ERJ Express. **Published on May 12, 2011 as doi: 10.1183/09031936.00189010**

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

Authors: Cristóbal Esteban MD, Inmaculada Arostegui, MCs, PhD; Javier Moraza, MD; Myriam Aburto MD, PhD; Jose M. Quintana, MD, PhD; Julio Pérez-Izquierdo MD; Susana Aizpiri, MD; Alberto Capelastegui, MD, PhD.

From the Service of Pneumology (C. Esteban, J. Moraza J, M. Aburto, J Pérez-Izquierdo, S. Aizpiri , A. Capelastegui); Research Unit - CIBER Epidemiología y Salud Pública (CIBERESP) (J.M. Quintana) of Hospital de Galdakao-Usansolo, Galdakao, Bizkaia, Spain; Department of Applied Mathematics and Statistics and Operations Research, University of the Basque Country (UPV/EHU) - CIBER Epidemiología y Salud Pública (CIBERESP) (I. Aróstegui).

Correspondence and requests for reprints should be addressed to:
Dr. Cristóbal Esteban, Service of Pneumology, Hospital de Galdakao-Usansolo, Barrio Labeaga s/n.48960 Galdakao, Bizkaia, Spain.
Telephone number +34-944007002, Fax: +34-94 400 7132
e-mail: cristobal_esteban@yahoo.es

Running title: Classification tree for COPD severity and prognosis.

Abstract word count: 200.

Manuscript word count: 2977.

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 COPD.

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 years. CART analysis was used to predict 5-year mortality risk using the following covariates from the derivation cohort: age, forced expiratory volume in the first second as a percentage of predicted ($FEV_{1\%}$), dyspnea, physical activity, general health, and number of hospital admissions for COPD exacerbations in the prior 2 years.

Age ($\geq$75 years or <75) provided the first branch of the COPD-CART. The highest mortality risk (0.74) was seen in patients older than 75 with higher levels of dyspnea and with $FEV_{1\%}$ < 50%. Patients with the lowest risk of 5-year mortality (0.04) were those under age 55 years with $FEV_{1\%}$ > 35% and with 1 or no recent hospitalizations for COPD exacerbations.

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

Keywords: COPD; Decision trees; Severity score.

INTRODUCTION

Determining the likelihood that a patient with chronic obstructive pulmonary disease (COPD) will experience an adverse event such as hospitalization or death can help guide a clinician to establish a hierarchy of risk useful for designing a therapeutic strategy. Until 2004, when Celli et al published the BODE-index (Body mass index, Obstruction, Dyspnea, Exercise-index) [1], the way to determine the severity and prognosis of patients with COPD was based largely on the FEV_1 (forced expiratory volume in the first second). 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 centers 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 summarizing 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 5 outpatient clinics affiliated with a teaching hospital in the interior district of Bizkaia, 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 years.

Patients were consecutively included in one of the cohorts if they had been diagnosed with COPD for at least six months, had been under treatment for at least six months, and had been clinically stable (no increase in respiratory symptoms or changes in treatment) for the six weeks prior to inclusion. Other inclusion criteria were FEV_1 < 80% of the predicted value, with a quotient FEV_1/FVC < 70%, and a bronchodilation test with FEV_1 change <200 mL and under 15% of the baseline value. Patients were not eligible for the study if they had been diagnosed with asthma, had extensive residual pulmonary tuberculosis, 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 Health-Activity-Dyspnoea-Obstruction (HADO-score) [2] and the Age, Dyspnea, and airflow Obstruction (ADO-index): [3] dyspnea, $FEV_{1\%}$, level of physical activity, overall health status, age, body

mass index, and 6-minute walking test (6MWT). We also evaluated the impact of the number of COPD-related admissions to the hospital in the prior 2 years.

Patients' perceptions of their dyspnea were categorized using a 4-degree scale adapted from Fletcher: [8] degree 1 “dyspnea 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 dyspnea but incapable of walking at the same pace as people my own age;” and degree 4 “dyspnea after walking slowly for 100 meters” or “dyspnea when resting or after slight effort, such as getting dressed.”

