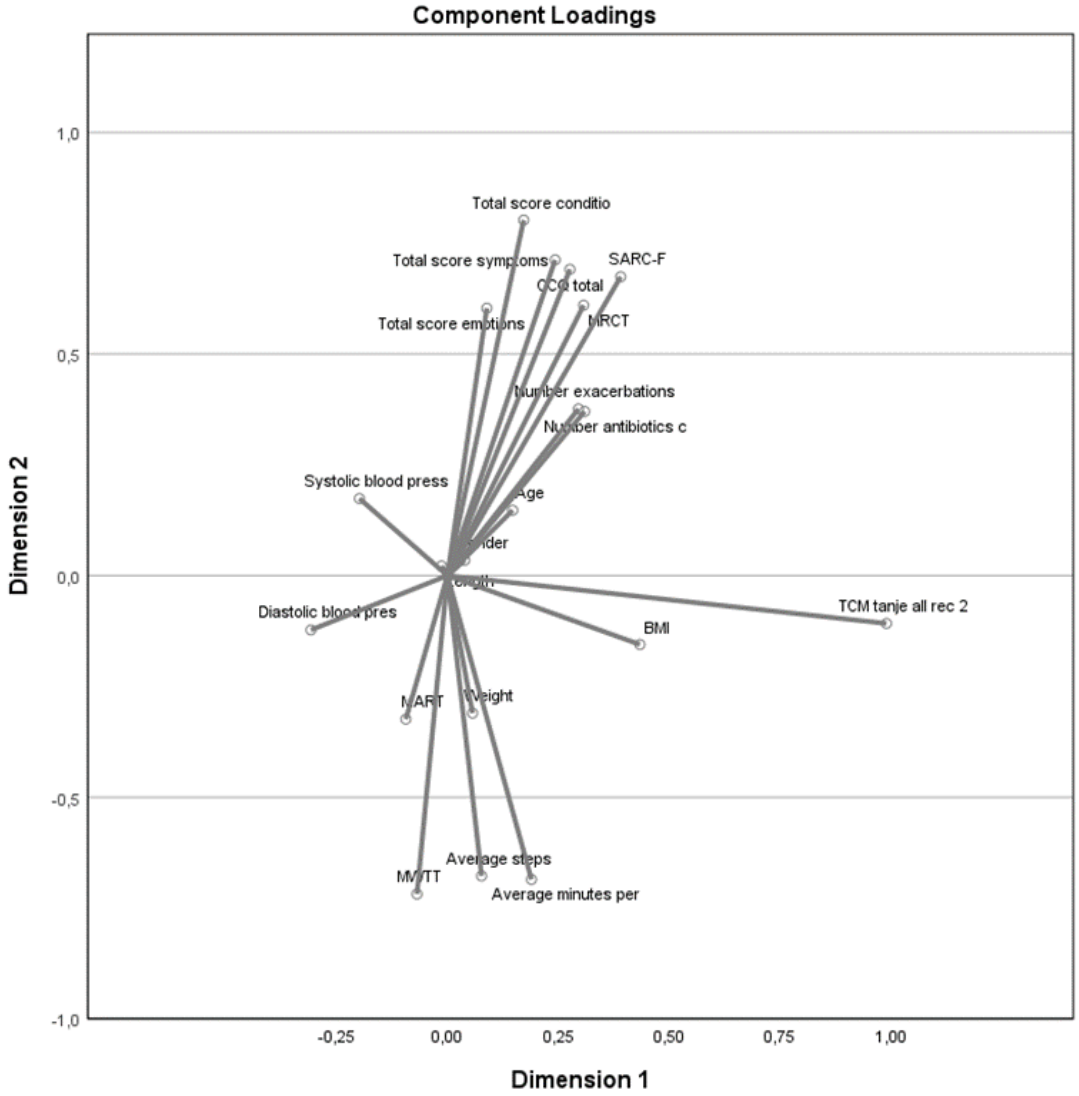


Figure 22 Loading plot forced classification TCM expert subtype weight 10, first and second dimension.

The loading plot of the first and third dimension is similar to the plot resulting from the clinical variables without forced classification (see Figure 23).

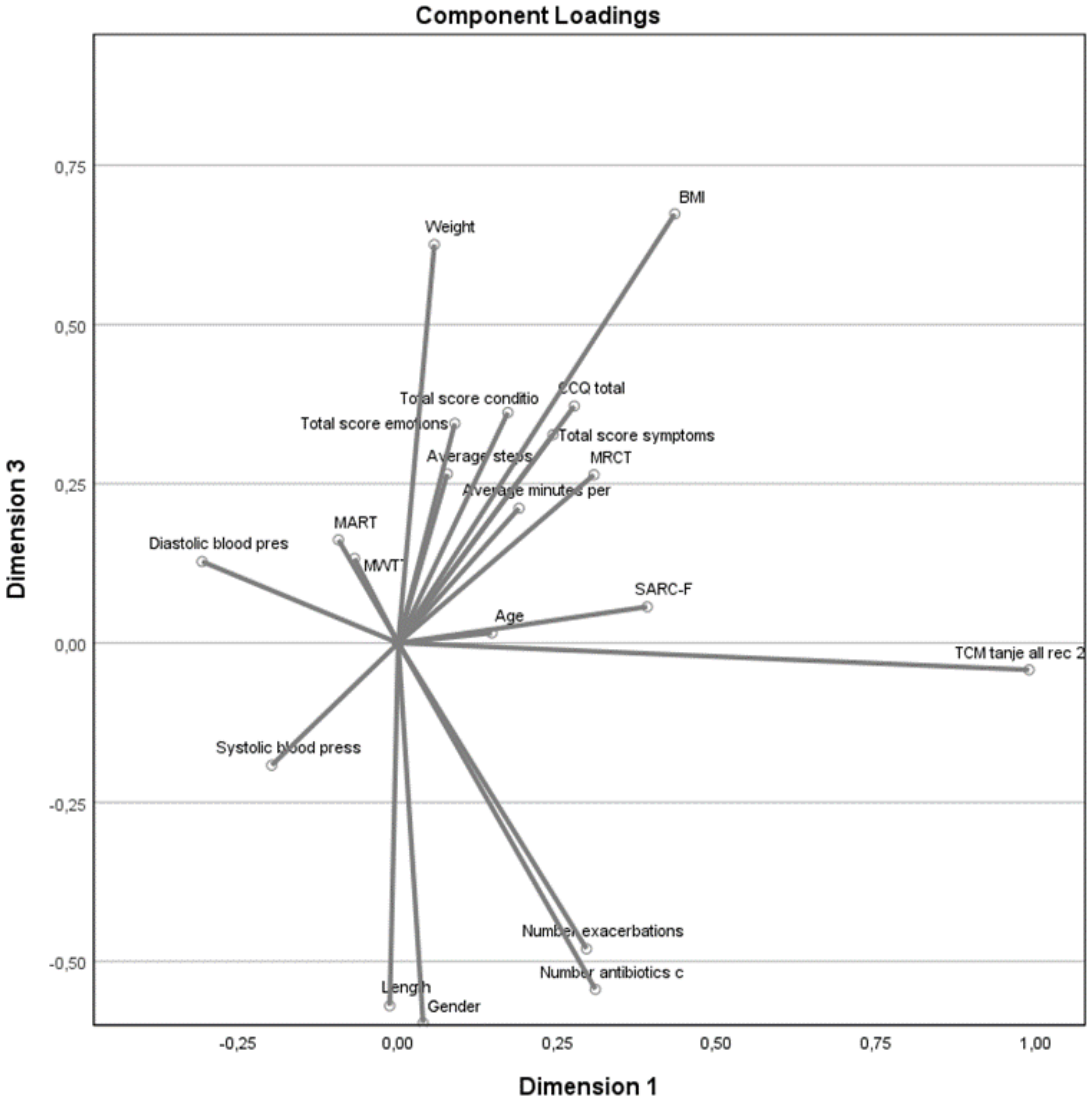


Figure 23 Loading plot forced classification TCM expert subtype weight 10, first and third dimension.

4.1.6 **Network analysis of the symptoms and clinical variables**

A network view was constructed to get a different visual representation of how the various TCM symptoms and clinical variables are related to each other (Figure 24). All correlations of transformed variables >0.40 and < -0.40 were selected. A perfuse force directed layout was used to organize the network in such a manner that the most connected variables are positioned close together. In the network visualization the clinical symptoms are represented with an hexagon, the TCM symptoms with a circle. Blue lines are negative correlations, red ones are positive correlations. The TCM symptoms are colored to represent the TCM subtype they are related to according to theory (1 = red, 2 = grey, 3 = green, 4 = blue).

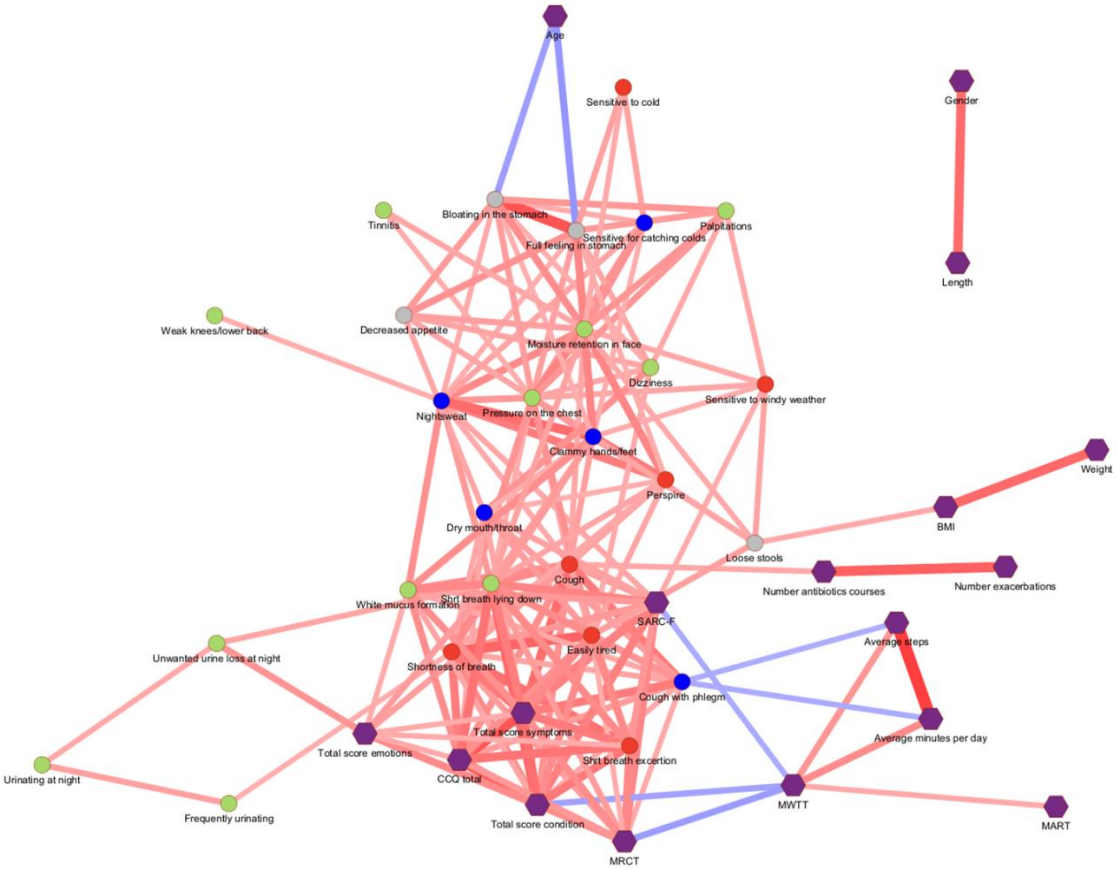


Figure 24 Network visualisation of the TCM symptoms (circles) and clinical variables (hexagons). Red lines are positive correlations, blue lines are negative correlations. Symptoms are colored according to TCM subtype (1 = red, 2 = grey, 3 = green, 4 = blue).

Figure 24 shows that the clinical variables are positioned together, except age. The TCM symptoms also group together, indicating that the clinical and TCM perspective are different and contribute unique information about COPD. However, there are also symptoms and clinical variables connected to each other, indicating the overlap between the TCM and clinical perspective. The TCM symptoms of subtype 1 are grouped together, which is the mildest variation of COPD, those with Lung Qi deficiency only. The grey symptoms are also relatively close together, representing Lung and Spleen Qi deficiency. Subtype 1 is close to the clinical variables, subtype 2 is in the upper part of the network, while subtype 3 and 4 (related to Kidney deficiency) are more in the center.

