International Journal of Chronic Obstructive Pulmonary Disease Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

# GOLD Classifications, COPD Hospitalization, and All-Cause Mortality in Chronic Obstructive Pulmonary Disease: The HUNT Study

This article was published in the following Dove Press journal:
*International Journal of Chronic Obstructive Pulmonary Disease*



Laxmi Bhatta[1] Linda Leivseth[2] Xiao-Mei Mai[1] Anne Hildur Henriksen[3,4] David Carslake[5,6] Yue Chen[7] Arnulf Langhammer[8,*] Ben Michael Brumpton[4,5,9,*]

[1]Department of Public Health and Nursing, NTNU Norwegian University of Science and Technology, Trondheim, Norway; [2]Centre for Clinical Documentation and Evaluation (SKDE), Northern Norway Regional Health Authority, Tromsø, Norway; [3]Department of Circulation and Medical Imaging, NTNU Norwegian University of Science and Technology, Trondheim, Norway; [4]Clinic of Thoracic and Occupational Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway; [5]Medical Research Council Integrative Epidemiology Unit, Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK; [6]Population Health Sciences, University of Bristol, Bristol, UK; [7]School of Epidemiology and Public Health, University of Ottawa, Ottawa, Canada; [8]HUNT Research Centre, Department of Public Health and Nursing, NTNU Norwegian University of Science and Technology, Levanger, Norway; [9]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU Norwegian University of Science and Technology, Trondheim, Norway

*These authors contributed equally to this work

Correspondence: Laxmi Bhatta
Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU Norwegian University of Science and Technology, P.O. Box 8905, MTFS, Trondheim NO-7491, Norway
Email laxmi.bhatta@ntnu.no

**Purpose:** The Global Initiative for Chronic Obstructive Lung Disease (GOLD) has published three classifications of COPD from 2007 to 2017. No studies have investigated the ability of these classifications to predict COPD-related hospitalizations. We aimed to compare the discrimination ability of the GOLD 2007, 2011, and 2017 classifications to predict COPD hospitalization and all-cause mortality.

**Patients and Methods:** We followed 1300 participants with COPD aged ≥40 years who participated in the HUNT Study (1995–1997) through to December 31, 2015. Survival analysis and time-dependent area under receiver operating characteristics curves (AUC) were used to compare the discrimination abilities of the GOLD classifications.

**Results:** Of the 1300 participants, 522 were hospitalized due to COPD and 896 died over 20.4 years of follow-up. In adjusted models, worsening GOLD 2007, GOLD 2011, or GOLD 2017 categories were associated with higher hazards for COPD hospitalization and all-cause mortality, except for the GOLD 2017 classification and all-cause mortality ($p_{trend}$=0.114). In crude models, the AUCs (95% CI) for the GOLD 2007, GOLD 2011, and GOLD 2017 for COPD hospitalization were 63.1 (58.7–66.9), 60.9 (56.1–64.4), and 56.1 (54.0–58.1), respectively, at 20-years' follow-up. Corresponding estimates for all-cause mortality were 57.0 (54.8–59.1), 54.1 (52.1–56.0), and 52.6 (51.0–54.3). The differences in AUCs between the GOLD classifications to predict COPD hospitalization and all-cause mortality were constant over the follow-up time.

**Conclusion:** The GOLD 2007 classification was better than the GOLD 2011 and 2017 classifications at predicting COPD hospitalization and all-cause mortality.

**Keywords:** GOLD grades, ABCD groups, COPD hospitalization, mortality, area under curve, AUC

## Introduction

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) published classifications of COPD in 2007, 2011 and recently in 2017.[1] In 2007, GOLD grades were introduced based on the severity of airflow limitation. In 2011, ABCD groups were introduced by combining severity of airflow limitation with exacerbation history and symptom burden.[2,3] In 2017, the ABCD groups were updated to include only exacerbation history and symptom burden, and used severity of airflow limitation separately.[1]

Although the GOLD classifications were meant to guide therapy, clinicians use the prognostic value of these classifications for risk classification at an individual

level.[4] Studies have compared the discrimination abilities of the GOLD 2007 and GOLD 2011 classifications to predict exacerbation and mortality.[4–8] Johannessen et al[5] found that the GOLD 2007 and GOLD 2011 classifications predicted respiratory hospitalization similarly well; however, Lange et al[6] and Chen et al[7] found that the GOLD 2011 classification was better than the GOLD 2007. In a pooled analysis of 22 cohorts, the GOLD 2007 and GOLD 2011 classifications did not differ significantly in predicting mortality.[4]

Since the publication of the GOLD 2017 classification, the prognostic value of this classification has been debatable with less clear findings.[9–12] To our knowledge, only one study has compared all three GOLD classifications and found that the GOLD 2017 classification predicted respiratory and all-cause mortality similarly well as the GOLD 2007 and GOLD 2011 classifications.[9] No previous studies have investigated the risk of COPD hospitalization using the GOLD 2017 classification or compared their discrimination ability with the previous GOLD classifications to predict COPD hospitalization. Criner et al[11] found that the prediction ability of the GOLD 2011 and GOLD 2017 classification was similar for exacerbation, however, exacerbations were used as binary rather than time-to-event data. They found that GOLD 2011 classification predicted mortality better than the GOLD 2017 classification.[11]

We aimed to compare the discrimination abilities of all three GOLD classifications to predict COPD hospitalization and all-cause mortality in a large Norwegian COPD cohort over a 20-year follow-up.

