



# Development of a decision tree to assess the severity and prognosis of stable COPD

C. Esteban\*, I. Arostegui<sup>#</sup>, J. Moraza\*, M. Aburto\*, J.M. Quintana<sup>†</sup>,  
J. Pérez-Izquierdo\*, S. Aizpiri\* and A. Capelastegui\*

**ABSTRACT:** The aim of this study was to develop and validate a new method: a classification and regression tree (CART) based on easily accessible measures to predict mortality in patients with stable chronic obstructive pulmonary disease (COPD).

This was a prospective study of two independent prospective cohorts: a derivation cohort with 611 recruited patients and a validation cohort with 348 patients, all followed for 5 yrs. CART analysis was used to predict 5-yr mortality risk using the following covariates from the derivation cohort: age, % predicted forced expiratory volume in 1 s (FEV<sub>1</sub>), dyspnoea, physical activity, general health and number of hospital admissions for COPD exacerbations in the previous 2 yrs.

Age ( $\geq 75$  or  $< 75$  yrs) provided the first branch of the COPD-CART. The highest mortality risk (0.74) was seen in patients  $> 75$  yrs of age with higher levels of dyspnoea and FEV<sub>1</sub>  $< 50\%$  pred. Patients with the lowest risk of 5-yr mortality (0.04) were  $< 55$  yrs of age with FEV<sub>1</sub>  $> 35\%$  pred and one or no recent hospitalisations for COPD exacerbations.

A simple decision tree that uses variables commonly gathered by physicians can provide a quick assessment of the severity of the disease, as measured by the risk of 5-yr mortality.

**KEYWORDS:** Chronic obstructive pulmonary disease, decision trees, severity score

Determining the likelihood that a patient with chronic obstructive pulmonary disease (COPD) will experience an adverse event, such as hospitalisation or death, can help guide a clinician in establishing a hierarchy of risk that is useful for designing a therapeutic strategy. Until 2004, when *CELLI et al.* [1] published the BODE (body mass index (BMI), obstruction, dyspnoea and exercise) index, the way to determine the severity and prognosis of patients with COPD was based largely on the forced expiratory volume in 1 s (FEV<sub>1</sub>). Since then, other multidimensional scales have been developed. These new scales aim to improve the prognostic ability of the BODE index while simultaneously making them easier to use in daily clinical practice [2–4].

The BODE index and similar instruments are typically created using linear and additive models. An alternative approach is to use a classification and regression tree (CART) [5]. Creating a CART centres around statistically optimal clustering of patients based on predictors that produce maximum separation among the subgroups and minimum variability within these subgroups with respect to the outcome. The result is a representation, in tree form, of a decision rule with a hierarchical sequential structure that can easily be applied in clinical practice. CART analysis is gaining widespread

popularity in the health sciences as a screening method for variables, for summarising large datasets, or as a means of devising prediction rules for rapid and repeated evaluation [6, 7].

To date, CART analysis has not been applied to gauging the severity of COPD. The aim of our study was to develop and validate a CART to predict mortality in patients with stable COPD.

## Methods

We surveyed all patients previously diagnosed with COPD who regularly visited five outpatient clinics affiliated with a teaching hospital in the interior district of Biscay, Spain with a catchment area of 300,000 rural and urban inhabitants. We prospectively assembled two cohorts of patients with stable COPD. The derivation cohort was recruited from February 1998 to February 1999. The validation cohort was recruited from January 2003 to January 2004. Patients in each cohort were followed for 5 yrs.

Patients were consecutively included in one of the cohorts if they had been diagnosed with COPD for  $\geq 6$  months, had been under treatment for  $\geq 6$  months and had been clinically stable (no increase in respiratory symptoms or changes in treatment) for the 6 weeks prior to inclusion. Other inclusion criteria were FEV<sub>1</sub>  $< 80\%$  predicted, with

## AFFILIATIONS

\*Service of Pneumology,  
<sup>#</sup>Dept of Applied Mathematics and Statistics and Operations Research, University of the Basque Country, CIBERESP, Bilbao, Spain, and  
<sup>†</sup>Research Unit, CIBER Epidemiología y Salud Pública (CIBERESP), Hospital de Galdakao-Usansolo, Galdakao.

## CORRESPONDENCE

C. Esteban  
Service of Pneumology  
Hospital de Galdakao-Usansolo  
Barrio Labeaga s/n.48960 Galdakao  
Bizkaia  
Spain  
E-mail: cristobal\_esteban@yahoo.es

## Received:

Dec 08 2010

Accepted after revision:

April 24 2011

First published online:

May 12 2011

European Respiratory Journal  
Print ISSN 0903-1936  
Online ISSN 1399-3003

a FEV<sub>1</sub>/forced vital capacity ratio of <70%, and a bronchodilation test with FEV<sub>1</sub> change of <200 mL and <15% of the baseline value. Patients were not eligible for the study if they had been diagnosed with asthma, had extensive residual pulmonary tuberculosis, or a neoplastic process, or were suffering from any problem that could prevent effective collaboration.