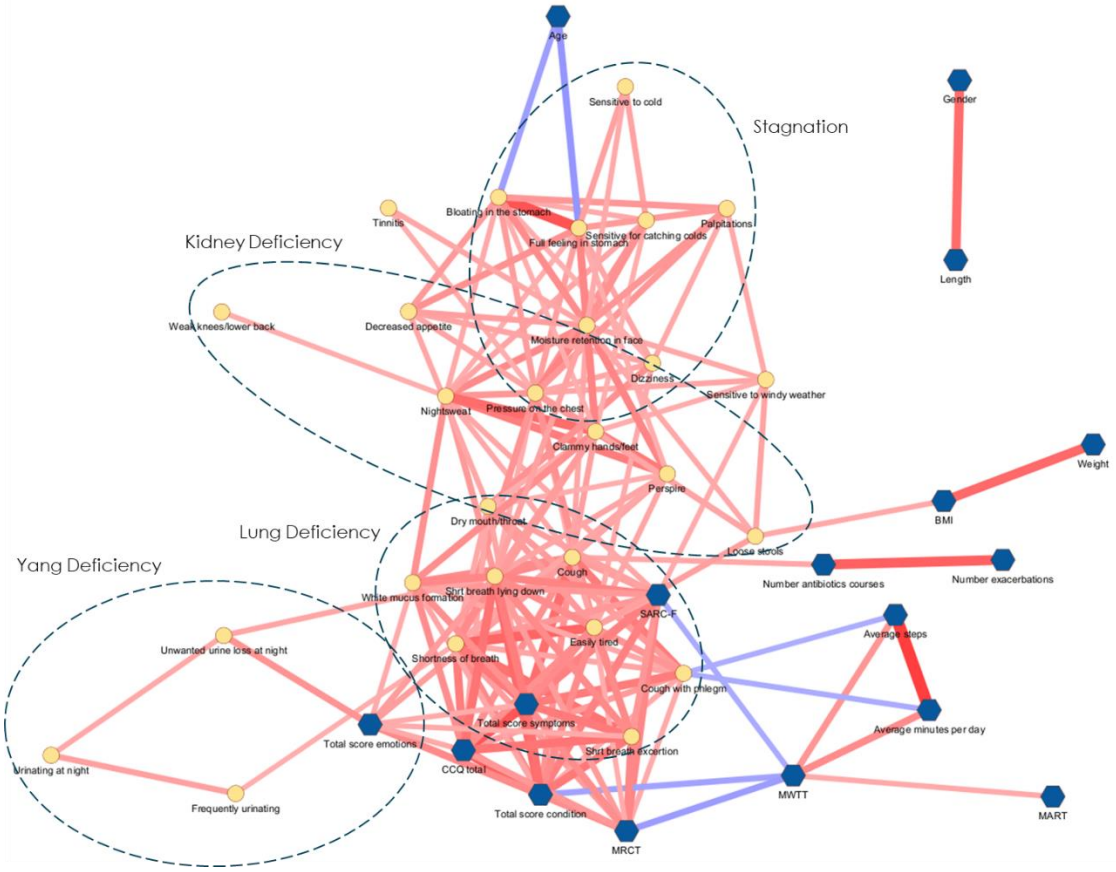


Figure 25 Network visualisation of the TCM symptoms (circles) and clinical variables (hexagons). Red lines are positive correlations, blue lines are negative correlations. Clusters of TCM symptoms are circled.

When the clustering of TCM symptoms is considered apart from the predefined TCM subtypes, a similar distribution can be observed as in the CATPCA analyses above. Stagnation related symptoms cluster in the top part, and are negatively correlated with age. Kidney Deficiency related symptoms are below the Stagnation symptoms. Yang Deficiency symptoms are in the bottom left part, while the Lung Deficiency symptoms are in the center lower part.

4.1.7 **Conclusions comparing Western and Chinese COPD subgroups**

The exploratory data analysis shows that there are no clear subgroups in the included population of COPD patients. The main sources of variation between the patients are: 1) severity of the symptoms (dimension 1), 2) KNGF profile related scores on 6 minute walk test, activity and CCQ (dimension 2), and 3) BMI, gender, antibiotics use and exacerbation related (dimension 3).

The analysis doesn't show groups of subjects according to GOLD stage, suggesting that there are no TCM symptoms or clinical variables clearly representing GOLD stage in the dataset. For the KNGF profile there is an indication for a relationship with TCM symptom severity and severity of the clinical variables. The TCM expert subtype 1 is associated with less scores on TCM symptoms overall, and is related to a relatively mild stage of COPD. The other three subtypes cannot be separated. The TCM algorithm seems to provide scores related to TCM symptom severity overall.

It is interesting to note that the TCM symptoms provide an additional source of variation to the COPD patients that is different from the KNGF profiles and GOLD stages routinely used. Both the unsupervised CATPCA analysis as well as the forced classification analyses and the network analysis point this out. The main two clusters of TCM symptoms are: Stagnation related symptoms and Deficiency related symptoms. One level deeper Yang Deficiency and Kidney Yin Deficiency patterns can be distinguished.

This finding is supported by several randomized clinical studies in which TCM medicines are provided in addition to conventional COPD medication based on TCM symptom patterns (Chen 2023, Wang 2014, Li 2012, Li 2013). Providing additional TCM medication based on symptom patterns has been found to improve several outcomes compared to conventional COPD treatment only, for instance Li et al. report lower frequency of acute exacerbations, shorter duration of exacerbations, larger FEV, and improved quality of life (Li 2012). Wang et al. report similar results and also improved 6 minute walk test results (Wang 2014). These clinical findings indicate that it could be of great benefit to COPD patients to implement TCM pattern diagnosis and personalized medication for these patterns in COPD treatment strategies.

4.2 Biological interpretation of COPD subgroups with metabolomics

4.2.1 GOLD stages are associated with oxylipins, lysophospholipids and bile acids

To explore which and how signaling lipids are associated with Western and Chinese COPD subgroups, diverse statistical methods were performed. Ordinal regression model results reveal the a positive relationship between 3 LPSs, 3 oxylipins and 1 sphingolipid and GOLD stages (Figure 25A). With the increased relative concentration of 8 targets, a tendency towards a higher GOLD stage can be observed. Moreover, different metabolites and metabolite ratios are found to vary significantly among the GOLD stages (Figure 25B) based on ANCOVA results. There is a trend of a mild decline followed by an increase observed in LPE (18:1) and DCA. Additionally, the 2 oxylipins ratios 16-HDoHE/DHA and HDoHEs/DHA display increased abundance with higher GOLD stage, in which the variation of HDoHEs/DHA shares the same trend as found in the ordinal regression analysis. For these 4 metabolites (LPE (18:1), DCA, 16HDoHE/DHA and HDoHEs/DHA), the plasma of patients from stage 4 exhibit obviously higher concentrations compared with other stages. On the contrary, 4 bile acids ratios and sphinganine-1-phosphate (18:0) are more enriched in stage 3 while stage 2 and stage 4 do not show a significant difference when comparing these 2 groups.

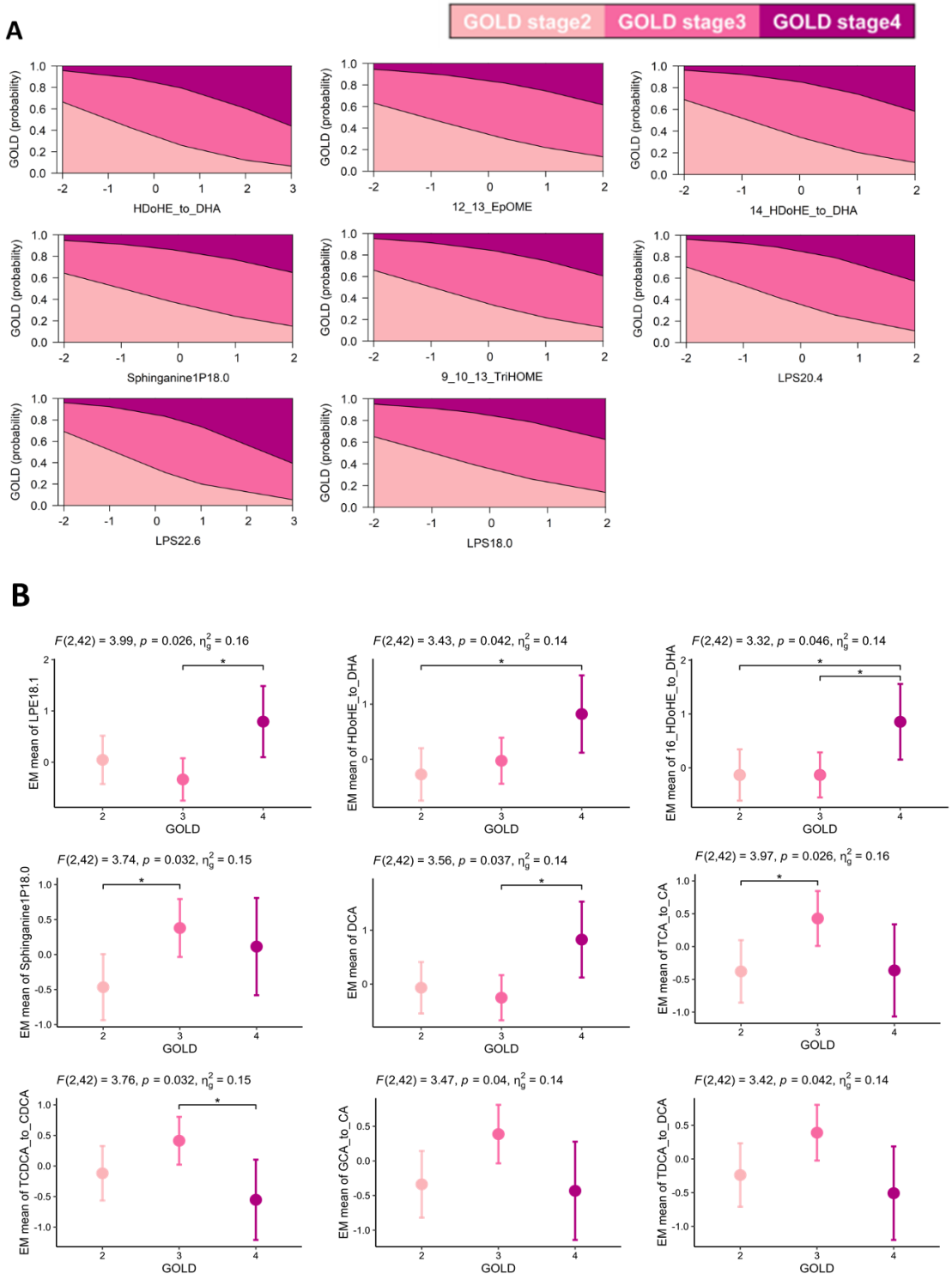


Figure 26 Signaling lipids associated with COPD GOLD stages. (A) The results of the Ordinal Regression Model revealed the contribution of signaling lipids to GOLD stage. The Ordinal Regression Model is presented by the probability distribution of different GOLD stages, indicating the likelihood of each patient being in different GOLD stages based on metabolite changes. (B) The results of the Analysis of Covariance indicated the variation of signaling lipids among GOLD stages. All the targets have been examined but only targets with a p-value less than 0.05 were displayed. Analysis of Covariance is presented by estimated marginal mean with the high and low end of confidence interval.

4.2.2 KNGF profiles are associated with oxylipins and lysophospholipids

To investigate the relationship between KNGF profiles and signaling lipids, the same statistical analyses were conducted as for the GOLD stage analysis. The ordinal regression models indicate that increases in LPA (14:0), 8,12-iso-iPF2a-VI, GUDCA, and TUDCA contribute to the possibility of being classified into a higher KNGF profile (Figure 26A). As for the ANCOVA results, the targets classified within the same lipid subclass exhibit similar variation patterns among different KNGF profiles (Figure 26B). For instance, the trend of increase followed by a decrease is found in LPE (16:0), LPE (18:0) and LPE (22:4). Regarding LPA (14:0) and LPA (16:0), the abundance of the targets is higher in profile 4 patients' plasma compared to other profiles. In addition, the relative concentrations of LPS (18:0) and LPS (20:4) increase in profile 2 but subsequently decrease in profile 4 and profile 5, which trend is also found in sphingosine (18:1). For the oxylipins, ratios of HEPEs/EPA and 14-HDoHE/DHA, and PGK2 exhibit tilde-shape shifting trends from lower to higher profiles except 8,12-iso-iPF2a-VI, which is constantly increased. Importantly, the LPA (14:0) and 8,12-iso-iPF2a-VI demonstrate a close association with KNGF profiles since they are identified in both statistical analyses.

