Epidemiology and evaluation of breathlessness - A data-driven approach

Olsson, Max

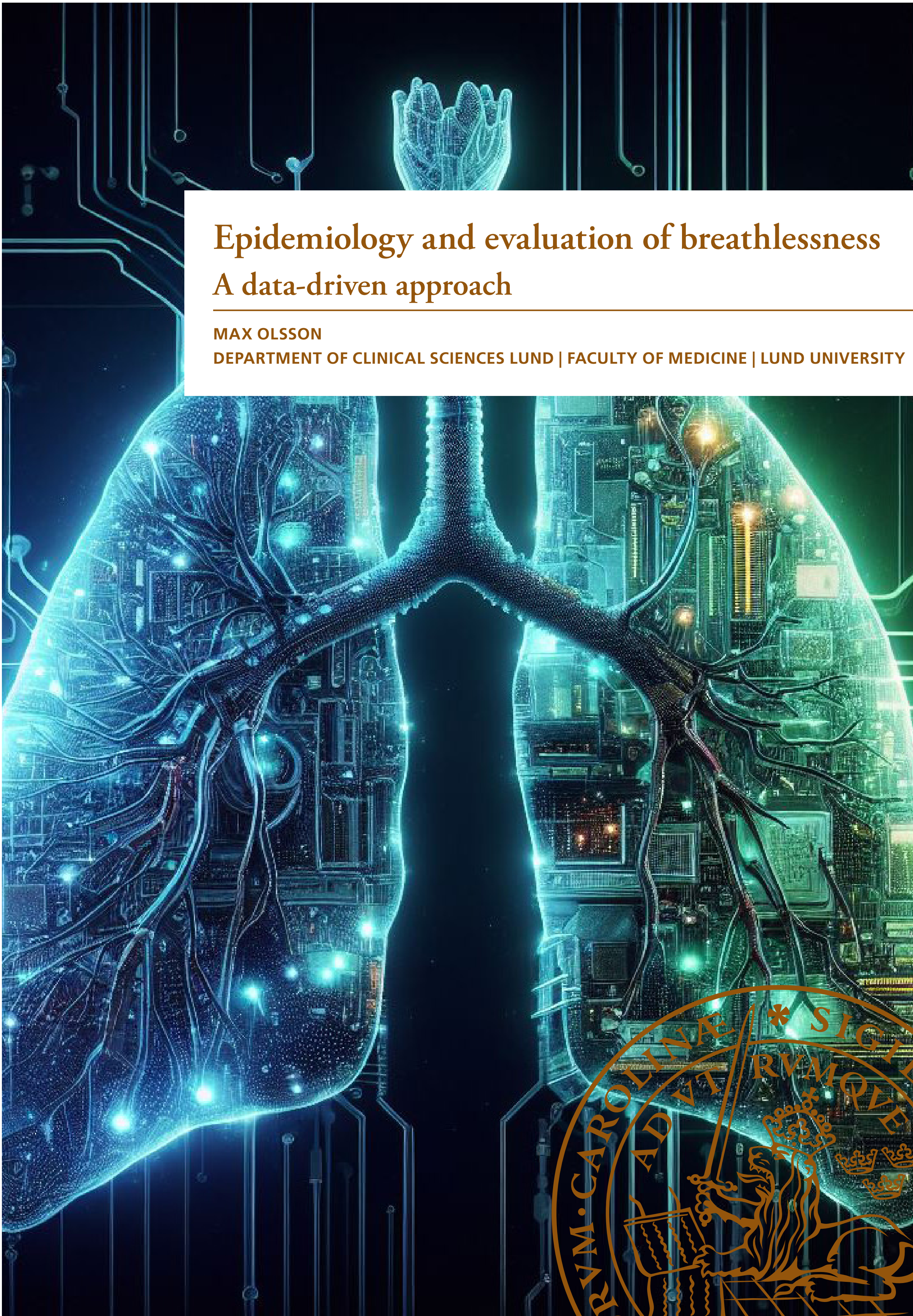
2024

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Olsson, M. (2024). *Epidemiology and evaluation of breathlessness - A data-driven approach*. [Doctoral Thesis (compilation), Department of Clinical Sciences, Lund]. Lund University, Faculty of Medicine.