# Materials and Methods

## Study Population

Trøndelag is a county in central Norway with a homogenous and stable population. The HUNT Study invited the entire adult population (≥20 years) of northern Trøndelag to attend clinical examinations and answer questionnaires.[13]

The study cohort included people aged ≥40 years who participated in HUNT2 (1995–1997, n=44,384, 75.2% participation). A 5% random sample (n=2300) and persons reporting asthma related symptoms, diagnosis or medication use (n=7123) were invited to perform spirometry.[14] Participants from rural municipalities and participants from urban municipalities having an airflow limitation [pre-bronchodilator (BD) $FEV_1$/forced vital capacity (FVC)<0.75 or percent-predicted forced expiratory volume in first second ($ppFEV_1$)<80 using the European Coal and Steel Community (ECSC) equations[15]] were invited to attend post-BD spirometry (n=5678). We used the airflow limitation criteria to allow for future changes in diagnosis and severity classification of COPD. Among those performing post-BD spirometry (n=4178, 73.6% of invited), 3840 (91.9%) had acceptable spirometry maneuvers. Participants with acceptable post-BD spirometry and COPD were included in the analysis (n=1300) (Supplementary Figure S1). COPD was defined as participants having post-BD $FEV_1$/FVC<0.70 and [respiratory symptoms (daily cough in periods, cough with phlegm, wheezing, and dyspnea) or self-reported doctor-diagnosed COPD].[1]

This study was conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the Regional Committee for Medical Research Ethics (2015/1461/REK midt). All participants gave informed written consent.

## Spirometry and Lung Function Classification

Spirometry was performed before and 30 mins after inhalation of 1 mg terbutaline according to the 1994 ATS-guidelines.[16,17] Quality assurance of spirometry is described in detail elsewhere.[16,18]

The GOLD 2007 classification (GOLD grades)[1] was categorized as grade 1 ($ppFEV_1$≥80), grade 2 (80>$ppFEV_1$ ≥50), grade 3 (50>$ppFEV_1$≥30), or grade 4 ($ppFEV_1$<30). The GLI-2012 reference equation was used to calculate $ppFEV_1$ and ppFVC.[16,19]

The GOLD 2011 (ABCD groups) and GOLD 2017 classification (new ABCD groups)[1,3] were categorized as group A, B, C, or D (Supplementary-Figure S2). We defined symptom burden as low or high based on our local dyspnea scale where "dyspnea when walking" ("Do you become short of breath when walking on flat ground at a normal pace?") corresponds to the 2nd scale on the mMRC dyspnea scale. All the questions on our local dyspnea scale were similar to those of the mMRC dyspnea scale but were phrased as individual "yes or no questions".[20,21] Exacerbation history was based on two questions: "Have you ever taken cortisone tablets for breathing problems/asthma?" and "Have you previously taken it in periods when your illness had worsened?" Participants answering 'yes' to both questions and having ≥2 cortisone courses during the last year or ≥3 courses during the last two years were categorized as having a high exacerbation risk.

## Clinical Examination and Questionnaires

From clinical examinations and questionnaires, information on age (years), sex, body mass index (BMI, kg/m$^2$), smoking status, smoking pack-years, physical activity, education, diabetes ever, asthma ever, cardiovascular disease, systolic blood pressure (mmHg), and non-fasting total serum cholesterol (mmol/L) were recorded.

Age was recorded to one decimal place. Height and weight were measured with light clothing and without shoes. Height was rounded to the nearest centimeter and weight to the nearest half kilogram.[14,22] Cardiovascular disease included self-reported angina pectoris, myocardial infarction, and stroke. From three measurements of systolic blood pressure, the mean of the last two measurements was used.[22]

## Follow-Up and Outcomes

Follow-up for both events (first unplanned COPD hospitalization and all-cause mortality) began at the date of participation in HUNT2 and ended at an event date or at the end of follow-up, 31 December 2015. Participants were right-censored on emigration (n=4). When COPD hospitalization was the event, observations were right-censored at the date of death. There was no other loss to follow-up. COPD hospitalizations were identified from the international statistical classification of disease and related health problems (ICD) codes in medical records (Supplementary-Table S1).[23] Information on date of COPD hospitalizations and all-cause mortality was obtained from the Norwegian Cause of Death Registry and The Nord-Trøndelag Hospital Trust, respectively.

## Statistical Analysis

COPD hospitalization rates and all-cause mortality rates per 1000 person-years and 95% confidence intervals (CI) were calculated. Cumulative incidence curves for COPD hospitalization and all-cause mortality were constructed. We used the GOLD classifications as continuous measures to test for trend.

Cox proportional hazard models were used to calculate hazard ratios (HRs) and 95% CIs for the association of the GOLD classifications with COPD hospitalization or all-cause mortality. We presented crude HRs (Model 1) and adjusted HRs (Model 2). Model 2 accounts for age (as a continuous variable), sex (women, men), smoking [never, former (<10, 10–19, ≥20 pack-years), current (<10, 10–19, ≥20 pack-years), unknown], body mass index (<25.0, 25.0–29.9, ≥30.0, unknown), and education (<10, ≥10 years, unknown). In supplementary analysis (Model 3) we additionally adjusted for physical activity (no, light exercise, hard exercise, unknown), cardiovascular diseases (no, yes, unknown), asthma ever (no, yes, unknown), diabetes ever (no, yes, unknown), systolic blood pressure (sex-specific quartiles, unknown), and cholesterol (sex-specific quartiles, unknown).

