# THREE-MINUTE CONSTANT RATE STAIR STEPPING (3-MIN CRSST) AND CONSTANT SPEED SHUTTLE TESTS (3-MIN CSST) TO ASSESS EXERTIONAL BREATHLESSNESS IN COPD: DEVELOPMENT AND PROSPECTIVE VALIDATION OF AN EQUATION TO SELECT EXERCISE INTENSITY FOR USE IN CLINICAL CARE AND RESEARCH

by

EMILY KOCH

B.Sc. Kinesiology, Brock University, 2018

A THESIS SUBMITTED TO McGILL UNIVERSITY IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS OF THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF EDUCATION

(Department of Kinesiology and Physical Education)

McGILL UNIVERSITY

MONTREAL, QUEBEC, CANADA

August 2020

## ACKNOWLEDGEMENTS

First off, I would like to offer my sincere gratitude to my supervisor Dr. Dennis Jensen for his constant guidance, enthusiasm and support. Thank you for not only making my thesis possible, but also for all the amazing opportunities you have provided over the past two years.

I would like to thank my fellow CERPL lab members for supporting me throughout this process. In my two short years as part of this incredible lab I have grown both personally and professionally and made lifelong connections. A special thanks Hayley Lewthwaite, Lauren Tracey and Leticia Fernandes Belo for their help in the data collection process.

I would also like to thank my thesis committee members, Dr. Jean Bourbeau and Dr. Benjamin Smith for their interest and valuable feedback on my project. Additional thanks to Dr. Bourbeau for helping facilitate the integration of my study at the RI-MUHC with the pulmonary rehabilitation program. I gratefully acknowledge the contributions of Pei Zhi Li who was instrumental in helping develop the equations for this study.

Finally, thank you to my friends and family for providing me with endless love and support throughout this journey. I am so grateful to have you all by my side, I would not have been able to do it without you.

**Table of Contents**

## LIST OF TABLES

## LIST OF FIGURES

## PREFACE AND CONTRIBUTION OF AUTHORS

Koch, E was the primary author of this thesis and played the principle role in data collection, data management, data analysis and data interpretation as well as the preparation of this document.

Lewthwaite, H played a role in the collection, analysis and interpretation of data and preparation of this thesis.

Belo, L played role in the data collection.

Li, P provided statistical support for the data analysis.

Finally, Jensen, D contributed to the conception, design and financial support of this study, in addition to the analysis and interpretation of data and the preparation of this thesis.

# ABSTRACT

**Introduction:** Breathlessness on exertion is the most prominent and debilitating symptom of chronic obstructive pulmonary disease (COPD). The 3-minute constant rate stair-stepping test (3-min CRSST) and the 3-minute constant speed shuttle test (3-min CSST) have been developed and prospectively validated to assess exertional breathlessness for people with COPD. These tests, for the first time, enable assessment of breathlessness during a standardized and individualized level of exertion, which mimic problematic activities of daily life. However, the current method used to determine an appropriate cadence (exercise intensity) for these tests in a given individual is time consuming and impractical for clinical use. **Objective:** The objective of this research was twofold: 1) to develop prediction equations to identify the "optimal" step rate and shuttle speed for an individual with COPD based on his/her unique personal and clinical characteristics; and 2) prospectively validate the prediction equations in adults with COPD. Based on the results of the prospective validation, we also aimed to determine which test may be better suited for use in the clinical care setting of COPD. **Methods:** Prediction equations were developed based on data available from people with COPD who completed a 3-min CRSST (n= 90 [74% male]; age= 66.6 ± 6.8 years; $FEV_1$%pred= 54.0 ± 20.6%) and/or a 3-min CSST (n= 112 [70% male]; age= 65.8 ± 6.9 years; $FEV_1$%pred= 50.4 ± 21.4%) as part of previous studies. Mixed-effects linear models estimated exercise intensity (step rate or shuttle speed) based on Borg CR10 scale breathlessness intensity rating. The prediction equations were validated prospectively in 18 people with COPD ($FEV_1$%pred= 45.9±17.2; age= 69.3±7.4 years). Participant's age, height and body mass were recorded, and each participant completed health status and symptom burden questionnaires as well as a 3-min CRSST and 3-min CSST at the stepping rate and shuttle speed estimated from the newly developed equations. The prediction equation was characterized as "successful" when the participant completed all 3-minutes of the test at the identified exercise intensity (step rate or shuttle speed) with a breathlessness intensity rating of $\geq 3$ Borg CR10 scale units. **Results**: Step rate (steps/min), sex (1=male, 0=female), age (years), body mass (kg), modified Medical Research Council (mMRC) dyspnea score and Global Initiative for Obstructive Lung Disease stage (GOLD I-IV) were significant predictors of the breathlessness intensity response to the 3-min CRSST: Breathlessness intensity (Borg CR10 scale units) = -13.918 + (0.653 * step rate) – (1.078 * sex) + (0.155 * age) – (0.006 * age * step rate) + (0.028 * body mass) + (0.635 * mMRC) + GOLD stage. Shuttle speed (km/hr), sex (1=male, 0=female), age (years), body mass index (BMI, $kg/m^2$), forced expiratory volume in 1-sec ($FEV_1$, L) and mMRC dyspnea score were significant predictors of the breathlessness intensity response to the 3-min CSST: Breathlessness intensity (Borg CR10 scale units) = -2.530 – (0.619 * shuttle speed) + (0.206 * shuttle speed$^2$) - (0.366 * sex) + (0.031 * age) + (0.094 * BMI) – (0.667 * $FEV_1$) + (0.748 * mMRC). For the prospective validation, 61% of participants finished all 3-minutes of the step or shuttle test at the exercise intensity predicted using the newly developed equations with at least moderately intense breathlessness. The physiological and perceptual responses to the step and shuttle test were comparable. **Conclusion:** The equations developed in this study successfully identified the step rate or shuttle speed in the majority of individuals with COPD and may represent a unique opportunity for healthcare providers and researchers to incorporate the 3-min CRSST and 3-min CSST into their practice as an individualized and standardized tool to track exertional breathlessness in people with COPD.

**RÉSUMÉ**

**Introduction:** L'essoufflement à l'effort est le symptôme le plus important et le plus débilitant de la maladie pulmonaire obstructive chronique (MPOC). Le test de marche d'escalier (CRSST) et marche à vitesse constante de 3 minutes (CSST) ont récemment été développés et validés de manière prospective pour évaluer l'essoufflement à l'effort chez les personnes atteintes de MPOC. Ces tests, pour la première fois, permettent d'évaluer l'essoufflement pendant un niveau d'effort standardisé et individualisé, qui imite les activités problématiques de la vie quotidienne. Cependant, la méthode existante utilisée pour déterminer une cadence appropriée (intensité de l'exercice) pour ces tests chez un individu donné est longue n'est pas pratique pour une utilisation clinique. **Objectif:** L'objectif de cette recherche était double: 1) développer des équations permettant d'identifier la cadence et la vitesse de navette "optimales" pour un individu avec le MPOC en fonction de ses caractéristiques personnelles et cliniques uniques; et 2) valider de manière prospective les équations chez les adultes avec le MPOC. Sur la base des résultats de la validation prospective, nous avons également cherché à déterminer quel test est le mieux adapté pour être utilisé dans le cadre des soins cliniques de la MPOC. **Méthodes**: Les équations de prédiction ont été élaborées sur la base des données disponibles de personnes avec le MPOC ayant suivi un CRSST (n= 90 [74 % d'hommes]; âge= 66,6 ± 6,8 ans; VEMSprévu= 54,0 ± 20,6 %) et/ou une CSST (n= 112 [70 % d'hommes]; âge= 65,8 ± 6,9 ans; VEMSprévu= 50,4 ± 21,4 %) dans les études précédentes. Des modèles linéaires à effets mixtes ont permis d'estimer l'intensité de l'exercice en fonction de la cote d'intensité de l'essoufflement sur la gamme CR10 de Borg. Les équations ont été validées de manière prospective dans 18 personnes avec le MPOC (VEMSprévu= 45,9 ± 17,2; âge= 69,3 ± 7,4 ans). L'âge, la taille et la masse corporelle des participants ont été enregistrés, et chaque participant a rempli des questionnaires sur l'état de santé et la charge des symptômes, ainsi qu'un CRSST et une CSST, à la cadence et à la vitesse de navette estimées à partir des équations nouvellement développées. L'équation de prédiction a été qualifiée de "réussie" lorsque le participant a effectué les 3 minutes du test à l'intensité d'exercice identifiée (rythme par paliers ou vitesse de navette) avec une note d'intensité d'essoufflement de ≥3 unités de la gamme Borg CR10. **Résultats**: Le taux de pas (pas/min), le sexe (1=homme, 0=femme), l'âge (années), la masse corporelle (kg), le score modifié de la dyspnée du Medical Research Council (mMRC) et le stade (I-IV) de l'Initiative Mondiale pour la bronchopneumopathie chronique obstructive (GOLD) étaient des prédicteurs significatifs de la réponse de l'intensité de l'essoufflement au CRSST: Intensité de l'essoufflement (unités de la gamme Borg CR10) = -13. 918 + (0,653 * taux d'échelon) - (1,078 * sexe) + (0,155 * âge) - (0,006 * âge * taux d'échelon) + (0,028 * masse corporelle) + (0,635 * mMRC) + stade GOLD. La vitesse de la navette (km/h), le sexe (1=homme, 0=femme), l'âge (années), l'indice de masse corporelle (IMC, kg/m2), le volume expiratoire forcé en 1 seconde (VEMS, L) et le marquer de dyspnée mMRC étaient des prédicteurs significatifs de la réponse de l'intensité de l'essoufflement à la CSST: Intensité de l'essoufflement (unités de la gamme Borg CR10) = -2. 530 - (0,619 * vitesse de la navette) + (0,206 * vitesse de la navette$^2$) - (0,366 * sexe) + (0,031 * âge) + (0,094 * IMC) - (0,667 * $VEM_1$) + (0,748 * mMRC). Pour la validation prospective, 61% des participants ont terminé les 3 minutes du test de l'étape ou de la navette à l'intensité d'exercice prévue à l'aide des équations nouvellement développées avec un essoufflement au moins modérément intense. Les réponses physiologiques et perceptuelles au test de pas et de la navette étaient comparables. **Conclusion:** Les équations développées dans cette étude ont permis d'identifier avec succès le taux de pas ou la vitesse de la navette chez la majorité des personnes avec le MPOC et peuvent représenter une opportunité unique pour les pourvoyeurs de soins de santé et les chercheurs d'intégrer le CRSST et le CSST dans leur pratique en tant qu'outil individualisé et standardisé pour suivre l'essoufflement à l'effort chez les personnes impacté par le MPOC.

# CHAPTER 1. Review of Literature

## 1.0 Chronic Obstructive Pulmonary Disease (COPD): Causes, Prevalence and Burden

Chronic obstructive pulmonary disease (COPD) is characterized by chronic and progressive expiratory airflow limitation (EFL) and deterioration of lung function. Airflow limitation is caused by narrowing of the bronchioles due to chronic inflammation, mucus hypersecretion, decreased elastic recoil of the lungs, and/or destruction of the alveoli/gas exchange surface area [1]. The primary cause of COPD is inhalation of noxious gases, typically from cigarette smoking; however, occupational exposures, genetic predisposition and 'childhood disadvantage factors' (for example, gestational influences) are also causal factors [1]. While there are various management strategies available to people once diagnosed with COPD, there is no known cure for this disease.

COPD is a prevalent chronic health condition; the estimated global prevalence of COPD in people aged ≥30 years is ~11.7% [2]. In Canada, 16.7% of adults aged ≥40 years recruited from the general population (n=3,042) met spirometric criteria for at least mild COPD [3]. The increasing prevalence of COPD in Canada and worldwide can be partly attributed to the aging population and increased exposure to risk factors (e.g., ambient air pollution) [4]. COPD causes approximately three million deaths per year, with this statistic expected to increase to 5.4 million by the year 2060 [4].

COPD places a heavy burden on those diagnosed; symptoms associated with the disease – particularly breathlessness (respiratory discomfort) – impair activities of daily living (ADL). As the disease progresses, people with COPD become increasingly breathless when performing benign daily tasks, such as dressing and undressing. While disease-modifying interventions improve lung function (e.g., inhaled bronchodilators and anti-inflammatory agents), many patients suffer from chronic and disabling breathlessness, which contributes to impaired health status and quality of life (QoL) [5]. Furthermore, COPD is a major financial burden on society; for example, the average annual cost per patient for COPD-related symptoms in Canada was estimated to be $4,147 [6]. Indirect costs associated with COPD (such as the presence of comorbidities) further augment the financial burden this disease places on society. These numbers are amplified in the United States where the direct and indirect health care costs of COPD are estimated to be $29.5 and $20.4 billion per year, respectively [7]. With the increasing prevalence of COPD and its burden on patients and the health care system, the only way to minimize the impact is to ensure COPD is properly diagnosed and that both the disease and its symptomatic manifestations are treated early and optimally.

### 1.1 Symptom Burden in COPD

Breathlessness (or, dyspnea) on exertion is the most prevalent and disabling symptomatic manifestation of COPD. The American Thoracic Society defines dyspnea as *"the subjective experience of breathing discomfort that consists of qualitatively distinct sensations that vary in intensity"* [8]. A cohort study of 49,438 people with COPD managed in primary care by Mullerova et al. [9] found that 82% of respondents experienced at least mild dyspnea (2 on the Medical Research Council [MRC] dyspnea scale or 1 on the modified MRC [mMRC] scale), while >40% of respondents reported moderate to severe dyspnea (MRC ≥3scale or mMRC scale ≥2) (**Fig. 1.1**). According to the Global Initiative for Obstructive Lung Disease (GOLD) [4], a mMRC dyspnea score of ≥2 is indicative of chronic and disabling breathlessness, meaning that people become breathless during simple/routine physical tasks such as getting dressed. In another cohort study of 1,689 people with COPD, 54% experienced disabling breathlessness (mMRC ≥2), with the majority (65%) of these people having GOLD III-IV (severe-to-very severe) COPD [10]. A longitudinal cohort study by Sundh and Ekstrom [10] reported that 74% of 1,689 people with COPD had chronic breathlessness syndrome, which is defined as breathlessness that persists despite optimal treatment of the underlying pathophysiology [5]. A study of 820 people with COPD by Annegarn et al. [11] identified walking and stair climbing to be the most problematic ADL. This study also showed that individuals in GOLD IV (very severe) COPD reported more difficulty during self-care tasks compared to those in GOLD I-III (mild-to-severe) COPD.

**Modified Medical Research Council Dyspnea Scale (mMRC)**

| Grade | |
|---|---|
| 0 | Not troubled with breathlessness except with strenuous exercise. |
| 1 | Troubled by shortness of breath when hurrying on the level or walking up a slight hill. |
| 2 | Walks slower than people of the same age on the level because of breathlessness or has to stop for breath when walking at own pace on the level. |
| 3 | Stops for breath after walking about 100 yards or after a few minutes on the level. |
| 4 | Too breathless to leave the house or breathless when dressing or undressing. |

**Note: This is a modified MRC scale that uses the same descriptions as the original MRC scale. These descriptions are numbered from 1 to 5. The modified MRC (0-4) scale is used to calculate the BODE index.**

**Figure 1.1:** Modified Medical Research Council (mMRC) dyspnea scale questionnaire [12].

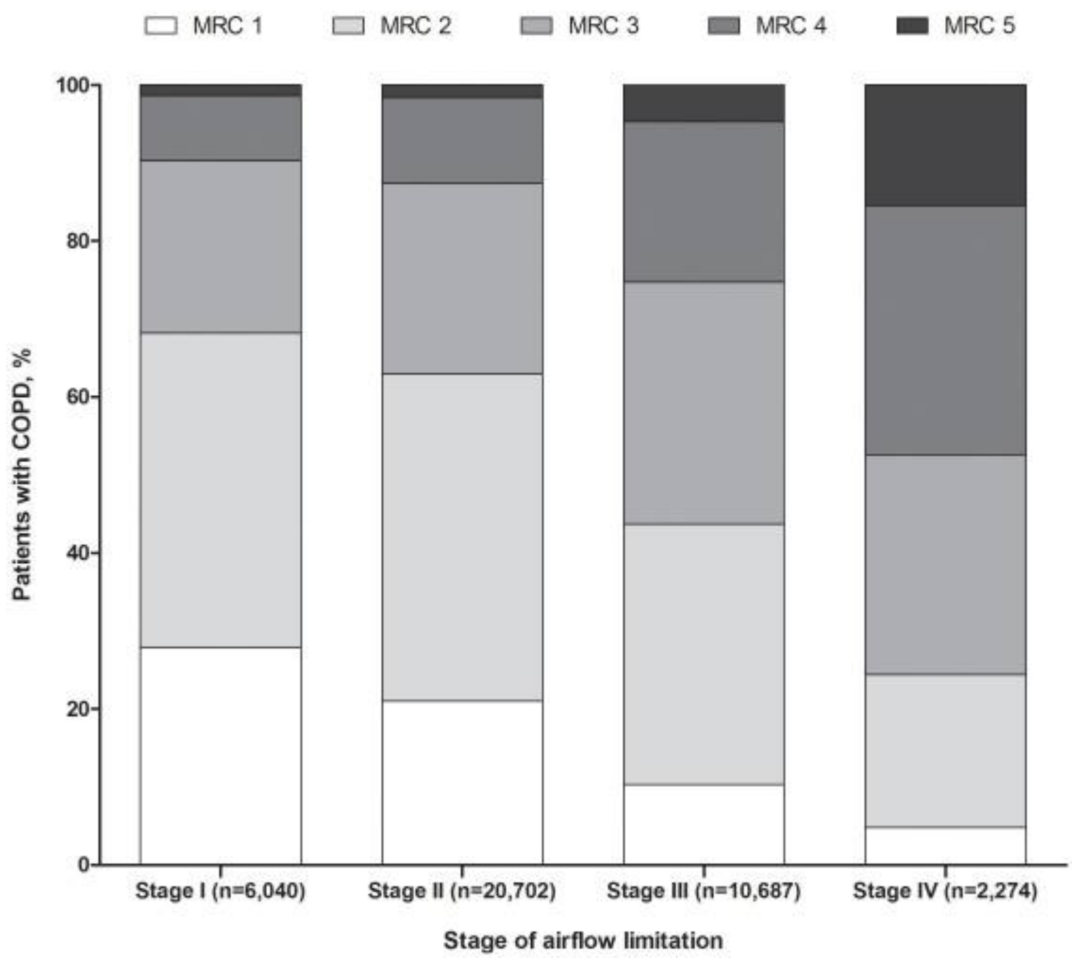



**Figure 1.2.** Percentage of patients with COPD in each MRC dyspnea grade based on severity of chronic airflow limitation, where stage I, II, III and IV represent mild, moderate, severe and very severe COPD according to criteria established by the Global Initiative for Obstructive Lung Disease [9].

Recently, an updated model outlining the relationship between dyspnea and physical inactivity has been developed and externally validated in people with COPD (**Fig. 1.2**). This model displays the continuous cycle of how progressive airflow limitation leads to lung hyperinflation with attendant dyspnea and exercise limitation, which, in turn, contributes to progressive worsening of dyspnea, physical inactivity and exercise intolerance in people with COPD [13]. A study by Watz et al. [14] showed that people with COPD who reported being too breathless to leave their house had a physical activity level just above being bed-ridden.

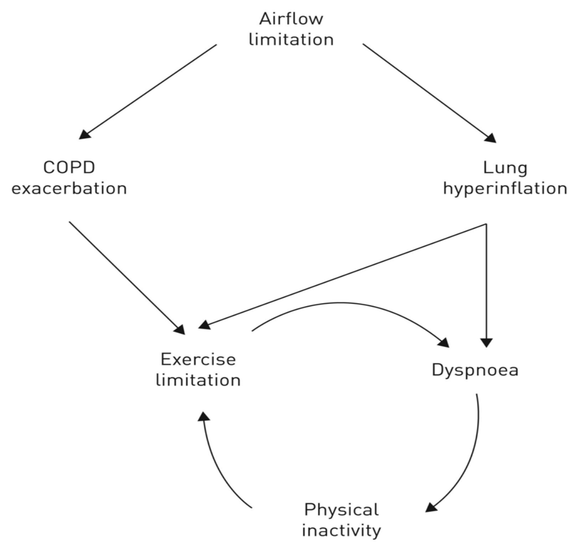



**Figure 1.3.** Vicious dyspnea-inactivity cycle in people with COPD [13].

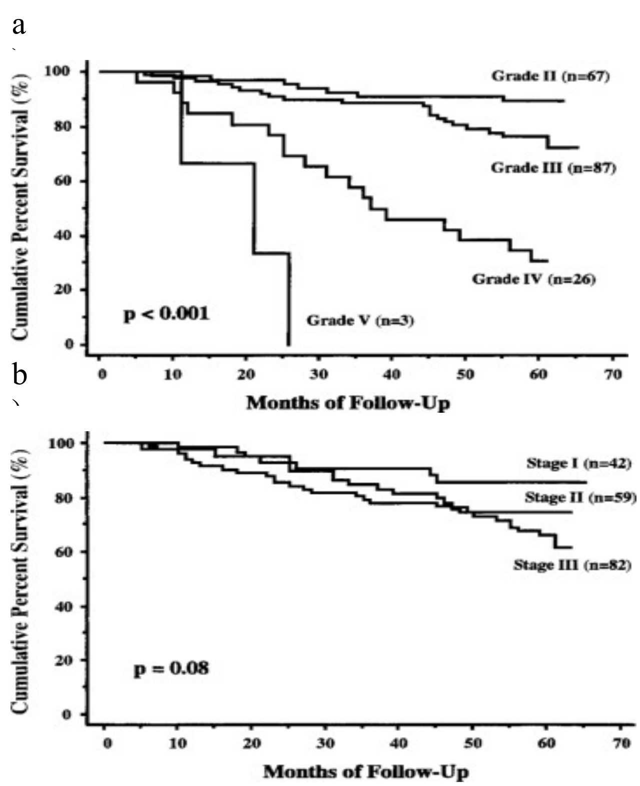



**Figure 1.4.** Five-year survival of people with COPD according to a) dyspnea and b) $FEV_1$ [15].

Breathlessness is becoming an increasingly important indicator of adverse health outcomes among people with COPD. Müllerová et al. [9] reported a positive correlation between exacerbation risk and MRC dyspnea score in COPD. This study also showed that 7% of respondents with a MRC score of 1 (dyspnea only with strenuous exertion) experienced a minimum of one severe exacerbation requiring hospitalization over the course of a year; this statistic increased to 24% among respondents with a MRC score of 5 (incapacitating dyspnea) [9]. Previous research has shown that the severity of activity-related dyspnea is a better independent predictor of 5-year mortality than $FEV_1$ in people with COPD (**Fig. 1.3**) [15, 16]. Considering the impact of breathlessness on physical (in)activity, exercise tolerance, health status, QoL, morbidity and mortality, it makes sense that breathlessness is a common target for therapy/disease management.

### 1.2 Treatment Algorithms Based on Breathlessness

Contemporary evidence-based guidelines for comprehensive management of COPD published by GOLD [4] and the Canadian Thoracic Society (CTS) [8] are focused on alleviating breathlessness, improving health status/QoL and preventing/treating exacerbations *via* various pharmacological and non-pharmacological interventions (**Fig. 1.4**). Despite these guidelines being heavily focused on breathlessness relief (symptom management), there are few evaluative tools currently available for routine clinical use to assess activity-related breathlessness and its response to therapy at a standardized and individualized level of exertion; such tools are needed to accurately assess interventional efficacy and optimize symptom management of people with COPD [17].

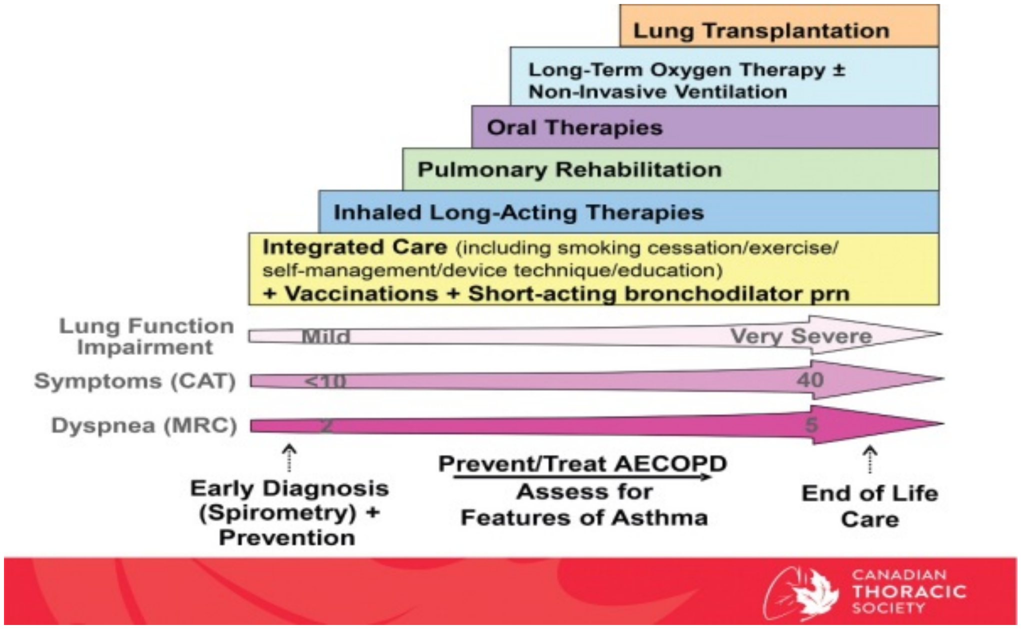



**Figure 1.5**. Algorithm for the comprehensive management of COPD advocated by the Canadian Thoracic Society [18]. **Abbreviations**: CAT, COPD Assessment Test; MRC, Medical Research Council dyspnea scale; AECOPD, acute exacerbation of Chronic Obstructive Pulmonary Disease; Short-acting bronchodilator prn, short-acting bronchodilator as needed.

### 1.3 Standardized Breathlessness Evaluation

Regulatory agencies like the US Food and Drug Administration (FDA) have recognized symptom relief as a justifiable goal for drug development [19]; however, the problem currently exists that there is no operationally feasible test that can accurately assess changes in exertional breathlessness in the primary care setting. Presently, there are tools capable of evaluating the impact of breathlessness on a patient's ADL, including (but certainly not limited to) the MRC dyspnea scale and the Baseline and Transition Dyspnea Indexes (B/TDI). However, these tools have substantial limitations (refer to *Section 5.1* below) and, as a result, the FDA made a statement declaring these tests are *"unsuitable for use as the sole or primary evidence of efficacy, and for supporting specific labeling claims"* [19, 20]. A new tool that can provide a more accurate, standardized and individualized assessment of exertional breathlessness needs to be developed to help optimize symptom-based management of COPD as per GOLD and CTS clinical practice guidelines.

The majority of tests most often used to assess exertional breathlessness (outlined in *section 1.4*) are not standardized for the level of exertion [21], which violates a fundamental principle of psychophysics: standardize the stimulus, evaluate the symptom response. Failure to standardize the level of exertion (stimulus) when assessing breathlessness (symptom response) in COPD has the potential to lead to misguided decisions about interventional efficacy, which may, in turn, compromise clinical care of people with COPD (e.g., failure to prescribe adequate medication or unnecessary dose escalation for breathlessness relief) [21].

