Mortality prediction in chronic obstructive pulmonary disease comparing GOLD2015 and GOLD2019 staging

García Castillo, Elena; Alonso Pérez, Tamara; Ancochea, Julio; Pastor-Sanz, Teresa María ; Almagro, Pere; Martinez-Camblor, Pablo; Rodriguez-Carballeira, Monica; Navarro, Annie; Lamprecht, Bernd; Sofia Ramírez, Ana; Kaiser, Bernhard; Alfageme, Inmaculada; Casanova, Ciro; Esteban, Cristobal; Soler Cataluña, Juan José; Celli, Bartolome; Marin, Jose Maria; Ter Riet, Gerben; Sobradillo, Patricia; Lange, Peter

DOI:
10.1183/13993003.congress-2020.966

Mortality prediction in chronic obstructive pulmonary disease comparing GOLD2015 and GOLD2019 staging:a pooled analysis of individual patient data

Journal:	*ERJ Open Research*
Manuscript ID	ERJOR-00253-2020.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	30-Jul-2020
Complete List of Authors:	García Castillo, Elena; Hospital Universitario de la Princesa, Pulmonology Alonso Pérez, Tamara; Hospital Universitario de la Princesa, Pulmonology department Ancochea, Julio; Universidad Autonoma de Madrid, Servicio de Neumología, Hospital La Princesa Pastor-Sanz, María Teresa; Instituto de Investigación Hospital Universitario de la Princesa (IISP), Universidad Autónoma de Madrid, Universidad Autónoma de Madrid Almagro, Pere; Universitary Hospital Mutua de Terrassa, Internal Medicine Martinez-Camblor, Pablo; Dartmouth College Geisel School of Medicine, The Dartmouth Institute for Health Policy and Clinical Practice Miravitlles, Marc; Hospital Universitari Vall d'Hebron, Pneumology Department Rodriguez-Carballeira, Monica; Universitary Hospital Mutua de Terrassa, Internal Medicine Navarro, Annie; Hospital Mútua Terrassa. University of Barcelona., Pneumology Lamprecht, Bernd; Kepler University Hospital, Department of Pulmonary Medicine; Johannes-Kepler-University, Faculty of Medicine Ramirez Garcia-Luna, Ana Sofia Kaiser, Bernhard; Paracelsus Medical University, Pulmonary Medicine Alfageme, Inmaculada; Unidad de Gestión Clínica de Neumología, Hospital Universitario Virgen de Valme, Universidad de Sevilla, Sevilla Casanova, Ciro; Unidad de Investigación , Respiratory ESTEBAN, CRISTÓBAL; GALDAKAO-USANSOLO HOSPITAL, PNEUMOLOGY Soler Cataluña, Juan José; Hospital Arnau de Vilanova, Servicio de Neumología de Torres, JP; Servicio de Neumología, Clínica Universitaria de Navarra, Pamplona, España Celli, Bartolome ; Brigham and Women's Hospital, Pulmonary and Critical Care Division Marin, Jose Maria; Hospital Universitario Miguel Servet, Respiratory Service ter Riet, Gerben; Department of General Practice , Academic Medical Center, Sobradillo, Patricia

	Lange, Peter; Herlev Hospital, ; Copenhagen University, Department of Public Health Garcia Aymerich, Judith; Instituto de Salud Global Barcelona, Campus Mar; Universitat Pompeu Fabra, Antó, Josep Maria; Instituto de Salud Global Barcelona, Turner, Alice M; 23. Institute of Applied Health Research, University of Birmingham , Edgbaston , UK Han, MeiLan; University of Michigan, Internal Medicine Langhammer, Arnulf; Norwegian University of Science and Technology, HUNT Research Centre, Department of Public Health and General Practice Vikjord, Sigrid Anna; Norwegian University of Science and Technology, Department of Public Health and Nursing Sternberg, Alice; Johns Hopkins University Bloomberg School of Public Health Leivseth, Linda; Centre Northern Norway Regional Health Authority, Centre for Clinical Documentation and Evaluation Bakke, Per ; University of Catania, Department of Clinical and Experimental Medicine Johannessen, Ane; Department of Global Public Health and Primary Care , University of Bergen , Bergen, Bergen , Norway Oga, Toru; Graduate School of Medicine, Kyoto University, Respiratory Care and Sleep Control Medicine Cosio, Borja; Hospital Universitario Son Espases, Respiratory Medicine-IdISBa; CIBERES, COPD CRP Echazarreta, Andrés; 32. Servicio de Neumonología, Hospital San Juan de Dios de La Plata , Buenos Aires , Argentina Roche, Nicolas; Respiratory Medicine, Cochin Hospital APHP, University Paris Descartes , Paris , France Burgel, Pierre Regis; Cochin Hospital, APHP, Pulmonary Department and Adult CF Centre Sin, Don; University of British Columbia and The Providence Heart and Lung Institute and the James Hogg Research Laboratories, St. Paul's Hospital, Department of Medicine (Division of Respirology) Puhan, Milo; 36. Epidemiology , Biostatistics and Prevention Institute, University of Zurich , Zurich , Switzerland Lopez-Campos, Jose Luis; Hospital Universitario Virgen del Rocío, Unidad Médico-Quirúrgica de Enfermedades Respiratorias; CIBER de Enfermedades Respiratorias (CIBERES), Carrasco, Laura; Hospital Universitario Virgen del Rocio, Unidad Médico-Quirúrgica de Enfermedades Respiratorias Soriano, Joan B; Instituto de Investigación Hospital Universitario de la Princesa (IISP), Universidad Autónoma de Madrid, Universidad Autónoma de Madrid;
Key Words:	COPD, mortality, Survival analysis

ERJOR-00253-2020

The full comments of the reviewers are as follows:

Reviewer: 1

Comments to the Author
Castillo et al. included a large sample size to predict the mortality in GOLD 2015 vs 2019. Some major limitations are mentioned in the discussion part.Some important questions occurred while reading the manuscript:

1.Was the study registered in Clinical Trials
RESPONSE #1: Negative. The 3CIA initiative is a compilation of individually published COPD clinical cohorts which started in 2015 [Lancet Respir Med. 2015 PMID: 25995071]. In retrospect, it was the internationalization of the Spanish-based COCOMICS study, which started in 2011 and produced its first publication in 2013 [Eur Respir J. 2013 PMID: 23222874]. At that time, it was not a journal requirement to register observational studies at https://clinicaltrials.gov or elsewhere. However, for the record most of the individual COPD clinical cohorts are/have been registered.

2. How many patients received oxygene (LTOT)?
RESPONSE #2: As suggested, we attach the data on the number and percentage of patients receiving long term oxygen therapy.

3. Can you provide pCO2 values in table 1.
RESPONSE #3: Regrettably, our protocol requested a set of core variables that did not include pCO_2 or any blood gases

4. Do you have data for emphysema or hyperinflation since this was described as predictors for mortality (Budweiser et al Journal of COPD Vol 11 2014)
RESPONSE #4: Similar to RESPONSE #3, same applied to CT-scan for the entire pooled 3CIA database

5. Is an inclusion bias possible since the inclusion of the centers seemed to be pretty different- please comment and ad some data if the inclusion is driven by a small number of centers.
RESPONSE #5: We kindly differ with Reviewer 1, as the small number of centres (22 cohorts of eleven countries) might be a restriction of representativity, but not a bias (neither selection, information nor confounding). However, the following text has been included in Limitations, page 13:

> "… or a primary care population. ***Indeed, the 22 cohorts from eleven countries in 3CIA within our initiative are only a sample representing the estimated 300 million COPD patients worldwide [new ref: Lancet Respir Med. 2020 Jun;8(6):585-596. doi: 10.1016/S2213-2600(20)30105-3. PMID: 32526187]*** Finally, the recruitment timeframe was long (10 years), so evolving treatment recommendations might influence results."

6. How did you define exacerbation- was the same definition used for every center? (page 7 line 6)
RESPONSE #6: Thanks for allowing us to highlight this most critical issue. We have identified a related, recent ERJ Open Res paper, and the following paragraph is now included in Discussion, page 10:

> "… ***How to define and stage COPD exacerbations is a matter of an intense, long debate [new ref: Mathioudakis AG, Moberg M, Janner J, Alonso-Coello P, Vestbo J. Outcomes reported on the management of COPD exacerbations: a systematic survey of randomised controlled trials. ERJ Open Res. 2019 May 10;5(2):00072-2019. doi: 10.1183/23120541.00072-2019. PMID: 31111041.]. Most 3CIA individual cohorts used subsequent iterations of GOLD-accepted definitions in their protocols, mostly based on Rodriguez-Roisin R, et al. seminal paper [new ref: Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. Chest. 2000 May;117(5 Suppl 2):398S-***

401S. doi: 10.1378/chest.117.5_suppl_2.398s. PMID: 10843984.], including mild (symptom-only) exacerbations. However, when pooling for 3CIA, we only focused on those COPD exacerbations that required health services use, ER, admission or death."

7. Since the comorbidities are missing mortality is a difficult approach from my perspective- please comment.
RESPONSE #7: In our study, we showed the number and percentage of patients with COPD who present comorbidities such as hypertension, diabetes mellitus, cardiac disease or asthma.

8. What is the clinical impact of the displayed data? Would the data influence your clinical decisions?
RESPONSE #8: Candidly, see our paragraphs on research relevance in pages 8 and 9. Indeed, … "the implication of progressively milder disease classifications for the treatment choices clinicians make in real life practice guided by GOLD is not yet clear, but will be important to monitor".

9. What is your suggestion to define a more precise prediction scale for mortality.
RESPONSE #9: The following text and references has been included in Discussion, page 13:
"… ***Ours should be considered a constructive exercise and a critical appraisal of the last two GOLD iterations defining COPD. Within 3CIA, we have already suggested several proposals for future COPD staging and grading classifications, by applying more evidence-based thresholds of evidence-based variables.*** *[New references.]"*

New references:
Almagro P, Martínez-Camblor P, Miravitlles M, Rodríguez-Carballeira M, Navarro A, Lamprecht B, Ramirez-Garcia Luna AS, Kaiser B, Alfageme I, Casanova C, Esteban C, Soler-Cataluña JJ, de-Torres JP, Celli BR, Marin JM, Ter Riet G, Sobradillo P, Lange P, Garcia-Aymerich J, Anto JM, Turner AM, Han MK, Langhammer A, Sternberg A, Leivseth L, Bakke P, Johannessen A, Oga T, Cosío B, Ancochea J, Echazarreta A, Roche N, Burgel PR, Sin DD, Puhan MA, Soriano JB; 3CIA collaboration. External Validation and Recalculation of the CODEX Index in COPD Patients. A 3CIAplus Cohort Study. COPD. 2019 Feb;16(1):8-17. doi: 10.1080/15412555.2018.1484440. Epub 2019 Mar 14.PMID: 30870059

Guerra B, Haile SR, Lamprecht B, Ramírez AS, Martinez-Camblor P, Kaiser B, Alfageme I, Almagro P, Casanova C, Esteban-González C, Soler-Cataluña JJ, de-Torres JP, Miravitlles M, Celli BR, Marin JM, Ter Riet G, Sobradillo P, Lange P, Garcia-Aymerich J, Antó JM, Turner AM, Han MK, Langhammer A, Leivseth L, Bakke P, Johannessen A, Oga T, Cosio B, Ancochea-Bermúdez J, Echazarreta A, Roche N, Burgel PR, Sin DD, Soriano JB, Puhan MA; 3CIA collaboration. Large-scale external validation and comparison of prognostic models: an application to chronic obstructive pulmonary disease..BMC Med. 2018 Mar 2;16(1):33. doi: 10.1186/s12916-018-1013-y.PMID: 29495970

Haile SR, Guerra B, Soriano JB, Puhan MA; 3CIA collaboration. Multiple Score Comparison: a network meta-analysis approach to comparison and external validation of prognostic scores.. BMC Med Res Methodol. 2017 Dec 21;17(1):172. doi: 10.1186/s12874-017-0433-2.PMID: 29268701

Burgel PR, Paillasseur JL, Janssens W, Piquet J, Ter Riet G, Garcia-Aymerich J, Cosio B, Bakke P, Puhan MA, Langhammer A, Alfageme I, Almagro P, Ancochea J, Celli BR, Casanova C, de-Torres JP, Decramer M, Echazarreta A, Esteban C, Gomez Punter RM, Han MK, Johannessen A, Kaiser B, Lamprecht B, Lange P, Leivseth L, Marin JM, Martin F, Martinez-Camblor P, Miravitlles M, Oga T, Sofia Ramírez A, Sin DD, Sobradillo P, Soler-Cataluña JJ, Turner AM, Verdu Rivera FJ, Soriano JB, Roche N; Initiatives BPCO, EABPCO, Leuven and 3CIA study groups. A simple algorithm for the identification of clinical COPD phenotypes. Eur Respir J. 2017 Nov 2;50(5):1701034. doi: 10.1183/13993003.01034-2017.

