ERJ open research
THE BEST IN OPEN ACCESS BASIC, TRANSLATIONAL & CLINICAL RESPIRATORY RESEARCH

Early View

Original research article

Reference equations for breathlessness during incremental cycle exercise testing

Viktor Elmberg, Linus Schiöler, Thomas Lindow, Kristofer Hedman, Andrei Malinovschi, Hayley Lewthwaite, Dennis Jensen, Lars Brudin, Magnus Ekström

Please cite this article as: Elmberg V, Schiöler L, Lindow T, *et al*. Reference equations for breathlessness during incremental cycle exercise testing. *ERJ Open Res* 2023; in press (https://doi.org/10.1183/23120541.00566-2022).

Word count for body of text: 3182

Word count for abstract: 242

Reference equations for breathlessness during incremental cycle exercise testing

Short title: Reference equations for breathlessness

Viktor Elmberg, MD[1,2]; Linus Schiöler, PhD[3]; Thomas Lindow, MD, PhD[4,5]; Kristofer Hedman, MD, PhD [6]; Andrei Malinovschi, MD, PhD[7]; Hayley Lewthwaite, PhD[8,9]; Dennis Jensen, PhD[10,11]; Lars Brudin, MD, PhD[12]; Magnus Ekström, MD, PhD[1]

[1] Department of Clinical Sciences, Division of Respiratory Medicine, Allergology and Palliative Medicine, Lund University, Lund, Sweden.

[2] Department of Clinical Physiology, Blekinge Hospital, Karlskrona, Sweden.

[3] Occupational and Environmental Medicine, School of Public Health and Community Medicine, Institute of Medicine, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden.

[4] Department of Clinical Physiology, Research and Development, Växjö Central Hospital, Region Kronoberg, Sweden

[5] Kolling Institute, Royal North Shore Hospital, University of Sydney, Sydney, Australia

[6] Department of Clinical Physiology in Linköping, and Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden,

[7] Department of Medial Sciences, Clinical Physiology, Uppsala University, Uppsala, Sweden,

[8] School of Environmental & Life Sciences, College of Engineering, Science and Environment, University of Newcastle, Ourimbah, Australia

[9] UniSA: Allied Health and Human Performance, Innovation, Implementation and Clinical Translation in Health, University of South Australia, Adelaide, Australia

[10] Clinical Exercise and Respiratory Physiology Laboratory, Department of Kinesiology and Physical Education, McGill University, Montreal, Quebec, Canada

[11] Research Institute of the McGill University Health Centre, Translational Research in Respiratory Diseases Program, Montreal, Quebec, Canada

[12] Department of Clinical Physiology, Kalmar County Hospital, Kalmar,Sweden

Corresponding author: Viktor Elmberg, Department of Clinical Physiology, Blekinge Hospital, Karlskrona, Sweden. Tel: +46(0)4551000. Email: viktore@gmail.com

Funding: VE was funded by an unrestricted grant from the Scientific Committee of Blekinge Region. TL was funded by an unrestricted grant from the Department of Research and Development, Region Kronoberg. DJ holds a Canada Research Chair, Tier II, in Clinical Exercise and Respiratory Physiology from the Canadian Institutes of Health Research. ME was supported by unrestricted grants from the Swedish Society for Medical Research and the Swedish Research Council (Dnr: 2019-02081).

Conflict of interest: No conflicts of interest exist for the authors.

Authorship: LS performed the data analyses. VE, LS, TL, KH, AM, HL, DJ, LB and ME contributed substantially to the study design, data analysis and interpretation, and the writing of the manuscript.

Acknowledgments: The authors acknowledge the efforts made by the staff at the Department of Clinical Physiology at Kalmar County Hospital for compiling the database and caring for the patients.

Number of tables: 1**; Number of figures:** 3**; Number of references:** 32**; Number of supplements:** 2

Abbreviations: BMI = body mass index, Borg CR10 = Borg Category-ratio scale (0-10, Borg RPE = Borg rating of perceived exertion scale, CPET = cardiopulmonary exercise testing, ECG = electrocardiogram, FEV1 = forced expiratory volume in 1-s, GEE = generalized estimating equations, IET = Incremental Exercise Test, mMRC = Medical Research Council, QIC = quasi-information criterion, SD = standard deviations, ULN = upper limit of normal, W = Watt, W_{max} = predicted peak power output, W_{peak} = peak power output, %predW_{max} = W % of predicted Wmax.

Data Availability: The data underlying this article will be shared on reasonable request to the corresponding author.

Take home message Reference equations for breathlessness intensity are of importance for research and clinical care. We present the first reference equations for breathlessness during incremental cycle exercise testing for use to assess and compare breathlessness severity.

ABSTRACT

Background: Exertional breathlessness is commonly assessed using incremental exercise testing (IET), but reference equations for breathlessness responses are lacking. We aimed to develop reference equations for breathlessness intensity during IET.

Methods: Retrospective, consecutive cohort study of adults undergoing IET in Sweden. Exclusion criteria included cardiac or respiratory disease, death or any of the aforementioned diagnoses within one year of the IET, morbid obesity, abnormally low exercise capacity, submaximal exertion, or an abnormal exercise test. Probabilities for breathlessness intensity ratings (Borg CR10) during IET in relation to power output (%predWmax), age, sex, height, and body mass were analyzed using marginal ordinal logistic regression. Reference equations for males and females were derived to predict the upper limit of normal (ULN), and the probability of different Borg CR10 intensity ratings.

Results: 2,581 participants (43% female) aged 18-90 years were included. Mean breathlessness intensity was similar between sexes at peak exertion (6.7±1.5 *vs.* 6.4±1.5 Borg CR10 units) and throughout exercise in relation to %predW_{max}. Final reference equations included age, height and %predW_{max} for males, whereas height was not included for females. The models showed a close fit to observed breathlessness intensity ratings across %predW_{max} values. Models using absolute W did not show superior fit. Scripts are provided for calculating the probability for different breathlessness intensity ratings, and the ULN by %predW_{max} throughout IET.

Conclusion: We present the first reference equations for interpreting breathlessness intensity during incremental cycle exercise testing in males and females aged 18-90 years.

Keywords: Breathlessness, dyspnea, exercise testing, exercise capacity, reference values, CPET

INTRODUCTION

Breathlessness is common and a cardinal symptom of cardiac or respiratory disease [1]. Breathlessness is often provoked by physical activity and may, in severe cases, be present during low-level exertion or at rest [1-3].

Assessment of breathlessness should account for the level of exertion, as people tend to exercise up to a similar intensity of breathlessness regardless of exercise capacity and/or health status [4-8]. Therefore, breathlessness intensity measured during daily activities, or near an individual's limit of exercise tolerance, may be similar even between healthy individuals and people with cardiac or respiratory disease [8-11]. In addition, people often decrease their physical activity to avoid breathlessness, a behavior of avoidance that can lead to a cycle of reduced physical activity, deconditioning, and worsening activity-related breathlessness.[12].

Intensity of exertional breathlessness is commonly assessed using a validated scale, such as Borg's 0-10 category ratio scale (Borg CR10), during symptom-limited incremental exercise testing (IET) in relation to a standardized relative exercise intensity [5-7, 13-15]. IET is a key test for evaluation of breathlessness severity in clinical care and research [6, 8, 14].

Reference equations to predict an individual's expected normal breathlessness intensity at different standardized submaximal power outputs during IET are lacking. This is despite symptom-limited IET being widely used in clinical practice for evaluation of breathlessness and exercise (in)tolerance [5, 7]. Killian *et al.* reported reference equations for breathlessness intensity in 460 healthy individuals aged 20-70 years with normal exercise capacity [16]. Their reference equations were limited by an assumption of normally distributed scores and predictions using linear regression (yielding predicted scores outside the Borg CR10 scale range) and have not been adopted in clinical care or research [16]. Neder *et al.* recently reported on the breathlessness intensity response to cardiopulmonary cycle exercise testing

(CPET) descriptively (n=275) including the 95th percentile, which could be used to determine the upper limit of normal (ULN) breathlessness intensity for a range of absolute power outputs and levels of minute ventilation. Approximate median values of breathlessness as well as ULN for breathlessness intensity ratings were also provided as tables. However, Neder *et al.* did not provide reference equations to predict probabilities for reporting different breathlessness Borg CR10 scores, deviation of a breathlessness intensity score from the predicted reference value, nor the ULN breathlessness response based on readily available participant characteristics [17]. Reference equations to predict normal and abnormal breathlessness (>ULN) intensity ratings during IET are important for the evaluation of breathlessness in research and clinical care, as well as to compare the severity of breathlessness between individuals or groups.[6, 14].

The aim of this study was to develop reference equations for breathlessness intensity ratings (Borg CR10) across submaximal power outputs during symptom-limited IET in a large cohort of ostensibly healthy adults.

MATERIAL AND METHODS

Study design and participants

This was a consecutive cohort study of people free of clinically apparent disease that was referred for IET at the Department of Clinical Physiology at Kalmar County Hospital, Sweden between May 2005 and October 2016. IETs were performed according to the Swedish protocol for standardized cycle exercise testing [18]. This database is the basis for the established Swedish reference values for predicted peak power output (W_{max}) and exertional systolic blood pressure response [19-22].

Inclusion criteria were: age ≥ 18 years; normal exercise capacity defined as peak power output (W_{peak}) within 75 - 125 % of the participant's predicted W_{max} [21]; and a normal IET as determined by the attending physician. Exclusion criteria were: 1) ischemic heart disease, heart failure, atrial fibrillation, asthma or chronic obstructive pulmonary disease recorded in patient records during the five years before the IET; 2) death or any of the aforementioned diagnoses within one year after the IET; 3) reason for referral to the IET being either dyspnea, aortic stenosis, post infarction and/or pacemaker; 4) exercise time <5 minutes or submaximal exertion defined as a rating of perceived exertion (RPE) <16 on Borg's 6-20 scale at end exercise [23]; 5) body mass index (BMI) >35 kg/m^2 , i.e., morbid obesity or higher; and 6) no Borg CR10 breathlessness intensity rating during the last two minutes of the IET. For participants who performed more than one IET during the study period, the most recent was included.