## 1.4 Existing Tools to Assess Breathlessness

A variety of tools/tests are commonly used to assess breathlessness and its response to therapy in people with COPD, most notably: task-based questionnaires; self-paced exercise tests; and incremental and constant work rate exercise tests. For the reasons outlined in **Table 1.1** and described in detail below, these tests are not particularly viable for use in the primary care setting of COPD to assess for changes in exertional breathlessness.

### *1.4.1 Respiratory Questionnaires*

There exist multiple respiratory questionnaires commonly used to assess breathlessness and its response to therapy in people with COPD. Some notable examples include the mMRC, the B/TDI, the Chronic Respiratory Disease Questionnaire (CRQ), the Multidimensional Dyspnea Profile (MDP), the Shortness of Breath with Daily Activities questionnaire (SOBDA), the

University of California, San Diego Shortness of Breath questionnaire (UCSD SOBQ), and the Dyspnoea-12 questionnaire (D12) [22, 23]. These questionnaires typically consist of questions aimed at understanding the level of breathlessness experienced by an individual during a given task, or on average on a day-to-day basis. There are many beneficial aspects of using these task-based questionnaires to assess breathlessness; they are simple, inexpensive, require no equipment, are responsive to intervention (both bronchodilation and pulmonary rehabilitation), and correlate with adverse clinical and patient-reported outcomes, including mortality [17, 24-27]. Benefits notwithstanding, these questionnaires lack the level of specificity needed to provide a standardized stimulus to accurately assess breathlessness on exertion and its response to therapy. Take the BDI for example; stage 1 of the Baseline Magnitude of Effort is classified by the patient as *"[Becoming] short of breath with little effort. Tasks performed with little effort or more difficult tasks performed with frequent pauses and requiring 50-100% longer to complete than the average person might require"* [24, 28]. What qualifies as "little effort"? What classifies someone as an "average person"? This description is too vague to accurately quantify the breathlessness severity. Taking all this into consideration, questionnaires are a useful tool to assess the impact of breathlessness on a person's daily life; however, they do not directly quantify the intensity of exertional breathlessness and therefore need to be cross-referenced with exercise tests that quantify breathlessness at a standardized exercise stimulus [29].

*1.4.2 Self-Paced Exercise Tests*

One of the most common self-paced exercise tests used in the context of COPD is the 6-minute walk test (6MWT) [30]. This test is performed on a flat 30m course where the participant is instructed to walk as far as they can in a period of 6-minutes. Results are measured in meters or feet traveled in the span of 6-minutes (i.e., 6-min walk distance [6MWD]). Measures of breathlessness and leg fatigue are taken prior to beginning and at the end of 6-minutes; arterial oxygen saturation ($SpO_2$) and heart rate (HR) are monitored throughout the 6-minutes. The 6MWT is responsive to both pharmacological and non-pharmacological interventions in COPD with a MCID of 25-33m [30]. By all accounts, the 6MWT is a test of functional capacity, with 6MWD correlating well with cardiopulmonary exercise test derived measures of peak exercise capacity (e.g., peak rate of $O_2$ consumption [$V'O_{2peak}$]); however, for the reasons outlined below, the 6MWT is not well designed to assess for changes in breathlessness [29].

While the 6MWT requires little experience, money and equipment to conduct, it fails to adhere to the principles of psychophysics and should not be used to assess for changes in breathlessness. Because this test is self-paced, the level of exertion is not standardized and consequently, improvements in breathlessness cannot be attributed to interventional efficacy. For example, when COPD patients start bronchodilator therapy, their 6MWD typically improves (i.e., they walk faster and at a higher intensity), while breathlessness intensity ratings are often similar at end exercise because it is a maximal or near maximal test. During the 6MWT, breathlessness is only assessed at peak exercise when, despite differences (increases) in the distance walked in response to therapy, intensity ratings of breathlessness are often unchanged following treatment. Due to this limitation, the 6MWT may compromise clinical decision-making, rendering it inappropriate for use in the clinical care setting of COPD. Furthermore, significant variability in 6MWD has been observed between the first and second 6MWT performed by people with COPD hospitalized for an acute exacerbation [31]; thus, it is recommended that the 6MWT be performed twice when assessing exercise capacity due to the presence of a learning effect.

*1.4.3 Incremental Exercise Tests*

Incremental work-rate tests (INCR) are laboratory-based exercise tests that provide practitioners with insight into the underlying mechanisms of exercise intolerance; they are currently considered the “Gold Standard” for cardiopulmonary exercise testing (CPET) [32]. INCRs are typically performed on an electronically braked cycle ergometer or motorized treadmill, with the intensity being increased incrementally (e.g., 10 W/min) until the point of volitional fatigue or symptom limitation. Cardiac, metabolic and ventilatory parameters are simultaneously measured breath-by-breath using a metabolic cart, while intensity ratings of breathlessness and leg fatigue are collected at regular intervals throughout the test. The rate of $O_2$ consumption (V’$O_2$ peak) and work rate (WR peak) at the symptom limited peak of INCR CPET are responsive to both pharmacological and non-pharmacological interventions in people with COPD [30]. While INCR is considered the “Gold Standard” for CPET, it is not pragmatic for use in the primary care setting of COPD because its conduct requires a substantial amount of time, equipment, technical expertise and financial resources. Additionally, INCR CPET performed on the cycle ergometer does not mimic the most problematic ADLs reported by people with COPD (i.e., walking, stair climbing) and lacks an established MCID [30].

The incremental shuttle walking test (ISWT) is another symptom limited exercise test commonly used to assess the exercise tolerance of people with COPD [37]. During this test, people are instructed to walk between two pylons placed at 0.5m and 9.5m of a 10m course (to allow for a 0.5m turning radius) at an externally paced audio signal at which they are expected to have reached the pylon. The test continues until 12 minutes are complete or until symptom limitation (i.e., breathlessness and/or leg fatigue becomes intolerable). Performance is quantified by the distance covered or the time achieved. Intensity ratings of breathlessness and leg fatigue are often collected prior to beginning and at cessation of the test, while HR and $SpO_2$ are often monitored continuously throughout the entire duration of the test. This test is typically responsive to pulmonary rehabilitation as well as various pharmacological interventions in people with COPD [30]. The ISWT is used in both clinical care and research settings of COPD as it requires minimal facilities, expertise and financial resources to conduct, and also because it provides useful prognostic information [30]. However, like the 6MWT, the ISWT has a learning effect (meaning the test must be performed at least twice) [30, 33], which limits its use in clinical practice.

The primary limitation with incremental tests (i.e., INCR and ISWT) arises when analyzing the outcome measures; as the test itself is not an isotime measure, breathlessness is not the main outcome. The main outcome measure is exercise tolerance (quantified by time, distance covered, $V'O_2$ peak, WR peak). As previously mentioned, symptom relief is one of the primary goals of COPD management [1, 18], meaning the INCR and ISWT are not ideal tools for use in clinical practice. Nevertheless, this limitation can be overcome if measures are taken at predetermined submaximal work rates (INCR) or time points (ISWT) throughout the test; this allows changes in breathlessness to be tracked pre-to-post intervention where a standardized level of exertion is assumed at each time point.

### *1.4.4 Constant Work Rate Tests*

Constant work-rate exercise tests (CWR) are commonly used to assess interventional efficacy on exercise endurance and exertional breathlessness [30]. Like the INCR, CWR are typically performed on an electronically braked cycle ergometer or motorized treadmill. Prior to performing CWR, an INCR must be completed to determine the appropriate work rate (exercise intensity); typically 75-85% of the WR peak achieved during the INCR [30]. Cardiac, metabolic and ventilatory parameters are often simultaneously measured breath-by-breath using a metabolic cart, while intensity ratings of breathlessness and leg fatigue are often collected at regular intervals throughout the test using validated scales (e.g., Borg 0-10 category ratio scale [Borg CR10]). Together, these measurements provide insight into the pathophysiological mechanisms underlying exercise intolerance and exertional symptoms [30]. CWR isotime inspiratory capacity, isotime breathlessness and time to the limit of tolerance (or exercise endurance time) are responsive to numerous non-pharmacological and pharmacological interventions in the COPD population [30]. Again, when performed on a cycle ergometer, the CWR does not mimic the most problematic ADLs reported by people with COPD (i.e., walking, stair climbing).

A common CWR test used within the COPD population is the endurance shuttle walking test (ESWT). The ESWT follows the same general procedure as the ISWT (i.e., walking along a 10m course at an externally paced audio signal for 20 minutes or until symptom limitation) but the external pace of the audio signal stays consistent throughout the test (hence its classified as a CWR test). Prior to performing this test, an ISWT must be completed to gather baseline data needed to determine the appropriate walking speed (exercise intensity), which is 80% of peak ISWT speed [30]. ESWT results are quantified by duration or distance covered. Like the ISWT, intensity ratings of breathlessness and leg fatigue are often collected prior to beginning and at cessation of the ESWT, while HR and $SpO_2$ are often monitored continuously throughout the entire duration of the test. The outcomes of endurance time and distance are responsive to pulmonary rehabilitation, oxygen therapy and bronchodilator therapy in COPD [30].

The CWR and ESWT are generally considered to be the “Gold Standard” in evaluating interventional efficacy on exertional breathlessness in COPD as they adhere to the principles of psychophysics and are responsive to both pharmacological and non-pharmacological interventions [30]. Nevertheless, these tests have aspects that limit their operational feasibility for routine use in primary care as well as in large clinical (therapeutic) trials that assess for changes in exertional breathlessness. Specifically, not only do CWR/ESWT require prior INCR/ISWT testing to establish the appropriate exercise intensity, their conduct also requires considerable time, space, equipment and technical expertise.

**Table 1.1**. Properties of different exercise tests and questionnaires used to assess exertional breathlessness

| | 6-MWT | INCR CPET | CWR CPET | ISWT | ESWT | Questionnaires | | | | | | | 3-min CSST | 3-min CRSST |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | mMRC | BDI/TDI | D-12 | SOBDA | CRQ | MDP | UCSD SOBQ | | |
| **Standardized level of exertion** | No[30] | Yes[30] | Yes[30] | Yes[30] | Yes[30] | N/A | N/A | N/A | N/A | N/A | N/A | N/A | Y[34] | Y[35] |
| **Safety** | Yes[30] | Yes[30] | Yes[30] | Yes[30] | Yes[30] | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Y[34] | Y[35] |
| **Stands alone** | Yes[30] | N/A | No[30] | N/A | No[30] | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Y[36] | Y[36] |
| **Mimic ADL?** | Yes[30] | Yes (T) No (C) [30] | Yes (T) No (C)[30] | Yes[30] | Yes[30] | N/A | N/A | N/A | N/A | N/A | N/A | N/A | Y[34] | Y[35] |
| **MCID** | Yes[30] | No[30] | Yes[30] | Yes[30] | Yes[30] | Yes[37] | Yes[27] | No[23] | Yes[23] | Yes[17] | No[38] | Yes[39, 40] | Y[41] | Y[41] |
| **Reproducible** | Yes[30] | Yes[30] | Yes[30] | N/A[30] | Yes[30] | LI | Yes (I) [42] | Yes[23] | Yes[43] | Yes[44, 45] | Yes[46, 47] | Yes[42] | Y[36] | Y[36] |
| **Responsive to BD** | Yes[30] | Yes (I) [30] | Yes[30] | Yes (I) [30] | Yes (I) [30] | No[48] | Yes[27, 49] | Yes | Yes[49, 50] | No[51] | Yes[52] | LI | Y[34] | Y[35] |
| **Responsive to Rehab** | Yes[30] | Yes[30] | Yes[30] | Yes[30] | Yes[30] | No[48] | Yes[27, 53] | Yes | LI | Yes[54] | Yes[55] | Yes[53, 56] No[57] | Y[41] | Y(I) [41] |
| **Time** | Low[30] | High[30] | High[30] | Medium [30] | High[30] | Low | Low | Low | Low | Low | Low | Low | Low[34] | Low[35] |
| **Equipment** | Low[30] | High[30] | High[30] | Low[30] | Low[30] | Low | Low | Low | Low | Low | Low | Low | Low[34] | Low[35] |
| **Expertise** | Low[30] | High[30] | High[30] | Low[30] | Medium [30] | Low | Low | Low | Low | Low | Low | Low | Low[34] | Low[35] |

**Abbreviations**: 3-min CSST, three-minute constant speed shuttle test; 3-min CRSST, three-minute constant rate stair stepping test; 6-MWT, six-minute walk test; ADL, activities of daily living; BDI, Baseline Dyspnea Index; C, cycle ergometer; CPET, cardiopulmonary exercise testing; CRQ, chronic respiratory disease; CWR, constant work rate; D-12, Dyspnea-12; ESWT, endurance shuttle walk test; I, inconsistent; INCR, incremental; ISWT, incremental shuttle walk test; LI, limited information; MCID, minimal clinically important difference; MDP, Multidimensional Dyspnea Profile; mMRC, modified Medical Research Council dyspnea score; N/A, not applicable; SOBDA, Shortness of Breath with Daily Activities; SOBQ, Shortness of Breath Questionnaire; T, treadmill; TDI, Transition Dyspnea Index; UCSD; University of California, San Diego

### 1.5 Declaration

In light of the stated limitations (both methodological and practical) of existing tests to assess for changes in exertional breathlessness (refer to **Table 1.1**), there is a need for tests that standardize and individualize the level of exertion faced during physical activities known to be the most problematic for people with COPD, notably walking and stair stepping [11]. Using a broader view, it becomes clear that other areas of medicine have equitable tests to assess interventional efficacy on important clinical and/or patient-reported outcomes that are consistently used by healthcare providers. Classic examples are the use of blood pressure cuffs in people with hypertension and the oral glucose tolerance test used in people with type II diabetes. Both of these tests provide valid, standardized and responsive measurements [58-60] that are seemingly irreplaceable in the primary care setting of their respective fields. The lack of such a standardized test for exertional breathlessness has the potential to lead to misguided decisions in the clinical care of COPD. Consider a scenario where a person with COPD modifies their lifestyle to participate in less physical activity and/or purposefully avoids situations that may cause them to become breathless, despite their healthcare provider switching them from mono to dual bronchodilator therapy. When filling out any one of the aforementioned task-based breathlessness questionnaires before and 6 weeks after their change in bronchodilator regime, this person may report experiencing less impact of breathlessness on their ADLs following the change in bronchodilator therapy. Under these circumstances, the healthcare provider may conclude that the change in bronchodilator regime was effective at alleviating their patient's exertional breathlessness, when, in fact, relief of exertional breathlessness may be due to the patient's lower levels of physical activity participation. This situation could lead to failure of the healthcare provider to better optimize pharmacotherapy. By contrast, it is possible for a person with COPD to increase their level of daytime physical activity and report no change in their perceived breathlessness during ADLs when responding to a task-based questionnaire 6 weeks after a change in bronchodilator regime. Under these circumstances, the healthcare provider may conclude that the change in bronchodilator regime was ineffective at alleviating their patient's exertional breathlessness and unnecessarily modify their approach to symptom management (e.g., dose escalation, prescribe an opioid, etc.). These theoretical scenarios highlight the aforementioned limitation of using task-based questionnaires to assess for changes in exertional breathlessness, while simultaneously highlighting the need for a clinically relevant test of breathlessness that standardizes the level of exertion during the most problematic ADLs experienced by people with COPD.

### 1.6 Characteristics of an Ideal Test

In our opinion, an ideal test to assess for changes in exertional breathlessness in both clinical care and research settings of COPD should satisfy the following criteria:

1) The test should adhere to the principles of psychophysics [61], meaning the level of exertion (stimulus) should be standardized when quantifying the perception of breathlessness (response).

2) The test should be safe, cost effective, and require little expertise, time, space and equipment.

3) The test should stand-alone and not require prior testing to gather baseline information (i.e., should not require completion of an incremental exercise test to set the exercise intensity).

4) The test should mimic the most troublesome ADLs for people with COPD, most notably walking and stair stepping [11].

5) The test should provoke a level of breathlessness that is amendable to therapy (e.g., ≥3 Borg CR10 scale units).

6) The test should be responsiveness to intervention (both pharmacologic and non-pharmacologic), reproducible and have an established Minimal Clinically Important Difference (MCID).

### 1.7 3-Minute Constant Rate Stair Stepping Test (3-min CRSST) and 3-Minute Constant Speed Shuttle Test (3-min CSST)

In the past 10-15 years, two tests which meet the criteria of an ideal exertional breathlessness test outlined in *section 1.6* have been developed and prospectively validated [36]; the 3-minute constant rate stair-stepping test (3-min CRSST) and the 3-mintue constant speed shuttle test (3-min CSST).

The 3-min CRSST is performed on a wooden stair that rises 20cm off the ground. The person is instructed to 'step up' [indicating to place both feet up onto a single step, one foot after the other] and 'step down' [indicating to step back down to the floor, one foot after the other] by an automated cue played from a computer for 3-minutes; they maintain this step rate for the entire 3-minutes or until symptom limitation (i.e. breathlessness or leg fatigue gets too great for them to keep pace) [36].

The 3-min CSST follows a similar concept but participants are instructed to walk between two pylons placed at 0.5m and 9.5m of a 10m course (to allow for a 0.5m turning radius) at a constant and externally paced 'beep' provided by an audio signal. Walking starts upon hearing the audio instructions and participants walk the course for 3-minutes reaching the pylons on each 'beep' [36]. Again, people are instructed to match the pace to the best of their ability for the entire 3-minutes or until symptom limitation.

For both the 3-min CRSST and 3-min CSST, HR and $SpO_2$ are monitored throughout the entire 3-minutes; Borg CR10 scale intensity ratings of breathlessness and leg fatigue are taken at the beginning (pre-exercise rest), at regular intervals during the test, and end of each test (i.e., 3-minutes or at symptom limitation, whichever occurs first). The pace for these tests is predetermined; the current algorithm used to select the intensity of exercise (step rate or shuttle speed) and the limitations of this algorithm are discussed below in *section 1.7.4.*

Compared to the 6MWT, INCR, CWR, ISWT and ESWT, these two simplified field-based exercise tests have many advantages and meet the requirements for an ideal test outlined in *section 1.6*:

1) they require limited space, expertise, time and money to conduct;

2) they do not require baseline information from an incremental exercise test to determine stepping rate and shuttle speed;

3) they are externally paced eliminating the possibility of a learning effect;

4) they mimic the two most problematic ADLs for people with COPD;

5) they adhere to the laws of psychophysics allowing changes in breathlessness to be assessed at standardized and individualized levels of exertion (**Table 1.1**).

### *1.7.1. Reproducibility of the 3-min CRSST and 3-min CSST*

Given the standardized and individualized level of exertion these tests provide, the primary outcome measure is breathlessness intensity at the end of the 3$^{rd}$ minute of exercise (opposed to, for example, the 6MWT where the primary outcome is 6MWD) [36]. Because time and step rate/shuttle speed (exercise intensity) are standardized and individualized, the 3-min CRSST and

3-min CSST have the potential to provide accurate assessments of intra-individual changes in exertional breathlessness. Strong Pearson's and intra-class correlation coefficients were reported by Perrault et al. [36] for all variables during repeat step and shuttle tests; these findings indicate that the measurement techniques for physiological variables and more importantly breathlessness are reproducible (**Figs. 1.5 & 1.6**). Furthermore, the 3-min CRSST and 3-min CSST have a small measurement error in regards to intensity ratings of breathlessness of just ±0.5 Borg CR10 scale units from trial 1 to trial 2, which falls below the MCID of ±1 Borg CR10 scale unit [36]. The reproducibility was equally strong in the 3-min CRSST and 3-min CSST. During these trials, test-retest correlation coefficients were all greater than 0.93 and less than a 5% difference was reported from trials 1 to 2 for $V'O_2$, ventilation, tidal volume, HR and Borg CR10 scale breathlessness intensity ratings [36] (**Fig. 1.5 & 1.6**). The standardized level of exertion induced by these tests ensures a constant metabolic stress will be placed on the body from test-to-test; this standardizes the underlying physiology allowing intra-individual changes in breathlessness to be attributed to an intervention or change in disease status vs. the test itself [36].

The 3-min CRSST and 3-min CSST are arguably more ecologically valid than tests conducted on a cycle ergometer or motorized treadmill because they more closely mimic the most troublesome ADLs reported by people with COPD. In earlier studies conducted primarily in people with moderate to severe COPD, stepping rates and shuttle speeds of 14-32 steps/min and 1.5-6 km/hr were used during the 3-min CRSST and 3-min CSST, respectively [34-36].

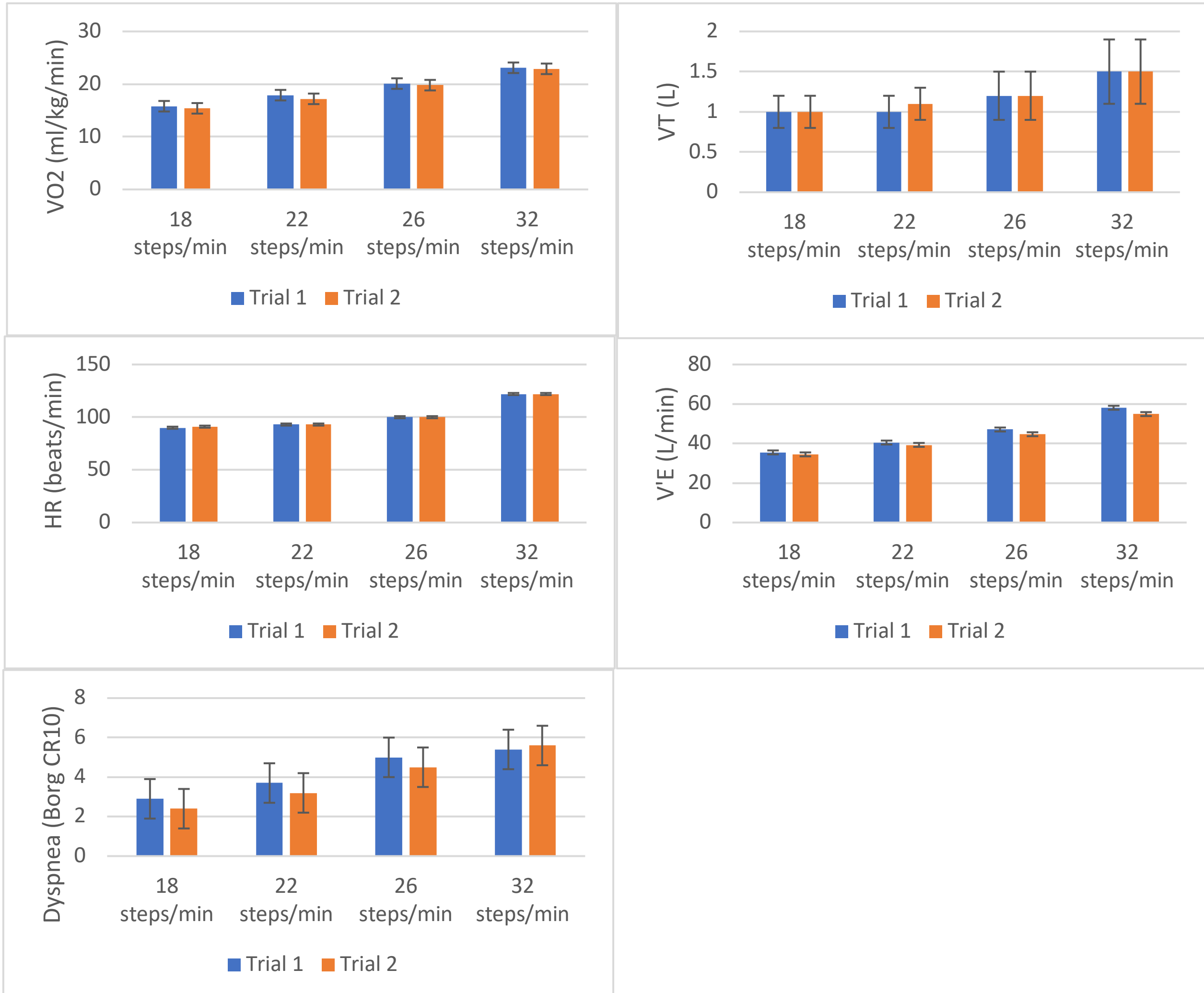



**Figure 1.6.** Oxygen consumption ($V'O_2$), ventilation ($V'_E$), tidal volume ($V_T$), HR and BORG dyspnea intensity ratings during trials 1 and 2 of the 3-min CRSST at each stepping rate. Adapted from Perrault et al. [36]

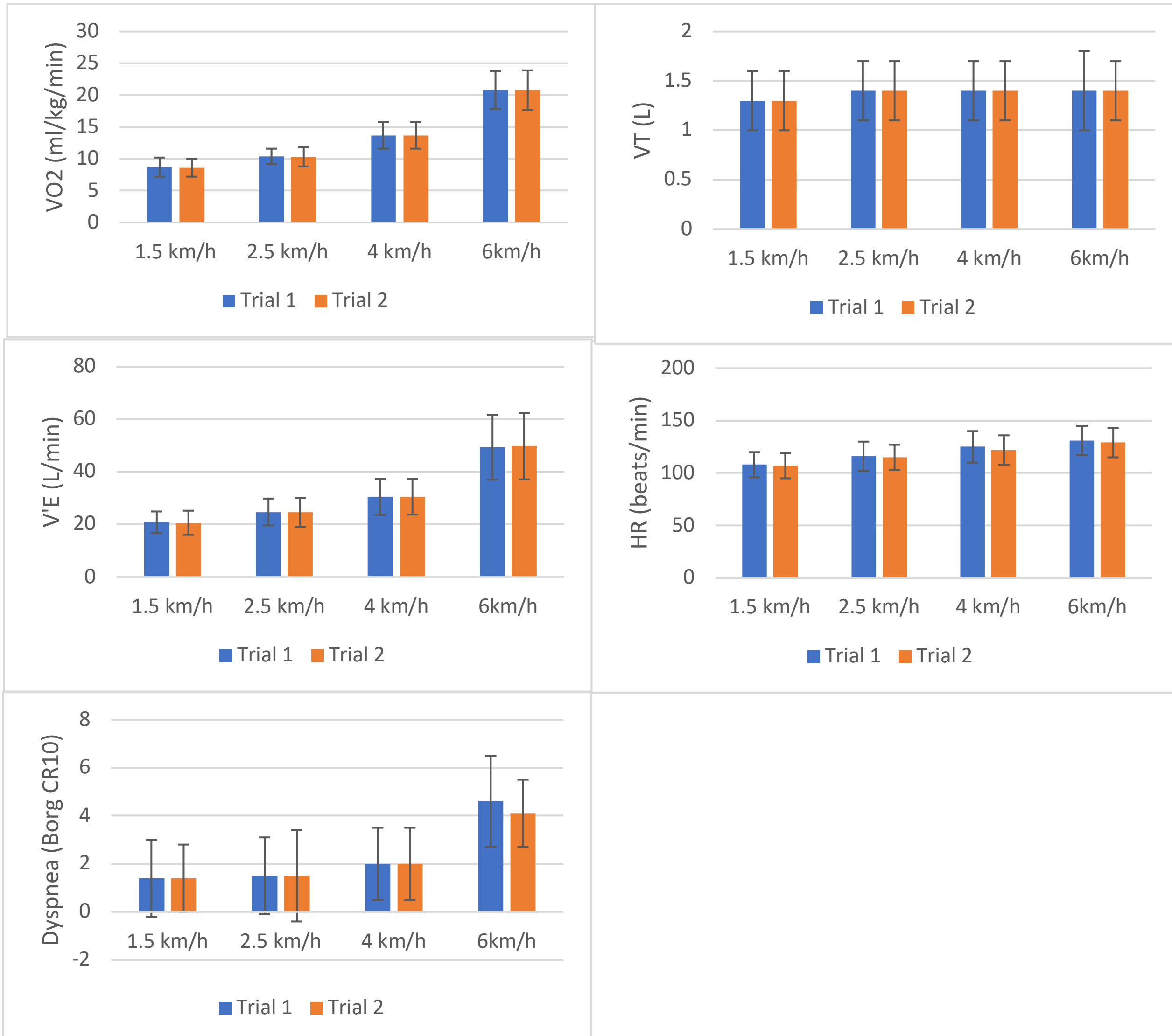



**Figure 1.7.** Oxygen consumption (V'O$_2$), ventilation (V'$_E$), tidal volume (V$_T$), HR and BORG dyspnea intensity ratings during trials 1 and 2 of the 3-mn CSST at each shuttle speed. Adapted from Perrault et al. [36]

*1.7.2. Responsiveness of the 3-min CRSST and 3-min CSST*

*3-min CRSST*

A randomized, double-blind, placebo-controlled, crossover study by Borel et al. [35] confirmed the responsiveness of the 3-min CRSST to detect relief of exertional breathlessness following acute bronchodilator therapy in 40 people with COPD [35]. Compared to placebo, single-dose administration of ipratropium bromide (500 µg) / salbutamol (2.5 mg) significantly decreased breathlessness intensity ratings by an average of 0.6-0.7 Borg CR10 scale units at the end of the 3-min CRSST performed at rates of 14 and 16 steps/min, but not at 20 and 24 steps/min (**Fig. 1.7**). Borel et al. [35] speculated that changes in breathlessness intensity ratings were not significant at higher stepping rates due to the intensity of the test. These higher stepping rates elicited levels of ventilation between 85-91% of peak INCR CPET values, suggesting that participants were most likely breathing on the upper alinear (non-compliant) portion of the respiratory systems' sigmoid pressure-volume curve, where improvements in dynamic lung function associated with bronchodilation may not result in proportional decreases in breathlessness [35]. These results indicated a possible ceiling effect with the 3-min CRSST, in which as the intensity gets too high, the test becomes unresponsive.

