Personalized medicine for patients with COPD: where are we?

Abstract: Chronic airflow limitation is the common denominator of patients with chronic obstructive pulmonary disease (COPD). However, it is not possible to predict morbidity and mortality of individual patients based on the degree of lung function impairment, nor does the degree of airflow limitation allow guidance regarding therapies. Over the last decades, understanding of the factors contributing to the heterogeneity of disease trajectories, clinical presentation, and response to existing therapies has greatly advanced. Indeed, diagnostic assessment and treatment algorithms for COPD have become more personalized. In addition to the pulmonary abnormalities and inhaler therapies, extra-pulmonary features and comorbidities have been studied and are considered essential components of comprehensive disease management, including lifestyle interventions. Despite these advances, predicting and/or modifying the course of the disease remains currently impossible, and selection of patients with a beneficial response to specific interventions is unsatisfactory. Consequently, non-response to pharmacologic and non-pharmacologic treatments is common, and many patients have refractory symptoms. Thus, there is an ongoing urgency for a more targeted and holistic management of the disease, incorporating the basic principles of P4 medicine (predictive, preventive, personalized, and participatory). This review describes the current status and unmet needs regarding personalized medicine for patients with COPD. Also, it proposes a systems medicine approach, integrating genetic, environmental, (micro)biological, and clinical factors in experimental and computational models in order to decipher the multilevel complexity of COPD. Ultimately, the acquired insights will enable the development of clinical decision support systems and advance personalized medicine for patients with COPD.

Keywords: chronic obstructive pulmonary disease, personalized medicine, systems medicine, review

Introduction

Chronic obstructive pulmonary disease (COPD) is one of the leading global causes of morbidity and mortality among the non-communicable chronic diseases (NCD).1 Despite preventive measures to reduce exposure to risk factors and therapeutic advances, the worldwide burden of COPD is estimated to increase in the next decades,2 and the disease cannot be cured.3 Chronic airflow limitation assessed by spirometry remains the defining feature of COPD in subjects with respiratory symptoms and a history of exposure to risk factors. However, it is well-recognized that different phenotypes can be distinguished4 and that disease trajectories differ between patients.5 Patients with a comparable degree of airflow limitation may differ considerably in symptoms, functional capacity, and other patient-related outcomes (PROs),6 implicating that additional factors, such as age and
extra-pulmonary conditions, contribute to the burden of the disease in individual patients. Comprehensive lung function assessment, including the degrees of hyperinflation and diffusion impairment, blood gases and mouth pressures, does not adequately reflect these PROs in COPD. Also, differentiation of COPD from asthma is challenging, and these conditions may coexist. In the last decades, the diagnostic workup of COPD has been revised in order to account for the heterogeneity of the disease and thereby enable a more patient-tailored treatment. Current strategies, however, neglect the potential of personalized measures for disease prevention and mainly focus on individualized pharmacotherapies. In order to reduce the burden of COPD and change the natural course of the disease, it is essential to advance our understanding of the complex pathophysiology, develop diagnostic tools that reflect the heterogeneity of the disease and its associated conditions, and be able to predict the response to comprehensive interventions in individual patients. Also, from a health economic perspective, better selection of probable responders for specific or costly interventions is warranted. However, this might not automatically result in a reduction in costs, if diagnostic tools to identify these responders are advanced and expensive.

Systems medicine is the integrated study of disease networks at multiple levels, ranging from the molecular level, through cells, organs, to the population level, in order to yield a comprehensive understanding of disease. This approach is required to provide care that is predictive, preventive, personalized, and incorporates patient participation (P4 medicine). Indeed, a systems medicine approach, integrating genetic, (micro)biological, radiological, clinical, and environmental factors in experimental and computational models may advance personalized treatment of COPD. Recently, the “Systems Medicine-based clinical decision support for COPD patients” (SysMed-COPD) project was initiated by an international and interdisciplinary consortium combining clinicians, clinical and basic scientists, computational and systems biology researchers, and bioinformatics engineers.

The first part of this review focuses on the current status regarding personalized management of COPD, and identifies the unmet needs in the domains of prevention, diagnosis, and assessment and treatment. In the second part, we propose a systems medicine approach towards COPD and highlight its potential to unravel the complex pathophysiology of the disease and provide the basis for P4 medicine in COPD.

Development of COPD
Lung development starts prenatally and continues after birth until maximal lung function is attained in early adulthood. After a plateau phase of approximately a decade, this is ensued by an individual lung function decline. In some individuals, a unique combination of genetic, lifestyle, and environmental factors may prevent reaching the normal plateau or result in accelerated decline, which may eventually lead to the onset of COPD at around 40–50 years of age.

Pathobiology of COPD
Fully-developed COPD is characterized by a combination of individual degrees of (small) airway disease (bronchitis) and destruction of alveolar tissue (emphysema). In most cases, its development is initiated by long-term inhalation of oxidative and cytotoxic substances (eg, cigarette smoke) which induce epithelial and endothelial cell apoptosis, pro-inflammatory signaling, and recruitment of circulating monocytes and neutrophils to the lungs. The activated immune cells secrete proteolytic enzymes including neutrophil elastase, which cause extracellular matrix degradation. Some individuals appear able to maintain normal lung structure and function by mounting an appropriate repair response. In contrast, an insufficient repair response results in disintegration of the lung parenchyma and a dysregulated response in remodeling of the small airways, including smooth muscle cell, goblet cell, and mucus gland metaplasia, as well as subepithelial extracellular matrix deposition. Eosinophils seem to contribute to innate airway inflammation in a sub-group of COPD patients. During later stages of COPD, autoimmune responses may become activated, as indicated by elevated Th17-lymphocytes, B lymphocytes, and antibodies against self-antigens that become exposed during tissue injury.

Once established, the loop of tissue damage, inflammation, and remodeling may self-perpetuate, leading to progressive airflow limitation, even without continued exposure to environmental triggers.