Total number of authors:
1



Epidemiology and evaluation of breathlessness

A data-driven approach

MAX OLSSON

DEPARTMENT OF CLINICAL SCIENCES LUND | FACULTY OF MEDICINE | LUND UNIVERSITY

Epidemiology and evaluation of breathlessness – A data-driven approach.

Epidemiology and evaluation of breathlessness

A data-driven approach

Max Olsson, BSPH.

LUND UNIVERSITY

DOCTORAL DISSERTATION

By due permission of the Faculty of Medicine, Lund University, Sweden.

To be defended at St. Algatan Hörsal, St. Algatan 4, Lund, 13 September 2024 at 13.00.

Faculty opponent
Christine Jenkins, professor, Respiratory Group, The George Institute for Global Health, Sydney, Australia.

Organisation: LUND UNIVERSITY

Document name: DOCTORAL DISSERTATION **Date of issue: 13 September 2024.**

Author(s): Max Olsson

Sponsoring organisation: None

Title and subtitle: Epidemiology and evaluation of breathlessness – A data-driven approach.

Abstract

Background. Breathlessness is defined as the subjective experience of breathing discomfort and is a prevalent condition in the general population. The symptom compose different dimensions such as the physical, emotional, and affective dimensions, but the instruments measuring these dimensions have not been validated for use in epidemiological studies. Owing to the lack of validated instruments, knowledge of the epidemiology of breathlessness in different dimensions is very limited. Multiple underlying factors can contribute to breathlessness, but previous studies have not compared the importance of different contributing factors in the general population. In clinical practice, identifying the underlying condition(s) contributing to breathlessness among patients is challenging and costly, and new diagnostic pathways are needed.

Methods. Five studies were included in this dissertation. Study I involved the recruitment of an older male population for Studies II and III. In Study II, the psychometric properties of the multidimensional breathlessness instruments, the Dyspnoea-12 (D12) and the Multidimensional Dyspnea Profile (MDP), were evaluated. In Study III, the prevalence of different dimensions of breathlessness among older men in the population was evaluated. Study IV involved the use of machine learning to identify which factors associated the most to breathlessness in the middle-aged general population. Study V used reinforcement learning to develop a diagnostic pathway to identify as many underlying conditions contributing to breathlessness with as low economic costs as possible.

Results and conclusions. The D12 and MDP are valid and reliable tools for use in epidemiological studies. Different dimensions of breathlessness were highly prevalent in the older male population. Over one-third of the study population reported that their breathlessness experiences were unpleasant, and approximately one-fifth experienced breathlessness in a physical and/or emotional domain. In the middle-aged general population, the most important factors associated with breathlessness are obesity, reduced lung function, and a sedentary lifestyle. The developed diagnostic pathway suggests that it is most valuable to perform a primary care visit, followed by examinations of the lungs and the heart.

Keywords: breathlessness, psychometrics, epidemiology, machine learning, reinforcement learning.

Classification system and/or index terms (if any) Supplementary bibliographical information

Language: English **Number of pages**: 80

ISSN and key title: **1652-8220 Lund University, Faculty of Medicine Doctoral Dissertation Series 2024:101**

ISBN: 978-91-8021-597-8

Recipient's notes Price Security classification

Signature Date 2024-07-02

Epidemiology and evaluation of breathlessness

A data-driven approach

Max Olsson, BSPH.

LUND UNIVERSITY

Coverphoto by author in collaboration with Bing Designer.

Faculty of Medicine

Department of Clinical Sciences Lund

ISBN **978-91-8021-597-8**

ISSN **1652-8220**

Printed in Sweden by Media-Tryck, Lund University

Lund 2024

NORDIC SWAN ECOLABEL
Printed matter
3041 0903

Media-Tryck is a Nordic Swan Ecolabel
certified provider of printed material.
Read more about our environmental
work at www.mediatryck.lu.se