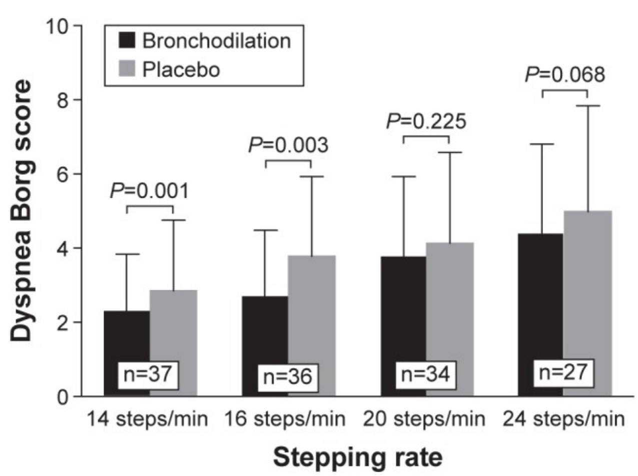



**Figure 1.8.** Effect of bronchodilation vs. placebo on BORG CR10 dyspnea intensity ratings at end exercise during the 3-min CRSST in people with COPD [35].

In addition, the Borel et al. [35] study showed that breathlessness intensity ratings decreased by ≥1 Borg CR10 scale unit (indicating a clinically significant improvement) in 45%, 45%, 33% and 28% of participants completing all 3-min of stair stepping exercise at 14, 16, 20 and 24 steps/min, respectively [35]. The findings of this study provided scientific justification for use of the 3-min CRSST (particularly at step rates <20 steps/min) in clinical care and research settings of COPD to evaluate interventional efficacy on exertional breathlessness.

*3-min CSST*

A randomized, double-blind, placebo-controlled, crossover study by Sava et al. [34] confirmed the responsiveness of the 3-min CSST to detect relief of exertional breathlessness following acute bronchodilator therapy in 39 people with moderate to severe COPD. In this study, shuttle speeds of 2.5, 4 and 6 km/hr were used and results showed that at standardized measurement time points of 2, 2.5 and 3.0 min, Borg CR10 scale intensity ratings of breathlessness were reduced following single-dose inhalation of ipratropium bromide (500 μg) compared to placebo (**Fig. 1.8**) [34]. Out of the 39 participants, 27 (69%) reported a clinically meaningful decrease in Borg CR10 scale intensity ratings of breathlessness of ≥1 unit at the end of the 3$^{rd}$ minute of exercise.

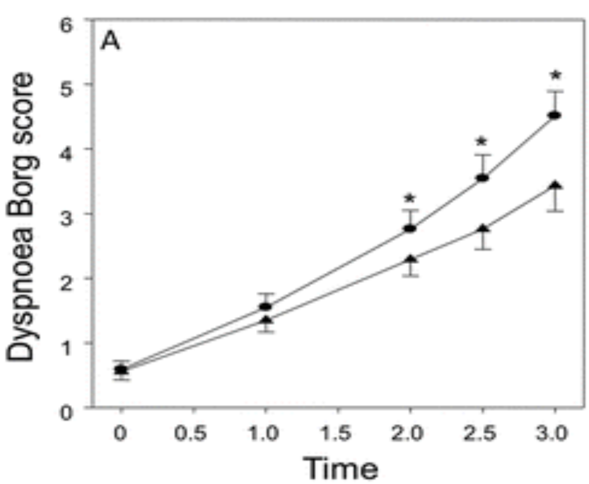



**Figure 1.9.** BORG CR10 dyspnea intensity ratings taken throughout the fastest completed 3-min CSST for each participant under both experimental conditions (ipratropium bromide [triangles] and placebo [circles]) [34].

More recently, Maltais et al. [62] provided additional experimental support of the responsiveness of the 3-min CSST to detect relief of exertional breathlessness following bronchodilator therapy in people with COPD. This randomised, double-blind, two-period crossover study showed that, compared with baseline, Borg CR10 scale intensity ratings of exertional breathlessness decreased following 6-weeks of treatment with 5 µg tiotropium (monotherapy) and 5 µg tiotropium/ 5 µg olodaterol (combination therapy) (**Fig. 1.9**). This study also found that the 3-min CSST was sensitive to detect modest but statistically significant differences in exertional breathlessness after 6-weeks of mono *versus* dual bronchodilator therapy (**Fig. 1.9**) [62]. The findings of these studies by Sava et al. [34] and Maltais et al. [62] provided scientific justification for use of the 3-min CSST in clinical care and research settings of COPD to assess interventional efficacy on exertional breathlessness.

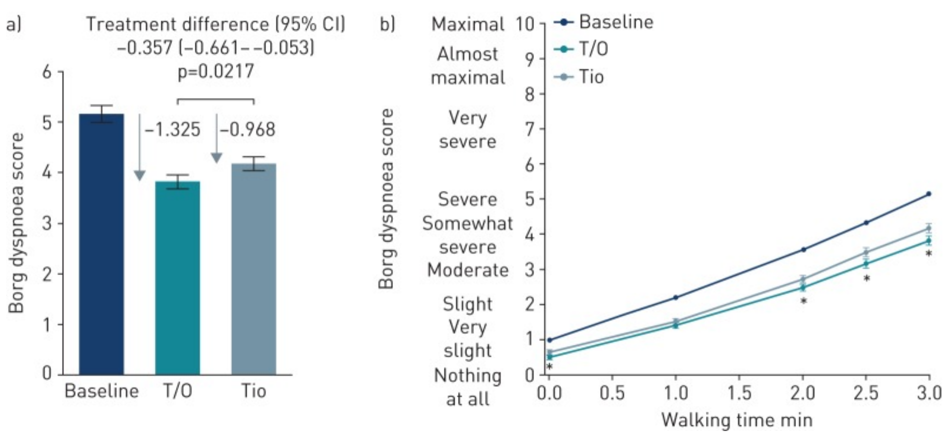



**Figure 1.10.** Borg CR10 scale dyspnea intensity scores during (b) and upon completion (a) of the 3-min CSST in people with COPD under the three experimental conditions; baseline, tiotropium/olodaterol (T/O) and tiotropium (Tio) [62].

To date, no published studies have assessed the responsiveness of the 3-min CRSST or 3-min CSST to detect relief of exertional breathlessness following non-pharmacological therapies; however, preliminary data from our laboratory by Tracey et al. [41] assessed the responsiveness of these tests to 8-12 weeks of pulmonary rehabilitation in 56 people with COPD (mean±SD $FEV_1$%predicted = 57.8±22.9). The preliminary results showed the PR-induced change in breathlessness intensity ratings was -1.0±1.5 Borg CR10 scale units (mean±SD) for the 3-min CSST (p=0.03) and -0.6±2.1 Borg CR10 scale units for the 3-min CRSST (p=0.24), suggesting that the 3-min CSST is responsive to PR whereas the 3-min CRSST is not. When these analyses were restricted to individuals who had baseline (or pre PR) breathlessness intensity ratings of >3 Borg CR10 scale units, the PR-induced change in breathlessness intensity ratings was -0.9±1.6 and -0.8±1.8 Borg CR10 scale units for the 3-min CSST (p=0.01) and the 3-min CRSST (p=0.04), respectively (**Fig. 1.10**) [41]. These preliminary findings suggest that both of these exercise tests are responsive to PR, provided breathlessness intensity ratings during the initial assessment period are of at least "moderate" intensity on Borg's CR10 scale. We suspect that the tests lack responsiveness when breathlessness intensity ratings during the initial assessment period are <3 Borg CR10 units because there is not much room for symptomatic improvement when the perception of breathlessness at baseline is not particularly intense and/or bothersome. This indicates a possible floor effect of these tests wherein if the intensity of breathlessness is too low, the test lacks responsiveness.

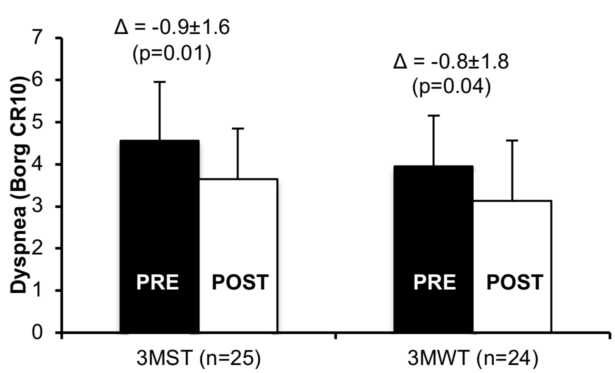



**Figure 1.11**. Pulmonary rehabilitation-induced changes in dyspnea intensity ratings with analyses restricted to patients with a baseline BORG rating of >3 units [41].

### *1.7.3. MCID of the 3-min CRSST and 3-min CSST*

An established MCID is an important aspect of tests to be used in clinical practice and research. Using a distribution-based analysis (i.e., 0.5*SD of the PR-induced change in breathlessness intensity ratings), preliminary results from Tracey et al. [41] suggest that a change in breathlessness intensity ratings of ±0.9 and ±1.0 Borg CR10 units is clinically meaningful for the 3-min CRSST and 3-min CSST in people with COPD, respectively. These estimates are comparable to other studies reporting MID values of approximately ±10mm on the 100mm Visual Analog Scale, which is roughly equivalent to a ±1 unit change on Borg's CR10 scale [40, 63].

### *1.7.4. Limitations of the 3-min CRSST and 3-min CSST*

Despite advancements of the step and shuttle tests, their use in the clinical care and research setting of COPD is limited by the method currently used to identify the exercise intensity (stepping rate or shuttle speed) needed to provoke a level of breathlessness amenable to therapy in an individual. The current approach is algorithmic with all participants starting at the same stepping rate/shuttle speed and subsequently increasing or decreasing the stepping rate/shuttle speed based on whether or not they finish all 3-min of the initial test [62]. For the 3-min CSST, participants start at a shuttle speed of 4 km/hr. If they are unable to complete all 3-minutes of exercise, the speed is decreased to 3.25 km/hr; however, if all 3-mintues of exercise are completed, the speed is increased to 5 km/hr (**Fig. 1.11**). For the 3-min CRSST, participants start stair stepping at 16 steps/minute. If they are unable to complete all 3-minutes of exercise, the step rate is decreased to 14 steps/minute; however, if all 3-minutes of exercise are completed, the step rate is increased to 20 steps/minute. This pattern of adjustment continues until the participant successfully completes all 3-minutes of shuttle walking or stair stepping exercise and rates their intensity of perceived breathlessness as ≥4 Borg CR10 units at the end of the 3rd minute of exercise. While functional and successful in 97% of participants studied [34, 62], application of this algorithm is time consuming and burdensome to both the participant and healthcare provider (or researcher) as a minimum of two step or shuttle tests must be conducted to determine the "optimal" stepping rate or shuttle speed [62]. Specifically, the current approach requires a minimum 20-30 minutes per patient with COPD to identify the optimal step rate or shuttle speed. It follows that the nature of its design makes this algorithm operationally infeasible for routine use of the 3-min CRSST and 3-min CSST in clinical practice as well as in large-scale clinical (therapeutic) trials, where time is not readily available for either physician/researcher or patient.

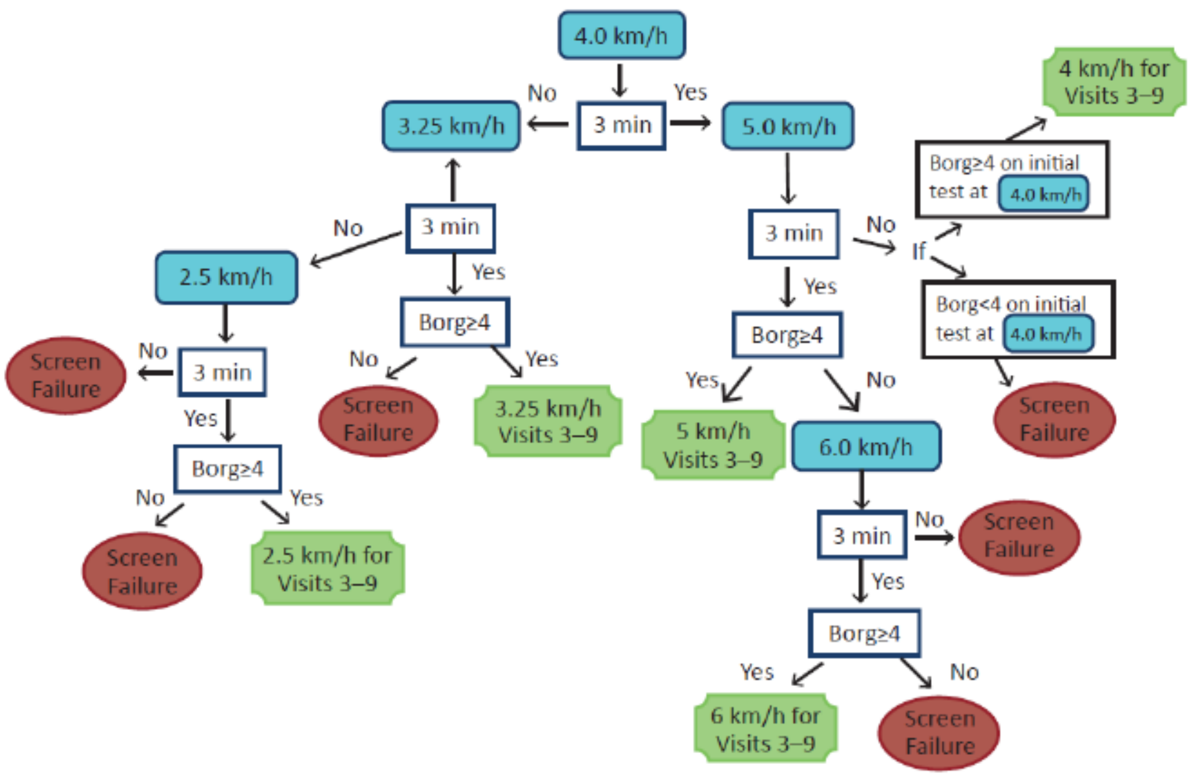



**Figure 1.12.** Example of algorithmic approach used by Maltais et al. [62] to identify the optimal (individualized) shuttle speed for people with COPD during the 3-min CSST.

Another potential problem with the algorithm displayed in **Fig 1.11** is the size of the increments between shuttle speeds. We contend that, for some people, the increments may be too large to truly individualize the exercise intensity that allows for the entire 3-min CSST to be completed with breathlessness intensity ratings of ≥4 Borg units. For example, someone may complete all 3-minutes of exercise at 4.5 km/hr with a breathlessness intensity rating of 3 Borg CR10 units; however, increasing the shuttle speed to 6 km/hr may elicit more severe breathlessness that prevents them from completing all 3-minutes of exercise. Alternatively, someone that failed to complete all 3-minutes of exercise by, for example, 15 seconds with a breathlessness intensity rating of 4 Borg CR10 units at a shuttle speed of 4 km/hr, may easily complete all 3-min of exercise at 3.25 km/hr but with a breathlessness intensity rating of <4 Borg CR10 units. In both cases, smaller increments in shuttle speed (e.g., 0.5 km/hr) may have been beneficial to find the most appropriate individualized exercise intensity. The same problem arises for the 3-min CRSST where the increments in step rate of 18, 22, 26 and 32 steps/min used by Perrault et al. [36] and of 14, 16, 20 and 24 steps/min used by Borel et al. [35] might have been too large for some people with COPD. Indeed, with increasing exercise intensity (stepping rate or shuttle speed), the percentage of patients who completed the entire three minutes of exercise decreased for both tests. A total of 93%, 90%, 85%, 68%, 70%, 40% of participants completed the entire 3-min CRSST at 14, 16, 20, 24, 26 and 32 steps/min, respectively [37, 38]. Similarly, 100%, 93%, 83% and 33% of participants completed the entire 3-min CSST at 1.5, 2.5, 4.0 and 6.0 km/hr, respectively [38].

### 1.8. Prediction Equation for the 3-min CRSST and 3-min CSST

In order for these tests to be more operationally feasible for implementation and use in clinical care and research settings, there needs to be a more effective and time efficient way to determine the "optimal" stepping rate/shuttle speed that provokes a level of breathlessness amenable to intervention at the end of the 3$^{rd}$ minute of exercise in an individual patient with COPD. Ideally, this issue would be resolved by development and prospective validation of an equation that would allow healthcare providers and researchers to identify (predict) the "optimal" stepping rate/shuttle speed for an individual patient based on their personal and clinical characteristics that are readily available in most medical charts (e.g., age, sex, body mass, $FEV_1$, mMRC dyspnea rating).

### 1.9 Objective

The objective of the proposed research was twofold: 1) to develop equations to predict the "optimal" stair stepping rate (3-min CRSST) and shuttle speed (3-min CSST) for an individual with COPD based on their unique personal and clinical characteristics; and 2) prospectively validate these newly developed equations in adults with COPD. Based on the results of the prospective validation, we also aimed to identify which of these tests, the 3-min CRSST or the 3-min CSST, and their respective prediction equations, may be better suited for use in the clinical care setting of COPD.

## **CHAPTER 2:** Three-minute constant rate stair stepping (3-min CRSST) and constant speed shuttle tests (3-min CSST) to assess exertional breathlessness in COPD: Development and prospective validation of an equation to select exercise intensity for use in clinical care and research.

## 2.1 Abstract

**Introduction:** The three-minute constant rate stair stepping test (3-min CRSST) and three-minute constant speed shuttle test (3-min CSST) are standardized breathlessness assessment tools which lack feasibility for use in clinical practice due to the burdensome nature of determining an appropriate exercise intensity. The objective of this research was twofold: 1) develop equations to select the step rate and shuttle speed for an individual with chronic obstructive pulmonary disease (COPD); and 2) prospectively validate these equations in adults with COPD. We also aimed to determine which of these tests may be better suited for use in clinical practice. **Methods:** The equations were derived retrospectively from people with COPD who completed one 3-min CRSST (n=90) and/or 3-min CSST (n=112) as part of previous studies; mixed-effects linear models estimated exercise intensity based on Borg CR10 scale breathlessness intensity ratings. Once developed, the equations were tested prospectively in 18 people with COPD. A successful prediction was characterized by the participant finishing all 3-minutes of exercise with a breathlessness intensity rating ≥3 Borg CR10 scale units. **Results**: Step rate/shuttle speed, sex, age, modified Medical Research Council dyspnea score, body mass (3-min CRSST) or body mass index (3-min CSST), and Global Initiative for Obstructive Lung Disease stage (3-min CRSST) or forced expiratory volume in 1-sec ($FEV_1$) %predicted (3-min CSST) were significant predictors of the breathlessness intensity response to the tests and were included in the final models. Prospective validation testing revealed that 61% of our participants finished the step or shuttle test at the exercise intensity identified using the prediction equations with a breathlessness intensity ≥3 Borg units. Physiological and perceptual responses to the tests were comparable. **Conclusion:** The equations developed and prospectively validated will permit use of the 3-min CRSST and 3-min CSST as standardized tools to assess breathlessness in people with COPD.

## 2.2 Introduction

Breathlessness on exertion is the most prominent and debilitating symptom of chronic obstructive pulmonary disease (COPD) [9]. Breathlessness is independently associated with exercise intolerance, physical activity limitation, exacerbation risk, adverse health status and premature death [1, 9, 15]. A primary aim of evidence-based guidelines for the comprehensive management of COPD published by the Global Initiative for Obstructive Lung Disease (GOLD) and the Canadian Thoracic Society (CTS) is to alleviate breathlessness via pharmacological and/or non-pharmacological intervention [4, 18]. Nevertheless, few assessment tools exist to measure breathlessness at a standardized and individualized level of exertion in clinical practice [21, 29].

The ability to quantify breathlessness in response to a standardized and individualized exercise stimulus is critical in clinical care and research settings of COPD to assess interventional efficacy in regard to symptom burden. Until recently, there has been no readily available exertional breathlessness evaluation test that adheres to the laws of psychophysics [30] and that is capable of eliciting reproducible levels of breathlessness in people with COPD during their most problematic activities of daily life (i.e., walking and stepping up the stairs) [11]. Two novel exercise tests that elicit a standardized and reproducible level of exertion, with respect to both intensity and duration, have been developed for use in COPD; the three-minute constant rate stair stepping test (3-min CRSST) and the three-minute constant speed shuttle test (3-min CSST) [29, 34-36].

These tests are reproducible, reliable, and responsive to both bronchodilator therapy and pulmonary rehabilitation in people with COPD [34-36, 41, 62]. Additionally, we recently estimated a minimal clinically important difference of ±1 Borg 0-10 scale unit in breathlessness intensity ratings for both step and shuttle tests [41]. The limitation to implementation and routine use of these tests for clinical and research purposes is the method of determining the step rate and shuttle speed (exercise intensity) required to elicit a breathlessness intensity rating that is amenable to intervention. Currently, an arbitrary step rate or shuttle speed is chosen and subsequently increased or decreased depending on whether or not the individual completes all 3-minutes of exercise [62]. This pattern continues until the individual completes all 3-minutes of the step or shuttle test breathlessness intensity rating of ≥4 Borg 0-10 scale units. This algorithm, while functional, is burdensome and time consuming for both those administering and performing the tests. This indicates the need for an approach to predict the optimal step rate and shuttle speed for an individual with COPD based on data readily available in a standard medical record (e.g., age, sex, body mass, forced expiratory volume in 1-sec [$FEV_1$], modified Medical Research Council [mMRC] dyspnea rating).

The objective of the proposed research was twofold: 1) develop prediction equations to identify the stair stepping rate (3-min CRSST) and shuttle speed (3-min CSST) that provokes at least moderately intense breathlessness at the end of the 3-min test for an individual with COPD based on their unique personal and clinical characteristics; and 2) prospectively validate the newly developed equations in a cohort of adults with mild-to-very severe COPD. Based on the results from the prospective validation, we also sought to identify which of these tests, the 3-min CRSST or the 3-min CSST, and their respective prediction equations, may be better suited for use in clinical practice.

## 2.3. Methods

**Figure 2.1** outlines the methodological approach of this thesis, including both Objective 1 (derivation of prediction equations) and Objective 2 (prospective validation of prediction equations).

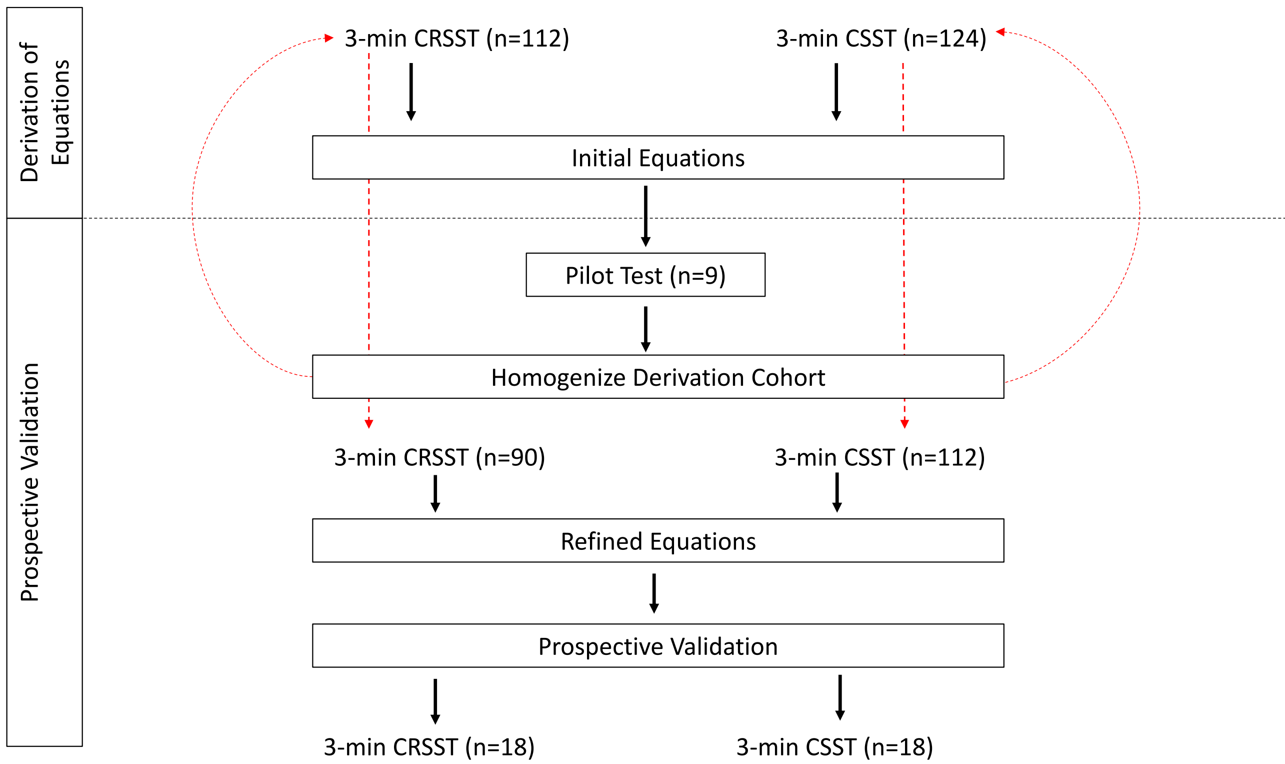



**Figure 2.1**. Layout for the methods section of this thesis.

### 2.3.1 Methods

*Derivation of Prediction Equations (Objective 1)*

2.3.1.1. Participants. This retrospective analysis included ambulatory and clinically stable adults aged >40 years with at least mild COPD according to GOLD criteria [7] who participated in earlier studies by Perrault et al. [36] (n=43), Sava et al. [34] (n=39), Borel et al. [35] (n=40) and Tracey et al. [41] (n=48), had spirometry data available and completed all 3-minutes of at least one 3-min CRSST and/or at least one 3-min CSST.

