

RESEARCH PAPER

Are GOLD ABCD groups better associated with health status and costs than GOLD 1234 grades? A cross-sectional study

*Melinde RS Boland^{1,2}, Apostolos Tsiachristas^{1,2}, Annemarije L Kruis³, Niels H Chavannes³, Maureen PMH Rutten-van Mölken^{1,2}

¹ Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands

² Department of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

³ Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands

Originally received 12th July 2013; resubmitted 31st October 2013; revised 25th November 2013; further revision 30th November 2013; accepted 30th November 2013; online 21st January 2014

Abstract

Aims: To investigate the association of the GOLD ABCD groups classification with costs and health-related quality of life (HR-QoL) and to compare this with the GOLD 1234 grades classification that was primarily based on lung function only.

Methods: In a cross-sectional study, we selected patients diagnosed with chronic obstructive pulmonary disease (COPD) from electronic medical records of general practices. Multi-level analysis was used with costs (medication, primary care, healthcare, societal), disease-specific and generic HR-QoL as independent variables. Either the new or the old GOLD stages were included in the analysis together with several covariates (age, gender, living situation, co-morbidity, self-efficacy, smoking, education, employment).

Results: 611 patients from 28 general practices were categorised as GOLD-A (n=333), GOLD-B (n=110), GOLD-C (n=80) and GOLD-D (n=88). Patients in the GOLD-B and GOLD-D groups had the highest prevalence of co-morbidities and the lowest level of physical activity, self-efficacy, and employment. The models with GOLD ABCD groups were more strongly related to and explained more variance in costs and in disease-specific and generic HR-QoL than the models with GOLD 1234 grades. The mean Clinical COPD Questionnaire score worsened significantly, with scores 1.04 (GOLD-B), 0.4 (GOLD-C) and 1.21 (GOLD-D) worse than for patients in GOLD-A. Healthcare costs per patient were significantly higher in GOLD-B (72%), GOLD-C (74%) and GOLD-D (131%) patients than in GOLD-A patients.

Conclusions: The GOLD ABCD groups classification is more closely associated with costs and HR-QoL than the GOLD 1234 grades classification. Furthermore, patients with GOLD-C had a better HR-QoL than those with GOLD-B but the costs of the two groups did not differ.

© 2014 Primary Care Respiratory Society UK. All rights reserved.

MRS Boland *et al.* *Prim Care Respir J* 2014; 23(1): 30-37

<http://dx.doi.org/10.4104/pcrj.2014.00002>

Keywords chronic obstructive pulmonary disease, economics, primary care

See linked editorial by Soriano and Román-Rodríguez on pg 5

The full version of this paper, with online appendices, is available online at www.thepcrj.org

Introduction

The assessment, monitoring, and treatment of chronic obstructive pulmonary disease (COPD) have long been driven by the level of airflow limitation. However, emphasis has recently shifted towards a better understanding of the heterogeneity among patients with

COPD. This led to the revision of the strategy document of the Global Initiative for Chronic Obstructive Lung Disease (GOLD) in 2011.¹ In the 2011 GOLD strategy, COPD is not only classified with the spirometric 1234 grades but also with the ABCD groups.¹ According to the ABCD groups, the assessment of COPD is based on symptoms and the risk of experiencing an exacerbation in addition to lung function.¹ Low or high symptom burden can be determined using the modified Medical Research Council (mMRC) dyspnoea scale² (mMRC <2 versus mMRC ≥2) or the COPD Assessment Test (CAT) (CAT <10 versus CAT ≥10).³ A low or high risk is based on the history of exacerbations (<2 or ≥2 exacerbations in the previous 12

* **Corresponding author:** Mrs Melinde Boland, Erasmus University Rotterdam - Health Policy and Management, Institute for Medical Technology Assessment, Burgemeester Oudlaan 50, Rotterdam 3015 EJ, The Netherlands. Tel: 0031-104089778 E-mail: boland@bmg.eur.nl

months) or the forced expiratory volume in one second (FEV₁) as a percentage of the predicted value ($\geq 50\%$ or $< 50\%$), whichever results in a higher risk. Based on this information, the patient is placed into one of the four ABCD groups. Ultimately, the GOLD classification should estimate current health status, predict future health outcomes, and guide therapy.¹

Several studies have investigated the association between the GOLD ABCD groups and health outcomes.⁴⁻⁷ Lange *et al.*⁴ investigated the ability of the GOLD ABCD groups to predict the risk of exacerbations and mortality in two Danish general populations combined. However, the FEV₁% predicted was based on pre-bronchodilator values, which can overestimate the prevalence of severe COPD. Han *et al.*⁵ evaluated the influence of symptom instrument choice on patient category assignment and prospective exacerbation risk by GOLD group in the COPDGene cohort with patients recruited in hospitals. Soriano *et al.*⁶ compared the distribution and prognostic validity of the GOLD ABCD groups with the GOLD 1234 grades as predictors of mortality in 11 small Spanish cohorts combined. Agusti *et al.*⁷ used the ECLIPSE data to investigate the ability of the GOLD ABCD groups to predict mortality, exacerbations, and hospitalisations. Furthermore, two studies compared the ability of the GOLD 1234 grades and GOLD ABCD groups to predict mortality^{8,9} and hospitalisations.⁸ However, none of these studies has used such a wide variation of different disease-specific and generic health-related quality of life (HR-QoL) measurements as we did in our study. Moreover, none has

investigated the association with costs. The latter is important information for decision-analytical models that estimate the cost-effectiveness of COPD treatments in different subgroups.

The aim of our study was to investigate the association of the GOLD ABCD groups classification with a wide range of HR-QoL outcomes and cost categories. We compared this with the GOLD 1234 grades classification, which is based on lung function only. The study was performed in patients with COPD recruited in primary care, where the vast majority of patients with COPD are treated.

Methods

Setting

Patients diagnosed with COPD were selected from electronic medical records of general practices (EMR-GP) in the Western part of the Netherlands between September 2010 and September 2011. Participants had a post-bronchodilator FEV₁/forced vital capacity (FVC) ratio of < 0.7 . Patients with terminal illnesses, dementia, cognitive impairment, hard drug abuse, alcohol abuse, or those who were unable to complete Dutch questionnaires were excluded. The general practitioners verified that the included patients fulfilled the inclusion and exclusion criteria. All participants provided written informed consent before participation and the study was approved by the Medical Ethics Committee of the Leiden University Centre.

Data

Participants completed the following questionnaires with the support of research nurses: Clinical COPD Questionnaire (CCQ),^{10,11} St George's

Table 1. Input and outcome variables used in the analysis

Input variable		
GOLD	1234 grades ABCD groups	GOLD-1, GOLD-2, GOLD-3, GOLD-4. Reference group is GOLD-1 GOLD-A, GOLD-B, GOLD-C, GOLD-D. Reference group is GOLD-A
Patient characteristics	Co-morbidity Age Gender Low education Single Employment Self-efficacy Smoking	Charlson co-morbidity index ¹⁸ Age of the participants in years at the time of the interview 0 female; 1 male 0 high education; 1 low education (defined as no or only primary education) 0 not single (married/living together); 1 single (never married, divorced, widow(er)) 0 not employed; 1 employed Component of the Self-Management Ability Scale from 0 (worst) to 100 (best) Never smoker; former smoker and smoker. Reference group is never smoker
Outcome variable		
Disease specific HR-QoL	CCQ	0 (best) to 6 (worst), with 3 sub scores: Symptoms, Functional state and Mental state. The minimum clinically important difference is -0.4 points ^{10,11}
	SGRQ	Scale from 0 (best) to 100 (worst), with 3 sub scores: symptom, activity, and impact. The minimum clinically important difference is 4 points ¹¹
Generic HR-QoL	EQ5D-VAS	Scale from 0 (worst imaginable health state) to 100 (best imaginable health state) ¹⁶
	SF-36 physical	Scale from 0 (worst) to 100 (best) ^{12,14}
	SF-36 mental	Scale from 0 (worst) to 100 (best) ^{12,14}
Cost	Medication	Average medication costs in 3 months (including non-COPD medication)
	Primary care Healthcare costs	Costs of GP, dietician, physiotherapist, podiatrist, occupational therapist, home care Costs of medication, primary care, hospital days, intensive care days, specialist visits, ED visits and pulmonary rehabilitation
	Costs from societal perspective	Healthcare, travel and productivity costs

*HR-QoL=health-related quality of life, CCQ=Clinical COPD Questionnaire, SGRQ=St George's Respiratory Questionnaire, EQ5D-VAS=EuroQoL 5 dimensions-visual analogue scale

Respiratory Questionnaire (SGRQ),¹² MRC Dyspnoea Scale,² Short Form-36 (SF-36),^{13,14} EuroQoL-5 dimensions (EQ5D) including both the descriptive part and the visual analogue scale (EQ5D-VAS),^{15,16} and a questionnaire asking about healthcare utilisation, travel expenses, and absence from paid work in the three months prior to completion of the questionnaire. Information on drug prescriptions was extracted from the EMR-GP. EQ5D utilities were estimated using the Dutch value set.¹⁷ We collected patient characteristics such as age, gender, co-morbidity,¹⁸ education level, marital status, employment status, self-efficacy, and smoking status. Further details are shown in Table 1.