Proportional hazards assumptions were assessed with log-log survival curves and Schoenfeld residuals tests.[24] Multicollinearity was tested where the variance inflation factor (VIF) was less than 1.2 in all models.[25,26] As a measure of goodness of fit, Akaike Information Criteria (AIC) for each model were estimated (Supplementary-Table S2).[27]

Incident/dynamic time-dependent area under the receiver operating characteristic curves (AUCs) were used to assess the discrimination ability of the GOLD classifications to predict COPD hospitalization and all-cause mortality.[28–30] We used crude models to compare AUCs, as the clinical decision does not explicitly take into account other factors.[4,31] We used 10,000 bootstrap iterations to calculate 95% CI for the AUCs.[32] A general bootstrap algorithm (gBA) was applied to compare the AUCs.[33]

Statistical analyses were performed using R 3.5.0 software (http://www.r-project.org) and Stata 15.1 software (StataCorp., College Station, Texas).

# Results

In this population-based COPD cohort with over 20.4 years of follow-up, 522 were hospitalized due to COPD and 896 people died. Using the GOLD 2007 classification, 31.9%, 54.2%, 12.5%, and 1.4% were classified as grades 1–4, respectively. The number of people in the highest two categories was slightly higher using the GOLD 2011 (16.8% in groups C or D) than the GOLD 2007 (13.9% in grades 3 or 4) but considerably lower using the GOLD 2017 (4.1% in groups C or D) (Table 1, Figure 1).

We observed increasing COPD hospitalization and all-cause mortality with worsening categories of the GOLD 2007, GOLD 2011, or GOLD 2017 classifications ($p_{trend}$<0.001), except for the GOLD 2017 classification and all-cause mortality ($p_{trend}$=0.411) (Table 1). Similar results were observed in unadjusted cumulative incidence curves (Supplementary-Figure S3) and for HRs in adjusted models (Table 2: Model 2; Supplementary-Table S3: Model 3).

In crude models, the AUCs (95% CI) for COPD hospitalization at 20 years of follow-up for the GOLD 2007, GOLD 2011, and GOLD 2017 classification were 63.1 (58.7–66.9), 60.9 (56.1–64.4), and 55.1 (54.0–58.1),

**Table 1** Characteristics of COPD Participants Aged ≥40 Years in the HUNT Study (N=1300)

| Characteristics | GOLD 2007 ¶ | | | | GOLD 2011 ¥ | | | | GOLD 2017 † | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Grade 1 (n=415) | Grade 2 (n=704) | Grade 3 (n=163) | Grade 4 (n=18) | Group A (n=771) | Group B (n=311) | Group C (n=98) | Group D (n=120) | Group A (n=847) | Group B (n=400) | Group C (n=22) | Group D (n=31) |
| Participants (%) | 31.9 | 54.2 | 12.5 | 1.4 | 59.3 | 23.9 | 7.6 | 9.2 | 65.1 | 30.8 | 1.7 | 2.4 |
| Women (%) | 35.4 | 38.2 | 36.2 | 38.9 | 34.0 | 43.4 | 33.7 | 43.3 | 33.5 | 42.0 | 50.0 | 61.3 |
| Age(years) (Mean±SD) | 64.9±11.3 | 65.4±10.4 | 68.1±9.6 | 65.2±10.5 | 65.0±10.6 | 65.8±11.2 | 68.3±9.2 | 66.6±9.8 | 65.3±10.5 | 66.2±11.0 | 69.4±7.7 | 64.0±8.7 |
| BMI (Mean±SD) | 26.8±4.0 | 26.7±4.6 | 24.7±4.0 | 23.2±2.6 | 26.3±4.0 | 27.8±5.2 | 24.5±3.4 | 25.4±4.4 | 26.2±4.0 | 27.1±5.1 | 25.1±2.1 | 27.5±4.3 |
| Smoking pack-years (Mean±SD) | 22.9±14.7 | 25.0±15.2 | 27.3±16.4 | 29.5±25.3 | 24.4±14.8 | 24.3±16.1 | 25.9±14.9 | 27.1±18.6 | 24.7±14.8 | 25.4±17.2 | 22.1±14.5 | 20.4±11.3 |
| Current smoker (%) | 44.1 | 51.6 | 44.6 | 22.2 | 49.9 | 48.4 | 45.3 | 35.7 | 50.1 | 46.2 | 23.8 | 27.6 |
| Physically inactive (%) | 9.6 | 11.3 | 9.9 | 20.0 | 7.6 | 17.4 | 12.3 | 12.5 | 8.1 | 15.9 | 7.1 | 19.1 |
| Education ≥10 years (%) | 39.5 | 35.4 | 29.1 | 33.3 | 40.4 | 28.9 | 28.4 | 31.0 | 39.4 | 29.4 | 26.3 | 30.4 |
| CVD (%) | 19.8 | 22.2 | 20.4 | 23.5 | 18.5 | 29.0 | 20.6 | 19.3 | 18.6 | 27.9 | 22.7 | 6.5 |
| Asthma ever (%) | 52.3 | 60.5 | 77.9 | 94.4 | 54.2 | 60.8 | 77.6 | 86.7 | 55.7 | 65.8 | 100.0 | 96.8 |
| Diabetes ever (%) | 5.3 | 5.3 | 4.3 | 5.6 | 4.5 | 7.2 | 4.1 | 5.0 | 4.5 | 6.6 | 5.6 | 6.5 |
| SBP (mean, SD) | 147±22.6 | 147.2±21.8 | 149.8±24.7 | 136.7±20.3 | 147.1±22.1 | 147.2±22.1 | 150.5±25.2 | 145.9±22.9 | 147.4±22.3 | 147.1±22.1 | 151.9±27.2 | 143.4±25.0 |
| Cholesterol (Mean±SD) | 6.3±1.2 | 6.3±1.2 | 6.3±1.2 | 6.3±1.0 | 6.3±1.2 | 6.4±1.3 | 6.3±1.3 | 6.3±1.2 | 6.3±1.2 | 6.4±1.3 | 7.0±1.2 | 6.3±1.2 |
| COPD hospitalization (at least one) (n) | 104 | 293 | 109 | 16 | 248 | 128 | 64 | 82 | 296 | 191 | 16 | 19 |
| COPD hospitalization rate (per | 18.6 | 39.2 | 111.1 | 364.4 | 26.7 | 37.7 | 100.6 | 106.5 | 30.3 | 49.6 | 99.5 | 61.0 |
| 1000 person-years) | (15.4–22.6) | (35.0–44.0) | (92.1–134.0) | (223.2–594.8) | (23.6–30.3) | (31.7–44.8) | (78.7–128.5) | (85.8–132.3) | (27.1–34.0) | (43.0–57.1) | (61.0–162.5) | (38.9–95.6) |
| All-cause mortality (n) | 234 | 499 | 146 | 17 | 503 | 208 | 86 | 99 | 571 | 289 | 18 | 18 |
| All-cause mortality rate | 38.5 | 56.0 | 101.5 | 97.5 | 48.0 | 52.2 | 93.3 | 81.3 | 51.3 | 60.7 | 68.0 | 41.1 |
| (per 1000 person-years) | (33.9–43.8) | (51.3–61.1) | (86.3–119.4) | (60.6–156.8) | (44.0–52.4) | (45.6–59.8) | (75.6–115.3) | (66.7–99.0) | (47.3–55.7) | (54.1–68.1) | (42.9–108.0) | (25.9–65.2) |

