BMJ open accessible medical research

Large-scale international validation of the ADO index in subjects with COPD: an individual subject data analysis of 10 cohorts

Milo A Puhan,[1,2] Nadia N Hansel,[3] Patricia Sobradillo,[4] Paul Enright,[5] Peter Lange,[6] DeMarc Hickson,[7,8] Ana M Menezes,[9] Gerben ter Riet,[10] Ulrike Held,[2] Antonia Domingo-Salvany,[11,15] Zab Mosenifar,[12] Josep M Antó,[11,13,14,15] Karel G M Moons,[16] Alphons Kessels,[17] Judith Garcia-Aymerich,[11,13,14,15] for the International COPD Cohorts Collaboration Working Group

To cite: Puhan MA, Hansel NN, Sobradillo P, *et al*. Large-scale international validation of the ADO index in subjects with COPD: an individual subject data analysis of 10 cohorts. *BMJ Open* 2012;**2**:e002152. doi:10.1136/bmjopen-2012-002152

▸ Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2012-002152).

Received 22 September 2012
Revised 6 November 2012
Accepted 12 November 2012

MAP and JGA contributed equally to this study.

For numbered affiliations see end of article.

Correspondence to
Dr Judith Garcia-Aymerich; jgarcia@creal.cat

ABSTRACT

Background: Little evidence on the validity of simple and widely applicable tools to predict mortality in patients with chronic obstructive pulmonary disease (COPD) exists.

Objective: To conduct a large international study to validate the ADO index that uses age, dyspnoea and FEV_1 to predict 3-year mortality and to update it in order to make prediction of mortality in COPD patients as generalisable as possible.

Design: Individual subject data analysis of 10 European and American cohorts (n=13 914).

Setting: Population-based, primary, secondary and tertiary care.

Patients: COPD GOLD stages I–IV.

Measurements: We validated the original ADO index. We then obtained an updated ADO index in half of our cohorts to improve its predictive accuracy, which in turn was validated comprehensively in the remaining cohorts using discrimination, calibration and decision curve analysis and a number of sensitivity analyses.

Results: 1350 (9.7%) of all subjects with COPD (60% male, mean age 61 years, mean FEV_1 66% predicted) had died at 3 years. The original ADO index showed high discrimination but poor calibration ($p<0.001$ for difference between predicted and observed risk). The updated ADO index (scores from 0 to 14) preserved excellent discrimination (area under curve 0.81, 95% CI 0.80 to 0.82) but showed much improved calibration with predicted 3-year risks from 0.7% (95% CI 0.6% to 0.9%, score of 0) to 64.5% (61.2% to 67.7%, score of 14). The ADO index showed higher net benefit in subjects at low-to-moderate risk of 3-year mortality than FEV_1 alone.

Interpretation: The updated 15-point ADO index accurately predicts 3-year mortality across the COPD severity spectrum and can be used to inform patients about their prognosis, clinical trial study design or benefit harm assessment of medical interventions.

ARTICLE SUMMARY

Article focus

- We aimed to conduct a large international study to validate the ADO index that uses age, dyspnoea and FEV_1 to predict 3-year mortality and to update it in order to make prediction of mortality in chronic obstructive pulmonary disease (COPD) patients as generalisable as possible.

Key messages

- The updated 15-point ADO index accurately predicts 3-year mortality across the COPD severity spectrum (GOLD stage I–IV), settings (general population, primary care and specialised care) and geographical area.
- The updated ADO index can be used to inform patients, clinical trial study design and benefit harm assessment of medical interventions on a population level or individual level.
- In addition, the ADO index could serve as a reference standard for risk prediction against which the additional value of various biomarkers to predict mortality could be assessed.

Strengths and limitations of this study

- The study includes a large sample size from 10 European and American cohorts and covers the entire COPD severity spectrum, which increases external validity.
- The study uses information readily available in routine clinical practices.
- Focus on mortality and easily available predictors.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is among the leading causes of death worldwide.[1 2] Although the substantial excess mortality associated with COPD is well

recognised from a public health standpoint, relatively little evidence is available on how to estimate the risk of mortality for an individual patient. Tools to accurately project the clinical course of the disease, including prediction of outcomes such as mortality, exacerbations or quality of life, would inform patients and their caregivers about prognosis and allow for a better understanding of the benefits and harms of possible treatments.[3–6] Also, tools that incorporate prognostic information from easily available parameters could serve as reference against which the additional prognostic value of biomarkers could be assessed.

Current international COPD guidelines provide little guidance on how to assess a patient's prognosis.[7 8] This is likely due, in large part, to the scarcity of evidence on how to accurately estimate prognosis in patients with COPD; however, this is in contrast to other chronic disease guidelines that have clear recommendations on the use of prognostic indices to inform patients and to guide treatment decisions.[9–13] Prognostic indices for COPD have recently received increased attention, but have seen little application in clinical practice. This may be because indices, to date, have either required information not readily available in routine clinical practice,[14] do not provide explicit outcome risks[15 16] or have received minimal validation.[15–17] A recently developed index, the ADO index, combines age, dyspnoea and airflow obstruction to predict the risk of mortality. It may have great potential for widespread application because of its simplicity. However, formal testing of its accuracy across a variety of COPD patient cohorts following standard methods has not yet been done.[6 18–21] The original index was derived in a cohort of moderate-to-severe COPD patients from specialised care, therefore, it requires validation in larger and notably more diverse COPD populations. We conducted such large-scale international validation of the ADO index to determine how well it predicts mortality for individual subjects with COPD from diverse settings, and updated the index as needed.

METHODS

Cohorts and patients

Investigators from 10 COPD and population-based cohort studies in Europe and the Americas agreed to collaborate in the International COPD Cohorts Collaboration Working Group. These cohorts include the Barmelweid cohort (Switzerland, clinic based),[17] the Basque study (Spain, clinic based),[22] the Cardiovascular Health Study (CHS, USA, population based),[23] the Copenhagen City Heart Study (CCHS, Denmark, population based),[24] the Jackson Heart Study (JHS, USA, population based),[25] the Lung Health Study (LHS, USA, clinic based),[26] the cohort from which patients for the National Emphysema Treatment Trial were recruited (NETT, USA, clinic based),[27] the Phenotype and Course of COPD PAC-COPD Study (PAC-COPD, Spain, clinic based),[28] the PLATINO study (Uruguay, population based),[29] and the Quality of Life of COPD Study Group (SEPOC, Spain, clinic based).[30] Details about the cohorts are provided in the online supplement (see online supplementary appendix 1). From this international pool of cohorts we selected participants with at least 40 years of age and with COPD defined by spirometry as a post-bronchodilator (BD) $FEV_1/FVC \leq 0.7$, except for the CHS and CCHS cohorts where post-BD was not available and pre-BD values were used. Thus, our large pool of cohorts represents a heterogeneous group of subjects, combining (1) COPD patients from clinical cohorts and (2) subjects with evidence of airway limitation from the population-based cohorts, but without a confirmed diagnosis of COPD. Ethics Board approval was obtained in all cohorts.

Mortality and candidate predictors of mortality

All-cause mortality at 3 years was defined as the outcome. It was obtained from personal follow-up of patients or relatives, national registries, or hospital records, yielding no missing information with respect to mortality. We considered potential predictors of mortality which are easy to obtain across diverse medical settings. These variables included age, sex, smoking status, prebronchodilator or postbronchodilator FEV_1 as available, dyspnoea score (Medical Research Council Dyspnea scale), respiratory signs and symptoms (cough, sputum and wheezing), body mass index (BMI), asthma and cardiovascular disease (CVD, which included ischaemic heart disease, stroke, congestive heart failure or peripheral vascular disease). As in previous analyses,[17] we explicitly excluded potential predictors of mortality which are more burdensome to measure such as exercise capacity (eg, 6 min walked distance) or arterial blood gases, since these are unlikely to be available consistently in clinical practice outside academic centres. Missing values were imputed using 10-fold multiple imputation for each cohort, using the remaining variables as predictors.[31 32] Methods used for collecting and harmonising data, and for handling missing data are detailed in the appendix (see online supplementary appendices 2–4).

Statistical analysis

A detailed version of statistical analysis including sample size assessment is available in the appendix (see online supplementary appendix 5).

We first validated the original ADO index[17] through the assessment of its discrimination (area under curve) and calibration (comparison of predicted vs observed risk) properties in all subjects except for those included in the original derivation cohort (ie, the Barmelweid study).

In order to make the risk estimation tool as generalisable to different international populations as possible, we then updated the ADO index following standardised procedures that first included an updating or adjustment

of the intercept only followed by, if necessary, more extensive updates including model revision (refitting the predictor-outcome associations) and model extension (adding new predictors).[21 32] Model refitting of the ADO index was performed using all subjects from the CCHS, LHS, NETT, PLATINO and PAC-COPD cohorts (update cohort, n=10 221), applying logistic regression with death as the outcome variable and age, dyspnoea and FEV_1 as predictors. Then the validation (discrimination and calibration) of the final updated ADO index was done with the subjects from the Barmelweid study, CHS, Basque Study, JHS and SEPOC cohorts (validation cohort, n=3693). Thus, both update and validation sets included a large number of subjects with COPD or airflow limitation, diverse in terms of disease severity (GOLD I–IV) and settings (general population, primary care and specialised care). We translated the final model into a simple-to-use 15-point scale.[33]

To further quantify the accuracy of the updated ADO index, we performed a decision curve analysis that compares the net benefit of different approaches. Net benefit is defined as the difference between the proportion of subjects that are correctly identified to be at or above a certain risk threshold (eg, 5% risk) and the proportion of subjects incorrectly identified to be at or above that threshold. We focused on subjects with COPD at low-to-moderate risk for 3-year mortality (<20%) where most uncertainty about the balance between benefits and harms of treatments may exist so that risk thresholds may be specifically useful.[34 35]

Finally, we explored whether adding new predictors (eg, CVD, BMI and sex) improved the updated (refitted) models' discrimination and calibration and we conducted three sensitivity analyses that tested how susceptible our results were to analytical approaches taken. All analyses were repeated: (1) using multilevel (rather than conventional) logistic regression analysis; (2) excluding subjects with mild COPD (GOLD stage I) and (3) excluding subjects with a physician diagnosis of asthma from cohorts where only prebronchodilator spirometry was available. We also considered restricting the analyses to subjects with an FEV_1/FVC ratio below their lower limit of normal level according to local prediction equations, but the number of subjects not fulfilling this criterion was very low (<1%).