Predisposing factors
Genetic factors have been estimated to account for 40–55% of the variance in adult lung function. Polymorphisms in genes that influence smoking behavior, the protease-antiprotease balance, the oxidant-antioxidant balance, inflammatory processes, or bronchodilator-response contribute to COPD risk. The most well-known gene associated with COPD is SERPINA1, which codes for the serine protease inhibitor alpha1-antitrypsin. Alpha1-antitrypsin inhibits neutrophil elastase and thereby prevents tissue damage. Mutations and polymorphisms of SERPINA1 that lead to
decreased expression or activity of alpha1-antitrypsin are associated with early onset COPD with a strongly emphysematous phenotype, especially in smokers.22

Next to genetics, factors encountered during embryonic development and childhood also influence adult lung function and, therefore, COPD risk.23,24 For instance, maternal smoking during pregnancy, pre-term birth or low birth-weight, and childhood asthma are associated with low lung function in early adulthood.25 In a longitudinal prospective cohort study, 44% of children with severe asthma at age 10 were diagnosed with COPD by the age of 50, irrespective of their smoking history, corresponding to a 32-fold relative risk.26 Lower respiratory tract infections during early childhood are also associated with COPD.27 However, it remains to be clarified whether this is a causative relationship or whether the risks for respiratory infections and COPD are both elevated due to an underlying immune deficiency.

Initiating factors
In the majority of cases, COPD development is initiated by exposure to noxious particles or gases, with the most important risk factor being cigarette smoking.28 Related respiratory irritants arising eg, from passive exposure to tobacco smoke, occupational sources, indoor open fires, and outdoor pollution, are thought to account for a large fraction of the smoking-independent incidence of COPD.29 Most respiratory toxicants are complex mixtures containing up to several thousand chemically distinct compounds, and their exact composition varies considerably depending on their mode of generation.30,31 Therefore, pathological mechanisms are likely to differ according to the differences in toxicant composition. Indeed, the type of exposure may influence the clinical phenotype of COPD.32

Modulating factors
A number of factors do not directly contribute to the predisposition or initiation of COPD, yet influence the course of the disease. Gender is likely to be an important modulating factor, as women develop more severe airflow limitation and emphysema than men with a comparable smoking history.33,34 Yet, when comparing women and men with a similar degree of airflow limitation, women have less emphysema and better oxygenation, but more small airway involvement, more frequent exacerbations, and a poorer quality-of-life.33,34 Also, they have fewer and different comorbidities than men.34 Lifestyle factors, such as diet and physical activity, also modulate COPD development. A number of epidemiological studies have found that intake of diets rich in vitamin C, vitamin E, and β-carotenes (eg, fruit, vegetables, oily fish, whole grains) is positively associated with lung function and, therefore, protects against COPD.35 This may be attributed to the ability of these micronutrients to diminish oxidative stress and oxidative stress-induced inflammation. Higher cardiorespiratory fitness is associated with lower incidence of COPD in the general population,36 and it was shown that improving fitness during childhood and adolescence is associated with greater adult lung volumes.37 Concerning physical activity, active smokers with moderate or high physical activity show attenuated lung function decline and reduced risk of COPD during long-term follow-up compared to those with low levels of physical activity.38 Moreover, COPD patients with regular physical activity have lower rates of hospital admissions and mortality.39

Finally, it has recently been established that the lungs have a microbiome, which may influence the pathogenesis of COPD.40 While it is still unclear whether and how smoking affects lung microbiome composition, the microbiome of COPD patients has a decreased stability and diversity, promoting a disproportionate proliferation of potentially pathogenic species, such as Haemophilus influenzae, Moraxella catarrhalis, and Streptococcus pneumoniae in the lower respiratory tract.40 This outgrowth in turn promotes airway inflammation and, along with viral infections, constitutes an important trigger of acute exacerbations. Importantly, it is known that microbiome composition influences the inflammatory profile: when the microbiome is dominated by Proteobacteria (eg, Haemophilus spp.) or Firmicutes (eg, Streptococcus spp.), this is associated with mediators of neutrophilic or eosinophilic inflammation, respectively.41

Personalized prevention of COPD
In individuals with normal early-adult lung function, the most important strategy for preventing COPD is avoidance of exposure to respiratory irritants. Depending on the source of exposure, this may be achieved by occupational safety procedures such as breathing masks, policies for reduction of air pollution, providing alternatives to indoor open fire for cooking and heating and – most importantly – smoking prevention and cessation. Lifestyle interventions, including prevention of toxic exposures, dietary changes, and increased physical activity, would not only decrease the incidence of COPD but also that of other, often comorbid, chronic diseases.42 The challenge here is to motivate
those at risk to actively participate in preventive measures and to overcome the high dropout rate of lifestyle interventions. A personalized motivational approach, taking into account the individual psychosocial background, would likely improve the intervention success.\(^{43}\) In addition, specific preventive measures against comorbid diseases (for example, treatment of hypertension or dyslipidaemia) have to be taken into account, and alpha 1 antitrypsin augmentation therapy may be considered in deficient individuals.

Preventive measures are particularly important for vulnerable individuals, for instance with a family history of COPD, alpha 1-antitrypsin-deficiency, or who have experienced considerable early life disadvantage (maternal smoking, low birth weight, asthma, frequent and severe respiratory infections, etc.). It can be hypothesized that influenza and pneumococcal vaccination, by avoiding lower respiratory tract infections and associated inflammation, may attenuate lung function decline and lower COPD incidence; however, evidence is currently lacking to support this hypothesis.

**Unmet needs**

Currently, we are lacking effective screening tools to identify people at risk of developing COPD at an early stage. Ideally, a vulnerable population should be identified using a risk score comprising information on the family history of COPD, relevant early life factors, lung function in early adulthood,\(^{42}\) and lifestyle. Second, we need a comprehensive understanding of the different molecular and clinical disease subtypes, as well as correlated specific (companion) diagnostic or therapeutic measures. Patients and people at risk need to be screened over time for activation of pathobiological modules, such as oxidative stress, extracellular matrix degradation, neutrophilic or eosinophilic inflammation, autoimmune effects, and microbiome dysbiosis, to name just a few. Before this is possible, we need to expand our knowledge on the pathobiological modules and to identify corresponding biomarkers. These markers should be measurable with a cost-efficient test that allows screening of large at-risk populations, and should be detectable in biological specimens that can be collected easily during routine visits to the general practitioner, such as exhaled breath, saliva, sputum, or blood. We also need to identify strategies for modification of the pathobiological modules, including those that contribute to comorbidities. Besides the above-mentioned lifestyle interventions, these may include antioxidant and anti-inflammatory drugs, but also anti- or probiotics to induce shifts in the microbiome composition.