Participants were classified according to both 1234 grades and ABCD groups.¹ Low symptom burden was defined as mMRC <2 and high symptom burden as mMRC ≥2. Spirometry data were obtained from the EMR-GP. If spirometry data were unavailable, a respiratory nurse contacted the individual participants for a spirometry test. The number of exacerbations in the previous 12 months was determined according to International Classification of Primary Care (ICPC) registrations in the EMR-GP. Exacerbations were defined as prescriptions for: (1) prednisone with or without antibiotic, unless the ICPC explicitly indicated a reason other than COPD exacerbation; or (2) antibiotics alone with an ICPC description of a COPD exacerbation. A gap of at least 21 days between prednisone prescriptions was used to distinguish subsequent exacerbations and to define different events.

Patient subgroups were created within GOLD-C and GOLD-D (i.e. those at high risk due to low lung function (GOLD-C₁ and D₁)), a history of frequent exacerbations (GOLD-C₂ and D₂), or both (GOLD-C₃ and D₃).

We defined four different categories of costs (medication costs, costs in the primary care sector, total healthcare costs, and total costs from a societal perspective). Medication costs included COPD- and non-COPD-related prescriptions. Primary care costs included costs of consulting with a GP, nurse, dietician, physiotherapist, podiatrist, or occupational therapist, and costs of home care provided by home care organisations. Total healthcare costs included medication costs, costs of primary care plus hospital admissions, consultations with medical specialists, emergency department visits, and pulmonary rehabilitation costs. In addition to the total healthcare costs, the costs from a societal perspective included travel costs of patients to healthcare providers and costs of productivity losses.

Standard unit costs were obtained from the Dutch manual for costing research¹⁹ inflated to 2012 values using the general consumer price index.²⁰ The costs of medications were obtained from the GIP-Databank, including taxes and pharmacist dispensing fees.²¹ The productivity costs were estimated using the friction cost approach, which assumes that productivity loss only occurs during the time it takes to replace a sick employee (the friction period).²² We used a friction period of 115 days.¹⁹

Statistical analyses

Descriptive statistics of patient characteristics, clinical measures, HR-QoL, and costs were provided. Continuous variables were summarised as means and standard deviation and categorical variables as frequencies. We used ANOVAs for normally distributed variables, Kruskal-Wallis tests for non-normally distributed

continuous variables, and Pearson's χ^2 tests for categorical variables to test for differences between the four 1234 grades and ABCD groups; p values <0.05 were considered statistically significant.

We modelled disease-specific and generic HR-QoL as functions of the 1234 grades and the ABCD groups controlling for patient characteristics using linear mixed-effects models. We accounted for the hierarchical nature of the data in these multi-level models with level 1 patient and level 2 general practices. This model was specified for six different outcome variables: CCQ, SGRQ, EQ5D, EQ5D-VAS, SF-36 mental component, and SF-36 physical component. The inclusion of a covariate was based on the likelihood ratio test that is suitable to compare the goodness-of-fit of nested models.²³ The following covariates were considered for inclusion: Charlson co-morbidity index, age, gender, low education, employment status, self-efficacy, and smoking status.

We also modelled costs as a function of the 1234 grades and the ABCD groups. We used a generalised linear mixed model suitable for analysing the right-skewed cost data. The distribution and link function were selected after comparing the goodness-of-fit of models with different specifications of the distribution and link functions. Models that had the lowest Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) were selected. Four cost categories were used as outcome variables: healthcare costs, medication costs, primary care costs, and total costs from a societal perspective.

Figure 1. Flowchart showing inclusion of patients in the study

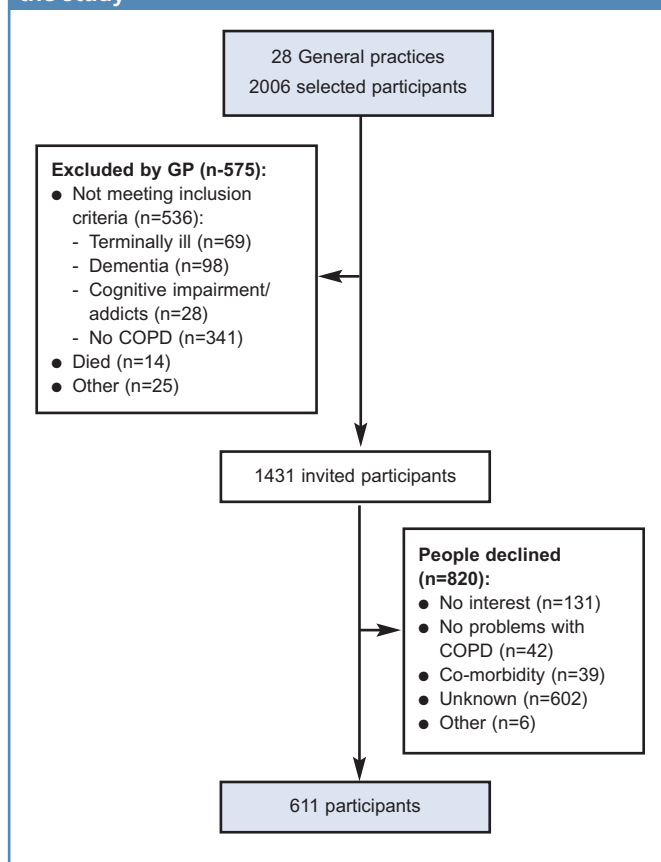


Table 2. Patient sample distributed according to the GOLD 1234 grades and ABCD groups classification

	GOLD-1	GOLD-2	GOLD-3	GOLD-4	Total
GOLD-A	105	228	0	0	333 (55%)
GOLD-B	30	80	0	0	110 (18%)
GOLD-C	10	14	51	5	80 (13%)
GOLD-C ₁ (lung function)	0	0	38	5	43
GOLD-C ₂ (exacerbation)	10	14	0	0	24
GOLD-C ₃ (lung function + exacerbation)	0	0	13	0	13
GOLD-D	4	13	59	12	88 (14%)
GOLD-D ₁ (lung function)	0	0	49	8	57
GOLD-D ₂ (exacerbation)	4	13	0	0	17
GOLD-D ₃ (lung function + exacerbation)	0	0	10	4	14
Total	149 (24%)	335 (55%)	110 (18%)	17 (3%)	611 (100%)

In all models we used an unstructured covariance matrix since this does not impose a particular pattern on the covariance. To correct for incorrect specifications of the covariance matrix we used sandwich estimators. The AIC compared the goodness-of-fit of the model that included the ABCD groups with the model that included the 1234 grades. Furthermore, the models were compared with respect to the proportion of explained variance measured at the

patient level (R_1^2).²⁴ Finally, we tested whether the four ABCD groups statistically differed from each other in the four cost models and the six HR-QoL models.