MADE IN SWEDEN

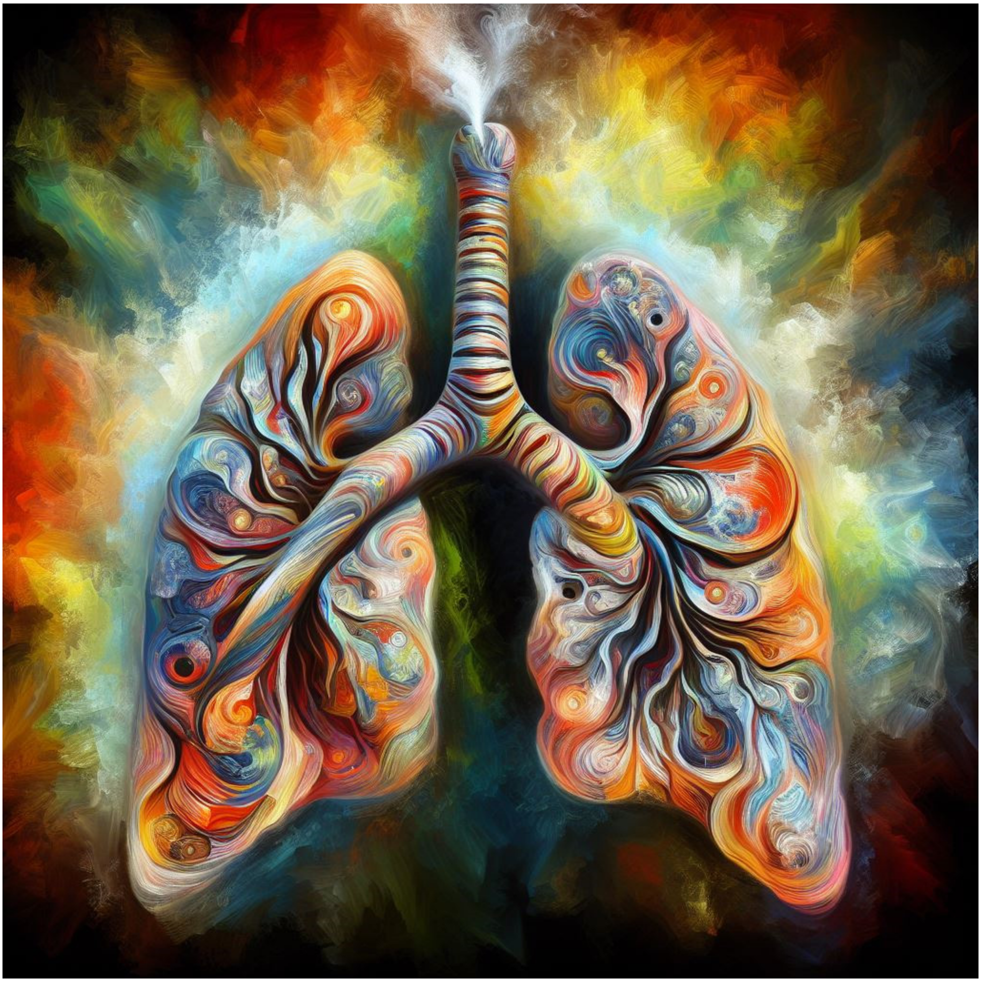


Image created by the author in collaboration with Bing Designer.

Table of Contents

Abstract

Background: Breathlessness is defined as the subjective experience of breathing discomfort and is a prevalent condition in the general population. The symptom compose different dimensions such as the physical, emotional, and affective dimensions, but the instruments measuring these dimensions have not been validated for use in epidemiological studies. Owing to the lack of validated instruments, knowledge of the epidemiology of breathlessness in different dimensions is very limited. Multiple underlying factors can contribute to breathlessness, but previous studies have not compared the importance of different contributing factors in the general population. In clinical practice, identifying the underlying condition(s) contributing to breathlessness among patients is challenging and costly, and new diagnostic pathways are needed.

Methods: Five studies were included in this dissertation. Study I involved the recruitment of an older male population for Studies II and III. In Study II, the psychometric properties of the multidimensional breathlessness instruments, the Dyspnoea-12 (D12) and the Multidimensional Dyspnea Profile (MDP), were evaluated. In Study III, the prevalence of different dimensions of breathlessness among older men in the population was evaluated. Study IV involved the use of machine learning to identify which factors associated the most to breathlessness in the middle-aged general population. Study V used reinforcement learning to develop a diagnostic pathway to identify as many underlying conditions contributing to breathlessness with as low economic costs as possible.