Soriano JB, Lamprecht B, Ramírez AS, Martinez-Camblor P, Kaiser B, Alfageme I, Almagro P, Casanova C, Esteban C, Soler-Cataluña JJ, de-Torres JP, Miravitlles M, Celli BR, Marin JM, Puhan MA, Sobradillo P, Lange P, Sternberg AL, Garcia-Aymerich J, Turner AM, Han MK,

Langhammer A, Leivseth L, Bakke P, Johannessen A, Roche N, Sin DD. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. Lancet Respir Med. 2015 Jun;3(6):443-50. doi: 10.1016/S2213-2600(15)00157-5. Epub 2015 May 17.PMID: 25995071

Please ad some thoughts in the discussion.
RESPONSE #10: We have significantly expanded our Discussion section with the clarifications previously exposed and 8 new references

Associate Editor: Andreas, Stefan
Comments to the Author:
The mortality data on GOLD 2015 vs 2019 classification are well presented in a large, international cohort. Please add in table 4: Accuracy for predicting mortality…
RESPONSE #11: As requested, we modify the table 4 title.

Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2015 and GOLD 2019 staging: a pooled analysis of individual patient data

Elena García Castillo[1] *, Tamara Alonso Pérez[1] *, Julio Ancochea[1], Maria Teresa Pastor Sanz[1], Pere Almagro[2], Pablo Martínez-Camblor[3], Marc Miravitlles[4], Mónica Rodríguez- Carballeira[5], Annie Navarro[6], Bernd Lamprecht[7], Ana S Ramírez-García Luna[8], Bernhard Kaiser[9], Inmaculada Alfageme[10], Ciro Casanova[11], Cristóbal Esteban[12], Juan J Soler-Cataluña[13], Juan P de-Torres[14], Bartolomé R Celli[15], Jose M Marín[16], Gerben ter Riet[17], Patricia Sobradillo[18], Peter Lange[19], Judith García-Aymerich[20,21,22], Josep M Anto[20,21,22], Alice M Turner[23], MeiLan K Han[24], Arnulf Langhammer[25], Sigrid Anna Aalberg Vikjord[25], Alice Sternberg[26], Linda Leivseth[27], Per Bakke[28], Ane Johannessen[29], Toru Oga[30], Borja Cosío[31], Andrés Echazarreta[32], Nicolás Roche[33], Pierre-Régis Burgel[34], Don D Sin, Milo[34,35], Milo A Puhan[36], Jose Luis López Campos[37], Laura Carrasco[38], Joan B Soriano[39] & for the 3CIA collaboration.

* Shared first authors

Author information:

1. Pneumology Department, Hospital Universitario de la Princesa. Instituto de Investigación Hospital Universitario de la Princesa (IISP), Universidad Autónoma de Madrid, Madrid, Spain.
2. Multimorbidity Patients Unit, Internal Medicine Dept, University Hospital Mutua de Terrassa, University of Barcelona, Barcelona,Spain.
3. Geisel School of Medicine at Dartmouth, Hanover, NH, USA.
4. Pneumology Department, Hospital Universitary Vall d'Hebron, CIBER de Enfermedades Respiratorias (CIBERES), Barcelona , Spain
5. Multimorbidity Patients Unit, Internal Medicine, Hospital Universitari Mutua de Terrassa, Universitat de Barcelona, Barcelona , Spain.
6. Pneumology Service, Hospital Universitari Mútua Terrassa, Barcelona, Spain
7. Department of Pulmonary Medicine, Kepler-University-Hospital. Faculty of Medicine , Johannes-Kepler-University Linz , Austria.
8. Facultad de Medicina, Universidad Autónoma de San Luis Potosí, San Luis Potosí, Mexico.
9. Department of Pulmonary Medicine , Paracelsus Medical University Hospital , Salzburg, Austria.
10. Departamento de Medicina. Universidad de Sevilla. HU Virgen de Valme. Sevilla. Spain.
11. Hospital Universitario Nuestra Señora de La Candelaria, Universidad de La Laguna, Tenerife, Spain.

12. Pulmonology Department. Hospital Galdakao-Usansolo, Galdakao , Bizkaia , Spain.
13. Servicio de Neumología, Hospital Arnau de Vilanova, Valencia , Spain.
14. Clinica Universidad de Navarra, Pamplona, Spain.
15. Pulmonary and Critical Care Medicine , Harvard University, Brigham and Women's Hospital , Boston , MA , USA
16. Hospital Universitario Miguel Servet, Zaragoza, and CIBER de Enfermedades Respiratorias (CIBERES), Spain
17. Department of General Practice , Academic Medical Center, University of Amsterdam (AMC) , Amsterdam , The Netherlands.
18. Hospital Universitario Araba, Sede Txagorritxu, Vitoria, Spain for Universitary Hospital of Cruces in Barakaldo, Spain.
19. Section of Social Medicine, Department of Public Health , Copenhagen University, Copenhagen City Heart Study, Frederiksberg Hospital, Frederiksberg , Copenhagen , Denmark
20. Global, Centre for Research in Environmental Epidemiology (CREAL) , Barcelona , Spain.
21. Department of Experimental and Health Sciences , Universitat Pompeu Fabra (UPF) , Barcelona, Spain.
22. CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain.
23. Institute of Applied Health Research, University of Birmingham , Edgbaston , UK.
24. Internal Medicine, Division of Pulmonary and Critical Care Medicine, University of Michigan , Ann Arbor , MI , USA.
25. Department of Public Health and Nursing , NTNU, Norwegian University of Science and Technology , Trondheim , Norway.
26. Department of Epidemiology , Johns Hopkins Bloomberg School of Public Health , Baltimore , MD , USA.
27. Centre for Clinical Documentation and Evaluation , Northern Norway Regional Health Authority, Tromso , Norway
28. Department of Clinical Science , University of Bergen , Bergen , Norway.
29. Department of Global Public Health and Primary Care , University of Bergen , Bergen, Bergen , Norway.
30. Department of Respiratory Care and Sleep Control Medicine , Kyoto University , Kyoto , Japan.
31. Department of Respiratory Medicine , Hospital Son Espases-IdISPa, Ciberes, Mallorca , Spain.
32. Servicio de Neumonología, Hospital San Juan de Dios de La Plata , Buenos Aires , Argentina.
33. Respiratory Medicine, Cochin Hospital APHP, University Paris Descartes , Paris , France.
34. James Hogg Research Centre , University of British Columbia , Vancouver , BC , Canada.
35. Division of Respiratory Medicine, Department of Medicine , St Paul's Hospital , Vancouver , BC , Canada.
36. Epidemiology , Biostatistics and Prevention Institute, University of Zurich , Zurich , Switzerland
37. Pulmonology Department. Hospital Universitario Virgen del Rocío, Sevilla, España
38. Pulmonology Department. Hospital Universitario Virgen del Rocío, Sevilla, España.

39. Instituto de Investigación Hospital Universitario de la Princesa (IISP), Universidad Autónoma de Madrid, Madrid, Spain

Address for correspondence:

Dr. Joan B Soriano

Hospital Universitario de la Princesa, Diego de León 62, Neumología 6ª planta, 28006-Madrid, Spain

Email: jbsoriano2@gmail.com

Cellular: +34 618 86 77 69

File information:

Date: July, 2020

Word count: 2,838 words

References: 39 references

Illustrations: 4 tables and 4 figures

Keywords: COPD, mortality, survival

Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2015 and GOLD 2019 staging: a pooled analysis of individual patient data

Abstract

In 2019, The Global Initiative for Chronic Obstructive Lung Disease (GOLD) modified the grading system for patients with chronic obstructive pulmonary disease (COPD) creating 16 subgroups (1A-4D). As part of the COPD Cohorts Collaborative International Assessment (3CIA) initiative, we aim to compare mortality prediction of 2015 and 2019 COPD GOLD staging systems.

We studied 17 139 COPD patients from the 3CIA study, selecting those with complete data. Patients were classified by the 2015 and 2019 GOLD ABCD systems and we compared the predictive ability for 5-year mortality of both classifications.

17 139 patients with COPD were enrolled in 22 cohorts of 11 countries between 2003 to 2017; 8 823 of them had complete data and were analyzed. Mean age was 63.9 years (SD 9.8) and 62.9% were male. GOLD 2019 classified the patients in milder degrees of COPD. For both classifications, group D had higher mortality. Five-year mortality did not differ between groups B and C in GOLD 2015; in GOLD 2019, mortality was greater for group B than C. Patients classified as group A and B had better sensitivity and positive predictive value with GOLD 2019 classification than GOLD 2015. GOLD 2015 had better sensitivity for group C and D than GOLD 2019. The AUC for 5-year mortality were only 0.67 (95% CI 0.66-0.68) for GOLD 2015 and 0.65 (95% CI 0.63-0.66) for GOLD 2019.

The new GOLD 2019 classification does not predict mortality better than the previous GOLD 2015 system.

Take Home message:

GOLD 2019 staging system created 16 subgroups

GOLD 2015 and GOLD 2019 are not strong predictors of mortality and don't have sufficient discriminatory power to be used as a tool for risk classification of mortality in patients with COPD

Introduction

Chronic obstructive pulmonary disease (COPD) is a common cause of morbidity and mortality in the world. COPD affects approximately 328 million people worldwide, and COPD related deaths amount to 4 million deaths every year[1]. Assessment of disease severity is essential to predict prognosis and to standardize treatment regimes. The Global Initiative for Chronic Obstructive Lung Disease (GOLD) document is the most widely used treatment guide for the staging and management of COPD. The GOLD grading system for COPD has significantly evolved since first publication in 2001 to the current version in 2019. Initially, in the GOLD 2007 classification, only post-bronchodilator airflow limitation based on spirometry forced expiratory volume in 1 second (FEV_1) was used to grade the severity of COPD[2]. Later on, some criticism arosed on this grading score because it relied only upon FEV_1 which is not a good predictor of dyspnoea, quality of life or exercise tolerance. Further, other important variables to evaluate prognosis in COPD, such as sub-phenotypes, exacerbations, dyspnoea severity or comorbidities have been proposed, among others. Therefore, GOLD 2011 proposed a classification system of four groups, ABCD, combining FEV_1 and two clinical parameters: history of exacerbations and respiratory symptoms measured by the modified Medical Research Council (mMRC) dyspnoea score, or the COPD Assesment Test score (CAT) [3]. The 2011 ABCD classification was considered an improvement in the management of patients with COPD, providing an opportunity to further guide the individualized care of these patients. GOLD 2011 predicted future exacerbations better than GOLD 2007 but there was no difference in mortality predictions or respiratory outcomes [4-7]. In 2015 an updated report was published with the same measurement parameters (FEV_1, dyspnoea and exacerbations) than 2011 classification[8].

The latest GOLD update, published in 2019 uses a composite of spirometry, symptoms and exacerbations, but importantly separating the spirometric 1-4 staging from the ABCD groups[9]. This separation is relevant because it is known there are differences in the rate of exacerbations for the most severe COPD patients, depending on whether the risk is based on pulmonary function tests, on the history of exacerbations or both[10]. All these classifications were initially designed not to assess prognosis, but to aid clinicians in creating optimal treatment

regimes for patients. Thus, the prognostic ability of GOLD 2019 compared to previous classifications is largely unknown, with only a few published studies[11,12]. To address this issue, we used pooled data from 17,139 patients of 22 COPD cohorts and eleven countries and compared the prognostic capacity of the 2019 versus 2015 GOLD staging classifications to predict mortality.