2.3.1.2. Data Analysis. Using SAS 9.4 software, prediction equations for both step rate and shuttle speed during the 3-min CRSST and 3-min CSST were developed with the help of a biostatistician (Mrs. Pei Zhi Li). Patterns of association between the step rate/shuttle speed completed by each participant and the end-exercise Borg CR10 scale intensity rating were first assessed with scatterplots, profile plots and generalized additive models with three degrees of freedom. The generalized additive models indicated significant ($p<0.05$) linear and nonlinear trends between stepping rate/shuttle speed and end exercise breathlessness intensity ratings.

Mixed-effects linear models with compound symmetry structure to account for the within-subjects variance-covariance were then used to estimate predictors of step rate or shuttle speed. Models were fit to look at the change in step rate/shuttle speed including the within subject predictor (Borg CR10 breathlessness intensity) and between subject covariates (sex, age, height, body mass, BMI, mMRC, resting lung function test parameters). Linear, quadratic and cubic association between step rate/shuttle speed and breathlessness intensity were explored; no significant quadratic or cubic trends were observed, and these were removed from the final models.

Covariates and their interaction terms were included in the final models when significant ($p<0.05$). For the 3-min CRSST, four models were explored, with the following variables considered: Borg CR10 breathlessness intensity rating, participant sex (male/female), age (years), height (cm), body mass (kg), BMI ($kg/m^2$), mMRC dyspnea score (0-4), $FEV_1$ (L), $FEV_1$ (%predicted NHANES [64]) and $FEV_1$ (%predicted GLI [65]). Breathlessness intensity, sex, age, and mMRC dyspnea score were included in all four models, with these variables forming Model 1. $FEV_1$ (%predicted GLI) and body mass were added in Model 2; $FEV_1$ (%predicted NHANES) and body mass in Model 3; and $FEV_1$ (%predicted NHANES) and BMI in Model 4.

For the 3-min CSST, four models were explored, with the following variables considered: Borg CR10 breathlessness intensity rating, sex (male/female), age (years), height (cm), body mass (kg), BMI, mMRC dyspnea score, $FEV_1$ (L), $FEV_1$ (%predicted NHANES [64]) and $FEV_1$ (%predicted GLI [65]). Breathlessness intensity, sex, age, and mMRC dyspnea score were included in each model, with these variables forming Model 1, in addition to height, body mass and $FEV_1$ (%predicated GLI). BMI and $FEV_1$ (%predicated GLI) were added in Model 2; body mass, height and $FEV_1$ (%predicated NHANES) in Model 3, and; BMI and $FEV_1$ (%predicated NHANES) in Model 4.

Akaike Information Criterion (AIC) was used to assess model fit, with a smaller AIC value indicating a more optimal model fit. Final predictor variables and their interaction terms were retained in the final model when significant ($p<0.05$) and if their addition decreased the AIC value. The prediction equations were assessed using mean error (Mean Absolute Error, Root Mean Squared Error and Mean Percentage Error [MPE], where the MPE provides information on whether the model systematically underestimates [negative error] or overestimates [positive error]). Details of each analysis are provided in **Appendices A and B**.

### 2.3.2. Results

*Derivation of Prediction Equations (Objective 1)*

*2.3.2.1. Participant Characteristics*

The baseline characteristics of the derivation cohort for development of the 3-min CRSST and 3-min CSST equations are presented in **Table 2.1**. Participants were predominantly older males with moderate airflow obstruction, pulmonary gas trapping, lung hyperinflation and moderate activity-related breathlessness.

**Table 2.1.** Baseline characteristics of the 3-min constant rate stair stepping test (3-min CRSST) and 3-min constant speed shuttle test (3-min CSST) derivation cohorts.

| | **3-min CRSST** | | **3-min CSST** | |
|---|---|---|---|---|
| | **N [# of Tests]** | **Total (n=112)** | **N [# of Tests]** | **Total (n=124)** |
| Sex, male gender, n (%) | 112 | 79 (71) | 124 | 84 (68) |
| Age, in year | 112 | 66.6 ± 6.6 | 124 | 65.9 ± 6.7 |
| Height (cm) | 112 | 167.1 ± 9.7 | 124 | 166.4 ± 10.0 |
| Weight (kg) | 112 | 73.6 ± 16.3 | 124 | 74.3 ± 17.3 |
| BMI ($kg/m^2$) | 112 | 26.3 ± 5.0 | 124 | 26.7 ± 5.3 |
| GLI: Post-BD $FEV_1$ (% predicted) | 112 | 53.0 ± 18.1 | 124 | 50.2 ± 19.9 |
| Post-BD $FEV_1$ | 112 | 1.5 ± 0.5 | 124 | 1.4 ± 0.6 |
| $FEV_1$/ FVC (%) | 112 | 44.6 ± 10.9 | 124 | 44.1 ± 12.3 |
| TLC (% predicted) | 107 | 131.1 ± 32.2 | 119 | 130.6 ± 31.1 |
| FRC (% predicted) | 107 | 148.5 ± 51.4 | 118 | 149.5 ± 50.7 |
| RV (% predicted) | 107 | 164.4 ± 57.4 | 119 | 168.3 ± 58.2 |
| GLI: DLCO (% predicted) | 108 | 64.1 ± 18.4 | 120 | 59.7 ± 18.8 |
| mMRC (0-4) | 112 | 1.4 ± 0.9 | 124 | 1.7 ± 1.1 |
| Step Rate (steps/min) | 112 [262] | 20.7 ± 4.6 [12-32] | - | - |
| Walking Speed (km/hr) | - | - | 124 [228] | 3.4 ± 1.3 [1.5-6] |
| Breathlessness intensity (Borg CR10) | 112 | 3.6 ± 1.9 | 124 | 2.5 ± 2.0 |

Values are mean ± standard deviation unless otherwise specified. Repeat measures were used; some subjects had more complete step and walk tests than others and therefore contributed more data. **Abbreviations:** BMI, body mass index; BD, bronchodilator; TLC, total lung capacity; FRC, functional residual capacity; RV, residual volume; DLCO, diffusion capacity; IC, inspiratory capacity; mMRC, modified Medicine Research Council dyspnea scale; CR10, 0-10 category ratio scale. Breathlessness intensity ratings (Borg CR10) were taken upon test completion (i.e. end of the 3rd minute of exercise).

*2.3.2.2. 3-min CRSST*

Out of the four models explored to predict the breathlessness intensity response to the 3-min CRSST, model 2 was the best fit, with a mean absolute error of 2.98 and a mean percentage error of 2.2% (see **Appendix A**). The resultant prediction equation was:

Step Rate (steps/min) = 26.646 + (1.332 * Borg CR10 breathlessness intensity) + (3.394 * sex) – (0.162 * age) – (0.040 * body mass) + (0.053 * $FEV_1$%predicted GLI) – (1.407 * mMRC), *where male sex = 1 and female sex = 0.*

*2.3.2.3. 3-min CSST*

Out of the four models explored to predict the breathlessness intensity response to the 3-min CSST, model 2 was the best fit, with a mean absolute error of 0.93 and a mean percentage error of 15.2% (see **Appendix B**). The resultant prediction equation was:

Shuttle Speed (km/hr) = 4.935 + (0.358 * Borg CR10 breathlessness intensity) – (0.363 * sex) – (0.020 * age) – (0.047 * BMI) + (0.012 * $FEV_1$%predicted GLI) – (0.305 * mMRC), *where male sex = 1 and female sex = 0.*

**2.3.3. Methods**

*Prospective Validation of the Prediction Equations (Objective 2)*

2.3.3.1. Study design. In this single-centre, prospective, uncontrolled trial, eligible participants completed experimental testing during a single 1.5-2.0 hour visit to the McConnell Centre for Innovative Medicine of the McGill University Health Centre's Research Institute (RI-MUHC). The equations generated from Part 1 and reported in *sections 2.3.2.2 and 2.3.2.3* were pilot tested in a sample of nine people with COPD (**Figure 2.1**). This pilot test was conducted to evaluate preliminarily the prediction equations to see if they were successful at predicting step rates and shuttle speeds which allowed participants to complete all 3-minutes of the CRSST or CSST with a breathlessness intensity rating of ≥3 Borg CR10 units upon completion; this provided an opportunity to refine the initial equations prior to their prospective validation in 100 people with COPD (**Figure 2.1**).

2.3.3.2. Participants. Men and women aged ≥40 years with GOLD I-IV COPD [1] were recruited from: the COPD outpatient clinic at the Montreal Chest Institute (MCI); the outpatient pulmonary rehabilitation program of the MCI; and *via* contact with people with COPD who consented to be enlisted in a database of prospective research study participants. Participants were included in this study if their post-bronchodilator spirometry test results (i.e., $FEV_1$ and $FEV_1$-to-forced vital capacity ratio [$FEV_1$/FVC]) were available in their electronic medical record, were collected within one year of study participation, and met GOLD diagnostic criteria for COPD (i.e., $FEV_1$/FVC <0.70). Exclusion criteria were: inability to read and write in English or French; change in medication dosage and/or frequency of administration in preceding two weeks; exacerbation and/or hospitalization in preceding six weeks; history of asthma; presence of medical condition(s) other than COPD that were unstable and/or that could contribute to breathlessness and exercise intolerance, including neuromuscular and/or musculoskeletal disease(s); and/or use of domiciliary oxygen. The Biomedical Research Ethics Committee of the RI-MUHC approved the study protocol and consent form (2020-5914). Written informed consent was obtained from all participants before any assessments were performed.

2.3.3.3. Protocol. Participant's age, height and body mass were first recorded. Each participant then completed health status and symptom burden questionnaires followed by a 3-min CRSST and a 3-min CSST at the stepping rate and shuttle speed estimated from the newly developed equations. Step and shuttle tests were in a randomized for each participant and separated by a rest period of

15-30 minutes. In this way, step and shuttle tests (and their respective prediction equations) were administered in the way that we believed they would most likely be used in clinical practice and in large clinical trials.

2.3.3.4. Data Extraction. Pulmonary function test data (spirometry, plethysmographic lung volumes, pulmonary diffusing capacity for carbon monoxide) was extracted from the participant's medical records. Each participant's $FEV_1$ was expressed as a percentage of their age, sex and height predicted normal value [65].

2.3.3.5. Anthropometry, health status and symptom burden. Height and body mass were measured using an electric scale with stadiometer. Health status and symptom burden were assessed using the oxygen cost diagram [66], baseline dyspnea index [24], COPD assessment test [67], and mMRC dyspnea scale [12].

2.3.3.6. 3-min constant rate stair stepping (3-min CRSST) and 3-min constant speed shuttle tests (3-min CSST). The 3-min CRSST consisted of a single 3-min bout of exercise performed at the step rate identified by the newly developed prediction equation *(see section 2.3.2.2*) needed to provoke a breathlessness intensity rating of 4 Borg CR10 scale units (*see section 2.3.3.7*) for a given participant. Participants were instructed to "step-up" (indicating to place both feet up onto a single step, one foot after the other) and "step-down" (indicating to step back down to the floor, one foot after the other) a single 20-cm high step in response to audible instructions for the entire 3-minutes of the test or until they became symptom-limited [36, 62].

The 3-min CSST consisted of a single 3-min bout of exercise performed at the shuttle speed identified by the newly developed prediction equation *(*see *section 2.3.2.3)* needed to provoke a breathlessness intensity rating of 4 Borg CR10 scale units (see *section 2.3.3.7*) for a given participant. Participants were instructed to start walking on hearing a 'beep' generated by an audio signal (indicating to start walking between two pylons placed 9m apart); and to walk at a constant speed that allowed them to reach the pylons upon each 'beep' of the audio signal for the entire 3-minutes of the test or until they became symptom-limited [34, 36, 62].

Using Borg's CR10 scale [62], participants rated the intensity (*"how strong"*) and unpleasantness (*"how bad"*) of their breathlessness, as well as the intensity of their leg discomfort at rest and at the end of each 1-minute interval of exercise (or at the point of symptom-limitation if they stopped before completing all 3-min of the test). Participants who stopped exercise prior to the end of 3$^{rd}$ minute due to symptom-limitation were asked to identify their main reason(s) for stopping exercise (breathlessness, leg discomfort, combination of breathlessness and leg discomfort, other) and to quantify the percentage contribution of breathlessness and leg discomfort to exercise cessation. Heart rate and peripheral oxygen saturation ($SpO_2$) were monitored continuously at rest and during exercise by portable finger-tip pulse oximeter.

Participants were asked to fill out a questionnaire prior to completing the 3-min CRSST and 3-min CSST to identify the activity (or activities) that induce breathlessness in their daily life as well as to identify which of those activities are the most difficult (**Appendix D**). After completing both the 3-min CRSST and 3-min CSST, participants were asked to complete another questionnaire to identify which test they preferred, which test was more difficult, and if each exercise test provided an accurate representation of the breathlessness they experience on a daily basis (**Appendix E**). These measures were obtained to differentiate between the 3-min CRSST and 3-min CSST in regard to participant preferences and ecological validity.

2.3.3.7. Determining stepping rate/shuttle speed and target breathlessness intensity rating. The increments in exercise intensity used for the 3-min CRSST and 3-min CSST were 2 steps/min [range: 10-30 steps/min] and 0.5 km/hr [range: 1.5-9 km/hr]. When the predicted step rate/shuttle speed fell between these increments, the exercise intensity used for testing was rounded up to the nearest 2 steps/min or 0.5 km/hr increment. This decision was made based on the assumption that it was better to over predict the step rate/shuttle speed and elicit a higher breathlessness intensity response, than it was to under predict and elicit a breathlessness intensity response not likely amendable to intervention (i.e., <3 Borg CR10 scale units).

After pooling the mean results from 20 randomized controlled (therapeutic) trials in people with mild-to-very severe COPD (n=731), we calculated a median and mean breathlessness intensity rating during constant work rate treadmill or cycle ergometry exercise at isotime (i.e., highest submaximal exercise time completed by a given participant following active and control/placebo interventions) of ~5 Borg CR10 scale units (severe) in the placebo arm of the trials (see **Appendix C)** [21, 76-99]. The preliminary results of a study from our laboratory by Tracey et al. [41] indicated that the 3-min CRSST and 3-min CSST are responsive to detect relief of exertional breathlessness following 8-12 weeks of pulmonary rehabilitation in people with COPD provided breathlessness intensity ratings during the initial (baseline or pre PR) assessment are >3 Borg CR10 units. The published results of Maltais et al. [62] indicated that the 3-min CSST is responsive to detect relief of exertional breathlessness when breathlessness intensity ratings during the initial (baseline or pre-treatment) assessment are ≥4 Borg CR10 units. For the purpose of this study, we decided to target a breathlessness intensity rating of 4 Borg CR10 scale units (somewhat severe) as it was the targeted value used most recently by Maltais et al. [62]. However, we thought that it was reasonable to define a breathlessness intensity rating of ≥3 Borg CR10 scale units as successful when prospectively validating the 3-min CSST and 3-min CRSST prediction equations.

2.3.3.8. Definition of successful performance of the equation to identify exercise intensity. Although a breathlessness intensity rating of 4 Borg CR10 units was used in our equations to identify the exercise intensity (step rate or shuttle speed) for an individual with COPD, the prediction equations were considered "successful" when a participant completed all three minutes of exercise and rated their breathlessness intensity ≥3 Borg CR10 units at the end of 3-min. The prediction equations were, therefore, considered "unsuccessful" when a participant: 1) completed all 3-min of exercise but rated their breathlessness intensity <3 Borg CR10 units at end exercise; ***or*** 2) did not complete all 3-min of exercise, regardless of the breathlessness intensity rating at the point of exercise cessation.

2.3.3.9. Statistical Analyses. Participants were described by basic demographic and health characteristics. For both the 3-min CRSST and 3-min CSST, participants were grouped according to successful or unsuccessful completion of the tests, and reason for unsuccessful completion:

1. Completing all 3-min of exercise at the predicted stepping rate or shuttle speed with a breathlessness intensity rating of ≥3 Borg CR10 units at end exercise (successful);

2. Completing all 3-min of exercise at the predicted stepping rate or shuttle speed with a breathlessness intensity rating of <3 Borg CR10 units at end exercise (unsuccessful, reason 1);
3. Failing to complete all 3-min of exercise at the predicted stepping rate or shuttle speed with a breathlessness intensity rating of ≥3 Borg CR10 units at end exercise (unsuccessful, reason 2);
4. Failing to complete all 3-min of exercise at the predicted stepping rate or shuttle speed with a breathlessness intensity rating of <3 Borg CR10 units at end exercise (unsuccessful, reason 3).

Equations developed by Lewthwaite et al. [68] (*see below*) were used to estimate the rate of $O_2$ consumption (estV'$O_2$) at end exercise for the 3-min CRSST and 3-min CSST, and the ratio of breathlessness intensity to estV'$O_2$ was calculated. The equations used were:

*3-min CRSST end-exercise estV'$O_2$ (L/min)* = [0.015286 * body mass (kg)] + [0.035605 * step rate (steps/min)] – 0.698449

*3-min CSST end-exercise* est*V'$O_2$ (L/min)* = [0.012039 * body mass (kg)] + [0.217654 * shuttle speed (km/hr)] – 0.691760

Frequency and descriptive statistics were then used to describe groups by baseline characteristics, breathlessness intensity rating at end exercise, and breathlessness/estV'$O_2$ ratio at end exercise. Paired samples t-tests were used to compare physiological and symptom responses between the 3-min CRSST and 3-min CSST (HR, $SpO_2$, breathlessness intensity, unpleasantness and leg discomfort ratings at end-exercise, estV'$O_2$, breathlessness/estV'$O_2$ ratio). Pearson's correlation coefficients were calculated to explore associations between breathlessness intensity, unpleasantness and leg discomfort ratings at end-exercise for the 3-min CRSST and the 3-min CSST; these correlative analyses were only performed using data from those participants who completed all 3-mintues of both step and shuttle tests (n=15).

Participant responses to the following questions were reported as proportion of respondents: (i) 'Which test did you find more difficult?'; (ii) 'Which test did you prefer?'; and (iii) 'Which test best mimics the breathlessness you experience in daily life?'. The proportion of participants who listed walking up a hill, walking on a flat surface and stair climbing as activities which (i) induce breathlessness; (ii) are the most difficult ADL; (iii) are the 2nd most difficult ADL, and; (iv) are the 3rd most difficult ADL were reported.

Statistical significance was set at $p<0.05$ and data are reported as mean ± standard deviation (SD) unless otherwise stated. Data were analyzed using IBM SPSS Statistics (version 24) software.

**2.3.4. Results**

*Pilot Test of Initial Equations*

*2.3.4.1. Participant Characteristics*

The baseline characteristics of the nine pilot test participants and according to outcome groups for the 3-min CRSST and 3-min CSST are presented in **Table 2.2, 2.3 & 2.4**, respectively. Participants were predominantly older individuals with severe airflow obstruction, pulmonary gas trapping and mild-to-moderately severe breathlessness during their daily lives

**Table 2.2.** Pilot Test: participant characteristics.

| | | All |
|---|---|---|
| n (M: F) | | 9 (5: 4) |
| Age (years) [range] | | 69.8 ± 9.3 [51-85] |
| BMI (kg/m$^2$) [range] | | 25.5 ± 2.7 [20.8-29.1] |
| $FEV_1$ %predicted [range] | | 41.1 ± 20.1 [17.4-81] |
| $FEV_1$/FVC (%) | | 47.9 ± 23.3 [1.8-78.8] |
| mMRC dyspnea (n) [%] | | |
| | 0 | 0 [0] |
| | 1 | 5 [56] |
| | 2 | 1 [11] |
| | 3 | 3 [33] |
| | 4 | 0 |

Values are mean ± SD unless otherwise stated. **Abbreviations**: BMI, body mass index; $FEV_1$, forced expiratory volume in 1-sec; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Disease; mMRC, modified Medical Research Council Dyspnea scale; CAT, COPD assessment test; BDI, baseline dyspnea index.

**Table 2.3.** Pilot Test: participant characteristics according to outcome group for the 3-min CRSST.

| | Completed 3-CRSST Borg ≥3 | Completed 3-CRSST Borg <3 | Did not Complete 3-CRSST Borg ≥3 | Did not Complete 3-CRSST Borg <3 |
|---|---|---|---|---|
| n (M: F) | 2 (0:2) | 1 (1:0) | 5 (3:2) | 1 (1:0) |
| Age (years) | 68.5 ± 3.54 | 72 | 70.8 ± 12.62 | 65 |
| BMI (kg/m$^2$) | 24.5 ± 5.2 | 29.1 | 24.9 ± 1.6 | 27.0 |
| $FEV_1$ %predicted | 36.7 ± 12.1 | 81.0 | 36.7 ± 17.8 | 32.1 |
| $FEV_1$/FVC (%) | 43.4 ± 2.6 | 78.8 | 41.4 ± 27.4 | 58.8 |
| mMRC dyspnea | 2.0 ± 1.4 | 2 | 1.8 ± 1.1 | 1 |

Values are mean ± SD unless otherwise stated. **Abbreviations**: BMI, body mass index; $FEV_1$, forced expiratory volume in 1-sec; FVC, forced vital capacity; mMRC, modified Medical Research Council Dyspnea scale.

**Table 2.4.** Pilot Test: participant characteristics according to outcome group for the 3-min CSST.

| | Completed 3-CRSST Borg ≥3 | Completed 3-CRSST Borg <3 | Did not Complete 3-CRSST Borg ≥3 |
|---|---|---|---|
| n (M: F) | 5 (2:3) | 3 (3:0) | 1 (0:1) |
| Age (years) | 71.8 ± 8.0 | 64.0 ± 11.4 | 77 |
| BMI (kg/m²) | 25.4 ± 2.9 | 25.7 ± 3.4 | 25.5 |
| $FEV_1$ %predicted | 42.4 ± 14.0 | 41.5 ± 34.5 | 33.6 |
| $FEV_1$/FVC (%) | 37.9 ± 21.4 | 60.8 ± 26.2 | 59.6 |
| mMRC dyspnea | 1.8 ± 1.1 | 1.3 ± 0.6 | 3 |

Values are mean ± SD unless otherwise stated. **Abbreviations**: BMI, body mass index; $FEV_1$, forced expiratory volume in 1-sec; FVC, forced vital capacity; mMRC, modified Medical Research Council Dyspnea scale.

*2.3.4.2. 3-min CRSST (Pilot)*

Of the nine participants who performed the step test, the predicted step rate was considered successful for two (22%) participants and unsuccessful for seven (78%) participants (**Table 2.3**). On average, the prediction equation for the 3-min CRSST over predicted step rates (20.2 ± 2.9 [16-24] steps/min) with an average test duration of 2:18 min (range: 1:05-3:00 min).

*2.3.4.3. 3-min CSST (Pilot)*

Of the nine participants who performed the shuttle test, the predicted shuttle speed was considered successful for five (56%) participants and unsuccessful for four (44%) participants (Table 2.4). On average, the prediction equation for the 3-min CSST under predicted shuttle speeds (3.8 ± 0.4 [3.5-4.5] km/hr) with an average test duration of 2:56 min (range: 2:26-3:00 min).

### 2.3.5. Methods

*Development of the 2nd Equations based on Results from the Pilot Test (Objective 1)*

2.3.5.1. Refinement of the derivation cohort. The results of the pilot test indicated a need to refine the prediction equations to increase their ability to successfully identify step rates and shuttle speeds for the 3-min CRSST and 3-min CSST in people with COPD (**Figure 2.1**). After a consultation with clinical scientists and MSc thesis advisor committee members, Dr. Jean Bourbeau and Dr. Benjamin Smith, we decided to homogenize the derivation cohort by excluding people who reported being asymptomatic, as indicated by a mMRC dyspnea rating of 0. The 3-min CRSST and 3-min CSST are exercise tests designed to assess the symptom of breathlessness and therefore people without self-reported symptoms are not who these tests were designed for, nor are they the people who would most likely be performing these tests. This refinement led to a total of 22 people being removed from the 3-min CRSST cohort, changing the sample size from 112 to 90; and a total of 12 people being removed from the 3-min CSST cohort, changing the sample size from 124 to 112 (**Figure 2.1**).

2.3.5.2. Outcome measure of the equations. For the original equations, we used step rate/shuttle speed as the outcome variable and Borg CR10 breathlessness intensity rating as the predictor variable in hopes of creating a tool that would allow researchers and clinicians to accurately target specific breathlessness intensity ratings in people with COPD. However, after pilot testing, we decided to develop the equations with step rate/shuttle speed as the predictor variable and Borg CR10 breathlessness intensity rating as the outcome variable. This modification allowed us to assess how each variable in the equation affects breathlessness intensity and more importantly, the equations reflect the primary outcome measure of the 3-min CRSST and 3-min CSST –

breathlessness intensity. After the equations were recreated, they were rearranged them so that step rate/shuttle speed was the isolated variable to make the equations simpler to use (i.e. the target Borg CR10 breathlessness intensity rating was inputted into the equation, and the step rate/shuttle speed needed to elicit that level of breathlessness was calculated).

2.3.5.3. Models for the new equations. The methods for developing the 2nd set of equations were the same as described in Methods Part 1 (see *section 2.3.1*). Some of the variables that came out significant in the univariate analysis differed from the original derivation cohort and therefore the equation models were also different. The models considered for the 2nd set of equations were as follows: Four models were analyzed for the CRSST, with the following variables considered for use in each model: step rate (steps/min), sex (male/female), age (years), height (cm), body mass (kg), BMI (kg/$m^2$), mMRC dyspnea score (1-4), GOLD stage (1-4), $FEV_1$ (L), $FEV_1$ (%predicted NHANES [64]) and $FEV_1$ (%predicted GLI [65]). Step rate, sex, age, body mass and mMRC dyspnea score were included in every model, with the addition of $FEV_1$ (L) in Model 1; $FEV_1$ (%predicted GLI) in Model 2; $FEV_1$ (%predicted NHANES) in Model 3 and; GOLD stage in Model 4. Details this analysis are provided in **Appendix F**.

Five models were analyzed for the CSST, with the following variables considered for use in each model: shuttle speed (km/hr), sex, age, height, body mass, BMI, mMRC dyspnea score, GOLD stage (1-4), $FEV_1$ (L), $FEV_1$ (%predicted NHANES [64]) and $FEV_1$ (%predicted GLI [65]). Shuttle speed, sex, age, body mass and mMRC dyspnea score were included in every model, with the addition of $FEV_1$ in Model 1; nothing in Model 2; BMI and $FEV_1$ in Model 3; BMI and $FEV_1$

(%predicted NHANES) in Model 4; and BMI and $FEV_1$ (%predicated GLI) in Model 5. Details of this analysis are provided in **Appendix G**.

