LUND UNIVERSITY

Patient Heterogeneity in Health Economic Decision Models for Chronic Obstructive Pulmonary Disease
Are Current Models Suitable to Evaluate Personalized Medicine?

Hoogendoorn, Martine; Feenstra, Talitha L.; Asukai, Yumi; Briggs, Andrew H.; Borg, Sixten; Dal Negro, Roberto W.; Hansen, Ryan N.; Jansson, Sven Arne; Leidl, Reiner; Risebrough, Nancy; Samyshkin, Yevgeniy; Wacker, Margarethe E.; Rutten-van Mölken, Maureen P M H

Patient heterogeneity in health economic decision models for chronic obstructive pulmonary disease (COPD): are current models suitable to evaluate personalized medicine?

Running title: patient heterogeneity in COPD models

Martine Hoogendoorn [1], Ph.D.

Talitha L. Feenstra [2, 3], Ph.D.

Yumi Asukai [4], M.Sc.

Andrew H. Briggs [5], DPhil.

Sixten Borg [6-8], M.Sc.

Roberto W. Dal Negro [9], MD, Ph.D.

Ryan N. Hansen [10], Pharm.D., Ph.D.

Sven-Arne Jansson [11], Ph.D.

Reiner Leidl [12], Ph.D.

Nancy Risebrough [13], Ph.D. ?

Yevgeniy Samyshkin [4], M.Sc.

Margarethe E. Wacker [12], M.Sc

Maureen P.M.H. Rutten-van Mölken [1], Ph.D.

[1] Institute for Medical Technology Assessment (iMTA), Erasmus University Rotterdam, Rotterdam, The Netherlands

[2] Department for Prevention and Health Services Research, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

[3] Department of Epidemiology, University Medical Centre Groningen, Groningen, The Netherlands

[4] IMS Health, Health Economics and Outcomes Research and Real-World Evidence Solutions, London, UK

[5] Institute of Health & Wellbeing, University of Glasgow, Glasgow, UK

[6] The Swedish Institute for Health Economics, Lund, Sweden

[7] Health Economics Unit, Department of Clinical Sciences in Malmö, Lund University, Sweden

[8] Evidera, London, UK.

[9] National Center for Respiratory Pharmacoeconomics and Pharmacoepidemiology (CESFAR), Verona, Italy.

[10] Pharmaceutical Outcomes Research and Policy Program, School of Pharmacy, University of Washington, Seattle, WA, USA

[11] Department of Public Health and Clinical Medicine, Occupational and Environmental Medicine, The OLIN Unit, Umeå University, Umeå, Sweden

[12] Helmholtz Zentrum München, Institute of Health Economics and Health Care Management, Member of the German Center for Lung Research, Comprehensive Pneumology Center Munich, Neuherberg, Germany

[13] ICON Health Economics, Toronto, Canada

Correspondence and request for reprints:

Martine Hoogendoorn

Institute for Medical Technology Assessment

Erasmus University Rotterdam,

P.O. Box 1738

3000 DR Rotterdam

The Netherlands

Phone: *31 10 4088871

Email: hoogendoorn@imta.eur.nl

Key words:

COPD, model, patient heterogeneity, validation

Patient heterogeneity in health economic decision models for chronic obstructive pulmonary disease (COPD): are current models suitable to evaluate personalized medicine?

Abstract

Objectives: To assess how suitable current COPD cost-effectiveness models are to evaluate personalized treatment options for COPD by exploring the type of heterogeneity included in current models and by validating outcomes for subgroups of patients.

Methods: A consortium of COPD modelling groups completed three tasks. First, they reported all patient characteristics included in the model and provided the level of detail in which the input parameters were specified. Second, groups simulated disease progression, mortality, quality-adjusted life-years (QALYs) and costs for hypothetical subgroups of patients that differed in terms of gender, age, smoking status and lung function (FEV$_1$% predicted). Finally, model outcomes for exacerbations and mortality for subgroups of patients were validated against published subgroup results of two large COPD trials.

Results: Nine COPD modelling groups participated. Most models included gender (seven), age (nine), smoking status (six) and FEV$_1$% predicted (nine), mainly to specify disease progression and mortality. Trial results showed higher exacerbation rates for females (found in one model), higher mortality rates for males (two models), lower mortality for younger patients (four models), and higher exacerbation and mortality rates in patients with severe COPD (four models).

Conclusions: The majority of currently available COPD cost-effectiveness models are able to evaluate the cost-effectiveness of personalized treatment based on gender, age, smoking and FEV$_1$% predicted. Treatment in COPD is however, more likely to be personalized based on clinical parameters. Two models include several clinical patient characteristics and are therefore most suitable to evaluate personalized treatment, although some important clinical parameters are still missing.

Introduction

Personalized medicine has the potential to improve the (cost-)effectiveness of treatments and contribute to health care cost containment. Hence, interest in personalized medicine has increased exponentially in the past decade [1]. The concept of personalized medicine has different definitions [2, 3]. Some definitions focus on the use of genetics, proteomic, cytomic and/or metabolic biomarkers to define patient geno- and phenotypes most likely to benefit most from a certain treatment. Other definitions are broader and refer to customizing treatment to the individual characteristics, needs, and preferences of a patient during all stages of care, including prevention, diagnosis, treatment, and follow-up [3]. The expectations of personalized medicine are high, but so far, cost-effectiveness of such technologies on average has not been found to be any better than that of non-personalized interventions [4, 5]. Thus, economic evaluation studies are needed to show whether the personalized approach indeed leads to greater efficiency in health care compared to the one size fits all approach. If based on health economic models, such economic evaluations of personalized treatment strategies based on demographic and clinical patient characteristics, require models that are able to address the heterogeneity due to those patient characteristics appropriately. However, the majority of currently available cost-effectiveness models are cohort Markov or state-transition models that have only limited capability for addressing patient heterogeneity. The majority of published cost-effectiveness analysis similarly focused on results for the average patient. Current guidelines for health economic modelling recognize the importance of patient heterogeneity, but there is a lack of consensus on how to best address this in a modelling framework [6].

One of the major chronic diseases for which treatment is increasingly personalized is chronic obstructive pulmonary disease (COPD). In 2011 the new Global initiative for Obstructive Lung Disease (GOLD) guidelines proposing a new classification system for COPD patients based on lung function, symptoms and exacerbations. Although this new classification is a step forward to personalized treatment for COPD, proposed treatment guidelines for the subgroups of patients still need to be validated [7]. Manufacturers of drugs and devices focus are increasingly focused on specific phenotypes of patients, such as the patients with frequent exacerbations, a rapid decline in lung function, persistent systemic inflammation, raised eosinophil concentrations, bacterial colonization etc [8, 9]. In addition, caregivers providing multidisciplinary integrated care programs develop different treatment modules for subgroups of patients who need physical reactivation,

4

smoking cessation support, nutritional interventions, treatment of depression etc. In contrast, the majority of decisions about the reimbursement of new treatments is still made for large groups of patients and based on evaluations that do not consider patient heterogeneity. This is expected to change rapidly because biologics for COPD are under development [10] and payers are likely to limit the reimbursement of these expensive drugs to specified subgroups. Moreover, new drugs for COPD that have recently been launched or are still under development are often quite similar to the currently marketed drugs which makes it increasingly difficult to demonstrate their value for money, unless one can specify the subgroup for which they are particularly beneficial.

In the past years several health economic decision models for COPD have been published. In 2011 our modelling group took the initiative to establish a network of people involved in COPD modelling around the world (COPD modelling teams, pharmaceutical companies interested in COPD modelling, epidemiologists, clinicians etc). Up to now, we have organized three one-day meetings in 2011, 2012 and 2014 with the aim to compare the different available models with respect to model structure and input parameters and to cross validate the models against each other [11]. Proceedings of the third meeting regarding patient heterogeneity are described in the current paper.

The aim of the paper is to assess how suitable current models are to evaluate the cost-effectiveness of personalized treatment options for COPD. Firstly, we explored which type of patient heterogeneity is included in currently available COPD cost-effectiveness models to see whether the models are able to evaluate subgroups of COPD patients that are considered clinically relevant. Second, we investigated the impact of patient characteristics on the outcomes. Finally, we validated the outcomes of specific subgroup analyses with the models against subgroup results of clinical trials to assess whether the models are suitable for performing subgroup analyses. These questions are relevant because most of the models are initially built to evaluate treatment options for a large group of COPD patients.