Results

Descriptive statistics

Figure 1 shows the study flowchart. From the 2,006 potential

Table 3. Sample characteristics according to GOLD ABCD groups classification

		GOLD-A (N=333)	GOLD-B (N=110)	GOLD-C (N=80)	GOLD-D (N=88)	Average (N=611)	p Value
Disease	FEV ₁	73.2 (14.6)	72.1 (16.4)	50.3 (19.9)	43.6 (16.4)	65.7 (19.8)	
	FER	58.8 (8.3)	58.2 (9.1)	47.1 (12.8)	44.2 (13.3)	55.1 (11.6)	
	Exacerbation rate	0.14 (0.34)	0.25 (0.43)	1.51 (1.71)	1.56 (2.02)	0.54 (1.20)	
	MRC dyspnoea	1.26 (0.67)	3.45 (0.67)	1.40 (0.54)	3.64 (0.68)	2.01 (1.24)	
	Charlson Index	2.14 (1.11)	2.61 (1.34)	2.08 (1.11)	2.52 (1.35)	2.27 (1.21)	0.002
	Severe cardiac problems, %	10%	25%	18%	18%	15%	0.002
	Diabetes, %	12%	19%	12%	21%	15%	0.091
Patient	Depression, %	7%	14%	12%	14%	10%	0.067
	Age	67 (11)	72 (11)	69 (10)	69 (10)	68 (11)	0.001
	Men, %	58%	48%	60%	55%	56%	0.279
	Low education, %	34%	56%	51%	49%	43%	<0.001
	Single, %	33%	51%	33%	36%	37%	0.018
	Employment, %	36%	15%	29%	13%	28%	<0.001
	Self-efficacy	68.5 (16.4)	61.1 (17.1)	69.8 (16.0)	57.8 (17.3)	65.7 (17.1)	<0.001
Health behaviour	Smoker, %	35%	34%	45%	27%	35%	0.166
	Former smoker, %	56%	53%	50%	66%	56%	
	Non-smokers, %	8%	13%	5%	7%	8%	
	MET-min	3876 (5331)	1211 (1717)	3112 (4906)	1304 (2564)	2923 (4630)	<0.001
	High/moderate physical activity, %	15%	4%	14%	3%	11%	0.001
HR-QoL	Low physical activity, %	85%	96%	86%	97%	89%	
	CCQ	1.04 (0.70)	2.20 (0.99)	1.41 (0.73)	2.42 (0.91)	1.50 (0.98)	<0.001
	SGRQ	25.5 (14.6)	51.8 (17.4)	33.9 (16.6)	56.3 (16.8)	35.8 (20.3)	<0.001
	EQ5D	0.81 (0.22)	0.62 (0.30)	0.76 (0.24)	0.64 (0.28)	0.74 (0.26)	<0.001
	EQ5D-VAS	73.6 (14.0)	58.1 (17.0)	68.3 (14.4)	54.9 (17.7)	67.4 (17.0)	<0.001
	SF36 Physical	42.7 (9.4)	32.6 (10.0)	40.0 (8.9)	28.9 (8.6)	38.5 (10.8)	<0.001
	SF36 Mental	50.8 (9.2)	45.9 (10.6)	48.9 (10.4)	45.3 (11.1)	48.9 (10.1)	<0.001
Cost	Medication	219 (209)	336 (365)	427 (691)	361 (253)	288 (353)	<0.001
	Primary care	173 (232)	289 (299)	213 (298)	425 (457)	236 (307)	<0.001
	Healthcare costs	653 (1375)	1124 (1737)	1143 (1497)	1877 (2778)	978 (1772)	<0.001
	Costs from societal perspective	917 (1857)	1164 (1749)	1640 (3022)	2195 (4674)	1240 (2633)	<0.001

*Values are mean (SD) unless stated otherwise.

FEV₁=forced expiratory volume in 1 second, FER=forced expiratory ratio, MRC=Medical Research Council Dyspnoea scale, MET-min=metabolic equivalent time, HR-QoL=health-related quality of life, CCQ=Clinical COPD Questionnaire, SGRQ-C=St George's Respiratory Questionnaire, EQ5D-VAS=visual analogue scale

participants, 575 (29%) were excluded by their GP, most of whom (59%) were misdiagnosed. From the remaining 1,431 potential participants, 57% refused to participate in the study. Most of these indicated no reason for refusing (73%), while others expressed no interest (16%) or reported not having troublesome COPD symptoms (5%). In total, we included 611 participants diagnosed with COPD from 28 general practices; 55% were classified in GOLD-A, 18% in GOLD-B, 13% in GOLD-C, and 14% in GOLD-D (Table 2). Patients with low exacerbation risks (GOLD-A and GOLD-B) were mainly former GOLD-2 patients (70%). Patients with high exacerbation risks (GOLD-C and GOLD-D) included those from all former GOLD classes, but were mainly from GOLD-3 (65%). Table 2 shows that the major criterion for classifying patients as high-risk was low lung function.

The sample characteristics, according to GOLD group, are presented in Table 3. Patients with GOLD-B were older, more often single, and had a lower level of education than patients in other groups. GOLD-C had the highest percentage of current smokers (45%) while GOLD-B and GOLD-D had the greatest prevalence of cardiac morbidities, diabetes, and depression in addition to the lowest employment levels, self-efficacy scores, and levels of physical activity.

Patients with GOLD-A had the least impaired disease-specific and generic HR-QoL, followed by patients with GOLD-C, GOLD-B, and GOLD-D for all HR-QoL measures. Patients with GOLD A had the lowest costs while patients with GOLD D had the highest.

Association between GOLD classifications and HR-QoL

Table 4 shows the results of the multi-level models for HR-QoL. The AIC of the models with the GOLD ABCD groups classification were

lower than those with the GOLD 1234 grades classification, indicating a better model fit. Furthermore, these models explained more variance between patients than models with the GOLD 1234 grades classification, as indicated by a higher R^2 .

Patients with GOLD-B, GOLD-C, and GOLD-D had significantly worse HR-QoL than those with GOLD-A across all questionnaires after controlling for patient characteristics. For example, the mean CCQ score of patients with GOLD-D was 1.21 worse than the mean CCQ score of patients with GOLD-A. Appendix 1 (available online at www.pcrj.org) shows the pairwise differences from the HR-QoL models, first with GOLD-A as reference category, then GOLD-B, etc. The HR-QoL was significantly different between GOLD-B and GOLD-C, and between GOLD-C and GOLD-D, except for the SF-36 mental component. The difference in HR-QoL between patients with GOLD-B and GOLD-D was only significant for the SF-36 physical component. The GOLD group order, by increasingly impaired HR-QoL, was GOLD-A, GOLD-C, GOLD-B, and GOLD-D.

The GOLD 1234 grades classification showed greater impairment of HR-QoL as GOLD grade increased, but the differences were not statistically significant between GOLD-1 and GOLD-2 in the SF-36 physical component, between GOLD-1, GOLD-2, and GOLD-3 in the SF-36 mental component, or between GOLD-1, GOLD-2, GOLD-3, and GOLD-4 in the EQ5D score.

Association between GOLD classifications and costs

Cost models with a log normal distribution and identity link function showed the best fit (Appendix 2, available online at www.pcrj.org). The estimated β -coefficients from our cost models were transformed into the exponential form, which allows us to interpret them as the percentage of change in costs. Table 5 shows the results of these