We conducted all analyses using Stata for Windows (V.11.1, College Station, Texas, USA) and R, V.2.12 (R Foundation for Statistical Computing, Vienna, Austria, 2011).

RESULTS

In total, 13 914 subjects with COPD (60% men) were included in the analysis (table 1). On average, subjects were approximately 61 years old, with moderate airflow limitation and mild dyspnoea; however, there was a wide range of disease severity within and across cohorts. The majority of subjects were former or current smokers (89%) and 22% had concomitant CVD. After 3 years 1350 (9.7%) subjects had died.

The original ADO index showed high discrimination (see online supplementary appendix 6) but poor calibration with a substantial mismatch between predicted and observed risks across the entire risk spectrum. Updating the intercept only did not substantially improve this miscalibration. Therefore, we decided to update the original ADO index.

In the update cohort, the updated ADO model showed very good agreement between predicted and observed 3-year mortality risk across 10 equally sized groups of subjects with increasing predicted risk (figure 1: mean predicted risk 9.1%). More importantly, in the validation cohort, the updated index still had good prediction across all risk categories, in particular in subjects at mortality risks below 20%. There was only a slight overprediction among subjects at very high risk. This validation did not indicate a need for further adjustment of the intercept or regression coefficients of the updated ADO model, which indicated good generalisability across countries and settings. Discrimination was, as expected, somewhat lower in the validation cohort but still 0.73 (95% CI 0.70% to 0.76%). Further extensions of the updated ADO index by adding CVD, BMI and sex did not substantially improve the model's discrimination or calibration, even though all three predictors were significantly associated with mortality in the multivariate model (all p values <0.05). The area under the curve remained 0.85 in the update cohort and 0.74 in the validation cohort, and the calibration also remained good (see online supplementary appendix 7).

Tables 2 and 3 show the updated ADO index where the strength of association of age, dyspnoea and FEV_1 with 3-year mortality is reflected in the regression coefficients and the corresponding integer point score. The 3-year risks of mortality associated with ADO scores are shown in table 4 and range from 0.7% (95% CI 0.6% to 0.9%) with a score of zero to 64.5% (95% CI 61.2% to 67.7%) at a point score of 14. The area under the curve of the updated ADO index is 0.81 (95% CI 0.80% to 0.82%).

Figure 2A shows that, from 1% to 20% risk of 3-year mortality, using the updated ADO index (regression equation) is consistently more accurate to classify patients correctly above or below certain risk thresholds than using either of the three predictors alone. Figure 2B shows the consequences of more accurate risk classification. For example at a risk threshold of 5%, using the ADO index would result in a reduction of the number of patients classified incorrectly to be above 5% by 33/100 subjects compared with considering all patients to be above 5% (ie, without using any predictors), and compared with using only FEV_1 (18 per 100 subjects), age (24 per 100 subjects) or dyspnoea (10 per 100 subjects). At higher risk thresholds, the updated ADO index and FEV_1 perform similarly.

Discrimination, calibration and the analysis of accuracy for risk thresholds remained essentially

Table 1 Description of sociodemographic and clinical characteristics of 13 914 subjects with COPD from the cohorts

	Total n=13914	Barmelweid cohort Switzerland, Europe n=231	Basque study Spain, Europe n=106	Cardio-vascular Health Study USA, North America n=2619	Copenhagen City Heart Study Denmark, Europe n=2287	Jackson Heart Study USA, North America n=419	Lung Health Study USA, North America n=5167	National Emphysema Treatment Trial USA, North America n=2252	PAC-COPD Study Spain, Europe n=342	PLATINO study Uruguay, South America n=173	SEPOC study Spain, Europe n=318
Age (years), mean (SD)	60.8 (11.6)	72.4 (8.8)	70.5 (8.9)	73.6 (5.9)	60.7 (9.4)	62.4 (11.0)	50.1 (5.7)	66.7 (6.3)	67.9 (8.6)	67.2 (11.3)	65.2 (9.2)
Sex: male, n (%)	8324 (60)	138 (60)	104 (98)	1341 (51)	1235 (54)	184 (44)	3223 (62)	1366 (61)	318 (93)	97 (56)	318 (100)
Working status: active, n (%)	5297 (63)	n.a.	n.a.	n.a.	n.a.	413 (99)	4538 (88)	178 (8)	61 (18)	63 (36)	51 (16)
Smoking											
Never, n (%)	1452 (11)	9 (5)	0 (0)	930 (36)	215 (9)	215 (52)	0 (0)	12 (1)	2 (1)	57 (33)	12 (5)
Former, n (%)	4751 (34)	138 (73)	82 (77)	1245 (48)	443 (19)	102 (25)	73 (1)	2240 (100)	220 (67)	60 (35)	148 (58)
Current, n (%)	7590 (55)	41 (22)	24 (23)	444 (17)	1626 (71)	99 (24)	5094 (99)	0 (0)	109 (33)	56 (32)	97 (38)
Body mass index(kg/m^2), mean(SD)	25.7 (4.5)	25.9 (6.1)	26.1 (4.9)	26.2 (4.8)	25.0 (4.2)	29.6 (7.5)	25.6 (3.9)	24.9 (4.2)	28.2 (4.7)	27.4 (5.3)	26.4 (4.2)
Dyspnoea (MRC, 0–4), mean (SD)	1.1 (1.3)	2.2 (1.2)	2.0 (0.6)	0.8 (1.1)	1.1 (1.3)	0.2 (0.6)	0.6 (0.8)	2.7 (1.0)	2.2 (1.0)	0.6 (0.6)	2.1 (1.5)
Cough, n (%)	4009 (59)	n.a.	n.a.	444 (17)	n.a.	108 (26)	3259 (100)	n.a.	138 (41)	60 (35)	n.a.
Sputum, n (%)	4289 (52)	n.a.	n.a.	656 (25)	908 (40)	107 (26)	2400 (100)	n.a.	172 (51)	46 (27)	n.a.
Wheeze, n (%)	4352 (53)	n.a.	n.a.	187 (9)	n.a.	73 (18)	3897 (75)	n.a.	125 (37)	70 (40)	n.a.
FEV_1 (% pred), mean (SD)*	65.9 (24.8)	45.1 (16.1)	46.9 (11.4)	77.3 (22.4)	70.5 (23.7)	71.2 (20.5)	77.8 (9.1)	27.5 (8.9)	52.4 (16.2)	84.3 (18.1)	45.0 (18.3)
Inhaler steroid use, n (%)	1833 (33)	-	103 (100)	55 (2)	n.a.	n.a.	n.a.	1376 (61)	266 (79)	33 (19)	n.a.
6-min walk distance, mean (SD)	357.8 (111.7)	363.1 (127.0)	442.9 (95.4)	n.a.	n.a.	n.a.	n.a.	340.6 (106.3)	435.5 (90.6)	n.a.	n.a.
Asthma,† n (%)	920 (8)	0 (0)	0 (0)	148 (6)	246 (11)	74 (18)	373 (7)	n.a.	30 (9)	49 (28)	0 (0)
Diabetes,† n (%)	262 (7)	41 (18)	n.a.	n.a.	60 (3)	71 (19)	n.a.	n.a.	65 (19)	7 (4)	18 (6)

Continued

Table 1 Continued

	Total n=13914	Barmelweid cohort Switzerland, Europe n=231	Basque study Spain, Europe n=106	Cardio-vascular Health Study USA, North America n=2619	Copenhagen City Heart Study Denmark, Europe n=2287	Jackson Heart Study USA, North America n=419	Lung Health Study USA, North America n=5167	National Emphysema Treatment Trial USA, North America n=2252	PAC-COPD Study Spain, Europe n=342	PLATINO study Uruguay, South America n=173	SEPOC study Spain, Europe n=318
Cardiovascular disease,†‡ n (%)	3007 (22)	165 (71)	25 (29)	747 (29)	813 (34)	79 (19)	278 (5)	704 (31)	85 (25)	37 (21)	74 (23)
Death during 3-year follow-up, n (%)	1350 (10)	79 (34)	16 (15)	232 (9)	186 (8)	29 (7)	58 (1)	632 (28)	41 (12)	16 (9)	61 (19)

*Prebronchodilator FEV_1 used where postbronchodilator FEV_1 not available (Cardiovascular Health Study and Copenhagen City Heart Study).
†Cocomorbidities are self-reported, self-report of a doctor diagnosis, or doctor diagnosed (according to medical chart and physical examination) depending on the cohort.
‡Cardiovascular disease is defined as at least one of the following: ischaemic heart disease, stroke, congestive heart failure or peripheral vascular disease.
COPD, chronic obstructive pulmonary disease.

unchanged in all sensitivity analyses (see online supplementary appendix 8).

DISCUSSION

Our study showed that the updated ADO index, ranging from 0 to 14, accurately predicts 3-year mortality in subjects with COPD. We found that adding CVD, BMI or sex does not significantly improve prediction of mortality when added to age, dyspnoea and FEV_1. Importantly, these results were consistent in sensitivity analyses and across very diverse COPD populations. Based on these results, the updated ADO index has the potential to provide COPD patients with accurate prognostic information on mortality.

The interest in prognostic assessment of COPD has resulted in several prognostic tools.[15–17] The latter had, however, little impact on clinical guidelines or practice so far. The current study overcomes potentially important barriers to the use of previously published prognostic tools by providing an extensive, international validation of a simple tool. The first step of our analysis, the large-scale validation of the original ADO index, showed that mortality could not be predicted accurately but that the combination of age, dyspnoea and FEV_1 is highly discriminative. Therefore, a more extensive update than just an adjustment for different underlying risks was necessary and we updated the entire regression model in our very diverse update cohort that represented the entire disease spectrum. The resulting updated ADO index showed excellent calibration and discrimination in both the update and validation cohorts. An additional adjustment was not deemed necessary for the validation cohort, which may be due to the great diversity of the update and validation cohorts in terms of disease severity, clinical setting and geographical area. Our results confirm that CVD and low BMI are important comorbidities in COPD patients and are significantly associated with mortality; however, they did not provide additional accuracy in risk prediction when added to age, dyspnoea and FEV_1, as shown by the fact that the performance of the ADO index was not improved by adding CVD, sex and BMI to the statistical model.