Methods

Procedure

In March 2014 the COPD modelling groups that participated in previous meetings as well as new groups identified through publications or other participants within the network were contacted to

explore their interest in participation in the third COPD modelling meeting and in running the model simulations required for this heterogeneity analysis. Modelling groups were first asked to specify which patient characteristics are currently included in their model and which input parameters are specified by subgroup. Second, they were asked to run their model for hypothetical patients that differed in terms of patient and disease characteristics to explore the impact of these characteristics on the outcomes of the models. Third, they were asked to simulate the outcomes for subgroups of patients that had similar characteristics as the subgroups of two large randomized controlled clinical trials to validate the model outcomes. Results were returned to the organizers of the meeting in a structured format in Microsoft Excel two weeks before the meeting. The combined results of the models were circulated to all participants of the COPD modelling meeting one week before the meeting to give participants time to reflect on the outcomes. The results were presented during the meeting and explanations for the differences in outcomes between the models were discussed.

Participating models

Nine different COPD models participated in the model simulations. Six of these nine models also participated in the previous meeting in 2012 [12-17]. A short description of these six models can be found in the publication about the proceedings of this second COPD modelling meeting [11]. Three new models also participated. The simulation model of Asukai et al was published in 2013 [18]. The other two models have not yet been published, but have been presented at international conferences. The model of Briggs et al (GALAXY COPD model) was presented at the ISPOR Annual International meeting 2013 and the ISPOR 17th Annual European Congress 2014 [19, 20]. The model of Dal Negro et al was also presented at the ISPOR 17th annual European Congress 2014 [21].

Content of the modelling challenge

For the first part of the modelling challenge groups reported all patient characteristics that are currently included in their models. Furthermore, groups provided the level of detail in which the following input parameters had been specified in the models: disease progression, exacerbation frequency, mortality, case-fatality of an exacerbation, utilities during stable disease, utilities during exacerbations, maintenance costs and exacerbation-related costs.

In part two, modelling groups simulated the outcomes for hypothetical subgroups of patients that differ in terms of gender, age, smoking status and level of $FEV_1\%$ predicted to see how patient heterogeneity affected effects and costs within one model and between models. Gender, age, smoking status and $FEV_1\%$ predicted were chosen because these are the factors that are included in most models (Table 2). In the first simulation, outcomes for a 65-year old, ex-smoking, male patient with severe COPD were calculated. In each of the following four simulations one of the patient characteristics was changed: female instead of male patient, 75-year old instead of 65-year old patient, smoking instead of ex-smoking patient and patient with moderate COPD instead of severe COPD. The outcomes of each of these four simulations were compared with the results of the first simulation.

In part three, the validity of the subgroup analyses with the models was investigated by comparing the model outcomes against published subgroup results of two large randomized controlled clinical trials, the four-year UPLIFT trial and the three-year TORCH trial [22, 23]). In the UPLIFT trial almost 6000 patients with moderate-to-very severe COPD were randomized to placebo, defined as all regular respiratory medication except inhaled anticholinergics, or tiotropium 18 µg plus all regular respiratory medication except other inhaled anticholinergics [22]. In the TORCH trial slightly more than 6000 patients with moderate-to-very severe COPD were randomized to placebo, defined as all COPD medications except for long-acting bronchodilators and inhaled corticosteroids, salmeterol 50µg, fluticasone 500 µg or the combination of salmeterol 50 µg and fluticasone 500 µg [23]. For both the UPLIFT and the TORCH trial several subgroup analyses have been published mainly by gender, age, smoking status and level of $FEV_1\%$ predicted [24-27]. The modelling groups were asked to adjust the starting population of the model to the baseline characteristics of the specific subgroups within the placebo group of the trial in terms of percentage of males, mean age, percentage current smokers and mean $FEV_1\%$ predicted (or distribution over the GOLD severity stages moderate, severe or very severe COPD). The baseline characteristics of the subgroups in the trial are given in Table 1. For comparison with the UPLIFT placebo arm, groups were asked to run the simulations assuming that patients received all regular respiratory medication except for inhaled anticholinergics. Simulation of the TORCH placebo arm was done assuming that patients did not use long-acting bronchodilators and/or inhaled corticosteroids.

Outcomes

For part two of the challenge, the following outcomes were reported: disease progression defined as decline in lung function, mortality, quality-adjusted life years per patient, and total costs per patient. These outcomes were calculated for both one-year and lifetime horizons. For part three of the challenge, the total number of exacerbations per patient-year and all-cause mortality were calculated including the uncertainty around the outcomes if possible. The time horizon of the model simulations was similar to the duration of the clinical trials: three years for the TORCH trial and four years for the UPLIFT trial.

Results

Table 2 shows the patient characteristics included in the nine participating cost-effectiveness models. The majority of models include gender, age, smoking status and FEV_1% predicted. Newer models also included patient characteristics such as previous exacerbations, BMI and co-morbidities (Asukai_simulation model, Briggs). Table 3 shows the level of specification for the most important input parameters. The level of detail in which the input parameters are specified varies greatly between the models. Table 4 shows the results for the model simulations of hypothetical patients that differ in terms of gender, age, smoking status and FEV_1% predicted. The validation of subgroup analysis in the models against the outcomes of the subgroup analyses in the clinical trials is presented in figures 1 to 4. Below, the results will be discussed separately for the four patient characteristics that most of the models have in common: gender, age, smoking status and FEV_1% predicted. For each characteristic we first present the results for the simulations of hypothetical patients followed by the simulations of the subgroups in the clinical trials.

Gender

Seven models included gender as patient characteristic (Asukai_Markov, Asukai_simulation, Briggs, Dal Negro, Hoogendoorn, Samyshkin and Wacker). Gender was mainly included to specify disease progression (six models) and mortality (six models) (Asukai_Markov, Asukai_simulation, Briggs, Dal Negro, Hoogendoorn, Samyshkin) (Table 3). Results of the comparison of hypothetical subgroups of

8

patients showed however, that three of the models that included gender reported a disease progression rate that was similar for males and females (Asukai_Markov, Dal Negro, Hoogendoorn) (Table 4), while three models had faster disease progression rates for females (Asukai_simulation, Briggs, Samyshkin). Five models had a higher one-year mortality rate for male patients compared to female patients (Asukai_markov, Asukai_simulation, Brigs, Hoogendoorn, Wacker), while in one model mortality in males and females was comparable (Samyshkin). A few models specified utilities during stable disease and/or COPD-related maintenance costs by gender (Table 3) (Asukai_simulation, Briggs, Hoogendoorn). The simulation model of Asukai et al specifying utility weights by gender reported a lower number of QALYs for a female compared to a male patient, despite a lower female mortality, while the model of Briggs et al had a slightly higher number of QALYs for females. Of the two models specifying maintenance costs by gender one used higher costs for females (Hoogendoorn), while the other model (Briggs) included lower costs for females compared to males. Validation of the model results with respect to gender against empirical data showed that in the UPLIFT study male patients had less exacerbations (RR=0.89 (95% CI: 0.81; 0.98)) compared to female patients, while this was only found in the model of Briggs (Figure 1A). The other models reported rates equal for females and males. Four-year all-cause mortality in the UPLIFT trial was reported to be higher for males (RR=1.49 (95% CI: 1.21;1.84)), which was found in two models (Hoogendoorn and Wacker) (Figure 1B).

Age

All nine participating models included age as characteristic mainly to specify disease progression (seven models) and mortality (nine models). In one model, disease progression was lower for a 75-year old compared to a 65-year patient (Asukai_Markov)). In the models of Dal Negro and Hoogendoorn disease progression increased with increasing age, while in the remaining models disease progression was about equal for both ages (Table 4). In all models one-year mortality was higher for a 75-year old compared to a 65-year old (1.5-3.3 times higher). Only one model reported higher QALYs for an older patient (Briggs). Three of the four models that specified costs by age (Briggs, Dal Negro, Hoogendoorn and Wacker) reported higher costs for an older patient (Briggs, Dal Negro and Hoogendoorn). Validation of the model results with respect to age against the outcomes of the UPLIFT study showed that in all models except one, the exacerbation rate in the subgroup of

patients younger than 50 years was equal to the rate in the total population (mean age 65 years), which was in accordance with the trial results (Figure 2A). Only in the model of Briggs the subgroup of patients below 50 years of age had a significantly lower exacerbation rate compared to the total population. In the subgroup of younger patients in the UPLIFT study four-year all-cause mortality was significantly lower compared to the total population (RR=0.43 (95% CI: 0.25; 0.74), which was also found in four models (Briggs, Hansen, Hoogendoorn, Wacker) (Figure 2B).