Table 4. Health-related quality of life models

Input variable	Outcome variable	CCQ		SGRQ		EQ5D		EQ5D-VAS		SF-36 physical		SF-36 mental	
		ABCD groups	1234 grades	ABCD groups	1234 grades	ABCD groups	1234 grades	ABCD groups	1234 grades	ABCD groups	1234 grades	ABCD groups	1234 grades
	Intercept	2.54**	2.52**	37.53**	41.15**	0.43**	0.45**	44.23**	44.53**	41.55**	38.61**	20.42**	22.20**
GOLD	GOLD-B	1.04**		23.08**		-0.18**		-13.52**		-8.83**		-2.99*	
	GOLD-C	0.36**		8.38**		-0.07**		-5.53**		-2.62*		-2.33*	
	GOLD-D	1.21**		27.67**		-0.13**		-15.17**		-11.90**		-2.93*	
	GOLD-2		0.30**		7.72**		0.00		-4.31**		-2.16		-0.86
	GOLD-3		0.68**		15.52**		-0.02		-7.74**		-6.02**		-1.95
	GOLD-4		1.06**		29.65**		-0.14		-15.14**		-9.91**		-7.85**
Patient	Charlson index	0.13**	0.18**	2.15**	3.48**	-0.04**	-0.05**	-1.25*	-2.02**	-1.82**	-2.28**	-0.63	-0.70
	Age	-0.02**	-0.01*			0.00**	0.00*	0.17*	0.11			0.23**	0.20**
	Low education	0.16**	0.25**	3.17**	5.32**	0.00	-0.02	-0.69	-1.98	-0.29	-1.12	-0.94	-1.33
	Single	0.15	0.24**	1.14	4.10**	-0.03	-0.04					-2.15**	-2.45**
	Employment	-0.17*	-0.29*	-0.27	-3.56			1.16	2.94*	0.83	2.29*		
	Self-efficacy	-0.01**	-0.02**	-0.25**	-0.35**	0.00**	0.00**	0.29**	0.35**	0.07**	0.11**	0.25**	0.26**
	Smoker											-0.66	-0.41
	Former smoker											0.31	0.73
	R^2 , %		47.5	31.2	51.1	32.4	22.4	16.4	31.5	22.2	32.1	19.7	30.5
AIC		1176.6	1321.9	4341.6	4518.6	-38.2	1.4	4392.2	4457.2	3810.8	3901.0	3797.7	3797.9

* $p < 0.05$, ** $p < 0.01$.

CCQ=Clinical COPD Questionnaire, SGRQ=St George's Respiratory Questionnaire, EQ5D-VAS=visual analogue scale, AIC=Akaike's information criterion, the variable gender did not have an added value for the models, thus single is not presented in this table.

Table 5. Cost models

Input variable	Outcome variable	Healthcare costs		Medication costs		Primary care costs		Costs from societal perspective	
		ABCD groups	1234 grades	ABCD groups	1234 grades	ABCD groups	1234 grades	ABCD groups	1234 grades
		Exp(β)	Exp(β)	Exp(β)	Exp(β)	Exp(β)	Exp(β)	Exp(β)	Exp(β)
	Intercept	296.52**	312.09**	70.90**	71.07**	87.90**	94.78**	351.50**	414.30**
GOLD	GOLD-B	1.72**		1.28*		1.53**		1.53**	
	GOLD-C	1.74**		1.64**		1.20		1.70**	
	GOLD-D	2.31**		1.64**		2.37**		2.11**	
	GOLD-2		1.09		1.05		1.03		1.10
	GOLD-3		1.71**		1.50**		1.54**		1.62**
	GOLD-4		2.93**		2.01**		2.56*		3.20**
Patient	Charlson index	1.16**	1.18**	1.16**	1.17**	1.13**	1.16**	1.12**	1.14**
	Age	1.00	1.00	1.01*	1.01**				
	Men	0.87	0.82					0.84	0.82*
	Low education	1.21*	1.29**					1.13	1.19
	Employment	0.95	0.89	1.04	1.02	0.86	0.71*		
	Self-efficacy	1.00	1.00	1.00	1.00			1.00	1.00
	Smoker	0.84	0.80					0.97	0.93
	Former smoker	1.24	1.18					1.37*	1.32
R ² , %		25.8	22.1	12.1	9.8	10.6	7.4	13.3	10.9
AIC		1475.8	1500.1	1405.6	1415.5	1538.3	1556.9	1608.9	1622.6

*p<0.05, **p<0.01.

generalised linear mixed cost models. The AIC of cost models with the GOLD ABCD groups were slightly lower, indicating a better model fit, and they explained more variance than those with the GOLD 1234 grades. This suggests that models using the GOLD ABCD groups classification performed better.

Patients with GOLD-B, GOLD-C, and GOLD-D had significantly higher costs than those with GOLD-A, except for primary care costs. Total healthcare costs were significantly higher for patients with GOLD-B (72%), GOLD-C (74%), and GOLD-D (131%) than for those with GOLD-A. Costs of patients with GOLD-B were similar to those with GOLD-C, but significantly lower than those with GOLD-D (Appendix 3, available online at www.pcrj.org). Overall, the rank ordering of GOLD group by increasing costs was GOLD-A, GOLD-B, GOLD-C, and GOLD-D.

The GOLD 1234 grades classification also showed that healthcare costs increased with increasing GOLD grade (9% higher in GOLD-2, 71% higher in GOLD-3, and 193% higher in GOLD-4 compared with GOLD-1).

Discussion

Main findings

We found that the GOLD ABCD groups classification was more strongly related to HR-QoL (disease-specific and generic) and costs than the GOLD 1234 grades classification. Multi-level analysis of HR-QoL and costs using (generalised) linear mixed models incorporating the GOLD ABCD groups classification consistently explained a higher proportion of the variance between patients and had better goodness-of-fit than models incorporating the GOLD 1234 grades classification.

The rank order of GOLD ABCD groups from best to worst HR-QoL found in this primary care population was: GOLD-A, GOLD-C,

GOLD-B, GOLD-D. HR-QoL was better in patients with GOLD-C than GOLD-B, despite higher exacerbation rates and lower lung function. This can be attributed to the higher level of symptoms in GOLD-B, since the results were corrected for a higher score on the Charlson co-morbidity index, lower physical activity level, lower self-efficacy scores, lower educational level, lower employment rate, and a higher proportion of single patients with GOLD-B than with GOLD-C. Moreover, we found that the GOLD 1234 grades classification could not distinguish between GOLD-1 and GOLD-2 with respect to generic HR-QoL. Reclassifying patients with GOLD-1 and GOLD-2 into high (GOLD-B) and low symptom burden (GOLD-A) (Table 2) led to clearer differences in disease-specific and generic HR-QoL. Similarly, the separation of patients with GOLD-3 and GOLD-4 into GOLD-C and GOLD-D led to clearer distinctions in HR-QoL. Patients with lower risk levels had a significantly better HR-QoL than those with a higher risk level, and adding symptom burden to lung function improved the discriminative capabilities of the GOLD classification.

The rank order of GOLD groups by increased costs was: GOLD-A, GOLD-B, GOLD-C, GOLD-D. The GOLD 1234 grades classification could not distinguish between GOLD-1 and GOLD-2, but the GOLD ABCD groups classification did show that costs in GOLD-B, GOLD-C, and GOLD-D differed from costs in GOLD-A. Despite the fact that the model with the GOLD groups explained somewhat more variation in costs, none of the cost categories showed a statistically significant difference between GOLD-B and GOLD-C. Apparently, the higher level of symptoms in GOLD-B did not lead to more healthcare use.

Strengths and limitations of this study

Apart from the wide range of HR-QoL measures and costs used, the strength of our study lies in the large number of patients with COPD

from many general practices. The sample covers the entire range of COPD severity and is representative of the Dutch COPD population because we applied few inclusion and exclusion criteria. Although we have recruited patients in primary care, the sample also included patients seen by pulmonary physicians in hospitals. This is because patients referred to secondary care are still included in the GP's records. Our severity distribution (24% GOLD-1, 55% GOLD-2, 18% GOLD-3, 3% GOLD-4) was comparable to the severity distribution previously reported in primary care (27% GOLD-1, 55% GOLD-2, 15% GOLD-3, and 2% GOLD-4).²⁵ Moreover, like other primary care studies,^{26,27} approximately one-third of our study population had a high symptom burden (mMRC ≥ 2).

This study has some limitations. One is that healthcare utilisation was self-reported except for prescriptions, which may have led to an underestimation of costs. If this occurred to an equal extent in all four GOLD classes, differences between the four classes are not biased.²⁸ Another limitation is that the exacerbation history was retrospectively obtained from prednisone and antibiotic prescriptions in combination with ICPC codes recorded in the EMR-GP, so mild exacerbations cannot be included. Although these mild exacerbations are unlikely to affect costs, they may affect HR-QoL. A third limitation was that, despite its size (>600 patients), our sample size was still insufficient to allow subgroup analysis within GOLD-C and GOLD-D because most patients had either GOLD-A or GOLD-B. The GOLD-C and GOLD-D groups contain a mixture of patients with a high risk due to a low level of lung function, a history of frequent exacerbations, or both (Table 2). The latter patients are likely to have a worse HR-QoL and higher healthcare costs, but we were unable to confirm this.