Informing patients about their prognosis is a core task of clinicians. Patients with chronic disease are particularly interested in the potential course of their disease in order to better understand what a diagnosis such as COPD implies for them. Important outcomes that characterise prognosis are exacerbations, quality of life and mortality.[36] With the ADO index, we now provide a simple tool that clinicians from any setting can use to estimate the risk of mortality. We propose that such multivariable tools can also be used to balance the benefits and harms of possible treatments since the benefit harm balance depends heavily on the patients' prognosis.[4 37] Thus, estimation of prognosis is of key importance for patients but also for policy makers, regulatory agencies and clinical guideline developers. Our data suggest that

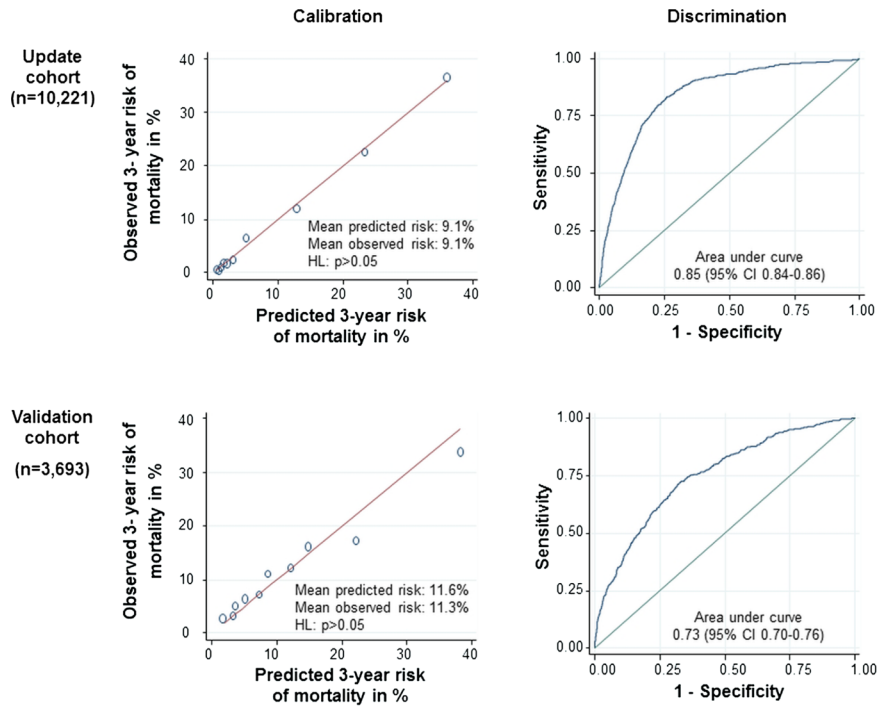


Figure 1 Update and validation of the ADO index in 13 914 subjects with chronic obstructive pulmonary disease (COPD). The upper part of the figure shows the predictive performance of the updated ADO index in 10 221 subjects with COPD from the Copenhagen City Heart Study, Lung Health Study, National Emphysema Treatment Trial, PLATINO and the Phenotype and Course of COPD Study. The calibration plot shows the predicted and observed risks for 10 equally sized group with increasing risk of 3-year mortality. The discrimination plot shows the area under the curve. The lower part of the figure shows the predictive performance of the updated ADO index in the validation cohort with 3693 subjects from the Cardiovascular Health Study, Basque COPD study, Jackson Heart Study, Barmelweid Study and the Quality of Life of Chronic Obstructive Pulmonary Disease Study (SEPOC).

Table 2 Regression coefficients and development of updated ADO index

Variable	Regression coefficients β_s per unit increase	Categories	Reference values W_{ij} (mid point)	β_s* (W_{ij}-*$W_{1reference}$)	Risk score=β_s* (W_{ij}-*$W_{1reference}$)/B†
FEV_1 (% predicted)	−0.0288 (SE 0.0023, p<0.0001)	≥81	87.0 ($W_{2reference}$)		0
		≥65–80	72.5	0.418	1
		≥50–64	57.0	0.864	2
		≥36–49	42.5	1.282	3
		≤35	25.0	1.786	4
Dyspnoea (mMRC, 0–4)	0.2585 (SE 0.0406, p<0.0001)	0	0 ($W_{3reference}$)		0
		1	1	0.259	1
		2	2	0.517	1
		3	3	0.776	2
		4	4	1.034	3
Age (years)	0.0703 (SE 0.0048, p<0.0001)	40–49	44.5 ($W_{4reference}$)		0
		50–59	54.5	0.703	2
		60–69	64.5	1.406	4
		70–79	74.5	2.109	5
		≥80	84.5	2.812	7

†1 Point is assigned per 15% in FEV_1=coefficient of 0.40. Points rounded to the next integer. Constant of regression equation=−5.640.
MRC, Medical Research Council.

Table 3 Assignment of points for the updated ADO index, compared with the original ADO index

Assignment of points	0	1	2	3	4	5	7
Updated ADO index							
FEV_1 (% predicted)	≥81	65–80	51–64	36–50	≤35		
Dyspnoea (mMRC, 0–4)	0	1–2	3	4			
Age (in years)	40–49		50–59		60–69	70–79	≥80
Original ADO index							
FEV_1 (% predicted)	≥65	36–64	≤35				
Dyspnoea (mMRC, 0–4)	0–1	2	3	4			
Age (years)	40–49	50–59	60–69	70–79	80–89	≥90	

the ADO index classifies patients more correctly above or below certain risk thresholds than only FEV_1, and this gain is especially relevant in subjects with very low and low risk of 3-year mortality where the benefit harm balance may be unfavourable.[3 4] Although most COPD treatments are not prescribed to modify mortality risk, but to reduce exacerbations, and improve symptoms and quality of life, similar estimates for the benefit harm balance could be made for patients at low risk for exacerbations. Therefore, in the future, the ADO index should be complemented by other widely validated risk tools that make accurate projections about the risk for important outcomes in COPD, including exacerbations or worsening quality of life, in order to balance the benefits and harms of possible treatments.

In addition, accurate prediction of mortality by the ADO index can be used in clinical trials to base sample size calculations on realistic estimates of expected event rates, to target treatments to specific risk groups, for prestratification or to adjust for potential baseline imbalances. Also, the ADO index could be useful for the evaluation of biomarkers. Currently, major studies are being carried out to identify biomarkers that might help to improve outcome prediction and response to treatments. The use of such biomarkers in clinical practice seems justified if they add significantly to the prediction based on easily available information. The ADO index is a simple tool that could serve as reference against which the additional value of biomarkers to predict mortality could be assessed. Therefore, the ADO index is likely to be useful for both medical practice and research.

Table 4 The ADO index—prediction of 3-year mortality in chronic obstructive pulmonary disease subjects

Three-year risk of mortality per ADO score in % (95% CI)	
0	0.7 (0.6 to 0.9)
1	1.0 (0.9 to 1.2)
2	1.6 (1.3 to 1.8)
3	2.3 (2.0 to 2.6)
4	3.4 (3.0 to 3.7)
5	4.9 (4.5 to 5.4)
6	7.2 (6.7 to 7.7)
7	10.3 (9.7 to 10.9)
8	14.5 (13.8 to 15.3)
9	20.1 (19.1 to 21.1)
10	27.2 (25.8 to 28.6)
11	35.7 (33.7 to 37.7)
12	45.1 (42.6 to 47.7)
13	55.0 (52.0 to 58.0)
14	64.5 (61.2 to 67.7)

OR per 1 point increase in ADO index: 1.48 (95% CI 1.45 to 1.52).
Area under the curve: 0.81 (95% CI 0.80 to 0.82).

A limitation to the current study is the use of mortality as the only assessed outcome, since COPD morbidity includes other relevant outcomes such as exacerbations, hospital admissions, or quality-of-life. Thus, our study should be considered a simplification of the clinical setting but it may pave the way for similar research evaluating risk prediction of additional outcomes. Once risk tools for various important outcomes and improved evidence about benefit and harm of treatments to modify these risks are available, informed decisions for providing the most appropriate care can be better supported. Our study was confined to a limited number of readily available predictors; therefore, variables with potential relevance to mortality risk such as exacerbation frequency or measures of exercise capacity were not included. This may also be perceived as strength of our study because it uses information readily available in routine clinical practices, including primary care settings, where most COPD patients are treated. Additional strengths of our study include the already mentioned large sample size and diversity of the populations. This increases external validity, which in this context means that recalibrations in populations different from those included in our analyses do not seem necessary. Lastly, by using decision curve analysis we looked beyond standard metrics for the performance of risk tools (discrimination and calibration) by providing an interpretation of the risk model that refers to different risk thresholds that may be used to inform treatment decisions.