Spirometry was conducted following Spanish Respiratory Society criteria. [9] Theoretical values were those prescribed by the European Community for Steel and Coal. [10] COPD severity was defined according to the four levels established by the American Thoracic Society (ATS) based on $FEV_{1\%}$. [11] The four categories were ordered from $FEV_{1\%}$ <35%; 35-49; 50-64 to ≥ 65%.

In the baseline personal interview, patients were asked about the types of physical activity (PA) they did. 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 (doesn’t leave the house, life is limited to the bed or armchair or to doing some domestic chores, or leaves the house but walks less than 100 meters), low (engaging in light physical

activity such as walking for less than 2 hours/week), medium (engaging in light physical activity such as walking for 2-4 hours/week), and high (engaging in light physical activity such as walking for more than 4 hours/week), as had been done in previous studies. [12, 13]

The 6MWT was carried out following ATS rules. [14] Two tests were conducted with a 30-minute break in between; the best result was selected.

Body mass index (BMI) was calculated by dividing weight in kilograms by the square of height in meters.

Health-related quality of life (HRQoL) was assessed using the St. George's Respiratory Questionnaire (SGRQ), an instrument designed specifically to assess HRQoL among 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 characterize your health?" The 4 possible answers were bad, fair, good and very good or excellent.

The study was approved by the research committee of our hospital. 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 on 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-square test and Student's t-test. Univariate analysis between mortality and the covariates was performed using Pearson's chi-square test. Logistic regression was applied to obtain the area under the receiving operating curve (AUC) for each covariate, with 5-year mortality as the dependent variable. [17]

CART analysis was used to predict mortality risk in a 5-year period with the following covariates: age, FEV_1, dyspnea, physical activity, general health, and number of hospital admissions in the previous 2 years. *CART-COPD was calculated by applying a recursive partitioning algorithm to the data. The algorithm looks 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, negative predictive value, and AUC. These measures were also calculated in the validation sample for the BODE index, the HADO index, and the ADO index. *Various cut points were used to estimate the sensitivity, specificity, PPV, NPV and AUC changes according to the selected cut point for each instrument. Those cut points were chosen by being some of the originally considered by the authors of the score because they complemented the previous.*

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 so 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 HADO, BODE and ADO scores by the Hosmer and Lemeshow's test. The AUCs of the three severity scores were compared to the AUC of the CART model. Analysis of variance was performed to compare BODE, HADO, ADO, and SGRQ across CART quartiles of mortality risk. The Scheffe test was performed for multiple comparisons.

All effects were considered statistically significant at $\alpha = 0.05$. Data analyses were performed using SAS for Windows version 9.2 (SAS Institute Inc., 1994) [18] and R (R development Core Team, 2005, version 2.11.1). [19]

RESULTS

The derivation cohort included 611 patients; the validation cohort included 348 patients. Statistically significant differences between the cohorts were observed for gender, age, dyspnea, $FEV_{1\%}$, and previous hospital admissions for COPD exacerbations in the preceding 2 years. 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-year mortality were evaluated in a univariate analysis (Table 2). All of the variables except gender 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: dyspnea, $FEV_{1\%}$, level of PA, overall health status, age, and hospital admissions for COPD exacerbations in the preceding 2 years. 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 years and older had a high and increasing risk of mortality depending on their level of dyspnea, $FEV_{1\%,}$ 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 under age 75 followed various paths. Those with severe flow limitation ($FEV_{1\%} < 35\%$) had a significant risk of death, except patients with only moderate to mild dyspnea (degree 1 or 2) and no hospitalizations in the previous 2 years for COPD