Ethical considerations

The study was conducted in accordance with the amended Declaration of Helsinki and was approved by the Regional Ethical Review Board in Linköping (DNr: 2018/141-31). As this was an observational study using data collected from IETs in routine clinical practice, individual participant consent was waived. The study is reported in accordance with the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) statement [24].

Exercise test protocol and assessments

The protocol for the standardized symptom-limited IET has been detailed previously [19, 21, 22]. All tests were performed using an electrically braked cycle ergometer (Rodby Inc, Karlskoga, Sweden). The initial power output and ramp increment of 10, 15 or 20 W/min was selected depending on the participant's predicted W_{peak}, aiming at an exercise duration of 8-12 minutes [19, 21, 22]. Differences in power output increment between individuals are accounted for by the Swedish reference values for W_{max} [21].

Before IET, resting 12-lead electrocardiogram (ECG), body mass and height were recorded. The Borg RPE 6-20 and CR10 scales and their anchors (for pain, exertion, and breathlessness) between the extremes were explained. Participants were invited to rate their overall intensity of breathlessness on the Borg CR10 scale, which was also shown, from 0= no breathlessness to 10= the most intense breathlessness that you've experienced or could imagine experiencing [13]. During IET, ECG was recorded continuously, whereas systolic blood pressure, RPE (Borg 6-20 scale), breathlessness intensity (Borg CR10), and chest pain (Borg CR10) were measured every 2 minutes.

Statistical analyses

Characteristics were tabulated and compared using descriptive statistics. Associations between breathlessness intensity (Borg CR10) during IET and potential determinants (W % of predicted W_{max} [%predW_{max}], age, sex, height, and body mass) based on the literature [16, 17, 21] were evaluated using heatmaps and locally estimated scatterplot smoothing (LOESS) curves. Percentiles of calculated breathlessness intensity were plotted across %predW_{max}.

Modeling the reference equations involved the following considerations. Usual inferential statistics, such as p-values and confidence intervals were considered less useful, since we were not interested in making statements regarding any particular estimand. The aim was also not to produce the best prediction possible for each individual, but to describe the distribution among a group of people free of clinically apparent disease with similar characteristics. Thus, the calibration of the model was considered as the most important measure.

The reference equations for breathlessness intensity ratings were modeled using marginal ordinal logistic regression with a cumulative logit link [25], estimated by generalized estimating equations (GEE). This approach was chosen due to breathlessness being rated on the discrete Borg CR10 scale, and hence models assuming continuous, conditionally normally distributed values are unsuitable and may yield predictions outside the 0 – 10 scale range. Instead, the current model predicted the probability of rating each score on Borg's CR10 scale by %predW_{max} and values of the other covariates. Initially, we included age, height, body mass and %predW_{max} in separate models for males and females, with all variables as natural/restricted cubic splines [26] with knots placed at the 5th, 35th , 65th and 95th percentiles. After evaluation, height was excluded from the model for females and kept as a linear variable for males, as this resulted in an improved quasi-information criterion (QIC) [27] for the models (supplemental **Table A1**). Furthermore, we included the linear interaction between age and %predW_{max}, as plots indicated a lack of fit in the upper age ranges, which improved model QIC. Body mass was excluded, even though it improved QIC slightly for men, as calibration was affected negatively in the upper range (supplemental **Figure A1**). We also conducted all analyses by including %W_{peak} instead of %predW_{max} in the models, which resulted in similar findings without any improvements in model fit. We investigated the proportional odds assumption by comparing the odds ratios from the proportional odds model to odds ratios calculated by binary logistic regression models for each cut-point (i.e. setting the outcome to breathlessness rating <i for i=1,..10). We found no major violations of the assumption (supplemental **Figure A2**).

From the regression models, we calculated the probability of having a specific Borg CR10 breathlessness rating, as well as the probability of having an equal or greater breathlessness rating (interpreted as the percentage of the population having an equal or greater value). Z-scores, reflecting the deviation of the observed breathlessness rating from the predicted reference rating, expressed in standard deviations (SDs), were calculated by entering the probability of having a rating less than the observed value into the standard normal quantile function. The ULN value for breathlessness intensity was calculated using linear interpolation of the linear predictor of the responses closest to below and above a probability of 0.95.

Model fit was evaluated using calibration curves, a LOESS curve of the observed proportion having a specific Borg CR10 breathlessness intensity rating versus the predicted probabilities of the model (supplemental **Figure A3**). Ideally, these should be identical; for example, 20% of the participants with a predicted probability of 0.2 having a breathlessness intensity rating of 7 Borg CR10 scale units should have that rating observed. The analyses showed that the models were well calibrated. Coefficients for the final models are presented in supplemental **Table A2** and the knots used for splines in supplemental **Table A3**. In addition, predicted breathlessness intensity ratings for each participant were plotted against %predW_{max}, with loess curves of both predicted and observed breathlessness intensity ratings.

As a first step towards validation of the model, the total population was randomly split into a 70% and 30% group. The model was then re-estimated using the larger subgroup with the smaller group used as a validation sample.

RESULTS

A total of 2,581 participants (43% female) aged 18-90 years were included (**Table 1**). Nearly all were refereed to IET for either suspect stable coronary syndrome (n=1,844), occupational reasons (n=269), suspect arrhythmia (n=226) or determination of exercise capacity (n=186). The mean ± SD age was 46.6 ± 14.5 years for males and 54.0 ± 13.7 years for females. By design, all participants had an exercise capacity within the normal predicted range (defined as a W_{peak} 75-125% of expected W_{max}) [21], with males and females achieving a W_{peak} of 241 ± 44 W (98 ± 12 % predicted) and 142 ± 28 W (99 ± 13 % predicted), respectively. Breathlessness intensity ratings at peak exercise were similar between males and females, 6.7 ± 1.5 vs. 6.4 ± 1.5 Borg CR10 scale units.

For the final models, the relationship between increasing breathlessness intensity ratings and increasing power outputs during IET by age and sex for the whole cohort are shown in

Figure 1 (absolute power output in W is used instead of %predW_{max} for comparability with, for example, the study by Neder *et al*) [28]. Breathlessness intensity ratings increased with increasing power output during IET across all groups, with younger participants and males reaching a higher mean absolute W_{peak} than older participants and females, respectively. Breathlessness intensity in relation to %predW_{max} was similar between males and females (**Figure 2**).

The final reference equations incorporated age, height (males), and %predW_{max}. Predicted breathlessness intensity ratings generated by the models showed a close fit with observed breathlessness ratings throughout the IET (**Figure 2**).

The split group validation resulted in similar estimates (regression coefficients) as when running the model on the whole population. The model was well calibrated when used in the test sample with a close fit with observed breathlessness intensities (supplemental **Figure A4**).

A script of the model and an excel file are provided as a supplement (**Supplement 2**) for calculation of: 1) probabilities for different breathlessness intensity ratings by %predW_{max}, age, sex (and height in males); and 2) the ULN for breathlessness intensity ratings across different %predW_{max} values. As an illustration, **Figure 3** shows the ULN for breathlessness intensity ratings across %predW_{max} for males or females aged 70 years with a height of 180 cm. The ULN for breathlessness intensity ratings in relation to %predW_{max} was similar for other ages, but tended to be slightly lower with increasing age. Including %W_{peak} instead of %predW_{max} in the model resulted in an inferior model fit (lower QIC) and calibration as compared to %pred$W_{max.}$

DISCUSSION

Main findings

This study presents the first reference equations that enable prediction of the normal and ULN for breathlessness intensity (Borg CR10) ratings at different %predW_{max} values during IET based on readily available participant characteristics. Our findings are based on a large Swedish cohort that forms the basis for the Swedish reference values for exercise capacity (W_{max}) and systolic blood pressure response during IET [19-22].

These reference equations can be used to calculate: 1) the ULN for breathlessness intensity, for a given %predW_{max}; 2) probabilities for reporting each breathlessness Borg CR10 score, and 3) deviation of a breathlessness intensity score from the predicted reference value. The equations enable breathlessness evaluation throughout levels of exertion and using submaximal tests.

We used %predW_{max} in the models, to facilitate comparisons between people and groups. We suggest that W_{max} should be calculated using the most representative reference material for the individual and/or population under investigation. This approach is similar to spirometry reference values, where lung function can be graded and compared (as % predicted) using the best reference material for each individual and/or population. Although this cannot be inferred from our study, we think a similar approach for exertional breathlessness assessment could facilitate comparisons of symptom severity between populations and should be validated in future studies. It is sometimes suggested that breathlessness intensity should be evaluated in relation to the person's achieved peak power output (%W_{peak}) during the IET. However, as the achieved W_{peak} is influenced by the underlying health status, assessing breathlessness severity in relation to W_{peak} is less informative than in relation to the predicted maximal power output. This is especially the case for tests that are symptom-limited, where most people discontinue exercise at breathlessness intensities of 6-8 Borg CR10 units, similarly across healthy people and those with illness [4, 9, 11]. In the current study, although males and females stopped exercising at significantly different W_{peak}, the temporal breathlessness intensity response in relation $\%predW_{max}$ were similar between males and females. Using $\%predW_{max}$ instead of $\%W_{peak}$ in our models resulted in superior model fit (QIC) and better calibration, while also making differences in breathlessness intensity easier to visualize.

The present study extends the methodology used in the studies by Killian and Satia [16, 29], only requiring basic characteristics and yielding only predictions within the Borg CR10 scale range. Neder *et al.* recently reported data for 275 healthy adults performing incremental cycle CPET [28], including breathlessness intensity ratings (Borg CR10) by W_{peak}, absolute ventilation (L/min), age and sex. Tables with approximate median values as well as ULN for breathlessness intensity ratings were provided. We build on this by developing reference equations to predict individual breathlessness intensity ratings based on readily available participant characteristics at any point during IET, from lowest to maximal exertion. The trajectories of breathlessness intensity ratings during symptom-limited incremental cycle CPET reported by Neder *et al.* in healthy people [28] closely mirror those in the present study.