**Notes:** ¶Grade 1 – percent-predicted forced expiratory volume in first second (ppFEV$_1$) ≥80; grade 2 –80>ppFEV$_1$≥50; grade 3 –50>ppFEV$_1$≥30; grade 4 –ppFEV$_1$<30. ¥Group A – ppFEV$_1$≥50 and exacerbation history <2 and modified Medical Research Council dyspnoea scale (mMRC) <2; group B – ppFEV$_1$≥50 and exacerbation history <2 and mMRC ≥2; group C – ppFEV$_1$<50 or exacerbation history ≥2 and mMRC <2; group D – ppFEV$_1$<50 or exacerbation history ≥2 and mMRC ≥2. †group A – mMRC <2 & exacerbation history <2; group B – mMRC ≥2 & exacerbation history <2; group C – mMRC <2 & exacerbation history ≥2; group D – mMRC ≥2 & exacerbation history ≥2.

**Abbreviations:** HUNT, Nord-Trøndelag Health Study 1995–1997; SD, standard deviation; BMI, body mass index; CVD, cardiovascular disease; SBP, systolic blood pressure; GOLD, global initiative for chronic obstructive lung disease.

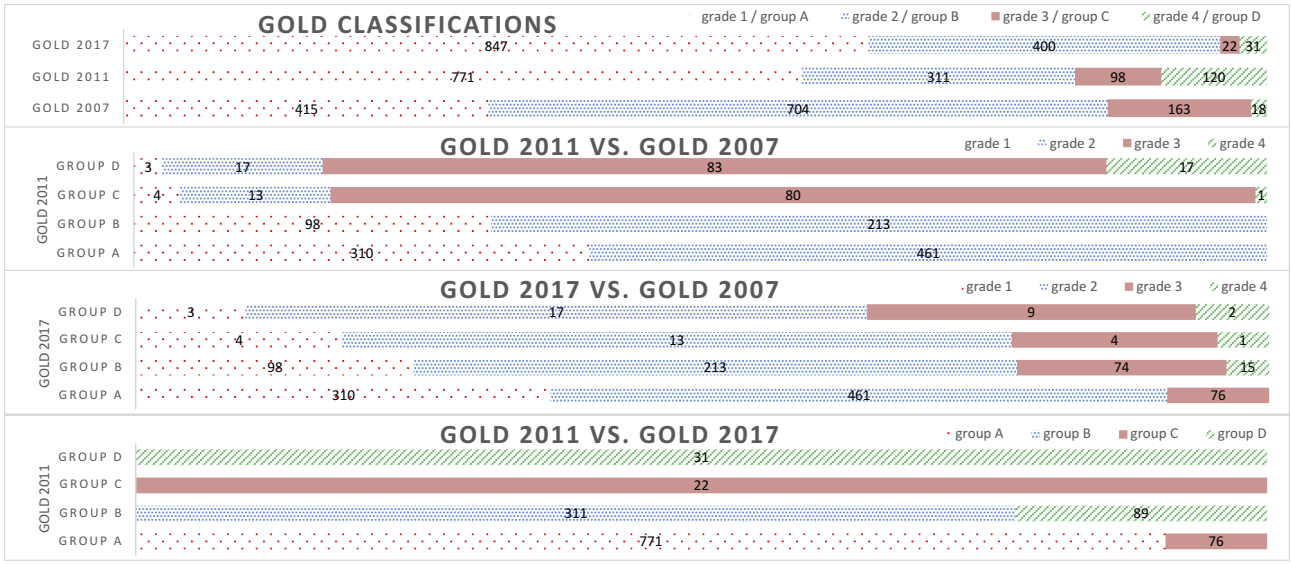


**Figure 1** Distribution of participants in different GOLD classifications.

respectively (Table 3; Figure 2). A similar pattern was observed over the follow-up time (Figure 2). For all-cause mortality, the corresponding estimates were 57.0 (54.8–59.1), 54.1 (52.1–56.0), and 52.6 (51.0–54.3) at 20 years of follow-up (Table 3; Figure 2). We observed similar results over follow-up time (Figure 2).