exacerbations. The risk of mortality for patients with $FEV_{1\%}$ between 35%-64% and prior admissions for COPD exacerbations of 2 or more had a high risk of mortality depending on the level of physical activity of the patients (0.27 and 0.65), except when $FEV_{1\%}$ was ≥65%, in which case the risk of death was lower (0.14). For those with 0 or 1 previous hospitalizations for COPD exacerbations, the risk depended on age: patients younger than 55 had the lowest mortality risk (0.04), While those between 55 and 74 years with $FEV_{1\%}$ > 64% also had low mortality risk (0.07). Among patients with FEV_1 between 35% and 64%, mortality risk depended on 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 index, ADO index, and HADO score 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-points for each (Table 3). The COPD-CART provides acceptable sensitivity or specificity, depending on the cut-point chosen, and is similar to or better than the other indices. AUCs for the different cut-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 with the total score on the SGRQ (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 to all the others. With the SGRQ scores, differences were found between the two more severe categories with the two more moderate.

DISCUSSION

A classification and regression tree built from five easily obtained parameters in clinical practice provides an accurate decision model, expressed as a tree, for 5-year mortality among patients with stable COPD. Patient age (75 years or greater, or under 75 years) provided the first branch. Next came $FEV_{1\%}$ and dyspnea status, followed by the number of admissions for COPD exacerbations in the previous 2 years and level of PA. Such a decision tree may help physicians estimate prognoses for their COPD patients and potentially influence clinical decision-making.

In the COPD-CART, age established a distinct cut-point (age <75 years vs. age $\geq$75 years) with significantly different probabilities of mortality. This is in keeping with previous studies that have established age as a predictor of mortality. [20-23] In the ADO index, for example, age was included with $FEV_{1\%}$ and dyspnea as an important predictive variable. [3] In a 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_1 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_1 was an important predictor of mortality and provided a branch of the COPD-CART.

In a study by Nishimura and colleagues, dyspnea proved to be a more effective predictor of mortality than $FEV_{1\%}$ among COPD patients. [23] This finding was not replicated by our group, we found that dyspnea was a more powerful predictor than $FEV_{1\%}$ for all-cause mortality but not for respiratory mortality, while $FEV_{1\%}$ was a stronger predictor than dyspnea for respiratory mortality. [22] Thus, it appears that both $FEV_{1\%}$ and dyspnea are good predictors but for different types of mortality. The importance of these measures is underscored by the fact that FEV_1 and dyspnea are the only two factors included in all four multidimensional prognostic scores. [1-4]

On relation to the choice of previous hospital admission, Soler et al. demonstrated that the risk of mortality among patients with COPD increases with the frequency of severe exacerbations requiring hospital admission. [25] 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 (C statistic 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. showed that even slight physical activity (equivalent to walking or bicycling for 2 hours per week) reduced the risk of respiratory mortality among patients with COPD. [12] Our results also highlighted the importance in some cases of the level of physical activity that the patient referred 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 are statistical and other 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 other model constructions. CART algorithm will itself identify the most significant variables and eliminate non-significant ones. CART can easily handle outliers and is less dependent on missing values in the independent variables. Then, from the use of the CART point of view it can be highlighted its simplicity to interpret the results since the reader just have to follow the branches that fulfil the patient under study characteristics. It provides with a global vision not only of what can be expected for a particular patient but also if the patient status changes. Among CART disadvantages are that the clusters created may not necessarily agree with established diagnostic combinations or have any rational clinical meaning. On the other hand, scales developed by logistic regression models (as the BODE, HADO and ADO) may use continuous variables, all relevant variables are considered in the score and will have more power since always uses the whole data base.*

Other authors have used this kind of recursive partitioning-based model to develop clinical practice guidelines for high-dose chemotherapy therapy and bone marrow transplant in chronic myelogenous leukemia compared to ANOVA [26] and to logistic

regression. [27] McConnochie et al. compared classification trees to logistic regression for developing prediction rules and concluded that the use of both techniques is complementary. [28] 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 score of the SGRQ, 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-points chosen, and specificities ranged from 83.9 to 47.5. This compares favorably with the sensitivities and specificities of BODE, ADO, HADO since having quite similar results, or even better, the prognosis ability of our tool seem to be, at least, similar to most of the previous tools.