Strengths and limitations

Strengths of the present study include its use of a large database of males and females across a wide age-span (18-90 years), with standardized measurement of breathlessness intensity using Borg's CR10 scale. Our reference equations only incorporate widely available data – age, sex, height — and do not require lung function assessment. As breathlessness intensity may be assessed at any $\%predW_{max}$, evaluation is possible even without maximal exertion, which is important in many clinical populations. The reference equations showed a very good fit with observed values throughout exercise as shown in **Figure 2**. There was a small deviation towards the extremes of breathlessness intensity likely caused by relatively few included participants reaching these high workloads. As an example, this results in a small overestimation of predicted breathlessness intensity for older males, and a small underestimation for younger males at the highest relative workloads (supplemental **Figure A5**). This will likely not have any practical implications when using the normal values as the deviation is small and only applies to few individuals.

A limitation is that the database was not derived from a random healthy population sample. However, the database has been extensively validated and forms the basis for the current Swedish reference values for W_{max} and systolic blood pressure response during IET [19, 21]. An extensive set of exclusion criteria was applied to restrict the sample to people with normal W_{peak} and without known or likely disease. We did not have information on the presence of possible mental health comorbidities such as anxiety and/or depression that might influence the perception of, and risk of developing, breathlessness [30, 31]. However, any reference material must always be put into a clinical context with individual assessment of whether its applicable to a specific person. Physiological data were limited, with no access to data on pulmonary function tests, ventilation or gas exchange, limiting our ability to ensure that participants provided maximal effort during the IET (i.e., respiratory exchange ratio). However, mean peak heart rate was within 2% of the predicted maximal value and mean peak Borg RPE ratings were 17-18 implying maximal effort [23]. Further, we had no data on smoking or heart function. An independent Swedish database of Borg CR10 scale breathlessness intensity ratings during IET in people without known disease for external validation was not available. Internal validation using a split sample approach, demonstrated good fit and robustness of the models. In addition, the intensity ratings of exertional breathlessness were similar to reports on healthy people undergoing symptom-limited incremental cycle CPET [28]. However, external validation of the reference values would be valuable. Lastly, our results from cycle IETs (non-weight bearing exercise) are likely not applicable to exercise tests performed on a treadmill (weight bearing exercise) given known differences in cardiac, metabolic and ventilatory responses between the two modes of exercise as well as greater difficulty quantifying external power output during treadmill compared with cycle exercise [32, 33].

Importance and future directions

Reference values for intensity ratings of breathlessness during IET are important for research and clinical care as they allow: (a) determination of the normality of a reported breathlessness intensity during IET; (b) unmasking of abnormally high burden of exertional breathlessness, which may not be apparent at rest or during daily activities, particularly when assessed using task-based questionnaires such as the modified Medical Research Council dyspnea scale [34] or when the level of exertion is not standardized (e.g., six-minute walk test); c) comparison of the level of exertional breathlessness between individuals and/or between populations; and d) selection and stratification of participants for clinical trials based on their severity of exertional breathlessness. The prognostic and clinical implication of having a breathlessness intensity-%predW_{max} response to IET >ULN should be evaluated in further work.

Conclusion

Reference equations are presented for breathlessness intensity and the ULN (Borg CR10) for different %predW_{max} during IET in Swedish males or females aged 18-90 years.

REFERENCES

1. Parshall MB, Schwartzstein RM, Adams L, Banzett RB, Manning HL, Bourbeau J, Calverley PM, Gift AG, Harver A, Lareau SC, Mahler DA, Meek PM, O'Donnell DE, American Thoracic Society Committee on D. An official American Thoracic Society statement: update on the mechanisms, assessment, and management of dyspnea. *Am J Respir Crit Care Med* 2012: 185(4): 435-452.
2. Simon ST, Higginson IJ, Benalia H, Gysels M, Murtagh FE, Spicer J, Bausewein C. Episodic and continuous breathlessness: a new categorization of breathlessness. *J Pain Symptom Manage* 2013: 45(6): 1019-1029.
3. Johnson MJ, Yorke J, Hansen-Flaschen J, Lansing R, Ekström M, Similowski T, Currow DC. Towards an expert consensus to delineate a clinical syndrome of chronic breathlessness. *Eur Respir J* 2017: 49(5).
4. Faisal A, Alghamdi BJ, Ciavaglia CE, Elbehairy AF, Webb KA, Ora J, Neder JA, O'Donnell DE. Common Mechanisms of Dyspnea in Chronic Interstitial and Obstructive Lung Disorders. *American journal of respiratory and critical care medicine* 2016: 193(3): 299-309.
5. Puente-Maestu L, Palange P, Casaburi R, Laveneziana P, Maltais F, Neder JA, O'Donnell DE, Onorati P, Porszasz J, Rabinovich R, Rossiter HB, Singh S, Troosters T, Ward S. Use of exercise testing in the evaluation of interventional efficacy: an official ERS statement. *The European respiratory journal* 2016: 47(2): 429-460.
6. Ekstrom M, Elmberg V, Lindow T, Wollmer P. Breathlessness measurement should be standardised for the level of exertion. *Eur Respir J* 2018: 51(5).
7. ATS/ACCP Statement on cardiopulmonary exercise testing. *American journal of respiratory and critical care medicine* 2003: 167(2): 211-277.
8. Lewthwaite H, Jensen D, Ekström M. How to Assess Breathlessness in Chronic Obstructive Pulmonary Disease. *Int J Chron Obstruct Pulmon Dis* 2021: 16: 1581-1598.
9. Schaeffer MR, Ryerson CJ, Ramsook AH, Molgat-Seon Y, Wilkie SS, Dhillon SS, Mitchell RA, Sheel AW, Khalil N, Camp PG, Guenette JA. Effects of hyperoxia on dyspnoea and exercise endurance in fibrotic interstitial lung disease. *The European respiratory journal* 2017: 49(5).
10. O'Donnell DE, Elbehairy AF, Faisal A, Webb KA, Neder JA, Mahler DA. Exertional dyspnoea in COPD: the clinical utility of cardiopulmonary exercise testing. *Eur Respir Rev* 2016: 25(141): 333-347.
11. Abdallah SJ, Wilkinson-Maitland C, Saad N, Li PZ, Smith BM, Bourbeau J, Jensen D. Effect of morphine on breathlessness and exercise endurance in advanced COPD: a randomised crossover trial. *The European respiratory journal* 2017: 50(4).
12. Ramon MA, Ter Riet G, Carsin AE, Gimeno-Santos E, Agusti A, Anto JM, Donaire-Gonzalez D, Ferrer J, Rodriguez E, Rodriguez-Roisin R, Puhan MA, Garcia-Aymerich J. The dyspnoea-inactivity vicious circle in COPD: development and external validation of a conceptual model. *Eur Respir J* 2018: 52(3).
13. Borg GAV. Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise* 1982: 14(5): 377-381.
14. Ekstrom M. Why treatment efficacy on breathlessness in laboratory but not daily life trials? The importance of standardized exertion. *Current opinion in supportive and palliative care* 2019: 13(3): 179-183.
15. Palange P, Ward SA, Carlsen K-H, Casaburi R, Gallagher CG, Gosselink R, O'Donnell DE, Puente-Maestu L, Schols AM, Singh S, Whipp BJ. Recommendations on the use of exercise testing in clinical practice. *European Respiratory Journal* 2007: 29(1): 185-209.
16. Killian KJ, Summers E, Jones NL, Campbell EJ. Dyspnea and leg effort during incremental cycle ergometry. *Am Rev Respir Dis* 1992: 145(6): 1339-1345.
17. Lansing RW, Gracely RH, Banzett RB. The multiple dimensions of dyspnea: review and hypotheses. *Respir Physiol Neurobiol* 2009: 167(1): 53-60.
18. Jorfeldt L, Pahlm O. Kliniska arbetsprov: metoder för diagnos och prognos. *Studentlitteratur* 2013.

19. Hedman K, Lindow T, Elmberg V, Brudin L, Ekström M. Age- and gender-specific upper limits and reference equations for workload-indexed systolic blood pressure response during bicycle ergometry. *European journal of preventive cardiology* 2020: 2047487320909667.
20. Hedman K, Lindow T, Carlen A, Cauwenberghs N, Elmberg V, Brudin L, Ekstrom M. Lower peak systolic blood pressure during exercise testing predicts higher risk of all-cause mortality even when accounting for exercise capacity and other confounders. *European Heart Journal* 2020: 41(Supplement_2).
21. Brudin L, Jorfeldt L, Pahlm O. Comparison of two commonly used reference materials for exercise bicycle tests with a Swedish clinical database of patients with normal outcome. *Clin Physiol Funct Imaging* 2014: 34(4): 297-307.
22. Lindow T, Brudin L, Elmberg V, Ekström M. Long-term follow-up of patients undergoing standardized bicycle exercise stress testing: new recommendations for grading of exercise capacity are clinically relevant. *Clinical physiology and functional imaging* 2020: 40(2): 83-90.
23. Peterman JE, Arena R, Myers J, Harber MP, Bonikowske AR, Squires RW, Kaminsky LA. Reference Standards for Peak Rating of Perceived Exertion During Cardiopulmonary Exercise Testing: Data from FRIEND. *Medicine and science in sports and exercise* 2022.
24. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of clinical epidemiology* 2008: 61(4): 344-349.
25. Agresti A. Categorical data analysis. Wiley, Hoboken, NJ, 2013.
26. Hastie TTRFJH. The elements of statistical learning : data mining, inference, and prediction, 2009.
27. Pan W. Akaike's information criterion in generalized estimating equations. *Biometrics* 2001: 57(1): 120-125.
28. Neder JA, Berton DC, Nery LE, Tan WC, Bourbeau J, O'Donnell DE. A frame of reference for assessing the intensity of exertional dyspnoea during incremental cycle ergometry. *Eur Respir J* 2020: 56(4).
29. Satia I, Farooqi MAM, Cusack R, Matsuoka M, Yanqing X, Kurmi O, O'Byrne PM, Killian KJ. The contribution of FEV(1) and airflow limitation on the intensity of dyspnea and leg effort during exercise. Insights from a real-world cohort. *Physiological reports* 2020: 8(8): e14415.
30. Neuman Å, Gunnbjörnsdottir M, Tunsäter A, Nyström L, Franklin KA, Norrman E, Janson C. Dyspnea in relation to symptoms of anxiety and depression: A prospective population study. *Respiratory Medicine* 2006: 100(10): 1843-1849.
31. Scano G, Gigliotti F, Stendardi L, Gagliardi E. Dyspnea and emotional states in health and disease. *Respiratory Medicine* 2013: 107(5): 649-655.
32. Muscat KM, Kotrach HG, Wilkinson-Maitland CA, Schaeffer MR, Mendonca CT, Jensen D. Physiological and perceptual responses to incremental exercise testing in healthy men: effect of exercise test modality. *Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme* 2015: 40(11): 1199-1209.
33. Porszasz J, Casaburi R, Somfay A, Woodhouse LJ, Whipp BJ. A treadmill ramp protocol using simultaneous changes in speed and grade. *Med Sci Sports Exerc* 2003: 35(9): 1596-1603.
34. 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): 581-586.