**Diagnosis and assessment of COPD**

Obstructive post-bronchodilator spirometry is the defining criterion for COPD in subjects with chronic respiratory symptoms.\(^{3}\) The Global initiative for Obstructive Lung Disease (GOLD) recommends a fixed ratio of forced expiratory volume in one second (FEV\(_1\)), and forced vital capacity (FVC) <0.7 to diagnose the disease and FEV\(_1\) as percent predicted is used for spirometric classification of severity of airflow limitation.\(^{3}\) Since 2011, the current severity of symptoms and the history of moderate or severe exacerbations within the previous 12 months are assessed to allocate patients to groups A/B/C/D and guide therapy.\(^{3}\) Although the classification according to these three domains of the disease enables more personalized treatment recommendations, substantial heterogeneity in clinical characteristics is observed in patients within the same GOLD quadrant.\(^{44,45}\) Also, the GOLD A/B/C/D system does not discriminate between patients in which other factors, for example age or comorbidities, contribute to the burden of disease, and does not adequately differentiate patients with respect to mortality risk.\(^{46}\) Alternative tools to predict outcomes, including the BODE (body mass index, airflow obstruction, dyspnea, and exercise capacity) and ADO (age, dyspnea, and airflow obstruction) indices have been developed and validated in COPD, but cannot be used to guide therapy. In order to overcome the dynamic complexity and heterogeneity of the disease and facilitate treatments targeting the individual needs of patients, the “treatable traits” concept was recently proposed.\(^{47}\) In this concept, all manageable pulmonary and extra-pulmonary features of COPD are identified in individual patients and then treated.

**Unmet needs**

While management of stable COPD became more personalized over the last decades, by incremented understanding of the determinants of the individual burden of disease, current disease management does not meet the demands of P4 medicine. Predicting a beneficial or adverse individual response to specific interventions remains impossible for most; despite our increased understanding of the pathophysiology of COPD, primary or secondary prevention remains problematic. Also, it is not feasible to assess all treatable traits in every individual with COPD, and this
will result in substantial costs for additional diagnostics and treatments, so there is a need for diagnostic algorithms for subgroups of homogeneous patients.48

Another unmet need is the high rate of underdiagnosis of COPD in the general population. While approximately half of patients with severe airflow limitation have a physician-based diagnosis of COPD, only 5% of those with mild lung function abnormality reported a diagnosis of COPD.49 Timely identification and pharmacologic treatment may ameliorate the progressive decline in lung function in mild disease.50 In addition, there is a significant group of symptomatic smokers with preserved spirometry, formerly referred to as GOLD stage 0.51 While only a minority of this group will progress to develop chronic airflow limitation,52 these symptomatic smokers have significant morbidity and healthcare utilization.53 Also, radiologic abnormalities including emphysema and air trapping are common in these subjects.54 Since emphysematous abnormalities predict accelerated loss of lung function55 and increased mortality,56 there is a need for early identification of these subjects. Computed tomography-based measurement of total airway count (TAC) was recently identified as a potential biomarker for airway-related disease changes in early/mild COPD.57 However, other than smoking cessation, no effective treatments are currently available for GOLD stage 0.

Exacerbations
Diagnostics and current treatment
COPD exacerbations are defined as acute worsenings of respiratory symptoms resulting in additional therapy.5 There is large variability in the frequency of exacerbations between patients.58 While a frequent-exacerbator phenotype was previously identified, the proportion of patients with two or more events in consecutive years is very small.59 This observation limits the clinical value of the assessment of future exacerbation risk as currently proposed by GOLD.60 Currently, a history of exacerbations and poor health status remain the best clinical predictors of future events59 and, thus, may drive pharmacotherapeutic decisions.

Although exacerbations are considered pivotal events in the natural course of disease, until recently, there was very little focus on characterizing these events and their triggers. The heterogeneous nature of exacerbations is increasingly recognized; distinct biological clusters have been identified based on sputum biomarkers, namely those associated with bacteria, viruses, or eosinophilic airway inflammation.61 In addition, a further cluster associated with limited changes in the inflammatory profile was identified. Although the pathophysiology of this cluster is largely unknown, it is very well possible that these “pauci-inflammatory” exacerbations are in fact other events mimicking the symptomatology of exacerbations, such as acute cardiovascular diseases, pulmonary embolism, acute psychological distress, or mechanical ventilatory constraints. It is increasingly advocated, therefore, that exacerbations must be differentiated from other events in order to enable a precision medical management of a patient presenting with a flare up of respiratory symptoms.62

Personalized exacerbation management
Bronchodilators, systemic corticosteroids, and antibiotics are central interventions in the management of exacerbations,66 and are broadly prescribed. Strategies to personalize the treatment of exacerbations are not available; sputum cultures may be used to guide antibiotics, but are not feasible in most clinical settings. Use of biomarkers of airway infection, including C-reactive protein (CRP) and procalcitonin, in the decision for antibiotics is not recommended.60 Blood eosinophilia is a promising biomarker for biomarker-directed systemic corticosteroid therapy during exacerbations,63 but warrants further investigation. Obviously, microbiologic, inflammatory, and clinical characterization of exacerbations, more details on pathophysiological interactions and potential treatment targets are needed to tailor more specific, individualized treatment strategies. Moreover, increased understanding of individual genetic variants modulating the immune response involved in exacerbation susceptibility64 and response to treatment65 may enable personalized exacerbation management in future patients.