Strengths of the study: Besides the previously cited strengths, our study, with 2 large samples of patients with stable COPD which allowed us to validate our proposal, outlines the importance of some well known variables (age, FEV1, dyspnea) but also of others (physical activity, previous hospitalizations) not included in previous rules. The latest variables make the difference in terms of mortality in some cases.

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

these results. For example, both cohorts were predominately male, reflecting the history of smoking patterns in northern Spain, which was traditionally restricted to men.

Another question is if CART results might be applied to other cohorts in lower risk settings such as primary care. The facts that in our two cohorts at least 50% of patients have a FEV1 higher than 50% make that a important proportion of the patients sample was very similar to that what is managed in primary care.

Physical activity is a variable sometimes difficult to measure and exposed to measurement error. The effect of measurement error for the physical activity was checked with a result of less than 5% of significant misclassification and a non statistically significant 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^2, which has been shown to constitute an independent predictor of COPD-related mortality. [29] In our study, s*ome other non statistically significant variables in the univariate analysis, but considered in other COPD prediction rules, were considered as the BMI and the general health of the patient. None of them entered in 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 relations 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 minimizes 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 memorize 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-year 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.

Acknowledgements

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

REFERENCES

1. Celli BC, Cote CG, Marín JM, Casanova C, Montes de Oca M, Mendez RA, Pinto Plata V, Cabral HJ. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. NEJM 2004; 350: 1005-1012.

2. Esteban C, Quintana JM, Aburto M, Moraza J, Capelastegui A. A simple score for assessing stable chronic obstructive pulmonary disease. QJM 2006; 99: 751-759.

3. Puhan MA, Garcia-Aymerich J, Frey M, ter Riet G, Antó JM, Agustí AG, Gómez FP, Rodríguez-Roisín R, Moons KG, Kessels AG, Held U. 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.

4. Soler-Cataluña JJ, Martínez-García MA, Sánchez LS, Tordera MP, Sánchez PR. Severe exacerbations and BODE index: two independent risk factors for death in male COPD patients. Respir Med 2009 ;103: 692-699.

5. Zhang H and Singer B (1999). *Recursive Partitioning in the Health Sciences*. New York: Springer-Verlag.

6. Santos M, Braga JU, Gomes RV, Werneck GL. Predictive factors for pneumonia onset after cardiac surgery in Rio de Janeiro, Brazil. Infect Control Hosp Epidemiol 2007; 28: 382-388.

7. Hansell AL, Lam KA, Richardson S, Visick G, Soriano JB. Medical event profiling of COPD patients. Pharmacoepidemiol Drug Saf 2004; 13:547-555.

8. 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.
9. Sanchis J. Normativa para la espirometría forzada: Grupo de trabajo de la SEPAR para la práctica de la espirometría clínica. Arch Bronconeumol 1989; 25: 132-142.
10. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. 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: 5-40.

11. Medical Section of the American Lung Association. Standardization of spirometry: 1994 update. Am Rev Respir Dis 1994; 152: 1107-1136.
12. Garcia-Aymerich J, Lange P, Benet M, Schnohr P, Antó JM. Regular physical activity reduces hospital admission and mortality in chronic obstructive pulmonary disease: a population based cohort study. Thorax 2006; 61:772-778.
13. Schnohr P, Scharling H, Jensen JS. Changes in leisure time physical activity and risk of death: An observational study of 7000 men and women. Am J Epidemiol 2003; 158: 639-.
14. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002; 166:111-117.

15. Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. Respir Med 1991: 85 (Suppl B): 25-31.

16. Ferrer M, Alonso J, Prieto L, Plaza V, Monsó E, Marrades R, Aguar MC, Khalaf A, Antó JM. 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.

17. Hosmer DH and Lemeshow S. (2000). *Applied Logistic Regression*, 2nd Ed. New York: Wiley and Sons.

18. SAS Institute Inc. (1994). *SAS Procedures Guide*, Version 6. Cary, N.C.: SAS Institute.

19. R development Core Team. (2005). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

20. Anthonisen NR, Wright EC, Hodgkin JE. Prognosis in Chronic obstructive pulmonary disease. Am Rev Respir Dis 1986; 133: 14-20.