TABLES

Table 1. Baseline characteristics and incremental exercise test variables

	Males N = 1,483 (57%)	**Females** N = 1,098 (43%)
Age	46.6 ± 14.5	54.0 ± 13.7
Height (cm)	180 ± 7	165 ± 6
Body mass (kg)	84.7 ± 11.6	69.4 ± 10.8
Body mass index (kg/m^2)	26.1 ± 3.1	25.4 ± 3.7
IET values at peak exercise		
Power output (W)	241 ± 44	142 ± 28
Power output (%predW_{max})	98 ± 12	99 ± 13
Breathlessness (Borg CR10)	6.7 ± 1.5	6.4 ± 1.5
Exertion (Borg RPE)	17.6 ± 0.9	17.4 ± 0.9
Heart rate (beats/min)	171 ± 17	162 ± 16
Heart rate (% of 220-age)	98.7 ± 7.3	97.9 ± 7.9
Exercise time (minutes)	11.5 ± 2.6	10.8 ± 2.6

Data presented as means ± standard deviation. *Abbreviations:* IET = Incremental Exercise Test; Borg CR10 = Borg Category-ratio scale (0-10); %predW_{max} = W % of predicted W_{max}; Borg RPE = Borg rating of perceived exertion scale (6-20); W = Watt.

SUPPLEMENTAL MATERIAL

Supplement 1.

Supplemental figures and tables.

Supplement 2.

Excel file for calculation of predicted breathlessness intensity ratings, upper limit of normal (ULN) and probabilities for different breathlessness intensity ratings by %predWmax, age, sex (and height in males).

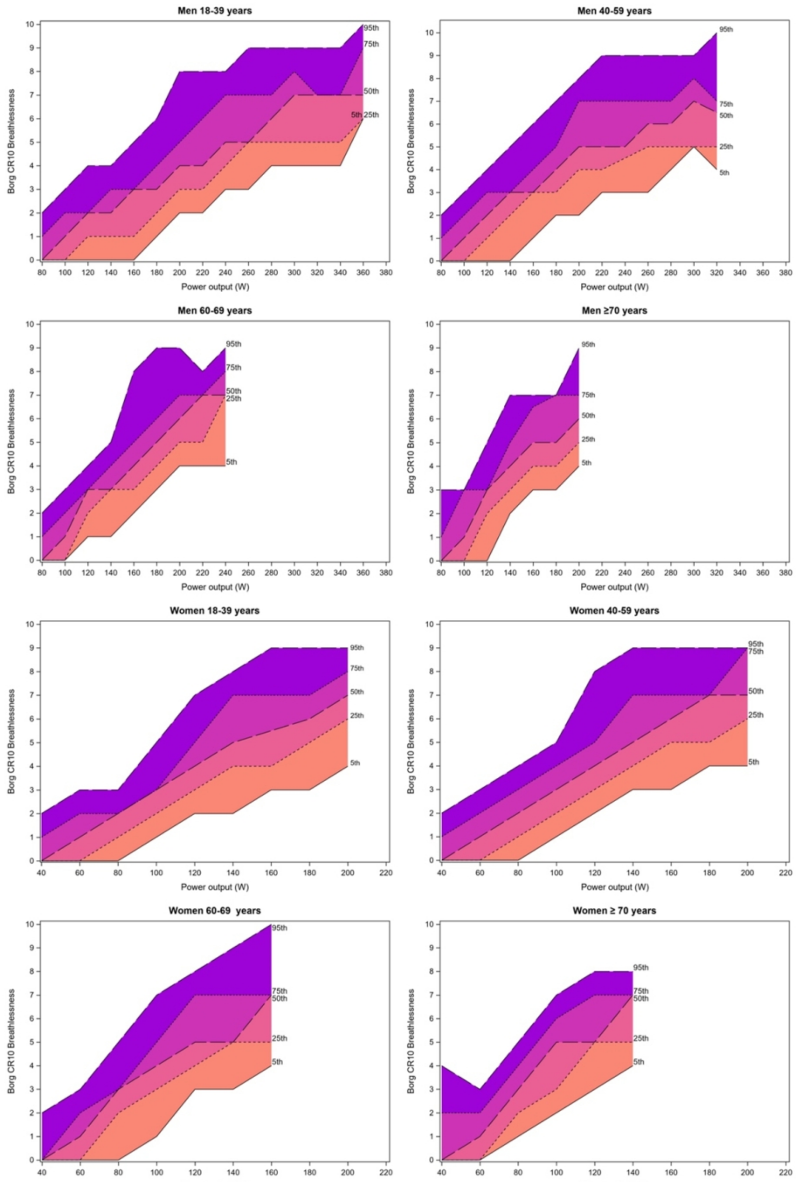


Figure 1. Breathlessness intensity (Borg CR10) during IET related to power output (W), by sex and age. Females: n =1098; males: n = 1483; 18–39 years: n = 657; 40–59 years: n = 1244; 60–69 years: n = 485; ≥70 years: n =195.

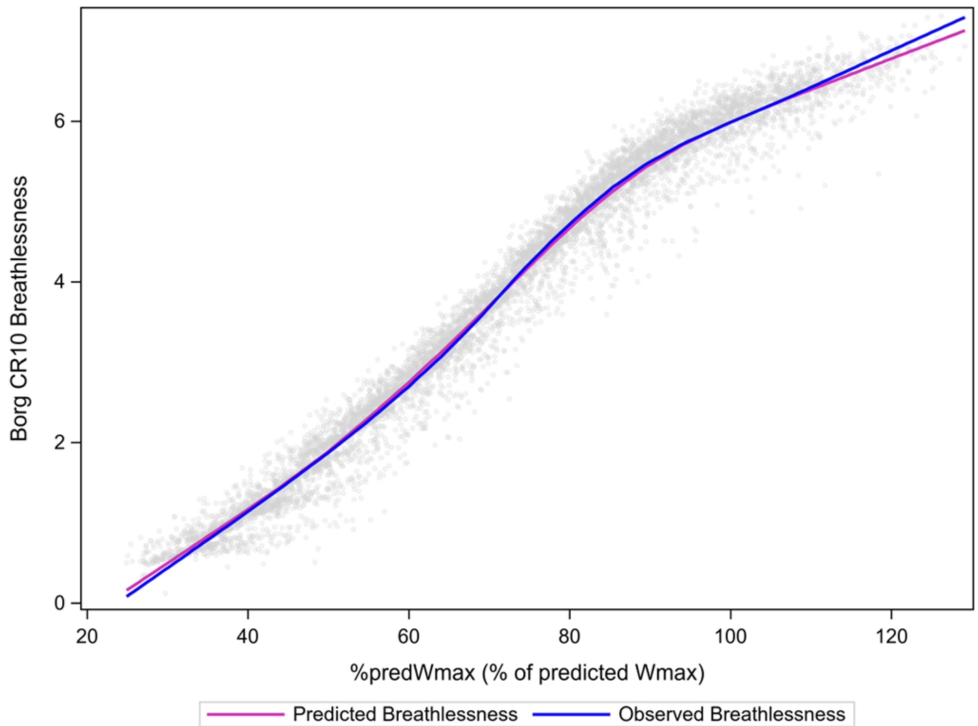


Figure 2 a. Fit of the model in males and females separately. Loess fit of predicted breathlessness intensity and observed breathlessness intensity (Borg CR10) plotted against %predWmax (W % of predicted Wmax) in a) males, and b) females. Grey points are individual predicted values.