Smoking status

Six of the nine models included smoking status as patient characteristic (Asukai_Markov, Asukai_simulation, Briggs, Hansen, Hoogendoorn and Wacker). In all six models disease progression was specified by smoking status (Table 3) as smokers having a higher disease progression compared to ex-smokers (Table 4). Three models specified mortality by smoking status (Briggs, Hoogendoorn, and Wacker). One-year mortality in these models was higher for smokers (Table 4). Only in the model of Briggs et al smoking status was modelled to have an association with other input parameters such as utilities and costs (Table 3).. As a result, the one-year QALYs and costs were (almost) equal for a smoking compared to an ex-smoking patient in the models (Table 4). The UPLIFT study showed no difference in exacerbations between smokers and ex-smokers, which was in line with the results of most models, except for two (Briggs and Wacker) that reported a higher exacerbation rate for ex-smokers. Four-year mortality in the UPLIFT trial was also not found to be different between smokers and ex-smokers. Three models however, reported a significantly higher mortality rate for ex-smokers compared to smokers (Hansen, Hoogendoorn and Wacker), most likely because ex-smokers in the UPLIFT trial were older than the smokers and not because of their difference in smoking status. This is confirmed by Table 4, which isolates the impact of smoking status from the impact of age.

FEV$_1$% predicted

All nine models included FEV$_1$% predicted. In all models except the Markov model of Asukai disease progression was specified by FEV$_1$% predicted. All models used FEV$_1$% predicted to specify exacerbation frequency, mortality, utilities and maintenance costs. In all nine models one-year mortality and costs were lower for a patient with moderate COPD (FEV$_1$% predicted about 65%) compared to a patient with severe COPD (FEV$_1$% predicted about40%), while the number of QALYs

was higher in all models (Table 4). Five models also specified utilities during exacerbations by $FEV_1\%$ predicted (Briggs, Borg, Dal Negro, Hansen, Wacker). Exacerbation costs were specified by $FEV_1\%$ predicted by four models (Briggs, Dal Negro, Hansen, Wacker) (Table 3).

Validation of the model results with respect to $FEV_1\%$ predicted against the outcomes of the three-year TORCH study showed that in three models (Borg, Briggs and Wacker) the exacerbation rate in patients with severe COPD was significantly higher than in patients with moderate COPD (Figure 4A) which was in line with the trial results. In the TORCH trial, mortality in patients with severe COPD was higher than in patients with moderate COPD (RR=1.33 ((95% CI:1.09;1.63). Four models reported a higher mortality rate in severe COPD (Borg, Hansen, Hoogendoorn and Wacker).

Discussion

By comparing patient heterogeneity in currently available COPD cost-effectiveness models this study aimed to assess how suitable current models are to evaluate the cost-effectiveness of personalized treatment options for COPD. The patient characteristics that almost all nine models included were gender, age, smoking status and $FEV_1\%$ predicted.

Results with respect to gender showed that the empirical data supported a difference in exacerbations between female and male patients. Several studies in the literature confirm that female patients have a higher exacerbation rate compared to males [28-31]. Only one of the current COPD models included this relationship. Two models reported a significantly higher mortality rate in male patients, which was also found in the UPLIFT trial. However, it should be noted that in the trial female patients were on average three years younger and had a slightly higher $FEV_1\%$ predicted, but they were also more likely to smoke. The difference in mortality is less clear from the literature. Some studies reported lower mortality rates for females [32-34], while others found no difference [35-38]. Only two models specified quality of life by gender, while several studies found a lower quality of life for female patients [39-41].

All models included a strong association between age and mortality, which was in line with studies found in the literature [32, 33, 35-38]. Validation against trial results also showed a difference in

mortality between young and old patients for the trial as well as most models. The majority of models did not specify quality of life by age, while several studies have shown that older COPD patients have lower quality of life [39, 42]. Three models specified COPD-related costs by age. Several studies investigated the association between age and COPD-related costs and found no significant impact of age on costs [43-46]. This may be because age and $FEV_1\%$ predicted are correlated and in three of these studies $FEV_1\%$ was also included in the multivariate model and found to be a significant predictor.

The impact of smoking on disease progression was included in six of the models. Smokers were modelled to have a higher disease progression compared to ex-smokers, which was in line with the literature [47-49]. Three models included a higher all-cause mortality rate for smokers compared to ex-smokers in their input data (Table 4). No difference in mortality rate between smokers and ex-smokers was found in the UPLIFT trial. Validation of model outcomes against trial results showed that three models reported a higher mortality rate for ex-smokers. This was probably because ex-smokers in the trial were on average five years older than smokers and had a lower $FEV_1\%$ predicted. Of seventeen studies on predictors of mortality found in the literature four studies found a higher mortality rate for smokers, six studies found no association and ten studies did not investigate this association.

With respect to $FEV_1\%$ predicted the models are in line with the empirical data and the literature. All models include an association between $FEV_1\%$ predicted and exacerbation frequency, quality of life, mortality and costs. In some models the differences in exacerbations and mortality are however very small and not significant due to the large uncertainty around the estimates. The literature also showed that disease progression in terms of lung function decline seems to decrease when the FEV_1 decreases [47, 48, 50]. This association was not included in all models, because five of the models used the Lung Health Study as (one of the) data source (s) for disease progression, which is a study in mild-to-moderate COPD patients showing the opposite relation (decline increases when FEV_1 decreases) [49]. It is not surprising that the model performance with respect to patient heterogeneity was best for $FEV_1\%$ predicted, because in almost all models the $FEV_1\%$ predicted is the key parameter. Five out of the nine models are cohort Markov models with Markov states defined

according to the 2007 GOLD classification, i.e. the severity of COPD was defined by the $FEV_1\%$ predicted.

Outcomes of the different models were compared to published subgroup results of two large trials: the UPLIFT and TORCH trial. Some of the COPD models were built to extrapolate the results of clinical trials (Asukai-Markov, Samyshkin) and were populated with input data from these trials, while other models are population models using a wide range of different data sources as input (Hoogendoorn, Hansen and Wacker). For the latter models comparison with trial results may therefore be less valid because it is well known that patients included in clinical trials are a subgroup of the total COPD population. This was, for example, shown in a study of Kruis et al that found that patients participating in large clinical trials sponsored by pharmaceutical companies are on average younger, more likely to be male, have a lower $FEV_1\%$ predicted and tend to have more exacerbations compared to patients in primary care [51]. In the current analysis this selection bias issue was most likely the explanation that no difference in mortality was found between smokers and ex-smokers in the UPLIFT trial. Patients who continue to smoke are probably less likely to be included in clinical studies, because they progress faster to more severe stages, have more co-morbidities and are more likely to die due to other smoking-related diseases. On the other hand, among the patients who continue to smoke despite having COPD there may be relatively more patients that are less susceptible to smoking-related diseases. The true impact of trial populations being a selective group of patients on the outcomes of this study is difficult to assess. In the current study we mainly focused on the differences in outcomes between subgroups within one model or trial and therefore the impact might be limited.

The results of this study showed that all currently available models are capable of running simulations for different age- and COPD severity classes. The majority of the models also have the ability to run analyses separately for males and females and for smokers and ex-smokers. The validity of these subgroup analyses within the models is questionable, because important input parameters have not been specified by gender, age or smoking status. For example, the specification of higher exacerbations for females or lower quality of life with older age, is not included in the majority of models. Most models are developed to evaluate treatment options for the total general COPD population or the average COPD trial population.

Although it is not unlikely that future treatment will be increasingly tailored to age, gender and smoking status, treatment is more likely to be personalized on the basis of clinical parameters especially when considering ethical debates and societal preferences. Exacerbations for example are highly associated with the number of previous exacerbations [29-31, 52]. Mortality in COPD is associated with age and FEV_1, but also with co-morbidities, BMI, dyspnea and several other clinical parameters [32, 34, 36, 38, 53]. BMI also seems to be associated with quality of life in COPD [39, 54] and co-morbidities may affect COPD-related costs [44, 45, 55, 56]. Only the more recently developed models of Asukai (simulation) and Briggs included these types of clinical parameters and are therefore more suitable to evaluate personalized treatment that the other models . Recent reviews on phenotyping in COPD showed that currently about four different phenotypes have been defined in COPD that may require a different treatment strategy: emphysema, COPD with chronic bronchitis, COPD combined with asthma (ACOS) and COPD with frequent exacerbations [57-60]. Information on effectiveness and cost-effectiveness of treatment options for these subgroups is needed to guide clinical guideline development and decisions for reimbursement. Three of the nine models included information on one of the phenotypes, i.e. COPD with frequent exacerbations, but none of the current models is able to evaluate treatment options for the other three phenotypes. Future models should include all clinical patient characteristics currently considered to influence disease severity, prognosis and treatment response in COPD.