21. 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.

22. Esteban C, Quintana JM, Aburto M, Moraza J, Egurrola M, España PP, Pérez-Izquierdo J, Capelastegui A. Predictors of mortality in patients with stable COPD. J Gen Intern Med 2008; 23:1829-1834.

23. Nishimura K, Takateru I, Tsukino M, Oga T. Dyspnea is a better predictor of 5 year survival than airway obstruction in patients with COPD. Chest 2002; 121: 1434-1440.

24. Celli BR, MacNee W, Agusti A, Anzueto, B. Berg, A.S. Buist, P.M.A. Calverley, N. Chavannes, T. Dillard, B. Fahy, A. Fein, J. Heffner, S. Lareau, P. Meek, F. Martinez, W. McNicholas, J. Muris, E. Austegard, R. Pauwels, S. Rennard, A. Rossi, N. Siafakas, B. Tiep, J. Vestbo, E. Wouters and R. ZuWallack. 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.

25. Soler-Cataluña JJ, Martinez-Garcia MA, Román Sánchez P, Salcedo E, Navarro M, Ochando R. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. Thorax 2005; 60: 925-931.

26. Gale RP, Park RE, Dubois RW, Herzing GP, Hocking WG, Horowitz MM, Keating A, Kenpin S, Linker CA, Schiffer CA, Wiernik PH, Weisdorf DJ, Rai KR. Delphi-panel analysis of appropriateness of high-dose therapy and bone marrow transplant in chronic myelogenous leukemia in chronic phase. Leukemia Research, 1999; 23; 817-26.

27. Wietlisbach V, Vader JP, Porchet F, Constanza MC, Burnand B. statistical approaches in the development of clinical practice guidelines from expert panels. Medical Care, 1999; 37: 785-97.

28. McConnochie KM, Roghmann KJ, Pasternack J. Developing prediction rules and evaluating observation patterns using categorical clinical markers: two complementary procedures. Medical Decision Making, 1993; 13: 30-42

29. . Soriano JB, Ancochea J, Miravitlles M, García-Río F, Duran E, Muñoz L, Jiménez-Ruiz CA, Masa JF, Viejo JL, Villasante C, Fernández-Fau L, Sánchez G, Sobradillo-Peña V. Recent trends in COPD prevalence in Spain: a repeated cross-sectional survey 1997-2007. Eur Respir J 2010; 36:758-765.

30. Steyerberg EW. (2009). Clinical Prediction Models: A Practical Approach to Development, Validation and Updating. New York: Springer-Verlag.

Table 1: Description of derivation and validation cohorts.

	Derivation cohort	Validation cohort	*p-value*
n	611	348	
Gender: *n* (%)			0.03
Male	597 (97.7)	331 (95.1)	
Female	14 (2.3)	17 (4.9)	
Age: mean (SD)	66.7 (8.5)	68.0 (8.5)	0.02
BMI: mean (SD)	27.8 (4.3)	28.2 (4.6)	0.16
Physical activity: *n* (%)			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)	
$FEV_{1\%}$ *n* (%)			<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)	
Level of dyspnea: *n* (%)			<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)	
General health: *n* (%)			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)	
Hospitalizations for COPD in the preceding 2 years: *n* (%)			0.90
0	446 (73.0)	247 (71.0)	
1	111 (18.2)	68 (19.5)	
>1	54 (8.8)	33 (9.5)	
Mortality: *n* (%)	166 (27.2)	112 (32.2)	0.10
BODE	--	2.87 (1.86)	
ADO		3.70 (1.47)	
HADO: mean (SD)	6.22 (2.09)	6.78 (2.39)	<0.001

BMI= body mass index; BODE = Body mass index, Obstruction, Dyspnea, Exercise (BODE) index; ADO = Age, Dyspnoea, and airflow Obstruction (ADO) index; HADO = Health-Activity-Dyspnoea-Obstruction (HADO) score; SD = standard deviation.