Interpretation of findings in relation to previously published work

Since the GOLD strategy revision in 2011, several studies have investigated the distribution and characteristics of the GOLD ABCD groups and their association with outcomes.^{4,9} Lange *et al.*⁴ and Leivseth *et al.*⁹ recruited patients in the general population while Han *et al.*⁵, Soriano *et al.*⁶, Augusti *et al.*⁷, and Johannessen *et al.*⁸ recruited patients mostly in hospital settings. We recruited patients from primary care, which is most likely the reason why we found more GOLD-C patients and fewer GOLD-D patients than in the latter four studies.⁵⁻⁸ Despite the difference in recruitment setting, our finding that co-morbidities were particularly prevalent in the two 'high symptom' groups (GOLD-B and GOLD-D) is in line with previous studies.^{7,8} Furthermore, like previous reports,^{4,5} we found that patients were often categorised at high risk of exacerbation because of lung function rather than exacerbation history. Only 30% of GOLD-C patients and 19% of GOLD-D patients were assigned to high-risk groups because of their exacerbation history. Also consistent with earlier studies^{5,26} was our finding that HR-QoL for GOLD-C patients was better than for GOLD-B patients. Three studies found that three-year survival was similar for GOLD-B and GOLD-C patients,^{4,6,7} but the results seem to change in the long term because one study found that patients in GOLD-B had a better survival than those in GOLD-C after 10 years of follow-up.⁶ Two studies concluded that the GOLD ABCD groups classification did not differ significantly

from the GOLD 1234 grades classification in terms of predicting mortality^{6,8} and hospitalisations.⁸ Moreover, in one study the GOLD 1234 grades predicted mortality better than the GOLD ABCD groups.⁹

Debates on the GOLD ABCD groups classification are ongoing. For example, it has been suggested that one severe exacerbation requiring hospitalisation should qualify a patient as being at high risk.^{29,30} In our sample there were only three hospitalisations for COPD exacerbations among the low-risk group in the past three months, one in a patient with GOLD-A and two in patients with GOLD-B. Moving these patients to the high-risk groups would not have changed our results.

Implications for future research, policy, and practice

Whether the GOLD ABCD groups classification will lead to major changes in daily practice is questionable. The GOLD ABCD groups classification certainly raises awareness among physicians that assessment of COPD severity should include multiple components other than lung function. It stimulates patient-centred outcome thinking and is a step towards personalised medicine, although this would also require an integrated assessment of, for example, risk factors, biomarkers, exercise capacity, nutritional status, multi-morbidities, personal goals, illness perceptions, and coping behaviour as is currently advocated by the Dutch Care Standard for COPD.³¹

The GOLD ABCD groups classification may have consequences for decision-analytical models that aim to assess the cost-effectiveness of COPD treatments. Currently, most of these models are state transition models in which the states are defined by FEV₁% predicted and exacerbations. However, these models are also moving towards the simulation of individual patients with different characteristics. Adopting the GOLD ABCD groups classification may be a step towards this, although it still needs to be demonstrated that the cost-effectiveness of treatments indeed varies between the GOLD ABCD groups.

Conclusions

The GOLD ABCD groups classification stimulates a multi-dimensional assessment of COPD severity that appears to be more strongly related to HR-QoL (disease-specific and generic) and costs than the GOLD 1234 grades classification, which was mainly based on lung function. Reclassifying former GOLD-1 and GOLD-2 patients into the GOLD ABCD groups classification clearly led to a greater difference in HR-QoL (disease-specific and generic) and costs between groups. Similarly, separating former GOLD-3 and GOLD-4 patients into GOLD-C and GOLD-D groups led to a clearer difference in HR-QoL. Furthermore, patients with GOLD-C had a better HR-QoL than those with GOLD-B, but their costs did not differ.

Handling editor Arnulf Langhammer

Statistical review Gopal Netuveli

Conflicts of interest The authors declare that they have no conflicts of interest in relation to this article. NHC is an Associate Editor of the *PCRJ*, but was not involved in the editorial review of, nor the decision to publish, this article.

Contributorship MRSB was involved in the study design, data collection, statistical analysis, and writing of the manuscript. AT and MPMHR contributed to the study design, data collection, interpretation of the data, and writing of the manuscript. NHC and ALK participated in the data collection, interpretation of the

data and revision of the manuscript. All authors read and approved the final manuscript.

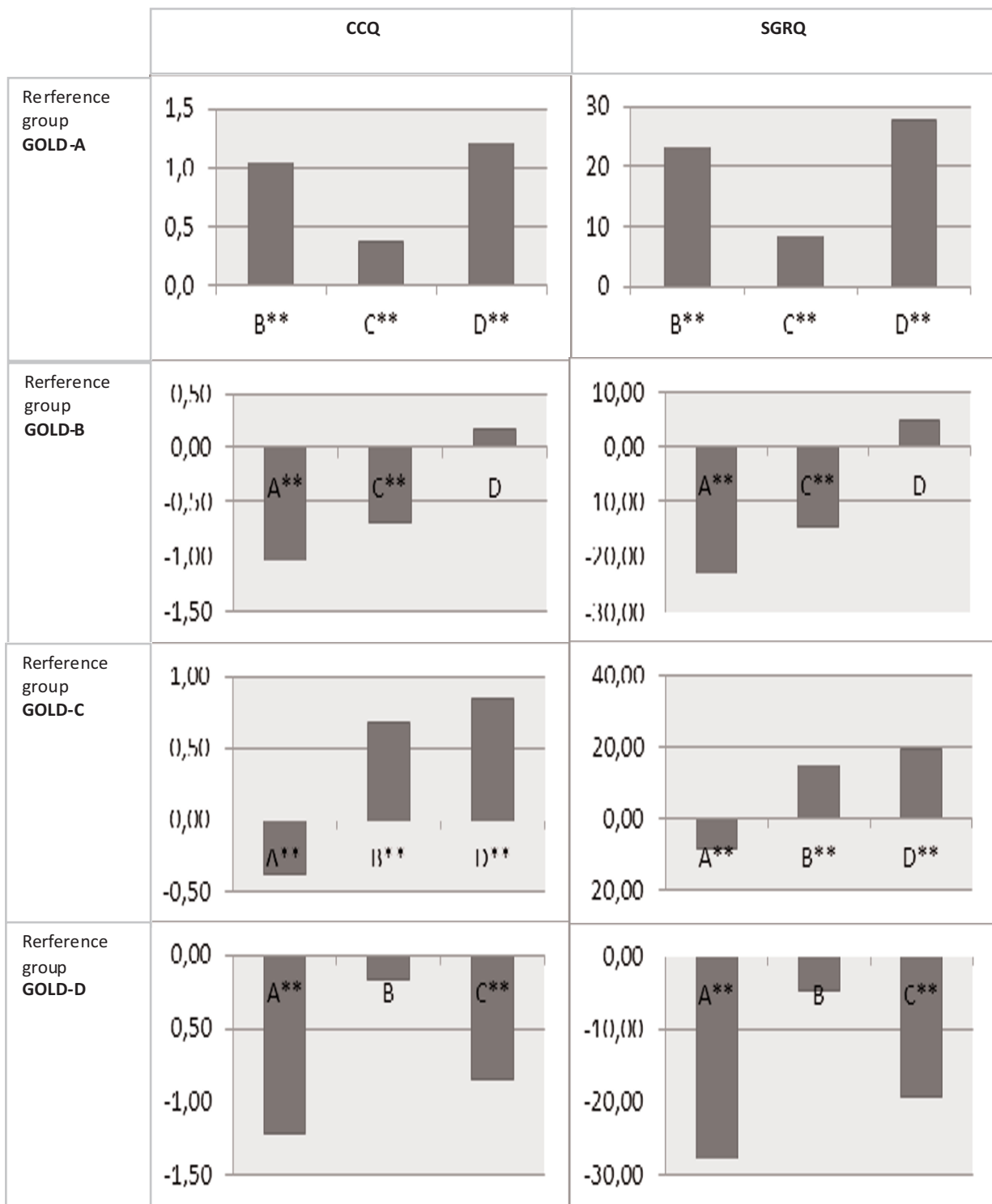
Funding This study was supported by grants from Stichting Achmea, a Dutch Healthcare Insurance Company, and the Netherlands Organisation for Health Research and Development (Zon-MW), subprogramme Effects & Costs (project number 171002203). The funding agencies (Achmea Zorgverzekeringen and Zon-MW) had no influence on the analysis and writing of the papers.