Comorbidities
Concomitant diseases are common in patients with COPD, and contribute to the burden of disease and mortality in individual patients. They comprise a wide range of diseases, including cardiovascular disease in terms of atherosclerosis, heart failure, arrhythmias, hypertension, peripheral vascular disease and stroke, metabolic disorders including hyperglycemia and diabetes, hyperlipidemia, hyperuricemia, as well as gastroesophageal reflux disease, osteoporosis, rheumatoid diseases, cachexia, and mental disorders such as anxiety and depression.66–71 In addition, there is a substantial overlap with asthma, bronchiectasis, pneumonia, sleep disordered breathing, as well as an
increased lung cancer risk. Almost all patients with COPD exhibit at least one comorbidity, while the majority of patients have four or more other diseases.

In particular, coexistence of COPD and cardiovascular disease is associated with a worse prognosis, ie, increased morbidity and mortality. Moreover, reduction of FEV$_1$ was shown as an independent unfavorable prognostic predictor in decompensated heart failure. Vice versa, the presence of COPD also worsens the prognosis of cardiac disease, as shown in patients hospitalized for acute myocardial infarction or decompensated heart failure.

Recently, it was shown that plasma troponin I concentrations are an indicator of future cardiovascular events and cardiovascular death in patients with COPD and an increased cardiovascular risk, even when in the normal clinical range. Noteworthy, inhaled therapies affect neither troponin concentrations nor mortality or cardiovascular outcomes.

### Current concepts and management of lung-heart interactions

Occurrence of an increased right heart load in COPD is well known since decades and referred to as *Cor pulmonale*. In contrast, also a reduced right heart size has been reported recently, named *Cor pulmonale parvus*. Right ventricular volumes were lower and associated with the degree of pulmonary emphysema. In stable COPD, the degree of airflow limitation and hyperinflation are linked to left ventricular volume, wall stress, and an impaired diastolic filling. Potential mechanisms are a reduced left ventricular mass, and increased thoracic pressure oscillations during breathing in COPD, aggravated at exercise.

Current interventional COPD studies on inhaled corticosteroid (ICS)/long-acting beta2-agonist (LABA) therapy showed short-term increases of left and right ventricular volumes and an improved ventricular strain. The CLAIM study on dual bronchodilation, using combined LABA and long-acting muscarinic antagonist (LAMA) therapy, showed an improved cardiac function, as indicated by left-ventricular end-diastolic volume. Recently, it was shown in this cohort with marked hyperinflation that dual bronchodilation led to an improved pulmonary microvascular blood flow and regional ventilation, as assessed by magnetic resonance imaging. However, it remains to be shown whether these interventional short-term effects translate into persistent cardiac changes during long-term treatment of COPD.

### Personalized management of comorbidities and unmet needs

The list of comorbidities observed in patients with COPD is extensive. Currently, most comorbidities remain undiagnosed and untreated however. Also, in contrast to international recommendations, comorbidities are differently treated in COPD. For example, COPD patients with decompensated heart failure are prone to receive an inappropriate heart failure therapy, which is associated with worse long-term prognosis. Particularly, prescription rates of betablockers are insufficient, although their beneficial effects in COPD patients with cardiovascular comorbidity have been shown.

While it is not feasible to screen for all comorbidities in every patient, there is a need for guidance in the diagnostic approach towards these. Identification of subgroups of patients with an increased risk of specific comorbidities may be helpful. Five patient clusters with distinct comorbidity profiles were previously reported in COPD, including metabolic, cardiovascular, cachectic, psychological, and less comorbidity clusters. While the degree of airflow limitation was comparable, there were significant differences in other lung function parameters, body weight, fat-free mass, and health status between clusters which may guide risk stratification of comorbidities. For example, the proportion of patients with comorbid osteoporosis was highest in the cachectic cluster, while health status was worst in patients in the psychological cluster. Neither circulating inflammatory biomarkers nor pack years of smoking differed between clusters, indicating that the origin of these profiles is multifactorial and largely unknown. Integrated analyses of not only the clinical disease network of COPD, but also incorporating endogenous (genetics, inflammation, oxidative stress, microbiology, aging, repair mechanisms), environmental (smoking, air pollution, physical activity, diet), and socio-economic factors (preterm birth, family size, employment) is warranted. A systems medicine approach for COPD provides this opportunity.

### Pharmacologic management

#### Historic and current treatment algorithms

In the context of this document, it seems preferable to focus not so much on specific drugs but on treatment concepts. These are nicely reflected in the GOLD documents. In 2001, the first GOLD statement was published. The committee suggested to base treatment recommendations primarily on the level of airflow limitation. Regarding treatment...
options, the following proceedings were proposed: first, bronchodilators come first. Combinations of bronchodilators with different mechanisms of action may be preferable. Second, inhaled corticosteroids (ICS) can be given if significant symptoms and a significant lung function response (in a 6 week to 3 months trial with ICS) and/or repeated exacerbations are present.

In the 2007 Executive Summary, a stage dependent step up of treatments was suggested. Now, the use of ICS should only be considered in patients with an FEV₁ below 50% of predicted and a history of repeated exacerbations. This recommendation was mostly based on studies with fixed-dose LABA-ICS combinations. The abovementioned ICS test was no longer recommended.

In the 2011 document, the new assessment system based on spirometry, symptom load and exacerbation history was introduced and patients were categorized into four groups (A, B, C, D) with separate treatment recommendations for each group. In this document the goals of treatment were defined as reduction of symptoms and reduction of future risk, in particular exacerbations.

In the 2017 version, the assessment system was refined. The A, B, C, D groups that define the pharmacological treatment were based on symptoms and exacerbation history only. The most relevant reason for this change was that the exacerbation history proved to be a better predictor of future risk of exacerbations than spirometric impairment. Now for each group treatment algorithms were introduced, including escalation and de-escalation (for ICS) concepts. Besides, the role of non-pharmacologic treatments including pulmonary rehabilitation (PR) and other measures that may increase physical activity is emphasized. It is stated that these treatments are equally important as drugs and that combining adequate pharmacological and non-pharmacological therapies is mandatory.