Results and conclusions: The D12 and MDP are valid and reliable tools for use in epidemiological studies. Different dimensions of breathlessness were highly prevalent in the older male population. Over one-third of the study population reported that their breathlessness experiences were unpleasant, and approximately one-fifth experienced breathlessness in a physical and/or emotional domain. In the middle-aged general population, the most important factors associated with breathlessness are obesity, reduced lung function, and a sedentary lifestyle. The developed diagnostic pathway suggests that it is most valuable to perform a primary care visit, followed by examinations of the lungs and the heart.

Populärvetenskaplig sammanfattning

I vanliga fall så märker du inte av din andning, den sköter sig själv för det mesta. Om du drabbas av en sjukdom kan du dock börja lida av andfåddhet. Andfåddhet består i flera olika upplevelser eller dimensioner som till exempel att lungorna känns trånga eller att ens andning känns obehaglig. Det finns olika enkäter för att mäta dessa dimensioner, men det är okänt hur bra dessa enkäter är att användas i befolkningsstudier. Eftersom enkäterna ej är testade i befolkningsstudier så finns det ej studier över hur vanligt förekommande olika dimensioner av andfåddhet är i befolkningen. Det är okänt vilka faktorer som är starkast associerade till andfåddhet i befolkningen eftersom tidigare studier mest fokuserat på att undersöka enstaka sjukdomars association till andfåddhet. Att identifiera vilka tillstånd och sjukdomar som bidrar till en patientens andfåddhet är svårt och oftast behövs många tester vilket kostar sjukvården stora mängder pengar.

Syftet med den här avhandlingen var att utvärdera enkäterna Dyspnoea-12 (D12) och Multidimensional dyspea profile (MDP) användbarhet i befolkningsstudier, undersöka förekomsten av olika dimensioner av andfåddhet i befolkningen, undersöka vilka faktorer som är starkast associerade till andfåddhet, samt utveckla en utredningsgång för att identifiera underliggande sjukdomar som bidrar till en patients andfåddhet.

Studie I beskrev datainsamlingen och rekryteringen av deltagare till Studie II och III. Studie II utvärderade användbarheten av andfåddhetsenkäterna i befolkningsstudier. Studie III kartlagde förekomsten av olika dimension av andfåddhet i en äldre manlig population. Studie IV använde maskininlärning för att identifiera faktorer som är associerade till andfåddhet i en medåldersbefolkning. Studie V använde artificiell intelligens (AI) för att utveckla en utredningsgång för andfåddhet med så låga kostnader som möjligt.

Genom dessa studier kunde vi visa att D12 och MDP är användbara enkätinstrument i befolkningsstudier om andfåddhet. Över en tredjedel av äldre män upplever sin andfåddhet som obehaglig. Övervikt, lungfunktion och stillasittande livsstil var de starkast associerade faktorerna till andfåddhet. Den AI-utvecklade utredningsgången visade oss att när underliggande faktorer till andfåddhet ska identifieras är det mest värdefullt att börja med ett primärvårdsbesök, följt av utredning av lungorna, och sedan utredning av hjärtat.

Abbreviations

A2C	Advantage actor critic
AI	Artificial intelligence
BMI	Body mass index
CFA	Confirmatory factor analysis
CT	Computed tomography
D12	Dyspnoea-12
DLCO	Diffusing capacity for carbon monoxide
ECG	Electrocardiogram
ER	Emotional response
FEV_1	Forced expiratory volume in 1 second
FVC	Forced vital capacity
IP	Immediate perception
MDP	Multidimensional Dyspnea Profile
ESAS-R	Edmonton Symptom Assessment System-revised
mMRC	modified Medical Research Council breathlessness scale
ML	Machine learning
RL	Reinforcement learning
SCAPIS	Swedish CArdioPulmonary bioImage Study
SF12	Short form-12
VASCOL	VAScular and Chronic Obstructive Lung study
XGBoost	eXtreme gradient boosting
QoL	Quality of life

Image created by the author in collaboration with Bing Designer.