Table 2: Univariate relation between mortality and the studied covariables.

Variables	N (%)	*p-value*	Area under the ROC curve
Age (years)		<.0001	0.64
≤59	13 (11.3)		
60-69	57 (23.8)		
70-79	87 (36.4)		
≥ 80	9 (50.0)		
Gender		0.77	0.50
Male	163 (27.3)		
Female	3 (21.4)		
Dyspnea		<.0001	0.65
4	16 (57.1)		
3	86 (36.9)		
2	63 (20.6)		
1	1 (2.3)		
$FEV_{1\%}$		<.0001	0.65
<35	51 (50.5)		
35-49	60 (29.1)		
50-64	40 (20.8)		
≥65	15 (13.4)		
Physical activity		<.0001	0.66
Very low	45 (47.4)		
Low	33 (33.0)		
Moderate	65 (28.4)		
High	23 (12.3)		
Hospitalizations for COPD in the preceding 2 years		<.0001	0.59
0	78 (20.2)		
1	40 (30.5)		
>1	48 (51.6)		
General Health		0.0003	0.60
Bad	32 (45.7)		
Fair	91 (28.3)		
Good	40 (19.9)		
Very good/Excellent	3 (16.7)		
BMI		0.047	0.52
>21	154 (26.4)		
≤21	12 (44.4)		

BMI = body-mass index

Table 3: Ability to predict 5-year mortality measured by the area under the curve (AUC) and confidence interval in the validation sample for the COPD-CART decision tree and the three other severity measures, BODE index, ADO index, and HADO score (*n* = 348).

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

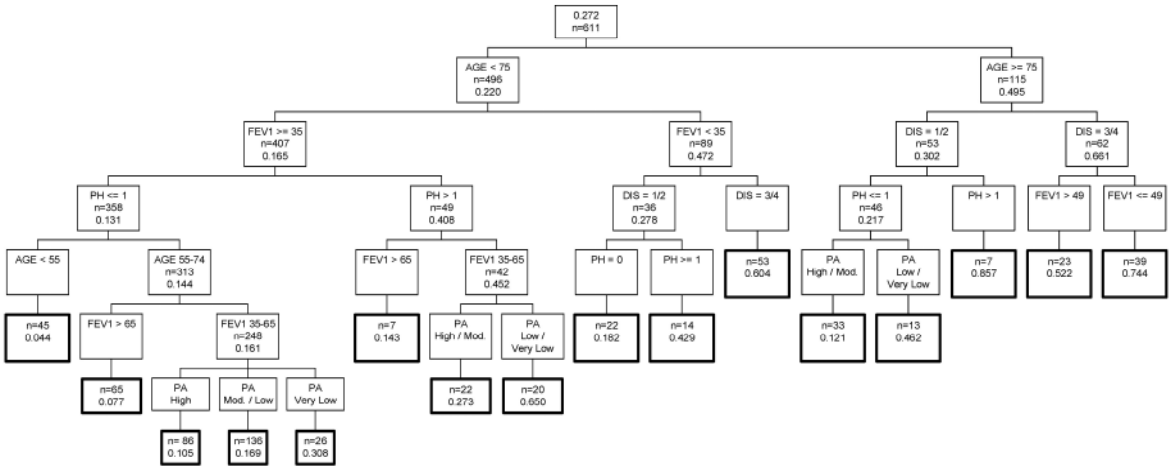
BODE = Body mass index, Obstruction, Dyspnea, Exercise (BODE) index; ADO = Age, Dyspnoea, and airflow Obstruction (ADO) index; HADO = Health-Activity-Dyspnoea-Obstruction (HADO) score.
Calibration by the Hosmer and Lemeshow's test.

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 CART analysis.