References

- Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Revised 2011. http://www.goldcopd.org/uploads/users/files/GOLD2011_Summary.pdf
- Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999;**54**(7):581-6. <http://dx.doi.org/10.1136/thx.54.7.581>
- Jones PW, Harding G, Berry P, et al. Development and first validation of the COPD assessment test. *Eur Respir J* 2009;**34**(3):648-54. <http://dx.doi.org/10.1183/09031936.00102509>
- 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):975-81. <http://dx.doi.org/10.1164/rccm.201207-1299OC>
- Han M, Muellerova H, Curran-Everett D, et al. GOLD 2011 disease severity classification in COPDGene: a prospective cohort study. *Lancet Respir Med* 2013;**1**:43-50. [http://dx.doi.org/10.1016/S2213-2600\(12\)70044-9](http://dx.doi.org/10.1016/S2213-2600(12)70044-9)
- Soriano JB, Alfageme I, Almagro P, et al. Distribution and prognostic validity of the new GOLD grading classification. *Chest* 2013;**143**:694-702.
- Agusti A, Edwards LD, Celli B, et al. Characteristics, stability and outcomes of the GOLD 2011 COPD groups in the ECLIPSE cohort. *Eur Respir J* 2013;**42**:636-46. <http://dx.doi.org/10.1183/09031936.00195212>
- 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):51-9. <http://dx.doi.org/10.1164/rccm.201212-2276OC>
- Leivseth L, Brumpton BM, Nilsen TI, Mai XM, Johnsen R, Langhammer A. GOLD classifications and mortality in chronic obstructive pulmonary disease: the HUNT study, Norway. *Thorax* 2013;**68**(10):914-21. <http://dx.doi.org/10.1136/thoraxjnl-2013-203270>
- van der Molen T, Willemse BW, Schokker S, ten Hacken NH, Postma DS, Juniper EF. Development, validity and responsiveness of the Clinical COPD Questionnaire. *Health Qual Life Outcomes* 2003;**1**:13. <http://dx.doi.org/10.1186/1477-7525-1-13>
- Kocks JW, Tuinenga MG, Uil SM, van den Berg JW, Stahl E, van der Molen T. Health status measurement in COPD: The minimal clinically important difference of the clinical COPD questionnaire. *Respir Res* 2006;**7**:62. <http://dx.doi.org/10.1186/1465-9921-7-62>
- Jones PW. St. George's Respiratory Questionnaire: MCID. *COPD* 2005;**2**(1):75-9. <http://dx.doi.org/10.1081/COPD-200050513>
- Aaronson NK, Muller M, Cohen PD, et al. Translation, validation, and norming of the Dutch language version of the SF-36 health survey in community and chronic disease populations. *J Clin Epidemiol* 1998;**51**(11):1055-68. [http://dx.doi.org/10.1016/S0895-4356\(98\)00097-3](http://dx.doi.org/10.1016/S0895-4356(98)00097-3)
- Ware JE, Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;**30**(6):473-83. <http://dx.doi.org/10.1097/00005650-199206000-00002>
- Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ* 2004;**13**(9):873-84. <http://dx.doi.org/10.1002/hec.866>
- Robinson A, Dolan P, Williams A. Valuing health status using VAS and TTO: What lies behind the numbers? *Soc Sci Med* 1997;**45**(8):1289-97. [http://dx.doi.org/10.1016/S0277-9536\(97\)00057-9](http://dx.doi.org/10.1016/S0277-9536(97)00057-9)
- Lamers LM, McDonnell J, Stalmeier PF, Krabbe PF, Busschbach JJ. The Dutch tariff: Results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ* 2006;**15**(10):1121-32. <http://dx.doi.org/10.1002/hec.1124>
- Charlson ME, Charlson RE, Peterson JC, Marinopoulos SS, Briggs WM, Hollenberg JP. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *J Clin Epidemiol* 2008;**61**(12):1234-40. <http://dx.doi.org/10.1016/j.jclinepi.2008.01.006>
- Hakkaart-van Roijen L, Tan S, Bouwmans C. Handleiding voor kostenonderzoek, methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. 2010.
- Consumentenprijzen: inflatie vanaf 1963. <http://statline.cbs.nl/StatWeb/>
- GIPdatabank. <http://www.gipdatabank.nl/>
- Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ* 1995;**14**(2):171-89. [http://dx.doi.org/10.1016/0167-6296\(94\)00044-5](http://dx.doi.org/10.1016/0167-6296(94)00044-5)
- Verbeke G, Molenberghs G. Linear mixed models for longitudinal data. New York: Springer, 2005.
- Snijders TAB, Bosker RJ. Modeled variance in two-level models. *Sociol Methods Res* 1994;**22**(3):342-63. <http://dx.doi.org/10.1177/0049124194022003004>
- Hoogendoorn M, Feenstra TL, Schermer TR, Hesselink AE, Rutten-van Molken MP. Severity distribution of chronic obstructive pulmonary disease (COPD) in Dutch general practice. *Respir Med* 2006;**100**(1):83-6. <http://dx.doi.org/10.1016/j.rmed.2005.04.004>
- Steuten LM, Creutzberg EC, Vrijhoef HJ, Wouters EF. COPD as a multicomponent disease: Inventory of dyspnoea, underweight, obesity and fat free mass depletion in primary care. *Prim Care Respir J* 2006;**15**(2):84-91. <http://dx.doi.org/10.1016/j.pcrj.2005.09.001>
- Chavannes NH, Grijzen M, van den Akker M, et al. Integrated disease management improves one-year quality of life in primary care COPD patients: a controlled clinical trial. *Prim Care Respir J* 2009;**18**(3):171-6. <http://dx.doi.org/10.3132/pcrj.2009.00003>
- Hoogendoorn M, van Wetering CR, Schols AM, Rutten-van Molken MP. Self-report versus care provider registration of healthcare utilization: impact on cost and cost-utility. *Int J Technol Assess Health Care* 2009;**25**(4):588-95. <http://dx.doi.org/10.1017/S0266462309990432>
- Wedzicha JA. GOLD and ABCD - a good start, but now for the evidence? *Lancet Respir Med* 2013;**1**(1):4-5. [http://dx.doi.org/10.1016/S2213-2600\(12\)70045-0](http://dx.doi.org/10.1016/S2213-2600(12)70045-0)
- 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):208-14. <http://dx.doi.org/10.1111/crj.12002>
- LAN. Zorgstandaard COPD. Amersfoort: Long Alliantie Nederland; 2010.