### 2.3.6. Results

*Development of the 2nd Prediction Equations based on Results of the Pilot Test (Objective 1)*

*2.3.6.1. Participant characteristics*

The participant characteristics of the refined derivation cohort for both the 3-min CRSST and 3-min CSST prediction equations are presented in **Table 2.5**. Participants were predominantly older males with moderate-to-severe airflow obstruction, pulmonary gas trapping, lung hyperinflation and moderately severe activity-related breathlessness.

**Table 2.5.** Baseline characteristics of the 3-min CRSST and 3-min CSST derivation cohorts.

| | **3-min CRSST** | | **3-min CSST** | |
|---|---|---|---|---|
| | **N [# of Tests]** | **Total (n=90)** | **N [# of Tests]** | **Total (n=112)** |
| Sex, male gender, n (%) | 90 | 67 (74.4) | 112 | 78 (69.6) |
| Age, in year | 90 | 66.6 ± 6.8 | 112 | 65.8 ± 6.9 |
| Height (cm) | 90 | 166.7 ± 9.5 | 112 | 166.2 ± 10.1 |
| Body mass (kg) | 90 | 74.1 ± 16.7 | 112 | 74.8 ± 17.4 |
| BMI (kg/m$^2$) | 90 | 26.6 ± 5.1 | 112 | 27.0 ± 5.4 |
| GLI: Post-BD $FEV_1$ (% predicted) | 90 | 57.2 ± 20.5 | 112 | 53.2 ± 21.8 |
| Post-BD $FEV_1$ | 90 | 1.5 ± 0.5 | 112 | 1.4 ± 0.6 |
| $FEV_1$/FVC (%) | 90 | 44.7 ± 11.2 | 112 | 43.8 ± 12.5 |
| GOLD stage | 90 | | 112 | |
| GOLD 1, n (%) | | 10 (11.1) | | 11 (9.8) |
| GOLD 2, n (%) | | 44 (48.9) | | 49 (43.8) |
| GOLD 3, n (%) | | 31 (34.4) | | 37 (33.0) |
| GOLD 4, n (%) | | 5 (5.6) | | 15 (13.4) |
| TLC (% predicted) | 87 | 132.2 ± 33.2 | 108 | 131.2 ± 31.8 |
| FRC (% predicted) | 87 | 149.6 ± 52.6 | 107 | 150.6 ± 52.0 |
| RV (% predicted) | 87 | 164.1 ± 59.1 | 108 | 169.8 ± 60.0 |
| GLI: $D_LCO$ (% predicted) | 86 | 64.6 ± 25.9 | 108 | 59.3 ± 19.0 |
| mMRC (1-4) | 90 | 1.6 ± 0.8 | 112 | 1.9 ± 0.9 |
| Step rate (steps/min) | 90 [203] | 17.6 ± 2.7 [12-32] | - | - |
| Shuttle speed (km/hr) | - | - | 112 [204] | 2.9 ± 1.2 [1.5-6] |
| Breathlessness intensity at end-exercise (Borg CR10) | 90 | 3.7 ± 1.7 | 112 | 2.5 ± 2.0 |

Values are mean ± standard deviation unless otherwise specified. Repeat measures were used; some subjects had more complete step and shuttle tests than others and therefore contributed more data. **Abbreviations:** BMI, body mass index; BD, bronchodilator; TLC, total lung capacity; FRC, functional residual capacity; RV, residual volume; DLCO, diffusion capacity; IC, inspiratory capacity; mMRC, modified Medicine Research Council dyspnea scale; CR10, 0-10 category ratio scale. Breathlessness (Borg CR10) ratings were taken upon test completion (i.e. end of the 3$^{rd}$ minute of exercise).

### *2.3.6.2. 3-min CRSST*

Out of the four models for the step test, model 4 was the best fit, with a mean absolute error of 1.22 and a mean percentage error of -26.9% (see **Appendix F**). The resultant prediction equation was:

Breathlessness intensity rating (Borg CR10) = -13.918 + (0.653 * step rate) – (1.078 * sex) + (0.155 * age) – (0.006 * age * steps) + (0.028 * body mass) + (0.635 * mMRC) + GOLD*, where male sex = 1 and female sex = 0; and GOLD 1 = 0, GOLD 2 = 0.450, GOLD 3 = 1.617, and GOLD 4 = 0.732.*

*2.3.6.3. 3-min CSST*

Out of the five models for the shuttle test, model 3 was the best fit, with a mean absolute error of 1.28 and a mean percentage error of -49.2% (see **Appendix G**). The resultant prediction equation was:

Breathlessness intensity rating (Borg CR10) = -2.530 – (0.619 * shuttle speed) + (0.206 * shuttle speed$^2$) – (0.366 * sex) + (0.031 * age) + (0.094 * BMI) – (0.667 * $FEV_1$) + (0.748 * mMRC), *where male sex = 1 and female sex = 0.*

### 2.3.7. Methods

*Prospective Validation of the Refined Equations (Objective 2)*

2.3.7.1. Methodological changes after pilot testing. The results from the pilot test showed that the prediction equations consistently overestimated step rates and underestimated shuttle speeds (**Table 2.3 & 2.4**). Based on these results, we decided to change the way in which we rounded the predicted cadences from rounding *up* to the nearest 2 steps/min or 0.5 km/hr increment (i.e. number only ever rounded up), to rounding to the *nearest* 2 steps/min or 0.5 km/hr increment (i.e. number could be rounded up *or* down). In this way, we would be testing the step rate and shuttle speed that most closely targeted a breathlessness intensity rating of 4 Borg CR10 units. In our view, this revised strategy would help to ensure that breathlessness intensity ratings at the end of each 3-min exercise test were ≥3 Borg CR10 units, which Tracey et al. [41] identified as important for both the 3-min CRSST and 3-min CSST to be able to detect a change in exertional breathlessness following pulmonary rehabilitation in people with COPD.

For details on the prospective validation methods, refer to *section 2.3.3* above.

## 2.4. Results

*Prospective Validation of the Prediction Equations (Objective 2)*

### *2.4.1. Participant characteristics*

Participant characteristics are presented in **Table 2.6**. Participants were older adults with severe airflow obstruction, pulmonary gas trapping, impaired health status, and severe breathlessness during their activities of daily life, as indicated by mean mMRC dyspnea and BDI focal score ratings of 2.6 and 5.1, respectively.

**Table 2.6**. Baseline characteristics of the study participants.

| | | All |
|---|---|---|
| n (M: F) | | 18 (11: 7) |
| Age (years) [range] | | 69.3 ± 7.4 [52-85] |
| BMI (kg/m$^2$) [range] | | 27.0 ± 3.5 [17.6-32.8] |
| $FEV_1$ %predicted [range] | | 45.9 ± 17.2 [17.2-79.7] |
| $FEV_1$/FVC (%) | | 48.8 ± 15.3 |
| GOLD Stage (n) [%] | | |
| | 1 | 0 |
| | 2 | 6 [33] |
| | 3 | 8 [44] |
| | 4 | 4 [22.2] |
| mMRC dyspnea (n) [%] | | 2.6 ± 1.2 |
| | 0 | 1 [5.5] |
| | 1 | 3 [16.7] |
| | 2 | 2 [11] |
| | 3 | 8 [44] |
| | 4 | 4 [22] |
| CAT total score | | 18.3 ± 8.2 |
| BDI focal score | | 5.7 ± 3.5 |
| *Pack years [range] | | 36.1 ± 20.1 [2-78] |

Values are mean ± standard deviation unless otherwise specified.
**Abbreviations**: BMI, body mass index; $FEV_1$, forced expiratory volume in 1-sec; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Disease; mMRC, modified Medical Research Council Dyspnea scale; CAT, COPD assessment test; BDI, baseline dyspnea index. *n=16

### *2.4.2. Success of the 3-min CRSST and 3-min CSST prediction equations*

The equations for the 3-min CRSST and 3-min CSST were equally successful in that 61% of the time the equations predicted step rates/shuttle speeds that resulted in successful tests (**Fig. 2.2 & 2.3**). Thirty-nine percent of the time the step rates/shuttle speeds resulted in unsuccessful tests (**Fig. 2.2**); some participants finished all 3-minutes of step (n=6) or shuttle exercise (n=4) with a breathlessness intensity rating <3, while other participants were unable to complete all 3-minutes of step (n=1) or shuttle (n=3) exercise test due to severe breathlessness (**Fig. 2.3**).

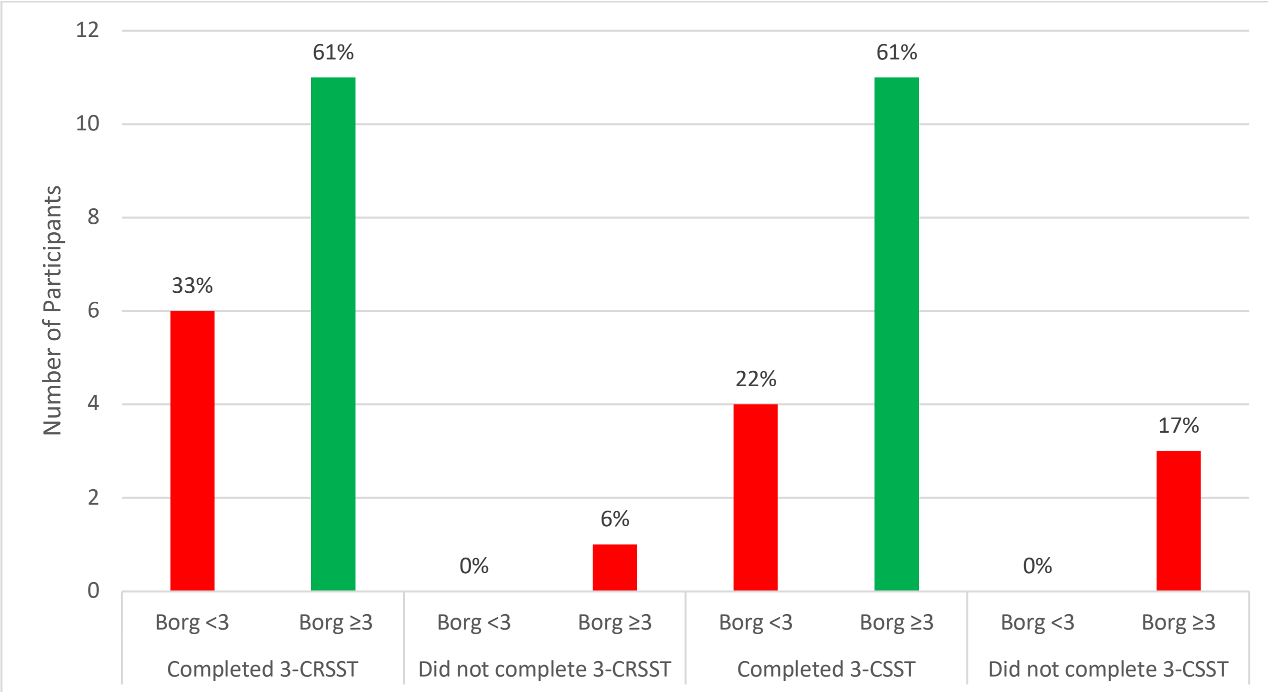



**Fig. 2.2**. Outcome Groups for the 3-min CRSST and 3-min CSST.

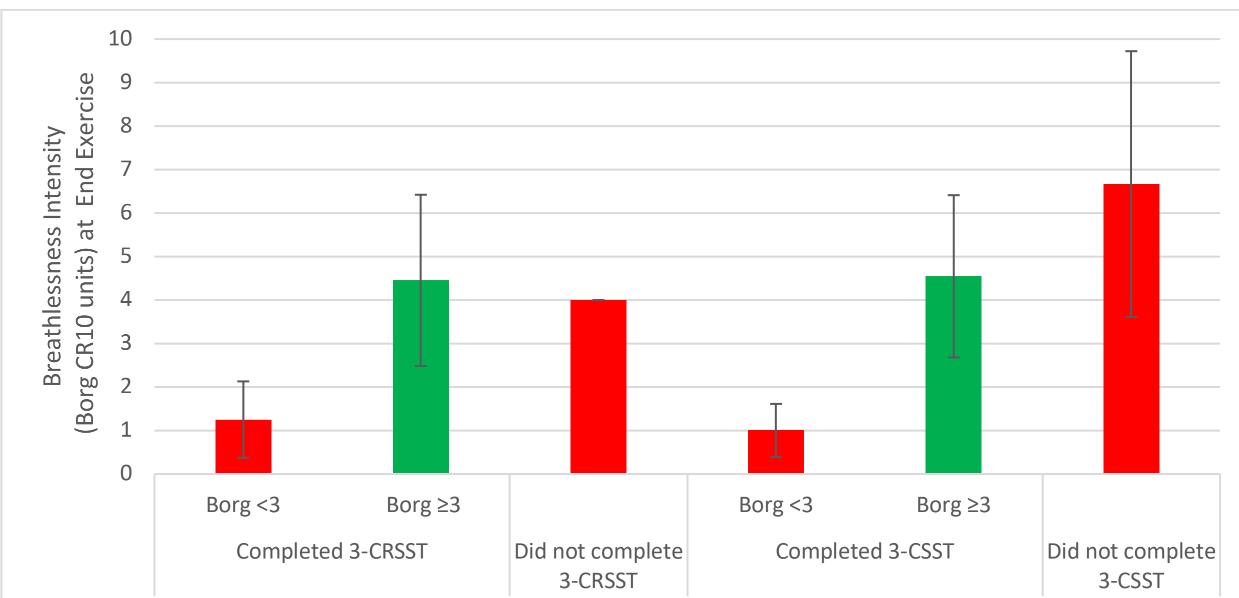



**Fig. 2.3**. Average Borg CR10 Breathlessness Intensity Rating at End Exercise in each Outcome Group for the 3-min CRSST and 3-min CSST.

### *2.4.3. Characteristics of each outcome group for the 3-min CRSST and 3-min CSST*

Participant characteristics according to outcome groups for the 3-min CRSST and 3-min CSST are presented in **Table 2.7 and 2.8**, respectively. On average, the participants that completed all 3-min of exercise with a breathlessness intensity rating of ≥3 Borg CR10 units (i.e., successful group) had more severe chronic airflow obstruction with greater symptom burden than those who completed all 3-min of exercise with a breathlessness intensity rating <3 Borg CR10 units (**Tables 2.7 and 2.8**). The few participants who did not complete all 3-min of step (n=1) or shuttle exercise (n=3) with breathlessness intensity ratings ≥3 Borg CR10 units were those with the most severe chronic airflow obstruction and particularly notable cigarette smoking histories (**Tables 2.7 and 2.8**).

**Table 2.7.** Baseline participant characteristics according to outcome group for the 3-min CRSST.

| | | Completed 3-CRSST Borg ≥3 | Completed 3-CRSST Borg <3 | Did not Complete 3-CRSST Borg ≥3 |
|---|---|---|---|---|
| n (M: F) | | 11 (6:5) | 6 (4:2) | 1 (1:0) |
| Age (years) | | 69.2 ± 8.9 | 69.5 ± 5.2 | 70 |
| BMI (kg/m$^2$) | | 26.6 ± 4.0 | 28.4 ± 2.0 | 22.8 |
| $FEV_1$ %predicted | | 43.9 ± 18.0 | 52.3 ± 15.6 | 29.8 |
| $FEV_1$/FVC (%) | | 50.3 ± 17.3 | 51.1 ± 6.1 | 25 |
| GOLD Stage (n) [%] | | - | - | - |
| | 1 | 0 | 0 | 0 |
| | 2 | 4 [22] | 2 [11] | 0 |
| | 3 | 4 [22] | 4 [22] | 0 |
| | 4 | 3 [16.7] | 0 | 1 [5.5] |
| mMRC dyspnea (n) [%] | | 3.1 ± 0.8 | 2.0 ± 1.4 | 1 |
| | 0 | 0 | 1 [5.5] | 0 |
| | 1 | 1 [5.5] | 1 [5.5] | 1 [5.5] |
| | 2 | 0 | 2 [11] | 0 |
| | 3 | 7 [39] | 1 [5.5] | 0 |
| | 4 | 3 [16.7] | 1 [5.5] | 0 |
| CAT total score | | 20.9 ± 6.7 | 13.3 ± 9.5 | 20 |
| BDI focal score | | 4.2 ± 2.6 | 8.5 ± 3.7 | 6 |
| *Pack years | | 28.1 ± 19.7 | 36.7 ± 16.5 | 78 |
| Exacerbations (12mo) | | 0.5 ± 0.8 | 1.2 ± 1.3 | 0 |
| Dyspnea/ estV'$O_2$ Ratio | | 6.1 ± 5.0 | 1.4 ± 1.1 | 3.8 |
| Locus of Symptom Limitation (n) | | | | |
| | Breathlessness | - | - | 0 |
| | Leg Discomfort | - | - | 0 |
| | Combination | - | - | 1 |

Values are mean ± standard deviation unless otherwise specified. **Abbreviations**: BMI, body mass index; $FEV_1$, forced expiratory volume in 1-sec; FVC, forced vital capacity; mMRC, modified Medical Research Council Dyspnea scale; CAT, COPD assessment test; BDI, baseline dyspnea index; estV'$O_2$, estimated rate of $O_2$ consumption. *n=16

**Table 2.8.** Baseline participant characteristics according to outcome group for the 3-min CSST.

| | | Completed 3-CRSST Borg ≥3 | Completed 3-CRSST Borg <3 | Did not Complete 3-CRSST Borg ≥3 |
|---|---|---|---|---|
| n (M: F) | | 11 (6:5) | 4 (3:1) | 3 (2:1) |
| Age (years) | | 68.5 ± 8.5 | 69.8 ± 7.4 | 72 ± 1.7 |
| BMI (kg/m$^2$) | | 26.5 ± 4.0 | 29.5 ± 0.9 | 25.3 ± 2.8 |
| $FEV_1$ %predicted | | 44.7 ± 15.0 | 58.6 ± 20.0 | 33.3 ± 14.9 |
| $FEV_1$/FVC (%) | | 48.2 ± 10.1 | 62.7 ± 19.2 | 34.8 ± 12.5 |
| GOLD Stage (n) [%] | | | | |
| | 1 | 0 | 0 | 0 |
| | 2 | 4 [22] | 2 [11] | 0 |
| | 3 | 5 [28] | 2 [11] | 1 [5.5] |
| | 4 | 2 [11] | 0 | 2 [11] |
| mMRC dyspnea (n) [%] | | 2.7 ± 1.4 | 2.5 ± 0.6 | 2.3 ± 1.2 |
| | 0 | 1 [5.5] | 0 | 0 |
| | 1 | 2 [11] | 0 | 1 [5.5] |
| | 2 | 0 | 2 [11] | 0 |
| | 3 | 4 [22] | 2 [11] | 2 [11] |
| | 4 | 4 [22] | 0 | 0 |
| CAT total score | | 19.6 ± 8.1 | 11.5 ± 5.7 | 22.7 ± 7.4 |
| BDI focal score | | 4.9 ± 3.6 | 8.0 ± 3.8 | 5.7 ± 2.5 |
| *Pack years | | 23.8 ± 15.8 | 43.3 ± 17.8 | 65.5 ± 17.7 |
| Exacerbations (12mo) | | 0.7 ± 1.1 | 1.0 ± 1.2 | 0.3 ± 0.6 |
| Breathlessness/ estV'$O_2$ Ratio | | 4.7 ± 2.8 | 1.0 ± 0.8 | 7.5 ± 4.9 |
| Locus of Symptom Limitation (n) | | | | |
| | Breathlessness | - | - | 1 |
| | Leg Discomfort | - | - | 0 |
| | Combination | - | - | 2 |

Values are mean ± standard deviation unless otherwise specified. **Abbreviations**: BMI, body mass index; $FEV_1$, forced expiratory volume in 1-sec; FVC, forced vital capacity; mMRC, modified Medical Research Council Dyspnea scale; CAT, COPD assessment test; BDI, baseline dyspnea index; estV'$O_2$, estimated rate of $O_2$ consumption. *n=16

### *2.4.4. Physiological and Perceptual Responses to the Predicted Step Rates and Walk Speeds*

The physiological response observed at end exercise for the 3-min CRSST and 3-min CSST were similar. The HR at end-exercise was 99 ± 12 bpm (range: 79-125 bpm) for the step test and 103 ± 14 bpm (range: 70-125 bpm) for the shuttle test (p=0.42). The $SpO_2$ at end-exercise was 91 ± 4% (range: 81-95%) for the step test and 91 ± 4% (range: 78-96%) for the shuttle test (p=0.30).

$SpO_2$ decreased from rest to end-exercise by an average of 4% and 5% during the step and shuttle test, respectively ($p=0.46$).

The perceptual response observed at end exercise for the 3-min CRSST and 3-min CSST were similar. Within subject breathlessness intensity ratings ($r = 0.90$ $p < 0.001$), breathlessness unpleasantness ratings ($r = 0.88$, $p < 0.001$), and intensity ratings of leg discomfort ($r = 0.58$, $p=0.023$) were positively correlated between the step and shuttle test in those 15 participants who completed all 3-minutes of both tests (**Fig. 2.4**).

There was a significant difference between estV'$O_2$ for the step test (0.92 ± 0.28 L/min) compared to the shuttle test (1.05 ± 0.26 L/min; $p=0.001$). That is, the shuttle speed predicted to elicit a moderate breathlessness response had a higher estV'$O_2$ than the step rate estimated to elicit the same breathlessness intensity response. Despite this, there was no difference between the breathlessness intensity rating at end-exercise for the step test [3.4 ± 2.2 (range: 0-9) Borg CR10 units] compared to the shuttle test [4.1± 2.6 (range: 0.5-10) Borg CR10 units; $p=0.11$], or the median breathlessness/estV'$O_2$ ratio for the step test [3.8 ± 4.5 (range: 0-19.4) Borg CR10 units/L/min] compared to the shuttle test [3.7 ± 3.5 (range: 0.3-13.1) Borg CR10 units/L/min; $p=0.91$].

There was no difference between breathlessness unpleasantness ratings at end exercise for the step (3.2 ± 2.1 Borg CR10 units) compared to shuttle test (3.6 ± 2.8 Borg CR10 units, $p = 0.41$) nor between intensity ratings of leg discomfort at end exercise for the step (2.0 ± 2.1 Borg CR10 units) compared to shuttle test (2.0 ± 2.0 Borg CR10 units, $p = 0.88$).

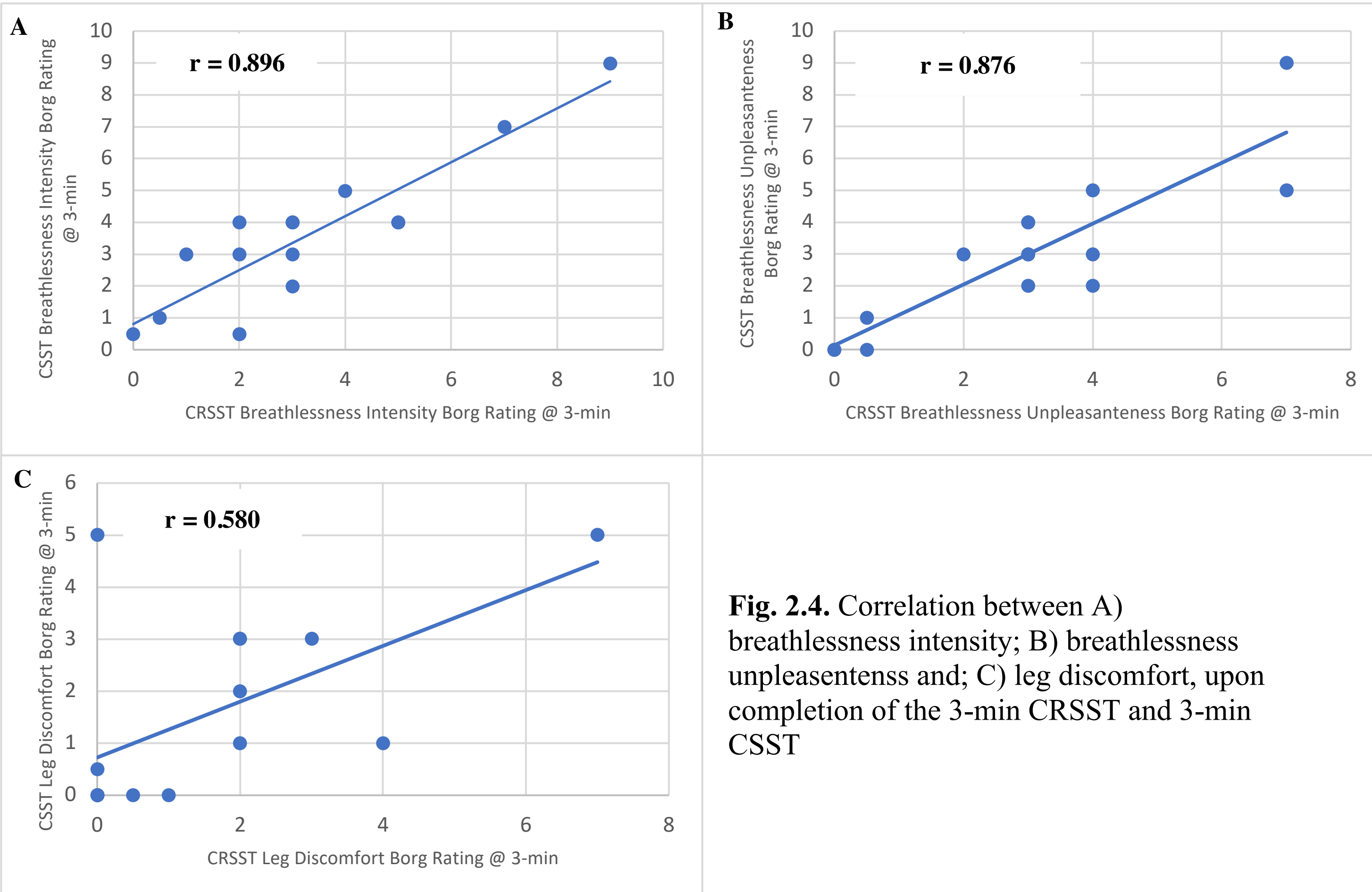



**Fig. 2.4.** Correlation between A) breathlessness intensity; B) breathlessness unpleasentenss and; C) leg discomfort, upon completion of the 3-min CRSST and 3-min CSST

### *2.4.5. Participant Reported Preferences between the 3-min CRSST and 3-min CSST*

On a self-administered questionnaire, 22% of participants identified walking on a flat surface as an ADL which induced breathlessness, but none of these participants identified it as their most difficult ADL. All participants identified walking uphill as an ADL which induced breathlessness, with 33% identifying it as their most difficult ADL. 94% percent of participants identified stair climbing as an ADL which induced breathlessness, with 17% identifying it as their most difficult ADL (**Table 2.9**).

**Table 2.9.** Participant reported most difficult activities of daily living due to breathlessness.

| Activities of Daily Living | Most Difficult | 2nd Most Difficult | 3rd Most Difficult | Induces Breathlessness |
|---|---|---|---|---|
| Walking on a flat surface | 0% | 0% | 5.6% | 22.2% |
| Walking up a hill | 33.3% | 11.1% | 27.8% | 100% |
| Stair climbing | 16.7% | 50% | 5.6% | 94.4% |

Prior to performing the 3-min CRSST and 3-min CSST, over 80% of participants identified stair stepping to be more difficult than walking on a flat surface. Nevertheless, after completing the 3-min CRSST and 3-min CSST the perceived difficulty was relatively equally distributed (**Fig. 2.5**): 33% of participants preferred the shuttle test; 11% preferred the step test; and 56% indicated not having a preference between the two tests.