In conclusion, the updated 15-point ADO index is a simple tool that can be used in diverse settings to inform patients and their caregivers about prognosis. Using risk tools in clinical COPD research may also help to design clinical trials and to inform policy makers, regulatory agencies or guideline developers when estimating the benefit harm balance and to serve as a reference standard for risk prediction against which the

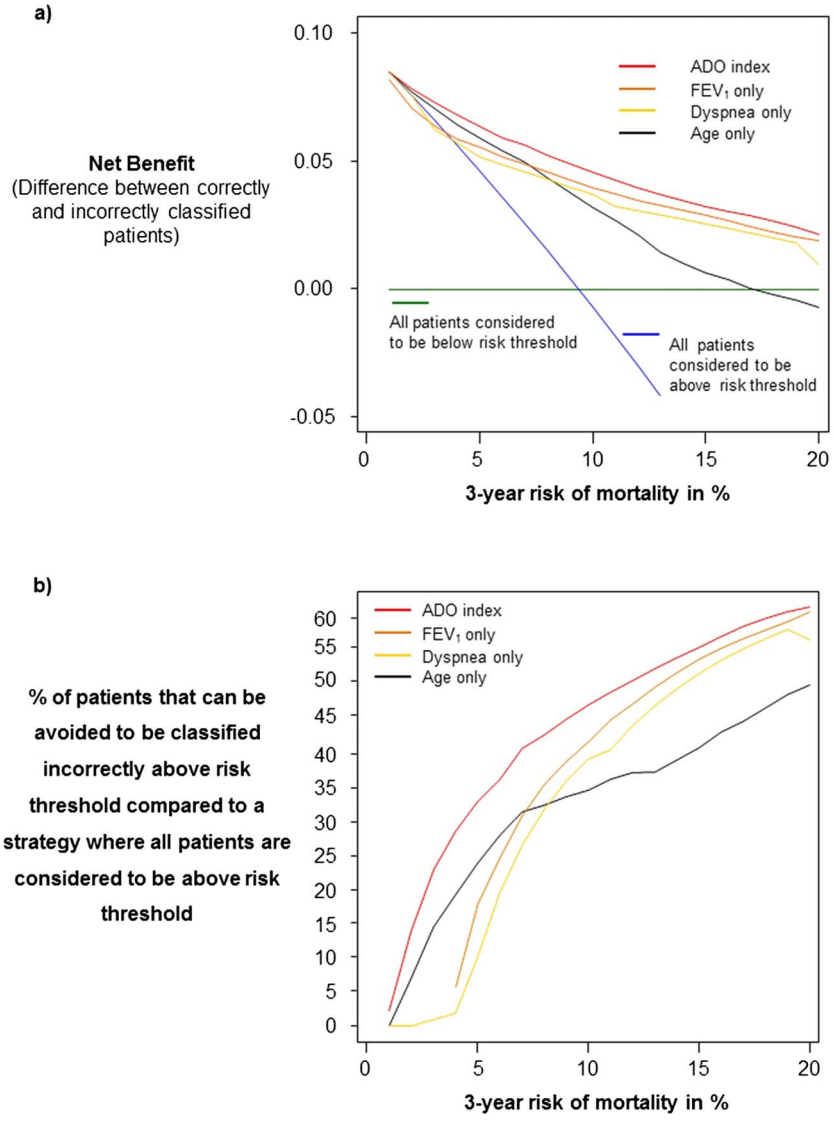


Figure 2 Accuracy of four strategies to classify subjects into risk categories. The upper part (A) of the figure shows the net benefit of six strategies to classify subjects with chronic obstructive pulmonary disease. The higher the values for net benefit the more patients are correctly classified. Two strategies do not use any predictors but assume that all patients would be above or below a risk threshold. The four other strategies use the ADO index, age, dyspnoea or FEV_1 and associated risks of 3-year mortality to classify patients. Net benefit is defined as the difference between the proportion of correctly classified subjects and the proportion of subjects classified incorrectly to be at or above a risk threshold (eg, 5% risk). The line for considering all patients to be above a risk threshold crosses that line for considering all patients to be below a risk threshold at the death rate observed (9.7%). The lower part of the graph (B) shows that, for example at a threshold of 5% mortality risk, using the ADO index would result in a net benefit similar to the reduction of 33 incorrectly classified patients per 100 subjects compared to considering all patients to be above a 5% mortality risk. Using age, dyspnoea or FEV_1 only would reduce it by only 24, 10 and 18 per 100 subjects, respectively. The graph is restricted to subjects at low to moderate risk for 3-year mortality (<20%) where most uncertainty about the balance between benefits and harms from treatment may exist.

additional value of various biomarkers to predict mortality could be assessed.

Author affiliations
[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA
[2]Horten Centre, University of Zurich, Zurich, Switzerland
[3]Division of Pulmonary & Critical Care Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA
[4]CIBER Enfermedades Respiratorias (CIBERES), Fundación Caubet-Cimera, Mallorca, Spain
[5]College of Public Health, The University of Arizona, Tucson, Arizona, USA
[6]Pulmonary Division, Hvidovre Hospital, Hvidovre, Danmark
[7]Jackson Heart Study, Coordinating Center, Jackson State University, Jackson, Mississippi, USA
[8]Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi, USA
[9]Federal University of Pelotas, Pelotas, Brazil
[10]Department of General Practice, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands
[11]Hospital del Mar Research Institute (IMIM), Barcelona, Spain
[12]Cedars Sinai Medical Center, David Geffen School of Medicine at UCLA, Los Angeles, California, USA
[13]Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
[14]Universitat Pompeu Fabra, Barcelona, Spain
[15]CIBER de Epidemiologia y Salud Pública (CIBERESP), Barcelona, Spain

[16]Department of Epidemiology, Julius Centre for Health Sciences and General Practice, University Medical Center, Utrecht, The Netherlands
[17]Clinical Epidemiology, and Medical Technology Assessment, University Hospital Maastricht, Maastricht, The Netherlands

Collaborators Cohorts and investigators of the International COPD Cohorts Collaboration Working Group: Barmelweid cohort (Switzerland)—Martin Frey, MD, Klinik Barmelweid, Switzerland; Ulrike Held, PhD, Horten Centre, University of Zurich, Switzerland; Milo A. Puhan, MD, PhD, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA and Horten Centre, University of Zurich, Switzerland. Basque study (Spain)—Milagros Iriberri Pascual, MD, Cruces Hospital, Baracaldo, Basque Country, Spain; Patricia Sobradillo, MD, CIBER Enfermedades Respiratorias (CIBERES), Fundación Caubet-Cimera, Mallorca, Spain. Cardiovascular Health Study (USA)—R. Graham Barr, MD, Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, USA; Paul Enright, MD, College of Public Health, The University of Arizona, Tucson, Arizona, USA; Jerry Krishnan, MD, Department of Medicine, Section of Pulmonary and Critical Care, University of Chicago, Chicago, USA; Tony Wilson, Department of Biostatistics, University of Washington, Seattle, USA; Sachin Yende, MD, Department of Critical Care Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. Copenhagen City Heart City study (Denmark)—Peter Lange, MD, Pulmonary Division, Hvidovre Hospital, Hvidovre, Danmark. Jackson Heart Study (USA)—DeMarc Hickson, PhD, MPH, Jackson State University, Jackson Heart Study, Coordinating Center, USA. Department of Medicine, University of Mississippi Medical Center, USA; Wendy White, PhD, MPH, Tougaloo College, Jackson Heart Study, USA. Lung Health Study (USA)—Nadia Hansel, MD, MPH, and Robert A. Wise, MD, Division of Pulmonary & Critical Care Medicine, Johns Hopkins School of Medicine, Baltimore, USA. National Emphysema Treatment Trial (USA)—Fernando Martinez, MD, Division of Pulmonary and Critical Care Medicine, University of Michigan Health System, Ann Arbor, USA; Zab Mosenifar, MD, Cedars Sinai Medical Center, David Geffen School of Medicine at UCLA, Los Angeles, USA; Andrew Ries, MD, Department of Medicine, University of California, San Diego, USA; James K. Stoller, MD, MS, Respiratory Institute, Department of Pulmonary and Critical Care Medicine, Cleveland Clinic Foundation, Cleveland, USA; Robert A. Wise, MD, see above. PAC-COPD (Spain)—Alvar Agustí, MD, Federico P Gómez, MD, and Roberto Rodriguez-Roisin, MD, Department of Pneumology, Thorax Institute, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, and CIBER Enfermedades Respiratorias (CIBERES), Spain; Josep M Antó, MD, PhD, and Judith Garcia-Aymerich, MD, PhD, (1) Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain, (2) Hospital del Mar Research Institute (IMIM), Barcelona, Spain, (3) Universitat Pompeu Fabra, Barcelona, Spain, (4) CIBER de Epidemiologia y Salud Pública (CIBERESP), Spain. PLATINO study team (Brazil, Uruguay)—Ana MB Menezes, MD, PhD, Federal University of Pelotas, Pelotas, Brazil; Maria Victorina Lopez, MD, and Adriana Muiño, MD, Universidad de la República, Montevideo. Uruguay. SEPOC study (Spain)—Josep M Antó, MD, PhD, see above; Antonia Domingo-Salvany, MD, PhD, Hospital del Mar Research Institute (IMIM), Barcelona, Spain; Judith Garcia-Aymerich, MD, PhD, see above; Methodological and Statistical team; Judith Garcia-Aymerich, MD, PhD, see above; Nadia Hansel, MD, MPH, see above; Ulrike Held, PhD, see above; Alphons Kessels, MD, Clinical Epidemiology, and Medical Technology Assessment, University Hospital Maastricht, Maastricht, The Netherlands; Karel GM Moons, PhD, Julius Centre for Health Sciences and General Practice, University Medical Center, Utrecht, The Netherlands; Milo A. Puhan, MD, PhD, see above; Gerben ter Riet, Department of General Practice, Academic Medical Centre, University of Amsterdam, The Netherlands; Ignasi Serra, Center for Research in Environmental Epidemiology, Barcelona, Spain.

Contributors MP, NH and JGA conceived the study idea, contributed to the statistical analysis and drafted the first version of the manuscript. MF, UH and MP oversaw all activities related to the conduct of the Barmelweid cohort (Switzerland). MIP and PS oversaw all activities related to the conduct of the Basque study (Spain). RGB, PE, JK and TW oversaw all activities related to the conduct of the Cardiovascular Health Study (USA). PL oversaw all activities related to the conduct of the Copenhagen City Heart City study (Denmark). DMH and WW oversaw all activities related to the conduct of the Jackson Heart Study (USA). NH and RW oversaw all activities related to the conduct of the Lung Health Study (USA). FM, ZM, AR and RW oversaw all activities related to the conduct of the National Emphysema Treatment Trial (USA). AA, FPG, RRR, JMA and JGA oversaw all activities related to the conduct of the PAC-COPD (Spain). AMM, MVL and AM oversaw all activities related to the conduct of the PLATINO study team (Brazil, Uruguay). JMA, ADS and JGA oversaw all activities related to the conduct of the SEPOC study (Spain). UH, AK, KGM, GtR and IS contributed to the statistical analysis. All authors revised the manuscript and accepted the final version.