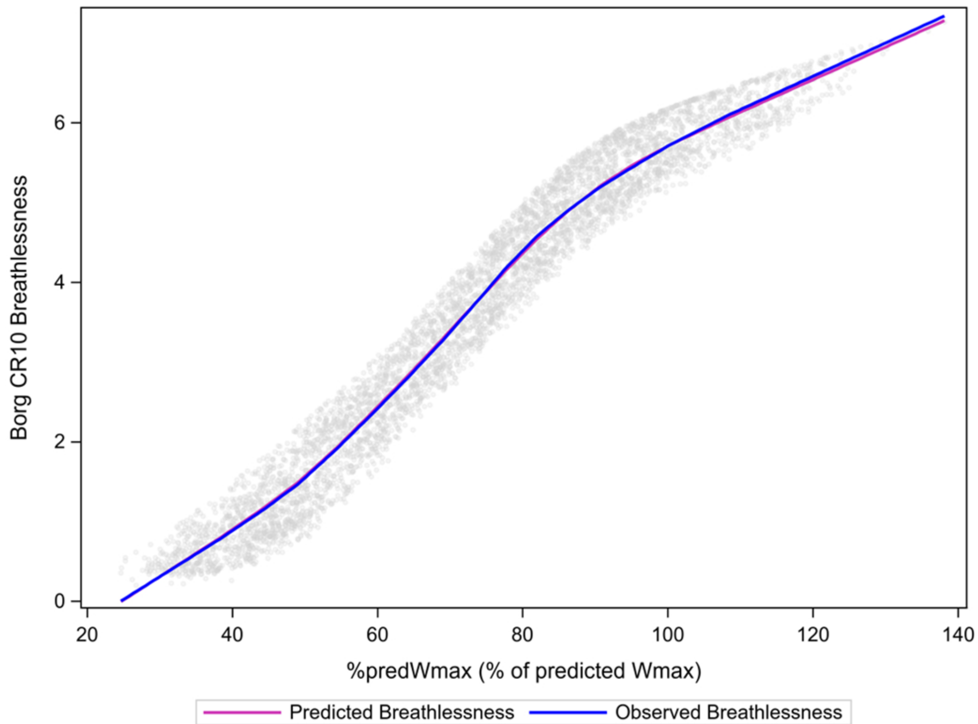


Figure 2 b. Fit of the model in males and females separately. Loess fit of predicted breathlessness intensity and observed breathlessness intensity (Borg CR10) plotted against %predWmax (W % of predicted Wmax) in a) males, and b) females. Grey points are individual predicted values.

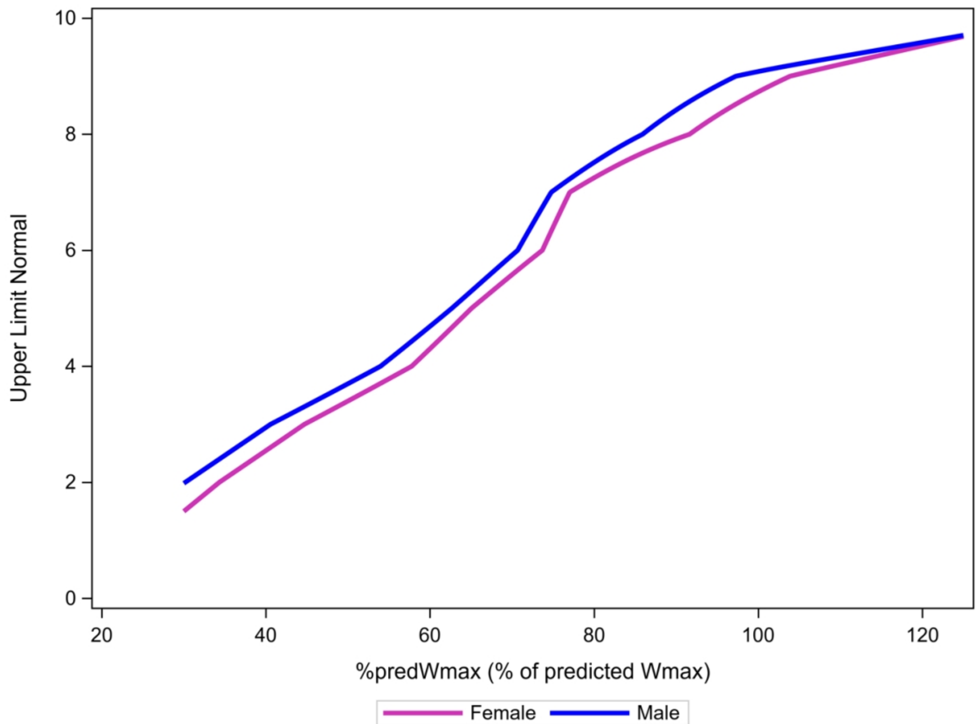


Figure 3. Upper limit of normal (ULN) for breathlessness intensity in relation to power output %predWmax by sex. Breathlessness intensity (Borg CR10 scale) is presented for a 70-year-old, 180 cm tall male and female. The ULN values were calculated using linear interpolation of the linear predictor of the responses closest to below and above a probability of 0.95 using marginal ordinal logistic regression, estimated by generalized estimating equations (GEE).

SUPPLEMENTAL MATERIAL 1

Reference equations for breathlessness during incremental cycle exercise testing

Table A1. Quasi Information Criterion for different models of breathlessness. Restricted cubic splines with knots placed at the 5^{th}, 35^{th} 65^{th} and 95^{th} percentiles are written as f(x), i.e. f(age) denotes the restricted cubic spline of age. Lower values indicate a better fitting model (bold).

Variables in model	**QIC Men**	**QIC Women**
f(%predW_{max})	25165	16911
f(%predW_{max})+ f(age)	25011	16626
f(%predW_{max}) + f(age)+f(height)	24963	16643
f(%predW_{max}) + f(age) + height	24957	16632
f(%predW_{max}) + f(age) + height + weight	24962	16632
f(%predW_{max}) + f(age) + height + f(weight)	24953	16635
f(%predW_{max}) + f(age) + height + f(weight) + %predW_{max}*age	**24935**	16618
f(%predW_{max}) + f(age) + height + %predW_{max}*age	**24939**	**16614**
f(%predW_{max}) + f(age) + %predW_{max}*age	24993	**16608**

Table A2. Parameter estimates.

	Men	Women
Parameter	**Estimate**	**Estimate**
Intercept1	7.105994524	2.823125
Intercept2	8.466031224	4.014411
Intercept3	9.825848726	5.44524
Intercept4	11.61099378	7.29478
Intercept5	12.82448793	8.37268
Intercept6	13.9804337	9.643334
Intercept7	14.54650183	10.10118
Intercept8	15.83592933	11.62911
Intercept9	16.66525731	12.36298
Intercept10	18.4985691	13.74351
%predSpl_1	-0.109693986	-0.10982
%predSpl_2	-0.000438626	-0.00041
%predSpl_3	0.002369292	0.002344
AgeSpl_1	0.046027134	0.039811
AgeSpl_2	-0.001408948	0.001028
AgeSpl_3	0.004509744	-0.00237
Height	-0.025396237	
lastW_of_pred*Age	-0.000302552	-0.0004

Table A3. Spline knots

Variable	**Men**	**Women**
Age	21.92, 40.20, 52.33, 69.03	27.57, 48.37, 59.21, 73.48
%predW_{max}	36.54, 62.51, 83.25, 108.84	36.75, 62.28, 83.70, 110.73

See Hastie, et al 2009 for calculation of splines.

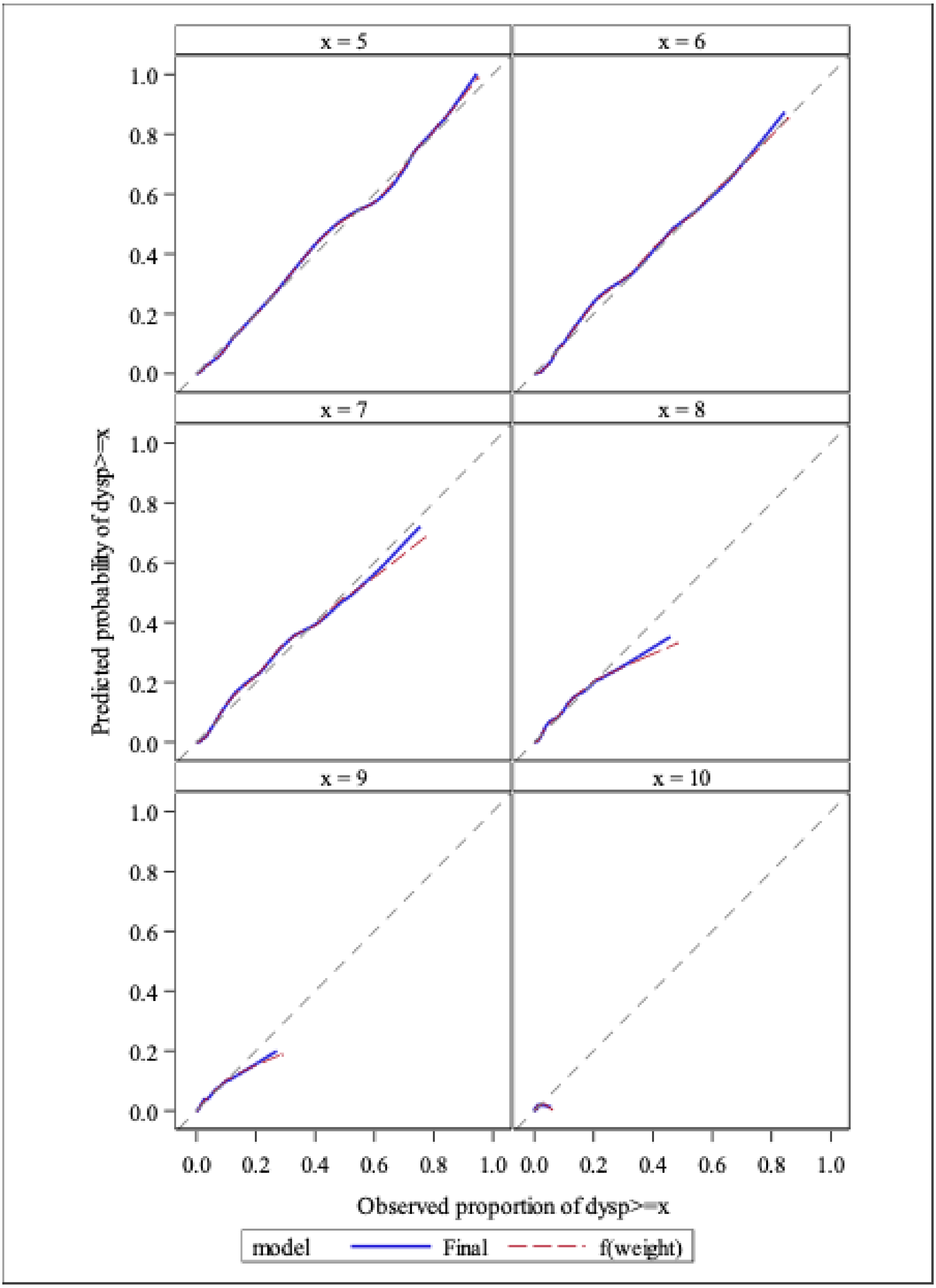
Figure A1. Comparison of the final model and the model with the lowest QIC for men. Calibration was indistinguishable between models for x=1-4 and is not shown.