Methods

Study population

In this international study, we assessed 17,139 patients from the COPD Cohorts Collaborative International Assessment (3CIA) initiative. All were prospective cohorts that recruited patients within the period 1999 to 2017, except one of them that was a population based cohort. All patients had a definition of COPD characterized by spirometry, that is post-bronchodilator FEV_1 to forced vital capacity ratio (FVC) ratio less than 0.7 and a clinical diagnosis of COPD. Spirometry was performed using the standards provided by the American Thoracic Society and European Respiratory Society[13]. The primary investigators of each of the participating 3CIA cohorts provided individual patient's data for pooled analysis. We obtained a minimum individual dataset including the vital status (up to death, right truncation or 2017), age, sex, smoking status, pre-bronchodilator and post-bronchodilator FEV_1 and FVC and dyspnoea measured with the modified mMRC, among others[14]. Only in a number of 3CIA cohorts were data of number of exacerbations in the previous year available. For the current study, we selected exclusively those cohorts in which the number of exacerbations in the previous year were available in the database, since this variable is required to calculate the GOLD 2015 and 2019 grading systems. Fifteen out of a total of 22 cohorts contained data on history of exacerbations, so that the final number of patients available to be classified according to GOLD 2015 and GOLD 2019 was 8,823. Symptoms were assessed using the mMRC dyspnoea scale. To determine the risk descriptor in the 2015 grouping system, we used exacerbations history and GOLD spirometry stages. The combination of symptoms (mMRC) and the worse risk descriptor (spirometry or exacerbation history) were used to classify patients by GOLD 2015 system. Participants were

classified using the GOLD 2019 system into four grades (1-4) based on postbronchodilator FEV_1 percentage of prediction as stage 1 ($FEV_1 \geq 80$), stage 2 (FEV_1 79-50), stage 3 (FEV_1 30-49) and stage 4 ($FEV_1 < 30$). Groups ABCD were defined by self-reported severity of dyspnea (mMRC) and number of exacerbations in the previous year.

All participants provided informed written consent, and each study was conducted with the formal approval of the local ethics institutional committees following the principles of the Declaration of Helsinki.

Outcomes

The primary outcome was the prediction ability of all-cause mortality in the individuals by the two GOLD systems.

Statistical analysis

The 3CIA database manager quality-controlled all data centrally and created a clean database with a data dictionary. All implausible or missing variables were queried with the original study investigators, and datum were removed from the central database if errors could not be corrected. Because the cohorts had different follow-up times, patients were right-censored at five years of follow-up.

Descriptive statistics used mean and standard deviation for continuous variables and the number of cases and percentages for categorical variables. Comparisons between groups were performed with the Chi^2 - test for categorical variables and the T-test for continuous variables.

We estimated 5-year all-cause mortality, according to GOLD 2015 and 2019 staging systems, using Kaplan-Meier survival statistics. Statistical comparisons were performed using the Log-Rank test Receiver Operation Characteristic (ROC) curve analyses, area under the curve (AUC) and the 95% CI of the AUC were calculated to measure the predictive accuracy for mortality. Also, we compared the prediction ability of mortality on both classifications using sensitivity, positive predictive value and the Youden's index with Epidat 3.1 program.

Results

We pooled data from 22 COPD cohorts with a total of 17,139 patients, finally including 8,823 patients from 15 cohorts that had all complete variables to be classified as GOLD 2015 and GOLD 2019. A comparison of baseline characteristics of included and not included patients is presented in Table 1. There were statistically significant differences in many variables given the large size, but most should be considered not clinically relevant (Table 1).

The 8,823 included patients were 62.9% male, with a mean age of 63.9 (SD± 9.8) years, mean±SD BMI 27.0±5.8 kg/m^2 and mMRC dyspnoea score of 1.8±1.4. Postbronchodilator FEV_1 was 54.8%±22.3 of the predictive value, and 6-minute walk distance was 376.9±129.1 meters. Based on spirometry staging 1,153 (13%) had mild (stage 1), 3,711 (42.1%) had moderate (stage 2), 2,654 (30.1%) had severe (stage 3) and 1,301 (14.8%) had very severe (stage 4) disease.

The distribution of these 8,823 patients according to GOLD 2015 and GOLD 2019 is presented in Figure 1. With GOLD 2019 there is a shift towards the less severe staging of disease (absolute increase in stage A and B of 7.8% and 22.2% respectively; and absolute decrease in stage C and D of 7.8 and 22.2% respectively).

The overall 5-year mortality rate was 18.3%. The all-cause 5-year mortality rates according to both classifications are shown in Table 2. Figure 2 shows Kaplan-Meier curves for 5-year mortality according to GOLD 2015 (Figure 2A) and GOLD 2019 (Figure 2B). All-cause mortality at 5 years in the 2015 GOLD classification was higher in grade D, followed by grades B, C (with similar mortality), and finally grade A, log-rank test $p<0.001$ (Table 2 and Figure 2A). Grade D diverged from the beginning of follow up, while grades B and C diverged after 1 year of follow-up. In GOLD 2019, the four Kaplan-Meier curves diverge during the first year, but interestingly, grade B had higher mortality than grade C, so mortality was higher in groups B and D (more symptoms) than in groups A and C (fewer symptoms; Figure 2B). The degree of obstruction measured by FEV_1 % further subclassified patients into 16 subgroups with different mortality rates, increasing from 1A to 4D in GOLD 2019 grading system (Table 3). Figure 3 shows Kaplan-Meier curves for each of the spirometry strata. The higher mortality of group B over group C persisted in each of the strata with the exception of spirometry strata 1, with a higher

mortality in group C. Similarly to GOLD 2015, in GOLD 2019, grade A and D had the lowest and highest mortality, respectively, with very similar absolute numbers (Table 2, Figure 2).

The primary outcome, the prediction capacity as measured by the AUC of ROC curve for mortality up to 5 years was intermediate (<0.70) for both systems (Figure 4). GOLD 2015 exhibited slightly better discrimination in predicting mortality (AUC 0.67 [95% CI 0.66-0.68]) than GOLD 2019 classification (AUC 0.64 [95% CI 0.63-0.66]).

Regarding sensitivity parameters, both classifications had quite shallow values. GOLD 2019 had a higher sensitivity for predicting mortality on A and B groups vs 2015 classification (19.1 vs 11.1 and 43.6 vs 11.6). On the other hand, GOLD 2015 had a higher sensitivity on groups C and D vs 2019 classification (14.6 vs 6.6 and 62.5 vs 30.6) based on overlapping 95% CIs (Table 4). The positive predictive values were also higher in GOLD 2019 group A and B vs the same groups in GOLD 2015 (9.5 vs. 6.9 and 21.2 vs. 13.3); but not different (overlapping confidence intervals) in groups C and D. The Youden indices were quite low for both classifications, even with negative values, showing that it is not an optimal classification system to assess mortality.

Discussion

Our study evaluates mortality according to the last two GOLD classifications, and it is one of the most extensive to date. The most important finding is that the new GOLD 2019 classification (based on symptoms or exacerbations along with stages spirometry analysed altogether) did not predict 5-year mortality better than GOLD 2015 classification (based on spirometry, history of exacerbation and symptoms). How to define and stage COPD exacerbations is a matter of an intense, long debate[15]. Most 3CIA individual cohorts used subsequent iterations of GOLD-accepted definitions in their protocols, mostly based on Rodriguez-Roisin R, et al. seminal paper[16], including mild (symptom-only) exacerbations. However, when pooling for 3CIA, we only focused on those COPD exacerbations that required health services use, ER, admission or death. In GOLD 2019 classification, subgroups B and D had the worse mortality, highlighting that the higher burden of symptoms conveys a higher mortality independently of

spirometry (Table 2 and Figure 2B). Finally, we show an important shift in the proportions of patients between the 2015 and 2019 ABCD grades, with milder severity in GOLD 2019.

GOLD 2015 and GOLD 2019 classifications had a low discriminatory power as per the AUCs, ranging from 0.67 to 0.65, and similar to other studies (by consensus, AUC below 0.70 are considered low or weak)[11]. Sensitivity and positive predictive values indicate that the general performance of the two models is similar and very low. Also, the Youden's indices are very low with negative values that have no meaningful interpretation in practice. These findings support the results of other studies, suggesting that GOLD classification is not a good predictor for mortality[11-12]. There may be various reasons for these poor results. The main reason is that these classifications were conceived to guide treatment, so it is not surprising that their capacity for predicting mortality is low. There are clinical phenotypes such as the asthma-COPD, the frequent exacerbator with emphysema or chronic bronchitis comorbidities and different indexes, that are significant predictors of mortality and are not entirely included in the GOLD stages[17-21]. In our study, sensitivity and positive predictive value are higher in GOLD 2019 groups A and B. On the contrary, sensitivity is higher in GOLD 2015 in groups C and D. These results suggest that GOLD 2019 predicts slightly better mortality on low risk groups (A and B) and GOLD 2015 in high risk groups (C and D), requiring more studies to corroborate it.

The discriminatory power of GOLD 2019 was lower than GOLD 2015 as shown in AUCs values. The partition of FEV_1 as a direct classifier in GOLD 2019 reduced ability to discriminate survival, highlighting the need to consider the severity of airflow obstruction in assessing mortality risk. When using GOLD 2019 with a composite of spirometry, exacerbations and symptoms (16 subgroups 1A to 4D classification), we found an increase in all-cause mortality between GOLD 2019 stage 1 and GOLD 2019 stage 4 across grades A, B and D, highlighting the persisting importance of FEV_1 as a predictor of mortality (Table 3 and Figure 3). In group C, mortality was higher in spirometry stage 1 than in 2, probably due to a significant difference of the proportion in patients between the two stages.

Our study confirms that patients classified as GOLD A had the best survival, and patients with GOLD D, the higher mortality in both classifications[22-24]. Mortality of groups B and C in GOLD 2015 lied in between A and D groups and often overlapped. In GOLD 2019 mortality was significantly higher in group B patients than in group C. This finding is similar as other previous reports published[25-27], showing that group B is an intermediate-high risk group associated with more exacerbations and likely to other comorbidities that may cause dyspnoea (such as heart failure). Furthermore, we show that the burden of symptoms (represented by groups B and D) have a prognostic value additive but independent to spirometry.

The current study demonstrates that the use of GOLD 2019 classification scheme shifted patients with COPD to groups of milder severity compared with GOLD 2015. This happened in keeping with previous reports, but in a smaller proportion of patients (30% of patients reassigned towards group A or B compared to 53% in Lee et al study or 66% in Tan et al study) [28-29]. The further distribution of spirometric parameters from two categories in GOLD 2015 (FEV_1 less or higher than 50%) to four categories in GOLD 2019, was one of the possible reasons for the patient's shift from C and D in GOLD 2015 to A and B groups in GOLD 2019. This phenomenon is opposed to the one observed with the use of the revised GOLD 2011 classification, which shifted the patients from the GOLD 2007 towards more advanced stages of the disease (D group increased almost 3 times) [7]. Remarkably, in both classifications, group C was consistently the smallest group (Figure 1), as seen in other reports[27]. The implication of progressively milder disease classifications for the treatment choices clinicians make in real life practice guided by GOLD is not yet clear given the recent nature of the latest GOLD iteration, but will be important to monitor.

Strength of our study include a large sample size, the study design (a pooled-analysis of individual patient-data from several cohorts), and the different degrees of severity of patients from different cohorts, maximizing its high external validity. Prospective data collection of spirometry with a post-bronchodilator test, dyspnea by mMRC scale, history of exacerbation and mortality

enabled direct classification of patients by 2015 and 2019 GOLD staging schemes. We also have a significant representation of women, of which other COPD studies might not have achieved[30]. Our study has several limitations. First, although we started with 17,139 patients with COPD, a considerable number of patients were excluded because of missing information on key variables, mainly regarding the history of exacerbations. These missing data are unlikely to affect the validity of our results, as we can see in our analysis comparing included and non-included patients. Second, the mortality analysis used all-cause death, and we have no data regarding specific causes of death (this data was not collected consistently in all cohorts). Third, symptoms were only evaluated using mMRC dyspnea score, but not with the COPD Assessment Test, or other instruments[31]; however, that is in line with other reported cohorts[32-33]. Fourth, most patients came from hospital-based cohorts, so we likely have more patients with moderate to severe disease and less patients with mild and moderate disease compared to an outpatient setting, or a primary care population. Indeed, the 22 cohorts from eleven countries in 3CIA within our initiative are only a sample representing the estimated 300 million COPD patients worldwide[34]. Finally, the recruitment timeframe was long (10 years), so evolving treatment recommendations might influence results.