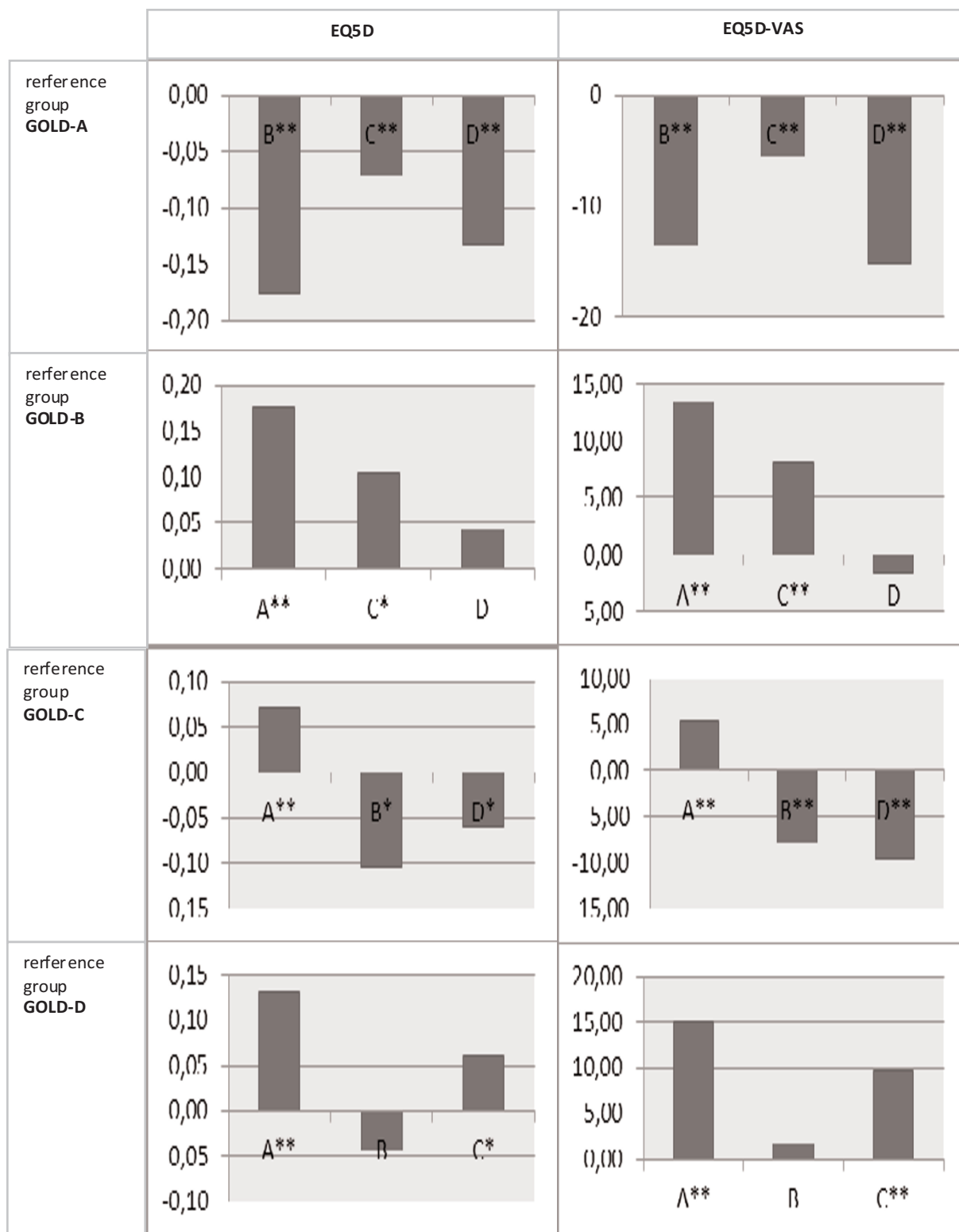
Available online at <http://www.thepcrj.org>

Appendix 1a. Differences in GOLD group estimated by the disease-specific HR-QoL models as shown in table 4 with the different GOLD groups as reference category



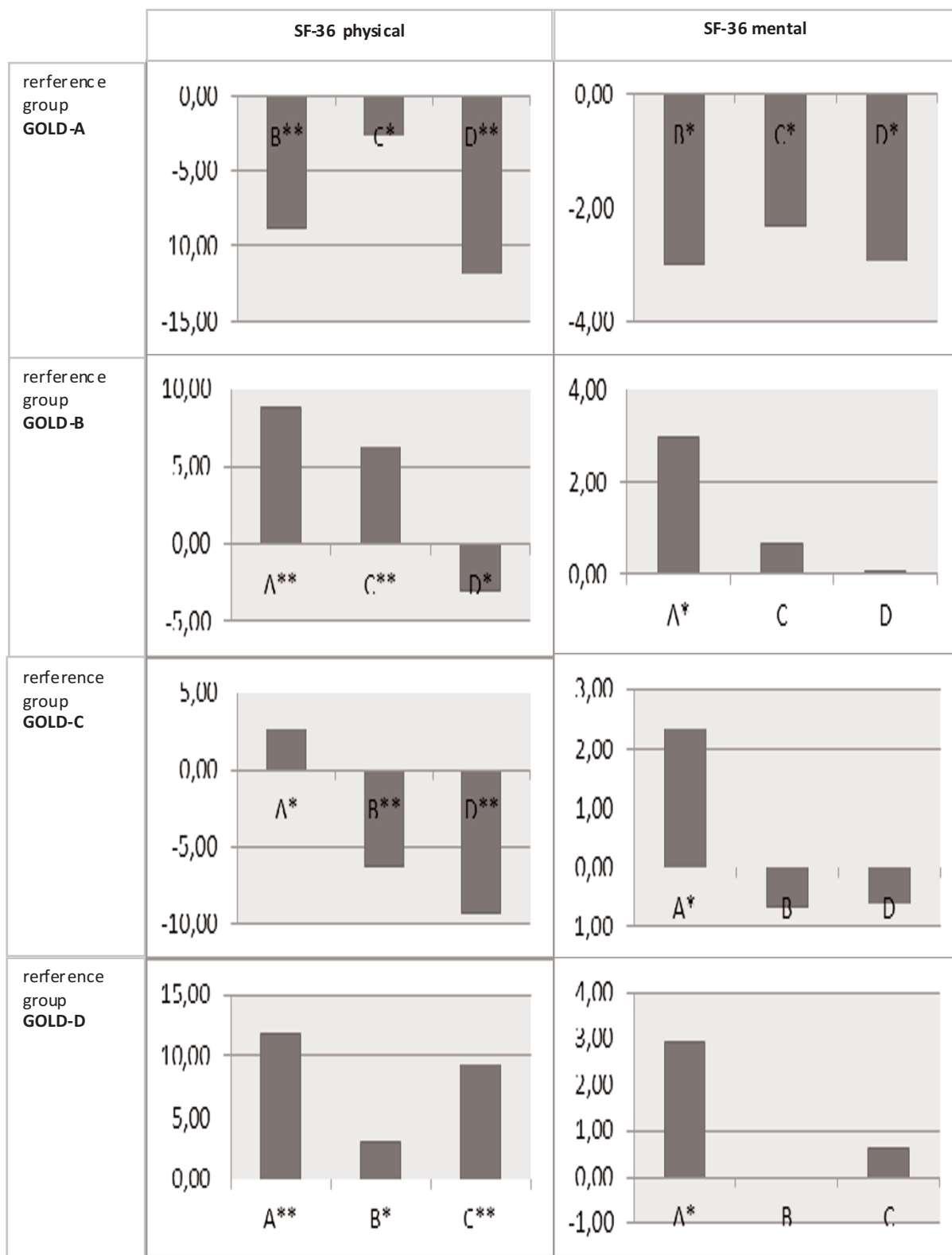
* Significant ($p < 0.05$), ** Significant ($p < 0.01$), CCQ= Clinical COPD Questionnaire, SGRQ= St George Respiratory Questionnaire

Appendix 1b. Differences in GOLD group estimated by the generic HR-QoL models as shown in table 4 with the different GOLD groups as reference category



* Significant (p<0.05), ** Significant (p<0.01) , EQ5D-VAS= visual analogue scale

Appendix 1c. Differences in GOLD group estimated by the generic HR-QoL models as shown in table 4 with the different GOLD groups as reference category