When asked to compare the breathlessness experienced during the 3-min CRSST and 3-min CSST to the breathlessness experienced during daily activities, nearly 80% of participants agreed that both the step and shuttle test provided an accurate representation of their breathlessness in daily life (**Fig. 2.6**), with 50% reporting that the shuttle test provided a more accurate representation compared to the step test (**Fig. 2.5**).

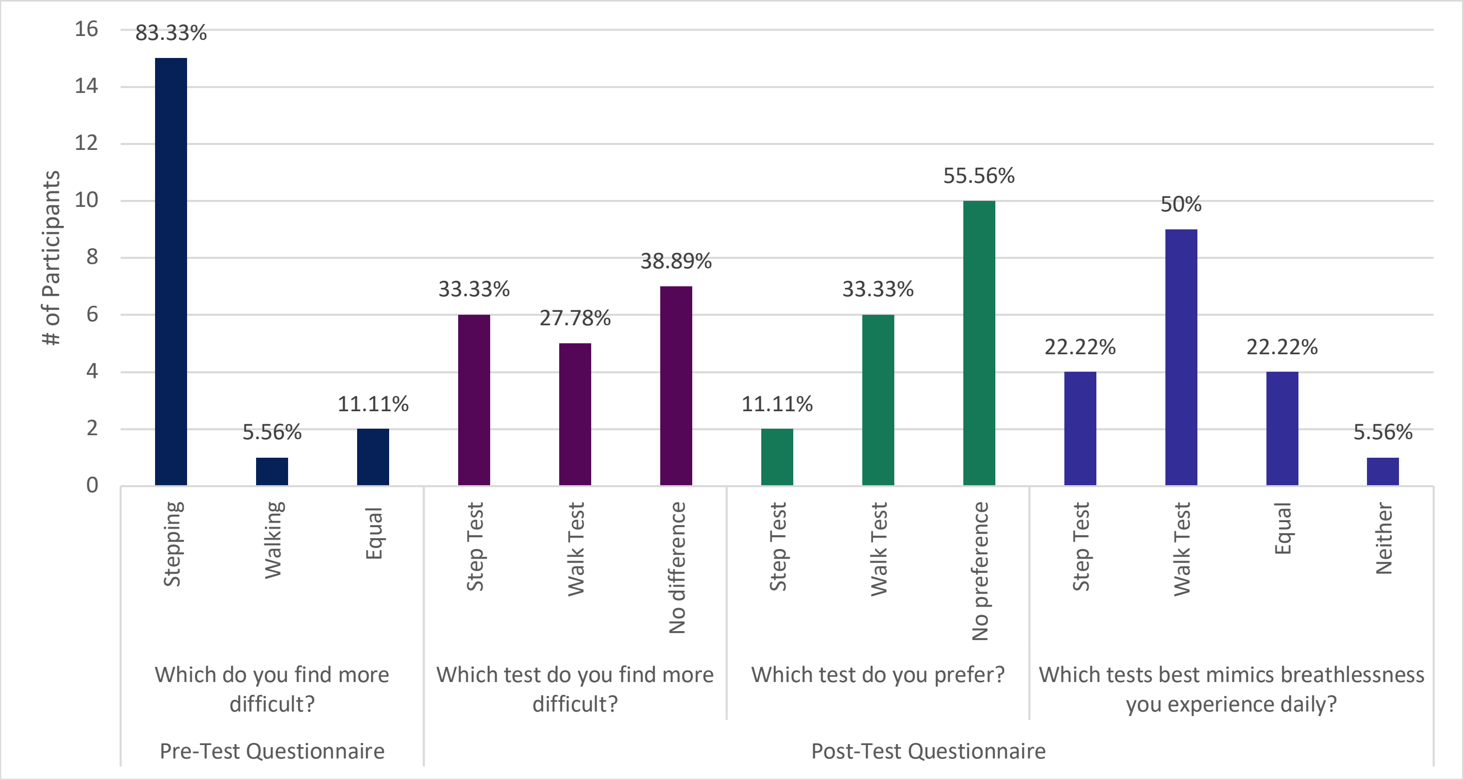



**Fig. 2.5.** Pre vs. Post 3-min CRSST and 3-min CSST Participant Reported Preferences: Stepping vs. Walking

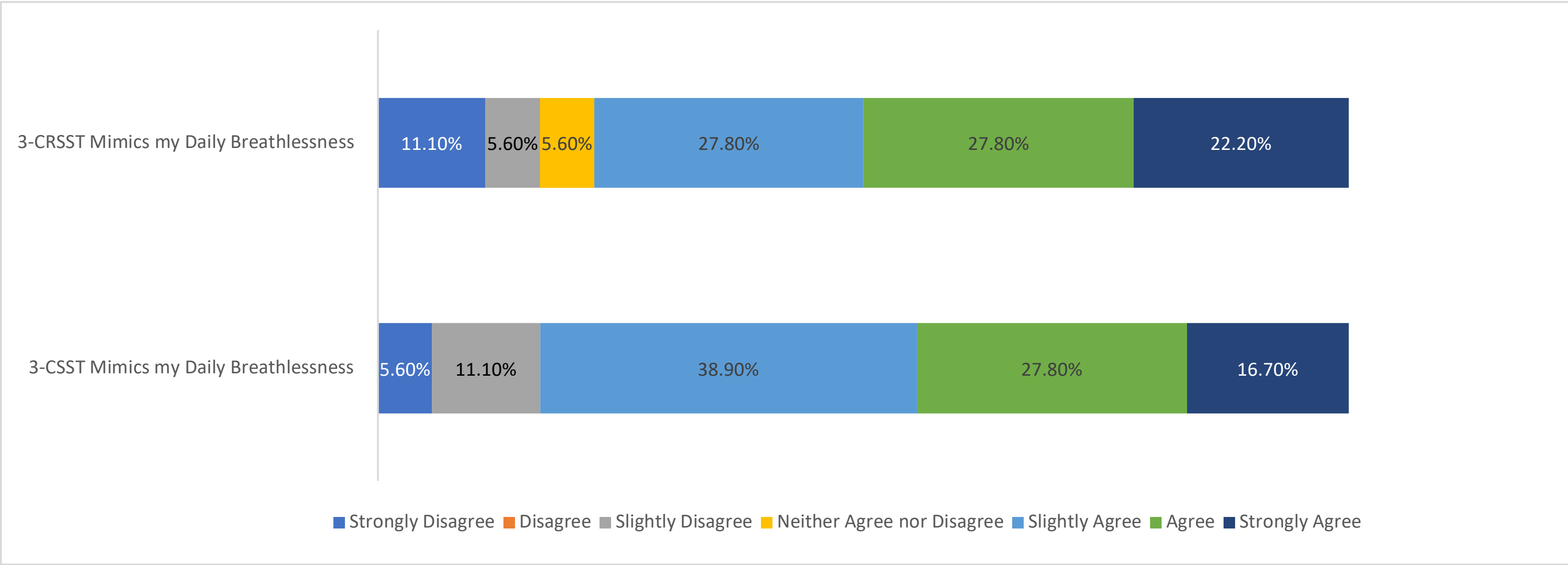



**Fig. 2.6.** Participant Reported Ecological Validity of the 3-min CRSST and 3-min CSST.

## 2.5. Discussion

### *2.5.1. Main findings*

The purpose of this study was to develop and prospectively validate prediction equations to identify the step rate and shuttle speed required to elicit at least a moderate level of breathlessness at the end of the 3-min CRSST and 3-min CSST in people with COPD. The primary finding of this study was that the two equations were able to be developed based on readily available participant characteristics and were successful with approximately 60% of the individually identified step rates and shuttle speeds eliciting a breathlessness intensity rating of ≥3 Borg CR10 units upon completion of all 3-min of step and shuttle testing.

### *2.5.2. Comparison with the existing literature*

Previous studies that have used the 3-min CRSST and/or 3-min CSST as a standardized and individualized breathlessness assessment tool, have implemented an algorithmic approach to determine the appropriate exercise intensity. A study conducted by Maltais et al. [62] screened 130 participants, four of which were excluded because the algorithm failed to identify a shuttle speed that elicited a breathlessness intensity rating of ≥4 Borg CR10 units at the end of the 3-min CSST; this indicates a 97% success rate. This approach required participants to complete a minimum of two 3-min CSSTs with at least a 30-minute break in between tests. Similarly, Sava et al. [34] used an algorithmic approach based on the shuttle speeds used in a previous study by Perrault et al. [36] and also reported a 97% success rate using the algorithm. These success rates are higher than those of the current study (60% success rate); however, both of these earlier studies required participants to perform multiple tests to identify the optimal exercise intensity. It is reasonable to assume that if our participants had completed at least one additional 3-min CRSST and/or 3-min CSST based on the results from the initial predicted step rate/shuttle speed, we would have also expected near perfect success rates. However, we chose to utilize the equations and assess their performance in the way that we felt they would most likely be used in clinical practice. While the algorithmic approach used in earlier clinical research studies was more effective than our equations at identifying an "optimal" exercise intensity for step and shuttle tests, both studies identified the burdensome nature of the algorithmic approach and the need for a more pragmatic way to determine the appropriate exercise intensity on an individual basis [34, 62]. Although additional research is needed to more accurately access the performance of the newly developed prediction equations, we believe that a preliminary success rate of 60% may support use of the step and/or shuttle tests and their respective prediction equations in clinical practice.

In past studies, step rates ranging from 14-32 steps/min have been used for the 3-min CRSST [35, 36]. The majority of participants in these earlier studies were able to complete the test at step rates ranging from 14-26 steps/min; approximately 60% of participants completed 26 steps/min, this percentage increased as step rate decreased [35, 36]. In our study, individualized step rates tended to be lower than those used in the existing literature; step rates ranged from 10-24 steps/min with the most common step rate being 10 steps/min for 44% of our participants. For the 3-min CSST, past studies have used shuttle speeds ranging from 1.5-6 km/hr [34, 36, 62]. The majority of participants in these earlier studies were able to complete the test at shuttle speeds ranging from 1.5-4 km/hr; 84% completed at 4 km/hr, this percentage increased as shuttle speed decreased [36]. In our study, individualized shuttle speeds were all within the range reported previously; shuttle speeds ranged from 2.5-6 km/hr, with the most common shuttle speed being 3.5 km/hr for 33% of our participants. Previous studies using the step and/or shuttle tests have implemented them primarily on people with moderate-to-severe COPD [34, 36, 62]. In the current study, participants had, on average, severe airflow obstruction; however, there were participants with a wide range of disease severities (range: 17-80 %pred $FEV_1$). Inclusion of participants in our study with very severe airflow obstruction may help to explain why some of the step rates identified by the 3-min CRSST prediction equation were lower than those reported in the existing literature.

Looking at the different outcome groups for the 3-min CRSST and 3-min CSST, there is an observable relationship between the groups and their lung function, health status and symptom burden. On average, participants who completed all 3-minutes of the step or shuttle test with a breathlessness intensity rating of ≥3 Borg CR10 units had more severe chronic airflow obstruction with greater symptom burden than those participants who completed all 3-minutes of the step or shuttle test with a breathlessness intensity rating of <3 Borg CR10 units. Taken together, our preliminary results suggest that the newly developed prediction equations worked quite well in people with greater symptom burden and more severe airflow obstruction; this is likely because these are shared characteristics with the derivation cohort, which was comprised mostly of symptomatic adults with GOLD II (moderate) and III (severe) COPD (Table 2.5). We speculate that these between-group differences will be significant and meaningful as more participants are enrolled in the trial. If these differences persist after further testing, as we expect, they could help inform how to further refine the equations to more accurately predict step rates and/or shuttle speeds for people in each GOLD stage of COPD, presumably by adding a greater number of people with GOLD I (mild) or IV (very severe) COPD to the derivation cohort.

*2.5.3. Which test is best?*

Previous studies have shown that both the 3-min CRSST and 3-min CSST are valid, reliable and responsive field-based exercise tests to assess breathlessness following bronchodilator therapy in people with COPD [34-36, 62]. To date, the majority of studies have employed the shuttle test, presumably because walking is more familiar and directly relevant to activities of daily life [34, 62]. In our study, the newly developed prediction equations identified step rates and shuttle speeds that allowed ~60% of participants to complete all 3-minutes of each test with a breathlessness intensity rating of ≥3 Borg CR10 units, suggesting that the equations performed equally well. With similar performance, the question becomes "which test is better suited for use in clinical practice - the step or shuttle test?"

There was a significant difference between the estV'$O_2$ (or oxygen cost of exercise) for the 3-min CRSST and 3-min CSST. On average, the estV'$O_2$ for the step test was lower than for the shuttle test (0.92 vs. 1.05 L/min, p=0.001). Breathlessness intensity ratings at end-exercise between the step and shuttle test was nevertheless similar (step: 3.4 ± 2.2 vs. shuttle: 4.1± 2.6 Borg CR10 units, p=0.10). According to the findings of Tracey et al., [41] this difference of 0.7 Borg CR10 units may not be clinically meaningful as it is less than the MCID of ~1 Borg CR10 unit. Taken together, this could explain why there were more people who (i) could not finish all 3-minutes of the shuttle test; and (ii) did not reach a breathlessness intensity rating of ≥3 Borg CR10 units during the step test (**Fig. 2.2**). Under the assumption that it is better to have an exercise test be more intense than less intense (and considering that the equations developed for the step and shuttle tests performed equally well), these results indicate a possible advantage to using the shuttle test over the step test.

Concerning the practicality of implementing these tests into clinical practice, however, the step test may be more feasible as it takes up less space and can be done in an examination room [35, 36]. By contrast, the shuttle test requires a 10m hallway that is both uninterrupted and wide enough to permit turning, conditions that don't likely exist in most physician's offices [34, 36, 62]. It is also reasonable to assume that the step test is better suited for people with COPD requiring supplemental oxygen, since the oxygen tank wouldn't need to be transported; and/or who require a walking aid but may be capable of performing the step test without use of an aid but assistance from hand rails if needed.

After performing both tests, participants reported a preference for the shuttle test and found that it more accurately mimicked the breathlessness they experienced in daily life (**Fig. 2.5**). Perceived difficulty was relatively equally distributed between the step and shuttle test (**Fig. 2.5**) with the majority of participants finding the tests equally difficult; this is likely because both equations targeted the same breathlessness intensity rating of 4 Borg CR10 units. With the 3-min CRSST and 3-min CSST performing equally well in regard to reliability, responsiveness, and success of the prediction equations, the decision of which test to use will ultimately be up to the clinician or researcher, and depend on the specific circumstances of evaluation (i.e., availability of space, patient preferences, use of ambulatory oxygen and/or walking aid). It is our belief that the step test may be better suited for use in clinical practice where space and time are limited, while the shuttle test (which more closely mimics exertional breathlessness in daily life) may be better suited for use in clinical research trials when there is often more time and space.

*2.5.4. Strengths and limitations*

A methodological strength of this study was the pilot test (n=9), which allowed for the newly developed prediction equations to be tested and subsequently refined to improve their performance. Initially, the equations were only successful in 22% and 56% of participants for the 3-min CRSST and 3-min CSST, respectively; the success rates increased after homogenizing the derivation cohort and refining the equations, particularly for the 3-min CRSST. It is reasonable to assume that recruiting to our originally planned sample size of 100 participants (post COVID-19) and refining the equations even further may boost their success rates above 60%. Both equations were developed using readily available patient-level data and were tested in a single visit. Qualitative measures were taken to supplement the quantitative results and gain a more wholistic view of the participant's breathlessness experience(s). This study was designed to be ecologically valid and was done with the goal of evaluating performance of the equations in the way that we believed they would most likely be implemented for use in clinical practice as well as clinical (therapeutic) research studies. Spirometry was not performed as a part of this study: lung function test parameters (i.e. $FEV_1$) were taken from electronic medical records and only simple variables required for the equations were measured on site (i.e. height, body mass, mMRC dyspnea rating). To further simplify the process, assessment was limited to a single step and shuttle test, without additional tests when the participant was unable to complete all 3-minutes and/or did not report at least a moderate intensity rating of breathlessness at the exercise intensity identified by the prediction equations.

The primary limitation of this study was the small sample size. Our goal was to complete testing of 100 participants but due to the unforeseen circumstances of COVID-19, there was an abrupt halt on participant recruitment and data collection. The remaining data collection is set to resume after the pandemic subsides. Furthermore, the equations were developed based on a relatively small derivation cohort of clinically stable and mostly symptomatic adults with moderate to severe COPD, without resting or exercise-induced desaturation. Thus, results may not apply to people with COPD with resting or exercise-induced oxygen desaturation and/or who are at either extreme of the disease severity spectrum. A larger sample size of participants in each GOLD stage is required to comment on the widespread applicability of the equations. These equations were developed specifically for people with COPD and should not be used in non-COPD populations that suffer from chronic breathlessness (e.g., interstitial lung disease, heart failure).

### *2.5.5. Implications & future directions*

The COPD-specific equations developed and prospectively validated in this study theoretically increase the feasibility of implementing the 3-min CRSST and/or 3-min CSST into clinical practice and research studies to assess the efficacy of breathlessness management strategies. These equations and their respective tests provide clinicians and researchers with a simple, inexpensive, standardized and individualized assessment method to quantify breathlessness and assess its response to intervention or disease progression in people with COPD.

To further increase the operational feasibility of implementing the 3-min CRSST and/or 3-min CSST for use in clinical practice or research, testing for this study needs to continue (post COVID-19) to see how the equations work in a larger cohort of people with mild-to-very severe COPD. If deemed necessary upon further data collection, the equations can/should be further refined to improve their performance and increase their generalizability. In addition, it will be important to compare the performance and responsiveness of the step and shuttle test to other commonly used exertional breathlessness assessment tests in the COPD population, specifically the endurance shuttle walk test and/or constant-load cardiopulmonary cycle/treadmill exercise test. As outlined by Lewthwaite et al. [29] and summarized in **Table 1.1**, existing exertional breathlessness assessments tests have unique limitations in terms of cost, time, space, personnel and expertise that limit their widespread use. In practice, a physician or researcher would identify one of the two exercise tests (step or shuttle) to assess breathlessness. In this way, we anticipate that it should take <10 minutes per patient to complete assessment of exertional breathlessness, including the time it takes to (i) identify the step rate or shuttle speed using the newly developed prediction equation, (ii) explain the test to the patient and familiarize them to Borg's CR10 scale, and (iii) equip the patient with a pulse oximeter to record HR and $SpO_2$.

Now that the prediction equations have been developed and once their performance has been assessed in a larger cohort of people with COPD, a logical next step would be to conduct an implementation study to assess the feasibility, performance and added value of the step and walk test in clinical practice. For example, studies conducted in the office of a primary care physician; in a COPD outpatient clinic of a secondary or tertiary health care centre; and/or as part of inpatient or outpatient pulmonary rehabilitation program. Finally, the validity, reliability and responsiveness of the 3-min CRSST and 3-min CSST should be assessed in other clinical populations burdened by breathlessness such as interstitial lung disease and heart failure. Once these test performance characteristics have been confirmed in these populations, prediction equations should be developed and prospectively validated.

## 2.6 Conclusion

The prediction equations developed and prospectively validated in this MSc thesis to identify the exercise intensity (step rate and shuttle speed) for use in adults with COPD, represent a unique opportunity for healthcare providers and researchers to incorporate the 3-min CRSST and/or 3-min CSST to assess breathlessness into their daily practice and clinical (therapeutic) trials. The predicted step rates and shuttle speeds resulted in comparable physiological and perceptual responses to exercise, providing no clear benefit to using one test over the other. Although additional research is required, the findings of this MSc thesis are promising and indicated that both the 3-min CRSST and 3-min CSST prediction equations performed well in 60% of our participants. In the absence of an established threshold for clinical feasibility, we believe that this success rate warrants future use of these tests and their respective equations to quantify breathlessness at an individualized and standardized level of exertion in people with COPD in clinical care and research settings.

## References


1. Vogelmeier, C.F., et al., *Global Strategy for the Diagnosis, Management and Prevention of Chronic Obstructive Lung Disease 2017 Report: GOLD Executive Summary.* Respirology, 2017. **195**(5): p. 557-582.
2. Adeloye, D., et al., *Global and regional estimates of COPD prevalence: Systematic review and meta-analysis.* J Glob Health, 2015. **5**(2): p. 020415.
3. Tan, W.C., et al., *Can age and sex explain the variation in COPD rates across large urban cities? A population study in Canada.* The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease, 2011. **15**(12): p. 1691-8.
4. *2020 Global Strategy for Prevention, Diagnosis and Managment of COPD. Availible from: https://goldcopd.org/gold-reports/.*
5. Waskiw-Ford, M., et al., *Effect of inhaled nebulized furosemide on breathlessness during exercise in the setting of abnormal restrictive constraints on tidal colume expansion: a randomized, double blind, placebo contolled, crossover, dose-response study.* Front. Physiol, 2018. **9**(86): p. 1-11.
6. Maleki-Yazdi, M.R., et al., *The burden of illness in patients with moderate to severe chronic obstructive pulmonary disease in Canada.* Can Respir J, 2012. **19**(5): p. 319-24.
7. *Morbidity and mortality : chartbook on cardiovascular, lung and blood diseases*. 2009: National Heart, Lung and Blood Institute.
8. Parshall, M.B., et al., *An official American Thoracic Society statement: update on the mechanisms, assessment, and management of dyspnea.* Am J Respir Crit Care Med, 2012. **185**(4): p. 435-52.
9. Mullerova, H., et al., *Prevalence and burden of breathlessness in patients with chronic obstructive pulmonary disease managed in primary care.* PLoS One, 2014. **9**(1): p. e85540.
10. Sundh, J. and M. Ekstrom, *Persistent disabling breathlessness in chronic obstructive pulmonary disease.* Int J Chron Obstruct Pulmon Dis, 2016. **11**: p. 2805-2812.
11. Annegarn, J., et al., *Problematic activities of daily life are weakly associated with clinical characteristics in COPD.* J Am Med Dir Assoc, 2012. **13**(3): p. 284-90.
12. Bestall, J.C., et al., *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): p. 581-6.
13. Ramon, M.A., et al., *The dyspnoea-inactivity vicious circle in COPD: development and external validation of a conceptual model.* Eur Respir J, 2018. **52**(3).
14. Watz, H., et al., *Physical activity in patients with COPD.* Eur Respir J, 2009. **33**(2): p. 262-72.
15. Nishimura, K., et al., *Dyspnea Is a Better Predictor of 5-Year Survival Than Airway Obstruction in Patients With COPD.* Chest, 2002. **121**(5): p. 1434-1440.
16. Oga, T., et al., *Analysis of the factors related to mortality in chronic obstructive pulmonary disease: role of exercise capacity and health status.* Am J Respir Crit Care Med, 2003. **167**(4): p. 544-9.
17. Schünemann, H.J., et al., *Measurement properties and interpretability of the Chronic respiratory disease questionnaire (CRQ).* COPD, 2005. **2**(1): p. 81-9.
18. Bourbeau, J., et al., *CTS position statement: Pharmacotherapy in patients with COPD-An update.* Canadian Journal of Respiratory, Critical Care, and Sleep Medicine, 2017. **1**(4): p. 222-241.
19. Mahler, D.A. and S. Sethi, *To Improve COPD Care: A New Instrument Is Needed to Assess Dyspnea.* Chest, 2018. **154**(2): p. 235-237.
20. Administration, U.F.D., *Chronic obstructive pulmonary disease: Developing drugs for treatment (Draft guidance for industry)*, in *Federal Register*. 2007. p. 31945-31946

21. Ekstrom, M., et al., *Breathlessness measurement should be standardised for the level of exertion.* Eur Respir J, 2018. **51**(5): p. e1800486.
22. Jones, P.W., et al., *Minimal clinically important differences in pharmacological trials.* Am J Respir Crit Care Med, 2014. **189**(3): p. 250-5.
23. Ekstrom, M., D.C. Currow, and M.J. Johnson, *Outcome measurement of refractory breathlessness: endpoints and important differences.* Curr Opin Support Palliat Care, 2015. **9**(3): p. 238-43.
24. Mahler, D.A., et al., *The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes.* Chest, 1984. **85**(6): p. 751-8.
25. Nouraei, S., et al., *Sensitivity and responsiveness of the Medical Research Council dyspnoea scale to the presence and treatment of adult laryngotracheal stenosis.* Clinical Otolaryngology, 2008. **33**(6): p. 575-580.
26. Casanova, C., et al., *Differential effect of modified medical research council dyspnea, COPD assessment test, and clinical COPD questionnaire for symptoms evaluation within the new GOLD staging and mortality in COPD.* Chest, 2015. **148**(1): p. 159-168.
27. Mahler, D.A. and T.J. Witek, Jr., *The MCID of the transition dyspnea index is a total score of one unit.* COPD, 2005. **2**(1): p. 99-103.
28. Fletcher, C.M., et al., *The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population.* Br Med J, 1959. **2**(5147): p. 257-66.
29. Lewthwaite, H., et al., *Standardized measurement of breathlessness during exercise.* Curr Opin Support Palliat Care, 2019. **13**(3): p. 152-160.
30. Puente-Maestu, L., et al., *Use of exercise testing in the evaluation of interventional efficacy: an official ERS statement.* Eur Respir J, 2016. **47**(2): p. 429-60.
31. Osadnik, C.R., et al., *Two 6-minute Walk Tests Are Required During Hospitalisation for Acute Exacerbation of COPD.* COPD, 2016. **13**(3): p. 288-92.
32. Borel, B., et al., *Responsiveness of Various Exercise-Testing Protocols to Therapeutic Interventions in COPD.* Pulm Med, 2013. **2013**: p. 1-11.
33. Singh, S.J., et al., *An official systematic review of the European Respiratory Society/American Thoracic Society: measurement properties of field walking tests in chronic respiratory disease.* Eur Respir J, 2014. **44**(6): p. 1447-78.
34. Sava, F., et al., *Detecting improvements in dyspnea in COPD using a three-minute constant rate shuttle walking protocol.* COPD, 2012. **9**(4): p. 395-400.
35. Borel, B., et al., *Three-minute constant rate step test for detecting exertional dyspnea relief after bronchodilation in COPD.* Int J Chron Obstruct Pulmon Dis, 2016. **11**: p. 2991-3000.
36. Perrault, H., et al., *Paced-walk and step tests to assess exertional dyspnea in COPD.* COPD, 2009. **6**(5): p. 330-9.
37. Bhatt, S.P., et al., *Association Between Expiratory Central Airway Collapse and Respiratory Outcomes Among Smokers.* JAMA, 2016. **315**(5): p. 498-505.
38. Williams, M.T., D. John, and P. Frith, *Comparison of the Dyspnoea-12 and Multidimensional Dyspnoea Profile in people with COPD.* Eur Respir J, 2017. **49**(3): p. e1600773.
39. Kupferberg, D.H., et al., *Minimal clinically important difference for the UCSD Shortness of Breath Questionnaire.* J Cardiopulm Rehabil, 2005. **25**(6): p. 370-7.
40. Ries, A.L., *Minimally clinically important difference for the UCSD Shortness of Breath Questionnaire, Borg Scale, and Visual Analog Scale.* COPD, 2005. **2**(1): p. 105-10.
41. Tracey, L., et al., *Constant-rate shuttle walking and stair stepping test to assess exertional breathlessness in COPD: Responsiveness to pulmonary rehabilitation and estimation of the minimal important difference [Unpublished Raw Data]*. 2019.