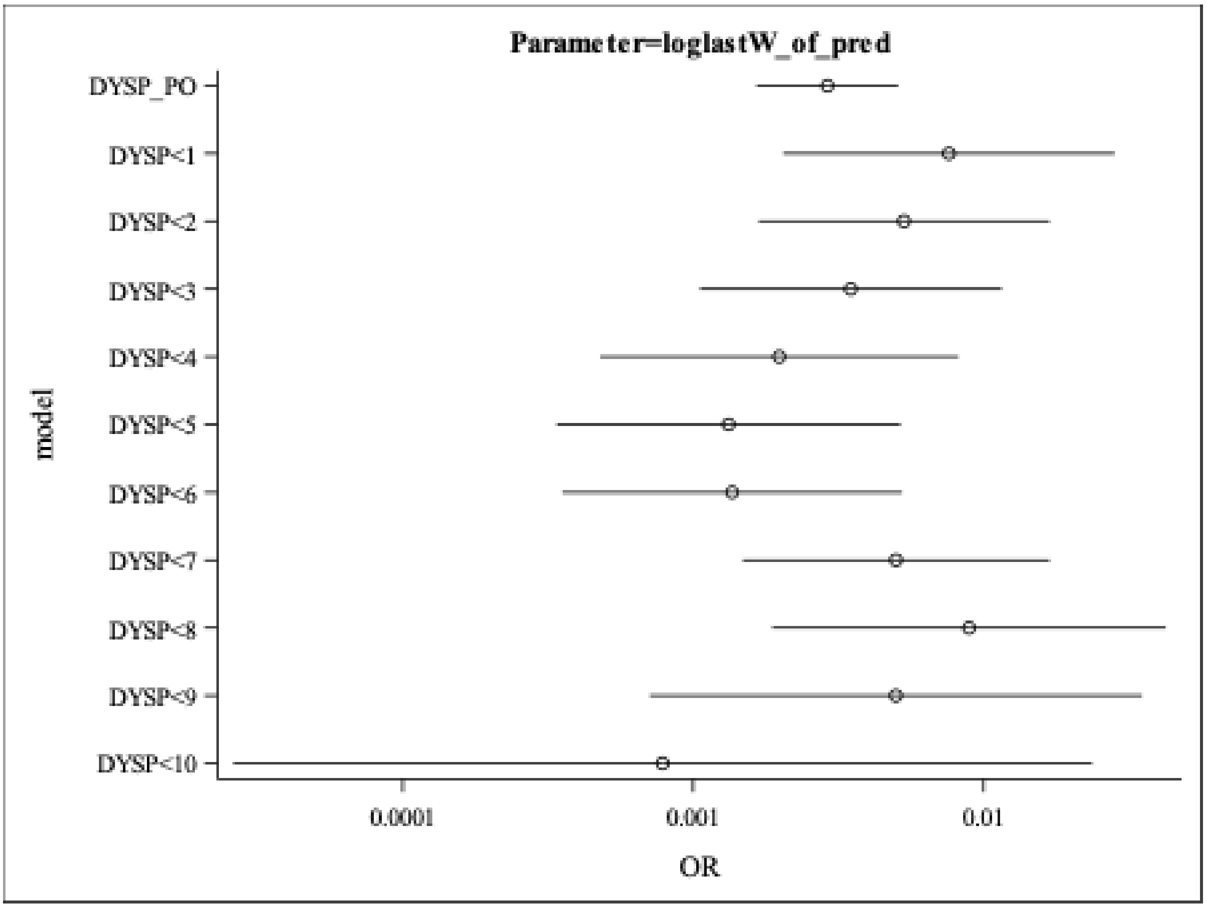
Figure A2. Odds ratios for log(%predWmax) for men. DYSP_PO is from the proportional odds model, the remaining shows binary logistic regression with different cut points. The proportional odds assumption was evaluated by calculating the odds ratio using binary logistic regression for each possible cut-point (i.e. setting the outcome to dysp<i for i=1,..10). For ease of interpretation, we modelled age and log(%predWmax) as linear rather than splines. Figure A2 shows the odds ratios for log(%predWmax) for men. Plots for other parameters looked slightly better, with the exception of the confidence interval for Dysp<10 being narrower in some cases, indicating the proportional odds assumption may be questionable for the highest value of breathlessness. We do not consider this a major problem, as we have few observations of Dysp=10.

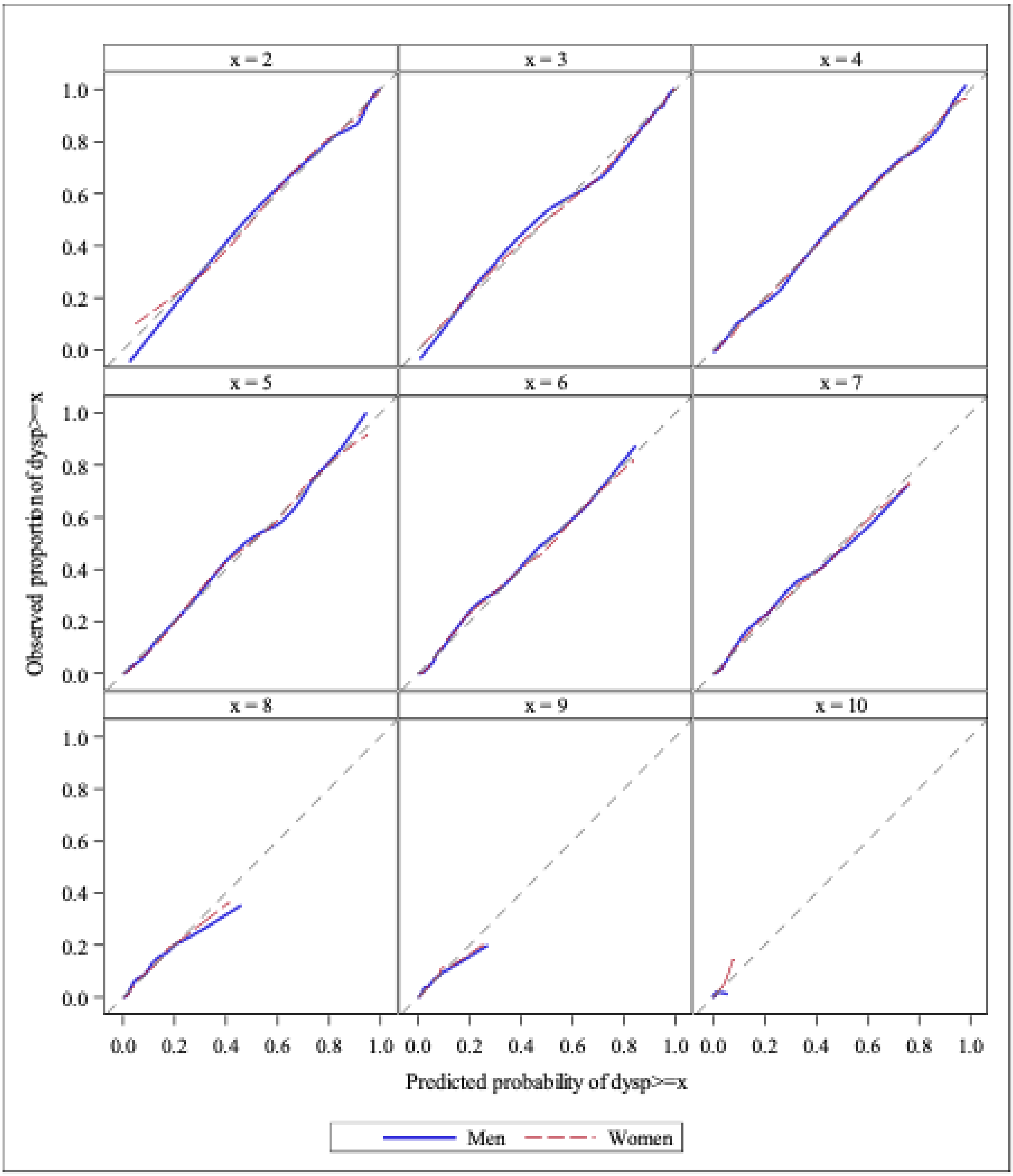
Figure A3. Calibration plots for the final models for men and women separately. Calibration plots were created by plotting a loess smoother of the observed Dysp≥x vs the predicted probability of Dysp≥x. Calibration suffers for larger probabilities of larger x, but the number of observations are relatively small, e.g. only 319 men and 65 women had a predicted probability larger than 0.25 of having Dysp≥8.