Variables evaluated for the development of the COPD-CART were those previously included in the BODE index and in two of its offshoots, the HADO (health, activity, dyspnoea and obstruction) score [2] and the ADO (age, dyspnoea and airway obstruction) index [3]: dyspnoea, FEV<sub>1</sub> % pred, level of physical activity (PA), overall health status, age, BMI and 6-min walking test (6MWT). We also evaluated the impact of the number of COPD-related admissions to the hospital in the prior 2 yrs.

Patients' perceptions of their dyspnoea were categorised using a four-degree scale adapted from that of FLETCHER *et al.* [8] (degree 1: "dyspnoea only with intense and strenuous exercise"; degree 2: "capable of walking at the same pace as other people my age on the level"; degree 3: "capable of walking on the level at my own speed without dyspnoea but incapable of walking at the same pace as people my own age"; and degree 4: "dyspnoea after walking slowly for 100 m" or "dyspnoea when resting or after slight effort, such as getting dressed").

Spirometry was conducted following Spanish Respiratory Society criteria [9]. Predicted values were those of the European Community for Steel and Coal [10]. COPD severity was defined according to the four levels established by the American Thoracic Society (ATS) (FEV<sub>1</sub> <35, 35–49, 50–64 or ≥65% pred) [11].

In the baseline personal interview, patients were asked about the types of PA they performed. Special emphasis was placed on walking, with questions about time spent walking and distance covered. The level of PA was defined as the time patients spent walking during their leisure time, and was classified as very low (does not leave the house, life is limited to the bed or armchair or to doing some domestic chores, or leaves the house but walks <100 m), low (engaging in light physical activity, such as walking, for <2 h·week<sup>-1</sup>), medium (engaging in light physical activity, such as walking, for 2–4 h·week<sup>-1</sup>) or high (engaging in light physical activity, such as walking, for >4 h·week<sup>-1</sup>), as in previous studies [12, 13].

The 6MWT was carried out according to ATS guidelines [14]. Two tests were conducted with a 30-min break in between; the best result was selected.

BMI was calculated by dividing weight in kilograms by the square of height in metres.

Health-related quality of life (HRQoL) was assessed using the St George's Respiratory Questionnaire (SGRQ), an instrument specifically designed to assess HRQoL in patients with COPD [15]. We used the version that has been translated into and validated in Spanish [16].

Overall health was assessed by means of a single question: "In general, how would you characterise your health?" The four possible answers were bad, fair, good and very good or excellent.

The study was approved by the research committee of our hospital (Hospital de Galdakao-Usansolo, Galdakao, Spain). Patients provided informed consent to take part in the study.

### Main outcomes

Vital status was initially determined by telephone calls made to patients or their next of kin. All reported deaths and dates of deaths were confirmed by reviewing medical reports, examining the hospital database and public death registries, or both. Deaths were considered confirmed if the record matched the subject for name, sex and date of birth. In all cases, the cause of death was based on the hospital reports and public death registries.

### Statistical analysis

Comparisons between the derivation and the validation sample were performed using Pearson's Chi-squared test and the unpaired t-test. Univariate analysis between mortality and the covariates was performed using Pearson's Chi-squared test. Logistic regression was applied to obtain the area under the receiver operating characteristic curve (AUC) for each covariate, with 5-yr mortality as the dependent variable [17].

CART analysis was used to predict mortality risk in a 5-yr period with the following covariates: age, FEV<sub>1</sub>, dyspnoea, physical activity, general health and number of hospital admissions in the previous 2 yrs. CART-COPD was calculated by applying a recursive partitioning algorithm to the data. The algorithm tests for statistically optimal clustering of patients based on predictors that produce maximum separation among the clusters and minimum variability within these clusters with respect to the outcome. The predictor causing the largest separation was situated at the top of the tree, followed by the predictor causing the next largest separation and so on. Splitting continues until the subgroups reach a minimum size or until no improvement can be obtained. Risk of death was the outcome of interest. Therefore, patients were clustered, depending upon the values of the predictors, on groups with similar predicted risk of death. The minimal and maximal scores are the minimal and maximal predicted risk of death by the CART model.

To test the predictive ability of the CART model created from the derivation cohort, we applied it to patients in the validation cohort. The predictive ability of the CART was evaluated by sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and AUC. These measures were also calculated in the validation sample for the BODE, HADO and ADO indices. Various cut-off points were used to estimate the sensitivity, specificity, PPV, NPV and AUC changes according to the selected cut-off point for each instrument. Those cut-off points were chosen as they were some of those originally considered by the authors of the scores, as they complemented previous studies.