Funding The Barmelweid cohort (Switzerland) was funded by the Swiss National Science Foundation (grant number 3233B0-115216) and by the Klinik Barmelweid. Basque study (Spain): No external funding. The Cardiovascular Health Study is supported by NHLBI Grant/Contract numbers N01-HC-85239, N01-HC-85079 through N01-HC-85086; N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133; HL080295, HL-075366; NIA Grant/Contract numbers AG-023269, AG-15928, AG-20098, and AG-027058; University of Pittsburgh Claude. D. Pepper Older Americans Independence Center grant number P30-AG-024827; with additional contribution from NINDS. See also http://www.chs-nhlbi.org/pi.htm The Copenhagen City Heart City study (Denmark) was supported by grants from The Danish Heart Foundation, The Danish Lung Association and Danish Medical Research Council. The Jackson Heart Study (JHS) is a collaborative study supported by the National Institutes of Health and the National Center on Minority Health and Health Disparities (study ID numbers: 5001; N01 HC95170; N01 HC95171; N01 HC95172) in partnership with Jackson State University, Tougaloo College, and University of Mississippi Medical Center. The Lung Health Study (USA) was supported by contract NIH/N01-HR-46002 from the National Heart, Lung and Blood Institute (NHLBI). The National Emphysema Treatment Trial (USA) is funded by the National Heart, Lung and Blood Institute, the Centers for Medicare and Medicaid Services, and the Agency for Healthcare Research and Quality. The PAC-COPD Study is funded by grants from Fondo de Investigación Sanitaria (FIS PI020541), Ministry of Health, Spain; Agència d'Avaluació de Tecnologia i Recerca Mèdiques (AATRM 035/20/02), Catalonia Government; Spanish Society of Pneumology and Thoracic Surgery (SEPAR 2002/137); Catalan Foundation of Pneumology (FUCAP 2003 Beca Marià Ravà); Red RESPIRA (RTIC C03/11); Red RCESP (RTIC C03/09), Fondo de Investigación Sanitaria (PI052486); Fondo de Investigación Sanitaria (PI052302); Fundació La Marató de TV3 (num. 041110); DURSI (2005SGR00392); and an unrestricted educational grant from Novartis Farmacèutica, Spain. CIBERESP and CIBERES are funded by the Instituto de Salud Carlos III, Ministry of Health, Spain. PLATINO study was funded by ALAT (Associacion Latino Americana del Tórax); Boehringer Ingelheim GmbH (BI), and GlaxoSmithKline (GSK). The SEPOC study (Spain) was supported by grants from Fondo de Investigación Sanitaria 99/0690 and CIRIT 1999SGR00240. Judith Garcia-Aymerich has a researcher contract from the Instituto de Salud Carlos III (CP05/00118), Ministry of Health, Spain. Karel G.M. Moons receives funding from the Netherlands Organisation for Scientific Research (project 9120.8004 and 918.10.615).

Competing interests None.

Ethics approval Ethics Committees of each participating hospital/reserach institute.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Lopez AD, Shibuya K, Rao C, *et al*. Chronic obstructive pulmonary disease: current burden and future projections. *Eur Respir J* 2006;27:397–412.
2. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med* 2006;3:e442.
3. Guo JJ, Pandey S, Doyle J, *et al*. A review of quantitative risk-benefit methodologies for assessing drug safety and efficacy-report of the ISPOR risk-benefit management working group. *Value Health* 2010;13:657–66.
4. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA* 2007;298:1209–12.

5. Politi MC, Han PK, Col NF. Communicating the uncertainty of harms and benefits of medical interventions. *Med Decis Making* 2007;27:681–95.
6. Moons KG, Royston P, Vergouwe Y, *et al*. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338:b375.
7. Global Strategy for Diagnosis, Management, and Prevention of COPD. 2010. http://wwwgoldcopdcom/GuidelinesResourcesasp?l1=2&l2=0 (accessed 10 Sept 2012).
8. Celli BR, MacNee W. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004;23:932–46.
9. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation* 2002;106:3143–421.
10. American Diabetes Association. Standards of medical care in diabetes. *Diabetes Care* 2011;34(Suppl 1):S11–61.
11. National Collaborating Centre for Cancer. NICE clinical guideline 58: Prostate cancer: diagnosis and treatment. National Institute for Health and Clinical Excellence. 2008. http://www.nice.org.uk/cg58 (accessed 10 Sept 2012).
12. National Collaborating Centre for Cancer. NICE clinical guideline 80: Early and locally advanced breast cancer: diagnosis and treatment. National Institute for Health and Clinical Excellence. 2009. http://www.nice.org.uk/cg80 (accessed 10 Sept 2012).
13. US Preventive Services Task Force. Aspirin for the prevention of cardiovascular disease: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;150:396–404.
14. Puhan MA, Zoller M, ter Riet G. COPD: more than respiratory. *Lancet* 2008;371:27–8.
15. Jones RC, Donaldson GC, Chavannes NH, *et al*. Derivation and validation of a composite index of severity in chronic obstructive pulmonary disease: the DOSE Index. *Am J Respir Crit Care Med* 2009;180:1189–95.
16. Celli BR, Cote CG, Marin JM, *et al*. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med* 2004;350:1005–12.
17. Puhan MA, Garcia-Aymerich J, Frey M, *et al*. Expansion of the prognostic assessment of patients with chronic obstructive pulmonary disease: the updated BODE index and the ADO index. *Lancet* 2009;374:704–11.
18. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;19:453–73.
19. Altman DG, Vergouwe Y, Royston P, *et al*. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009;338:b605.
20. Steyerberg EW, Bleeker SE, Moll HA, *et al*. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *J Clin Epidemiol* 2003;56:441–7.
21. Janssen KJ, Moons KG, Kalkman CJ, *et al*. Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* 2008;61:76–86.
22. Sobradillo P, Iriberri M, Gomez B, *et al*. *Validation of bode index as a predictor of mortality in COPD patients. 18th Annual Congress of the European Respiratory Society*. Berlin: European Respiratory Society, 2008:P531.
23. Fried LP, Borhani NO, Enright P, *et al*. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1991;1:263–76.
24. Appleyard M, Hansen A, Schnohr P. The Copenhagen City Heart Study: a book of tables with data from the first examination (1976–78) and a five years follow-up (1981–1983). *Scand J Soc Med* 1989;170:1–160.
25. Carpenter MA, Crow R, Steffes M, *et al*. Laboratory, reading center, and coordinating center data management methods in the Jackson Heart Study. *Am J Med Sci* 2004;328:131–44.
26. Connett JE, Kusek JW, Bailey WC, *et al*. Design of the Lung Health Study: a randomized clinical trial of early intervention for chronic obstructive pulmonary disease. *Control Clin Trials* 1993;14:3S–19S.
27. Fishman A, Martinez F, Naunheim K, *et al*. A randomized trial comparing lung-volume-reduction surgery with medical therapy for severe emphysema. *N Engl J Med* 2003;348:2059–73.
28. Garcia-Aymerich J, Gomez FP, Anto JM. Phenotypic Characterization and Course of Chronic Obstructive Pulmonary Disease in the PAC-COPD Study: design and methods. *Arch Bronconeumol* 2009;45:4–11.
29. Menezes AM, Perez-Padilla R, Jardim JR, *et al*. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. *Lancet* 2005;366:1875–81.
30. Domingo-Salvany A, Lamarca R, Ferrer M, *et al*. Health-related quality of life and mortality in male patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2002;166:680–5.
31. Moons KG, Donders RA, Stijnen T, *et al*. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 2006;59:1092–101.
32. Steyerberg E. Applications of prediction models. In: Steyerberg E (ed.). *Clinical prediction models—a practical approach to development, validation, and updating*. New York: Springer, 2010. pp. 281–310
33. Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: the Framingham Study risk score functions. *Stat Med* 2004;23:1631–60.
34. Vickers AJ. Decision analysis for the evaluation of diagnostic tests, prediction models and molecular markers. *Am Stat* 2008;62:314–20.
35. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565–74.
36. Cazzola M, MacNee W, Martinez FJ, *et al*. Outcomes for COPD pharmacological trials: from lung function to biomarkers. *Eur Respir J* 2008;31:416–69.
37. Gail MH. The estimation and use of absolute risk for weighing the risks and benefits of selective estrogen receptor modulators for preventing breast cancer. *Ann N Y Acad Sci* 2001;949:286–91.

Large scale international validation of the ADO index in subjects with COPD: an individual subject data analysis of 10 cohorts

ONLINE SUPPLEMENT

Appendix 1. Details of cohort studies

The **Barmelweid cohort (Switzerland)** enrolled, between 2004 and 2005, patients with moderate to very severe COPD in a secondary care hospital that provides both acute and rehabilitative care.[1] All patients were recruited after they had followed a pulmonary rehabilitation program. The Barmelweid cohort served as derivation cohort for the original ADO index.

The **Basque study (Spain)** is a is a prospective, observational study that included ambulatory patients visited from March 2004 to June 2005 in an outpatient clinic of Hospital Cruces, a university-affiliated, tertiary referral hospital in Bilbao.[2] Patients were followed up for a mean period of 4.5 years. In November 2009, mortality was assessed by exhaustive revision of clinical reports and by telephone contact.

The **Cardiovascular Health Study (CHS, USA)** is a population-based, longitudinal study of coronary heart disease and stroke in adults aged 65 years and older.[3] Eligible participants were sampled from Medicare eligibility lists in four areas in the US. All participants of the CHS were eligible for this analysis.

The **Copenhagen City Heart Study (CCHS, Denmark)** involves the study of an on-going prospective, random and age-stratified cohort of adults recruited in four waves from the general population (1976-1978, 1981-1983, 1991-1994 and 2000-2003).[4] The basis for this paper was all participants included in the second and third wave.

The **Jackson Heart Study (JHS, USA)** is a 12-year single-site observational study initiated in 2000 to investigate the etiology of cardiovascular, renal and respiratory disease in African American adults.[5, 6] All JHS participants were recruited from a tri-county area in central Mississippi, including Hinds, Rankin and Madison counties; participants with standardized spirometry (collected between 2000 and 2004) were eligible for this analysis.