The current study had some limitations. Firstly, not all models could perform uncertainty analysis around the results for exacerbations and all-cause mortality. Therefore, it was not always possible to determine whether predicted differences between subgroups are important within the range of uncertainty estimated. The information in the published papers of the two clinical trials was also sometimes not sufficient to calculate whether the difference between subgroups was statistically significant or not. Results of the trials by smoking status and by COPD severity did not present standard deviations or standard errors around the mean exacerbation rate [24, 27]. The exacerbation rates in the UPLIFT trial in smokers and ex-smokers were, however, almost equal, 0.77 versus 0.83, a difference that is most likely not significant [27]. By contrast, the difference in exacerbation rate in the TORCH trial between people with severe and moderate COPD was large, 1.24 versus 0.82 and therefore this difference seems significant given the population size [24].

A second limitation was that the starting population of the models was only adjusted with respect to gender distribution, age, smoking distribution and $FEV_1\%$predicted to mirror the population of either the TORCH or the UPLIFT trial. The most recent models include much more characteristics (e.g. BMI, co-morbidities) that could have been adjusted, to create model-populations that were even more similar to the trial-populations. However, this would have made a comparison with the other models impossible.

A third limitation of the current study was that we only assessed whether there were significant differences in outcomes between the subgroups simulated with the models and whether that result was in line with the trial results and the literature. We did not focus on the results of the models in absolute terms. However, the results of the different models showed substantial variation, especially for mortality. A more detailed external validation of model results against the outcomes of the trials for the total population will be topic of another paper.

Finally although this study aimed to compare cost-effectiveness models, it mainly focused on the comparison of clinical outcomes, not on costs. Only table 4 presents information on costs. Validation of costs was not possible because the trials did not present costs specified by subgroups. Clinical outcomes are important outcomes for cost-effectiveness models as well, because they have a high impact on quality of life and costs.

In conclusion, this study showed that the majority of currently available COPD cost-effectiveness models include the relevant patient characteristics gender, age, smoking status and $FEV_1\%$ predicted and are therefore able to evaluate the cost-effectiveness of personalized treatment based on these parameters. Most models however, do not include all important associations between these characteristics and input parameters. Furthermore, treatment in COPD is more likely to be personalized on the basis of clinical parameters. Two models also included several clinical patient characteristics, such as previous exacerbations, BMI and co-morbidities and seem to be more suitable to evaluate personalized treatment. Inclusion of other clinical parameters, such as emphysema, chronic bronchitis and co-existence of asthma is relevant to make the models suitable to evaluate the cost-effectiveness of treatment options for currently defined phenotypes in COPD.

References

1. O'Donnell JC. Personalized medicine and the role of health economics and outcomes research: issues, applications, emerging trends, and future research. Value Health 2013; 16: S1-3.

2. Redekop WK, Mladsi D. The faces of personalized medicine: a framework for understanding its meaning and scope. Value Health 2013; 16: S4-9.

3. Personalized Medicine. U.S. Food and Drug Administration. Available at: http://www.fda.gov/scienceresearch/specialtopics/personalizedmedicine/default.htm (accessed June 17th 2015)

4. Hatz MH, Schremser K, Rogowski WH. Is individualized medicine more cost-effective? A systematic review. Pharmacoeconomics 2014; 32(5):443-55.

5. Phillips KA, Ann Sakowski J, Trosman J, Douglas MP, Liang SY, Neumann P. The economic value of personalized medicine tests: what we know and what we need to know. Genet Med. 2014; 16(3):251-7.

6. Ramaekers BL, Joore MA, Grutters JP. How should we deal with patient heterogeneity in economic evaluation: a systematic review of national pharmacoeconomic guidelines. Value Health 2013; 16: 855-862.

7. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. http://www.goldcopd.com (accessed June 2013), 2013.

8. Han MK, Agusti A, Calverley PM, Celli BR, Criner G, Curtis JL, et al. Chronic obstructive pulmonary disease phenotypes: the future of COPD. Am J Respir Crit Care Med 2010; 182: 598-604.

9. Woodruff PG, Agusti A, Roche N, Singh D, Martinez FJ. Current concepts in targeting chronic obstructive pulmonary disease pharmacotherapy: making progress towards personalised management. Lancet. 2015 May 2;385(9979):1789-98.

10. Medicines in Development for Chronic Obstructive Pulmonary Disease (COPD). 2012 The Pharmaceutical Research and Manufacturers of America (PhRMA). Available at: http://www.phrma.org/sites/default/files/pdf/copd2012.pdf (accessed June17th 2015)

11. Hoogendoorn M, Feenstra TL, Asukai Y, Borg S, Hansen RN, Jansson SA, Samyshkin

Y, Wacker M, Briggs AH, Lloyd A, Sullivan SD, Rutten-van Mölken MP. Cost-effectiveness models for chronic obstructive pulmonary disease: cross-model comparison of hypothetical treatment scenarios. Value Health. 2014 Jul;17(5):525-36.

12. Price D, Gray A, Gale R, Asukai Y, Mungapen L, Lloyd A, et al. Cost-utility analysis of indacaterol in Germany: A once-daily maintenance bronchodilator for patients with COPD. Respir Med 2011.

13. Borg S, Ericsson A, Wedzicha J, Gulsvik A, Lundback B, Donaldson GC, et al. A computer simulation model of the natural history and economic impact of chronic obstructive pulmonary disease. Value Health 2004; 7: 153-67.

14. Hansen RN, Xu X, Sullivan SD. A dynamic cohort model of chronic obstructive pulmonary disease and its treatments. Value Health 2012; 15: A54.

15. Hoogendoorn M, Rutten-van Molken MP, Hoogenveen RT, Al MJ, Feenstra TL. Developing and applying a stochastic dynamic population model for chronic obstructive pulmonary disease. Value Health 2011; 14: 1039-1047.

16. Samyshkin Y, Kotchie RW, Mork AC, Briggs AH, Bateman ED. Cost-effectiveness of roflumilast as an add-on treatment to long-acting bronchodilators in the treatment of COPD associated with chronic bronchitis in the United Kingdom. Eur J Health Econ 2013.

17. Menn P, Leidl R, Holle R. A lifetime Markov model for the economic evaluation of chronic obstructive pulmonary disease. Pharmacoeconomics 2012; 30: 825-840.

18. Asukai Y, Baldwin M, Fonseca T, Gray A, Mungapen L, Price D. Improving clinical reality in chronic obstructive pulmonary disease economic modelling: development and validation of a micro-simulation approach. Pharmacoeconomics. 2013 Feb;31(2):151-61.

19. Briggs A, Lomas D, Rutten van-Molken M, Tabberer M, Chambers M, Muellerova H, Locantore N, Baker TM, Risebrough NA, Exuzides A, Colby C, Gonzalez McQuire S. Developing a new model of COPD: from conceptualization to implementation to validation. Value in Health 2013; 16(3): A234.

20. Risebrough NA, Briggs A, Baker TM, Exuzides A, Colby C, Rutten van-Molken M, Gonzalez McQuire S, Lomas D, Muellerova H, Tal-Singer R, Ismaila A. Validating a model to predict disease progression outcomes in patients with COPD. Value in Health 2014; 17(7): A560–A561.

21. Dal Negro RW. Current annual cost calculation is the best predictor of mortality at three years in COPD. Value in Health 2014; 17(7): A590-591

22. Tashkin DP, Celli B, Senn S, Burkhart D, Kesten S, Menjoge S, Decramer M; UPLIFT Study Investigators. A 4-year trial of tiotropium in chronic obstructive pulmonary disease. N Engl J Med. 2008 Oct 9;359(15):1543-54.

23. Calverley PM, Anderson JA, Celli B, Ferguson GT, Jenkins C, Jones PW, Yates JC, Vestbo J; TORCH investigators. Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. N Engl J Med. 2007 Feb 22;356(8):775-89.

24. Jenkins CR, Jones PW, Calverley PM, Celli B, Anderson JA, Ferguson GT, Yates JC, Willits LR, Vestbo J. Efficacy of salmeterol/fluticasone propionate by GOLD stage of chronic obstructive pulmonary disease: analysis from the randomised, placebo-controlled TORCH study. Respir Res. 2009 Jun 30;10:59.

25. Tashkin D, Celli B, Kesten S, Lystig T, Decramer M. Effect of tiotropium in men and women with COPD: results of the 4-year UPLIFT trial. Respir Med. 2010 Oct;104(10):1495-504.