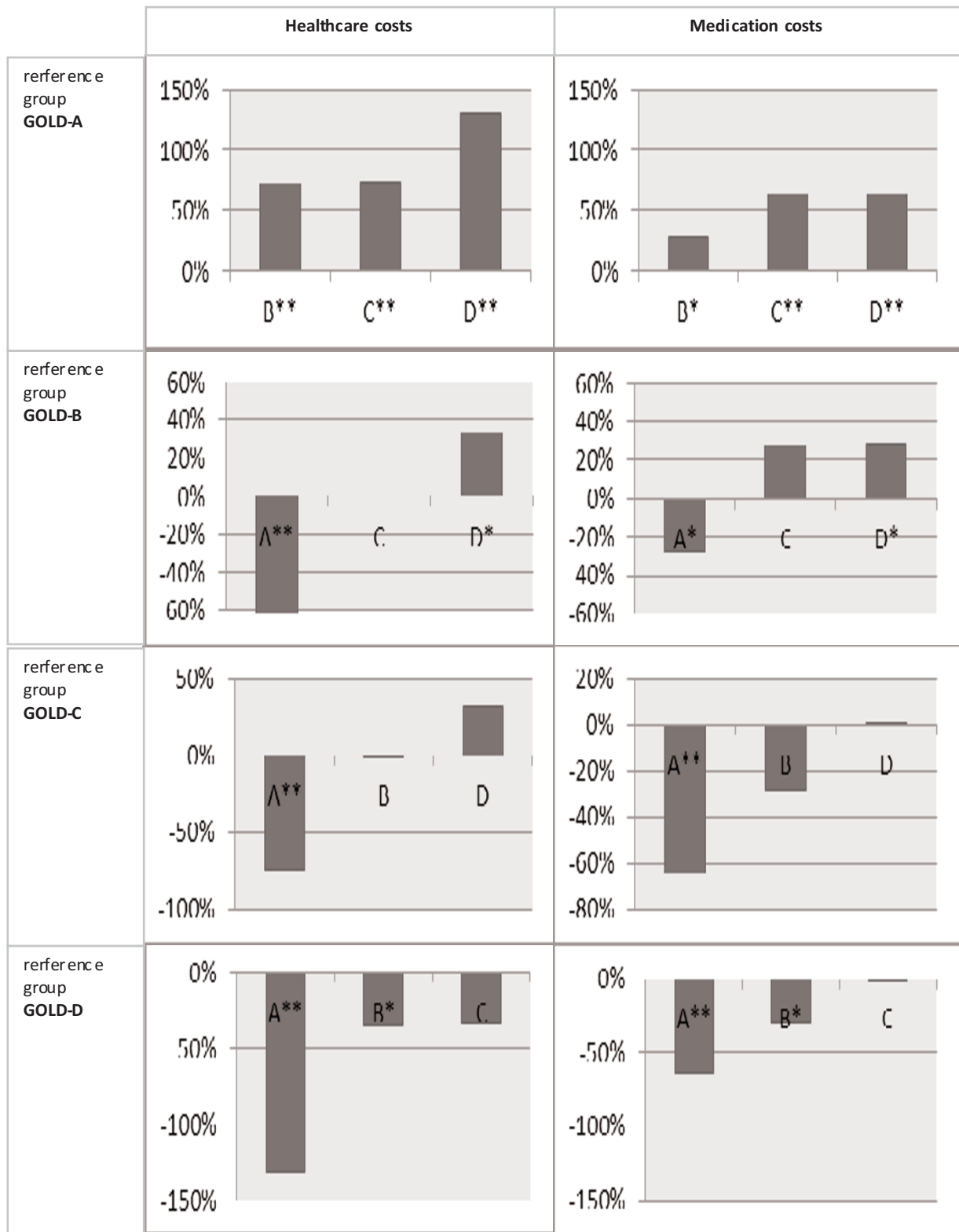


* Significant ($p < 0.05$), ** Significant ($p < 0.01$), SF-36= Short Form 36.

Appendix 2. Distribution and link function						
Outcome	Link function	Goodness-of-fit test	Distribution Gaussian	Gamma	Lognormal	Exponent
Healthcare costs	Identity	AIC	9375		1461	
		BIC	9395		1481	
	Log	AIC		8193		8200
		BIC		8213		8218
Medication costs	Identity	AIC	7182		1300	
		BIC	7201		1320	
	Log	AIC		6736		6906
		BIC		6756		6925
Primary care costs	Identity	AIC	7522		1469	
		BIC	7542		1489	
	Log	AIC		6335		6763
		BIC		6355		6781
Costs from societal perspective	Identity	AIC	9781		1557	
		BIC	9811		1577	
	Log	AIC		8431		8430
		BIC		8451		8448

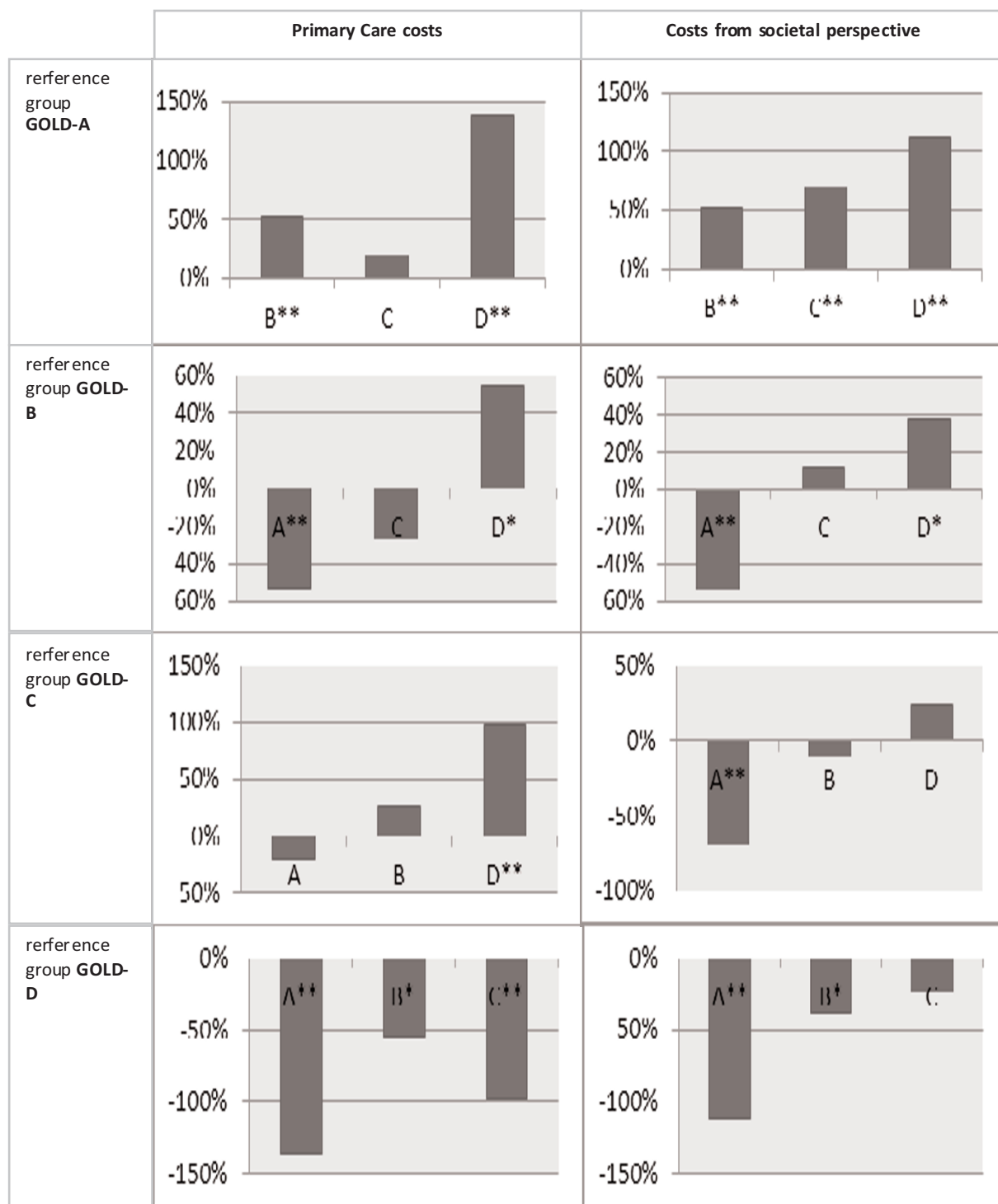
AIC=Akaike's information criterion, BIC=Bayesian information criterion.

Appendix 3a. Differences in GOLD group estimated by the cost models as shown in table 5 with the different GOLD groups as reference category



* Significant (p<0.05), ** Significant (p<0.01)

Appendix 3b Differences in GOLD group estimated by the cost models as shown in table 5 with the different GOLD groups as reference category



* Significant (p<0.05), ** Significant (p<0.01)