In the 2019 document, a further refinement was described. Groups A, B, C, D are now used for informing the initial treatment only. Regarding follow-up, two strata are suggested: patients that suffer primarily from symptoms (ie, dyspnoea) vs exacerbations. For both strata, separate treatment algorithms are proposed. Further, blood eosinophil counts are introduced as a biomarker for the likelihood that treating with an ICS may reduce the exacerbation rate. A management cycle (review – assess – adjust) is defined with the goal to identify reasons why patients are not doing well and adapt the treatments accordingly. Influenza vaccination is recommended as it reduces the incidence of lower respiratory tract infections and mortality in COPD, while pneumococcal vaccination decreases risk of pneumonia, especially in younger patients with severe COPD and those with comorbidities.

This brief history nicely shows how the pharmacological COPD treatment has evolved. The mainstay of pharmacological treatment was and is that bronchodilators come first for most patients – either as mono-treatment or as a combination of two bronchodilators with different mechanisms. Because of a superior efficacy for long-term treatment, long-acting bronchodilators are preferred. Next, ICS are only considered in individuals with exacerbations or a history of asthma. ICS containing regimens are combinations of long acting β2 agonists (LABA) and ICS or of long-acting muscarinic antagonists (LAMA), LABA, and ICS. Third, besides ICS, other anti-inflammatory treatments (roflumilast, macrolides) may be used in selected patients.

Coming from an all-comers concept where treatment decisions were mostly based on spirometry data, we now choose drugs using a more tailored patient stratification system a) that is based on symptoms and exacerbation risk, where b) treatment is adapted to the predominant clinical problem (dyspnea and/or exacerbations), c) blood eosinophil counts are used as a biomarker when the use of an ICS is considered, and d) the combination of pharmacological and non-pharmacological treatments is advocated.

**Unmet needs**

The treatment concepts of the future have to be personal or at least better tailored to the patient's needs. In this context, open research questions have been summarized in an ATS/ERS statement. Based on this, there are a number of unmet needs that need further research. The first need relates to risk factors; are there differences regarding treatment effects between COPD caused by biomass exposure and smoking induced COPD? Do patients with reduced lung growth have a different treatment response than individuals with normally developed lungs? What is the role of gender and age? The second need concerns the treatable traits; to what extent is it useful and feasible to implement elements of treatable traits in treatment decisions? So far only airflow obstruction, exacerbation history, and blood eosinophil counts are taken into account. What about chronic bacterial colonization, bronchiectasis, emphysema, or asthmatic features? In order to make progress in this field there is an urgent need for (more) biomarkers. Finally, several questions in the field of drugs and non-pharmacological treatments remain unanswered; what are
the benefits of pharmacological treatment of comorbidities (including cardiovascular disease and systemic inflammation) on respiratory outcomes and vice versa? What drugs are best to be combined with measures to improve physical activity for longer terms? What concepts regarding physical activation show the best results (hospital-based, ambulatory, internet-based)? How to incorporate patient preferences and priorities in the treatment algorithms for COPD?

Non-pharmacologic management
In addition to pharmacologic treatment, non-pharmacologic treatments play an important role in COPD management. Self-management education aims to adapt the health behaviour of patients and provide skills to personally manage the disease on an everyday basis. Reduction of behavioural risk factors, treatment adherence, and coping with symptoms are important aspects of self-management. Depending on the determinants of disease burden and complexity of interaction between these, additional non-pharmacological interventions may range from monodisciplinary interventions such as exercise training or nutritional supplementation to comprehensive interventions such as pulmonary rehabilitation (PR).99 Per definition, PR is a personalized intervention. Following a comprehensive assessment of the integrated health status of a patient with COPD, PR is a patient-tailored non-pharmacological intervention including exercise training, self-management education, nutritional counselling, psychological support, occupational therapy, and other treatments.100 PR is an evidence-based intervention associated with improvements in symptoms, health status, exercise tolerance, and healthcare utilization.100 Personalization of PR not only comprises variation in the components of the program between patients, but also flexibility of interventions within these components. While exercise training is considered the cornerstone of PR, this intervention may consist of endurance training, interval training, resistance training, nonlinear training, water-based training, neuromuscular electrical stimulation, or whole-body vibration training, based on the symptoms, degree of disability,101 and determinants of exercise intolerance in the individual patient.102

Unmet needs
As with pharmacological interventions in COPD, non-response to PR is common.103 Changes in exercise performance and health status are often used to qualify individuals with COPD as responders or non-responders to PR. However, it was shown that the response to PR may be differential; patients may improve health status without an improvement in exercise capacity or vice versa.104 Within the same domain, response may depend on the method of assessment.104 Using a non-parametric regression technique, four different clusters with a distinct multidimensional response to PR were identified,104 including groups of patients that show only a moderate (35.4%) or even a poor (10.5%) response. We need to better understand these groups as these may require an intensification or redesign of existing programs.

Interventional therapies
In COPD patients with pulmonary emphysema, air trapping and hyperinflation of the lungs contribute to dyspnoea, poor health status, and exercise intolerance.105 Despite optimal pharmacotherapy and PR, severe emphysema is associated with enormous refractory disability. Lung volume reduction is a treatment option for those with the most severe emphysema and results in improved health status, exercise tolerance, and even survival in carefully selected patients.106 As such, it is a nice example of personalized medicine in COPD. In contrast to other domains of disease management, predictors for beneficial response after lung volume reduction have been identified; patients with predominantly upper-lobe emphysema and low exercise capacity were shown to have a survival advantage following surgical lung volume reduction.107 In addition, absence of interlobar collateral ventilation is a predictor of response to endobronchial valve treatment.108

Unmet needs
Diagnostic workup of patients is complicated, techniques are expensive and require specific expertise, and most patients that are screened for lung volume reduction treatments do not meet the inclusion criteria for these interventions.108 However, the field of interventional pulmonology for COPD is expanding rapidly, and new techniques are being investigated.