**Table 2** Hazard Ratios for GOLD 2007, GOLD 2011, and GOLD 2017 Among Participants with COPD Aged ≥40 Years in the HUNT2 Study (N=1300)

| Outcomes | GOLD Categories ¶, ¥, † | GOLD 2007 ¶ | | GOLD 2011 ¥ | | GOLD 2017 † | |
|---|---|---|---|---|---|---|---|
| | | HR (95% CI) * | HR (95% CI) # | HR (95% CI) * | HR (95% CI) # | HR (95% CI) * | HR (95% CI) # |
| COPD hospitalization | Grade 1/group A | Reference | Reference | Reference | Reference | Reference | Reference |
| | Grade 2/group B | 2.15 (1.72–2.69) | 2.05 (1.63–2.57) | 1.41 (1.14–1.75) | 1.42 (1.14–1.76) | 1.64 (1.36–1.96) | 1.64 (1.36–1.97) |
| | Grade 3/group C | 6.40 (4.87–8.41) | 5.11 (3.85–6.78) | 3.89 (2.95–5.14) | 3.21 (2.42–4.27) | 3.27 (1.97–5.41) | 3.17 (1.88–5.32) |
| | Grade 4/group D | 22.56 (13.19–38.57) | 17.08 (9.77–29.86) | 4.09 (3.18–5.26) | 3.75 (2.88–4.88) | 2.01 (1.26–3.20) | 2.15 (1.33–3.46) |
| All-cause mortality | Grade 1/group A | Reference | Reference | Reference | Reference | Reference | Reference |
| | Grade 2/group B | 1.51 (1.29–1.76) | 1.56 (1.33–1.82) | 1.10 (0.93–1.29) | 1.03 (0.87–1.21) | 1.21 (1.05–1.39) | 1.16 (1.01–1.34) |
| | Grade 3/group C | 3.13 (2.54–3.86) | 2.88 (2.32–3.58) | 2.21 (1.75–2.78) | 1.74 (1.38–2.20) | 1.39 (0.87–2.22) | 1.16 (0.72–1.86) |
| | Grade 4/group D | 2.95 (1.80–4.84) | 3.87 (2.34–6.39) | 1.84 (1.48–2.28) | 2.04 (1.63–2.55) | 0.79 (0.49–1.26) | 1.05 (0.65–1.69) |

**Notes:** *Model 1 (crude). #Model 2 - adjusted for age, sex, smoking, body mass index, education. ¶Grade 1 – percent-predicted forced expiratory volume in first second ($ppFEV_1$) ≥80; grade 2 –80>$ppFEV_1$≥50; grade 3 –50>$ppFEV_1$≥30; grade 4 –$ppFEV_1$<30. ¥Group A – $ppFEV_1$≥50 and exacerbation history <2 and modified Medical Research Council dyspnea scale (mMRC) <2; group B – $ppFEV_1$≥50 and exacerbation history <2 and mMRC ≥2; group C – $ppFEV_1$<50 or exacerbation history ≥2 and mMRC <2; group D – $ppFEV_1$<50 or exacerbation history ≥2 and mMRC ≥2. †Group A – mMRC <2 & exacerbation history <2; group B – mMRC ≥2 & exacerbation history <2; group C – mMRC <2 & exacerbation history ≥2; group D – mMRC ≥2 & exacerbation history ≥2.

**Abbreviations:** HUNT2, Nord-Trøndelag Health Study 1995–1997; GOLD, global initiative for chronic obstructive lung disease; HR, Hazard ratio; CI, confidence interval.

**Table 3** AUC for GOLD 2007, GOLD 2011, and GOLD 2017 at 20 Years of Follow-Up Among Participants with COPD Aged ≥40 Years in the HUNT2 Study (N=1300)

| Outcomes | GOLD 2007 ¶ | GOLD >2011 ¥ | GOLD 2017 † |
|---|---|---|---|
| | AUC (95% CI) * | AUC (95% CI) * | AUC (95% CI) * |
| COPD hospitalization | 63.1 (58.7–66.9) | 60.9 (56.1–64.4) | 56.1 (54.0–58.1) |
| All-cause mortality | 57.0 (54.8–59.1) | 54.1 (52.1–56.0) | 52.6 (51.0–54.3) |

**Notes:** *Model 1 (crude model) - the Cox model included GOLD 2007, GOLD 2011, or GOLD 2017. ¶Grade 1 – percent-predicted forced expiratory volume in first second ($ppFEV_1$) ≥80; grade 2 –80>$ppFEV_1$≥50; grade 3 –50>$ppFEV_1$≥30; grade 4 –$ppFEV_1$<30. ¥Group A – $ppFEV_1$≥50 and exacerbation history <2 and modified Medical Research Council dyspnea scale (mMRC) <2; group B – $ppFEV_1$≥50 and exacerbation history <2 and mMRC ≥2; group C – $ppFEV_1$<50 or exacerbation history ≥2 and mMRC <2; group D – $ppFEV_1$<50 or exacerbation history ≥2 and mMRC ≥2. †Group A – mMRC <2 & exacerbation history <2; group B – mMRC ≥2 & exacerbation history <2; group C – mMRC <2 & exacerbation history ≥2; group D – mMRC ≥2 & exacerbation history ≥2.

**Abbreviation:** HUNT2, Nord-Trøndelag Health Study 1995–1997; AUC, area under receiver operating characteristics curves.