In conclusion, this study of COPD cohorts, including 8823 patients, showed that neither GOLD 2015 nor GOLD 2019 are strong predictors of mortality. GOLD 2019 predicted mortality better than GOLD 2015 in groups A and B but worse in groups C and D. However, none of the GOLD classifications has sufficient discriminatory power to be used as a tool for risk classification of mortality in patients with COPD. Ours should be considered a constructive exercise and a critical appraisal of the last two GOLD iterations defining COPD. Within 3CIA, we have already suggested several proposals for future COPD staging and grading classifications, by applying more evidence-based thresholds of evidence-based variables [35-39].

Declaration of interest

There are no direct or indirect conflicts of interest to disclose related with this manuscript

References

[1] GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Respir Med. 2017; 5(9): Sep 691-706.

[2] Rabe KF, Hurd S, Anzueto A, et al; Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med. 2007;176(6):532-55.

[3] Vestbo J, Hurd SS, Agustí AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med 2013; 187:347-65.

[4] Han MK, Muellerova H, Curran-Everett D, et al. GOLD 2011 disease severity classification in COPDGene: a prospective cohort study. Lancet Respir Med.2013;1(1):43-50.

[5] Soriano JB, Alfageme I, Almagro P, et al. Distribution and prognostic validity of the new Global Initiative for Chronic Obstructive Lung Disease grading classification. Chest 2013;143(3):694-702.

[6] Lange P, Marrot JL, Vestbo J, et al. Prediction of the clinical course of chronic obstructive pulmonary disease, using the new GOLD classification: a study of the general population. Am J Respir Crit Care Med.2012;186(10):975-981.

[7] Soriano JB, Lamprecht B, Ramírez AS, et al. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. Lancet Respir Med 2015; 3: 443-50.

[8] Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the diagnosis, management and prevention of COPD 2015. 2015 [accessed 2019 June 21]. Available from: http://www.goldcopd.org/.

[9] Singh D, Agusti A, Anzueto A, Barnes PJ, Bourbeau J, Celli BR et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease: the GOLD science committee report 2019. Eur Respir J. 2019 May 18;53(5).

[10] Goossens LM, Leimer I, Metzdorf N, Becker K, Rutten-van Mölken N. Does the 2013 GOLD classification improve the ability to predict lung function decline, exacerbations and mortality: a post-hoc analysis of the 4-year UPLIFT trial. BMC Pulm Med. 2014; 14: 163.

[11] Gedebjerg A, Szépligeti SK, Wackerhausen LH, Horváth-Puhó E, Dahl R, Hansen JG et al. Prediction of mortality in patients with chronic obstructive pulmonary disease with the new Global Initiative for Chronic Obstructive Lung Disease 2017 classification: a cohort study. Lancet Respir Med. 2018 Mar;6(3):204-212.

[12] Aramburu A, Arostegui I, Moraza J, Barrio I, Aburto M, García-Loizaga A et al. COPD classification models and mortality prediction capacity. International Journal of COPD 2019:14;605-613

[13] Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al; ATS/ERS Task Force. Standardisation of spirometry. Eur Respir J 2005;26:319-338.

[14] Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. Usefulness of the Medical Research Council (MRC) dyspnea scale as a measure of disability in patients with chronic obstructive pulmonary disease. Thorax 1999;54:581-586.

[15] Mathioudakis AG, Moberg M, Janner J, Alonso-Coello P, Vestbo J. Outcomes reported on the management of COPD exacerbations: a systematic survey of randomised controlled trials. ERJ Open Res. 2019 May 10;5(2):00072-2019.

[16] Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. Chest. 2000 May;117(5 Suppl 2):398S-401S.

[17] Golpe R, Suárez-Valor M, Martín – Robles I, Sanjuán-López P, Cano-Jiménez E, Castro-Añón O et al. Mortality in COPD patients according to clinical phenotypes. International Journal of COPD 2018:13;1433-1439.

[18] Marin JM, Alfageme I, Almagro P, Casanova C, Esteban C, Soler-Cataluña JJ et al. Multicomponent indices to predict survival in COPD: the COCOMICS study. Eur Respir J 2013; 42: 323–332.

[19] Izquierdo JL, Miravitlles M, Esquinas C, Pérez M, Calle M, López Campos JL et al. Characteristics of COPD Patients Managed in Respiratory Medicine Departments in Spain, According to GOLD Groups and GesEPOC Clinical Phenotypes. Arch Bronconeumol. 2018;54:559-567

[20] Radovanovic D, Contoli M, Marco FD, Sotgiu G, Pelaia G, Braido F et al. Clinical and Functional Characteristics of COPD Patients Across GOLD Classifications: Results of a Multicenter Observational Study. COPD. 2019;16:215-226.

[21] Almagro P, Martínez-Camblor P, Miravitlles M, Rodríguez-Carballeira M, Navarro A, Lamprecht B et al. External Validation and Recalculation of the CODEX Index in COPD Patients. A 3CIAplus Cohort Study. COPD. 2019;16:8-17.

[22] Han MZ, Hsiue TR, Tsai SH, Huang TH, Liao XM, Chen CZ. Validation of the GOLD 2017 and new 16 subgroups (IA-4D) classifications in predicting exacerbation and mortality in COPD patients. Int J Chron Obstruct Pulmon Dis. 2018; 13:3425-3433.

[23] Flynn RW, MacDonaldn TM, Chalmers JD, Schembri S. The effect of changes to GOLD severity stage on long term morbidity and mortality in COPD. Respiratory Research 2018 19:249.

[24] de Torres JP, Casanova C, Marin JM, Pinto-Plata V, Divo M, Zulueta JJ, et al. Prognostic evaluation of COPD patients: GOLD 2011 versus BODE and the COPD comorbidity index COTE. Thorax. 2014;69:799 –804.

[25] Lange P, Marott JL, Vestbo J, Olsen KR, Ingebrigtsen TS, Dahl M et al. Prediction of the clinical course of chronic obstructive pulmonary disease, using the new GOLD classification: a study of the general population. Am J Respir Crit Care Med. 2012;186:975–81.

[26] Lee SJ, Yun SS, Ju S, You JW, Cho JJ, Jeong YY et al. Validity of the GOLD 2017 classification in the prediction of mortality and respiratory hospitalization in patients with chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis. 2019; 14: 911-919.

[27] Flynn RWV, MacDonald TM, Chalmers JD, Schembri S. The effect of changes to GOLD severity stage on long term morbidity and mortality in COPD. Respir Res. 2018;19:249.

[28] Tan WC, Bourbeau J, Aaron SD, et al. GOLD 2017 classification and lung function decline in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2018;1,197(5):670-673.

[29] Criner RN, Labaki WS, Regan EA, Bon JM, Soler X, Bhatt SP et al. for the COPDGene Investigators. Mortality and exacerbations by Global Initiative for Chronic Obstructive Lung Disease Groups ABCD: 2011 versus 2017 in the COPDGene Cohort. Chronic Obstr Pulm Dis. 2019 Jan 10;6(1):64-73

[30] Kobayashi S, Hanagama M, Ishida M, Sato H, Ono M, Yamanda S et al. Clinical characteristics and outcomes in Japanese patients with COPD according to the 2017 GOLD classification: the Ishinomaki COPD Network Registry. Int J Chron Obstruct Pulmon Dis.2018:13;3947-3955.

[31] Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the Asthma Control Test: a survey for assessing asthma control. J Allergy Clin Immunol 2004;113:59-65.

[32] Johannessen A, Nilsen RM, Storebo M, Gulsvik A, Eagan T, Bakke P. Comparison of 2011 and 2007 global initiative for chronic obstructive lung disease guidelines for predicting mortality and hospitalization. Am J Respir Crit Care Med. 2013;188:51–9.

[33] Agusti A, Edwards LD, Celli B, Macnee W, Calverley PM, Mullerova H, Lomas DA, Wouters E, Bakke P, Rennard S, et al. Characteristics, stability and outcomes of the 2011 GOLD COPD groups in the ECLIPSE cohort. Eur Respir J. 2013;42:636–46.

[34] GBD Chronic Respiratory Disease Collaborators. Prevalence and attributable health burden of chronic respiratory diseases, 1990 – 2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Respir Med. 2020;8(6):585-96.

[35] Almagro P, Martínez-Camblor P, Miravitlles M, Rodríguez-Carballeira M, Navarro A, Lamprecht B, et al. External Validation and Recalculation of the CODEX Index in COPD Patients. A 3CIAplus Cohort Study. COPD. 2019 Feb;16(1):8-17.

[36] Guerra B, Haile SR, Lamprecht B, Ramírez AS, Martinez-Camblor P, Kaiser B, et al. Large-scale external validation and comparison of prognostic models: an application to chronic obstructive pulmonary disease. BMC Med. 2018 Mar 2;16(1):33.

[37] Haile SR, Guerra B, Soriano JB, Puhan MA; 3CIA collaboration. Multiple Score Comparison: a network meta-analysis approach to comparison and external validation of prognostic scores. BMC Med Res Methodol. 2017 Dec 21;17(1):172.

[38] Burgel PR, Paillasseur JL, Janssens W, Piquet J, Ter Riet G, Garcia-Aymerich J, et al. Initiatives BPCO, EABPCO, Leuven and 3CIA study groups. A simple algorithm for the identification of clinical COPD phenotypes. Eur Respir J. 2017 Nov 2;50(5):1701034.

[39] Soriano JB, Lamprecht B, Ramírez AS, Martinez-Camblor P, Kaiser B, Alfageme I, et al. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. Lancet Respir Med. 2015 Jun;3(6):443-50.

TABLES

	Excluded (n=8316)	**Included (n=8823)**	**p-value**
Age (years)	64.2 ± 10.7	63.9 ± 9.8	0.08
Male sex	6232 (74.9%)	5552 (62.9%)	< 0.001
BMI	26.5 ± 4.9	27.0 ± 5.8	< 0.001
Modified MRC dyspnea scale	1.5 ± 1.3	1.8 ± 1.4	< 0.001
0	1647 (23.4%)	1957(22.2%)	< 0.001
1	2261(32.1%)	1886(21.4%)	
2	1641(23.3%)	1772(20.1%)	
3	688(9.8%)	1951(22.1%)	
4	805(11.4%)	1257(14.3%)	
Six-minute walk distance (meters)	415.4 ± 108.8	376.9 ± 129.1	< 0.001
FEV_1 post BD (ml)	1.7 ± 0.7	1.6 ± 0.8	< 0.001
FEV_1 post BD (%)	60.8 ± 22.1	54.8 ± 22.3	< 0.001
Smoker			< 0.001
Former	3989 (49.1%)	5392 (61.4%)	
Current	3589 (44.2%)	3174 (36.1%)	
Never	542 (6.7%)	222 (2.5%)	
Pack-years	46.4 ± 28.8	42.1 ± 28.3	< 0.001
Cough	1103 (54.2%)	342 (43.9%)	< 0.001
Sputum	1159 (42.1%)	341 (43.9%)	0.353
Diabetes	354 (6.7%)	303(16.6%)	< 0.001
Cardiac disease	1072 (30.8%)	467 (25.9%)	< 0.001
Chronic bronchitis	166 (38.7%)	787 (69.5%)	< 0.001
Hypertension	454 (40.4%)	826 (44.8%)	0.028

Asthma	1243 (26.1%)	209 (10.7%)	< 0.001
Spirometry staging			< 0.001
1	1567(19%)	1153(13.1%)	
2	3892(47.1%)	3711(42.1%)	
3	2126(25.8%)	2654(30.1%)	
4	671(8.1%)	1301(14.8%)	
Long term oxygen therapy	119 (1.4%)	430 (4.8%)	0.259

Table 1: Comparison of demographic and clinical characteristics in 3CIA COPD patients included/excluded in this analysis

Data are n (%), mean (SD) or median (IQR). BMI: Body mass index (kg/m^2). MRC: Medical Research Council. FEV_1: forced expiratory volume in 1 second. BD: bronchodilator.