26. Morice AH, Celli B, Kesten S, Lystig T, Tashkin D, Decramer M. COPD in young patients: a pre-specified analysis of the four-year trial of tiotropium (UPLIFT). Respir Med. 2010 Nov;104(11):1659-67.

27. Tashkin DP, Celli B, Kesten S, Lystig T, Mehra S, Decramer M. Long-term efficacy of tiotropium in relation to smoking status in the UPLIFT trial. Eur Respir J. 2010 Feb;35(2):287-94

28. McGarvey L, Lee AJ, Roberts J, Gruffydd-Jones K, McKnight E, Haughney J. Characterisation of the frequent exacerbator phenotype in COPD patients in a large UK primary care population. Respir Med. 2015 Feb;109(2):228-37.

29. Müllerová H, Shukla A, Hawkins A, Quint J. Risk factors for acute exacerbations of COPD in a primary care population: a retrospective observational cohort study. BMJ Open. 2014 Dec 18;4(12):e006171.

30. Husebø GR, Bakke PS, Aanerud M, Hardie JA, Ueland T, Grønseth R, Persson LJ, Aukrust P, Eagan TM. Predictors of exacerbations in chronic obstructive pulmonary disease--results from the Bergen COPD cohort study. PLoS One. 2014 Oct 3;9(10):e109721.

31. Beeh KM, Glaab T, Stowasser S, Schmidt H, Fabbri LM, Rabe KF, Vogelmeier CF. Characterisation of exacerbation risk and exacerbator phenotypes in the POET-COPD trial. Respir Res. 2013 Oct 29;14:116.

32. Montserrat-Capdevila J, Godoy P, Marsal JR, Barbé-Illa F. [Risk factors for mortality in patients with chronic obstructive pulmonary disease]. Aten Primaria.2015 Mar 13. pii: S0212-6567(14)00392-8.

33. Schembri S, Anderson W, Morant S, Winter J, Thompson P, Pettitt D, MacDonald TM, Winter JH. A predictive model of hospitalisation and death from chronic obstructive pulmonary disease. Respir Med. 2009 Oct;103(10):1461-7.

34. de Torres JP, Cote CG, López MV, Casanova C, Díaz O, Marin JM, Pinto-Plata V, de Oca MM, Nekach H, Dordelly LJ, Aguirre-Jaime A, Celli BR. Sex differences in mortality in patients with COPD. Eur Respir J. 2009 Mar;33(3):528-35.

35. Patel MS, Natanek SA, Stratakos G, Pascual S, Martínez-Llorens J, Disano L, Terzis G, Hopkinson NS, Gea J, Vogiatzis I, Maltais F, Polkey MI. Vastus lateralis fiber shift is an independent predictor of mortality in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2014 Aug 1;190(3):350-2.

36. Stolz D, Meyer A, Rakic J, Boeck L, Scherr A, Tamm M. Mortality risk prediction in COPD by a prognostic biomarker panel. Eur Respir J. 2014 Dec;44(6):1557-70.

37. de Voogd JN, Wempe JB, Postema K, van Sonderen E, Ranchor AV, Coyne JC, Sanderman R. More evidence that depressive symptoms predict mortality in COPD patients: is type D personality an alternative explanation? Ann Behav Med. 2009 Oct;38(2):86-93.

38. Briggs A, Spencer M, Wang H, Mannino D, Sin DD. Development and validation of a prognostic index for health outcomes in chronic obstructive pulmonary disease. Arch Intern Med. 2008 Jan 14;168(1):71-9.

39. Tsiligianni I, Kocks J, Tzanakis N, Siafakas N, van der Molen T. Factors that influence disease-specific quality of life or health status in patients with COPD: a review and meta-analysis of Pearson correlations. Prim Care Respir J. 2011 Sep;20(3):257-68.

40. Fishwick D, Lewis L, Darby A, Young C, Wiggans R, Waterhouse J, Wight J, Blanc PD. Determinants of health-related quality of life among residents with and without COPD in a historically industrialised area. Int Arch Occup Environ Health. 2014 Dec 14. [Epub ahead of print]

41. Rutten-van Mölken MP, Oostenbrink JB, Tashkin DP, Burkhart D, Monz BU. Does quality of life of COPD patients as measured by the generic EuroQol five-dimension questionnaire differentiate between COPD severity stages? Chest. 2006 Oct;130(4):1117-28.

42. Punekar YS, Rodriguez-Roisin R, Sculpher M, Jones P, Spencer M. Implications of chronic obstructive pulmonary disease (COPD) on patients' health status: a western view. Respir Med. 2007 Mar;101(3):661-9.

43. Darnell K, Dwivedi AK, Weng Z, Panos RJ. Disproportionate utilization of healthcare resources among veterans with COPD: a retrospective analysis of factors associated with COPD healthcare cost. Cost Eff Resour Alloc. 2013 Jun 13;11:13.

44. Mapel DW, McMillan GP, Frost FJ, Hurley JS, Picchi MA, Lydick E, Spencer MD. Predicting the costs of managing patients with chronic obstructive pulmonary disease. Respir Med. 2005 Oct;99(10):1325-33.

45. Miravitlles M, Murio C, Guerrero T, Gisbert R. Costs of chronic bronchitis and COPD: a 1-year follow-up study. Chest. 2003 Mar;123(3):784-91.

46. Jansson SA, Andersson F, Borg S, Ericsson A, Jönsson E, Lundbäck B. Costs of COPD in Sweden according to disease severity. Chest. 2002 Dec;122(6):1994-2002.

47. Vestbo J, Edwards LD, Scanlon PD, Yates JC, Agusti A, Bakke P, Calverley PM, Celli B, Coxson HO, Crim C, Lomas DA, MacNee W, Miller BE, Silverman EK, Tal-Singer R, Wouters E, Rennard SI; ECLIPSE Investigators. Changes in forced expiratory volume in 1 second over time in COPD. N Engl J Med. 2011 Sep 29;365(13):1184-92.

48. Celli BR, Thomas NE, Anderson JA, Ferguson GT, Jenkins CR, Jones PW, Vestbo J, Knobil K, Yates JC, Calverley PM. Effect of pharmacotherapy on rate of decline of lung function in chronic obstructive pulmonary disease: results from the TORCH study. Am J Respir Crit Care Med. 2008 Aug 15;178(4):332-8.

49. Scanlon PD, Connett JE, Waller LA, Altose MD, Bailey WC, Buist AS, Tashkin DP; Lung Health Study Research Group. Smoking cessation and lung function in mild-to-moderate chronic obstructive pulmonary disease. The Lung Health Study. Am J Respir Crit Care Med. 2000 Feb;161(2 Pt 1):381-90.

50. Casanova C, de Torres JP, Aguirre-Jaíme A, Pinto-Plata V, Marin JM, Cordoba E, Baz R, Cote C, Celli BR. The progression of chronic obstructive pulmonary disease is heterogeneous: the experience of the BODE cohort. Am J Respir Crit Care Med. 2011 Nov 1;184(9):1015-21.

51. Kruis AL, Ställberg B, Jones RC, Tsiligianni IG, Lisspers K, van der Molen T, Kocks JW, Chavannes NH. Primary care COPD patients compared with large pharmaceutically-sponsored COPD studies: an UNLOCK validation study. PLoS One.2014 Mar 5;9(3):e90145.

52. Make BJ, Eriksson G, Calverley PM, Jenkins CR, Postma DS, Peterson S, Östlund O, Anzueto A. A score to predict short-term risk of COPD exacerbations (SCOPEX). Int J Chron Obstruct Pulmon Dis. 2015 Jan 27;10:201-9.

53. Schols AM, Slangen J, Volovics L, Wouters EF. Weight loss is a reversible factor in the prognosis of chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 1998 Jun;157(6 Pt 1):1791-7.

54. Villar Balboa I, Carrillo Muñoz R, Regí Bosque M, Marzo Castillejo M, Arcusa Villacampa N, Segundo Yagüe M. [Factors associated with the quality of life in patients with chronic obstructive pulmonary disease]. Aten Primaria. 2014 Apr;46(4):179-87.

55. Huber M, Wacker M, Vogelmeier CF, Leidl R (2015) Excess costs of comorbidities in chronic obstructive pulmonary disease: A systematic review. PLOS ONE 10(4): e0123292.

56. Huber M, Wacker M, Vogelmeier CF, Leidl R (2015) Comorbid influences on generic health-related quality of life in COPD: A systematic review. PLOS ONE (in press).