## Discussion

To our knowledge, this is the first study to investigate the association between the GOLD 2017 classification and COPD hospitalization and to compare its discrimination ability to the previous GOLD classifications (GOLD 2007 and GOLD 2011). We found that the GOLD 2007 classification (based on lung function) was better than the GOLD 2011 classification (based on lung function, symptom burden, and exacerbation history) and the GOLD 2017 classification (based on symptom burden and exacerbation history), where the GOLD 2017 classification was the worst at predicting COPD hospitalization and all-cause mortality.

We found that group D of the GOLD 2017 classification had lower hazards for COPD hospitalization and all-cause mortality than group C. However, the estimates were imprecise for these two groups due to low numbers and the confidence intervals for the HRs overlapped considerably. The GOLD 2017 classification might however misclassify the severity of some COPD cases because it accounts only for symptom burden and exacerbation history and does not include airflow limitations. This leads to 9% of group A and 10% of group D of the GOLD 2017 classification having severe (GOLD grade 3) and mild COPD (GOLD grade 1), respectively. Similar results to our study have been observed by other studies.[9,10] The PLATINO study showed no clear pattern for all-cause mortality across the categories of the GOLD 2017 classification.[34] No previous studies had investigated COPD hospitalization in relation to the GOLD 2017 classification. However, the COPDGene Cohort study reported the annual rate of exacerbation and found no clear pattern across the categories of the GOLD 2017 classification.[11]

Other studies have compared the GOLD 2017 with the previous GOLD classifications to predict COPD exacerbation. In contrast to our study, three studies[11,12,35] observed that the GOLD 2017 and the GOLD 2011 classifications predicted exacerbation similarly well. The discrepancies might be because these studies did not take exacerbations as time-to-event data when comparing the discrimination ability. Furthermore, we observed COPD hospitalization where COPD hospitalization and exacerbation are related;

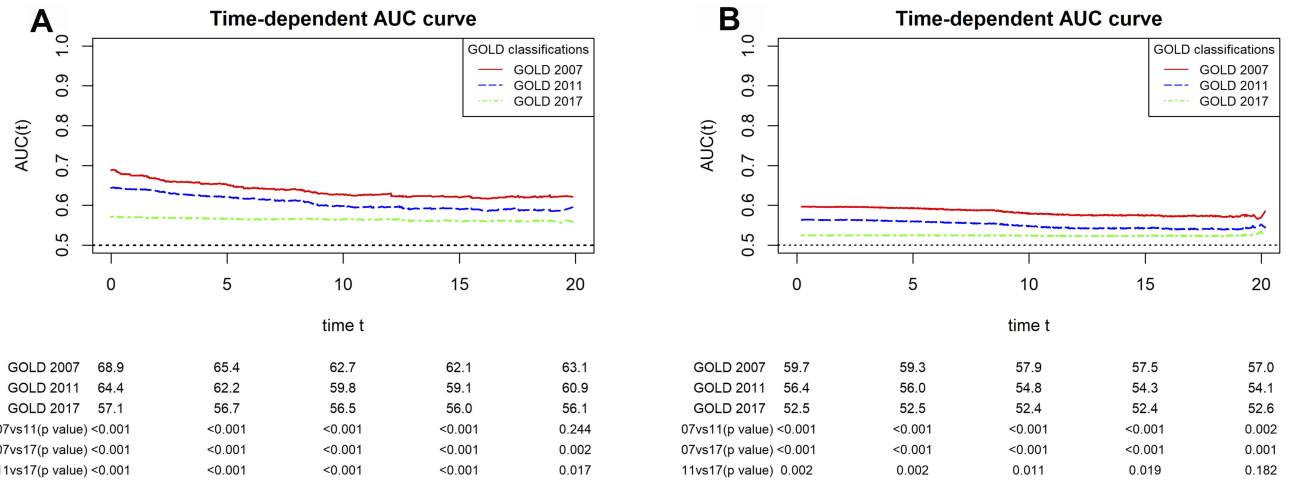


**Figure 2** AUC curves for the GOLD 2007, GOLD 2011, and GOLD 2017 classifications for (**A**) COPD hospitalization and (**B**) all-cause mortality over follow-up time (years) among participants with COPD aged ≥40 years in the HUNT2 study (N=1300).

however, not all exacerbations lead to hospitalizations. None of these studies included the GOLD 2007 classification. Similar to our study, a study by Cabrera Lopez et al[10] observed that all-cause mortality was better predicted by the GOLD 2011 than by the GOLD 2017 classification. In contrast to our study, Gedebjerg et al[9] found that the three GOLD classifications did not differ significantly in predicting respiratory and all-cause mortality. When they further divided the GOLD 2017 classification into a 16 sub-group classification by severity of airflow limitation, they found that the 16 sub-groups classification predicted mortality better than the GOLD classifications.[9] We attempted to repeat this analysis but we had sparse data for this classification with no observations in some sub-groups and highly imprecise estimates.