Systems medicine model for COPD
As described before, current therapies for stable COPD and exacerbations are frequently inadequate to halt disease progression and unable to cure the disease. In addition, we lack sufficient tools to predict an individual’s disease progression or response to therapies. Systems medicine promises to improve our understanding of lung health and COPD. Thus, it has the potential to facilitate development of effective, personalized and
Ideally preventive interventions. The definition of systems medicine has in recent years been extended beyond an approach to medical understanding and treatment that is based on multi-disciplinary healthcare teams and that integrates biomolecular, psychological, and social dimensions. Now, systems medicine includes novel techniques for systematic assessment from molecular omics to physiome and diseasome, and the formalization of mechanistic hypotheses into computational predictive models. In this focus on mechanistic understanding, systems medicine differs from current data-heavy machine learning approaches. In the following we will focus on the research oriented part of systems medicine combing clinical, epidemiological, and experimental research with mathematical modeling, computer science, and machine learning to model specific pathophysiological modules via computer simulation to test hypotheses, identify preventive measures, derive new biomarkers, and predict treatment outcomes.

Recent clinical trials have shown that targeted therapies with small chemical compounds such as kinase inhibitors and therapeutic antibodies can serve as highly specific tools to manipulate inter- and intracellular regulatory pathways. Combined, both developments – in principle – allow correcting dysregulated physiological modules (protease-antiprotease balance, oxidant-antioxidant balance, inflammation, etc.) in a personalized way for every patient or at least for patient sub-group. As discussed above, PR and integrated care as well as preventive measures are further aspects relevant to implement a systems medicine-based approach to COPD. The big challenge to translation from research is to develop, test, and implement strategies and tools to gain clinically meaningful insights for everyday patient care. The systems medicine research and translation strategy critically depend on the involvement of patients and patient representatives, as well as all relevant stakeholders in the health systems.

**Experimental models**

Observational and interventional clinical studies are prerequisites for systems medicine approaches. However, certain analyses are necessary for the modeling process, but cannot be performed in humans, eg, because they are difficult to observe (eg, very early events in COPD development), technically not possible (eg, in vitro cellular imaging of the lung, very detailed time course measurements in the lung), or would be unethical (eg, early drug testing, genetic manipulation). Therefore, experimental models are indispensable in systems medicine. New and improved molecular technologies, like single cell sequencing and genome editing by CRISPR technology, enable us to deepen our pathophysiological insight to a detailed level.

Frequent models in biomedical research are inbred mice. For COPD research, mouse models of acute or chronic exposure to cigarette smoke are established and under investigation. Advantages are the analysis of whole organisms and the existence of knock-out mice and numerous molecular tools. Disadvantages are relevant differences in pulmonary anatomy, immunology, and the smoke-induced clinical phenotype in comparison to COPD patients. Complementary models are ex vivo cultures of surgically removed lung tissue, or air–liquid interface cultures of epithelial airway cells as obtained by bronchoscopy, which undergo differentiation in culture to all relevant cells types, produce mucus, and display ciliary beating. These models allow access to “real” human COPD, are relatively easy to work with, but lack crosstalk to other organ systems. Additionally, new organ-on-a-chip and bioreactor models allow the cellularization of microfluidic devices or tissue scaffolds with epithelial and endothelial cells, as well as flow of air at the apical compartment and buffer, including immune cells, in the endothelial compartment. Finally, primary human cells or cell lines can be exposed in vitro to cigarette smoke or cigarette smoke extracts.

These models can be analyzed by manifold imaging (eg, high resolution lifetime microscopy), immunology (eg, flow cytometry), and molecular (eg, single-cell sequencing, proteomics, metabolomics) technologies. They complement clinical studies and provide comprehensive data sets for calibration, optimization, and validation of models derived from computer simulation.

**Computational models**

A computational model is a mathematical model aimed to investigate the behavior of a complex system. Due to the focus of systems medicine on the elucidation of medical questions, computational models usually have a certain predictive ability. Different sub-classes of computational models can be distinguished. Mechanistic models are traditionally based on the causal understanding of biological entities (eg, cells, proteins, organs) and their dynamic interactions. Basically, these models represent existing knowledge about biological systems in a form that can generate predictions on the behavior of the system. Statistical and
machine-learning based models are based on associations rather than causal mechanisms, and can provide direct and strong predictions of (clinically) relevant outcomes.

**Mechanistic models**

Due to the complexity of COPD, which affects multiple molecules, cell-types, organs, and psycho-social factors, only few attempts to apply mechanistic models to predictive, personalization of treatment have so far been reported. Many of these models have been based on ordinary differential equations (ODE) to calculate the change over time in certain factors. Mostly, these aim to improve prediction of progression or exacerbation risk, for example by coupling exposure to inhaled substances, biological activity in terms of inflammation, and the degree of airflow limitation to progression. By doing so, they provide personalized estimates of ventilation/perfusion heterogeneity in COPD patients, mechanistically explain etiology from an imbalance of protease-antiprotease, leading to a cascade of positive feedback loops in tissue-degeneration, macrophage and neutrophil inflammation, and alveolar epithelial cell apoptosis, or generate a physiological model of breathlessness. None of these models have so far resulted in clinical application. However, models based on computational fluid dynamics have recently contributed to treatment optimization by predicting particle deposition in the airway tree, and are supported by the FDA as prognostic biomarkers in clinical trials. All models described so far were based on a single modeling technology such as ODE to describe effects on a single level, eg, molecular or cellular. In contrast to these, several approaches recently embarked on the generation of multi-scale, multi-physics models which combine different technologies such as ODE, agent based model, or partial differential equations (PDE). One such approach described the lung in a patient-specific way, from its molecular constituents to its cell types and overall biophysical structure and properties. This multi-scale model already created clinical impact by debunking thermoplasty as biologically meaningful intervention to asthma. Coupling such a patient specific lung model to a multi-scale full-body model to explain muscle wasting in COPD and suggest specific PR approaches was recently reported by integrating the ventilation/perfusion lung model with a physiological gas-exchange and distribution model, and a model of muscle cell energy generation and regulation of re-modeling. In summary, mechanistic models make inroads to generate patient specific predictions to aid clinical decisions, but have not yet reached impact beyond clinical trial stage.