57. Miravitlles M, Soler-Cataluña JJ, Calle M, Molina J, Almagro P, Quintano JA, Riesco JA, Trigueros JA, Piñera P, Simón A, Rodríguez-Hermosa JL, Marco E, López D, Coll R, Coll-Fernández R, Lobo MÁ, Díez J, Soriano JB, Ancochea J. Spanish guideline for COPD (GesEPOC). Update 2014. Arch Bronconeumol. 2014 Jan;50 Suppl 1:1-16.

58. Segreti A, Stirpe E, Rogliani P, Cazzola M. Defining phenotypes in COPD: an aid to personalized healthcare. Mol Diagn Ther. 2014 Aug;18(4):381-8.

59. Rothe T. [Phenotype specific therapy of COPD]. Praxis (Bern 1994). 2014 Dec 10;103(25):1509-14.

60. Vestbo J. COPD: definition and phenotypes. Clin Chest Med. 2014 Mar;35(1):1-6.

Table 1: Baseline characteristics of subgroups within the placebo group of the UPLIFT and TORCH trial used as starting population of the model simulations [24-27]

Trial	UPLIFT trial					TORCH trial	
Subgroup	By gender		By age	By smoking status		By COPD severity	
	Males	Females	Age <50 years	Smokers	Ex-smokers	Moderate COPD	Severe COPD
N	2222	784	172	432	1791	535	775
Males %	100	0	58	62	78	72	76
Age, mean (SD)	65 (8)	62 (9)	<50	61 (9)	66 (9)	65 (9)	65 (8)
Current smokers %	26	40	57	100	0	47	43
Post-bronchodilator FEV1% predicted, mean (SD)	47 (13)	49 (13)	47 (15)	50 (15)	46 (15)	59 (7)	40 (6)
GOLD II: moderate COPD	44%	48%	48%	54%	41%	100%	0%
GOLD III: severe COPD	45%	43%	36%	38%	48%	0%	100%
GOLD IV: very severe COPD	10%	7%	15%	7%	10%	0%	0%

Table 2: Patient characteristics included in nine COPD cost-effectiveness models

Characteristic:	Asukai_ Markov model	Asukai_ Simulation model	Borg	Briggs	Dal Negro	Hansen	Hoogendoorn	Samyshkin	Wacker
Gender	X	X		X	X		X	X	X
Age	X	X	X	X	X	X	X	X	X
Smoking	X	X		X		X	X		X
FEV1 % predicted	X	X	X	X	X	X	X	X	X
Height		X							
Previous exacerbations		X	X	X					
BMI		X		X					
Co-morbidities		X		X	X				
Number of ER visits/ hospitalizations		X							
Rapid decline in FEV$_1$			X		X				
Dyspnea (mMRC)				X					
Disease specific quality of life (SGRQ)				X					
Utility (EQ-5D)				X					
Activity (6 MWD)				X					
Residual volume					X				
DLCO					X				
Lung transplantation									X

*FEV1=forced expiratory volume in one second, BMI=body mass index, ER=emergency room, mMRC=modified Medical Research Council dysnpnea scale, SGRQ=St. George's Respiratory Questionnaire, EQ-5D=EuroQol questionnaire with 5 dimensions, 6MWD=six-minute walking distance, RV=residual volume, DLCO= Diffusion Lung Capacity for carbon monoxide

Table 3: Level of specification of important input parameters

Model:	Disease progression specified by:	Exacerbation frequency specified by:	Mortality specified by:	Utilities during stable disease specified by:	Utilities during exacerbations specified by:	Maintenance costs specified by:	Exacerbation-related costs specified by:
				Type of input parameter			
Asukai_Markov model	Gender, age, smoking (indirectly)	FEV_1% pred.	Gender, age, FEV_1% pred.	FEV_1% pred.	Exacerbation severity	FEV_1% pred.	Exacerbation severity
Asukai_Simulation model	Gender, age, smoking, FEV_1% pred.	FEV_1% pred.	Gender, age, FEV_1% pred	Gender, age, FEV_1%pred., BMI, co-morbidities, # or ER visits/ hospitalizations	Exacerbation severity	FEV_1% pred.	Exacerbation severity
Borg	Age, FEV_1% pred., rapid decline	FEV_1% pred., frequent exacerbations	Age, FEV_1% pred.	FEV_1% pred.	Exacerbation severity, FEV_1% pred*	FEV_1% pred.	Exacerbation severity
Briggs	Gender, age, smoking, FEV_1% pred. + all other listed in table 2	Gender, age, smoking, FEV_1% pred. + all other listed in table 2	Gender, age, smoking, FEV_1% pred. + all other listed in table 2	Gender, age, smoking, FEV_1% pred. + all other listed in table 2	Gender, age, smoking, FEV_1% pred. + all other listed in table 2	Gender, age, smoking, FEV_1% pred. + all other listed in table 2	Gender, age, smoking, FEV_1% pred. + all other listed in table 2
Dal Negro	Gender, age, RV, DLCO, BODE, % FEV decline, co-morbidities	Age, RV, DLCO, BODE, % FEV decline, co-morbidities	Age, RV, DLCO, BODE, % FEV decline, co-morbidities	Age, RV, DLCO, BODE, % FEV decline, co-morbidities	Age, FEV_1% pred., RV, DLCO, BODE, % FEV decline, co-morbidities	Age, FEV_1% pred., RV, DLCO, BODE, % FEV decline, co-morbidities	Age, FEV_1% pred., RV, DLCO, BODE, % FEV decline, co-morbidities
Hansen	Smoking, FEV_1% pred.	FEV_1% pred.	Age, FEV_1% pred.	FEV_1% pred.	FEV_1% pred.	FEV_1% pred.	FEV_1% pred.

Hoogendoorn	Gender, age, smoking, FEV$_1$% pred.	FEV$_1$% pred.	Gender, age, smoking, FEV$_1$% pred.	FEV$_1$% pred.	Exacerbation severity	Gender, age, FEV$_1$% pred.	Exacerbation severity
Samyshkin	Gender, age, FEV$_1$% pred.	FEV$_1$% pred.	Gender, age, FEV$_1$% pred.	FEV$_1$% pred.	Exacerbation severity	FEV$_1$% pred.	Exacerbation severity
Wacker	Smoking, FEV$_1$% pred.	FEV$_1$% pred, lung transplant	Gender, age, smoking, FEV$_1$% pred., lung transplant	FEV$_1$% pred., lung transplant	FEV$_1$% pred., Exacerbation severity, lung transplant	Age, FEV$_1$% pred., lung transplant	Age, FEV$_1$% pred., exacerbation severity, lung transplant

* A disutility that depends on exacerbation severity is applied to the utility of stable disease.

Table 4: One-year model outcomes for a hypothetical reference patient (analysis 1) and for patients that differ on onepatient characteristic compared to the reference patient (analysis 2 to 5)

Model:	Asukai_ Markov	Asukai_ Simulation	Borg^	Briggs	Dal Negro	Hansen	Hoogendoorn	Samyshkin	Wacker
1. Male patient, 65 year, ex-smoking with severe COPD (=comparator)									
- Disease progression*	12.8%	0.4%	5.5%	-0.51% pred.	10.0%	5.2%	2.8%	11.9%	2.5%
- Mortality	4.0%	4.7%	7.2%	2.8%	10.0%	3.1%	7.1%	4.9%	12.4%
- QALYs	0.752	0.627	0.688	0.581	n/a	0.658	0.723	0.724	0.520
- Costs (2014 $)	1950	1560	2350	3830	2680-4020	2260	1680	2270	2480
2. **Female** patient, 65 year, ex-smoking with severe COPD									
- Disease progression*	12.8%	0.9%	-	-0.81% pred.	10.0%	-	2.8%	16.7%	2.5%
- Mortality	3.6%	3.3%	-	2.6%	10.0%	-	4.7%	4.9%	7.9%
- QALYs	0.754	0.406	-	0.623	n/a	-	0.723	0.721	0.530
- Costs (2014 $)	1950	1590	-	3090	2680-4020	-	2050	2320	2540
3. Male patient, **75 year**, ex-smoking with severe COPD									
- Disease progression*	11.4%	0.5%	5.6%	-0.57% pred.	20.0%	5.2%	3.6%	11.6%	2.1%
- Mortality	6.1%	13.8%	24.1%	5.5%	20.0%	9.3%	13.2%	12.2%	27.2%
- QALYs	0.742	0.598	0.624	0.593	n/a	0.643	0.723	0.696	0.470
- Costs (2014 $)	1900	1460	2090	4010	4020-5360	2200	2090	2180	2290
4. Male patient, 65 year, **smoking** with severe COPD									
- Disease progression*	12.8%	2.3%	-	-1.27% pred.	-	8.6%	3.6%	-	7.6%
- Mortality	4.0%	4.8%	-	3.6%	-	3.1%	8.1%	-	17.7%
- QALYs	0.752	0.627	-	0.548	-	0.657	0.723	-	0.500
- Costs (2014 $)	1950	1600	-	3840	-	2270	1680	-	2450
5. Male patient, 65 year, smoking with **moderate COPD#**									
- Mortality	2.5%	3.3%	5.3%	1.7%	8.0%	2.5%	5.0%	3.0%	6.3%
- QALYs	0.781	0.653	0.741	0.623	n/a	0.702	0.737	0.772	0.570
- Costs (2014 $)	330	450	1060	3910	2010-2680	1580	1270	1540	1460