The GOLD classifications are recommended to guide the treatment approach for individual patients in clinics.[1,4] The GOLD 2007 classification used airflow limitation but due to disease complexity that might not be fully explained by lung function in individual patients, the classification was upgraded in 2011 to include clinical parameters such as symptom burden and exacerbation history to classify patients into ABCD groups.[1,5] Furthermore, in 2017, the GOLD separates airflow limitation from ABCD groups.[1] In our study, we observed that from group C and D of GOLD 2011 classification 78% of group C were moved to group A of the GOLD 2017 classification and 74% of group D participants were moved to group B of the GOLD 2017 classification. In addition, we observed that using the GOLD 2017 classification, 9% of group A participants had severe COPD (GOLD grade 3) and 22% of group B participants had severe or very severe COPD (GOLD grades 3 or 4). Although the recently updated GOLD 2019 report[36] recommended use of the GOLD 2017 classification for prescribing an initial pharmacological treatment, the use of the GOLD 2017 classification to select a treatment approach should be cautioned for those in groups A or B. Further studies are required to understand the treatment effects related to prognosis of outcomes using the GOLD 2017 classification. In our study, the GOLD 2007 classification was better than the GOLD 2011 classification, followed by the GOLD 2017 classification to predict COPD hospitalization and all-cause mortality. Despite this observation, all classifications were generally still poor prognostic markers for COPD hospitalization and all-cause mortality.[37] The best prognostic marker might not necessarily be the best guide to pharmaceutical treatment because two individuals might have the same risk of mortality for different reasons, which would indicate different treatment strategies, therefore, a new classification based on other or additional criteria should be suggested to assess prognosis.

This study had several strengths. It is the first study to investigate the discrimination ability of all three GOLD classifications over a 20-year period to predict COPD hospitalization. The study is based on the HUNT2 Lung Study which had a reasonably high level of participation (76.0% of people invited to the HUNT2 Lung Study), limiting the potential for selection bias. We had complete information on mortality and there was no loss to follow-up other than very few emigrations (4 out of 1300 participants). To reduce measurement error, quality assurance of spirometry curves was performed.[16,18]

This study also had certain limitations. We had information on COPD hospitalizations only from the hospitals of study area (northern Trøndelag) and we lacked data from other hospitals of Norway. We lacked information on COPD hospitalizations prior to baseline; however, we used exacerbation history to classify groups in the GOLD 2011 and GOLD 2017 classification. The mMRC dyspnea questionnaire should have been used as a scale but we used them as individual questions when the survey was conducted.[20] However, a study has found close agreement between our dyspnea questions and mMRC dyspnea scale.[20,21] We do not have baseline information on treatment used for the people with COPD during 1995–1997. Due to this, there may be bias in the association of GOLD classifications and outcomes (COPD hospitalization and all-cause mortality). However, we do not expect this to have substantially influenced the comparison of discrimination abilities between GOLD classifications. There was missing information on some covariates, therefore to avoid sample loss in adjusted models, a missing indicator variable (missing information as unknown category) was used which might bias the association between the GOLD classifications and COPD hospitalization or all-cause mortality.

## Conclusion

We found that the GOLD 2007 classification was better than the GOLD 2011 and GOLD 2017 classifications at predicting COPD hospitalization and all-cause mortality.

## Acknowledgments

The Nord-Trøndelag Health Study (HUNT) is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Science, Norwegian University of Science and Technology NTNU), Nord-Trøndelag County Council and the Norwegian Institute of Public Health. The HUNT2

Lung Study was partly funded through a non-demanding grant from AstraZeneca Norway. We thank Nord-Trøndelag Hospital Trust for providing patient records.

## Author Contributions

All authors contributed to data analysis, drafting or revising the article, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

## Funding

This study was funded by ExtraStiftelsen Helse og Rehabilitering and Landsforeningen for hjerte-og-lungesyke (the Norwegian Extra Foundation for Health and Rehabilitation and the Norwegian Heart and Lung Patient Organization) (project number 2016/FO79031) and the liaison committee of the Central Norway Regional Health Authority – NTNU (Norwegian University of Science and Technology). Ben Brumpton works in a research unit funded by Stiftelsen Kristian Gerhard Jebsen; Faculty of Medicine and Health Sciences, NTNU; The Liaison Committee for education, research and innovation in Central Norway; the Joint Research Committee between St. Olavs Hospital and the Faculty of Medicine and Health Sciences, NTNU. Ben Brumpton and David Carslake work at Medical Research Council Integrative Epidemiology Unit at the University of Bristol which is supported by the Medical Research Council (MC_UU_00011/1) and the University of Bristol.

## Disclosure

LB reports personal fees for presenting the abstract of this study at Nordic Respiratory Science Forum (NRSF) organized by AstraZeneca in 2019. The authors report no other conflicts of interest in this work.