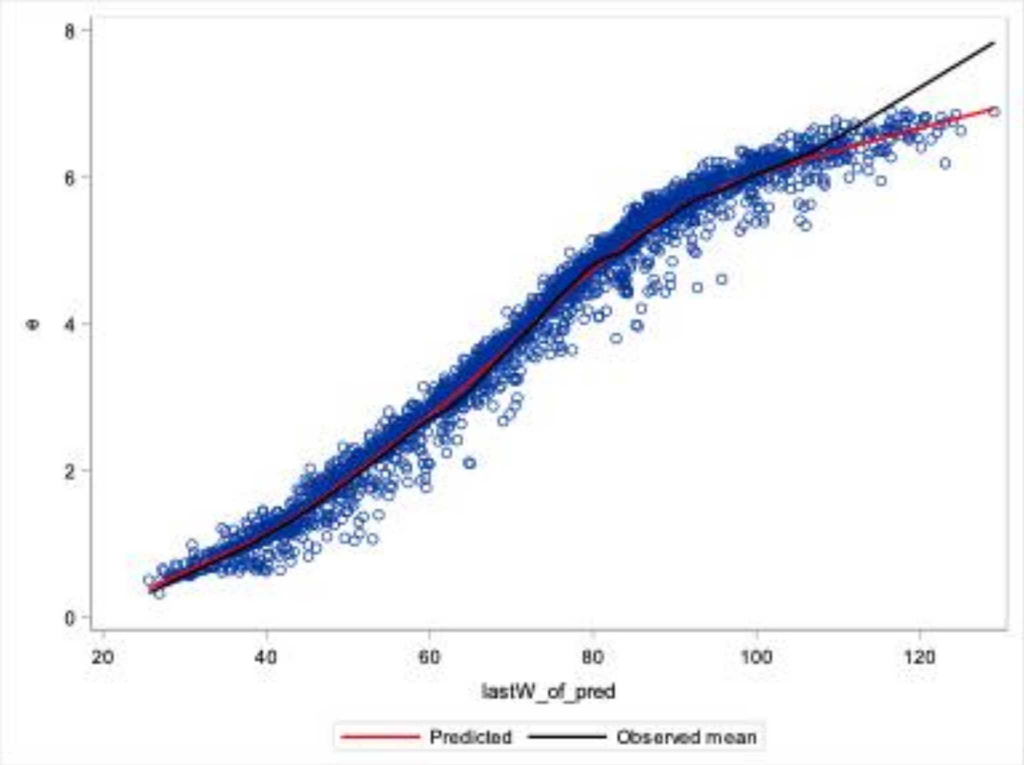
Figure A4 a Males. Fit of the reference equations modelled on the random 70% subgroup, compared with the 30% validation sample in males and females separately.
Figure A4 b Females. Fit of the reference equations modelled on the random 70% subgroup, compared with the 30% validation sample in males and females separately.

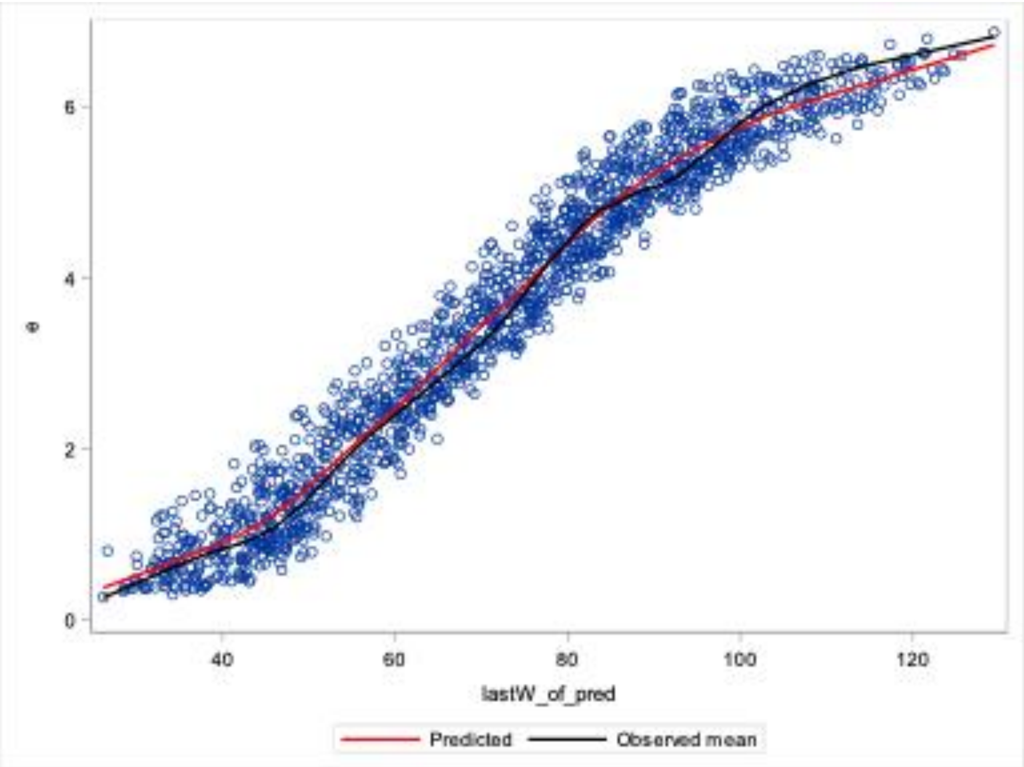
Figure A5 a Males. Fit of the model in males and females separately, by age group.
Figure A5 b Females. Fit of the model in males and females separately, by age group.

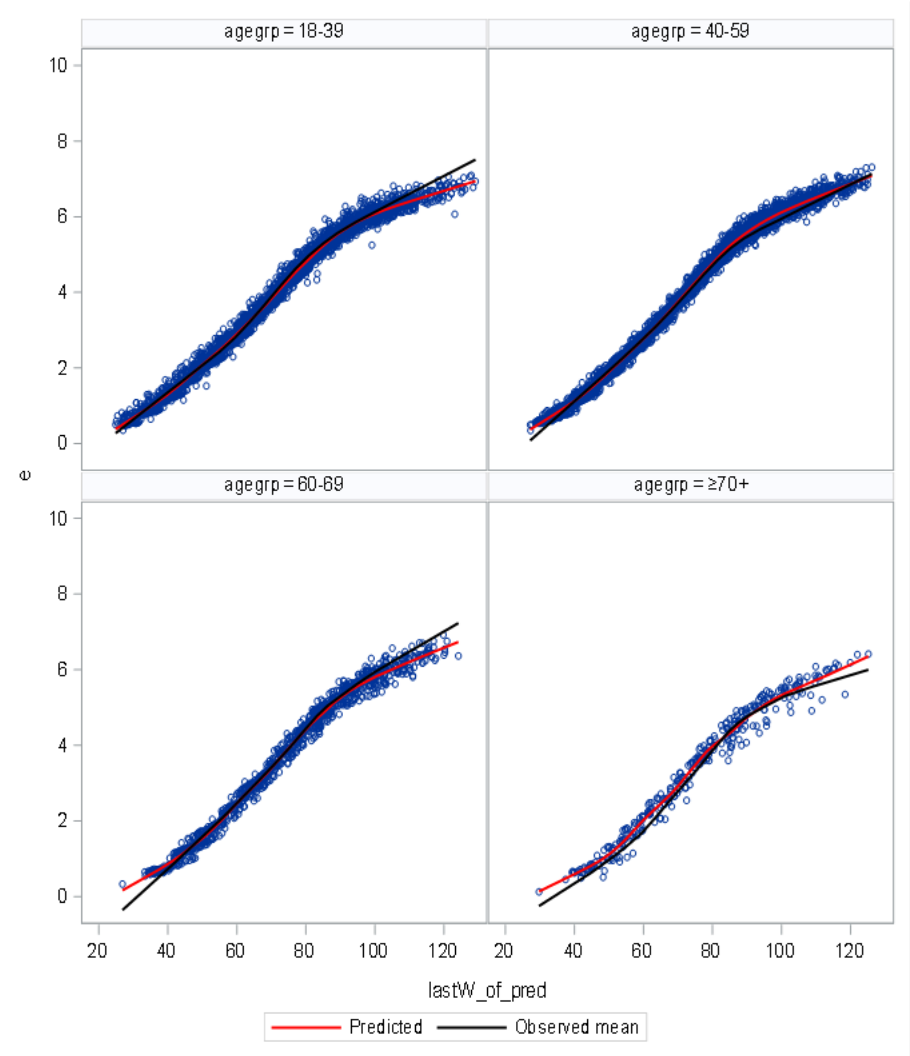

x = 5
x = 6
x = 7
x = 8
x = 9
x = 10
1.0
0.8
0.6
0.4
0.2
0.0
Predicted probability of dysp>=x
0.0 0.2 0.4 0.6 0.8 1.0
Observed proportion of dysp>=x
model
Final
f(weight)


Parameter=loglastW_of_pred
DYSP_PO
DYSP<1
DYSP<2
DYSP<3
DYSP<4
DYSP<5
DYSP<6
DYSP<7
DYSP<8
DYSP<9
DYSP<10
model
0.0001
0.001
0.01
OR