Table 2: Mortality risk among COPD patients according to GOLD 2019 and GOLD 2015 classifications

	GOLD 2015 5-year mortality	GOLD 2019 5-year mortality
Group A	5.8 %	7.5 %
Group B	13.8 %	23.2 %
Group C	14.1 %	14.8 %
Group D	30.8 %	32.8 %

Table 3: Five-year all-cause mortality among spirometry strata within GOLD 2019 ABCD classification

GOLD 2019	A	B	C	D
Spirometry I	3.4 %	9.8 %	16.2 %	5.5 %
Spirometry II	7 %	14.9 %	8.3 %	22.1 %
Spirometry III	13 %	23.7 %	17.7 %	32.6 %
Spirometry IV	16.7 %	39.1 %	29.5 %	46 %

Table 4: Accuracy for predicting mortality of the ABCD groups classifications by GOLD 2015 and GOLD 2019 schemes

CLASSIFICATION	Sensitivity (% 95 CI)	PPV (% 95 CI)	Indice de Youden (95% CI)
GOLD 2015 group A	11.1 (9.6-12.7)	6.9 (5.9-7.9)	- 0.23 (-0.25 – -0.22)
GOLD 2015 group B	11.6 (10-13,2)	13.3 (11.5-15.1)	-0.06 (-0.08 – -0.04)
GOLD 2015 group C	14.6 (12.8-16.4)	20.2 (17.9-22.6)	+0.01 (-0.01 – +0.03)
GOLD 2015 group D	62.5 (60.1-64.9)	29.5 (28-31.1)	+0.28 (+0.25 – +0.31)
GOLD 2019 group A	19.1 (17.2-21.1)	9.5 (8.4-10.5)	-0.23 (-0.25 – -0.21)
GOLD 2019 group B	43.6 (41.1-46.1)	21.2(19.8-22.6)	+0.06 (+0.03 – +0.09)
GOLD 2019 group C	6.6 (5.3-7.8)	21.7 (18-25.5)	+0.01 (0.0 – +0.02)
GOLD 2019 group D	30.6 (28.3-32.8)	32.8 (30.4-35.2)	+0.16 (+0.14 – +0.18)

PPV: Positive predictive value. CI: confidence interval

TABLES AND FIGURES:

Table 1: Comparison of demographic and clinical characteristics in 3CIA COPD patients included/excluded in this analysis

Table 2: Mortality risk among COPD patients according to GOLD 2019 and GOLD 2015 classifications

Table 3: Five-year all-cause mortality among spirometry strata within GOLD 2019 ABCD classification

Table 4: Accuracy for predicting mortality of the ABCD groups classifications by GOLD 2015 and GOLD 2019 schemes

Figure 1: Distribution of participants by classification in GOLD 2015 and GOLD 2019

Figure 2: Kaplan- Meier survival curves by GOLD 2015 and GOLD 2019

Figure 3: Kaplan- Meier survival curves by GOLD 2019 and spirometry subgroups (IA-IVD)

Figure 4: Receiver operating curves for all cause mortality at 5 years follow-up

Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2015 and GOLD 2019 staging: a pooled analysis of individual patient data

Elena García Castillo[1] *, Tamara Alonso Pérez[1] *, Julio Ancochea[1], Maria Teresa Pastor Sanz[1], Pere Almagro[2], Pablo Martínez-Camblor[3], Marc Miravitlles[4], Mónica Rodríguez- Carballeira[5], Annie Navarro[6], Bernd Lamprecht[7], Ana S Ramírez-García Luna[8], Bernhard Kaiser[9], Inmaculada Alfageme[10], Ciro Casanova[11], Cristóbal Esteban[12], Juan J Soler-Cataluña[13], Juan P de-Torres[14], Bartolomé R Celli[15], Jose M Marín[16], Gerben ter Riet[17], Patricia Sobradillo[18], Peter Lange[19], Judith García-Aymerich[20,21,22], Josep M Anto[20,21,22], Alice M Turner[23], MeiLan K Han[24], Arnulf Langhammer[25], Sigrid Anna Aalberg Vikjord[25], Alice Sternberg[26], Linda Leivseth[27], Per Bakke[28], Ane Johannessen[29], Toru Oga[30], Borja Cosío[31], Andrés Echazarreta[32], Nicolás Roche[33], Pierre-Régis Burgel[34], Don D Sin, Milo[34,35], Milo A Puhan[36], Jose Luis López Campos[37], Laura Carrasco[38], Joan B Soriano[39] & for the 3CIA collaboration.

* Shared first authors

Author information:

1. Pneumology Department, Hospital Universitario de la Princesa. Instituto de Investigación Hospital Universitario de la Princesa (IISP), Universidad Autónoma de Madrid, Madrid, Spain.
2. Multimorbidity Patients Unit, Internal Medicine Dept, University Hospital Mutua de Terrassa, University of Barcelona, Barcelona,Spain.
3. Geisel School of Medicine at Dartmouth, Hanover, NH, USA.
4. Pneumology Department, Hospital Universitary Vall d'Hebron, CIBER de Enfermedades Respiratorias (CIBERES), Barcelona , Spain
5. Multimorbidity Patients Unit, Internal Medicine, Hospital Universitari Mutua de Terrassa, Universitat de Barcelona, Barcelona , Spain.
6. Pneumology Service, Hospital Universitari Mútua Terrassa, Barcelona, Spain
7. Department of Pulmonary Medicine, Kepler-University-Hospital. Faculty of Medicine , Johannes-Kepler-University Linz , Austria.
8. Facultad de Medicina, Universidad Autónoma de San Luis Potosí, San Luis Potosí, Mexico.
9. Department of Pulmonary Medicine , Paracelsus Medical University Hospital , Salzburg, Austria.
10. Departamento de Medicina. Universidad de Sevilla. HU Virgen de Valme. Sevilla. Spain.
11. Hospital Universitario Nuestra Señora de La Candelaria, Universidad de La Laguna, Tenerife, Spain.

12. Pulmonology Department. Hospital Galdakao-Usansolo, Galdakao , Bizkaia , Spain.
13. Servicio de Neumología, Hospital Arnau de Vilanova, Valencia , Spain.
14. Clinica Universidad de Navarra, Pamplona, Spain.
15. Pulmonary and Critical Care Medicine , Harvard University, Brigham and Women's Hospital , Boston , MA , USA
16. Hospital Universitario Miguel Servet, Zaragoza, and CIBER de Enfermedades Respiratorias (CIBERES), Spain
17. Department of General Practice , Academic Medical Center, University of Amsterdam (AMC) , Amsterdam , The Netherlands.
18. Hospital Universitario Araba, Sede Txagorritxu, Vitoria, Spain for Universitary Hospital of Cruces in Barakaldo, Spain.
19. Section of Social Medicine, Department of Public Health , Copenhagen University, Copenhagen City Heart Study, Frederiksberg Hospital, Frederiksberg , Copenhagen , Denmark
20. Global, Centre for Research in Environmental Epidemiology (CREAL) , Barcelona , Spain.
21. Department of Experimental and Health Sciences , Universitat Pompeu Fabra (UPF) , Barcelona, Spain.
22. CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain.
23. Institute of Applied Health Research, University of Birmingham , Edgbaston , UK.
24. Internal Medicine, Division of Pulmonary and Critical Care Medicine, University of Michigan , Ann Arbor , MI , USA.
25. Department of Public Health and Nursing , NTNU, Norwegian University of Science and Technology , Trondheim , Norway.
26. Department of Epidemiology , Johns Hopkins Bloomberg School of Public Health , Baltimore , MD , USA.
27. Centre for Clinical Documentation and Evaluation , Northern Norway Regional Health Authority, Tromso , Norway
28. Department of Clinical Science , University of Bergen , Bergen , Norway.
29. Department of Global Public Health and Primary Care , University of Bergen , Bergen, Bergen , Norway.
30. Department of Respiratory Care and Sleep Control Medicine , Kyoto University , Kyoto , Japan.
31. Department of Respiratory Medicine , Hospital Son Espases-IdISPa, Ciberes, Mallorca , Spain.
32. Servicio de Neumonología, Hospital San Juan de Dios de La Plata , Buenos Aires , Argentina.
33. Respiratory Medicine, Cochin Hospital APHP, University Paris Descartes , Paris , France.
34. James Hogg Research Centre , University of British Columbia , Vancouver , BC , Canada.
35. Division of Respiratory Medicine, Department of Medicine , St Paul's Hospital , Vancouver , BC , Canada.
36. Epidemiology , Biostatistics and Prevention Institute, University of Zurich , Zurich , Switzerland
37. Pulmonology Department. Hospital Universitario Virgen del Rocío, Sevilla, España
38. Pulmonology Department. Hospital Universitario Virgen del Rocío, Sevilla, España.

39. Instituto de Investigación Hospital Universitario de la Princesa (IISP), Universidad Autónoma de Madrid, Madrid, Spain

Address for correspondence:

Dr. Joan B Soriano

Hospital Universitario de la Princesa, Diego de León 62, Neumología 6ª planta, 28006-Madrid, Spain

Email: jbsoriano2@gmail.com

Cellular: +34 618 86 77 69

File information:

Date: JulyApril, 2020

Word count: 23,83895 words

References: 391 references

Illustrations: 4 tables and 4 figures

Keywords: COPD, mortality, survival

Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2015 and GOLD 2019 staging: a pooled analysis of individual patient data

Abstract

In 2019, The Global Initiative for Chronic Obstructive Lung Disease (GOLD) modified the grading system for patients with chronic obstructive pulmonary disease (COPD) creating 16 subgroups (1A-4D). As part of the COPD Cohorts Collaborative International Assessment (3CIA) initiative, we aim to compare mortality prediction of 2015 and 2019 COPD GOLD staging systems.

We studied 17 139 COPD patients from the 3CIA study, selecting those with complete data. Patients were classified by the 2015 and 2019 GOLD ABCD systems and we compared the predictive ability for 5-year mortality of both classifications.

17 139 patients with COPD were enrolled in 22 cohorts of 11 countries between 2003 to 2017; 8 823 of them had complete data and were analyzed. Mean age was 63.9 years (SD 9.8) and 62.9% were male. GOLD 2019 classified the patients in milder degrees of COPD. For both classifications, group D had higher mortality. Five-year mortality did not differ between groups B and C in GOLD 2015; in GOLD 2019, mortality was greater for group B than C. Patients classified as group A and B had better sensitivity and positive predictive value with GOLD 2019 classification than GOLD 2015. GOLD 2015 had better sensitivity for group C and D than GOLD 2019. The AUC for 5-year mortality were only 0.67 (95% CI 0.66-0.68) for GOLD 2015 and 0.65 (95% CI 0.63-0.66) for GOLD 2019.

The new GOLD 2019 classification does not predict mortality better than the previous GOLD 2015 system.

Take Home message:

GOLD 2019 staging system created 16 subgroups

GOLD 2015 and GOLD 2019 are not strong predictors of mortality and don't have sufficient discriminatory power to be used as a tool for risk classification of mortality in patients with COPD

Introduction

Chronic obstructive pulmonary disease (COPD) is a common cause of morbidity and mortality in the world. COPD affects approximately 328 million people worldwide, and COPD related deaths amount to 4 million deaths every year[1]. Assessment of disease severity is essential to predict prognosis and to standardize treatment regimes. The Global Initiative for Chronic Obstructive Lung Disease (GOLD) document is the most widely used treatment guide for the staging and management of COPD. The GOLD grading system for COPD has significantly evolved since first publication in 2001 to the current version in 2019. Initially, in the GOLD 2007 classification, only post-bronchodilator airflow limitation based on spirometry forced expiratory volume in 1 second (FEV_1) was used to grade the severity of COPD[2]. Later on, some criticism arosed on this grading score because it relied only upon FEV_1 which is not a good predictor of dyspnoea, quality of life or exercise tolerance. Further, other important variables to evaluate prognosis in COPD, such as sub-phenotypes, exacerbations, dyspnoea severity or comorbidities have been proposed, among others. Therefore, GOLD 2011 proposed a classification system of four groups, ABCD, combining FEV_1 and two clinical parameters: history of exacerbations and respiratory symptoms measured by the modified Medical Research Council (mMRC) dyspnoea score, or the COPD Assesment Test score (CAT) [3]. The 2011 ABCD classification was considered an improvement in the management of patients with COPD, providing an opportunity to further guide the individualized care of these patients. GOLD 2011 predicted future exacerbations better than GOLD 2007 but there was no difference in mortality predictions or respiratory outcomes [4-7]. In 2015 an updated report was published with the same measurement parameters (FEV_1, dyspnoea and exacerbations) than 2011 classification[8].

The latest GOLD update, published in 2019 uses a composite of spirometry, symptoms and exacerbations, but importantly separating the spirometric 1-4 staging from the ABCD groups[9]. This separation is relevant because it is known there are differences in the rate of exacerbations for the most severe COPD patients, depending on whether the risk is based on pulmonary function tests, on the history of exacerbations or both[10]. All these classifications were

initially designed not to assess prognosis, but to aid clinicians in creating optimal treatment regimes for patients. Thus, the prognostic ability of GOLD 2019 compared to previous classifications is largely unknown, with only a few published studies[11,12]. To address this issue, we used pooled data from 17,139 patients of 22 COPD cohorts and eleven countries and compared the prognostic capacity of the 2019 versus 2015 GOLD staging classifications to predict mortality.