To test the calibration, we applied the Hosmer–Lemeshow goodness-of-fit test. As the risk of death predicted by the CART-COPD was subject to prediction error, predicted risk of death at each node was estimated by a random value at the 95% confidence interval of the predicted risk of death given by the CART-COPD at that node. Therefore, the estimated risk of death for each patient by the CART-COPD was compared to the observed mortality rate by the Hosmer–Lemeshow test. Calibration was also evaluated for the HADO, BODE and ADO scores by the Hosmer–Lemeshow test. The AUCs of the three severity scores were compared to the AUC of the CART model. ANOVA was performed to compare BODE, HADO, ADO and

**TABLE 1** Description of derivation and validation cohorts

	Derivation cohort	Validation cohort	p-value
<b>Subjects n</b>	611	348	
<b>Sex</b>			0.03
Male	597 (97.7)	331 (95.1)	
Female	14 (2.3)	17 (4.9)	
<b>Age yrs</b>	66.7±8.5	68.0±8.5	0.02
<b>BMI kg·m<sup>-2</sup></b>	27.8±4.3	28.2±4.6	0.16
<b>PA</b>			0.16
Very low	95 (15.6)	37 (10.6)	
Low	100 (16.4)	66 (19.0)	
Medium	229 (37.5)	140 (40.2)	
High	187 (30.5)	105 (30.2)	
<b>FEV<sub>1</sub> % pred</b>			<0.001
<35	101 (16.5)	29 (8.3)	
35–49	206 (33.7)	86 (24.7)	
50–64	192 (31.4)	138 (39.7)	
≥65	112 (18.3)	95 (27.3)	
<b>Level of dyspnoea</b>			<0.001
4	28 (4.6)	31 (8.9)	
3	233 (38.1)	105 (30.2)	
2	306 (50.1)	159 (45.7)	
1	44 (7.2)	53 (15.2)	
<b>General health</b>			0.26
Bad	70 (11.5)	33 (9.5)	
Fair	322 (52.7)	177 (50.9)	
Good	201 (32.9)	120 (34.5)	
Very good/excellent	18 (3.0)	18 (5.2)	
<b>Hospitalisations for COPD in preceding 2 yrs n</b>			0.90
0	446 (73.0)	247 (71.0)	
1	111 (18.2)	68 (19.5)	
>1	54 (8.8)	33 (9.5)	
<b>Mortality</b>	166 (27.2)	112 (32.2)	0.10
<b>BODE</b>		2.87±1.86	
<b>ADO</b>		3.70±1.47	
<b>HADO</b>	6.22±2.09	6.78±2.39	<0.001

Data are presented as n (%) or mean±SD, unless otherwise stated. BMI: body mass index; PA: physical activity; FEV<sub>1</sub>: forced expiratory volume in 1 s; % pred: % predicted; COPD: chronic obstructive pulmonary disease; BODE: BMI, obstruction, dyspnoea and exercise; ADO: age, dyspnoea and airflow obstruction; HADO: health, activity, dyspnoea and obstruction.

SGRQ across CART quartiles of mortality risk. The Scheffe test was performed for multiple comparisons.

All effects were considered statistically significant at an  $\alpha$ -value of 0.05. Data analyses were performed using SAS for Windows version 9.2 (SAS Institute Inc., Cary, NC, USA) [18] and R version 2.11.1 (R Development Core Team, Vienna, Austria) [19].

## RESULTS

The derivation cohort included 611 patients; the validation cohort included 348 patients. Statistically significant differences

**TABLE 2** Univariate relationship between mortality and the studied covariables

Variable	Subjects	p-value	AUC
<b>Age yrs</b>		<0.0001	0.64
≤59	13 (11.3)		
60–69	57 (23.8)		
70–79	87 (36.4)		
≥80	9 (50.0)		
<b>Sex</b>		0.77	0.50
Male	163 (27.3)		
Female	3 (21.4)		
<b>Dyspnoea</b>		<0.0001	0.65
4	16 (57.1)		
3	86 (36.9)		
2	63 (20.6)		
1	1 (2.3)		
<b>FEV<sub>1</sub> % pred</b>		<0.0001	0.65
<35	51 (50.5)		
35–49	60 (29.1)		
50–64	40 (20.8)		
≥65	15 (13.4)		
<b>PA</b>		<0.0001	0.66
Very low	45 (47.4)		
Low	33 (33.0)		
Moderate	65 (28.4)		
High	23 (12.3)		
<b>Hospitalisations for COPD in preceding 2 yrs n</b>		<0.0001	0.59
0	78 (20.2)		
1	40 (30.5)		
>1	48 (51.6)		
<b>General health</b>		0.0003	0.60
Bad	32 (45.7)		
Fair	91 (28.3)		
Good	40 (19.9)		
Very good/excellent	3 (16.7)		
<b>BMI kg·m<sup>-2</sup></b>		0.047	0.52
>21	154 (26.4)		
≤21	12 (44.4)		

Data are presented as n (%), unless otherwise stated. AUC: area under the receiver operating characteristic curve; FEV<sub>1</sub>: forced expiratory volume in 1 s; % pred: % predicted; PA: physical activity; COPD: chronic obstructive pulmonary disease; BMI: body mass index.

between the cohorts were observed for sex, age, dyspnoea, FEV<sub>1</sub> % pred and previous hospital admissions for COPD exacerbations in the preceding 2 yrs. In general, patients in the validation cohort had slightly more severe COPD than those in the derivation cohort (table 1).