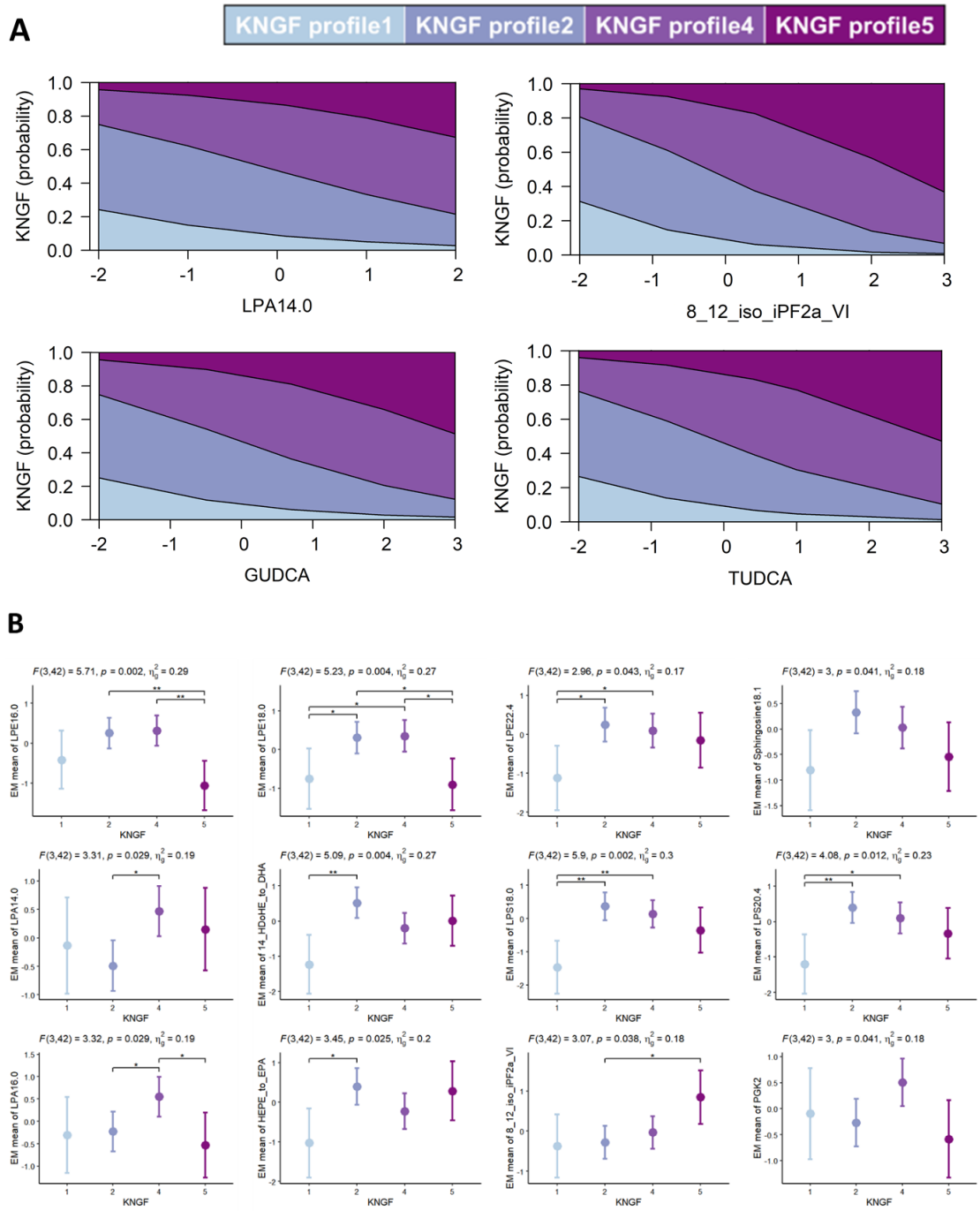


Figure 27 Signaling lipids associated with COPD KNGF profiles. (A) The results of the Ordinal Regression Model revealed signaling lipids contribution to the KNGF profiles. Ordinal Regression Model is presented by the probability distribution of different KNGF profiles, indicating the likelihood of each patient being in different KNGF profiles based on metabolite changes. (B) The results of the Analysis of Covariance indicated the variation of signaling lipids among the KNGF profiles. All the targets have been examined but only targets with a p-value less than 0.05 were displayed. Analysis of Covariance is presented by estimated marginal mean with the high and low end of confidence interval.

4.2.3 TCM subtypes are related to specific oxylipin targets and lysophospholipids

In addition to examining relations of lipids to GOLD stages and KNGF profiles, the possible link between signaling lipids and TCM algorithm subtypes was investigated. Figure 27A shows that LPS (18:0) and 4 oxylipins increase alongside the transformation from subtype 1 to subtype 3.

Meanwhile, ANCOVA analysis displays the variation of signaling lipids in different TCM algorithm subtypes (Figure 27B). For 13-HOTE, PGF3a, Sphingosine-1-phosphate (16:1) and thromboxane-B2, their relative concentrations in subtype 2 are higher than subtype 1 and subtype 3 while LPS (18:0) shows high abundance in both subtype 2 and subtype 3. Notably, LPS (18:0), 12-HETE, and 14-HDoHE consistently show a strong connection with the TCM algorithm subtypes across analyses.

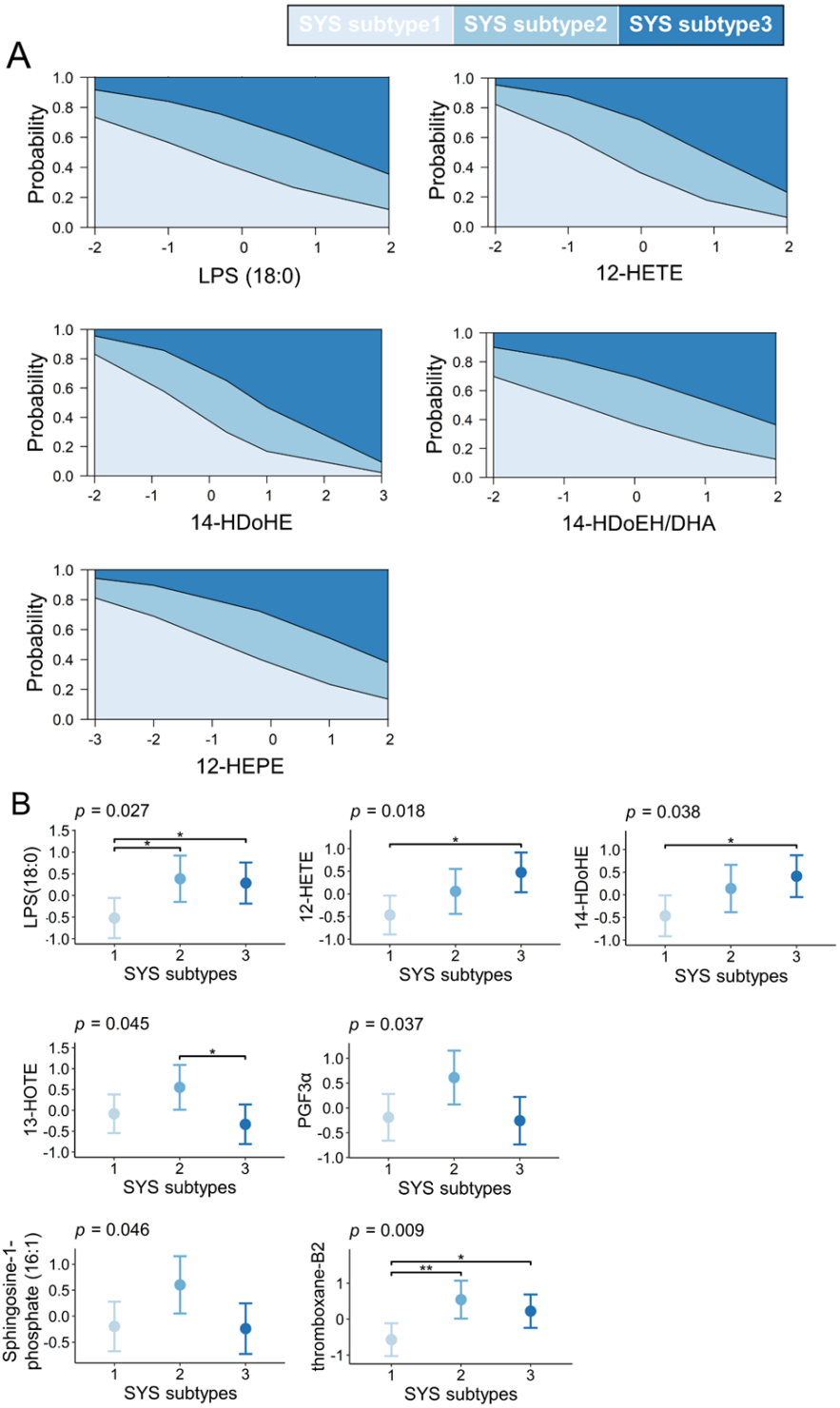


Figure 28 Signaling lipids associated with TCM algorithm subtypes. (A) Ordinal Regression Model revealed signaling lipids contribution to the TCM algorithm subtypes. Ordinal Regression Model is presented by the probability distribution of different TCM algorithm subtypes, indicating the likelihood of each patient being in different TCM algorithm subtypes based on metabolite changes. (B) Analysis of Covariance exhibited the variation of signaling lipids among the TCM algorithm subtypes. All the targets have been examined but only targets with a p-value less than 0.05 were displayed. Analysis of Covariance is presented by estimated marginal mean with the high and low end of confidence interval.

4.2.4 Correlations between metabolites, clinical variables and COPD subgroups

After the exploring the relevance between signaling lipids and different COPD subgroups, the correlations between the COPD clinical manifestation and parameters with both lipids and COPD subgroups were inspected (Figure 29). Regarding physical activity, a negative correlation with 1-OG & 2-OG is observed. KNGF profiles demonstrates moderate negative correlation with average steps, active minutes, and 6MWT. In addition, LPE (16:0), LPE (18:0), and lipoxin A5 exhibit a positive correlation with physical activity.

Furthermore, main COPD risk factors including sarcopenia, exacerbation, malnutrition and dyspnea were also examined. For sarcopenia, there is a positive correlation to cLPA (20:4), UDCA, KNGF profiles and TCM algorithm subtypes. Regarding exacerbation, a cluster of oxylipins is positively associated with lung attacks and antibiotic treatment, specifically HETEs derived from AA, HEPEs derived from EPA, and HDoHEs derived from DHA. Several lysophospholipids and bile acids show positive correlations with exacerbations, whereas, except for dyspnea, several bile acids and bile acid ratios negatively correlate with it. The last section of clinical indicators is Clinical COPD Questionnaires (CCQ), which is characterization of manifestations. There are only LPI (18:2) and 12-HEPE/EPA found to be weakly associated with the CCQ score emotion. The KNGF and TCM algorithm subtypes also demonstrate a close correlation with CCQ.

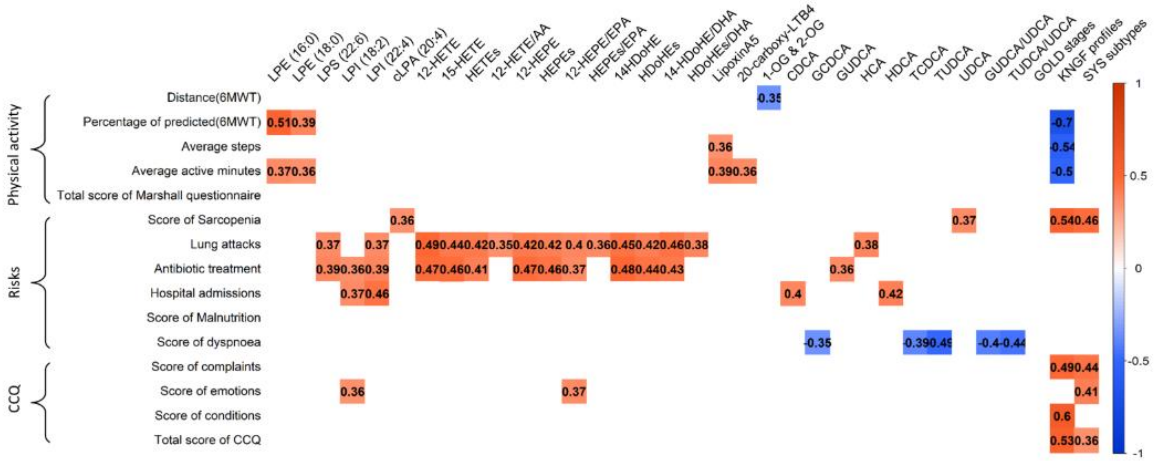


Figure 29 Spearman correlation analysis between COPD-related clinical indicators with signaling lipids and COPD subgroups. All the targets have been examined but only targets with a p-value less than 0.05 were displayed. The digit in the box indicates the correlation coefficients while the blank represents no significant correlation. Blue and red indicate positive and negative correlations respectively.

4.2.5 Interpretation of the metabolomics results

Our results of the metabolomics data analysis revealed that lysophospholipids, sphingolipids, oxylipins and bile acids are related to different COPD subgroups, as summarized in Figure 30.

ACV Analysis of Covariance
ORM Ordinal Regression Model

Metabolites correlated with clinical parameters

1. Lysophospholipids and sphingosines	GOLD stages	KNGF profiles	SYS subtypes
LPE (16:0)		ACV	
LPE (18:0)		ACV	
LPE (18:1)	ACV		
LPE (22:4)		ACV	
LPS (18:0)	ORM	ACV	ACV|ORM
LPS (20:4)	ORM	ACV	
LPS (22:6)	ORM		
LPA (14:0)		ACV|ORM	
LPA (16:0)		ACV	
Sphingosine (18:1)		ACV	
Sphinganine-1-phosphate (18:0)	ACV|ORM		
Sphingosine-1-phosphate (16:1)			ACV

2. Oxylipins	GOLD stages	KNGF profiles	SYS subtypes
12-HETE			ACV|ORM
12-HEPE			ORM
HEPEs/EPA		ACV	
14-HDoHE			ACV|ORM
14-HDoHE/DHA	ORM	ACV	ORM
16-HDoHE/DHA	ACV		
HDoHEs/DHA	ACV|ORM		
12,13-EpOME	ORM		
9,10,13-TriHOME	ORM		
13-HOTE			ACV
PGF3α			ACV
8,12-iso-iPF2a-VI		ACV|ORM	
thromboxane-B2			ACV

3. Bile acids	GOLD stages	KNGF profiles	SYS subtypes
DCA	ACV		
GCA/CA	ACV		
GUDCA		ORM	
TCA/CA	ACV		
TDCA/DCA	ACV		
TCDCA/CDCA	ACV		
TUDCA		ORM	

Figure 30 Summary of COPD subgroups related to signaling lipids. The targets identified by statistical analyses are listed per lipid category. The abbreviation indicates in which statistical method this target exhibited statistical significance. ACV for Analysis of Covariance, and ORM for Ordinal Regression Model. SYS subtypes indicate TCM algorithm based subgroups. The highlighting of targets presents relevance with clinical indicators.