Methods

Study population

In this international study, we assessed 17,139 patients from the COPD Cohorts Collaborative International Assessment (3CIA) initiative. All were prospective cohorts that recruited patients within the period 1999 to 2017, except one of them that was a population based cohort. All patients had a definition of COPD characterized by spirometry, that is post-bronchodilator FEV_1 to forced vital capacity ratio (FVC) ratio less than 0.7 and a clinical diagnosis of COPD. Spirometry was performed using the standards provided by the American Thoracic Society and European Respiratory Society[13]. The primary investigators of each of the participating 3CIA cohorts provided individual patient's data for pooled analysis. We obtained a minimum individual dataset including the vital status (up to death, right truncation or 2017), age, sex, smoking status, pre-bronchodilator and post-bronchodilator FEV_1 and FVC and dyspnoea measured with the modified mMRC, among others[14]. Only in a number of 3CIA cohorts were data of number of exacerbations in the previous year available. For the current study, we selected exclusively those cohorts in which the number of exacerbations in the previous year were available in the database, since this variable is required to calculate the GOLD 2015 and 2019 grading systems. Fifteen out of a total of 22 cohorts contained data on history of exacerbations, so that the final number of patients available to be classified according to GOLD 2015 and GOLD 2019 was 8,823. Symptoms were assessed using the mMRC dyspnoea scale. To determine the risk descriptor in the 2015 grouping system, we used exacerbations history and GOLD spirometry stages. The combination of symptoms (mMRC) and the worse risk descriptor (spirometry or

exacerbation history) were used to classify patients by GOLD 2015 system. Participants were classified using the GOLD 2019 system into four grades (1-4) based on postbronchodilator FEV_1 percentage of prediction as stage 1 ($FEV_1 \geq 80$), stage 2 (FEV_1 79-50), stage 3 (FEV_1 30-49) and stage 4 ($FEV_1 < 30$). Groups ABCD were defined by self-reported severity of dyspnea (mMRC) and number of exacerbations in the previous year.

All participants provided informed written consent, and each study was conducted with the formal approval of the local ethics institutional committees following the principles of the Declaration of Helsinki.

Outcomes

The primary outcome was the prediction ability of all-cause mortality in the individuals by the two GOLD systems.

Statistical analysis

The 3CIA database manager quality-controlled all data centrally and created a clean database with a data dictionary. All implausible or missing variables were queried with the original study investigators, and datum were removed from the central database if errors could not be corrected. Because the cohorts had different follow-up times, patients were right-censored at five years of follow-up.

Descriptive statistics used mean and standard deviation for continuous variables and the number of cases and percentages for categorical variables. Comparisons between groups were performed with the Chi^2 - test for categorical variables and the T-test for continuous variables.

We estimated 5-year all-cause mortality, according to GOLD 2015 and 2019 staging systems, using Kaplan-Meier survival statistics. Statistical comparisons were performed using the Log-Rank test Receiver Operation Characteristic (ROC) curve analyses, area under the curve (AUC) and the 95% CI of the AUC were calculated to measure the predictive accuracy for mortality. Also, we compared the prediction ability of mortality on both classifications using sensitivity, positive predictive value and the Youden's index with Epidat 3.1 program.

Results

We pooled data from 22 COPD cohorts with a total of 17,139 patients, finally including 8,823 patients from 15 cohorts that had all complete variables to be classified as GOLD 2015 and GOLD 2019. A comparison of baseline characteristics of included and not included patients is presented in Table 1. There were statistically significant differences in many variables given the large size, but most should be considered not clinically relevant (Table 1).

The 8,823 included patients were 62.9% male, with a mean age of 63.9 (SD± 9.8) years, mean±SD BMI 27.0±5.8 kg/m^2 and mMRC dyspnoea score of 1.8±1.4. Postbronchodilator FEV_1 was 54.8%±22.3 of the predictive value, and 6-minute walk distance was 376.9±129.1 meters. Based on spirometry staging 1,153 (13%) had mild (stage 1), 3,711 (42.1%) had moderate (stage 2), 2,654 (30.1%) had severe (stage 3) and 1,301 (14.8%) had very severe (stage 4) disease.

The distribution of these 8,823 patients according to GOLD 2015 and GOLD 2019 is presented in Figure 1. With GOLD 2019 there is a shift towards the less severe staging of disease (absolute increase in stage A and B of 7.8% and 22.2% respectively; and absolute decrease in stage C and D of 7.8 and 22.2% respectively).

The overall 5-year mortality rate was 18.3%. The all-cause 5-year mortality rates according to both classifications are shown in Table 2. Figure 2 shows Kaplan-Meier curves for 5-year mortality according to GOLD 2015 (Figure 2A) and GOLD 2019 (Figure 2B). All-cause mortality at 5 years in the 2015 GOLD classification was higher in grade D, followed by grades B, C (with similar mortality), and finally grade A, log-rank test $p<0.001$ (Table 2 and Figure 2A). Grade D diverged from the beginning of follow up, while grades B and C diverged after 1 year of follow-up. In GOLD 2019, the four Kaplan-Meier curves diverge during the first year, but interestingly, grade B had higher mortality than grade C, so mortality was higher in groups B and D (more symptoms) than in groups A and C (fewer symptoms; Figure 2B). The degree of obstruction measured by FEV_1 % further subclassified patients into 16 subgroups with different mortality rates, increasing from 1A to 4D in GOLD 2019 grading system (Table 3). Figure 3 shows Kaplan-Meier curves for each of the spirometry strata. The higher mortality of group B

over group C persisted in each of the strata with the exception of spirometry strata 1, with a higher mortality in group C. Similarly to GOLD 2015, in GOLD 2019, grade A and D had the lowest and highest mortality, respectively, with very similar absolute numbers (Table 2, Figure 2).

The primary outcome, the prediction capacity as measured by the AUC of ROC curve for mortality up to 5 years was intermediate (<0.70) for both systems (Figure 4). GOLD 2015 exhibited slightly better discrimination in predicting mortality (AUC 0.67 [95% CI 0.66-0.68]) than GOLD 2019 classification (AUC 0.64 [95% CI 0.63-0.66]).

Regarding sensitivity parameters, both classifications had quite shallow values. GOLD 2019 had a higher sensitivity for predicting mortality on A and B groups vs 2015 classification (19.1 vs 11.1 and 43.6 vs 11.6). On the other hand, GOLD 2015 had a higher sensitivity on groups C and D vs 2019 classification (14.6 vs 6.6 and 62.5 vs 30.6) based on overlapping 95% CIs (Table 4). The positive predictive values were also higher in GOLD 2019 group A and B vs the same groups in GOLD 2015 (9.5 vs. 6.9 and 21.2 vs. 13.3); but not different (overlapping confidence intervals) in groups C and D. The Youden indices were quite low for both classifications, even with negative values, showing that it is not an optimal classification system to assess mortality.

Discussion

Our study evaluates mortality according to the last two GOLD classifications, and it is one of the most extensive to date. The most important finding is that the new GOLD 2019 classification (based on symptoms or exacerbations along with stages spirometry analysed altogether) did not predict 5-year mortality better than GOLD 2015 classification (based on spirometry, history of exacerbation and symptoms). How to define and stage COPD exacerbations is a matter of an intense, long debate[15]. Most 3CIA individual cohorts used subsequent iterations of GOLD-accepted definitions in their protocols, mostly based on Rodriguez-Roisin R, et al. seminal paper[16], including mild (symptom-only) exacerbations. However, when pooling for 3CIA, we only focused on those COPD exacerbations that required health services use, ER, admission or death. In GOLD 2019 classification, subgroups B and D had the worse mortality,

highlighting that the higher burden of symptoms conveys a higher mortality independently of spirometry (Table 2 and Figure 2B). Finally, we show an important shift in the proportions of patients between the 2015 and 2019 ABCD grades, with milder severity in GOLD 2019.

GOLD 2015 and GOLD 2019 classifications had a low discriminatory power as per the AUCs, ranging from 0.67 to 0.65, and similar to other studies (by consensus, AUC below 0.70 are considered low or weak)[11]. Sensitivity and positive predictive values indicate that the general performance of the two models is similar and very low. Also, the Youden's indices are very low with negative values that have no meaningful interpretation in practice. These findings support the results of other studies, suggesting that GOLD classification is not a good predictor for mortality[11-12]. There may be various reasons for these poor results. The main reason is that these classifications were conceived to guide treatment, so it is not surprising that their capacity for predicting mortality is low. There are clinical phenotypes such as the asthma-COPD, the frequent exacerbator with emphysema or chronic bronchitis comorbidities and different indexes, that are significant predictors of mortality and are not entirely included in the GOLD stages[17-21]. In our study, sensitivity and positive predictive value are higher in GOLD 2019 groups A and B. On the contrary, sensitivity is higher in GOLD 2015 in groups C and D. These results suggest that GOLD 2019 predicts slightly better mortality on low risk groups (A and B) and GOLD 2015 in high risk groups (C and D), requiring more studies to corroborate it.

The discriminatory power of GOLD 2019 was lower than GOLD 2015 as shown in AUCs values. The partition of FEV_1 as a direct classifier in GOLD 2019 reduced ability to discriminate survival, highlighting the need to consider the severity of airflow obstruction in assessing mortality risk. When using GOLD 2019 with a composite of spirometry, exacerbations and symptoms (16 subgroups 1A to 4D classification), we found an increase in all-cause mortality between GOLD 2019 stage 1 and GOLD 2019 stage 4 across grades A, B and D, highlighting the persisting importance of FEV_1 as a predictor of mortality (Table 3 and Figure 3). In group C, mortality was higher in spirometry stage 1 than in 2, probably due to a significant difference of the proportion in patients between the two stages.

Our study confirms that patients classified as GOLD A had the best survival, and patients with GOLD D, the higher mortality in both classifications[20-24]. Mortality of groups B and C in GOLD 2015 lied in between A and D groups and often overlapped. In GOLD 2019 mortality was significantly higher in group B patients than in group C. This finding is similar as other previous reports published[25-27], showing that group B is an intermediate-high risk group associated with more exacerbations and likely to other comorbidities that may cause dyspnoea (such as heart failure). Furthermore, we show that the burden of symptoms (represented by groups B and D) have a prognostic value additive but independent to spirometry.

The current study demonstrates that the use of GOLD 2019 classification scheme shifted patients with COPD to groups of milder severity compared with GOLD 2015. This happened in keeping with previous reports, but in a smaller proportion of patients (30% of patients reassigned towards group A or B compared to 53% in Lee et al study or 66% in Tan et al study) [28-29]. The further distribution of spirometric parameters from two categories in GOLD 2015 (FEV_1 less or higher than 50%) to four categories in GOLD 2019, was one of the possible reasons for the patient's shift from C and D in GOLD 2015 to A and B groups in GOLD 2019. This phenomenon is opposed to the one observed with the use of the revised GOLD 2011 classification, which shifted the patients from the GOLD 2007 towards more advanced stages of the disease (D group increased almost 3 times) [7]. Remarkably, in both classifications, group C was consistently the smallest group (Figure 1), as seen in other reports[27]. The implication of progressively milder disease classifications for the treatment choices clinicians make in real life practice guided by GOLD is not yet clear given the recent nature of the latest GOLD iteration, but will be important to monitor.

Strength of our study include a large sample size, the study design (a pooled-analysis of individual patient-data from several cohorts), and the different degrees of severity of patients from different cohorts, maximizing its high external validity. Prospective data collection of spirometry with a post-bronchodilator test, dyspnea by mMRC scale, history of exacerbation and mortality enabled direct classification of patients by 2015 and 2019 GOLD staging schemes. We also have

a significant representation of women, of which other COPD studies might not have achieved[30,28].

Our study has several limitations. First, although we started with 17,139 patients with COPD, a considerable number of patients were excluded because of missing information on key variables, mainly regarding the history of exacerbations. These missing data are unlikely to affect the validity of our results, as we can see in our analysis comparing included and non-included patients. Second, the mortality analysis used all-cause death, and we have no data regarding specific causes of death (this data was not collected consistently in all cohorts). Third, symptoms were only evaluated using mMRC dyspnea score, but not with the COPD Assessment Test, or other instruments[31,29]; however, that is in line with other reported cohorts[32,0-33,1]. Fourth, most patients came from hospital-based cohorts, so we likely have more patients with moderate to severe disease and less patients with mild and moderate disease compared to an outpatient setting, or a primary care population. Indeed, the 22 cohorts from eleven countries in 3CIA within our initiative are only a sample representing the estimated 300 million COPD patients worldwide[34]. Finally, the recruitment timeframe was long (10 years), so evolving treatment recommendations might influence results.