*Disease progression was defined as the percentage of severe COPD patients moving to very severe COPD except for the model of Briggs et al in which disease progression was defined as the decline in FEV_1% pred.

a) Total exacerbations per patient year by gender

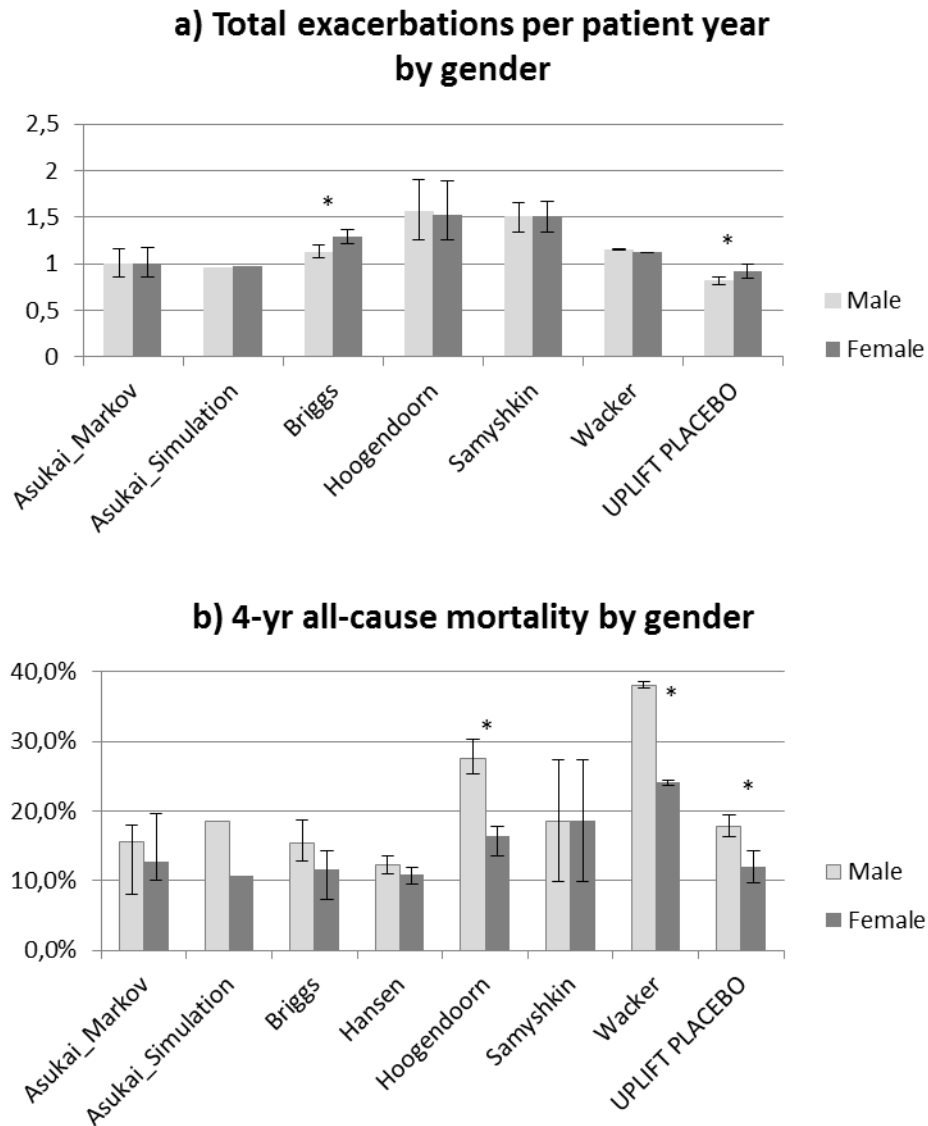## b) 4-yr all-cause mortality by gender

Figure 1: Comparison of model results for the subgroup gender with empirical results of the four-year UPLIFT trial for a) exacerbations and b) all-cause mortality. * indicates that the difference is statistically significant. Difference in the simulation model of Asukai could not be tested because no uncertainty was available around the estimates.
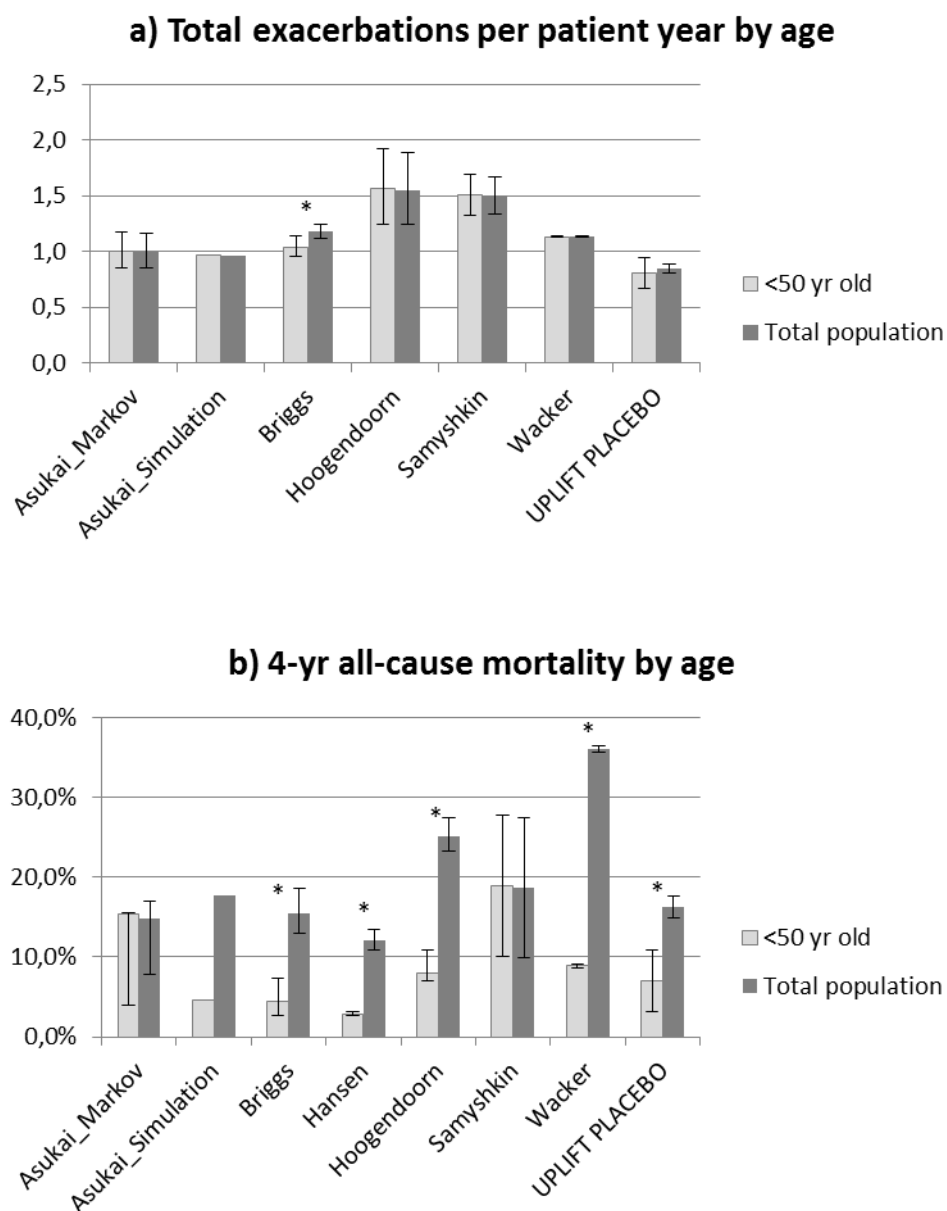
a) Total exacerbations per patient year by age

b) 4-yr all-cause mortality by age

Figure 2: Comparison of model results for the subgroup age with empirical results of the four-year UPLIFT trial for a) exacerbations and b) all-cause mortality. * indicates that the difference is statistically significant. Difference in the simulation model of Asukai could not be tested because no uncertainty was available around the estimates.