Lysophospholipids are a subgroup of biomembrane components which could be classified as lysoglycerophospholipids and lysosphingolipids based on their chemical backbones. In addition, the lysophospholipids could be further characterized according to their polar head group: lysophosphatidylethanolamine (LPE), lysophosphatidylserine (LPS), lysophosphatidylinositol (LPI), and lysophosphatidic acid (LPA), etc. (Kano 2022, Takagi 2022). In the current research, it has been shown that 4 LPEs, 3 LPSs, and 2 LPAs were positively associated with COPD subgroups, especially with more severe GOLD stages and more severe KNGF profiles. For LPEs, previous research has shown that LPE (16:0), LPE (18:0) and LPE (18:1) could activate MAPK-ERK1/2 pathway (Hisano 2021, Hisano 2021) which would contribute to mucus overproduction, cytokine expression, apoptosis and fibrosis (Saleem 2024). Furthermore, LPE (16:0) was reported to be related to smoking condition (Rivas Serna 2021), and LPE (18:0) is positively correlated with chronic stress and cardiovascular disease risk (Balasubramanian 2023). Apart from LPEs, it has been found that LPS, LPI and LPA mediated receptors which are mainly G-protein coupled receptors (GPRs) (Yaginuma 2023). Importantly, the LPS receptors expression is almost confined to immune tissue and immune cells. The activation of LPS receptors would trigger multiple immune processes, such as activation of RhoA-ROCK signaling (Inoue 2019), production of pro-inflammatory cytokines (Preissler 2015), enhancement of mast cell degranulation (Barnes 2015), etc. Moreover, LPA receptors, LPA1 and LPA2 have been illuminated to participate in the respiratory system and inflammatory response (Yanagida 2023). Elevated LPA levels in the bronchial lavage fluid were observed in acute lung injury (ALI) model (Mouratis 2015), whereas knockout of the LPA1 gene exerted significant protection effects against inflammation (Chen 2017).

In addition, 3 sphingolipids were upregulated within subgroups with more severe COPD in GOLD stage, KNGF profile and TCM algorithm subtype respectively. The receptors of sphingosine-1-phosphate (S1PR) are GPRs, S1PR1 to S1PR5. An extensive study has disclosed the tight association between pulmonary disorders and the sphingosine-1-phosphate (S1P) signaling axis (Mohammed 2017, Chen 2022). S1P was discovered to be increased in asthmatic patients (Ammit 2001) and administration of S1P would enhance airway resistance, CCR3 and IL-17 (Roviezzo 2010). Conversely, these effects could be hampered by the inhibitor of S1PR2 or Rho kinase (Chiba 2010). Even though most research about S1P refers to S1P (18:1), another S1P less discussed, S1P (16:1) was proven to activate S1PR1, S1PR2 and S1PR3 (Wang 2022). Further validation of the function of S1P (16:1) and other sphingosines in COPD or pulmonary diseases is needed.

Oxylipins are oxidation products of polyunsaturated fatty acids (PUFAs), such as linoleic acid (LA), arachidonic acid (AA), eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA). This study showed that 12 oxylipins and 4 oxylipin/PUFA ratios were correlated with different COPD subgroups. Specifically, 12-HETE, 12HEPE, 14-HDoHE and 14-HDoHE/DHA which are produced by 12-lipoxygenase (LOX), are all positively associated with TCM algorithm subtypes of COPD. Wang and co-workers reported increased 12-HETE in serum of COPD patients compared to non-COPD patients (Wang 2022), moreover, 12-HETE and 12-HEPE were increased in serum after exposure to

ambient air particulate matter (Wang 2021). In addition, many studies verified the significance of 12-LOX in pulmonary disorders. 12/15-LOX-deficient mice displayed augmented IL-33-induced lung inflammation while application of 14-HDoHE suppressed airway inflammation in which 14-HDoHE serves as precursor of an anti-inflammatory regulator, maresin-1 (Miyita 2021). Combined with the correlation findings in which 12-HETE, 12-HEPE, 15-HETE and 14-HDoHE are positively correlated with lung attacks, all the findings imply an essential role of these metabolites and 12-LOX in COPD.

Apart from the oxidation products of 12-LOX, the other LOXs, cyclooxygenase (COX) and cytochrome P450 (CYPs) mediated metabolites also exerted functions in pulmonary dysfunction. In this study, 12,13-EpOME and 9,10,13-TriHOME, derived from LA, displayed a positive association with GOLD stages and KNGF profiles. 12,13-EpOME is produced by inflammatory leukocytes. High levels of EpOMEs were observed in acute respiratory distress syndrome (ARDS) patients (Thompson 2007) and a female-dominated phenotype of COPD. Upregulated 9,10,13-TriHOME was found in asthma patients following provocation compared with healthy control (Lundstrom 2007). In addition, the elevation of thromboxane B2 (TXB2) was found in TCM algorithm subtype 2 in this study. TXB2 is the inactive form of TXA2, yet it could be regarded as an indicator of TXA2 generation (Firestein 2013). TXA2 was found to increase in allergic asthma subjects' urine and the therapeutic effects were also confirmed by inhibiting the TXA2 receptor (Suzuki 2022).

The last cluster of lipids discovered from the statistical analyses is bile acids. Bile acids are derived from 7/27-hydroxycholesterol, the primary bile acids are formed in the liver including CA and CDCA, then dehydroxylated into secondary bile acids DCA and UDCA/LCA in the intestine by the activity of the microbiome. To increase the solubility, these bile acids are conjugated with either glycine (G) or taurine (T). Therefore, both the ratios of bile acid conjugates to corresponding prototypes and secondary to primary bile acids are investigated. This study showed an increase in the ratios of 2 CA conjugates (GCA/CA and TCA/CA), a CDCA conjugate ratio (TCDCA/CDCA), and DCA and its conjugate ratio (TDCA/DCA) in GOLD stage 2. Meanwhile, 2 UDCA conjugates (GUDCA and TUDCA) showed a positive relationship with KNGF profiles. Among these bile acids, CA, CDCA and DCA are activators for Farnesoid X receptor (FXR) while UDCA is an antagonist (Jiang 2013). The overexpression of FXR led to airway remodeling and inflammation in COPD via epithelial-mesenchymal transition (Chen 2016). Contradictorily, the anti-inflammation effects of FXR activation were observed in ALI/ARDS. The administration of FXR agonist inhibits the release of proinflammatory cytokines and NF-κB pathway (Fei 2019). Therefore, further studies claiming the role of FXR in pulmonary diseases are needed.

From all the discovered lipids, different lipid clusters were found to connect with GOLD stage, KNGF profile and TCM algorithm subtype in various ways. KNGF profiles were closely associated with lysophospholipids, while TCM algorithm subtypes showed more variance in oxylipins. In KNGF profile, physical capacity and activity are the main criteria to distinguish profiles 2, 3, 4, and 5. A recent study showed that an intensive running exercise altered lipid metabolism including

phospholipids and lysophospholipids (Sakaguchi 2019), leading to a significant decrease in LPE (Nieman 2013), suggesting that KNGF profiles may reflect energy metabolism. TCM algorithm subtypes emphasize systemic symptoms beyond the pulmonary system, with oxylipins playing essential roles in diseases such as cardiovascular diseases, oncological diseases, diabetes, obesity, liver disease, neurological disorders, kidney diseases, and so on (Parchem 2014). Consequently, TCM algorithm subtypes may focus on systemic conditions. As for GOLD stage, since the GOLD stage is purely based on FEV1/FVC, and the lung capacity is influenced by diverse factors, there isn't a clear link with specific lipid clusters.

4.2.6 **Conclusions interpretation of COPD subgroups with metabolomics**

The metabolomics results revealed that specific lipid clusters, i.e. lysophospholipids, oxylipins and bile acids were related to different COPD categories in GOLD stages, KNGF profiles, and TCM algorithm subtypes. In general, GOLD stage was associated with LPSs, HDoHEs, CA and its derivative DCA. The relevant related proteins would be the LPS receptors, LOX enzymes, and FXR. KNGF profile was related with LPAs, isoprostanes and UDCA, where the associated proteins are LPA1/2 and FXR. As for the TCM algorithm subtype, it was linked to S1P (18:1), 12-LOX catalyzed products and TXB2, therefore the relevant related proteins are S1PR1-3, 12-LOX, and TXA2 receptor.

These results show that the current GOLD stage and KNGF profile classification of COPD patients can be supported with distinct metabolite profiles. Additionally, similar associations between metabolites and a TCM classification of the COPD patients is found. This indicates that the TCM classification system can also be supported by biological differences between COPD patients classified into the various TCM subtypes. Furthermore, it is interesting to note that each classification system has an overlapping but also significantly different set of metabolites associated with it. This indicates that each classification method provides a unique view on COPD and can be used together to optimize treatment for each individual.

5 Conclusions

5.1 Towards integrative medicine for COPD

Our research indicates that the overall severity of TCM symptoms experienced by COPD patients is related to the KNGF stage. However, the TCM symptom distribution in PCA loading plots seem to suggest that these represent additional sources of variation between COPD patients. The main clusters of TCM symptoms relate to Stagnation related symptoms and to Deficiency related symptoms. Additionally, on a deeper level Yang Deficiency symptoms and Kidney Yin Deficiency symptoms can be distinguished in separate clusters. These clusters of symptoms with corresponding nutritional, exercise and mental health advise based on TCM theory can be used to optimize treatment op COPD patients.

Our research also for the first time provides evidence for a relationship between TCM subtypes of COPD and plasma metabolite profiles. In particular, the oxylipins 12-HETE, 12HEPE, 14-HDoHE and 14-HDoHE/DHA which are produced by 12-lipoxygenase (LOX), are all positively associated with TCM algorithm subtypes of COPD. These results show that a TCM diagnosis of COPD patients using symptoms is related to distinct metabolic profiles, providing evidence for the validity of the TCM diagnostic system. Additionally, the metabolomics analysis also support the KNGF and GOLD classification systems with related metabolite patterns.

There is currently ample evidence for an additional benefit of TCM treatment on top of conventional treatment for COPD patients (Chen 2023). At least four meta-analyses have been published including dozens of studies and thousands of patients, showing that there are clear clinical improvements to be expected from including TCM treatment on top of conventional treatment (Xiong 2021, Wen 2023, Hu 2021, Huang 2022, Lou 2024). These benefits range from improved lung function (FEV), lower frequency and duration of exacerbations, and improved quality of life. No adverse events have been reported (Xiong 2021). In addition, several randomized controlled trials have shown benefits of using a personalized approach by subtyping patients using TCM symptom patterns and offering these various patient groups specific TCM treatments on top of their conventional COPD treatment (Li 2012, Li 2013, Wang 2014). In the study by Wang et al. (2014) patients were subtyped into Lung-Kidney Qi deficiency, Lung-Spleen deficiency, and Lung-Kidney Qi and Yin deficiency. Three different herbal medicines were given to the subgroups on top of conventional COPD treatment. Significant differences with the conventional medicine only group after 6 months were lower frequency and duration of exacerbations, better 6 minute walk test, better scores on the dyspnoea scale, and better scores on quality of life (Wang 2014). Also for acupuncture there is evidence for additional benefits for COPD patients as a recent meta-analysis including 17 RCTs with a total of 1165 participants describes (Luo 2024).

Since there is ample evidence for the benefits of TCM based therapies for COPD patients, future studies should focus on the implementation of these therapies in the Netherlands. The first step could be to design a guideline including nutrition, exercise and mental health advise for the two main subgroups of patients identified, the patients with Stagnation symptoms and those with only Deficiency related symptoms. The TCM symptom questionnaire currently used for this study could be shortened to focus on these two subgroups. Based on the questionnaire results, specific advise could be provided to COPD patients. In a next step this approach could be implemented in a primary care practice or a COPD department of a hospital for testing. A pragmatic evaluation would then need to be conducted to evaluate the effects of this approach in practice using patient centered outcome measures. When results are positive, the approach could be further implemented in other centers. Furthermore, a next level of personalization could then be introduced by considering the clusters Yang Deficiency and Kidney Yin Deficiency symptoms and designing appropriate lifestyle advise for these groups of patients.