**Machine learning models**

The increase in size and availability of patient datasets related to COPD enables us to employ statistical tools such as machine learning (ML) algorithms in order to solve scientific and clinical problems. ML algorithms use data to predict certain outcomes, and algorithms can learn and improve their predictions. ML methods are divided into two types. Supervised ML algorithms learn rules from annotated data (eg, labeled emphysema in CT-images) and apply them to classify new data samples (eg, unlabeled CT-image of emphysematous lung). Unsupervised ML algorithms use unlabeled data to detect patterns, eg, principle component analysis (PCA) or clustering. Unsupervised ML algorithms are well known in medical research, while supervised ML algorithms became popular in the last two decades and achieved promising results, for example in skin cancer classification and diabetic retinopathy detection.

One main advantage of using ML algorithms (instead of traditional mechanistic models) is that they are easily adaptable and do not require a deep understanding of the underlying mechanisms of the system. Supervised ML algorithms can learn with the arrival of new data and improve the outcomes with minimal effort. In contrast, dynamic models need to be re-examined, the mechanisms and pathways need changes, and these new models must be re-analyzed to understand the outcomes, including re-estimation of the parameters using the new data. This reconstruction of the mechanistic models is often a difficult and expensive process.

We review here a number of publications focusing on supervised ML algorithms that support clinical decisions for COPD (Table 1). The most prominently used ML algorithms are k-nearest neighbor (KNN), support vector machine (SVMs), and artificial neural network (ANN). KNNs predict the label of a new data sample by finding their closest data points (nearest neighbors) in the training dataset. The nearest neighbors then classify the new data point by vote (Figure 1A). SVMs try to find a line or plane that separates data samples with different labels. If that is not possible, the data points are transformed into a new dimension (Figure 1C). This makes it possible to separate the data samples with, for example, a two-dimensional plane. The ANNs are assembled by neurons organized in layers (Figure 1B). Each neuron sums up its weighted inputs and applies a non-linear activation function on the sum. The calculated value either functions as an input for the neurons
in the next layer or as an output of the ANN. Using non-linear activation functions, ANNs are able to find complex boundaries to separate differentially labeled data points.

Most of the reviewed publications report high performance results of their tested ML algorithms, for example a classification accuracy >90% or even >98%. However, we must consider several important issues in these COPD ML studies. First, many studies use additional “feature selection” to decrease the complexity of their data, and thus focus only on the relevant data variables. Feature selection methods are often used to remove irrelevant and redundant attributes that do not contribute to the accuracy of the outcome or may in fact decrease the accuracy of the predictive model. Feature selection, therefore, improves the performance of their ML algorithm, but it should be used on an independent dataset rather than the entire training data. This is not possible in most COPD ML studies that have access only to small datasets (<300 patients). Using feature selection on the entire training set may yield a model that is enhanced by the selected features over other models being tested to get seemingly better results, when in fact it is a biased result. Second, when we examined the system performance, we found that more than half of the studies did not report the use of an independent test dataset (data that was not used in the training procedure). A missing independent test dataset can lead to an overestimation in the performance of the ML algorithms and should always be implemented in the development and evaluation of a ML-based system. Also, to address the multi-component background of COPD, scientists should incorporate a variety of independent measurement techniques. Despite that, we found that most of the reviewed publications focused on one or two different data types (source of the data such as measurement type). Only some\textsuperscript{145–152} reported the use of a broader patient health assessment (eg, spirometry, blood analysis, CT images, etc.). Finally, it is generally known that a limited training dataset results in poor approximation, especially for complex ML algorithms and systems.\textsuperscript{153} Most of the COPD ML studies we reviewed have limited access to data, and their training datasets range from 16–300 patients. This may be one reason why none of the studies\textsuperscript{154,155} reported testing their algorithm in a real-world clinical setting, suggesting a gap between research in COPD ML algorithms and applications in daily clinical settings. To close this gap, a representative training dataset is needed. The data should also be based on a detailed patient assessment to reflect the multi-component background of COPD.

### Table 1: Overview of the reviewed publications focusing on machine learning in clinical decision support for COPD

<table>
<thead>
<tr>
<th>Aim</th>
<th>Type of data</th>
<th>ML algorithms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis (and classification) of COPD</td>
<td>Pulmonary function tests</td>
<td>LM, KNN, NB, DTREE, RF, GB, SVM, ANN, ANF, rule-based, ensembles</td>
</tr>
<tr>
<td>Detection of emphysema</td>
<td>CT images</td>
<td>LM, KNN, SVM, ANF, DTREE, ANF, rule-based, ensembles</td>
</tr>
<tr>
<td>Prediction of exacerbations</td>
<td>Respiratory sounds</td>
<td>RF, SVM</td>
</tr>
<tr>
<td></td>
<td>Patient health status</td>
<td>NB, BN, DTREE, RF, SVM, ANF</td>
</tr>
</tbody>
</table>

**Abbreviations:** AIS, artificial immune system; ANN, artificial neural network; BN, Bayesian network; DTREE, decision tree; GB, gradient boosting; LM, linear models; KNN, k-nearest neighbor; ML, machine learning; NB, Naïve Bayes; RF, random forest; SVM, support vector machine.
Statistical models with clinical application

Many approaches to the personalization of COPD risk assessment or treatment suggestion are based on multilevel models, combining classical statistical analysis and machine learning approaches to clinical data with mechanistic modeling of biomedical research data. For example, they use a Bayesian network algorithm to derive clinical variables predictive of exacerbation risk and suggest context-aware preventive action\(^{183}\) or identify survival risk factor attributes by univariate analysis to generate probability distribution models to predict ICU COPD mortality risk.\(^{184}\) Many of these combined models straddle the border between personalized health-behavior advice and public health preventive policy advice such as the personal air-exposure monitoring and exposure-health-association analysis and exacerbation risk prediction model for London which couples exposure-health-association to a time-activity exposure multi-scale ODE model.\(^{185}\) A linked-equation model enabled direct estimation of health service costs and quality-adjusted life years (QALYs) for COPD patients over their lifetimes and was validated for predicted annual exacerbation rate and annual decline in FEV\(_1\) with a 3-year longitudinal clinical study\(^{186,187}\) to focus treatment attention on relevant factors. Recently, a patient-level, health economic simulation model including a large number of patient characteristics and relevant outcomes was developed, which can be used to personalize treatment decisions in COPD.\(^{188}\) While influence of physical activity\(^{189}\) has resulted only in a few models, the personalized prediction of exacerbation risk is a major focus of statistical models.\(^{152,190–193}\)

Unmet needs

FAIR (findable, accessible, interoperable, reusable) models are required that bridge association-based (machine learning, statistical) and mechanism-based modeling.\(^{194}\) In addition, we need multi-scale, integrated computational models with medically relevant outputs on the level of the individual and focused on specific medical questions.