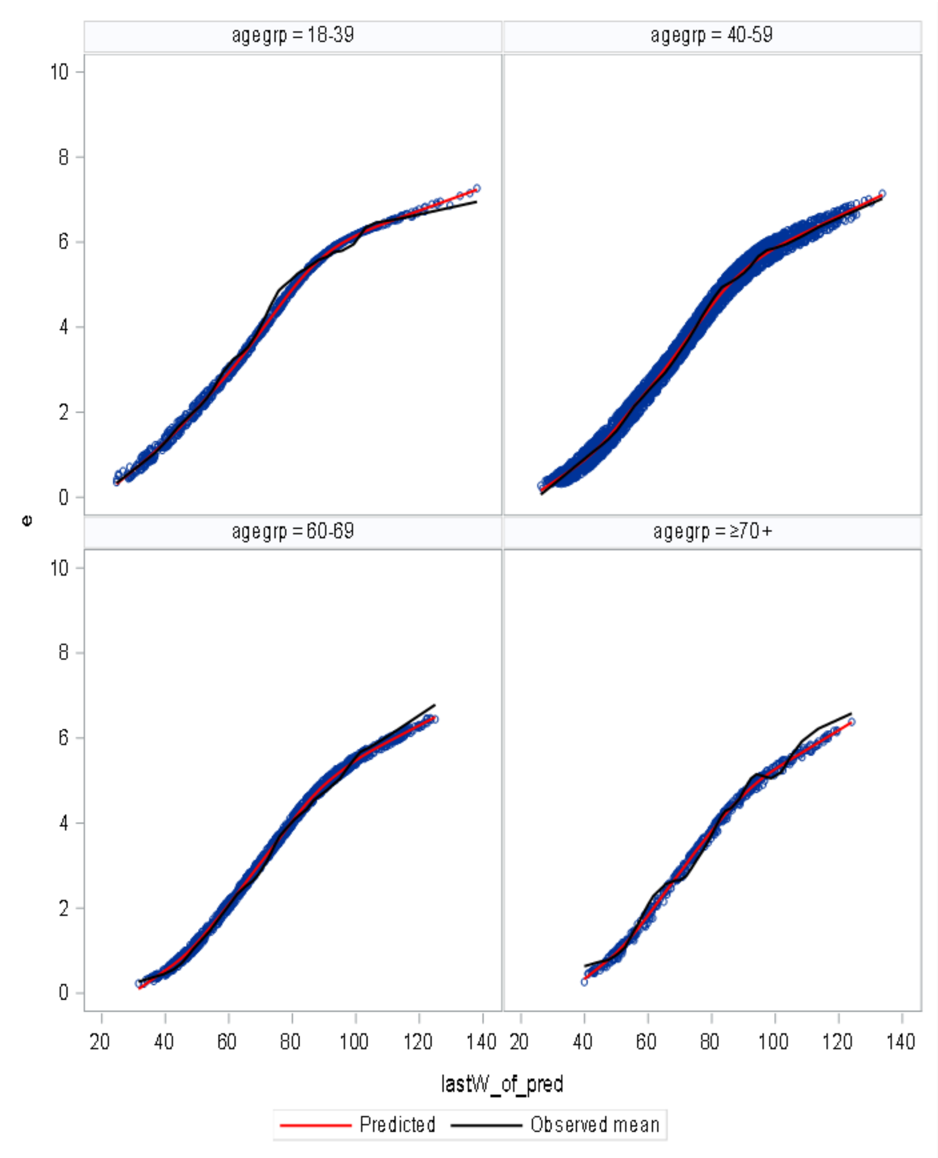

x = 2
x = 3
x = 4
x = 5
x = 6
x = 7
x = 8
x = 9
x = 10
0.0
0.2
0.4
0.6
0.8
1.0
Observed proportion of dysp>=x
Predicted probability of dysp>=x
Men
Women


8
6
4
2
0
e
20
40
60
80
100
120
lastW_of_pred
Predicted
Observed mean


6
4
2
0
e
40
60
80
100
120
lastW_of_pred
Predicted
Observed mean



agegrp = 18-39
agegrp = 40-59
agegrp = 60-69
agegrp = ≥70+
e
lastW_of_pred
Predicted
Observed mean

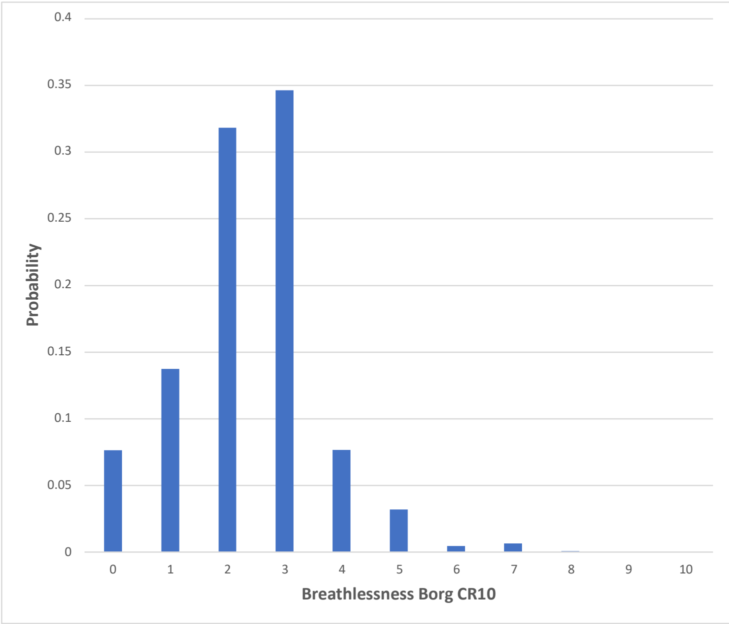

agegrp = 18-39
agegrp = 40-59
agegrp = 60-69
agegrp = ≥70+
e
lastW_of_pred
Predicted
Observed mean

Calculation of reference values for breathlessness intensity (Borg CR10) during incremental exercise testing (1)

Females

To be filled in before/after exercise test	
Age (years)	40
W as percentage of predicted	55
Observed breathlessness	3

P(breathlessness>=observed breathlessness	0.47
Upper Limit Normal breathlessness	4.90
Expected breathlessness	2.37



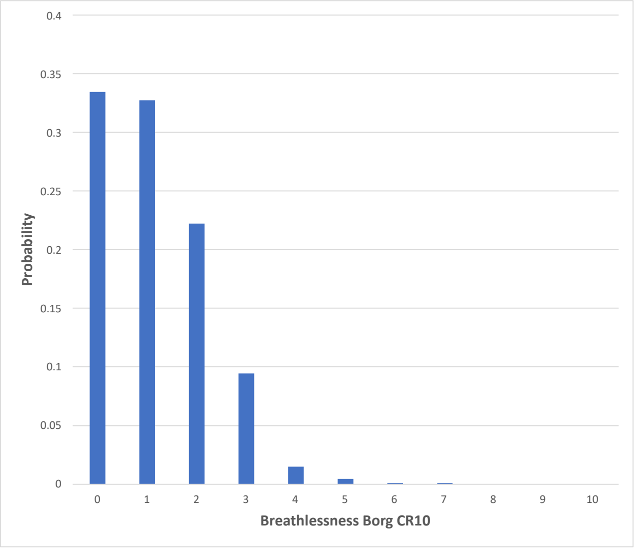
1. Elmberg V, Schiöler L, Lindow T, Hedman K, Malinovschi A, Lewthwaite H, Jensen D, Brudin L, Ekström M. Reference equations for breathlessness during incremental cycle exercise testing. ERJ Open Research 2023.

Calculation of reference values for breathlessness intensity (Borg CR10) during incremental exercise testing (1)

Males

To be filled in before/after exercise test	
Age (years)	38
Height (cm)	180
W as percentage of predicted	40
Observed breathlessness	5

P(breathlessness>=observed breathlessne	0.006
Upper Limit Normal breathlessness	3.51
Expected breathlessness	1.151462457



1. Elmberg V, Schiöler L, Lindow T, Hedman K, Malinovschi A, Lewthwaite H, Jensen D, Brudin L, Ekström M. Reference equations for breathlessness during incremental cycle exercise testing. ERJ Open Research 2023.

Male (1 if male, 0 if Female)	Age	%Pred	Height	Breathlessness	Pdysp=0	Pdysp=1	Pdysp=2	Pdysp=3	Pdysp=4	Pdysp=5	Pdysp=6	Pdysp=7	Pdysp=8	Pdysp=9	Pdysp=10	ULN	Expected Value	Pr(Higher Dysp)
0	55	89	167	8	0.002122743	0.004830043	0.021495495	0.12847959	0.1966459	0.307321658	0.09404424	0.179259805	0.033094176	0.024276078	0.008430233	8.397039908	5.143076365	0.032706311
1	56	105	185	8	0.000458631	0.001325979	0.00513157	0.032939969	0.0827141	0.184816156	0.131348369	0.300718517	0.127289008	0.109266104	0.023991566	9.584724254	6.500088818	0.13325767
1	57	62	179	8	0.033730887	0.085998874	0.226608462	0.413167363	0.154493	0.057236485	0.01222794	0.011927012	0.002593222	0.001693783	0.000322996	5.502596255	2.87894873	0.00201678
1	58	120	191	8	0.00022221	0.000643039	0.002496891	0.016348971	0.0436664	0.113561509	0.097710939	0.304252763	0.180169282	0.192632201	0.048295763	9.980106219	7.171193676	0.240927964
1	59	102	174	8	0.000748546	0.002161732	0.008331925	0.052226888	0.1222401	0.234432996	0.140527013	0.261820276	0.091443707	0.071233616	0.014833159	9.317351867	6.033519591	0.086066775
0	60	98	159	8	0.001427876	0.00325638	0.014618348	0.091903176	0.1576209	0.298283487	0.10723777	0.230805685	0.046953828	0.035402465	0.01249013	8.938302563	5.50832398	0.047892595
0	61	72	162	8	0.018716183	0.040351473	0.148877083	0.417373044	0.2053097	0.115243428	0.019189841	0.027144695	0.004037552	0.002809624	0.000947332	6.182820603	3.359893157	0.003756956
0	62	66	157	8	0.04522524	0.089648698	0.259800372	0.41095158	0.1184973	0.053352723	0.008155453	0.011215387	0.001637055	0.001134521	0.000381673	5.349980973	2.737421276	0.001516193
0	63	54	170	8	0.217319058	0.260177876	0.315117274	0.167853265	0.025719	0.009897892	0.001434867	0.001941446	0.000280359	0.000193849	6.51321E-05	3.867071111	1.573175761	0.000258981
1	88	74	180	9	0.035044571	0.088918503	0.231386308	0.411309051	0.1503998	0.055261853	0.011774246	0.011473719	0.00249326	0.001628221	0.00031047	5.468366948	2.852193011	0.00031047
0	87	124		5	0.000666361	0.001523496	0.006905106	0.046034824	0.091228	0.232875476	0.112028239	0.325262375	0.086110374	0.070958994	0.026406756	9.519797209	6.22331511	0.620766737

How to use the reference values

1.	Choose Female or Male (sheet)
2.	Input the required information in the green box: Age, years Height, centimeters (if male) W as percentage of predicted peak power output (point to evaluate breathlessness intensity at), %Pred Observed breathlessness intensity (at the specific %Pred), Borg CR10
3.	Upper limit of normal (ULN) and most likely breathlessness intensity for the specific %Pred is shown in the box below If observed breathlessness >ULN, breathlessness is considered abnormally increased Probabilities for different breathlessness intensity ratings is shown in the graph to the right
Alt:	The sheet "Multiple measurements" allows pasting of data (into the blue columns) from multiple individuals when working with larger volumes of data
	Most expected breathlessness intensity and upper limit of normal is shown in the green columns. It is important to not input data directly into the green columns as that will break the functionality