5.2 Summary of the conclusions

The current study shows that different medical paradigms can each contribute valuable knowledge and insights into disease processes and treatment options. In this case a Traditional Chinese Medicine perspective was introduced into the study by using a TCM symptom questionnaire and a TCM expert diagnosis. COPD patients were recruited and diagnosed using conventional diagnostic methods as well as TCM diagnostic tools. The study revealed that the variation between patients was not only related to COPD severity as established with the KNGF criteria, but also to specific clusters of TCM symptoms. The main clusters were Stagnation symptoms and Deficiency symptoms without Stagnation. This result means that the treatment of COPD could be more personalized when these differences in TCM symptoms between patients would be taken into account. Specific nutritional, exercise and mental health advise could be developed for these patient groups.

Furthermore, the current study offers a biological perspective on the two main classification systems for COPD (KNGF and GOLD) as well as several TCM symptom clusters. Specific groups of metabolites were found to be correlated to KNGF severity and to TCM symptom clusters. These results offer new insights in possible targets for interventions. Additionally, these results show that a diagnosis according to TCM theory can be supported by specific metabolic patterns corresponding to different TCM symptom clusters.

This study contributes a foundation for future implementation of personalized integrative medicine programs for COPD including specific lifestyle interventions tailored to TCM subgroups of COPD.

References

Agusti A, Calverley PM, Celli B, et al. Characterisation of COPD heterogeneity in the ECLIPSE cohort. Respir Res. 2010; 11: 122.

Agusti A, Sobradillo P and Celli B. Addressing the complexity of chronic obstructive pulmonary disease: from phenotypes and biomarkers to scale-free networks, systems biology, and P4 medicine. American journal of respiratory and critical care medicine. 2011; 183: 1129-37.

Agusti et al. Characterisation of COPD heterogeneity in the ECLIPSE cohort. Respiratory Research 2010, 11:122.

American Thoracic Society, Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: Guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166:111–117.

Ammit AJ, Hastie AT, Edsall LC, Hoffman RK, Amrani Y, Krymskaya VP, Kane SA, Peters SP, Penn RB, Spiegel S, Panettieri RA Jr. Sphingosine 1-phosphate modulates human airway smooth muscle cell functions that promote inflammation and airway remodeling in asthma. FASEB J. 2001 May;15(7):1212-4. doi: 10.1096/fj.00-0742fje. PMID: 11344091.

Balasubramanian R, Shutta KH, Guasch-Ferre M, Huang T, Jha SC, Zhu Y, Shadyab AH, Manson JE, Corella D, Fitó M, Hu FB, Rexrode KM, Clish CB, Hankinson SE, Kubzansky LD. Metabolomic profiles of chronic distress are associated with cardiovascular disease risk and inflammation-related risk factors. Brain Behav Immun. 2023 Nov;114:262-274. doi: 10.1016/j.bbi.2023.08.010. Epub 2023 Aug 8. PMID: 37557964.

Bang KM. Chronic obstructive pulmonary disease in nonsmokers by occupation and exposure: a brief review. Curr Opin Pulm Med. 2015; 21: 149-54.

Barabási, A.-L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. Nature Reviews. Genetics, 5(2), 101–113.

Barnes MJ, Li CM, Xu Y, An J, Huang Y, Cyster JG. The lysophosphatidylserine receptor GPR174 constrains regulatory T cell development and function. J Exp Med. 2015 Jun 29;212(7):1011-20. doi: 10.1084/jem.20141827. Epub 2015 Jun 15. PMID: 26077720; PMCID: PMC4493414.

Beasley R, Weatherall M, Travers J and Shirtcliffe P. Time to define the disorders of the syndrome of COPD. Lancet. 2009; 374: 670-2.

Beekman E, Mesters I, Spigt MG, et al. Phenotypic Variation in Patients with Chronic Obstructive Pulmonary Disease in Primary Care. Biomed Res Int. 2016; 2016: 8108717.

Bijlsma S, Bobeldijk I, Verheij ER, Ramaker R, MacDonald IA, et al. (2006) Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. Anal Chem 78(2): 567–74.

Blackstock FC. et al. Chronic Obstructive Pulmonary Disease Education in Pulmonary Rehabilitation. An Official American Thoracic Society/Thoracic Society of Australia and New Zealand/Canadian Thoracic Society/British Thoracic Society Workshop Report. Ann Am Thorac Soc. 2018 Jul;15(7):769-784. doi: 10.1513/AnnalsATS.201804-253WS.

Bolton, R. J., Hand, D. J., & Webb, A. R. (2003). principal component analysis. Statistics and Computing, 267–276.

Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. Eur Respir J Suppl 2003; 41:46s-53s.

Burgel PR, Paillasseur JL, Caillaud D, et al. Clinical COPD phenotypes: a novel approach using principal component and cluster analyses. The European respiratory journal. 2010; 36: 531-9.

Caramori G, Adcock IM, Papi A. Clinical definition of COPD exacerbations and classification of their severity. South Med J 2009;102(3):277-282.

Cazzola M. et al. A review of the most common patient-reported outcomes in COPD – revisiting current knowledge and estimating future challenges. Int J Chron Obstruct Pulmon Dis. 2015; 10: 725–738.

Chatila WM, Thomashow BM, Minai OA, Criner GJ, Make BJ. Comorbidities in chronic obstructive pulmonary disease. Proc Am Thorac Soc. 2008;5(4):549-55.

Chen B, You WJ, Xue S, Qin H, Zhao XJ, Zhang M, Liu XQ, Zhu SY, Jiang HD. Overexpression of farnesoid X receptor in small airways contributes to epithelial to mesenchymal transition and COX-2 expression in chronic obstructive pulmonary disease. J Thorac Dis. 2016 Nov;8(11):3063-3074. doi: 10.21037/jtd.2016.11.08. PMID: 28066584; PMCID: PMC5179405.

Chen X, Walther FJ, Laghmani EH, Hoogeboom AM, Hogen-Esch AC, van Ark I, Folkerts G, Wagenaar GT. Adult Lysophosphatidic Acid Receptor 1-Deficient Rats with Hyperoxia-Induced Neonatal Chronic Lung Disease Are Protected against Lipopolysaccharide-Induced Acute Lung Injury. Front Physiol. 2017 Mar 22;8:155. doi: 10.3389/fphys.2017.00155. PMID: 28382003; PMCID: PMC5360762.

Chen H, Wang J, Zhang C, Ding P, Tian S, Chen J, Ji G, Wu T. Sphingosine 1-phosphate receptor, a new therapeutic direction in different diseases. Biomed Pharmacother. 2022 Sep;153:113341. doi: 10.1016/j.biopha.2022.113341. Epub 2022 Jul 1. PMID: 35785704.

Chen R, Ni K, Ji C, Liu Z, Yu Y, Liu G, Yang J, Wang Z. Effects of co-application of tiotropium bromide and traditional Chinese medicine on patients with stable chronic obstructive pulmonary disease: a muilticenter, randomized, controlled trial study. Front Med (Lausanne). 2024 May 3;11:1289928. doi: 10.3389/fmed.2024.1289928. PMID: 38765259; PMCID: PMC11099264.

Chiba Y, Suzuki K, Uechi M, Kurihara E, Goto K, Sakai H, Misawa M. Downregulation of sphingosine-1-phosphate receptors in bronchial smooth muscle of mouse experimental asthma. Pharmacol Res. 2010 Oct;62(4):357-63. doi: 10.1016/j.phrs.2010.05.005. Epub 2010 Jun 8. PMID: 20554039.

Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, et al. (2007) Integration of biological networks and gene expression data using Cytoscape. Nat Protoc 2(10): 2366–82.

Decramer M, Rennard S, Troosters T, Mapel DW, Giardino N, Mannino D, Wouters E, Sethi S, Cooper CB. COPD as a lung disease with systemic consequences--clinical impact, mechanisms, and potential for early intervention. COPD. 2008 Aug;5(4):235-56. doi: 10.1080/15412550802237531. PMID: 18671149.

Elia, M. & Chairman of MAG and Editor (2003). The 'MUST' report. Nutritional screening of adults: a multidisciplinary responsibility. Development and use of the 'Malnutrition Universal Screening Tool' ('MUST') for adults. Redditch: BAPEN

Fei J, Fu L, Hu B, Chen YH, Zhao H, Xu DX, Li JB. Obeticholic acid alleviate lipopolysaccharide-induced acute lung injury via its anti-inflammatory effects in mice. Int Immunopharmacol. 2019 Jan;66:177-184. doi: 10.1016/j.intimp.2018.11.005. Epub 2018 Nov 20. PMID: 30468885.

Ferrer M, Alonso J, Morera J, Marrades RM, Khalaf A, Aguar MC, et al. Chronic obstructive pulmonary disease stage and health-related quality of life. The Quality of Life of Chronic Obstructive Pulmonary Disease Study Group. Ann Intern Med. 1997;127(12):1072-9.

Fingleton J, Weatherall M and Beasley R. Towards individualised treatment in COPD. Thorax. 2011; 66: 363-4.

Firestein GS, Budd RC, Gabriel, Sherine E., et al., Kelley's Textbook of Rheumatology, 2013. https://www.sciencedirect.com/book/9781437717389/kelleys-textbook-of-rheumatology.

Fox J, Hong J. Effect Displays in R for Multinomial and Proportional-Odds Logit Models: Extensions to the effects Package. J. Stat. Soft. [Internet]. 2009 Oct. 14 [cited 2024 Aug. 27];32(1):1-24. Available from: https://www.jstatsoft.org/index.php/jss/article/view/v032i01

Garcia-Aymerich J, Lange P, Benet M, Schnohr P and Anto JM. Regular physical activity reduces hospital admission and mortality in chronic obstructive pulmonary disease: a population based cohort study. Thorax. 2006; 61: 772-8.

Gehlenborg, N., O'Donoghue, S. I., Baliga, N. S., Goesmann, A., Hibbs, M. A., Kitano, H., … Gavin, A.-C. (2010). Visualization of omics data for systems biology. Nature Methods, 7(3 Suppl), S56-68.

Gibson GJ, Loddenkemper R, Lundback B, Sibille Y. Respiratory health and disease in Europe: the new European Lung White Book. Eur Respir J 2013; 42(3): 559-563.

Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis,

Han MK, Agusti A, Calverley PM, et al. Chronic obstructive pulmonary disease phenotypes: the future of COPD. American journal of respiratory and critical care medicine. 2010; 182: 598-604.

Han, M. K., Agusti, A., Calverley, P. M., Celli, B. R., Criner, G., Curtis, J. L., … Martinez, F. J. (2010). Chronic Obstructive Pulmonary Disease Phenotypes. American Journal of Respiratory Critical Care Medicine, 182, 598–604.

Hendriks MMWB, Smit S, Akkermans WLMW, Reijmers TH, Eilers PHC, et al. (2007) How to distinguish healthy from diseased? Classification strategy for mass spectrometry-based clinical proteomics. Proteomics 7: 3672–3680.

Hillas G, Perlikos F, Tsiligianni I, Tzanakis N. Managing comorbidities in COPD. Int J Chron Obstruct Pulmon Dis. 2015;10:95-109.

A. Inoue, F. Raimondi, F.M.N. Kadji, et al., Illuminating G-Protein-Coupling Selectivity of GPCRs, Cell 177(7) (2019) 1933-1947 e25. 10.1016/j.cell.2019.04.044

Hisano K, Kawase S, Mimura T, Yoshida H, Yamada H, Haniu H, Tsukahara T, Kurihara T, Matsuda Y, Saito N, Uemura T. Structurally different lysophosphatidylethanolamine species stimulate neurite outgrowth in cultured cortical neurons via distinct G-protein-coupled receptors and signaling cascades. Biochem Biophys Res Commun. 2021 Jan 1;534:179-185. doi: 10.1016/j.bbrc.2020.11.119. Epub 2020 Dec 6. PMID: 33298313.

Hisano K, Yoshida H, Kawase S, Mimura T, Haniu H, Tsukahara T, Kurihara T, Matsuda Y, Saito N, Uemura T. Abundant oleoyl-lysophosphatidylethanolamine in brain stimulates neurite outgrowth and protects against glutamate toxicity in cultured cortical neurons. J Biochem. 2021 Oct 12;170(3):327-336. doi: 10.1093/jb/mvab046. PMID: 33822960.

Hu H, Ji Z, Qiang X, Liu S, Sheng X, Chen Z, Liu F, Wang H, Zhang J. Chinese Medical Injections for Acute Exacerbation of Chronic Obstructive Pulmonary Disease: A Network Meta-analysis. Int J Chron Obstruct Pulmon Dis. 2021 Dec 17;16:3363-3386. doi: 10.2147/COPD.S335579. PMID: 34949918; PMCID: PMC8691136.

Huang P, Lin X, Liu Y, Hou Z. The efficacy and safety of combined traditional Chinese and western medicine in the treatment of chronic obstructive pulmonary disease complicated with respiratory failure: a systematic review and meta-analysis study. Ann Palliat Med. 2022 Mar;11(3):1102-1111. doi: 10.21037/apm-22-272. PMID: 35365040.

Kano K, Aoki J, Hla T. Lysophospholipid Mediators in Health and Disease. Annu Rev Pathol. 2022 Jan 24;17:459-483. doi: 10.1146/annurev-pathol-050420-025929. Epub 2021 Nov 23. PMID: 34813354; PMCID: PMC9641500.