The **Lung Health Study (LHS, USA)** was a multicenter (10 centers) randomized clinical trial carried out from October 1986 to April 1994, aimed to determine whether a program of smoking intervention and use of an inhaled bronchodilator could slow the rate of decline in

pulmonary function over a 5-year follow-up period.[7, 8] At baseline, all patients were active smokers between the ages of 35 and 60 with mild to moderate airflow obstruction defined as an FEV_1/FVC ratio less≤0.7, and an FEV_1 between 50-90% predicted.

The **National Emphysema Treatment Trial (NETT, USA)** was a randomized trial that compared lung volume reduction surgery versus medical care in 1,218 patients with severe COPD who had followed a pulmonary rehabilitation program.[9] In this analysis, we included the entire cohort (n=2,252) of patients who underwent spirometry as part of the eligibility assessment for NETT (1998 to 2002). Thus we included patients that were randomized subsequently (n=1,218) as well as patients with COPD who were not randomized (n=1,034).

The **Phenotype and Course of COPD Study (PAC-COPD, Spain)** is a prospective longitudinal study of 342 COPD patients hospitalized for the first time because of a COPD exacerbation in nine teaching hospitals in Spain between January 2004 and March 2006.[10, 11] They were followed prospectively for cause-specific hospitalizations and all-cause mortality during a 4-years follow-up.

The **PLATINO study (Uruguay)** is a multicenter population based study on the prevalence of COPD carried out in five centers of Latin America from 2002 to 2004 among people aged 40 years and over.[12] A follow-up after five years was completed in one of the five sites of Latin America: Montevideo, Uruguay. A total of 173 Uruguayan subjects with spirometric COPD criteria in 2003 were sought in 2008, performed the same procedures as before, including spirometry pre and post bronchodilator, and are eligible for this analysis.

The **Quality of Life of Chronic Obstructive Pulmonary Disease Study Group (SEPOC, Spain)** conducted a cohort study of 321 male patients with COPD, recruited in 1993–1994 at outpatient respiratory clinics of university hospitals in Barcelona and regional hospitals in nearby cities.[13]

Appendix 2. Methods for collecting and harmonizing candidate predictors of mortality

Smoking status and respiratory symptoms (cough, sputum, wheezing) were obtained from epidemiological questionnaires. Lung function was assessed through spirometry before and after a bronchodilator (BD) administration, following standardized procedures in all cohorts. Only pre-BD spirometry was available for CHS and CCHS. The ratio of forced expiratory volume in one second (FEV_1) to forced vital capacity (FVC) ≤0.70 was used for selecting subjects and to define airways limitation.[14] The measure considered for predicting mortality was the FEV_1, expressed as a percentage of its predicted value using local prediction equations (taking into account the age, height, and sex) for each cohort.

Dyspnea was assessed in most studies by using the Medical Research Council (MRC) questionnaire or the modified MRC questionnaire, yielding a score ranging from 0 to 4. The CCHS included a set of individual questions about dyspnea in daily life that were almost identical to the MRC dyspnea questions and allowed adaptation to the MRC scale. The Barmelweid cohort used the dyspnea domain of the chronic respiratory questionnaire, which was transformed into MRC scores (MRC 0: >6.0; MRC 1: >5.0 and ≤6.0; MRC 2: >4.0 and≤5.0; MRC 3: >2.5 and ≤4.0; MRC 4: ≤2.5). The NETT cohort used the Shortness of Breath questionnaire with scores ranging from 0 to 100, which were transformed into MRC scores (MRC 0: ≤40; MRC 1: >40 and ≤60; MRC 2: >60 and ≤80; MRC 3: >80 and ≤100; MRC 4: >100).

Co-morbidities, including asthma, diabetes, ischemic heart disease, stroke, congestive heart failure, peripheral vascular disease, or hypertension, were defined either as a self-report, a self-report of a doctor-diagnosis or a doctor diagnosis after physical examination and medical charts study.

As in previous analyses[1], we explicitly excluded potential predictors of mortality which were more burdensome to measure such as exercise capacity (e.g. six-minute walk distance test), arterial blood gases or lengthy quality of life questionnaires, since these are unlikely to be available consistently in clinical practice outside academic centers. Other variables such as socioeconomic

status, working status, or drug treatments were not consistently available across cohorts, and not possible to standardize.

Appendix 3. Number and percentage of missing data for each variable and cohort

	Barmelweid cohort *Switzerland, Europe* n=231	Basque study *Spain, Europe* N=106	Cardiovascular Health Study *USA, North America* n=2,619	Copenhagen City Heart Study *Denmark, Europe* n=2,287	Jackson Heart Study *USA, North America* n=419	Lung Health Study *USA, North America* n=5,167	National Emphysema Treatment Trial *USA, North America* n=2,252	PAC-COPD Study *Spain, Europe* n=342	PLATINO study *Uruguay, South America* n=173	SEPOC study *Spain, Europe* n=318
Age	0	0	0	0	0	0	0	0	0	0
Sex	0	0	0	0	0	0	0	0	0	0
Working status	231 (100%)	106 (100%)	2619 (100%)	2287 (100%)	0	0	0	0	0	0
Smoking	43 (18.6%)	0	0	3 (0.1%)	3 (0.7%)	0	0	11 (3.2%)	0	61 (19%)
Body mass index	0	0	8 (0.3%)	12 (0.5)	0	1 (0.02%)	0	0	0	3 (0.9%)
Dyspnea	0	0	411 (15.7%)	6 (0.3%)	0	59 (1.1%)	318 (14.1%)	4 (1.2%)	0	1 (0.3%)
Cough	231 (100%)	106 (100%)	10 (0.4%)	2287 (100%)	1 (0.2%)	1908 (36.9%)	2252 (100%)	4 (1.2%)	0	318 (100%)
Sputum	231 (100%)	106 (100%)	28 (1.1%)	3 (0.1%)	1 (0.2%)	2767 (53.6%)	2252 (100%)	4 (1.2%)	0	318 (100%)
Wheeze	231 (100%)	106 (100%)	438 (16.7%)	2287 (100%)	1 (0.2%)	0	2252 (100%)	6 (1.8%)	0	318 (100%)
Pre-BD FEV_1	231 (100%)	0	302 (11.5%)	10 (0.4%)	419 (100%)	0	0	8 (2.3%)	0	0
Post-BD FEV_1	0	19 (17.9%)	2619 (100%)	2287 (100%)	0	1 (0.02%)	3 (0.1%)	0	0	0
Inhaler steroid use	231 (100%)	3 (2.8%)	3 (0.1%)	2287 (100%)	419 (100%)	5167 (100%)	0	4 (1.2%)	0	318 (100%)
6-min walking distance	0	1 (0.9%)	2619 (100%)	2287 (100%)	419 (100%)	5167 (100%)	272 (12.1%)	33 (9.7%)	173 (100%)	318 (100%)
Asthma*	0	0	19 (0.7%)	113 (4.9%)	1 (0.2%)	0	2252 (100%)	4 (1.2%)	0	0
Diabetes*	0	106 (100%)	2619 (100%)	23 (1%)	35 (8.4%)	5167 (100%)	2252 (100%)	3 (0.9%)	0	0
Cardiovascular disease*,†	0	19 (17.9%)	15 (0.6%)	0	9 (2.2%)	0	0	0	0	0
Death during 3-y follow-up	0	0	0	0	0	0	0	0	0	0

* Co-comorbidities are self-reported, self-report of a doctor diagnosis, or doctor diagnosed (according to medical chart and physical examination) depending on the cohort.

† Cardiovascular disease is defined as at least one of the following: ischemic heart disease, stroke, congestive heart failure, or peripheral vascular disease (no hypertension).

Appendix 4. Handling of missing data

For each variable and per cohort, we determined the extent and pattern(s) of missing predictor and outcome variables. We had no missing data for the outcome (death), age and sex. The extent of missing data was small for most variables (<5%) except for dyspnea (14% in NETT and 16% in CHS cohorts), and smoking status (19% in Barmelweid and 19% in SEPOC cohort). To minimize bias and loss of power, missing data were imputed using multiple imputation (10-fold, "mi impute" command, Stata 11).[15, 16] Imputation techniques are based on the correlation between each variable with missing values and all other variables as estimated from the set of complete subjects. After imputation, all datasets were merged into one individual patient data dataset. All analyses, including figures, considered the (small) variability across the ten imputed datasets, using Rubin's rule.[17, 18]

Appendix 5. Statistical analysis – complete version

The original ADO index, ranging from 0 to 10, combined age, dyspnea and airflow obstruction to predict the 3 years risk of all-cause mortality in COPD patients. It was derived in the Barmelweid cohort and validated in the PAC-COPD cohort.[1] Using the original regression coefficients of the predictors in the ADO index[1], we first assessed its discrimination (area under curve) and calibration (predicted *versus* observed risk) in all subjects except for those included in the original derivation cohort (i.e. the Barmelweid study). We then followed a systematic approach for further updating and external validating the ADO index.[19] Given the characteristics of our original derivation cohort (moderate to severe COPD patients in a specialized secondary care setting[1]) we expected that at least an update of the intercept of the (original) ADO index would be necessary because of the different underlying baseline risks across the international studies. The intercept update was not sufficient to yield better agreement between predicted and observed risks, therefore we conducted a more extensive model revision with re-estimation of regression coefficients using logistic regression analysis with death as the outcome variable and age, dyspnea and FEV_1 as predictors.

We did not use formal sample size calculations because all the cohort studies are ongoing studies. Also, there are no generally accepted approaches to estimate the sample size requirements for derivation and validation studies of risk prediction models. Some have suggested having at least 10 events per candidate variable for the derivation of a model[20,21] and at least 100 events for validation studies.[22] Since many studies to develop and validate prediction models are small a potential solution is to have large scale collaborations as ours to derive stable estimates from regression models that are likely to generalize to other populations.[23] Our sample and the number of events far exceeds all approaches for determining samples sizes and, therefore, is expected to provide estimates that are very robust.