Associations between health variables and 5-yr mortality were evaluated in a univariate analysis (table 2). All of the variables except sex were statistically significantly related to mortality. The associations were highly significant except for BMI, which was just within the limit of significance.

Based on the results of the univariate analysis, we included the following variables in the COPD-CART: dyspnoea, FEV<sub>1</sub> %

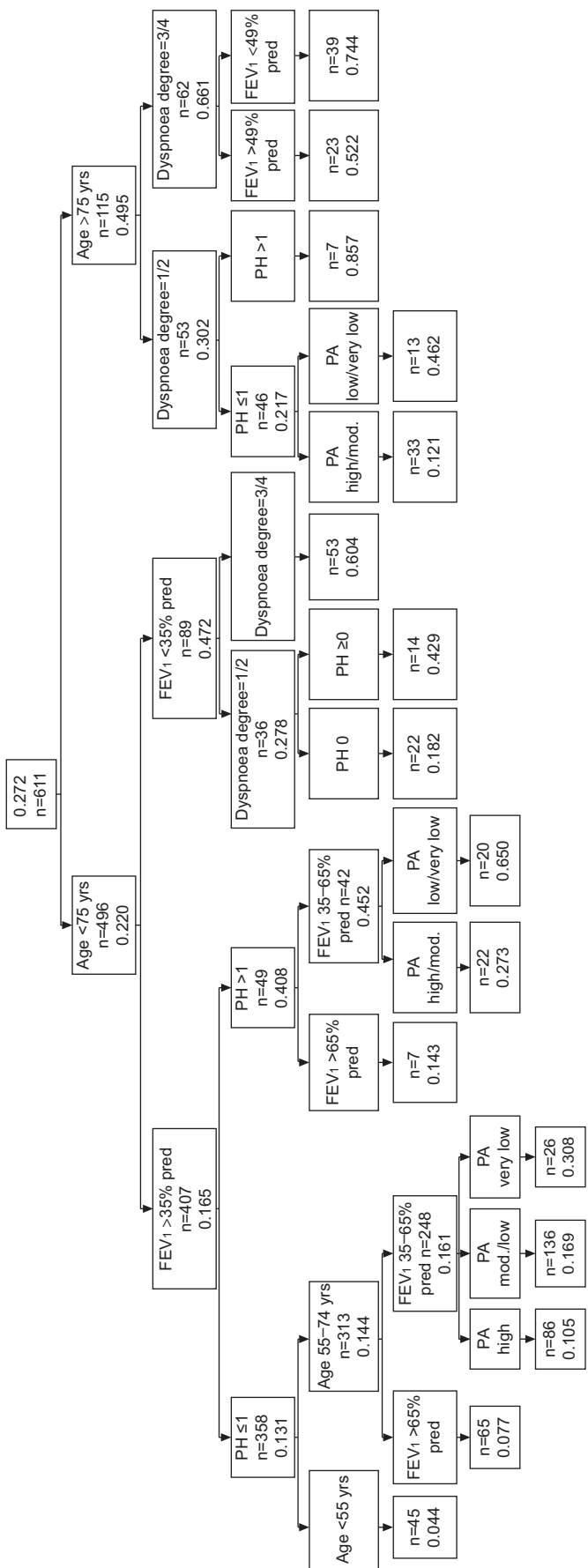


FIGURE 1. Decision tree. FEV<sub>1</sub>: forced expiratory volume in 1 s; % pred: % predicted; PH: prior hospitalisations; PA: physical activity.

pred, level of PA, overall health status, age and hospital admissions for COPD exacerbations in the preceding 2 yrs. All the variables were also selected as statistically significant by CART, except for overall health status. The CART created with data from the derivation cohort is shown in figure 1. Patients aged ≥75 yrs had a high and increasing risk of mortality depending on their level of dyspnoea, FEV<sub>1</sub> % pred and hospital admissions for COPD. The exception in this age group was patients with a moderate-to-high level of PA, who had a moderate risk of death (0.12). Patients aged <75 yrs followed various paths. Those with severe flow limitation (FEV<sub>1</sub> <35% pred) had a significant risk of death, except patients with only moderate-to-mild dyspnoea (degree 1 or 2) and no hospitalisations in the previous 2 yrs for COPD exacerbations. Patients with FEV<sub>1</sub> of 35–64% pred and two or more prior admissions for COPD exacerbations had a high risk of mortality depending on their level of PA (0.27 and 0.65 for high/moderate and low/very low PA, respectively), but when FEV<sub>1</sub> was ≥65% pred, the risk of death was lower (0.14). For those with no or one previous hospitalisation for COPD exacerbations, the risk depended on age: patients <55 yrs had the lowest mortality risk (0.04), while those between 55 and 74 yrs of age with FEV<sub>1</sub> >64% pred also had low mortality risk (0.07). Among patients with FEV<sub>1</sub> between 35 and 64% pred, mortality risk depended on the level of PA, ranging from a mortality risk of 0.1 among those with a high level of PA to 0.31 among those with a low level of PA.