Kitano, H. (2002). Systems biology: a brief overview. Science (New York, N.Y.), 295(5560), 1662–1664.

de Leeuw FA, Peeters CFW, Kester MI, Harms AC, Struys EA, Hankemeier T, van Vlijmen HWT, van der Lee SJ, van Duijn CM, Scheltens P, Demirkan A, van de Wiel MA, van der Flier WM, Teunissen CE. Blood-based metabolic signatures in Alzheimer's disease. Alzheimers Dement (Amst). 2017 Sep 6;8:196-207. doi: 10.1016/j.dadm.2017.07.006. PMID: 28951883; PMCID: PMC5607205.

Li F, Jiang C, Krausz KW, Li Y, Albert I, Hao H, Fabre KM, Mitchell JB, Patterson AD, Gonzalez FJ. Microbiome remodelling leads to inhibition of intestinal farnesoid X receptor signalling and decreased obesity. Nat Commun. 2013;4:2384. doi: 10.1038/ncomms3384. PMID: 24064762; PMCID: PMC6595219.

Li SY, Li JS, Wang MH, Xie Y, Yu XQ, Sun ZK, Ma LJ, Zhang W, Zhang HL, Cao F, Pan YC. Effects of comprehensive therapy based on traditional Chinese medicine patterns in stable chronic obstructive pulmonary disease: a four-center, open-label, randomized, controlled study. BMC Complement Altern Med. 2012 Oct 29;12:197. doi: 10.1186/1472-6882-12-197. PMID: 23107470; PMCID: PMC3528455.

Li JS, Li SY, Xie Y, Yu XQ, Wang MH, Sun ZK, Ma LJ, Jia XH, Zhang HL, Xu JP, Hou CX. The effective evaluation on symptoms and quality of life of chronic obstructive pulmonary disease patients treated by comprehensive therapy based on traditional Chinese medicine patterns. Complement Ther Med. 2013 Dec;21(6):595-602. doi: 10.1016/j.ctim.2013.09.006. Epub 2013 Sep 23. PMID: 24280466.

Linting, M., & van der Kooij, A. (2012). Nonlinear Principal Components Analysis With CATPCA: A Tutorial. Journal of Personality Assessment, 94(1), 12–25.

Longfonds. Zelf regie nemen. [The Lung Foundation Netherlands]. https://www.longfonds.nl/wat-doet-het-longfonds/zelf-regie-nemen (assessed 19-4-2019)

Luo Q, Sun M, Xu G, Tian H, Yang C, Huang L, Li X, Wang Z, Lu G, Yang Z, Ji L, Liang F. Exploration of quantitative-effectiveness association between acupuncture temporal parameters and stable chronic obstructive pulmonary disease: A systematic review and dose-response meta-analysis of randomized controlled trials. Complement Ther Med. 2024 Jun;82:103048. doi: 10.1016/j.ctim.2024.103048. Epub 2024 May 10. PMID: 38734186.

Lundström SL, Yang J, Källberg HJ, Thunberg S, Gafvelin G, Haeggström JZ, Grönneberg R, Grunewald J, van Hage M, Hammock BD, Eklund A, Wheelock ÅM, Wheelock CE. Allergic asthmatics show divergent lipid mediator profiles from healthy controls both at baseline and following birch pollen provocation. PLoS One. 2012;7(3):e33780. doi: 10.1371/journal.pone.0033780. Epub 2012 Mar 15. PMID: 22438998; PMCID: PMC3305349.

Maciocia, G. (2004). Diagnosis in Chinese medicine - a comprehensive guide (1st ed.). Elsevier.

Mahler DA, Wells CK: Evaluation of clinical methods for rating dyspnea. Chest. 1988, 93: 580-586.

Management, and Prevention of Chronic Obstructive Pulmonary Disease. 2019. https://goldcopd.org/gold-reports/ (accessed 18-4-2019).

Mannino DM and Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. Lancet. 2007; 370: 765-73.

Marshall AL, Smith B, Bauman AE, Kaur S. Reliability and validity of a brief physical activity assessment for use by family doctors. Br J Sports Med 2005;39:294-7.

Miyata J, Yokokura Y, Moro K, Arai H, Fukunaga K, Arita M. 12/15-Lipoxygenase Regulates IL-33-Induced Eosinophilic Airway Inflammation in Mice. Front Immunol. 2021 May 19;12:687192. doi: 10.3389/fimmu.2021.687192. PMID: 34093589; PMCID: PMC8170304.

Mohammed S, Harikumar KB. Sphingosine 1-Phosphate: A Novel Target for Lung Disorders. Front Immunol. 2017 Mar 14;8:296. doi: 10.3389/fimmu.2017.00296. Erratum in: Front Immunol. 2018 Jul 17;9:1628. doi: 10.3389/fimmu.2018.01628. PMID: 28352271; PMCID: PMC5348531.

Mouratis MA, Magkrioti C, Oikonomou N, Katsifa A, Prestwich GD, Kaffe E, Aidinis V. Autotaxin and Endotoxin-Induced Acute Lung Injury. PLoS One. 2015 Jul 21;10(7):e0133619. doi: 10.1371/journal.pone.0133619. PMID: 26196781; PMCID: PMC4509763.

Nieman DC, Shanely RA, Gillitt ND, Pappan KL, Lila MA. Serum metabolic signatures induced by a three-day intensified exercise period persist after 14 h of recovery in runners. J Proteome Res. 2013 Oct 4;12(10):4577-84. doi: 10.1021/pr400717j. Epub 2013 Sep 13. PMID: 23984841.

Nishisato S (1984) Forced classification: A simple application of a quantification method. Psychometrika 49: 25–36.

O'Donnell DE, Elbehairy AF, Faisal A, Webb KA, Neder JA, Mahler DA. Exertional dyspnoea in COPD: the clinical utility of cardiopulmonary exercise testing. Eur Respir Rev. 2016;25(141):333-47.

Parchem K, Letsiou S, Petan T, Oskolkova O, Medina I, Kuda O, O'Donnell VB, Nicolaou A, Fedorova M, Bochkov V, Gladine C. Oxylipin profiling for clinical research: Current status and future perspectives. Prog Lipid Res. 2024 Jul;95:101276. doi: 10.1016/j.plipres.2024.101276. Epub 2024 Apr 30. PMID: 38697517.

Perez De Souza L, Alseekh S, Brotman Y, Fernie AR. Network-based strategies in metabolomics data analysis and interpretation: from molecular networking to biological interpretation. Expert Rev Proteomics. 2020 Apr;17(4):243-255. doi: 10.1080/14789450.2020.1766975. Epub 2020 Jun 4. PMID: 32380880.

Pinto LM, Alghamdi M, Benedetti A, Zaihra T, Landry T and Bourbeau J. Derivation and validation of clinical phenotypes for COPD: a systematic review. Respir Res. 2015; 16: 50.

Pinto-Plata VM, Cote C, Cabral H, Taylor J, Celli BR. The 6-min walk distance: change over time and value as a predictor of survival in severe COPD. Eur Respir J. 2004;23(1):28-33.

Preissler J, Grosche A, Lede V, Le Duc D, Krügel K, Matyash V, Szulzewsky F, Kallendrusch S, Immig K, Kettenmann H, Bechmann I, Schöneberg T, Schulz A. Altered microglial phagocytosis in GPR34-deficient mice. Glia. 2015 Feb;63(2):206-15. doi: 10.1002/glia.22744. Epub 2014 Aug 20. PMID: 25142016.

Rivas Serna IM, Sitina M, Stokin GB, Medina-Inojosa JR, Lopez-Jimenez F, Gonzalez-Rivas JP, Vinciguerra M. Lipidomic Profiling Identifies Signatures of Poor Cardiovascular Health. Metabolites. 2021 Oct 29;11(11):747. doi: 10.3390/metabo11110747. PMID: 34822405; PMCID: PMC8624456.

Rostron BL, Chang CM and Pechacek TF. Estimation of cigarette smoking-attributable morbidity in the United States. JAMA Intern Med. 2014; 174: 1922-8.

Roviezzo F, D'Agostino B, Brancaleone V, De Gruttola L, Bucci M, De Dominicis G, Orlotti D, D'Aiuto E, De Palma R, Rossi F, Sorrentino R, Cirino G. Systemic administration of sphingosine-1-phosphate increases bronchial hyperresponsiveness in the mouse. Am J Respir Cell Mol Biol. 2010 May;42(5):572-7. doi: 10.1165/rcmb.2009-0108OC. Epub 2009 Jun 25. PMID: 19556602.

Sakaguchi CA, Nieman DC, Signini EF, Abreu RM, Catai AM. Metabolomics-Based Studies Assessing Exercise-Induced Alterations of the Human Metabolome: A Systematic Review. Metabolites. 2019 Aug 9;9(8):164. doi: 10.3390/metabo9080164. PMID: 31405020; PMCID: PMC6724094.

Saleem S. Targeting MAPK signaling: A promising approach for treating inflammatory lung disease. Pathol Res Pract. 2024 Feb;254:155122. doi: 10.1016/j.prp.2024.155122. Epub 2024 Jan 11. PMID: 38246034.

Schellevis FG. Multimorbidity in general practice: you don't notice until you figured it out. Huisarts en Wetenschap 2007;9:452-454.

Schols AM, Gosker HR. The pathophysiology of cachexia in chronic obstructive pulmonary disease.

Schroen, Y., Wietmarschen, H. A. van, Wang, M., Wijk, E. P. van, Hankemeier, T., Xu, G., & Greef, J. van der. (2014). East is East and West is West, and never the twain shall meet? Science, 346(6216), S10–S12.

Searle SR, Speed FM, & Milliken GA. Population Marginal Means in the Linear Model: An Alternative to Least Squares Means. The American Statistician. 1980, 34(4), 216–221. https://doi.org/10.1080/00031305.1980.10483031

Smid DE. et al. Redefining Cut-Points for High Symptom Burden of the Global Initiative for Chronic Obstructive Lung Disease Classification in 18,577 Patients With Chronic Obstructive Pulmonary Disease. Smid DEJ Am Med Dir Assoc. 2017 Dec 1;18(12):1097.e11-1097.e24. doi: 10.1016/j.jamda.2017.09.003.

Snoeck-Stroband JB, Schermer TRJ, Van Schayck CP, Muris JW, Van der Molen T, In 't Veen JCCM, Chavannes NH, Broekhuizen BDL, Barnhoorn MJM, Smeele I, Geijer RMM, Tuut MK. NHG-Werkgroep Astma bij volwassenen en COPD. NHG-Standaard COPD (derde herziening). Huisarts Wet 2015;58(4):198-211. [The Dutch College of General Practitioners (NHG) guidelines 'COPD']. https://www.nhg.org/dutch-college-general-practitioners and https://www.nhg.org/standaarden/volledig/nhg-standaard-copd (assessed 19-4-2019).

Spruit MA, Singh SJ, Garvey C, ZuWallack R, Nici L, Rochester C, et al.; ATS/ERS Task Force on Pulmonary Rehabilitation. An official American Thoracic Society/European Respiratory Society statement: key concepts and advances in pulmonary rehabilitation. Am J Respir Crit Care Med 2013;188:e13–e64.

Spruit MA. et al. Determinants of poor 6-min walking distance in patients with COPD: The ECLIPSE cohort. Respiratory Medicine (2010) 104, 849e857.

Suzuki T, Kropski JA, Chen J, Carrier EJ, Chen X, Sherrill TP, Winters NI, Camarata JE, Polosukhin VV, Han W, Rathinasabapathy A, Gutor S, Gulleman P, Sabusap C, Banovich NE, Tanjore H, Freeman ML, Tada Y, Young LR, Gokey JJ, Blackwell TS, West JD. Thromboxane-Prostanoid Receptor Signaling Drives Persistent Fibroblast Activation in Pulmonary Fibrosis. Am J Respir Crit Care Med. 2022 Sep 1;206(5):596-607. doi: 10.1164/rccm.202106-1503OC. PMID: 35728047; PMCID: PMC9716913.

Takagi Y, Nishikado S, Omi J, Aoki J. The Many Roles of Lysophospholipid Mediators and Japanese Contributions to This Field. Biol Pharm Bull. 2022;45(8):1008-1021. doi: 10.1248/bpb.b22-00304. PMID: 35908884.