a) Total exacerbations per patient year by smoking status

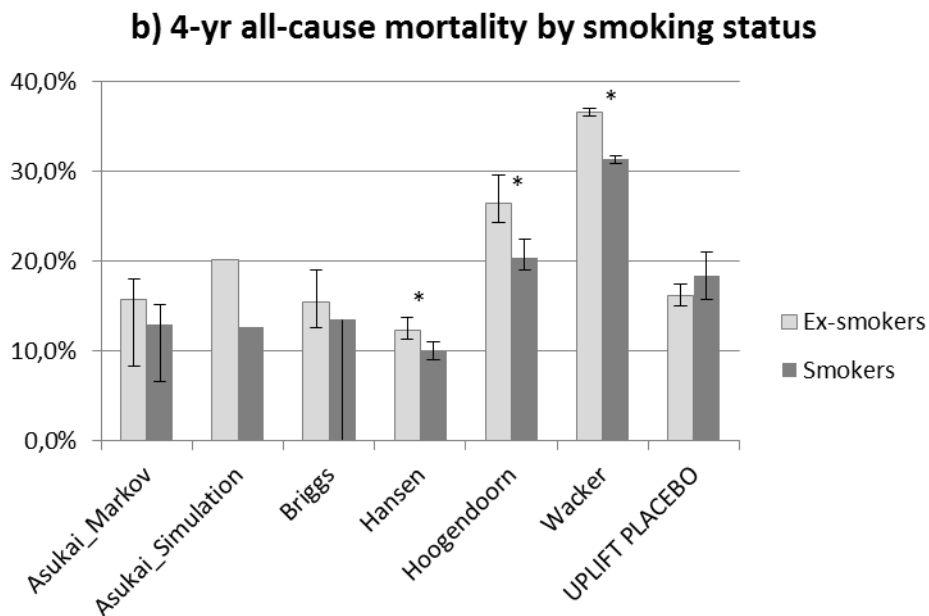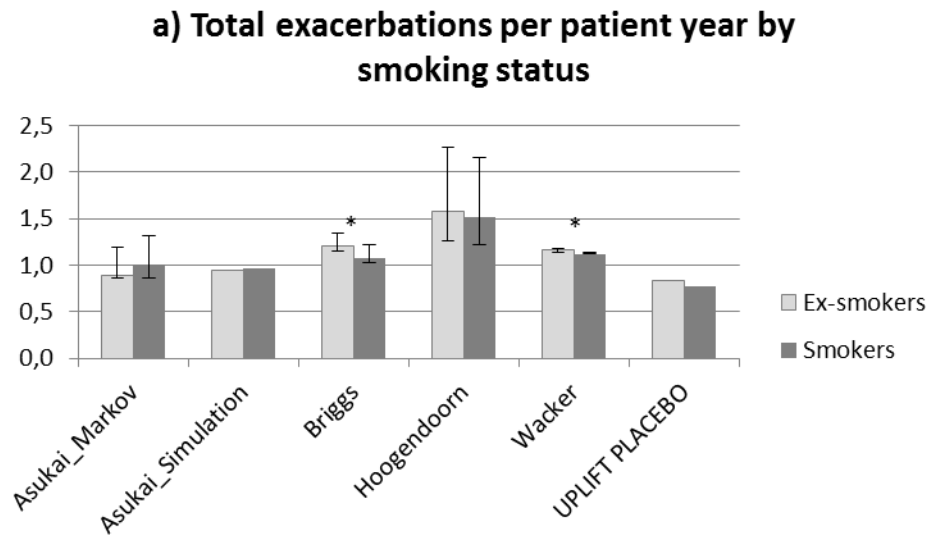b) 4-yr all-cause mortality by smoking status

Figure 3: Comparison of model results for the subgroup smoking status with empirical results of the four-year UPLIFT trial for a) exacerbations and b) all-cause mortality. * indicates that the difference is statistically significant. Difference in the simulation model of Asukai and the trial (exacerbations) could not be tested because no uncertainty was available around the estimates.
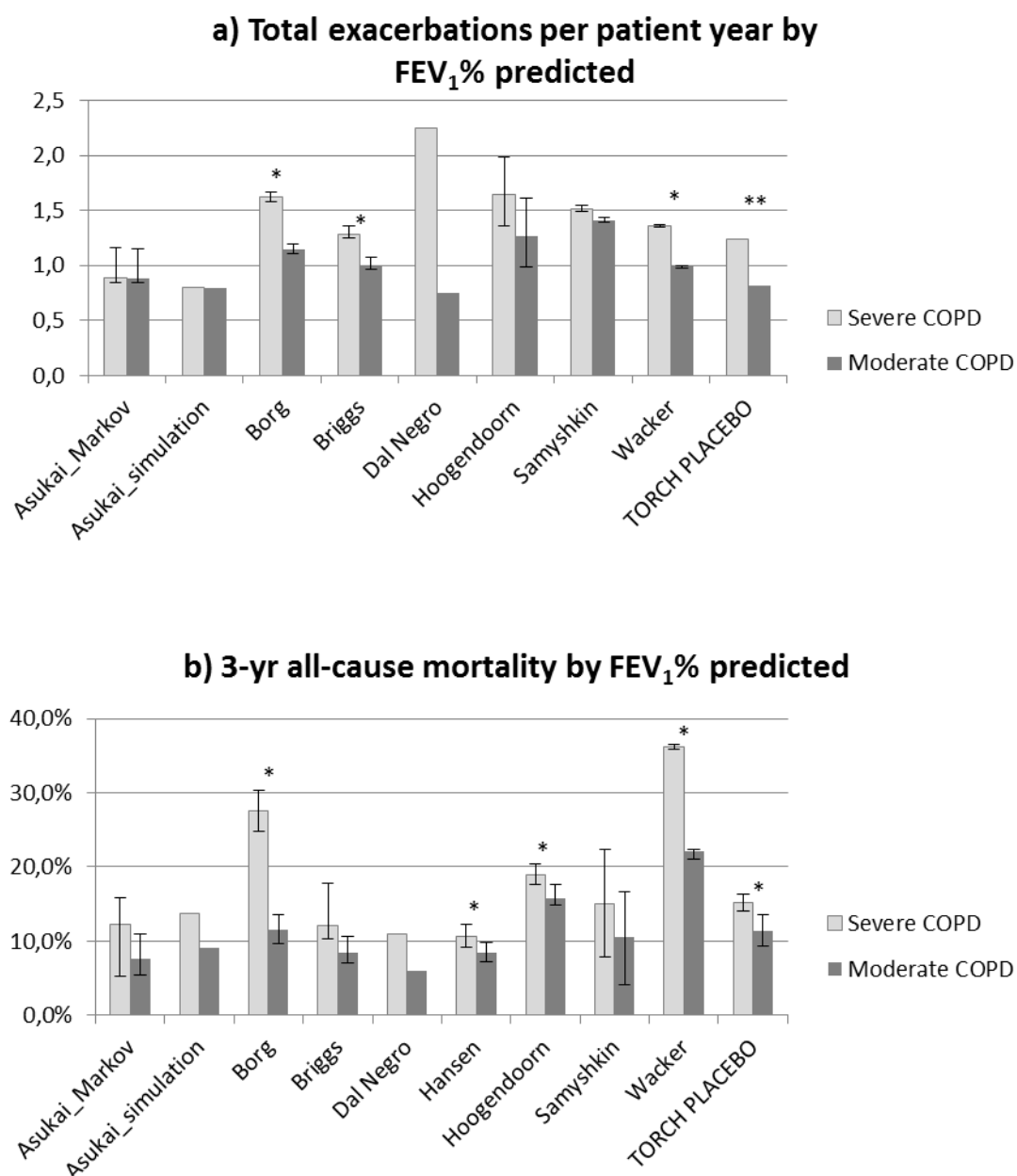
a) Total exacerbations per patient year by FEV$_1$% predicted

b) 3-yr all-cause mortality by FEV$_1$% predicted

Figure 4: Comparison of model results for the subgroup FEV$_1$% predicted with empirical results of the three-year TORCH trial for a) exacerbations and b) all-cause mortality. * indicates that the difference between moderate and severe COPD is statistically significant. ** indicates that the difference is most likely statistically significant, but standard errors were missing. Difference in the simulation model of Asukai and the model of Dal Negro could not be tested because no uncertainty was available around the estimates.