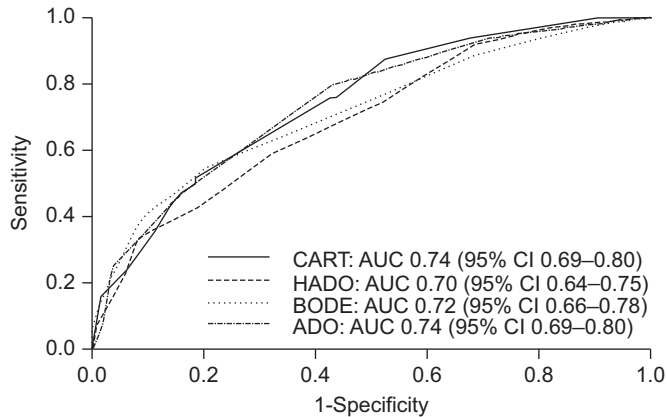
Figure 2 presents the AUC for the BODE, ADO and HADO scores as well as for the COPD-CART decision tree. The CART had the highest AUC (0.74), although the four values were not statistically different.

We compared the sensitivity, specificity and AUC of the three previously published COPD severity indices with the COPD-CART using three different cut-off points for each (table 3). The COPD-CART provided acceptable sensitivity or specificity, depending on the cut-off point chosen, and was similar to or better than the other indices. AUCs for the different cut-off points were 0.66–0.67 for the COPD-CART, which was slightly superior to the other rules in most cases.

To validate the COPD-CART, we compared its quartiles of severity with the three other indices and the total SGRQ score (table 4). Statistically significant differences were observed between the COPD-CART quartiles in the BODE and ADO scores, while differences in the HADO scores were observed between the two most severe categories compared with all the others. With the SGRQ scores, differences were found between the two more severe categories and the two more moderate.

**DISCUSSION**

A CART built from five easily obtained parameters in clinical practice provides an accurate decision model, expressed as a tree, for 5-yr mortality among patients with stable COPD. Patient age (≥75 or <75 yrs) provided the first branch. After age, in order of importance in the next branch, came FEV<sub>1</sub> % pred and dyspnoea status, followed by the number of admissions for COPD exacerbations in the previous 2 yrs and level of PA. Such a decision tree may help physicians estimate prognoses for their COPD patients and potentially influence clinical decision making.



**FIGURE 2.** Receiver operating characteristic curves for the classification and regression tree (CART), and the HADO (health, activity, dyspnoea and obstruction), BODE (body mass index, obstruction, dyspnoea and exercise) and ADO (age, dyspnoea and obstruction) scores. AUC: area under the curve.

In the COPD-CART, age established a distinct cut-off point (age <75 versus  $\geq$ 75 yrs) to give significantly different probabilities of mortality. This is consistent with previous studies that have established age as a predictor of mortality [20–23]. In the ADO index, for example, age was included with FEV<sub>1</sub> % pred and dyspnoea as an important predictive variable [3]. In our cohort of 611 COPD patients, age was the second most important predictor of all-cause mortality and the third most important predictor for respiratory mortality [22].

Guidelines for the management of COPD [24] highlight the importance of FEV<sub>1</sub> in the diagnosis, treatment and follow-up of patients with COPD. Several studies have established that the level of airway obstruction is related to disease prognosis [20–22]. It is no surprise, then, that FEV<sub>1</sub> was an important predictor of mortality and provided a branch of the COPD-CART.

In a study by NISHIMURA *et al.* [23], dyspnoea proved to be a more effective predictor of mortality than FEV<sub>1</sub> % pred among COPD patients. This finding was not replicated by our group; we found that dyspnoea was a more powerful predictor than FEV<sub>1</sub> % pred for all-cause mortality but not for respiratory mortality, while FEV<sub>1</sub> % pred was a stronger predictor than dyspnoea for respiratory mortality [22]. Thus, it appears that both FEV<sub>1</sub> % pred and dyspnoea are good predictors, but for different types of mortality. The importance of these measurements is underscored by the fact that FEV<sub>1</sub> and dyspnoea are the only two factors included in all four multidimensional prognostic scores [1–4].

In relation to the choice of previous hospital admission, SOLER-CATALUÑA *et al.* [25] demonstrated that the risk of mortality among patients with COPD increased with the frequency of severe exacerbations requiring hospital admission. When they replaced 6MWT in the BODE index with severe exacerbations, the new BODEx index had similar prognostic capacity for overall mortality (C statistic 0.74, 95% CI 0.65–0.83) compared with the original BODE index (0.75, 95% CI 0.66–0.84) [4].