Access to relevant, high quality data and gold standards for testing and validation is essential to build

---

**Figure 1** Different machine learning algorithms. (A) k-nearest neighbor (KNN); (B) artificial neural network (ANN); (C) support vector machine (SVMs). Different algorithms are explained in the "Machine learning models" section.

**Notes:** Class represents diagnostic classification, for example as ‘normal’ or ‘abnormal’ or representing different stages of disease. Test represents new cases entering classification.
association-based models. Large scale real-world data access efforts are ongoing in public health systems and will greatly improve our ability to learn.\textsuperscript{195} As data infrastructures become more important for the ability to improve healthcare, we believe clinics and health systems embracing them will flourish.

Beyond infrastructural and methodological tasks, we need clinical modelers. Too often, even in systems medicine, the disciplines are still separated due to experts staying in their own area; clinicians posing diagnostic and treatment questions and computational and mathematical experts focusing on methodological breakthroughs and fundamental insights. However, as long as “translators” intervene between the clinician and the model, we will lack the direct interaction giving rise to intuitive understanding and explorative generation of new questions, ideas, and hypotheses for clinical research. Therefore, we need models that are user friendly and accessible to clinical researchers to enable them to explore and pose their questions directly in-silico.

Towards a clinical decision support system for COPD

Given the availability of very large datasets and the increasing capability of machine learning approaches, clinical benefit could be optimized and patient risk minimized, developing a dynamic clinical decision support system (CDSS).\textsuperscript{196} In the clinical context, the term CDSS is often used for any computational system that provides direct aid to clinical decision-making, such as dashboards that present information in a comprehensive, actionable form and, therefore, help clinicians to integrate and prioritize multiple, diverse evidence.\textsuperscript{197,198}

However, in the context of systems medicine and computational modelling, we regard CDSS as software that matches patient characteristics to a computerized clinical knowledge base (KB) and then presents patient-specific assessments or recommendations to the clinician and/or the patient for direct clinical decision-making.\textsuperscript{199} These systems interpret or advise for action and, therefore, are, at least in Europe, regulated as “medical devices” (EC 2017/745). Existing systems range from CDSS for spirometry quality control,\textsuperscript{200} to integrated-care based applications creating individual treatment pathways from multiple intervention modules based on a broad biopsychosocial patient assessment.\textsuperscript{48,201} The results showed by the available CDSS for the detection and diagnosis of COPD are promising and can be used in combination with the existing protocols to facilitate disease management.\textsuperscript{202} However, none of these COPD-specific systems has yet reached regulatory approval for clinical practice. Outside the field of COPD, examples of systems start to appear which support complex intervention plans, such as multi-perturbation treatment\textsuperscript{203} or dynamic adaptation of biopsychosocial factors (lifestyle, environment, medication).\textsuperscript{204}

Too many CDSS have been proposed but never underwent dedicated validation trials at the point-of-care, and even less are enabled for continuous improvement on real world data from clinical practice. One of the difficulties to this is the lack of protocols and processes for regulatory approval of continuously adapting CDSS as a medical device.\textsuperscript{205} In addition, integration of CDSS in healthcare systems where patient participation and shared decision-making are increasingly important is a major challenge.\textsuperscript{206}

Conclusion

As a consequence of advances in our understanding of the pathophysiology of COPD at multiple levels and in comprehensive diagnostic and therapeutic strategies over the last decades, disease management transformed from a “one size fits all” towards a more personalized approach. However, additional, far-reaching efforts are required to transform care into actual P4 medicine for COPD, where the emphasis is not only on personalized management, but also on predictive and preventive healthcare in which individual patients are actively involved. A multidisciplinary systems medicine approach may reveal the multilevel complexity of COPD and fill current gaps in optimization of treatment for individuals at risk and those with established airflow limitation.

Acknowledgments

We would like to apologize to all colleagues whose excellent contributions to the field of personalized COPD medicine could not be included in this text due to space constraints. Part of this work has been funded by German Ministry for Education and Research (BMBF) (ERACoSysMed2 SysMed-COPD-FKZ 031L0140, JPIAMR Pneumo-AMR-Protect-FKZ 01KI1702, e:Med CAPSYS-FKZ 01X1304E/01ZX1304F) to BS and by a Kootstra Talent Fellowship from the Center for Research Innovation, Support and Policy (CRISP) of Maastricht University Medical Center + to BJB. ZonMW (ERACoSysMed
Disclosure
FMEF received personal fees for lectures and consultancies from AstraZeneca, Boehringer Ingelheim, Chiesi, TEVA, GlaxoSmithKline, and Novartis, outside of this work. He also received research grants from Novartis and MedImmune. BS received research funding from GlaxoSmithKline. SI reports grants from Austrian Science Fund FWF, during the conduct of the study. DM reports grants from Austrian Science Fund FWF, during the conduct of the study. MM reports grants from German Ministry for Education and Research (BMBF), during the conduct of the study. CFV and SI also received research grants from Novartis and TEVA. FMEF received personal fees for lectures and consultancies from AstraZeneca, Boehringer Ingelheim, Chiesi, TEVA, Cipla, Bayer Schering, MSD, and Pfizer, outside the submitted work. BS reports grants from BMBF and GlaxoSmithKline, during the conduct of the study. The authors report no other conflicts of interest in this work.

References


149. Deep Neural Networks for Prediction of Exacerbations of Patients with Chronic Obstructive Pulmonary Disease. Cham: Springer International Publishing; 19th International Conference, EANN 2018, September 3-5, 2018; Bristol, UK.


162. Classification of COPD with Multiple Instance Learning. *IEEE; 2014.22nd International Conference on Pattern Recognition, Stockholm, Sweden