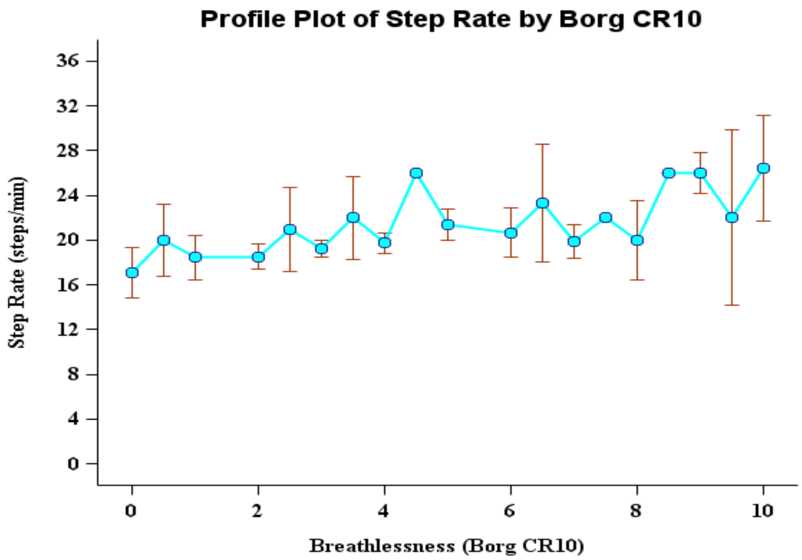
42. Eakin, E.G., et al., *Reliability and validity of dyspnea measures in patients with obstructive lung disease.* International Journal of Behavioral Medicine, 1995. **2**(2): p. 118-134.
43. Tabberer, M., J. Brooks, and T. Wilcox, *A meta-analysis of four randomized clinical trials to confirm the reliability and responsiveness of the Shortness of Breath with Daily Activities (SOBDA) questionnaire in chronic obstructive pulmonary disease.* Health Qual Life Outcomes, 2015. **13**: p. 1-10.
44. Puhan, M.A., et al., *Self-administration and interviewer-administration of the German Chronic Respiratory Questionnaire: instrument development and assessment of validity and reliability in two randomised studies.* Health Qual Life Outcomes, 2004. **2**(1): p. 1-9.
45. Williams, J.E., et al., *Development of a self-reported Chronic Respiratory Questionnaire (CRQ-SR).* Thorax, 2001. **56**(12): p. 954-9.
46. Meek, P.M., et al., *Reliability and validity of the multidimensional dyspnea profile.* Chest, 2012. **141**(6): p. 1546-1553.
47. Parshall, M.B., et al., *Test-retest reliability of multidimensional dyspnea profile recall ratings in the emergency department: a prospective, longitudinal study.* BMC Emerg Med, 2012. **12**(6): p. 1-10.
48. Mahler, D.A., et al., *Validity and Responsiveness of the Self-Administered Computerized Versions of the Baseline and Transition Dyspnea Indexes.* Chest, 2007. **132**(4): p. 1283-1290.
49. Celli, B., et al., *Once-Daily Umeclidinium/Vilanterol 125/25 µg Therapy in COPD : A Randomized, Controlled Study.* Chest, 2014. **145**(5): p. 981-991.
50. Watkins, M.L., et al., *Shortness of Breath with Daily Activities questionnaire: validation and responder thresholds in patients with chronic obstructive pulmonary disease.* BMJ Open, 2013. **3**(10): p. e003048.
51. Rutten-van Mölken, M., B. Roos, and J.A. Van Noord, *An empirical comparison of the St George's Respiratory Questionnaire (SGRQ) and the Chronic Respiratory Disease Questionnaire (CRQ) in a clinical trial setting.* Thorax, 1999. **54**(11): p. 995-1003.
52. Banzett, R.B., et al., *Multidimensional Dyspnea Profile: an instrument for clinical and laboratory research.* Eur Respir J, 2015. **45**(6): p. 1681-91.
53. Stulbarg, M.S., et al., *Exercise training improves outcomes of a dyspnea self-management program.* J Cardiopulm Rehabil, 2002. **22**(2): p. 109-21.
54. Puhan, M.A., et al., *Relative responsiveness of the Chronic Respiratory Questionnaire, St. Georges Respiratory Questionnaire and four other health-related quality of life instruments for patients with chronic lung disease.* Respiratory Medicine, 2007. **101**(2): p. 308-316.
55. Beaumont, M., et al., *Effects of inspiratory muscle training on dyspnoea in severe COPD patients during pulmonary rehabilitation: controlled randomised trial.* Eur Respir J, 2018. **51**(1): p. e1701107.
56. Major, S., et al., *Veterans with chronic obstructive pulmonary disease achieve clinically relevant improvements in respiratory health after pulmonary rehabilitation.* J Cardiopulm Rehabil Prev, 2014. **34**(6): p. 420-9.
57. Nguyen, L.P., et al., *Differential Pulmonary Rehabilitation Outcomes in Patients With and Without COPD: Role of Gender.* J Cardiopulm Rehabil Prev, 2017. **37**(5): p. 350-355.
58. *Standardization of the oral glucose tolerance test. Report of the Committee on Statistics of the American Diabetes Association June 14, 1968.* Diabetes, 1969. **18**(5): p. 299-307.
59. Climie, R.E., et al., *Validity and reliability of central blood pressure estimated by upper arm oscillometric cuff pressure.* Am J Hypertens, 2012. **25**(4): p. 414-20.
60. Kim, D.L., et al., *Is an Oral Glucose Tolerance Test Still Valid for Diagnosing Diabetes Mellitus?* Diabetes & metabolism journal, 2016. **40**(2): p. 118-28.

61. Zechman Jr, F.W. and R.L. Wiley, *Afferent inputs to breathing: respiratory sensation.* Comprehensive Physiology, 2011: p. 449-474.
62. Maltais, F., et al., *Dual bronchodilation with tiotropium/olodaterol further reduces activity-related breathlessness versus tiotropium alone in COPD.* Eur Respir J, 2019: p. e1802049.
63. Johnson, M.J., et al., *Clinically important differences in the intensity of chronic refractory breathlessness.* J Pain Symptom Manage, 2013. **46**(6): p. 957-63.
64. Hankinson, J.L., J.R. Odencrantz, and K.B. Fedan, *Spirometric reference values from a sample of the general U.S. population.* Am J Respir Crit Care Med, 1999. **159**(1): p. 179-87.
65. Quanjer, P.H., et al., *Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations.* Eur Respir J, 2012. **40**(6): p. 1324-43.
66. McGavin, C.R., et al., *Dyspnoea, disability, and distance walked: comparison of estimates of exercise performance in respiratory disease.* Br Med J, 1978. **2**(6132): p. 241-3.
67. Jones, P.W., et al., *Development and first validation of the COPD Assessment Test.* Eur Respir J, 2009. **34**(3): p. 648-54.
68. Lewthwaite, H., et al., *Predicting the rate of oxygen consumption during the 3-minute constant-rate stair stepping and shuttle tests in people with COPD.* Journal of Thoracic Disease, 2020. **12**(5): p. 2489-2498.

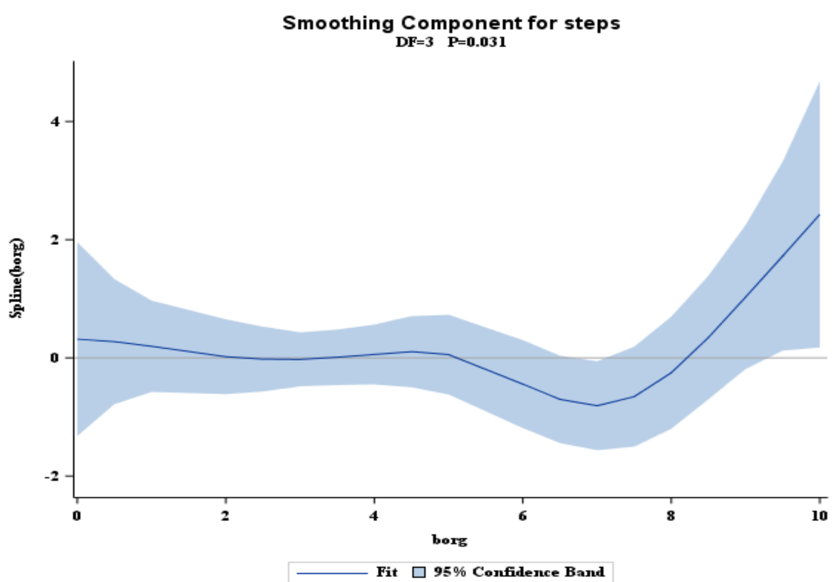
## Appendix A

**1st Set of Equations: Derivation of the 3-min CRSST Prediction Equation**

**Step 1: Check for patterns between stepping rate and breathlessness**




The profile plot showed a linear increasing trend of step rate with BORG dyspnea intensity (outlier at a step rate of 3.5 with a BORG of 3, was removed from the analysis).




The Generalized Additive Model also indicated a linearity; the p-value for the test of parameter for linear (BORG)<0.001, significant, but the nonlinear contribution was also significant, with a p=0.031. The above plot suggested the trend in BORG resembled quadratic and cubic, so quadratic and cubic variables were also put into the models.

## Step 2: Development of the prediction model

**Table 2. Change in Borg and step rate**

| | **Model 1** | | **Model 2** | | **Model 3** | | **Model 4** | |
|---|---|---|---|---|---|---|---|---|
| | **Coefficient (SE)** | **P value** | **Coefficient (SE)** | **P value** | **Coefficient (SE)** | **P value** | **Coefficient (SE)** | **P value** |
| **Parameter (Fixed effects)** | | | | | | | | |
| Intercept | 25.802 (3.093) | <0.001 | 26.646 (3.392) | <0.001 | -17.743 (21.599) | 0.4132 | -21.647 (21.593) | 0.3184 |
| Borg | 1.251 (0.133) | <0.001 | 1.332 (0.134) | <0.001 | 1.225 (0.131) | <0.001 | 1.222 (0.130) | <0.001 |
| Male sex | 2.725 (0.668) | <0.001 | 3.394 (0.690) | <0.001 | 3.369 (0.658) | <0.001 | 3.013 (0.622) | <0.001 |
| Age | -0.142 (0.044) | 0.002 | -0.162 (0.046) | 0.0007 | 1.203 (0.653) | 0.0682 | 1.336 (0.656) | 0.044 |
| Age*Age | | | | | -0.010 (0.005) | 0.0364 | -0.011 (0.005) | 0.0228 |
| Height | | | | | | | | |
| Weight | | | -0.040 (0.019) | 0.0384 | -0.037 (0.018) | 0.038 | | |
| BMI | | | | | | | -0.138 (0.059) | 0.0219 |
| mMRC (1-5) | -1.335 (0.322) | <0.001 | -1.407 (0.316) | <0.001 | -1.199 (0.310) | 0.0002 | -1.132 (0.310) | 0.0004 |
| FEV1, L | | | | | | | | |
| FEV1, % predicted-Nhanes | | | | | 0.052 (0.015) | 0.0006 | 0.060 (0.015) | 0.0002 |
| FEV1, % predicted-GLI | | | 0.053 (0.017) | 0.0029 | | | | |
| **Model Assessment** | | | | | | | | |
| **AIC** | 1493.2 | | 1451.4 | | 1463.5 | | 1460.1 | |

The model was fit to look at the change of step rate over BORG, including within-subject predictor (BORG) and significant between subject covariates. The analysis showed no significant quadratic and cubic trends, so the quadratic and cubic variables were removed from the model. Additionally, all the covariates, including their interaction terms were checked and only those which were significant were included in the final model; sex and age were forced into the final model. Absolute $FEV_1$ and plethysmography parameters were not significant predictors. Model 2 was the best model.

**Step 3: Evaluation of the prediction model**

Figure 1. Change in Step Rate by Breathlessness (Borg CR10)

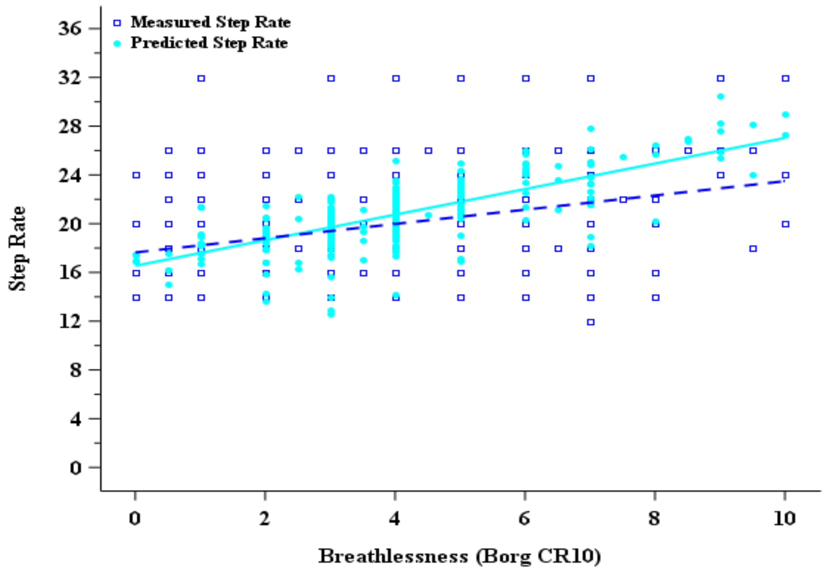
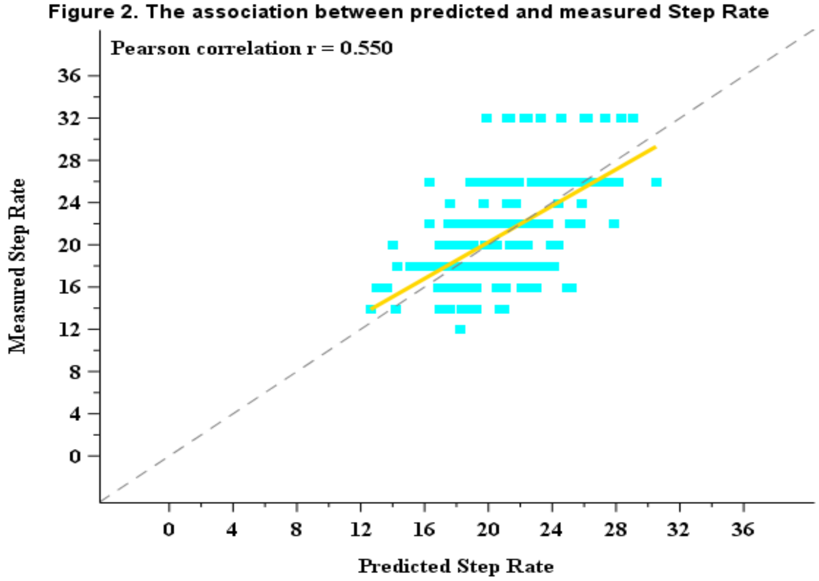



Figure 2. The association between predicted and measured Step Rate



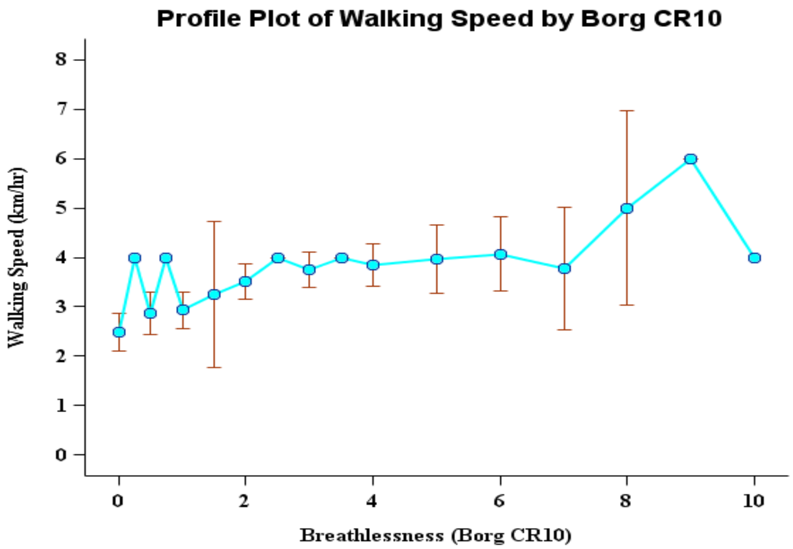

Mean Absolute Error = 2.98
Root Mean Squared Error = 3.74
Mean Percentage Error = 2.2%

All the error metrics suggest that, in general, the model did a fair job at predicting step rate based on breathlessness and several other factors however, the MPE indicates that the equation systematically overestimates (positive error) the step rate.

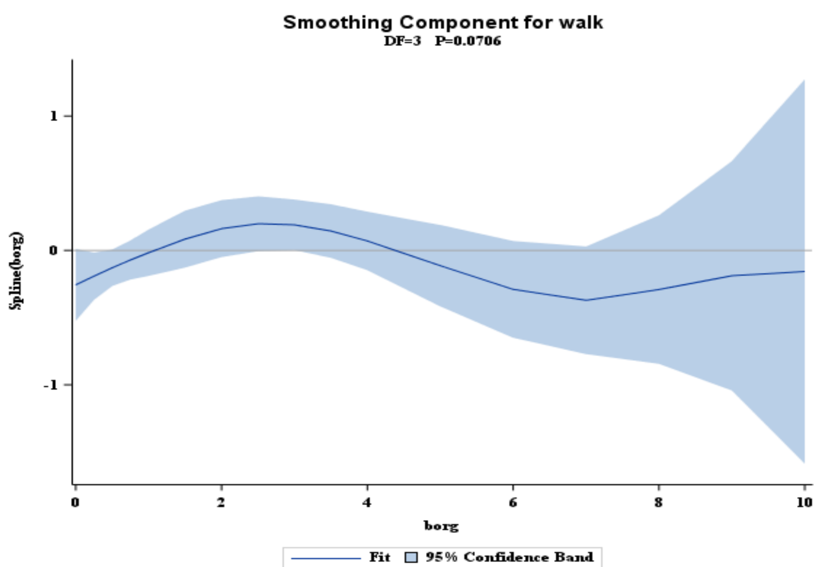
# Appendix B

## 1st Set of Equations: Derivation of the 3-min CSST Prediction Equation

**Step 1: Check for patterns between stepping rate and breathlessness**




The profile plot showed a linear increasing trend of walking speed with BORG dyspnea intensity ratings.




The Generalized Additive Model indicated a linearity; the p-value for the test of parameter for linear (BORG) <0.001, significant.

## Step 2: Development of the prediction model

**Table 2. Change in Borg and Walking Speed**

| | **Model 1** | | **Model 2** | | **Model 3** | | **Model 4** | |
|---|---|---|---|---|---|---|---|---|
| | **Coefficient (SE)** | **P value** | **Coefficient (SE)** | **P value** | **Coefficient (SE)** | **P value** | **Coefficient (SE)** | **P value** |
| **Parameter (Fixed effects)** | | | | | | | | |
| Intercept | -1.045 (1.828) | 0.5688 | 4.935 (0.874) | <0.001 | -1.817 (1.853) | 0.329 | 5.017 (0.871) | <0.001 |
| Borg | 0.339 (0.039) | <0.001 | 0.358 (0.039) | <0.001 | 0.340 (0.039) | <0.001 | 0.364 (0.039) | <0.001 |
| Male sex | 0.107 (0.200) | 0.5945 | 0.363 (0.173) | 0.0376 | -0.019 (0.011) | 0.0924 | 0.363 (0.172) | 0.0369 |
| Age | -0.018 (0.012) | 0.1122 | -0.020 (0.012) | 0.0936 | 0.053 (0.199) | 0.7887 | -0.021 (0.012) | 0.0807 |
| Height | 0.036 (0.010) | 0.0009 | | | 0.041 (0.011) | 0.0002 | | |
| Weight | -0.017 (0.005) | 0.0013 | | | -0.017 (0.005) | 0.0007 | | |
| BMI | | | -0.047 (0.015) | 0.0019 | | | -0.049 (0.015) | 0.0013 |
| FEV1, L | | | | | | | | |
| FEV1, % predicted-Nhanes | | | | | 0.014 (0.004) | 0.0005 | 0.012 (0.004) | 0.0036 |
| FEV1, % predicted-GLI | 0.014 (0.004) | 0.0024 | 0.012 (0.004) | 0.0067 | | | | |
| mMRC (1-5) | -0.273 (0.078) | 0.0007 | -0.305 (0.080) | 0.0002 | -0.265 (0.077) | 0.0008 | -0.307 (0.079) | 0.0002 |
| **Model Assessment AIC** | 733.1 | | **727.9** | | 732.6 | | 729.1 | |

The model was fit to look at the change of walking speed over BORG, including within-subject predictor (BORG) and significant between-subject co-variates. The analysis showed no significant quadratic and cubic trends, so the quadratic and cubic variables were removed from the model. Additionally, all the covariates, including their interaction terms were checked and only those which were significant were included in the final model; sex and age were forced into the final model. Absolute $FEV_1$ and plethysmography parameters were not significant predictors. Model 2 was the best model.

## Step 3: Evaluation of the prediction model

Figure 1. Change in Walking Speed by Breathlessness (Borg CR10)

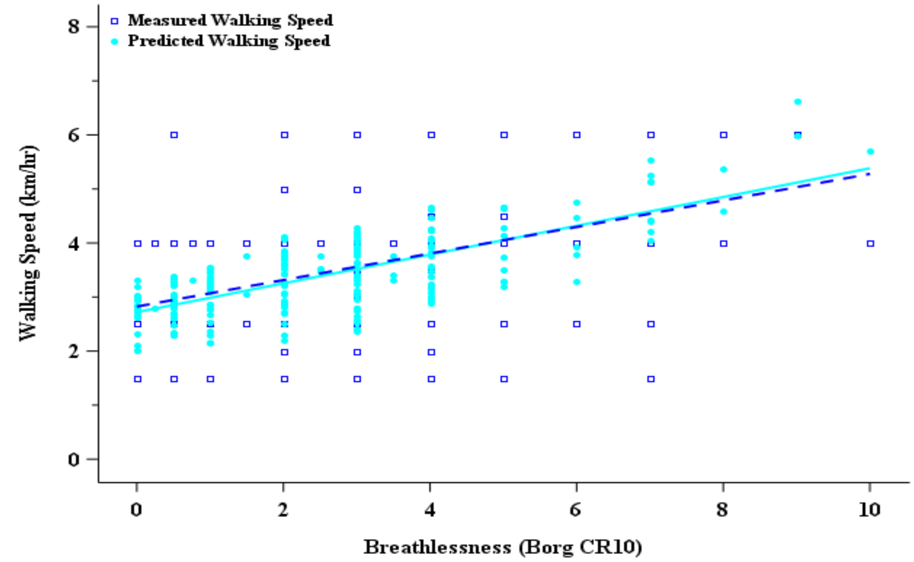
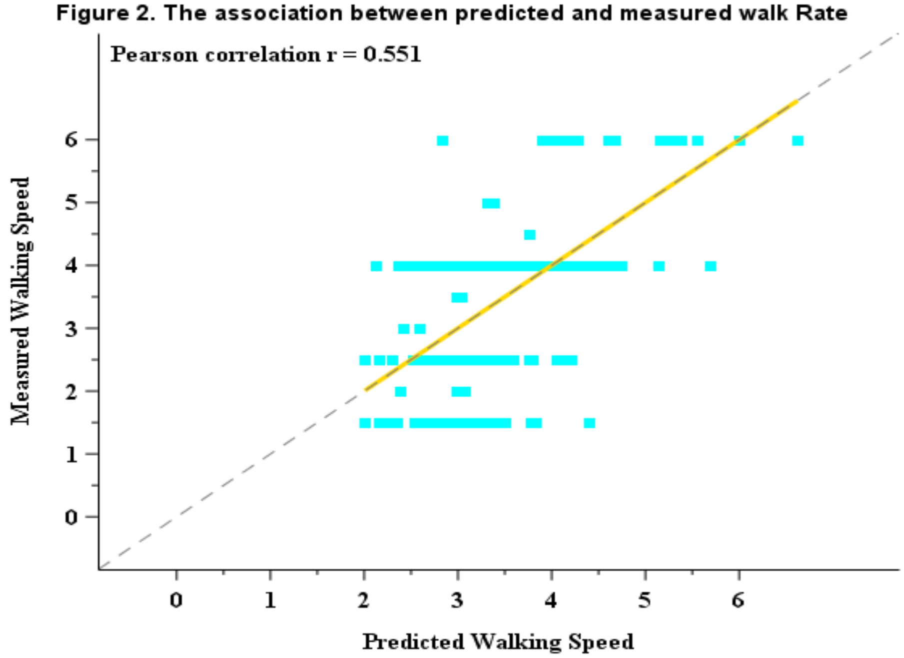



Figure 2. The association between predicted and measured walk Rate




Mean Absolute Error=0.927
Root Mean Squared Error=1.212
Mean Percentage Error=15.2%

All the error metrics suggest that, in general, the model did a fair job at predicting step rate based on breathlessness. The MPE indicates to us that it systematically overestimates (more positive error) the walking speed.

## Appendix C

**Median and mean breathlessness intensity ratings during constant work rate treadmill or cycle ergometry exercise at isotime in the placebo arm of randomized control trials.**

| First Author | Year | Journal | PubMed ID Number | Type of Intervention | Dyspnea Intensity (Borg CR10 scale units) during exercise at Isotime in the Placebo or Control Arm |
|---|---|---|---|---|---|
| Jensen | 2012 | Journal of Pain and Symptom Management | 22168961 | Inhaled nebulized fentanyl | 2.6 |
| Jensen | 2008 | Thorax | 18250181 | Inhaled nebulized furosemide | 4.9 |
| Beeh | 2011 | Journal of Chronic Obstructive Pulmonary Disease | 21793716 | BD | 6.7 |
| Beeh | 2014 | BMC Pulmonary Medicine | 25539654 | BD | 5.7 |
| Guennette | 2013 | Respiratory Medicine | 23421968 | BD | 3.2 |
| Man | 2004 | Thorax | 15170026 | BD | 4.62 |
| Neder | 2007 | Respiratory Medicine | 17658249 | BD | 7 |
| O'Donnell | 2004 | European Respiratory Journal | 15293609 | BD | 4.7 |
| O'Donnell | 2004 | European Respiratory Journal | 15218994 | BD | 5.4 |
| Peters | 2006 | Thorax | 16467067 | BD | 6.4 |
| Porszasz | 2013 | American Journal of Respiratory Critical Care Medicine | 23741986 | BD | 4.7 |
| Elbehairy | 2018 | Respiratory Physiology and Neurobiology | 29578103 | BD | 4.2 |
| O'Donnell | 2000 | American Journal of Respiratory and Critical Care Medicine | 11282762 | Hyperoxia | 5.2 |
| Eves | 2009 | Chest | 19017883 | Hyperoxia | 6 |
| Abdallah | 2018 | Annals of the American Thoracic Society | 30049223 | Cannibis | 2.6 |
| Abdallah | 2017 | European Respiratory Journal | 29051274 | Morphine | 4.2 |
| Cabral | 2015 | European Journal of Physical and Rehabilitative Medicine | 24691248 | Pursed Lip Breathing | 6.5 |
| Langer | 2018 | Journal of Applied Physiology | 29543134 | Inspiratory Muscle Training | 4 |
| Charususin | 2018 | Thorax | 29914940 | Inspiratory Muscle Training | 5.9 |
| Gass | 2017 | Respiratory Care | 28765494 | Positive Airway Pressure | 7.5 |
| | | | | **Mean** | **5.2** |
| | | | | **Median** | **5.4** |

## Appendix D

## Pre-Test Questionnaire

Pre-Test Questionnaire

Please identify the physical activity or activities that induce breathlessness in your everyday life (select all that apply):

- ☐ Walking on a flat surface
- ☐ Walking up a hill
- ☐ Stair climbing
- ☐ Cycling
- ☐ Carrying light objects on a flat surface (weighing 10 pounds or less)
- ☐ Carrying heavy objects on a flat surface (weighing 10 pounds or more)
- ☐ Carrying out low intensity physical activities related to your home (for example: sweeping, cleaning, preparing meals, doing dishes, doing laundry)
- ☐ Carrying out moderate to heavy intensity physical activities related to your home (for example: mowing the lawn, shoveling snow, gardening, vacuuming)
- ☐ Carrying our self-care activities (for example: showering, dressing and/or undressing, washing and/or drying parts of your body)
- ☐ Sports
- ☐ Activities with others (for example: children, grandchildren, friends, neighbours, coworkers)
- ☐ Other (please identify): ______________________________

Of the physical activities selected above, please identify the three activities that induce the worst (i.e., most severe and/or problematic) breathlessness in your everyday life, with 1 being the most severe/problematic, 2 being the second most severe/problematic, and 3 being the third most several/problematic:.