Physical activity has previously been shown to influence mortality among patients with COPD. In a population-based cohort, GARCIA-AYMERICH *et al.* [12] showed that even slight PA (equivalent to walking or bicycling for 2 h·week<sup>-1</sup>) reduced the

**TABLE 3** Ability to predict 5-yr mortality measured by the area under the receiver operating characteristic curve (AUC) and confidence interval in the validation sample for the chronic obstructive pulmonary disease classification and regression tree (CART) decision tree and the three other severity measures<sup>#</sup>

Severity tool	Sensitivity	Specificity	NPV	PPV	AUC	Calibration <sup>†</sup>
<b>CART</b>						0.372
$\geq 0.3$	47.3	83.9	77.0	58.2	0.66	
$\geq 0.2$	50.0	81.4	77.4	56.0	0.66	
$\geq 0.16$	75.9	57.2	83.3	45.7	0.66	
$\geq 0.12$	87.5	47.5	88.9	44.1	0.67	
<b>BODE</b>						0.692
$\geq 5$	39.3	91.1	76.0	67.7	0.65	
$\geq 4$	54.5	79.7	78.7	56.0	0.67	
$\geq 3$	69.6	57.6	80.0	43.8	0.64	
$\geq 2$	89.3	30.5	85.7	37.9	0.60	
<b>ADO</b>						0.888
$\geq 6$	25.0	96.2	73.0	75.7	0.61	
$\geq 5$	46.4	84.3	76.8	58.4	0.65	
$\geq 4$	79.5	57.2	85.4	46.8	0.68	
$\geq 3$	93.8	28.8	90.7	38.5	0.61	
<b>HADO</b>						0.138
$\leq 4$	33.9	91.5	74.5	65.5	0.63	
$\leq 5$	43.8	80.1	75.0	51.0	0.62	
$\leq 6$	58.9	67.8	77.7	46.5	0.63	
$\leq 7$	74.1	48.3	79.7	40.5	0.61	

NPV: negative predictive value; PPV: positive predictive value; BODE: body mass index, obstruction, dyspnoea and exercise; ADO: age, dyspnoea and obstruction; HADO: health, activity, dyspnoea and obstruction. <sup>#</sup>: n=348; <sup>†</sup>: Hosmer-Lemeshow test.

risk of respiratory mortality among patients with COPD. Our results also highlighted the importance in some cases of the level of PA in terms of mortality risk. Physical activity is also a component of the HADO score [2].

Our decision tree was built by recursive partitioning using CART. CART has some advantages over the other developed scores, some statistical and others in the way it can be used. Among the main statistical advantages are that CART is a nonparametric technique. Therefore, this method does not require specification of the nature of the relationship between predictors and outcome. Tree-based models allow complex interactions between the independent variables, which must be specified *a priori* in standard linear models. Interpretations of complex interactions are clear and often more easily understood than in other model constructions. The CART algorithm will itself identify the most significant variables and eliminate nonsignificant ones. CART can easily handle outliers and is less dependent on missing values in the independent variables. The simplicity of using the CART is also worth highlighting: the user merely has to follow the branches relevant to the patient under study. It provides a global vision not only of what can be expected for a particular patient but also what can be expected if

**TABLE 4** Values of three severity scores and one respiratory-specific quality of life score in the validation sample (n=348) for decision tree nodes classified by quartiles generated by the classification and regression tree (CART) analysis

CART quartiles of estimated mortality	Subjects n	Mortality rate	BODE	ADO	HADO	SGRQ
<0.10	83	8.43	1.36 <sup>#,*,†</sup> (1.11–1.61)	2.18 <sup>#,*,†</sup> (1.98–2.38)	8.46 <sup>#,†</sup> (8.03–8.89)	31.54 <sup>#,†</sup> (27.16–35.91)
0.10–0.15	79	25.32	2.24 <sup>†,‡,§</sup> (1.99–2.49)	3.30 <sup>†,‡,§</sup> (3.11–3.50)	8.18 <sup>#,†</sup> (7.88–8.47)	30.97 <sup>#,†</sup> (26.75–35.18)
0.16–0.29	95	33.68	3.06 <sup>#,†,§</sup> (2.80–3.33)	3.98 <sup>#,†,§</sup> (3.79–4.17)	6.43 <sup>#,†,§</sup> (6.13–6.73)	41.94 <sup>#,§</sup> (37.99–45.89)
≥0.30	91	58.24	4.59 <sup>#,†,§</sup> (4.19–5.00)	5.12 <sup>#,†,§</sup> (4.86–5.39)	4.42 <sup>#,†,§</sup> (3.97–4.86)	49.43 <sup>#,§</sup> (45.41–53.44)