CART quartiles of estimated mortality	n	Mortality rate	BODE	ADO	HADO
< 0.10	83	8.43	1.36 [b,c,d] (1.11, 1.61)	2.18 [b,c,d] (1.98, 2.38)	8.46 [c,d] (8.03, 8.89)
0.10 – 0.15	79	25.32	2.24 [a,c,d] (1.99, 2.49)	3.30 [a,c,d] (3.11, 3.50)	8.18 [c,d] (7.88, 8.47)
0.16 – 0.29	95	33.68	3.06 [a,b,d] (2.80, 3.33)	3.98 [a,b,d] (3.79, 4.17)	6.43 [a,b,d] (6.13, 6.73)
≥ 0.30	91	58.24	4.59 [a,b,c] (4.19, 5.00)	5.12 [a,b,c] (4.86, 5.39)	4.42 [a,b,c] (3.97, 4.86)

Nodes have been classified on quartiles depending on the estimated mortality: < 0.10, 0.10 – 0.15, 0.16 – 0.29, and ≥ 0.30.
ANOVA was performed to test statistical differences and Scheffe test for multiple comparisons for HADO, BODE, ADO, St. George and SF-36 scores. Difference for mortality rate were performed by the Pearson's chi-square test..
[a] Indicates statistically significant difference with estimated mortality < 0.10 group.
[b] Indicates statistically significant difference with estimated mortality between 0.10 and 0.15 group.
[c] Indicates statistically significant difference with estimated mortality between 0.16 and 0.29 group.
[d] Indicates statistically significant difference with estimated mortality ≥ 0.30.
BODE = Body mass index, Obstruction, Dyspnea, Exercise (BODE) index; ADO = Age, Dyspnoea, and airflow Obstruction (ADO) index; HADO = Health-Activity-Dyspnoea-Obstruction (HADO) score; SQRQ = St. George Respiratory Quotient

Figure 1


0.272
n=611
AGE < 75
n=496
0.220
AGE >= 75
n=115
0.495
FEV1 >= 35
n=407
0.165
FEV1 < 35
n=89
0.472
DIS = 1/2
n=53
0.302
DIS = 3/4
n=62
0.661
PH <= 1
n=358
0.131
PH > 1
n=49
0.408
DIS = 1/2
n=36
0.278
DIS = 3/4
PH <= 1
n=46
0.217
PH > 1
FEV1 > 49
FEV1 <= 49
AGE < 55
AGE 55-74
n=313
0.144
FEV1 > 65
FEV1 35-65
n=42
0.452
PH = 0
PH >= 1
n=53
0.604
PA
High / Mod.
PA
Low /
Very Low
n=7
0.857
n=23
0.522
n=39
0.744
n=45
0.044
FEV1 > 65
FEV1 35-65
n=248
0.161
n=7
0.143
PA
High / Mod.
PA
Low /
Very Low
n=22
0.182
n=14
0.429
n=33
0.121
n=13
0.462
n=65
0.077
PA
High
PA
Mod. / Low
PA
Very Low
n=22
0.273
n=20
0.650
n= 86
0.105
n=136
0.169
n=26
0.308

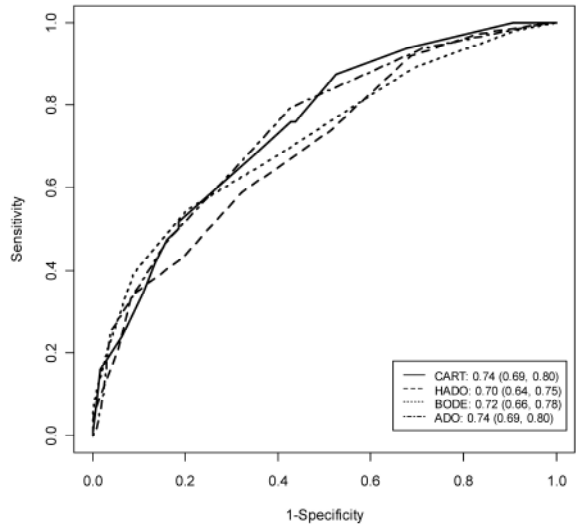

Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
CART: 0.74 (0.69, 0.80)
HADO: 0.70 (0.64, 0.75)
BODE: 0.72 (0.66, 0.78)
ADO: 0.74 (0.69, 0.80)

Figure 2