Our aim was to obtain a risk prediction model that would be as widely applicable as possible. To develop and validate the updated ADO models, we split our ten cohorts into two

groups of five cohorts that would represent two large COPD populations that are as diverse in terms of disease severity (GOLD I to IV), settings (general population, primary care and specialized care) and geographical area as possible. A priori, i.e. without conducting any exploratory analyses for how to split up the ten cohorts, we used all subjects from CCHS, LHS, NETT, PLATINO and PAC-COPD as update (or derivation) cohort (n=10,221) and subjects from the Barmelweid study, CHS, Basque Study, JHS and SEPOC as validation cohort (n=3,693). We explicitly did not apply a random split or equivalent cross-validation procedure, as these are rather internal than external validation methods.[24,25] The current international individual patient data analysis allowed for the most optimal method of external validation, i.e. geographical validation across countries, settings and disease severity.[24-27] For obtaining the updated ADO score, we fitted a multivariable logistic regression model with all-cause death at 3-years as outcome and age, dyspnoea, and FEV_1 as predictors. We translated the final statistical model into 15 point scale[28], and obtained its associated risks of 3-year mortality with the aim of making the ADO index simple to use in practice.

To obtain additional information about accuracy of the ADO index, we performed decision curve analysis, which explores whether the overall (net) benefit is positive or negative depending on the balance between benefit and harm.[29,30] The net benefit of different treatment decisions can be obtained for a range of outcome probabilities (e.g., 3-years mortality risk), under the assumption that each threshold of outcome probability (p_t) at which a patient (or his/her physician) would opt for (or recommend) treatment is informative of how the s/he weighs the relative harms of a false-positive and a false-negative classification and its corresponding treatment decision.[29,30] Net benefit is defined as:

$$\text{Net benefit} = \frac{\text{True positive count}}{n} - \frac{\text{False positive count}}{n}\left(\frac{p_t}{(1-p_t)}\right)$$

where the true positive count is the number of subjects correctly classified by a risk tool to be at or above p_t, the false positive count is the number of subjects incorrectly classified to be at or above p_t and who would only experience the harms from treatment, and n is the total number of

subjects. The proportion of false positives is multiplied by a term that reflects the weight that is put on false positive classifications (leading to unnecessary treatment) and on false negative classifications (leading to no treatment of patients who need treatment). For example, if p_t is 50% the threshold term is =1, which means that the benefits and harms from treatment are weighted equally. If the p_t is at 10% one would weigh false negatives (= patients not treated although they would need treatment) to be nine times more important than treating a patient unnecessarily (false positive).

For the current analysis we performed the decision curve analysis with a focus on subjects with COPD at low to moderate risk for 3-year mortality (<20%) where most uncertainty about the benefit harm balance may exist. We plotted the net benefit corresponding to six hypothetical strategies for classifying patients into risk categories: (1) using only age; (2) using only dyspnea; (3) using only FEV1 measurement; (4) using only the ADO index; (5) considering all COPD patients to be above a certain risk threshold; and (6) considering all COPD patients to be below a certain risk threshold . The latter two are reference scenarios in decision analysis not necessarily reflecting clinical practice. Decision curve analysis assumes that subjects would be treated if they are at a certain risk for the outcome (e.g. ≥5% risk of 3-year mortality) and calculates how many unnecessary treatments can be avoided.

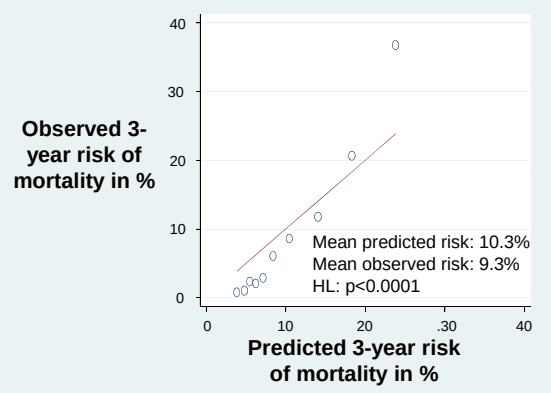
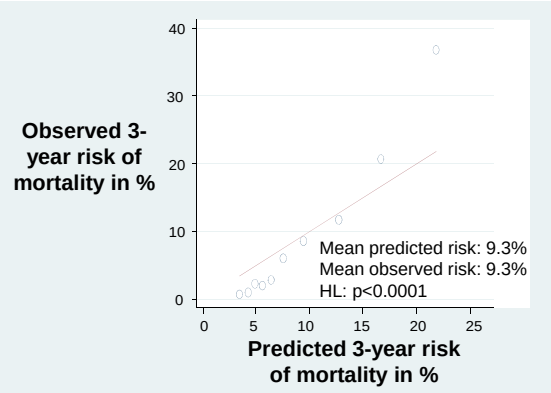
Additionally, we explored whether adding new predictors (e.g. CVD, BMI and sex) improved the updated (refitted) models' discrimination and calibration. For that purpose we repeated all the analyses above to assess if adding CVD, BMI and sex reduced risks on a continuous scale.

We conducted three sensitivity analyses to test how dependent our results were on some decisions taken. First, (sensitivity analysis 1) the analyses were repeated using multilevel (rather than conventional) logistic regression analysis. A random study-effect (related to the 3-year risk of mortality in each original study), and fixed effects for the covariates were used.[31-33] This approach takes into account that subjects within a single cohort are more likely to share some characteristics

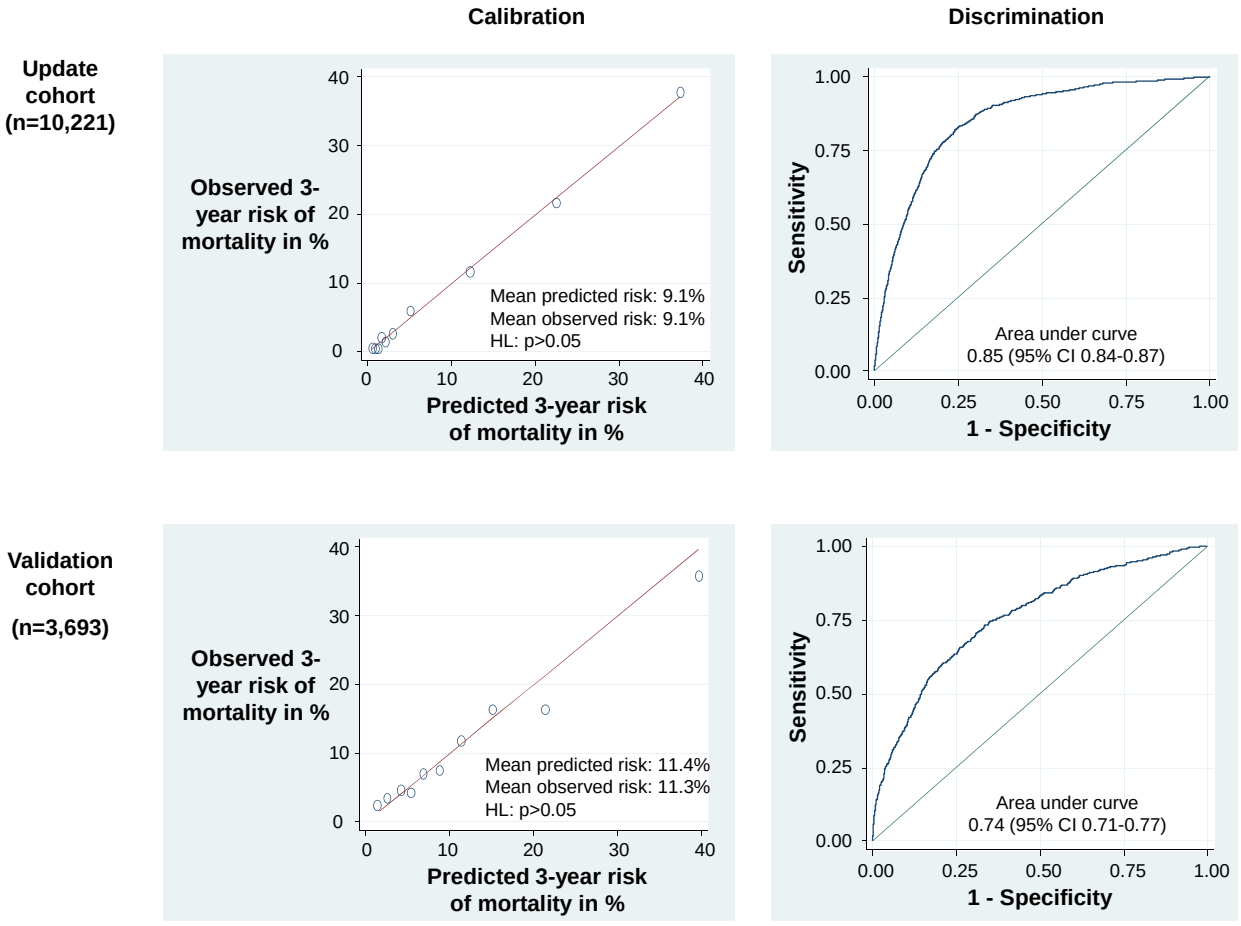
than two randomly chosen subjects from different cohorts. We compared the results from all analyses (discrimination, calibration and clinical usefulness) with those of our simpler models. Similarly, we repeated all analyses excluding subjects with GOLD stage I (sensitivity analysis 2), and excluding subjects with a physician diagnosis of asthma from cohorts where only pre-bronchodilator spirometry was available (sensitivity analysis 3). We also considered restricting analyses to subjects with a FEV_1/FVC ratio below their lower limit of normal level according to local prediction equations, thus taking into account the potential misclassification in older ages when defining COPD according to a fixed FEV_1/FVC ratio, but this analysis was not finally included because misclassification was very low (<1%). All analyses were conducted using Stata for Windows (version 11.1, College Station (TX), USA) and R 2.12 (R Foundation for Statistical Computing, Vienna, Austria, 2011).

Appendix 6. Validation of the original ADO index in 13,683 subjects with COPD

Barmelweid cohort not included because initial development cohort.