Data are presented as mean (95% CI), unless otherwise stated. Nodes have been classified into quartiles depending on the estimated mortality: <0.10, 0.10–0.15, 0.16–0.29, and ≥0.30. ANOVA was performed to test for statistical differences, and the Scheffe test was used for multiple comparisons for the BODE (body mass index, obstruction, dyspnoea and exercise), ADO (age, dyspnoea and obstruction), HADO (health, activity, dyspnoea and obstruction) and St George's Respiratory Questionnaire (SGRQ). Tests for statistical differences in mortality rate were performed with Pearson's Chi-squared test. #: statistically significant difference versus estimated mortality 0.10–0.15 group; †: statistically significant difference versus estimated mortality 0.16–0.29 group; ‡: statistically significant difference versus estimated mortality ≥0.30 group; §: statistically significant difference versus estimated mortality <0.10 group.

the patient's status changes. Among the disadvantages of the CART are that the clusters created may not necessarily agree with established diagnostic combinations or have any rational clinical meaning. However, scales developed using logistic regression models (as for the BODE, HADO and ADO indices) may use continuous variables; all relevant variables are considered in the scores and they will have more power, as they always use the whole database.

Other authors have used this kind of recursive partitioning-based model to develop clinical practice guidelines for high-dose chemotherapy and bone marrow transplant in chronic myelogenous leukaemia, compared to ANOVA [26] and logistic regression [27]. McCONNOCHIE *et al.* [28] compared classification trees to logistic regression for developing prediction rules and concluded that the use of both techniques is complementary. The three previously developed scores for COPD severity to which we compared the COPD-CART were built using logistic regression.

The ability of the COPD-CART to predict mortality, measured by AUC, was quite similar to that of the BODE index, HADO score, ADO index and total SGRQ score, supporting the validity of the COPD-CART.

Sensitivities for the COPD-CART for mortality risk ranged from 47.3% to 87.5%, depending on the cut-off points chosen, and specificities ranged from 83.9% to 47.5%. This compares favourably with the sensitivities and specificities of the BODE, ADO and HADO indices, and as it had quite similar results and, even better, prognostic ability, our tool seems to be at least similar to most of the previous tools.

Besides the previously cited strengths, our study, with two large samples of patients with stable COPD, which allowed us to validate our proposal, outlines the importance of some well-known variables (age, FEV1 and dyspnoea) but also of others (PA and previous hospitalisations) not included in previous rules. These latter variables make a difference, in terms of mortality, in some cases.

Several limitations of our study should be noted. The decision tree was derived from and validated in patients from a single hospital, which could limit the generalisability of these results.

For example, both cohorts were predominately male, reflecting the history of smoking patterns in northern Spain, where smoking was traditionally restricted to males.

Another question is whether the CART results might be applied to other cohorts in lower-risk settings, such as primary care. The fact that in our two cohorts, ≥50% of patients had a FEV1 >50% pred means that an important proportion of the patient sample was very similar to patients managed in primary care.

PA is a variable that is sometimes difficult to measure and at risk of measurement error. The effect of measurement error for PA was checked, with a <5% risk of significant misclassification and a statistically nonsignificant change of mean predicted risk of death (p=0.27).

The AUC of the COPD-CART is merely moderate, similar to those of other COPD indices. Thus, other variables are needed to improve the predictive capacity of the decision tree. One possibility is BMI <20 kg·m<sup>-2</sup>, which has been shown to constitute an independent predictor of COPD-related mortality [29]. In our study, some other variables that were statistically nonsignificant in the univariate analysis, but considered in other COPD prediction rules, were considered: BMI and the general health of the patient. None of them entered any branch of the developed trees. The main limitation of decision tree models is that including higher-order interactions without considering the main effects could lead to spurious relationships between predictors and overestimate the effect of some predictors. This is sometimes referred to as estimation bias [30]. However, our study validates the predictive capacity of the COPD-CART index in a different sample, which minimises that bias.

The utility of a clinical decision rule that is easily applied in daily practice cannot be understated. The main advantage of a decision tree is its simple representation. Some investigators believe that a decision tree represents how clinicians think, starting with the most important characteristic, followed by another characteristic depending on the answer of the first, and so on [24]. The main value of this predictive tree is that it can easily establish a prognosis without clinicians having to memorise the scores of different variables. The COPD-CART decision tree, which

employs measures generally gathered by physicians in the evaluation of COPD patients, provides a quick assessment of the severity of the disease, as measured by the risk of premature death, and by patient-perceived HRQoL, as measured by the SGRQ. The ability of this instrument to predict 5-yr mortality is as good as, or better than, that of other existing indices, but with greater ease of use in routine clinical practice. However, the COPD-CART must be tested in other settings and other populations, and other measures should be evaluated for inclusion to improve its predictive capacity, before it could become a widely used instrument.

### SUPPORT STATEMENT

This study was partially supported by grants from the Fondo de Investigación Sanitaria (PI97/0326 and PI020510), Departamento de Sanidad del Gobierno Vasco 200111002 and the Research Committee of the Hospital Galdakao-Usansolo.