Thompson DA, Hammock BD. Dihydroxyoctadecamonoenoate esters inhibit the neutrophil respiratory burst. J Biosci. 2007 Mar;32(2):279-91. doi: 10.1007/s12038-007-0028-x. PMID: 17435320; PMCID: PMC1904342.

van der Greef J, Martin S, Juhasz P, Adourian A, Plasterer T, Verheij ER, McBurney RN. The art and practice of systems biology in medicine: mapping patterns of relationships. J Proteome Res. 2007 Apr;6(4):1540-59. doi: 10.1021/pr0606530. Epub 2007 Mar 21. PMID: 17373844.

van der Greef J, van Wietmarschen H, Schroën J, Wang M, Hankemeier T, Xu G. Systems biology-based diagnostic principles as pillars of the bridge between Chinese and Western medicine. Planta Med. 2010 Dec;76(17):2036-47. doi: 10.1055/s-0030-1250450. Epub 2010 Oct 26. PMID: 20979015.

van der Molen T, Willemse BW, Schokker S, ten Hacken NH, Postma DS, Juniper EF. Development, validity and responsiveness of the Clinical COPD Questionnaire. Health Qual Life Outcomes. 2003;1:13.

van Manen JG, Bindels PJ, CJ IJ, van der Zee JS, Bottema BJ, Schade E. Prevalence of comorbidity in patients with a chronic airway obstruction and controls over the age of 40. J Clin Epidemiol. 2001;54(3):287-93.

van Wietmarschen HA, Dai W, van der Kooij AJ, Reijmers TH, Schroën Y, Wang M, Xu Z, Wang X, Kong H, Xu G, Hankemeier T, Meulman JJ, van der Greef J. Characterization of rheumatoid arthritis subtypes using symptom profiles, clinical chemistry and metabolomics measurements. PLoS One. 2012;7(9):e44331. doi: 10.1371/journal.pone.0044331. Epub 2012 Sep 12. PMID: 22984493; PMCID: PMC3440441.

van Wietmarschen HA, Reijmers TH, van der Kooij AJ, Schroën J, Wei H, Hankemeier T, Meulman JJ, van der Greef J. Sub-typing of rheumatic diseases based on a systems diagnosis questionnaire. PLoS One. 2011;6(9):e24846. doi: 10.1371/journal.pone.0024846. Epub 2011 Sep 16. PMID: 21949766; PMCID: PMC3174973.

van Wietmarschen H, Yuan K, Lu C, Gao P, Wang J, Xiao C, Yan X, Wang M, Schroën J, Lu A, Xu G, van der Greef J. Systems biology guided by Chinese medicine reveals new markers for sub-typing rheumatoid arthritis patients. J Clin Rheumatol. 2009 Oct;15(7):330-7. doi: 10.1097/RHU.0b013e3181ba3926. PMID: 20009967.

Vanfleteren LE, Spruit MA, Groenen M, Gaffron S, van Empel VP, Bruijnzeel PL, et al. Clusters of comorbidities based on validated objective measurements and systemic inflammation in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2013;187(7):728-35.

Vreeken H. Presentation Development of the new KNGF-guidelines on COPD. Week of the lungs [Week van de longen], 9-4-2019 at Ermelo. https://www.weekvandelongen.nl/nl/home (assessed 8-5-2019).

Wang M, Li J, Li S, Xie Y. Effects of comprehensive therapy based on traditional Chinese medicine patterns on older patients with chronic obstructive pulmonary disease: a subgroup analysis from a four-center, randomized, controlled study. Front Med. 2014 Sep;8(3):368-75. doi: 10.1007/s11684-014-0360-0. Epub 2014 Sep 9. PMID: 25204290.

Wang W, Xiang P, Chew WS, Torta F, Bandla A, Lopez V, Seow WL, Lam BWS, Chang JK, Wong P, Chayaburakul K, Ong WY, Wenk MR, Sundar R, Herr DR. Activation of sphingosine 1-phosphate receptor 2 attenuates chemotherapy-induced neuropathy. J Biol Chem. 2020 Jan 24;295(4):1143-1152. doi: 10.1074/jbc.RA119.011699. Epub 2019 Dec 27. PMID: 31882542; PMCID: PMC6983853.

Wang T, Chen X, Li H, Chen W, Xu Y, Yao Y, Zhang H, Han Y, Zhang L, Que C, Gong J, Qiu X, Zhu T. Pro-thrombotic changes associated with exposure to ambient ultrafine particles in patients with chronic obstructive pulmonary disease: roles of lipid peroxidation and systemic inflammation. Part Fibre Toxicol. 2022 Oct 24;19(1):65. doi: 10.1186/s12989-022-00503-9. PMID: 36280873; PMCID: PMC9590143.

Wang T, Han Y, Li H, Wang Y, Xue T, Chen X, Chen W, Fan Y, Qiu X, Gong J, Xu Y, Wang J, Li W, Zhu T. Changes in bioactive lipid mediators in response to short-term exposure to ambient air particulate matter: A targeted lipidomic analysis of oxylipin signaling pathways. Environ Int. 2021 Feb;147:106314. doi: 10.1016/j.envint.2020.106314. Epub 2020 Dec 14. PMID: 33326904.

Wei H, Pasman W, Rubingh C, Wopereis S, Tienstra M, Schroen J, Wang M, Verheij E, van der Greef J. Urine metabolomics combined with the personalized diagnosis guided by Chinese medicine reveals subtypes of pre-diabetes. Mol Biosyst. 2012 Apr;8(5):1482-91. doi: 10.1039/c2mb05445k. Epub 2012 Mar 13. PMID: 22414982. Westerhuis, J. J. a., Hoefsloot, H. H. C. J. H., Smit, S., Vis, D. D. J., Smilde, A. A. K., van Velzen, E. J. J., … Dorsten, F. a. (2008). Assessment of PLSDA cross validation. METABOLOMICS, 4(1), 81–89.

Wen G, Xuanlin LI, Hulei Z, Siyuan L, Yang X, Jiansheng LI. Effectiveness of Chinese herbal medicine combined with conventional medicine on acute exacerbation of chronic obstructive pulmonary disease: a systematic review and Meta-analysis. J Tradit Chin Med. 2023 Apr;43(2):212-220. doi: 10.19852/j.cnki.jtcm.2023.02.003. PMID: 36994509; PMCID: PMC10012211.

Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, Fung C, Nikolai L, Lewis M, Coutouly MA, Forsythe I, Tang P, Shrivastava S, Jeroncic K, Stothard P, Amegbey G, Block D, Hau DD, Wagner J, Miniaci J, Clements M, Gebremedhin M, Guo N, Zhang Y, Duggan GE, Macinnis GD, Weljie AM, Dowlatabadi R, Bamforth F, Clive D, Greiner R, Li L, Marrie T, Sykes BD, Vogel HJ, Querengesser L. HMDB: the Human Metabolome Database. Nucleic Acids Res. 2007 Jan;35(Database issue):D521-6. doi: 10.1093/nar/gkl923. PMID: 17202168; PMCID: PMC1899095.

World Health Organisation (WHO). Chronic Obstructive Pulmonary Disease (COPD). 2017. https://www.who.int/en/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd) (accessed 18-4-2019).

World Health Organisation (WHO). International Classification of Functioning, Disability and Health (ICF). 2001. https://www.who.int/classifications/icf/en/ (accessed 18-4-2019).

Woo J, Leung J, Morley JE. Validating the SARC-F: a suitable community screening tool for sarcopenia? J Am Med Dir Assoc. 2014 Sep;15(9):630-4. doi: 10.1016/j.jamda.2014.04.021. Epub 2014 Jun 16. PMID: 24947762.

Xiong C, Li Y, Zhuang G, Zeng Y, Wei H, Li C, Li L, Liao L, Fu J. Clinical efficacy and safety of Chinese herbal medicine versus placebo for the treatment of chronic obstructive pulmonary disease: A systematic review and meta-analysis. Complement Ther Med. 2021 Jun;59:102691. doi: 10.1016/j.ctim.2021.102691. Epub 2021 Feb 20. PMID: 33618010.

Yaginuma S, Omi J, Uwamizu A, Aoki J. Emerging roles of lysophosphatidylserine as an immune modulator. Immunol Rev. 2023 Aug;317(1):20-29. doi: 10.1111/imr.13204. Epub 2023 Apr 10. PMID: 37036835.

Yanagida K, Shimizu T. Lysophosphatidic acid, a simple phospholipid with myriad functions. Pharmacol Ther. 2023 Jun;246:108421. doi: 10.1016/j.pharmthera.2023.108421. Epub 2023 Apr 18. PMID: 37080433.

Yang W, Schoeman JC, Di X, Lamont L, Harms AC, Hankemeier T. A comprehensive UHPLC-MS/MS method for metabolomics profiling of signaling lipids: Markers of oxidative stress, immunity and inflammation. Anal Chim Acta. 2024 Apr 8;1297:342348. doi: 10.1016/j.aca.2024.342348. Epub 2024 Feb 10. PMID: 38438234.

Zhou A, Zhou Z, Zhao Y and Chen P. The recent advances of phenotypes in acute exacerbations of COPD. International journal of chronic obstructive pulmonary disease. 2017; 12: 1009-18.

Zitnik M, Li MM, Wells A, Glass K, Morselli Gysi D, Krishnan A, Murali TM, Radivojac P, Roy S, Baudot A, Bozdag S, Chen DZ, Cowen L, Devkota K, Gitter A, Gosline SJC, Gu P, Guzzi PH, Huang H, Jiang M, Kesimoglu ZN, Koyuturk M, Ma J, Pico AR, Pržulj N, Przytycka TM, Raphael BJ, Ritz A, Sharan R, Shen Y, Singh M, Slonim DK, Tong H, Yang XH, Yoon BJ, Yu H, Milenković T. Current and future directions in network biology. Bioinform Adv. 2024 Aug 14;4(1):vbae099. doi: 10.1093/bioadv/vbae099. PMID: 39143982; PMCID: PMC11321866.

Appendices

Appendix 1. Target list of detectable metabolites

Metabolites	Common_Name*	HMDB_ID	LIPID MAPS_ID
cLPA C14:0	Cyclic-Lysophosphatidic acid (14:0)		LMGP00000070
cLPA C16:1	Cyclic-Lysophosphatidic acid (16:1)		
cLPA C18:0	Cyclic-Lysophosphatidic acid (18:0)	HMDB0007004	LMGP00000055
cLPA C18:1	Cyclic-Lysophosphatidic acid (18:1)	HMDB0007006	LMGP00000056
cLPA C18:2	Cyclic-Lysophosphatidic acid (18:2)	HMDB0007007	
cLPA C20:4	Cyclic-Lysophosphatidic acid (20:4)		
LPA C14:0	Lysophosphatidic acid (14:0)	HMDB0062321	LMGP10050007
LPA C16:0	Lysophosphatidic acid (16:0)	HMDB07853/HMDB07849	LMGP10050006/LMGP10050042
LPA C16:1	Lysophosphatidic acid (16:1)	HMDB0062323	LMGP10050016
LPA C18:1	Lysophosphatidic acid (18:1)	HMDB07855/HMDB07851	LMGP10050008/LMGP10050014
LPA C18:2	Lysophosphatidic acid (18:2)	HMDB0007856/HMDB07852	LMGP10050017/LMGP10050044
LPA C20:3	Lysophosphatidic acid (20:3)	HMDB0062313	LMGP10050028
LPA C20:4	Lysophosphatidic acid (20:4)	HMDB0062312	LMGP10050013
LPA C20:5	Lysophosphatidic acid (20:5)	HMDB0062308	LMGP10050033
LPA C22:4	Lysophosphatidic acid (22:4)	HMDB0062310	LMGP10050020
LPA C22:5	Lysophosphatidic acid (22:5)	HMDB0114753/HMDB0114754	
LPA C22:6	Lysophosphatidic acid (22:6)	HMDB0114755	LMGP10050019
LPE C14:0	Lysophosphatidylethanolamine (14:0)	HMDB0011500/HMDB0011470	LMGP02050003/LMGP02050033
LPE C16:0	Lysophosphatidylethanolamine (16:0)	HMDB0011503/HMDB0011473	LMGP02050002/LMGP02050036
LPE C16:1	Lysophosphatidylethanolamine (16:1)	HMDB0011504 HMDB0011474	LMGP02050010 LMGP02050037
LPE C18:0	Lysophosphatidylethanolamine (18:0)	HMDB0011130/HMDB0011129	LMGP02050001/LMGP02050038
LPE C18:1	Lysophosphatidylethanolamine (18:1)	HMDB0011475/HMDB0011476 HMDB0011505/HMDB0011506	LMGP02050039/LMGP02050040 LMGP02050064/LMGP02050004
LPE C18:2	Lysophosphatidylethanolamine (18:2)	HMDB0011507/HMDB0011477	LMGP02050011/LMGP02050041
LPE C18:3	Lysophosphatidylethanolamine (18:3)	HMDB0011478/HMDB0011479 HMDB0011508/HMDB0011509	LMGP02050042/LMGP02050043 LMGP02050017/ LMGP02050029
LPE C20:3	Lysophosphatidylethanolamine (20:3)	HMDB0011516/HMDB0011486 HMDB0011514/HMDB0011484 HMDB0011515/HMDB0011485	LMGP02050022/LMGP02050050 LMGP02050065/LMGP02050048 LMGP02050066/LMGP02050049
LPE C20:4	Lysophosphatidylethanolamine (20:4)	HMDB0011517/HMDB0011487 HMDB0011518/HMDB0011488	LMGP02050009/LMGP02050051 LMGP02050067/LMGP02050052
LPE C20:5	Lysophosphatidylethanolamine (20:5)	HMDB0011519/HMDB0011489	LMGP02050027/LMGP02050053
LPE C22:4	Lysophosphatidylethanolamine (22:4)	HMDB0011523/HMDB0011493	LMGP02050014/LMGP02050057