## References

1. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary diesease. 2017. Available from: http://goldcopd.org/. Accessed 01 August 2017.
2. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary diesease. 2011. Available from: http://goldcopd.org/. Accessed 01 August 2016.
3. Vestbo J, Hurd SS, Rodriguez-Roisin R. The 2011 revision of the global strategy for the diagnosis, management and prevention of COPD (GOLD) – why and what? *Clin Respir J*. 2012;6(4):208–214. doi:10.1111/crj.12002
4. Soriano JB, Lamprecht B, Ramirez 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(6):443–450. doi:10.1016/S2213-2600(15)00157-5
5. 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(1):51–59. doi:10.1164/rccm.201212-2276OC
6. Lange P, Marott 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. doi:10.1164/rccm.201207-1299OC
7. Chen CZ, Ou CY, Hsu CH, Hsiue TR. Validation of the GOLD 2013 classification in predicting exacerbations and mortality in Taiwanese patients with chronic obstructive pulmonary disease. *J Formosan Med Assoc*. 2015;114(12):1258–1266. doi:10.1016/j.jfma.2014.12.003
8. Leivseth L, Brumpton BM, Nilsen TIL, Mai X-M, Johnsen R, Langhammer A. GOLD classifications and mortality in chronic obstructive pulmonary disease: the HUNT Study, Norway. *Thorax*. 2013;68(10):914. doi:10.1136/thoraxjnl-2013-203270
9. Gedebjerg A, Szepligeti SK, Wackerhausen LH, 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;6(3):204–212. doi:10.1016/S2213-2600(18)30002-X
10. Cabrera Lopez C, Casanova Macario C, Marin Trigo JM, et al. Comparison of the 2017 and 2015 global initiative for Chronic Obstructive Lung Disease reports. Impact on grouping and outcomes. *Am J Respir Crit Care Med*. 2018;197(4):463–469. doi:10.1164/rccm.201707-1363OC
11. Criner RN, Labaki WW, Regan EA; for the COPD Gene Investigators, et al. 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;6(1):64–73. doi:10.15326/jcopdf.6.1.2018.0130.
12. Han M-Z, Hsiue T-R, Tsai S-H, Huang T-H, Liao X-M, Chen C-Z. Validation of the GOLD 2017 and new 16 subgroups (1A-4D) classifications in predicting exacerbation and mortality in COPD patients. *Int J Chron Obstruct Pulmon Dis*. 2018;13:3425–3433. doi:10.2147/COPD.S179048
13. Krokstad S, Langhammer A, Hveem K, et al. Cohort profile: the HUNT study, Norway. *Int J Epidemiol*. 2013;42(4):968–977. doi:10.1093/ije/dys095
14. Bhatta L, Leivseth L, Mai X-M, et al. Prevalence and trend of COPD from 1995–1997 to 2006–2008: the HUNT study, Norway. *Respir Med*. 2018;138:50–56. doi:10.1016/j.rmed.2018.03.020
15. 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 tests, european community for steel and coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl*. 1993;16:5–40. doi:10.1183/09041950.005s1693
16. Langhammer A, Johannessen A, Holmen TL, et al. Global lung function initiative 2012 reference equations for spirometry in the Norwegian population. *Eur Respir J*. 2016;48(6):1602–1611. doi:10.1183/13993003.00443-2016
17. Standardization of Spirometry, 1994 Update. American Thoracic Society. *Am J Respir Crit Care Med*. 1995;152(3):1107–1136. doi:10.1164/ajrccm.152.3.7663792
18. Hankinson JL, Eschenbacher B, Townsend M, Stocks J, Quanjer PH. Use of forced vital capacity and forced expiratory volume in 1 second quality criteria for determining a valid test. *Eur Respir J*. 2015;45(5):1283–1292. doi:10.1183/09031936.00116814
19. Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3–95 year age range: the global lung function 2012 equations: report of the Global Lung Function Initiative (GLI), ERS Task Force to establish improved lung function reference values. *Eur Respir J*. 2012;40(6):1324–1343. doi:10.1183/09031936.00080312

20. Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest*. 1988;93(3):580–586. doi:10.1378/chest.93.3.580
21. Brogger JC, Bakke PS, Gulsvik A. Comparison of respiratory symptoms questionnaires. *Int J Tubercul Lung Dis*. 2000;4(1):83–90.
22. Holmen J, Midthjell K, Krüger Ø, et al. The Nord-Trøndelag Health Study 1995-97 (HUNT 2): objectives, contents, methods and participation. *Norsk Epidemiol*. 2003;13(1):19–32.
23. Northern Norway Regional Health Authority: Centre for Clinical Documentation and Evaluation (SKDE). COPD healthcare atlas. 2018. Available from: https://helseatlas.no/sites/default/files/copd-healthcare-atlas.pdf. Accessed 10 March 2019.
24. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994;81(3):515–526. doi:10.1093/biomet/81.3.515
25. Belsley DA, Kuh E, Welsch RE. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York (N.Y.): Wiley; 1980.
26. O'Brien RM. A caution regarding rules of thumb for variance inflation factors. *Qual Quant*. 2007;41(5):673–690. doi:10.1007/s11135-006-9018-6
27. Bozdogan H. Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions. *Psychometrika*. 1987;52(3):345–370. doi:10.1007/BF02294361
28. Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med Res Methodol*. 2017;17(1):53. doi:10.1186/s12874-017-0332-6
29. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61(1):92–105. doi:10.1111/biom.2005.61.issue-1
30. Saha P, Heagerty PJ. Time-dependent predictive accuracy in the presence of competing risks. *Biometrics*. 2010;66(4):999–1011. doi:10.1111/j.1541-0420.2009.01375.x
31. Miller MR, Pedersen OF. New concepts for expressing forced expiratory volume in 1 s arising from survival analysis. *Eur Respir J*. 2010;35(4):873–882. doi:10.1183/09031936.00025809
32. Bansal A, Heagerty PJ. A tutorial on evaluating the time-varying discrimination accuracy of survival models used in dynamic decision making. *Med Decis Making*. 2018;38(8):904–916. doi:10.1177/0272989X18801312
33. Martínez-Camblor P, Corral N. A general bootstrap algorithm for hypothesis testing. *J Stat Plan Inference*. 2012;142(2):589–600. doi:10.1016/j.jspi.2011.09.003
34. Menezes AM, Wehrmeister FC, Perez-Padilla R, et al. The PLATINO study: description of the distribution, stability, and mortality according to the global initiative for Chronic Obstructive Lung Disease classification from 2007 to 2017. *Int J Chron Obstruct Pulmon Dis*. 2017;12:1491–1501. doi:10.2147/COPD
35. Marcoa R, Rodrigues DM, Dias M, et al. Classification of Chronic Obstructive Pulmonary Disease (COPD) according to the new Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2017: comparison with GOLD 2011. *Copd*. 2018;15(1):21–26. doi:10.1080/15412555.2017.1394285
36. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary diesease. 2019. Available from: http://goldcopd.org/. Accessed 02 February 2019.
37. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thor Oncol*. 2010;5(9):1315–1316. doi:10.1097/JTO.0b013e3181ec173d