### STATEMENT OF INTEREST

None declared.

### REFERENCES

- Celli BC, Cote CG, Marín JM, *et al.* The body-mass index, airflow obstruction, dyspnoea, and exercise capacity index in chronic obstructive pulmonary disease. *New Engl J Med* 2004; 350: 1005–1012.
- Esteban C, Quintana JM, Aburto M, *et al.* A simple score for assessing stable chronic obstructive pulmonary disease. *QJM* 2006; 99: 751–759.
- 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: 704–711.
- Soler-Cataluña JJ, Martínez-García MA, Sánchez LS, *et al.* Severe exacerbations and BODE index: two independent risk factors for death in male COPD patients. *Respir Med* 2009; 103: 692–699.
- Zhang H, Singer B. *Recursive Partitioning in the Health Sciences*, New York, Springer-Verlag, 1999.
- Santos M, Braga JU, Gomes RV, *et al.* Predictive factors for pneumonia onset after cardiac surgery in Rio de Janeiro, Brazil. *Infect Control Hosp Epidemiol* 2007; 28: 382–388.
- Hansell AL, Lam KA, Richardson S, *et al.* Medical event profiling of COPD patients. *Pharmacoepidemiol Drug Saf* 2004; 13: 547–555.
- Fletcher CM, Elmes PC, Wood CH. The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *BMJ* 1959; 2: 257–266.
- Sanchis J. Normativa para la espirometría forzada: Grupo de trabajo de la SEPAR para la práctica de la espirometría clínica. [Guidelines for spirometry]. *Arch Bronconeumol* 1989; 25: 132–142.
- Quanjer PH, Tammeling GJ, Cotes JE, *et al.* Lung volumes and forced ventilatory flows. Report working party standardization of lung function test, European Community for Steel and Coal. Official statement of the European Respiratory Society. *Eur Respir J* 1993; 16: Suppl. 16, 5–40.
- American Thoracic Society. Standardization of spirometry: 1994 update. *Am Rev Respir Dis* 1994; 152: 1107–1136.
- Garcia-Aymerich J, Lange P, Benet M, *et al.* Regular physical activity reduces hospital admission and mortality in chronic obstructive pulmonary disease: a population based cohort study. *Thorax* 2006; 61: 772–778.
- Schnohr P, Scharling H, Jensen JS. Changes in leisure time physical activity and risk of death: an observational study of 7000 men and females. *Am J Epidemiol* 2003; 158: 639–644.
- ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002; 166: 111–117.
- Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. *Respir Med* 1991; 85: 25–31.
- Ferrer M, Alonso J, Prieto L, *et al.* Validity and reliability of the St George's Respiratory Questionnaire after adaptation to a different language and culture: the Spanish example. *Eur Respir J* 1996; 9: 1160–1166.
- Hosmer DH, Lemeshow S. *Applied Logistic Regression*, 2nd Edn. New York: Wiley and Sons, 2000.
- SAS Institute Inc. *SAS Procedures Guide*, Version 6. Cary: SAS Institute, 1994.
- R development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2005.
- Anthonisen NR, Wright EC, Hodgkin JE. Prognosis in chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1986; 133: 14–20.
- Traver GA, Cline MG, Burrows B. Predictors of mortality in chronic obstructive pulmonary disease. A 15 year follow up study. *Am Rev Respir Dis* 1979; 119: 895–902.
- Esteban C, Quintana JM, Aburto M, *et al.* Predictors of mortality in patients with stable COPD. *J Gen Intern Med* 2008; 23: 1829–1834.
- Nishimura K, Takateru I, Tsukino M, *et al.* Dyspnoea is a better predictor of 5 year survival than airway obstruction in patients with COPD. *Chest* 2002; 121: 1434–1440.
- Celli BR, MacNee W, Agusti A, *et al.* Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004; 23: 932–946.
- Soler-Cataluña JJ, Martínez-García MA, Román Sánchez P, *et al.* Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. *Thorax* 2005; 60: 925–931.
- Gale RP, Park RE, Dubois RW, *et al.* Delphi-panel analysis of appropriateness of high-dose therapy and bone marrow transplant in chronic myelogenous leukemia in chronic phase. *Leuk Res* 1999; 23: 817–826.
- Wietlisbach V, Vader JP, Porchet F, *et al.* Statistical approaches in the development of clinical practice guidelines from expert panels. *Medical Care* 1999; 37: 785–797.
- McConnochie KM, Roghmann KJ, Pasternack J. Developing prediction rules and evaluating observation patterns using categorical clinical markers: two complementary procedures. *Med Decis Making* 1993; 13: 30–42.
- Soriano JB, Ancochea J, Miravittles M, *et al.* Recent trends in COPD prevalence in Spain: a repeated cross-sectional survey 1997–2007. *Eur Respir J* 2010; 36: 758–765.
- Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation and Updating*. New York, Springer-Verlag, 2009.