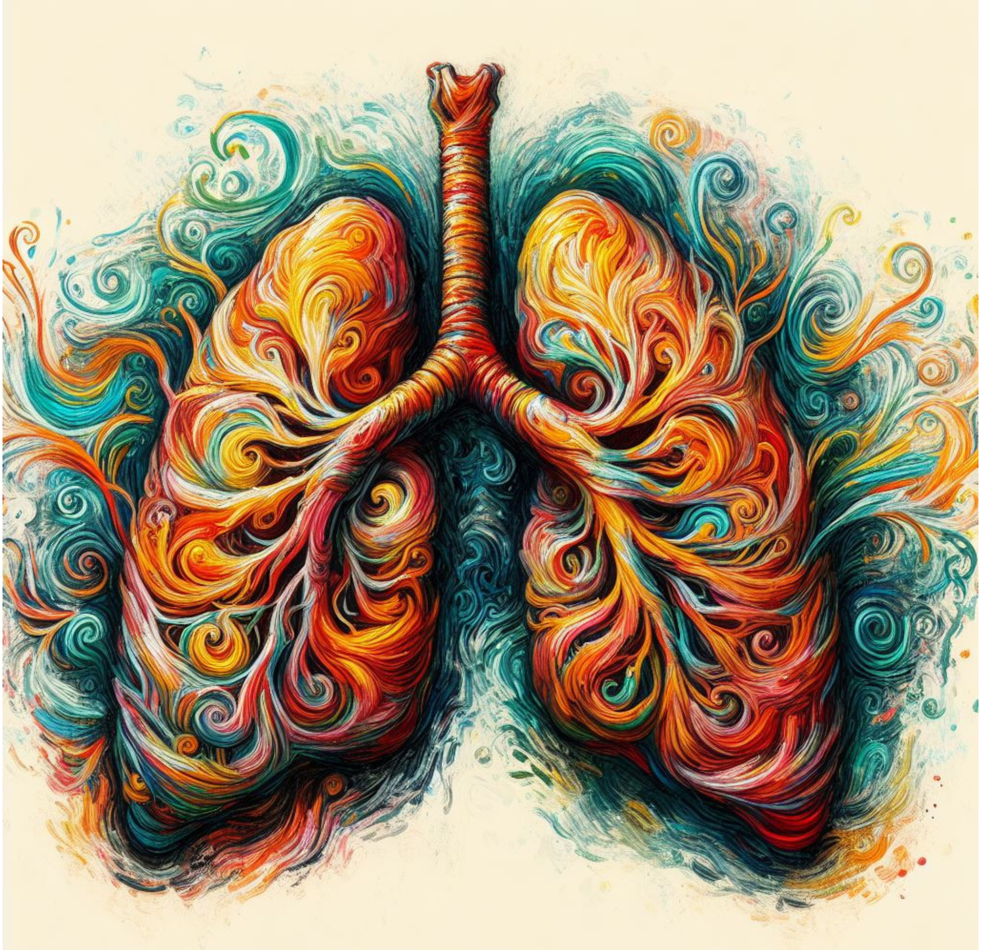


Image created by the author in collaboration with Bing Designer.

Preface

We need to breathe to live. Your mood can be affected through breathing, and just a simple deep breath can help you to calm down. However, there is no escape from breathing—even though you can hold your breath for a while sooner or later, you need to breath. When something that you need to do all the time to survive becomes unpleasant, it affects your whole life, which makes breathlessness an important topic to study. This is why I am motivated to study breathlessness, and ultimately, I hope this work can make a difference for all people who experience breathlessness.

I became interested in epidemiology while studying public health, and I became interested in breathlessness after completing an internship in my supervisor's laboratory. After graduating with my bachelor's degree in public health, I worked for a year as a database manager in an Alzheimer´s disease research group, where I started taking interest in artificial intelligence and machine learning.

The overall logic of the thesis is structured in steps that I find important when completing a larger research project. The steps include the measurement, evaluation of the prevalence and related factors, and examination of breathlessness. The first step is to ensure that breathlessness can be measured properly—without accurate measurements, we cannot capture a person's whole experience of their symptoms. The second step is to measure how common breathlessness is in the general population and focus on how people experience the symptoms. The third step is to study what factors are associated with breathlessness and understand the importance of different factors and their relationships with each other. The fourth step is to develop methods that could help relieve the burden of breathlessness.