Validation cohort (n=13,683)	Calibration	Discrimination
Before re-calibration of intercept		**Area under curve (95% CI)** In all subjects (13,683): 0.82 (0.81-0.83) In derivation cohort (10,221): 0.85 (0.83-0.86) In validation cohort: (3,462) 0.70 (0.67-0.73)
After re-calibration of intercept		HL: Hosmer Lemeshow test statistic

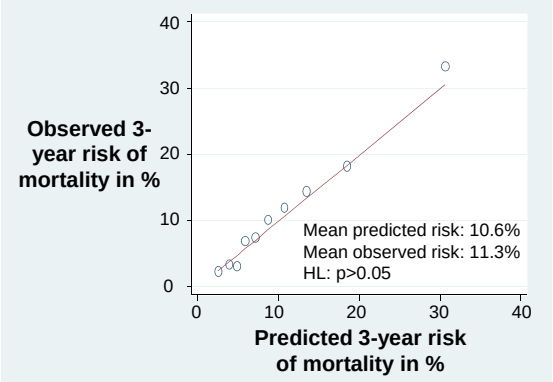
Appendix 7. Derivation and validation of the updated and extended (CVD, BMI and sex) ADO index



Appendix 8. Results of sensitivity analyses

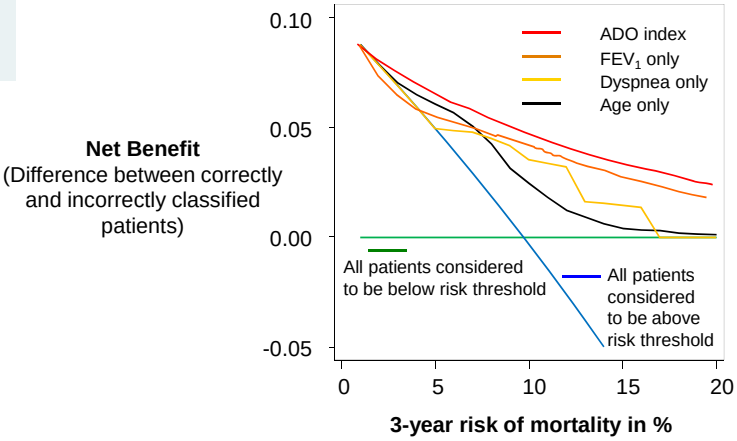
Sensitivity I: Multilevel model. Performance of updated ADO index in validation cohort (all 3,693 subjects included)

Calibration



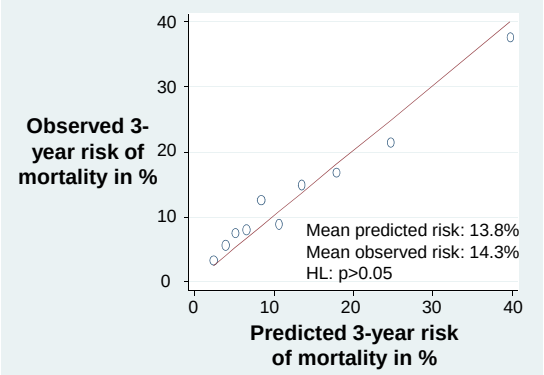
Discrimination

Area under curve
0.73 (0.70-0.75)



Sensitivity II: Performance of updated ADO index in subjects of validation cohort with GOLD stage ≥II (2,101 subjects included, 1,592 subjects with GOLD stage I excluded)

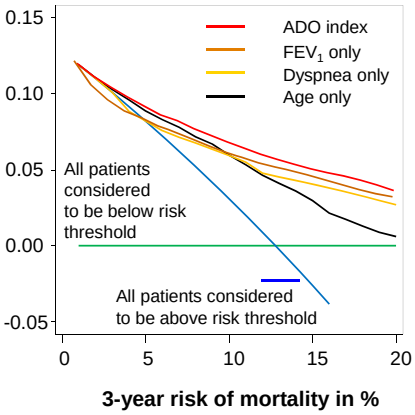
Calibration



Discrimination

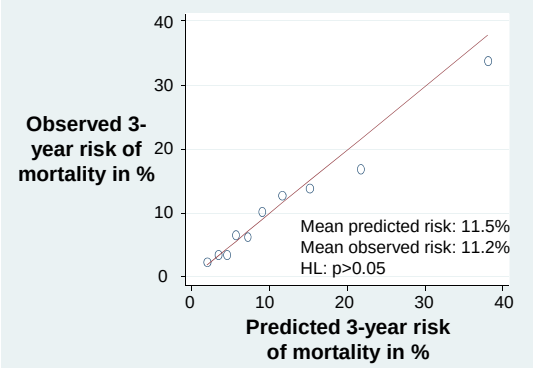
Area under curve
0.70 (0.67-0.73)

Net Benefit
(Difference between correctly and incorrectly classified patients)



Sensitivity III: Performance of updated ADO index in validation cohort after excluding subjects with a physician diagnosis of asthma from cohorts where only pre-bronchodilator spirometry was available (3,545 subjects included, 144 subjects excluded)

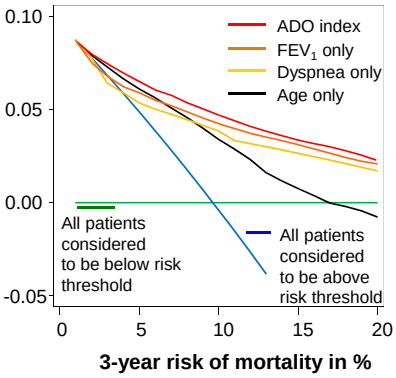
Calibration



Discrimination

Area under curve
0.73 (0.70-0.76)

Net Benefit
(Difference between correctly and incorrectly classified patients)



Appendix 9. Appendix references

1. Puhan MA, Garcia-Aymerich J, Frey M, et al. Expansion of the prognostic assessment of patients with chronic obstructive pulmonary disease: the updated BODE index and the ADO index. *Lancet* 2009;**374**:704-11.
2. Sobradillo P, Iriberri M, Gomez B, et al. Validation of bode index as a predictor of mortality in COPD patients. In: 18th Annual Congress of the European Respiratory Society. Berlin: European Respiratory Society; 2008:P531.
3. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1991;**1**:263-76.
4. Appleyard M, Hansen A, Schnohr P. The Copenhagen City Heart Study: a book of tables with data from the first examination (1976-78) and a five years follow-up (1981-1983). *Scand J Soc Med* 1989;**170**:1-160.
5. Carpenter MA, Crow R, Steffes M, et al. Laboratory, reading center, and coordinating center data management methods in the Jackson Heart Study. *Am J Med Sci* 2004;**328**:131-44.
6. Taylor HA, Jr., Wilson JG, Jones DW, et al. Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. *Ethn Dis* 2005;**15**:S6-4-17.
7. Anthonisen NR, Skeans MA, Wise RA, et al. The effects of a smoking cessation intervention on 14.5-year mortality: a randomized clinical trial. *Ann Intern Med* 2005;**142**:233-9.
8. Connett JE, Kusek JW, Bailey WC, et al. Design of the Lung Health Study: a randomized clinical trial of early intervention for chronic obstructive pulmonary disease. *Control Clin Trials* 1993;**14**:3S-19S.
9. Fishman A, Martinez F, Naunheim K, et al. A randomized trial comparing lung-volume-reduction surgery with medical therapy for severe emphysema. *N Engl J Med* 2003;**348**:2059-73.
10. Garcia-Aymerich J, Gomez FP, Anto JM. Phenotypic Characterization and Course of Chronic Obstructive Pulmonary Disease in the PAC-COPD Study: design and methods. *Arch Bronconeumol* 2009;**45**:4-11.
11. Garcia-Aymerich J, Gomez FP, Benet M, et al. Identification and prospective validation of clinically relevant chronic obstructive pulmonary disease (COPD) subtypes. *Thorax* 2011;**66**:430-7.
12. Menezes AM, Perez-Padilla R, Jardim JR, et al. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. *Lancet* 2005;**366**:1875-81.

13. Domingo-Salvany A, Lamarca R, Ferrer M, et al. Health-related quality of life and mortality in male patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2002;**166**:680-5.
14. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for Diagnosis, Management, and Prevention of COPD. http://wwwgoldcopdcom/GuidelinesResourcesasp?l1=2&l2=0 2010.
15. Moons KG, Donders RA, Stijnen T, et al. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 2006;**59**:1092-101.
16. Steyerberg E. Applications of Prediction Models. In: Clinical Prediction Models - A Practical Approach to Development, Validation, and Updating. New York: Springer; 2010.
17. Donders AR, van der Heijden GJ, Stijnen T, et al. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;**59**:1087-91.
18. Schafer JL. Analysis of Incomplete Multivariate Data. Boca Raton, FL: Chapman & Hall/CRC; 1997.
19. Steyerberg EW, Borsboom GJ, van Houwelingen HC, et al. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med* 2004;**23**:2567-86.
20. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;**49**:1373-1379.
21. Harrell FE Jr, Lee KL, Califf RM, et al. Regression modelling strategies for improved prognostic prediction. *Stat Med* 1984;**3**:143-52.
22. Vergouwe Y, Steyerberg EW, Eijkemans MJ, et al. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005;**58**:475-83.
23. Steyerberg EW: Study design for prediction models. In Clinical prediction models. Volume Chapter 3. New York: Springer; 2008
24. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;**338**:b375.
25. Royston P, Moons KG, Altman DG, et al. Prognosis and prognostic research: Developing a prognostic model. *BMJ* 2009;**338**:b604.
26. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;**19**:453-73.
27. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;**130**:515-24.

28. Sullivan LM, Massaro JM, D'Agostino RB, Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med* 2004;**23**:1631-60.
29. Vickers AJ. Decision analysis for the evaluation of diagnostic tests, prediction models and molecular markers. *Am Stat* 2008;**62**:314-20.
30. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;**26**:565-74.
31. Greenland S. Principles of multilevel modelling. *Int J Epidemiol* 2000;**29**:158-67.
32. Austin PC, Goel V, van Walraven C. An introduction to multilevel regression models. *Can J Public Health* 2001;**92**:150-4.
33. Urbach DR, Austin PC. Conventional models overestimate the statistical significance of volume-outcome associations, compared with multilevel models. *J Clin Epidemiol* 2005;**58**:391-400.