In conclusion, this study of COPD cohorts, including 8823 patients, showed that neither GOLD 2015 nor GOLD 2019 are strong predictors of mortality. GOLD 2019 predicted mortality better than GOLD 2015 in groups A and B but worse in groups C and D. However, none of the GOLD classifications has sufficient discriminatory power to be used as a tool for risk classification of mortality in patients with COPD. Ours should be considered a constructive exercise and a critical appraisal of the last two GOLD iterations defining COPD. Within 3CIA, we have already suggested several proposals for future COPD staging and grading classifications, by applying more evidence-based thresholds of evidence-based variables [35-39].

Declaration of interest

There are no direct or indirect conflicts of interest to disclose related with this manuscript

References

[1] GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Respir Med. 2017; 5(9): Sep 691-706.

[2] Rabe KF, Hurd S, Anzueto A, et al; Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med. 2007;176(6):532-55.

[3] Vestbo J, Hurd SS, Agustí AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med 2013; 187:347-65.

[4] Han MK, Muellerova H, Curran-Everett D, et al. GOLD 2011 disease severity classification in COPDGene: a prospective cohort study. Lancet Respir Med.2013;1(1):43-50.

[5] Soriano JB, Alfageme I, Almagro P, et al. Distribution and prognostic validity of the new Global Initiative for Chronic Obstructive Lung Disease grading classification. Chest 2013;143(3):694-702.

[6] Lange P, Marrot JL, Vestbo J, et al. Prediction of the clinical course of chronic obstructive pulmonary disease, using the new GOLD classification: a study of the general population. Am J Respir Crit Care Med.2012;186(10):975-981.

[7] Soriano JB, Lamprecht B, Ramírez AS, et al. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. Lancet Respir Med 2015; 3: 443-50.

[8] Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the diagnosis, management and prevention of COPD 2015. 2015 [accessed 2019 June 21]. Available from: http://www.goldcopd.org/.

[9] Singh D, Agusti A, Anzueto A, Barnes PJ, Bourbeau J, Celli BR et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease: the GOLD science committee report 2019. Eur Respir J. 2019 May 18;53(5).

[10] Goossens LM, Leimer I, Metzdorf N, Becker K, Rutten-van Mölken N. Does the 2013 GOLD classification improve the ability to predict lung function decline, exacerbations and mortality: a post-hoc analysis of the 4-year UPLIFT trial. BMC Pulm Med. 2014; 14: 163.

[11] Gedebjerg A, Szépligeti SK, Wackerhausen LH, Horváth-Puhó E, Dahl R, Hansen JG et al. Prediction of mortality in patients with chronic obstructive pulmonary disease with the new Global Initiative for Chronic Obstructive Lung Disease 2017 classification: a cohort study. Lancet Respir Med. 2018 Mar;6(3):204-212.

[12] Aramburu A, Arostegui I, Moraza J, Barrio I, Aburto M, García-Loizaga A et al. COPD classification models and mortality prediction capacity. International Journal of COPD 2019:14;605-613

[13] Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al; ATS/ERS Task Force. Standardisation of spirometry. Eur Respir J 2005;26:319-338.

[14] Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. Usefulness of the Medical Research Council (MRC) dyspnea scale as a measure of disability in patients with chronic obstructive pulmonary disease. Thorax 1999;54:581-586.

[15] Mathioudakis AG, Moberg M, Janner J, Alonso-Coello P, Vestbo J. Outcomes reported on the management of COPD exacerbations: a systematic survey of randomised controlled trials. ERJ Open Res. 2019 May 10;5(2):00072-2019.

[16] Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. Chest. 2000 May;117(5 Suppl 2):398S-401S.

[175] Golpe R, Suárez-Valor M, Martín – Robles I, Sanjuán-López P, Cano-Jiménez E, Castro-Añón O et al. Mortality in COPD patients according to clinical phenotypes. International Journal of COPD 2018:13;1433-1439.

[186] Marin JM, Alfageme I, Almagro P, Casanova C, Esteban C, Soler-Cataluña JJ et al. Multicomponent indices to predict survival in COPD: the COCOMICS study. Eur Respir J 2013; 42: 323–332.

[197] Izquierdo JL, Miravitlles M, Esquinas C, Pérez M, Calle M, López Campos JL et al. Characteristics of COPD Patients Managed in Respiratory Medicine Departments in Spain, According to GOLD Groups and GesEPOC Clinical Phenotypes. Arch Bronconeumol. 2018;54:559-567

[2018] Radovanovic D, Contoli M, Marco FD, Sotgiu G, Pelaia G, Braido F et al. Clinical and Functional Characteristics of COPD Patients Across GOLD Classifications: Results of a Multicenter Observational Study. COPD. 2019;16:215-226.

[2119] Almagro P, Martínez-Camblor P, Miravitlles M, Rodríguez-Carballeira M, Navarro A, Lamprecht B et al. External Validation and Recalculation of the CODEX Index in COPD Patients. A 3CIAplus Cohort Study. COPD. 2019;16:8-17.

[220] Han MZ, Hsiue TR, Tsai SH, Huang TH, Liao XM, Chen CZ. Validation of the GOLD 2017 and new 16 subgroups (IA-4D) classifications in predicting exacerbation and mortality in COPD patients. Int J Chron Obstruct Pulmon Dis. 2018; 13:3425-3433.

[231] Flynn RW, MacDonaldn TM, Chalmers JD, Schembri S. The effect of changes to GOLD severity stage on long term morbidity and mortality in COPD. Respiratory Research 2018 19:249.

[242] de Torres JP, Casanova C, Marin JM, Pinto-Plata V, Divo M, Zulueta JJ, et al. Prognostic evaluation of COPD patients: GOLD 2011 versus BODE and the COPD comorbidity index COTE. Thorax. 2014;69:799 –804.

[253] Lange P, Marott JL, Vestbo J, Olsen KR, Ingebrigtsen TS, Dahl M et al. Prediction of the clinical course of chronic obstructive pulmonary disease, using the new GOLD classification: a study of the general population. Am J Respir Crit Care Med. 2012;186:975–81.

[264] Lee SJ, Yun SS, Ju S, You JW, Cho JJ, Jeong YY et al. Validity of the GOLD 2017 classification in the prediction of mortality and respiratory hospitalization in patients with chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis. 2019; 14: 911-919.

[275] Flynn RWV, MacDonald TM, Chalmers JD, Schembri S. The effect of changes to GOLD severity stage on long term morbidity and mortality in COPD. Respir Res. 2018;19:249.

[286] Tan WC, Bourbeau J, Aaron SD, et al. GOLD 2017 classification and lung function decline in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2018;1,197(5):670-673.

[297] Criner RN, Labaki WS, Regan EA, Bon JM, Soler X, Bhatt SP et al. for the COPDGene Investigators. Mortality and exacerbations by Global Initiative for Chronic Obstructive Lung Disease Groups ABCD: 2011 versus 2017 in the COPDGene Cohort. Chronic Obstr Pulm Dis. 2019 Jan 10;6(1):64-73

[3028] Kobayashi S, Hanagama M, Ishida M, Sato H, Ono M, Yamanda S et al. Clinical characteristics and outcomes in Japanese patients with COPD according to the 2017 GOLD classification: the Ishinomaki COPD Network Registry. Int J Chron Obstruct Pulmon Dis.2018:13;3947-3955.

[3129] Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the Asthma Control Test: a survey for assessing asthma control. J Allergy Clin Immunol 2004;113:59-65.

[320] Johannessen A, Nilsen RM, Storebo M, Gulsvik A, Eagan T, Bakke P. Comparison of 2011 and 2007 global initiative for chronic obstructive lung disease guidelines for predicting mortality and hospitalization. Am J Respir Crit Care Med. 2013;188:51–9.

[331] Agusti A, Edwards LD, Celli B, Macnee W, Calverley PM, Mullerova H, Lomas DA, Wouters E, Bakke P, Rennard S, et al. Characteristics, stability and outcomes of the 2011 GOLD COPD groups in the ECLIPSE cohort. Eur Respir J. 2013;42:636–46.

[34] GBD Chronic Respiratory Disease Collaborators. Prevalence and attributable health burden of chronic respiratory diseases, 1990 – 2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Respir Med. 2020;8(6):585-96.

[35] Almagro P, Martínez-Camblor P, Miravitlles M, Rodríguez-Carballeira M, Navarro A, Lamprecht B, et al. External Validation and Recalculation of the CODEX Index in COPD Patients. A 3CIAplus Cohort Study. COPD. 2019 Feb;16(1):8-17.

[36] Guerra B, Haile SR, Lamprecht B, Ramírez AS, Martinez-Camblor P, Kaiser B, et al. Large-scale external validation and comparison of prognostic models: an application to chronic obstructive pulmonary disease. BMC Med. 2018 Mar 2;16(1):33.

[37] Haile SR, Guerra B, Soriano JB, Puhan MA; 3CIA collaboration. Multiple Score Comparison: a network meta-analysis approach to comparison and external validation of prognostic scores. BMC Med Res Methodol. 2017 Dec 21;17(1):172.

[38] Burgel PR, Paillasseur JL, Janssens W, Piquet J, Ter Riet G, Garcia-Aymerich J, et al. Initiatives BPCO, EABPCO, Leuven and 3CIA study groups. A simple algorithm for the identification of clinical COPD phenotypes. Eur Respir J. 2017 Nov 2;50(5):1701034.

[39] Soriano JB, Lamprecht B, Ramírez AS, Martinez-Camblor P, Kaiser B, Alfageme I, et al. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. Lancet Respir Med. 2015 Jun;3(6):443-50.

TABLES

	Excluded (n=8316)	**Included (n=8823)**	**p-value**
Age (years)	64.2 ± 10.7	63.9 ± 9.8	0.08
Male sex	6232 (74.9%)	5552 (62.9%)	< 0.001
BMI	26.5 ± 4.9	27.0 ± 5.8	< 0.001
Modified MRC dyspnea scale	1.5 ± 1.3	1.8 ± 1.4	< 0.001
0	1647 (23.4%)	1957(22.2%)	< 0.001
1	2261(32.1%)	1886(21.4%)	
2	1641(23.3%)	1772(20.1%)	
3	688(9.8%)	1951(22.1%)	
4	805(11.4%)	1257(14.3%)	
Six-minute walk distance (meters)	415.4 ± 108.8	376.9 ± 129.1	< 0.001
FEV_1 post BD (ml)	1.7 ± 0.7	1.6 ± 0.8	< 0.001
FEV_1 post BD (%)	60.8 ± 22.1	54.8 ± 22.3	< 0.001
Smoker			< 0.001
Former	3989 (49.1%)	5392 (61.4%)	
Current	3589 (44.2%)	3174 (36.1%)	
Never	542 (6.7%)	222 (2.5%)	
Pack-years	46.4 ± 28.8	42.1 ± 28.3	< 0.001
Cough	1103 (54.2%)	342 (43.9%)	< 0.001
Sputum	1159 (42.1%)	341 (43.9%)	0.353
Diabetes	354 (6.7%)	303(16.6%)	< 0.001
Cardiac disease	1072 (30.8%)	467 (25.9%)	< 0.001
Chronic bronchitis	166 (38.7%)	787 (69.5%)	< 0.001
Hypertension	454 (40.4%)	826 (44.8%)	0.028

Asthma	1243 (26.1%)	209 (10.7%)	< 0.001
Spirometry staging			< 0.001
1	1567(19%)	1153(13.1%)	
2	3892(47.1%)	3711(42.1%)	
3	2126(25.8%)	2654(30.1%)	
4	671(8.1%)	1301(14.8%)	
Long term oxygen therapy	119 (1.4%)	430 (4.8%)	0.259

Table 1: Comparison of demographic and clinical characteristics in 3CIA COPD patients included/excluded in this analysis

Data are n (%), mean (SD) or median (IQR). BMI: Body mass index (kg/m^2). MRC: Medical Research Council. FEV_1: forced expiratory volume in 1 second. BD: bronchodilator.