1.
2.
3.

## Appendix E

## Post-Test Questionnaire

Post-Test Questionnaire

1. a) Which of the two tests did you prefer? (please circle one)

   3-minute step test (3MST)     3-minute walk test (3MWT)     No preference

   b) Why?

   ______________________________________________________________

   ______________________________________________________________

   ______________________________________________________________

2. a) Which of the two tests did you find the most difficult? (please circle one)

   3-minute step test (3MST)     3-minute walk test (3MWT)     No difference

   b) Why?

   ______________________________________________________________

   ______________________________________________________________

   ______________________________________________________________

3. Which of the two tests best represents or mimics the breathlessness you experience most often in your daily life? (please circle one)

   3-minute step test (3MST)     3-minute walk test (3MWT)

   The tests are equally representative

   Neither test represents or mimics the breathlessness I most often experience in daily life.

4. a) If I ask you to focus on the 3-minute step test, what were some factors (if any) that affected your ability and/or confidence to perform this particular test?

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

b) If I ask you to focus on the 3-minute walk test, what were some factors (if any) that affected your ability and/or confidence to perform this particular test?

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

5. Using the 1 - 7 scale below, please indicate your how much you agree with the statement by putting an 'X' in the corresponding box.

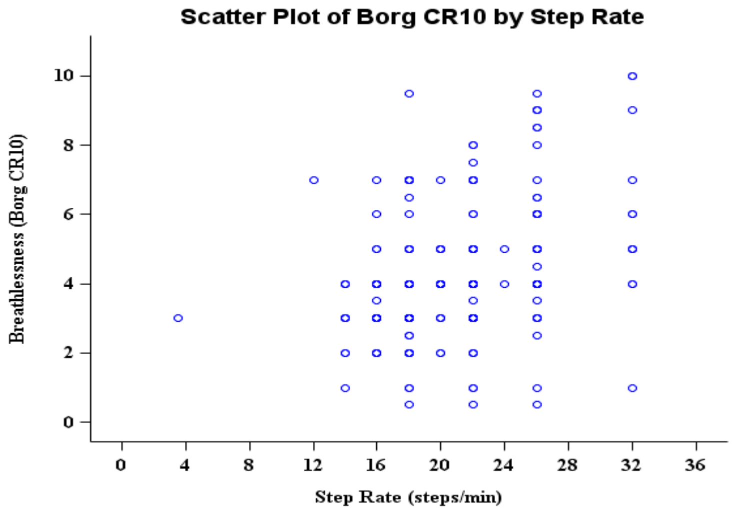
| | -3 Strongly disagree | -2 Disagree | -1 Slightly disagree | 0 Neither agree nor disagree | 1 Slightly agree | 2 Agree | 3 Strongly agree |
|---|---|---|---|---|---|---|---|
| The 3-minute step test mimics the breathlessness you experience on a daily basis. | ( ) | ( ) | ( ) | ( ) | ( ) | ( ) | ( ) |
| The 3-minute walk test mimics the breathlessness you experience on a daily basis. | ( ) | ( ) | ( ) | ( ) | ( ) | ( ) | ( ) |
| During a routine visit to my family or lung doctor, he/she would use the 3-minute step test to assess my breathlessness. | ( ) | ( ) | ( ) | ( ) | ( ) | ( ) | ( ) |
| During a routine visit to my family or lung doctor, he/she would use the 3-minute walk test to assess my breathlessness. | ( ) | ( ) | ( ) | ( ) | ( ) | ( ) | ( ) |

Additional Comments (if any)

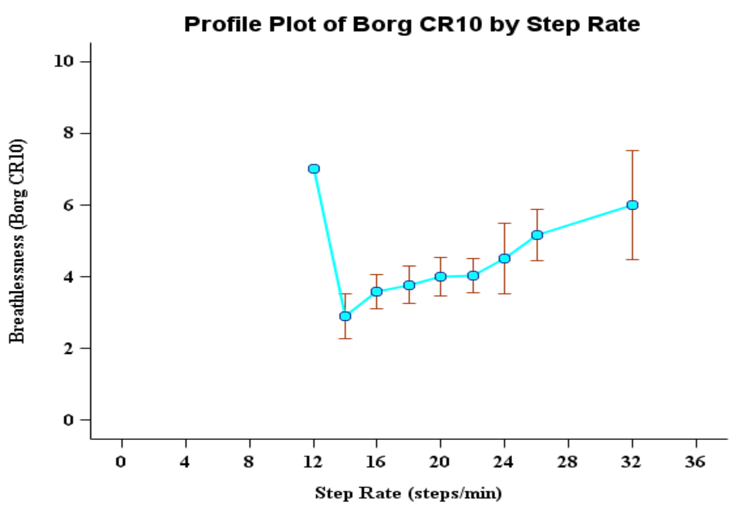
# Appendix F

## 2nd Set of Equations: Derivation of the 3-min CRSST Prediction Equation

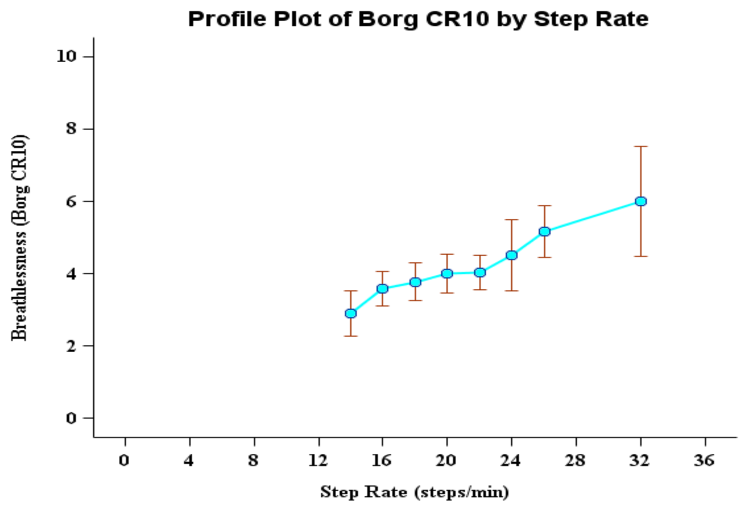
**Step 1: Check for patterns between stepping rate and breathlessness**



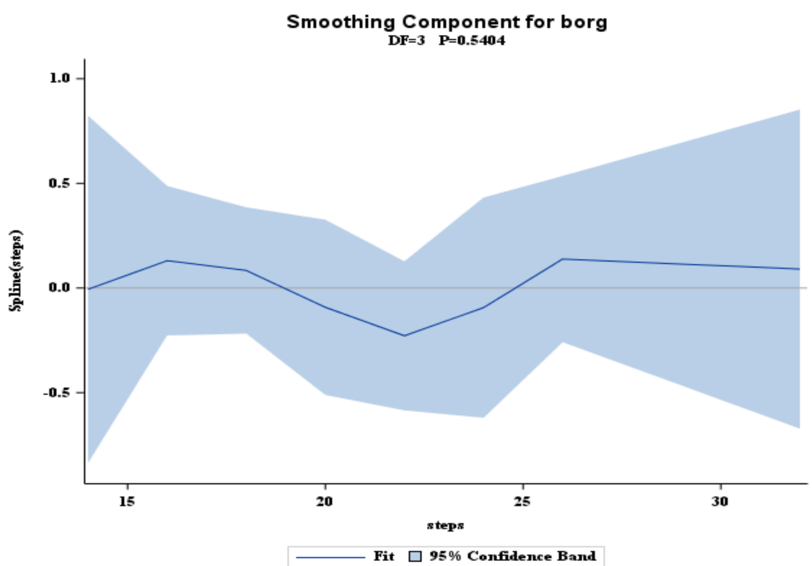

The scatter plot showed that there was an outlier at step rate = 3.5 with Borg = 3; this outlier was removed from the analysis.




From the above profile plot, we can see that there is another outlier at step rate = 12 and Borg = 7.0, which was also removed from the analysis.




After removing the outliers, the profile plot showed a linear increasing trend of Borg with step rate. Then we used 3 degrees of freedom Generalized Additive Model to check the patterns between Borg and step rate.




The Generalized Additive Model also indicated a linearity; the p-value for the test of parameter for linear (Steps) < 0.001, significant; the nonlinear contribution was non-significant, with a p = 0.5404.

## Step 2: Development of the prediction model

| Table2. Change in Step rate and Borg | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Model1 | | Model2 | | Model3 | | Model4 | |
| | Coefficient (SE) | P value | Coefficient (SE) | P value | Coefficient (SE) | P value | Coefficient (SE) | P value |
| **Parameter (Fixed effects)** | | | | | | | | |
| Intercept | -11.216 (4.148) | 0.0082 | -11.815 (4.134) | 0.0053 | -11.658 (4.129) | 0.0059 | -13.918 (4.239) | 0.0015 |
| Step rate | 0.644 (0.181) | 0.0005 | 0.646 (0.181) | 0.0005 | 0.643 (0.181) | 0.0006 | 0.653 (0.181) | 0.0005 |
| Male sex | -0.556 (0.442) | 0.2118 | -0.871 (0.424) | 0.0432 | -0.888 (0.423) | 0.0387 | -1.078 (0.427) | 0.0135 |
| Age | 0.140 (0.060) | 0.0228 | 0.158 (0.061) | 0.0107 | 0.157 (0.061) | 0.0111 | 0.155 (0.061) | 0.0119 |
| Age*Step rate | -0.006 (0.003) | 0.0269 | -0.006 (0.003) | 0.0264 | -0.006 (0.003) | 0.0274 | -0.006 (0.003) | 0.0234 |
| Height | | | | | | | | |
| Weight | 0.028 (0.012) | 0.0194 | 0.025 (0.011) | 0.0314 | 0.023 (0.011) | 0.0388 | 0.028 (0.011) | 0.0154 |
| BMI | | | | | | | | |
| mMRC (1-4) | 0.585 (0.235) | 0.0146 | 0.578 (0.233) | 0.0149 | 0.567 (0.232) | 0.0165 | 0.635 (0.230) | 0.0069 |
| FEV1, L | -0.814 (0.374) | 0.0315 | | | | | | |
| FEV1, % predicted-Nhanes | | | | | -0.023 (0.009) | 0.0105 | | |
| FEV1, % predicted-GLI | | | -0.023 (0.009) | 0.0153 | | | | |
| GOLD Stages | | | | | | | | 0.018 |
| GOLD1 | | | | | | | REF | |
| GOLD2 | | | | | | | 0.450 (0.590) | |
| GOLD3 | | | | | | | 1.617 (0.629) | |
| GOLD4 | | | | | | | 0.732 (0.955) | |
| **Model Assessment AIC** | 699.2 | | 705.3 | | 704.7 | | 691.3 | |

We fit the model to look at the change of Borg over step rate, including within-subjects predictor (Step rate) and significant between-subjects covariates. We checked all the covariates, including their interaction terms. For the final model we just kept the significant covariates and interaction terms. Sex and age were forced into the final model. No plethysmography parameters were significant predictors. Model 4 was the best model.

**Step 3: Evaluation of the prediction model**

Figure 1. Change in Breathlessness (Borg CR10) by Step Rate

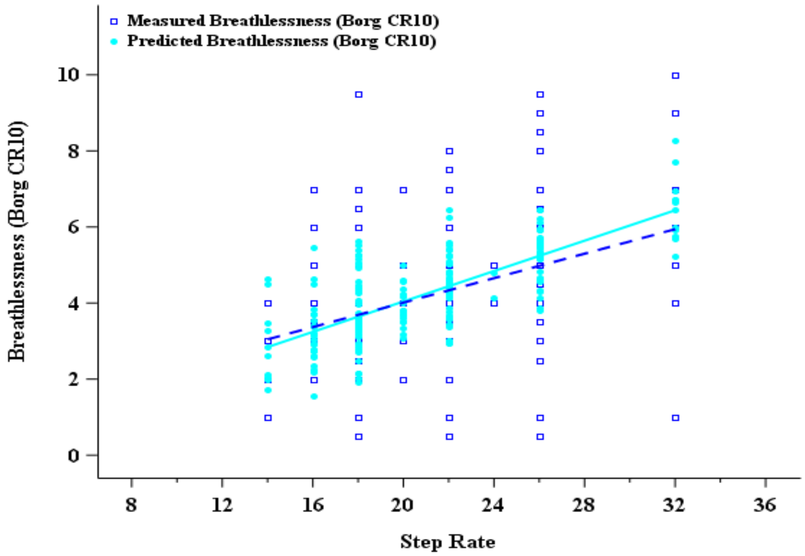
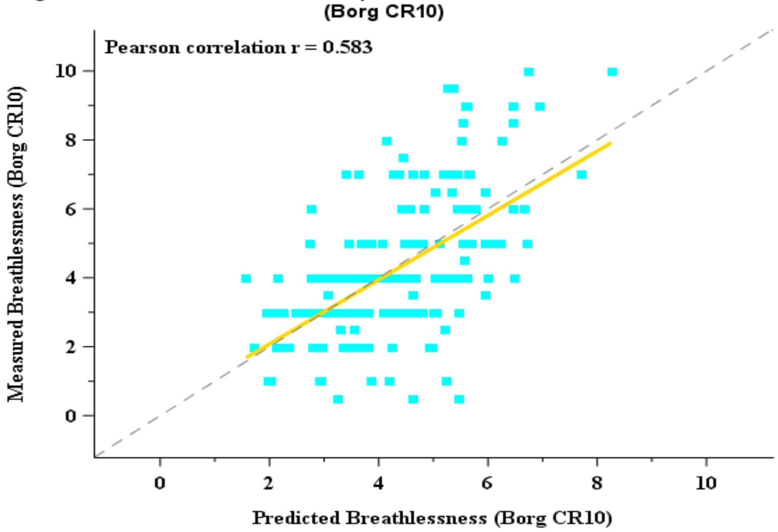



Figure 2. The association between predicted and measured Breathlessness (Borg CR10)



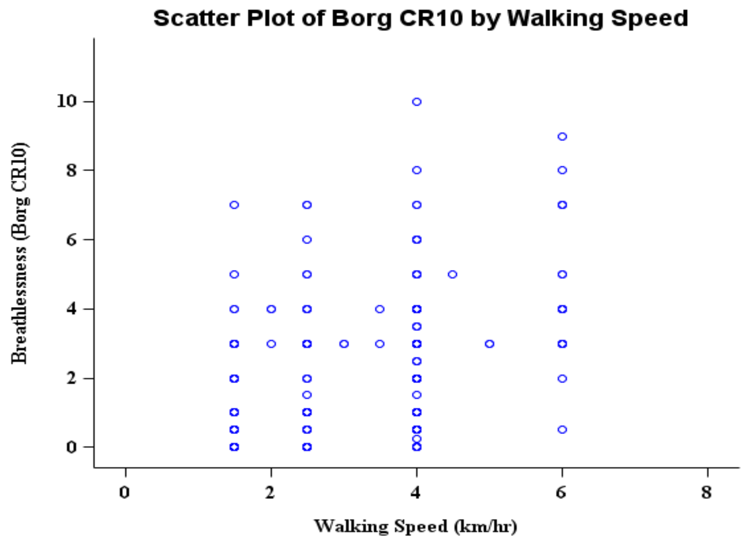

Mean Absolute Error = 1.22
Root Mean Squared Error = 2.46
Mean Percentage Error (MPE) = -26.9%

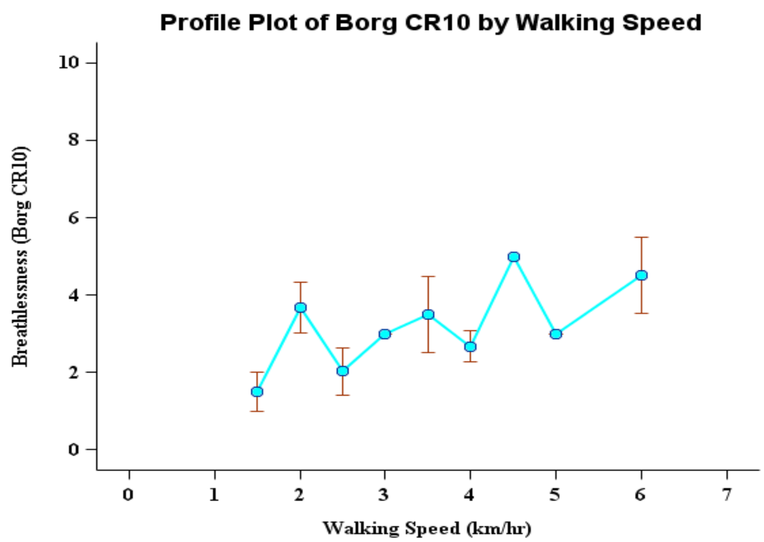
The MPE of -26.9% indicates that the equation systematically underestimates (negative error) the Borg rating.

## Appendix G

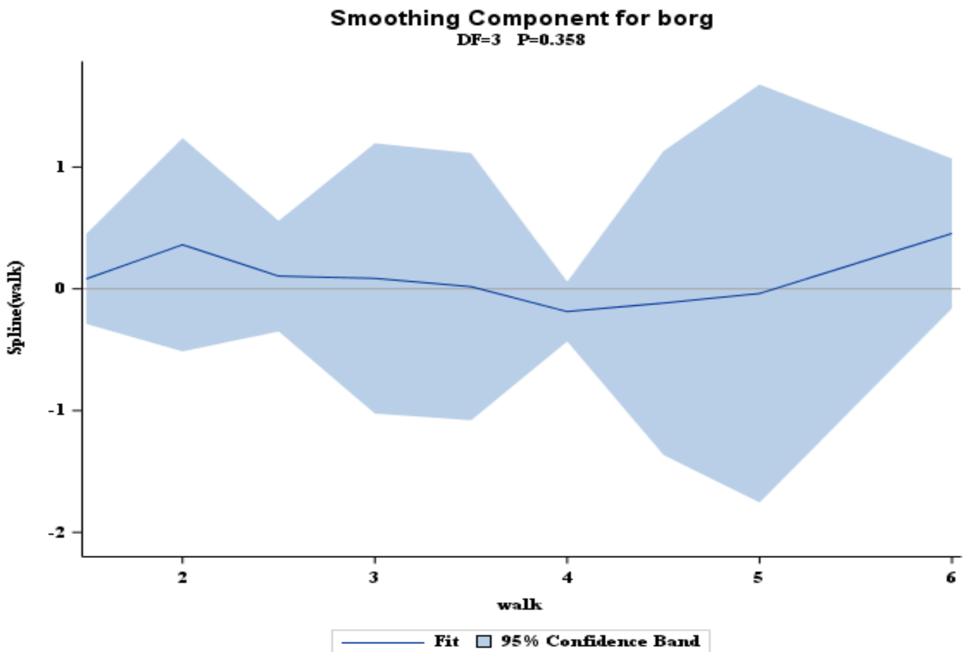
## 2nd Set of Equations: Derivation of the 3-min CSST Prediction Equation

**Step 1: Check for patterns between stepping rate and breathlessness**




The profile plot showed a linear increasing trend of Borg with walking speed. Then we used 3 degrees of freedom Generalized Additive Model to check the patterns between Borg and walking speed



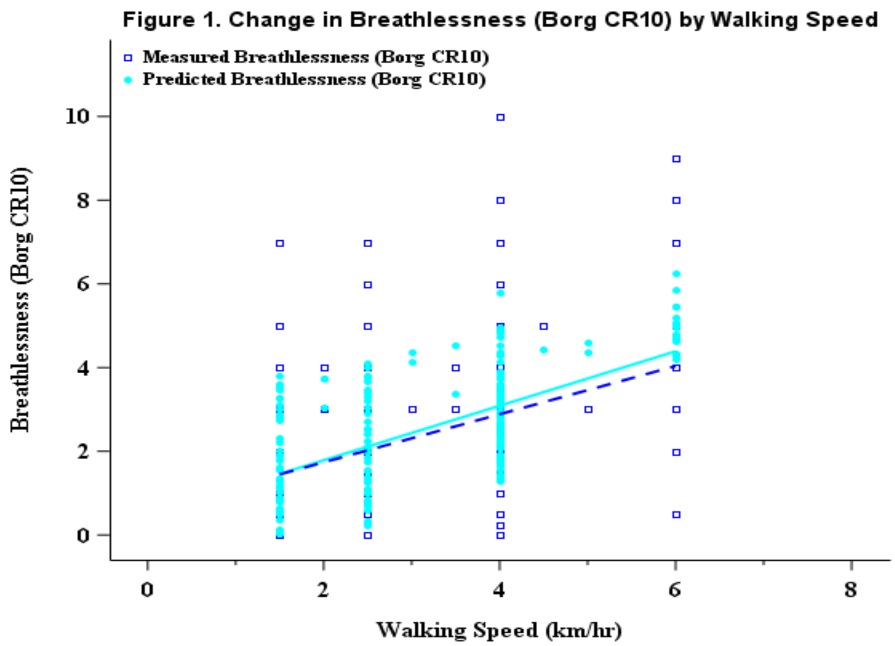

The Generalized Additive Model also indicated a linearity; the p-value for the test parameter for linear (Walk) < 0.001, significant; the nonlinear contribution was non-significant, with a p = 0.358.

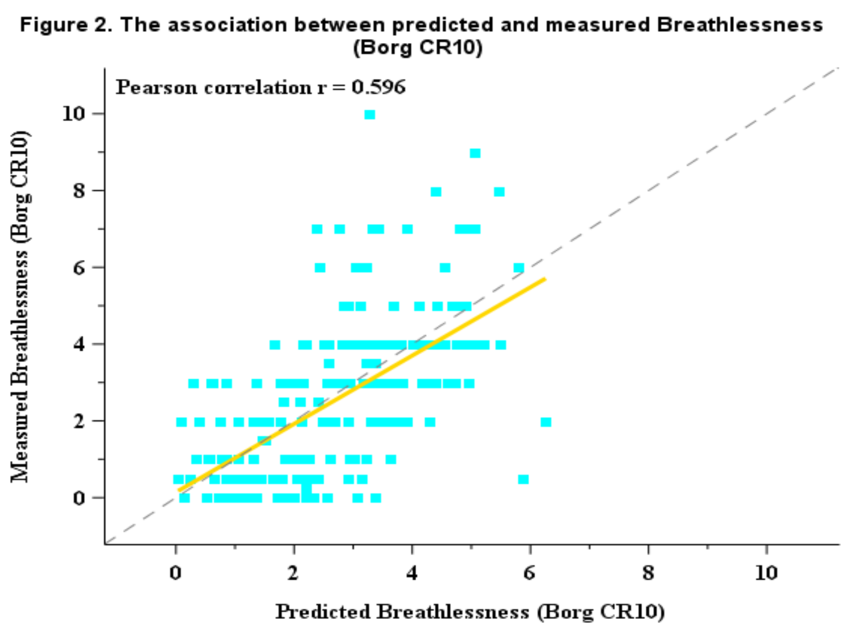
## Step 2: Development of the prediction model

| Table2. Change in Borg and Walking Speed | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Coefficient (SE) | P value | Coefficient (SE) | P value | Coefficient (SE) | P value | Coefficient (SE) | P value | Coefficient (SE) | P value |
| **Parameter (Fixed effects)** | | | | | | | | | | |
| Intercept | -2.129 (1.914) | 0.2684 | -2.453 (1.936) | 0.2076 | -2.530 (1.955) | 0.1984 | -3.308 (1.931) | 0.0895 | -3.242 (1.945) | 0.0983 |
| Walking Speed | -0.624 (0.263) | 0.0199 | -0.641 (0.264) | 0.0175 | -0.619 (0.261) | 0.0201 | -0.619 (0.263) | 0.0209 | -0.617 (0.263) | 0.0214 |
| Walking Speed*Walking Sp | 0.205 (0.036) | <0.001 | 0.206 (0.036) | <0.001 | 0.206 (0.036) | <0.001 | 0.205 (0.036) | <0.001 | 0.205 (0.036) | <0.001 |
| Male sex | -0.680 (0.361) | 0.0632 | -0.871 (0.355) | 0.0162 | -0.366 (0.369) | 0.3242 | -0.651 (0.344) | 0.0615 | -0.636 (0.346) | 0.0698 |
| Age | 0.036 (0.024) | 0.1387 | 0.034 (0.024) | 0.1623 | 0.031 (0.024) | 0.1998 | 0.046 (0.024) | 0.0621 | 0.045 (0.025) | 0.0747 |
| Height | | | | | | | | | | |
| Weight | 0.029 (0.010) | 0.0059 | 0.020 (0.010) | 0.0374 | | | | | | |
| BMI | | | | | 0.094 (0.032) | 0.0038 | 0.097 (0.031) | 0.0026 | 0.094 (0.032) | 0.0038 |
| FEV1, L | -0.726 (0.325) | 0.0278 | | | -0.667 (0.317) | 0.0376 | | | | |
| mMRC (1-4) | 0.740 (0.183) | <0.001 | 0.869 (0.176) | <0.001 | 0.748 (0.182) | <0.001 | 0.742 (0.179) | <0.001 | 0.756 (0.181) | <0.001 |
| FEV1, % predicted-Nhanes | | | | | | | -0.020 (0.008) | 0.0199 | | |
| FEV1, % predicted-GLI | | | | | | | | | -0.017 (0.008) | 0.0445 |
| GOLD Stages | | | | | | | | | | |
| GOLD1 | | | | | | | | | | |
| GOLD2 | | | | | | | | | | |
| GOLD3 | | | | | | | | | | |
| GOLD4 | | | | | | | | | | |
| **Model Assessment AIC** | 751.9 | | 756.4 | | 748.7 | | 754.8 | | 756.2 | |

We fit the model to look at the change of Borg over walking speed, including within-subject predictor (walking speed) and significant between-subject covariates. We checked all the covariates, including their interaction terms. For the final model we just kept the significant covariates and interaction terms. Sex and age were forced into the final model. No plethysmography parameters were significant predictors. The impact of including GOLD stage in the model was non-significant. Model 3 was the best fit.

**Step 3: Evaluation of the prediction model**




Mean Absolute Error= 1.28
Root Mean Squared Error= 2.71
Mean Percentage Error (MPE)= -49.2%

The MPE of -49.2% indicates that the equation systematically underestimates (more negative error) the Borg rating.