Metabolites	Common_Name*	HMDB_ID	LIPID MAPS_ID
LPE C22:5	Lysophosphatidylethanolamine (22:5)	HMDB0011494/HMDB0011495 HMDB0011524/HMDB0011525	LMGP02050058/LMGP02050059 LMGP02050069/LMGP02050070
LPE C22:6	Lysophosphatidylethanolamine (22:6)	HMDB0011526/HMDB0011496	LMGP02050013/LMGP02050060
LPG C14:0	Lysophosphatidylglycerol (14:0)		LMGP04050012
LPG C16:0	Lysophosphatidylglycerol (16:0)	HMDB0240601	LMGP04050008
LPG C16:1	Lysophosphatidylglycerol (16:1)		LMGP04050013
LPG C18:0	Lysophosphatidylglycerol (18:0)		LMGP04050009
LPG C18:1	Lysophosphatidylglycerol (18:1)	HMDB0240602	LMGP04050006
LPG C18:2	Lysophosphatidylglycerol (18:2)	HMDB0240600	LMGP04050014
LPG C18:3	Lysophosphatidylglycerol (18:3)		LMGP04050032/LMGP04050020
LPG C20:3	Lysophosphatidylglycerol (20:3)		LMGP04050025
LPG C20:4	Lysophosphatidylglycerol (20:4)		LMGP04050010
LPG C22:4	Lysophosphatidylglycerol (22:4)		LMGP04050017
LPG C22:6	Lysophosphatidylglycerol (22:6)		LMGP04050016
LPI C16:0	Lysophosphatidylinositol (16:0)	HMDB0061695	LMGP06050002
LPI C16:1	Lysophosphatidylinositol (16:1)		LMGP06050009
LPI C18:0	Lysophosphatidylinositol (18:0)	HMDB0240261	LMGP06050004
LPI C18:1	Lysophosphatidylinositol (18:1)	HMDB0061693	LMGP06050005
LPI C18:2	Lysophosphatidylinositol (18:2)	HMDB0240597	LMGP06050010
LPI C20:4	Lysophosphatidylinositol (20:4)	HMDB0062722	LMGP06050006
LPI C22:4	Lysophosphatidylinositol (22:4)		LMGP06050013
LPS C16:0	Lysophosphatidylserine (16:0)	HMDB0240605	LMGP03050002
LPS C18:0	Lysophosphatidylserine (18:0)	HMDB0240606	LMGP03050006
LPS C18:1	Lysophosphatidylserine (18:1)	HMDB0240603	LMGP03050001
LPS C18:2	Lysophosphatidylserine (18:2)	HMDB0240604	LMGP03050011
LPS C20:4	Lysophosphatidylserine (20:4)		LMGP03050007
LPS C22:4	Lysophosphatidylserine (22:4)		LMGP03050014
LPS C22:6	Lysophosphatidylserine (22:6)		LMGP03050013
PAF 18:2	PC(O-18:2(9Z,12Z)/2:0)		LMGP01020158
Sphinganine 18:0	Sphinganine 18:0	HMDB0000269	LMSP01020001
Sphinganine-1-phosphate 18:0	Sphinganine-1-phosphate (18:0)	HMDB0001383	LMSP01050002
Sphingosine 18:1	Sphingosine 18:1	HMDB00252	LMSP01010001
Sphingosine-1-phosphate 16:1	Sphingosine-1-phosphate (16:1)	HMDB0060061	LMSP01050005
Sphingosine-1-phosphate 18:1	Sphingosine-1-phosphate (18:1)	HMDB0000277	LMSP01050001
Sphingosine-1-phosphate 18:2	Sphingosine-1-phosphate (18:2)		
1-AG/2-AG	1/2-Arachidonoyl Glycerol	HMDB0011578/HMDB0004666	LMGL01010032/LMGL01010023
1-LG/2-LG	1/2-Linoleoyl Glycerol (18:2)	HMDB0011568/HMDB0011538	LMGL01010006/LMGL01010033
1-OG/2-OG	1/2-Oleoyl Glycerol (18:1)	HMDB0094684/HMDB0011537	LMGL01010005/LMGL01010024
11,12-DiHETrE	(+/-)11,12-DiHETrE	HMDB0002314	LMFA03050008
11-HDoHE	(+/-)-11-HDoHE		LMFA04000028
12,13-DiHODE	alpha-12,13-DiHODE	HMDB0010201	LMFA02000046
12,13-DiHOME	12,13-DiHOME	HMDB0004705	LMFA02000230
12,13-EpOME	12(13)-EpOME	HMDB0004702	LMFA02000038
12-HEPE	(+/-)-12-HEPE	HMDB10202	LMFA03070031
12-HETE	12-HETE	HMDB06111	LMFA03060088
12-HHTrE	12S-HHTrE;12-HHT	HMDB0012535	LMFA03050002

Metabolites	Common_Name*	HMDB_ID	LIPID MAPS_ID
13-HODE	Coriolic acid	HMDB0112194	LMFA02000154
13-HOTE	13-HOTE	HMDB0010203	LMFA02000029
14,15-DiHETrE	(+/-)14,15-DiHETrE	HMDB0002265	LMFA03050010
14-HDoHE	(+/-)-14-HDoHE		LMFA04000030
15-HETE	15-HETE	HMDB0003876	LMFA03060087
16-HDoHE	(+/-)-16-HDoHE		LMFA04000031
17,18-DiHETE	17,18-DiHETE	HMDB0010211	LMFA03060078
19,20-DiHDPA	19,20-DiHDPE	HMDB0010214	LMFA04000043
20-carboxy-LTB4	20-carboxy-LTB4	HMDB0006059	LMFA03020016
20-HETE	20-HETE	HMDB0005998	LMFA03060009
20-hydroxy-PGE2	20-hydroxy-PGE2	HMDB0003247	LMFA03010014
5,6-DiHETrE	(+/-)5,6-DiHETrE	HMDB0002343	LMFA03050004
5-iPF2a-VI	(+/-) 5-iPF2alpha-VI		LMFA03110011
8,12-iso-iPF2a-VI	8,12-iso-iPF2alpha-VI		
8,9-DiHETrE	(+/-)8,9-DiHETrE	HMDB0002311	LMFA03050006
8-iso-PGF2a	15-F2t-IsoP	HMDB0005083	LMFA03110001
9,10,13-TriHOME	9(S),10(S),13(S)-TriHOME		LMFA02000168
9,10-DiHOME	9,10-DiHOME		LMFA02000229
9,10-EpOME	9(10)-EpOME	HMDB0004701	LMFA02000037
9,12,13-TriHOME	9,12,13-TriHOME	HMDB0004708	LMFA02000014
9-HEPE	(+/-)-9-HEPE	HMDB0060053	LMFA03070029
9-HODE	9-HODE	HMDB0062652	LMFA02000151
9-HOTrE	9S-HOTrE; 9-HOTrE	HMDB0031934	LMFA02000024
9-HpODE	9S-HpODE	HMDB0062434	LMFA02000012
9-KODE	9-OxoODE	HMDB0004669	LMFA02000274
AEA	Anandamide;Anandamide (20:4, n-6)	HMDB0004080	LMFA08040001
CA	Cholic acid	HMDB0000619	LMST04010001
CDCA	Chenodeoxycholic acid	HMDB0000518	LMST04010032
DCA	Deoxycholic acid	HMDB0000626	LMST04010040
DHEA	Docosahexaenoyl Ethanolamide; Anandamide (22:6, n-3)	HMDB0013658	LMFA08040009
FA 18:1 ω9 (OA)	Oleic acid	HMDB0000207	LMFA01030002
FA 18:2 ω6 (LA)	Linoleic acid	HMDB0000673	LMFA01030120
FA 18:3 ω3 (ALA)	Alpha-Linolenic acid (alpha-LA)	HMDB0001388	LMFA01030152
FA 18:3 ω6 (GLA)	Gamma-Linolenic acid (gamma-LA)	HMDB0003073	LMFA01030141
FA 20:3 ω3 (DALA)	Dihomo-alpha-linolenic acid (DALA)	HMDB0060039	LMFA01030159
FA 20:3 ω6 (DGLA)	Dihomo-gamma-linolenic acid (DGLA)	HMDB0002925	LMFA01030158
FA 20:3 ω9 (MA)	5,8,11-Eicosatrienoic acid (Mead acid)	HMDB0010378	LMFA01030381
FA 20:4 ω6 (AA)	Arachidonic acid (AA)	HMDB0001043	LMFA01030001
FA 20:5 ω3 (EPA)	Eicosapentaenoic acid (EPA)	HMDB0001999	LMFA01030759
FA 22:4 ω6 (ADA)	Adrenic acid	HMDB0002226	LMFA01030178
FA 22:5 ω3 (ω3-DPA)	Docosapentaenoic acid (omega3-DPA);Clupanodonic acid	HMDB0006528	LMFA04000044

Metabolites	Common_Name*	HMDB_ID	LIPID MAPS_ID
FA 22:5 ω6 (ω6-DPA)	Docosapentaenoic acid (omega6-DPA);Osbond acid	HMDB0001976	LMFA04000064
FA 22:6 ω3 (DHA)	Docosahexaenoic acid (DHA)	HMDB0002183	LMFA01030185
GCA	Glycocholic acid	HMDB00138	LMST05030001
GCDCA	Glycochenodeoxycholic acid	HMDB0000637	LMST05030008
GDCA	Glycodeoxycholic acid	HMDB0000631	LMST05030006
GLCA	Glycolithocholic acid	HMDB0000698	LMST05030009
GUDCA	Glycoursodeoxycholic acid	HMDB0000708	LMST05030016
HCA	Hyocholic acid	HMDB0000760	LMST04010064
HDCA	Hyodeoxycholic acid	HMDB0000733	LMST04010024
LCA-3S	Lithocholic acid 3-sulfate	HMDB0000907	LMST05020015
LipoxinA5	Lipoxin A5		
NO2-OA	9/10-nitro-9E-octadecenoic acid	HMDB0062737/HMDB0062738	LMFA01120003/LMFA01120004
PEA	Palmitoyl Ethanolamide; Anandamide(16:0)	HMDB0002100	LMFA08040013
PGF3a	Prostaglandin F3alpha	HMDB0002122	LMFA03010138
PGK2	Prostaglandin K2		LMFA03010023
ResolvinD3	ResolvinD3	HMDB0257164	LMFA04030012
ResolvinE1	ResolvinE1	HMDB0010410	LMFA03140003
TCA	Taurocholic acid	HMDB0000036	LMST05040001
TCDCA	Taurochenodesoxycholic acid	HMDB0000951	LMST05040005
TDCA	Taurodeoxycholic acid	HMDB0000896	LMST05040013
thromboxane-B2	Thromboxane B2	HMDB0003252	LMFA03030002
TLCA	Taurolithocholic acid	HMDB00722	LMST05040003
TLCA-3S	Taurolithocholic acid 3-sulfate	HMDB0002580	LMST05020003
TUDCA	Tauroursodeoxycholic acid	HMDB0000874	LMST05040015
UDCA	Ursodeoxycholic acid	HMDB0000946	LMST04010033

** The different IDs for the same metabolite indicate the different binding sites between fatty acid and glycerol group*

Appendix 2. List of metabolomics variables

Sums of metabolites	Lipid subclass	Metabolites
ω-3 fatty acids	Fatty acid	FA 18:3 ω3 (ALA) FA 20:3 ω3 (DALA) FA 20:5 ω3 (EPA) FA 22:5 ω3 (ω3-DPA) FA 22:6 ω3 (DHA)
ω-6 fatty acids	Fatty acid	FA 18:2 ω6 (LA) FA 18:3 ω6 (GLA) FA 20:3 ω6 (DGLA) FA 20:4 ω6 (AA) FA 22:4 ω6 (ADA) FA 22:5 ω6 (ω6-DPA)
HETEs	Hydroxy-eicosatetraenoic acids	12-HETE 15-HETE 20-HETE
HEPEs	Hydroxy-eicosapentaenoic acids	9-HEPE 12-HEPE
HDoHEs	Hydroxy-docosahexaenoic acids	11-HDoHE 14-HDoHE 16-HDoHE

Metabolites ratios	Metabolites	Precursor metabolites
ω-6 fatty acids/ω-3 fatty acids	NA	NA
12-HETE/AA	12-HETE	FA 20:4 ω6 (AA)
15-HETE/AA	15-HETE	FA 20:4 ω6 (AA)
20-HETE/AA	20-HETE	FA 20:4 ω6 (AA)
HETEs/AA	HETEs	FA 20:4 ω6 (AA)
9-HEPE/EPA	9-HEPE	FA 20:5 ω3 (EPA)
12-HEPE/EPA	12-HEPE	FA 20:5 ω3 (EPA)
HEPEs/EPA	HEPEs	FA 20:5 ω3 (EPA)
11-HDoHE/DHA	11-HDoHE	FA 22:6 ω3 (DHA)
14-HDoHE/DHA	14-HDoHE	FA 22:6 ω3 (DHA)
16-HDoHE/DHA	16-HDoHE	FA 22:6 ω3 (DHA)
HDoHEs/DHA	HDoHEs	FA 22:6 ω3 (DHA)
DCA/CA	DCA	CA
GCA/CA	GCA	CA
GCDCA/CDCA	GCDCA	CDCA
GDCA/DCA	GDCA	DCA
GUDCA/UDCA	GUDCA	UDCA
TCA/CA	TCA	CA
TCDCA/CDCA	TCDCA	CDCA
TDCA/DCA	TDCA	DCA
TUDCA/UDCA	TUDCA	UDCA
UDCA/CDCA	UDCA	CDCA