Table 2: Mortality risk among COPD patients according to GOLD 2019 and GOLD 2015 classifications

	GOLD 2015 5-year mortality	GOLD 2019 5-year mortality
Group A	5.8 %	7.5 %
Group B	13.8 %	23.2 %
Group C	14.1 %	14.8 %
Group D	30.8 %	32.8 %

Table 3: Five-year all-cause mortality among spirometry strata within GOLD 2019 ABCD classification

GOLD 2019	A	B	C	D
Spirometry I	3.4 %	9.8 %	16.2 %	5.5 %
Spirometry II	7 %	14.9 %	8.3 %	22.1 %
Spirometry III	13 %	23.7 %	17.7 %	32.6 %
Spirometry IV	16.7 %	39.1 %	29.5 %	46 %

Table 4: Accuracy for predicting mortality of the ABCD groups classifications by GOLD 2015 and GOLD 2019 schemes

CLASSIFICATION	Sensitivity (% 95 CI)	PPV (% 95 CI)	Indice de Youden (95% CI)
GOLD 2015 group A	11.1 (9.6-12.7)	6.9 (5.9-7.9)	- 0.23 (-0.25 – -0.22)
GOLD 2015 group B	11.6 (10-13,2)	13.3 (11.5-15.1)	-0.06 (-0.08 – -0.04)
GOLD 2015 group C	14.6 (12.8-16.4)	20.2 (17.9-22.6)	+0.01 (-0.01 – +0.03)
GOLD 2015 group D	62.5 (60.1-64.9)	29.5 (28-31.1)	+0.28 (+0.25 – +0.31)
GOLD 2019 group A	19.1 (17.2-21.1)	9.5 (8.4-10.5)	-0.23 (-0.25 – -0.21)
GOLD 2019 group B	43.6 (41.1-46.1)	21.2(19.8-22.6)	+0.06 (+0.03 – +0.09)
GOLD 2019 group C	6.6 (5.3-7.8)	21.7 (18-25.5)	+0.01 (0.0 – +0.02)
GOLD 2019 group D	30.6 (28.3-32.8)	32.8 (30.4-35.2)	+0.16 (+0.14 – +0.18)

PPV: Positive predictive value. CI: confidence interval

TABLES AND FIGURES:

Table 1: Comparison of demographic and clinical characteristics in 3CIA COPD patients included/excluded in this analysis

Table 2: Mortality risk among COPD patients according to GOLD 2019 and GOLD 2015 classifications

Table 3: Five-year all-cause mortality among spirometry strata within GOLD 2019 ABCD classification

Table 4: Accuracy for predicting mortality of the ABCD groups classifications by GOLD 2015 and GOLD 2019 schemes

Figure 1: Distribution of participants by classification in GOLD 2015 and GOLD 2019

Figure 2: Kaplan- Meier survival curves by GOLD 2015 and GOLD 2019

Figure 3: Kaplan- Meier survival curves by GOLD 2019 and spirometry subgroups (IA-IVD)

Figure 4: Receiver operating curves for all cause mortality at 5 years follow-up

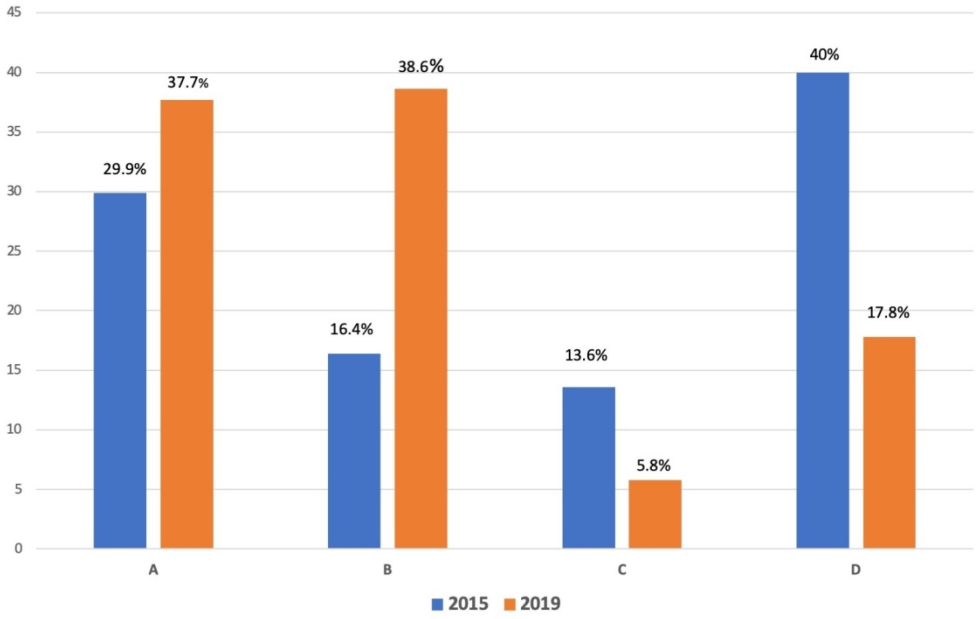


Figure 1: Distribution of participants by classification in GOLD 2015 and GOLD 2019

136x89mm (300 x 300 DPI)

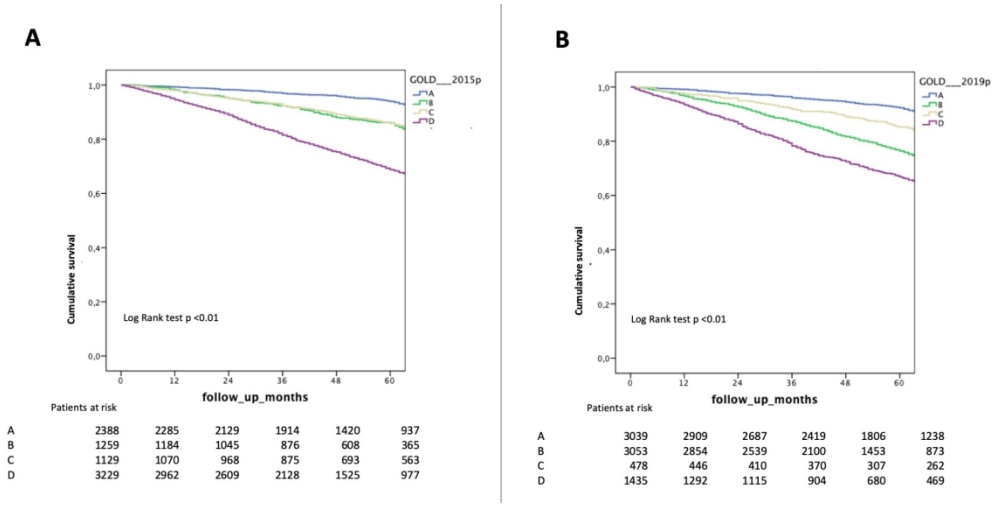


Figure 2: Kaplan- Meier survival curves by GOLD 2015 and GOLD 2019

319x159mm (169 x 169 DPI)

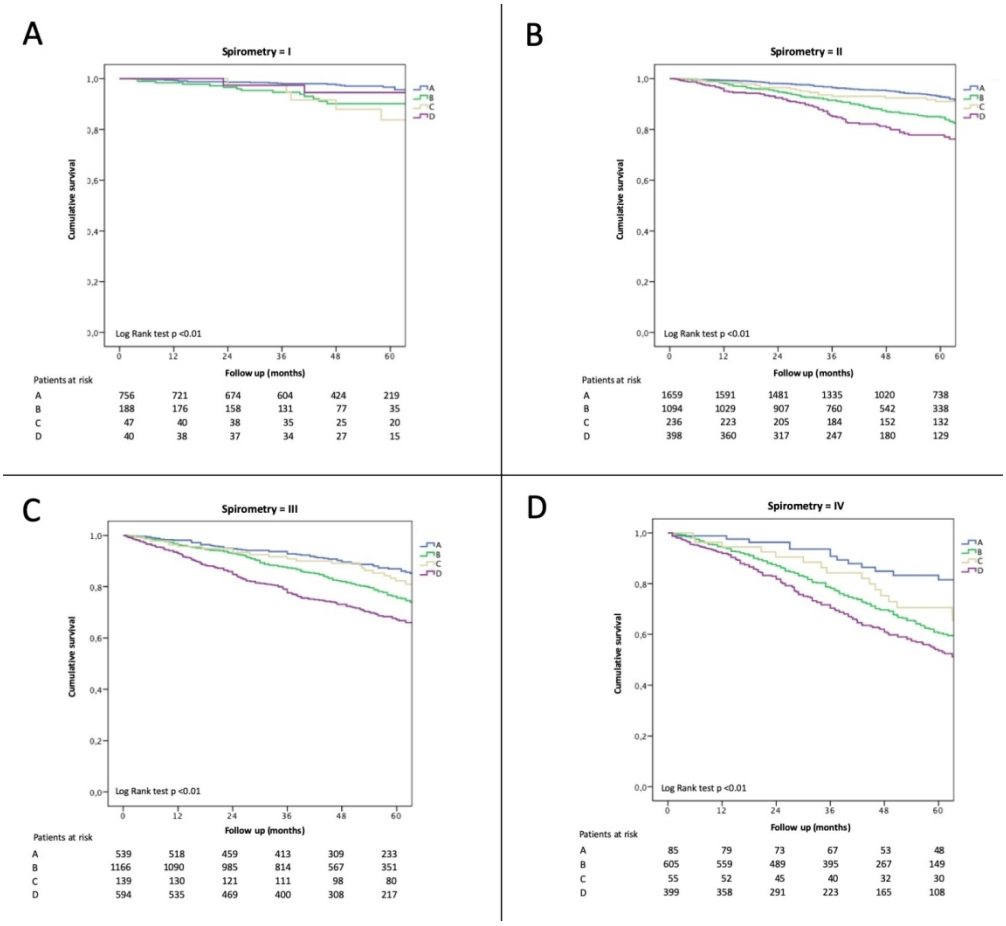
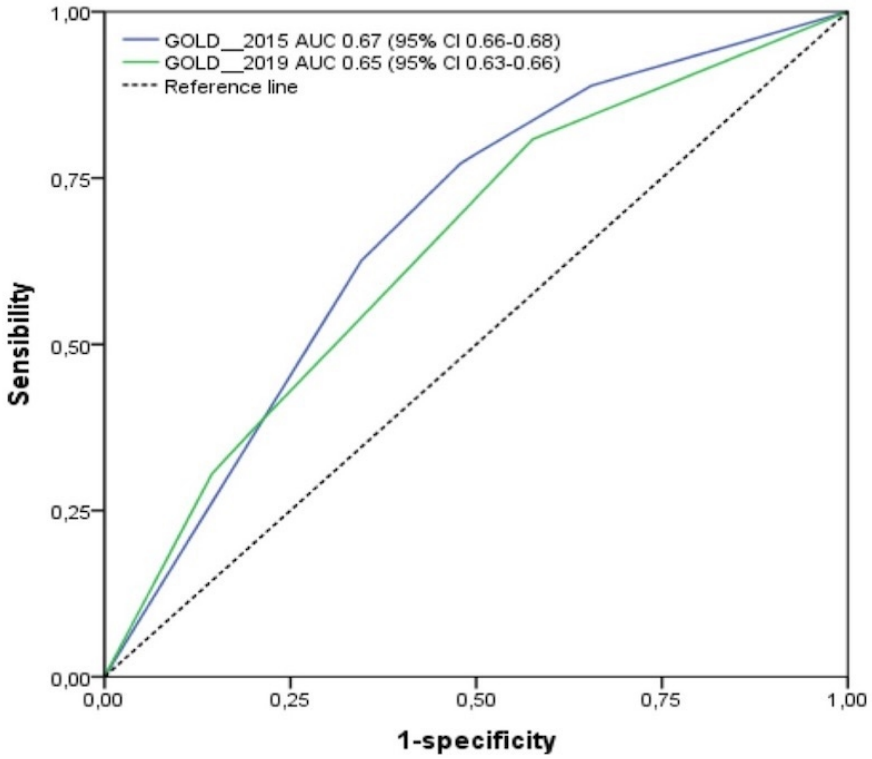


Figure 3: Kaplan- Meier survival curves by GOLD 2019 and spirometry subgroups (IA-IVD)

135x124mm (300 x 300 DPI)



AUC: area under the curve

Receiver operating curves for all cause mortality at 5 years follow-up

76x67mm (300 x 300 DPI)