Because my thoughts are often very visual, this dissertation includes multiple images, and you might have to flip through the dissertation a few times while reading to obtain a good understanding of the plots. If your thoughts are not as visual as mine—just read the text and trust me.

Now, take a deep breath, and enjoy this dissertation.

Introduction

This section first introduces the concept of breathlessness and then explains how breathlessness can be measured, the prevalence of breathlessness in the general population, and which factors are associated with breathlessness. The clinical evaluation of breathlessness is then introduced, followed by an introduction to artificial intelligence (AI).

Breathlessness

Breathlessness (or dyspnoea) is defined as the subjective feeling of breathing discomfort and is a common symptom of cardiovascular and respiratory diseases.[1] Breathlessness comprises multiple experiences, including physical experiences such as air hunger and chest tightness, as well as emotional and affective experiences such as anxiety and frustration.[2] This thesis focuses on chronic and persistent breathlessness, not temporary breathlessness caused by, for example, respiratory infection. It also focuses on unwanted breathlessness, not, for example, normal breathlessness due to exercise.

Under normal circumstances, humans breathe automatically without the need for any effort.[3] At the same time, breathing is something that humans can control if needed in comparison with many other bodily processes, such as the heartbeat. Breathlessness can be described as an interplay among the brain, body, and lungs.[3] Simply put, breathlessness occurs in individuals when the brain senses that the body needs more air than the lungs can provide and can result from a condition causing disturbances to the lungs or the body.[3] Breathlessness is a very subjective experience, and objective measurements of breathlessness severity often do not match the experienced severity.[4 5] Studies involving cycling on exercise bikes while wearing virtual reality googles have shown that the intensity of breathlessness can be altered on the basis of only the visual surroundings, for example, a hill,[6] without any change to the resistance of the bike (confirmed by the author in person). Past experiences and expectations of symptoms are very important for an individual's experience of breathlessness.[7]

Breathlessness is associated with multiple adverse health outcomes.[8 9] Breathlessness is associated with increased health-care seeking behaviour[10] and increased all-cause mortality in the general population,[11 12] and it is a better predictor of mortality than is lung function among people with chronic obstructive pulmonary disease (COPD).[13] The symptom is associated with lower physical capacity, decreased sexual well-being, anxiety, and depression, as well as overall lower quality of life (QoL).[14-18] In interview studies, persons affected by breathlessness described experiences of feeling helpless and a lack of control. [19] Often, the adverse effects of breathlessness lead to social isolation and inactivity, which ultimately leads to even more severe breathlessness.[20] Additionally, many patients do not want to discuss their breathlessness with their health professional[21], possibly because of the belief that it cannot be improved or that it is "deserved" because of lifestyle factors such as smoking. [22 23]

Today, symptomatic treatments for breathlessness are very limited,[19 24] and treating the underlying condition contributing to breathlessness is often the main focus for alleviating the symptom. Opioids have previously been suggested to reduce the symptom burden among people with breathlessness,[25 26] but later clinical trials failed to observe any improvements in breathlessness among COPD patients receiving low-dose, extended-release morphine.[24] Nonpharmacological treatments such as hand-held fan therapy,[27] low-intensity educational interventions,[28] behavioural strategies,[29] and breathlessness support services[30] have been shown to improve breathlessness, but studies have presented inconsistent and/or limited evidence of their effects.

Measuring breathlessness

The Medical Research Council breathlessness scale was introduced in 1952 to assess breathlessness among Welsh coal miners.[31 32] The scale has been modified over the years, and today, the *modified* Medical Research Council (mMRC)[33 34] breathlessness scale is the most commonly used instrument to evaluate breathlessness in population studies. The mMRC scale is scored on a categorical scale ranging from 0–4 (Table 1). For example, a mMRC grade of 0 indicates "*Breathless only with strenuous exercise,*" and the highest mMRC grade (4) indicates "*Too breathless to leave home or breathless when dressing*" (Table 1). In epidemiological studies of breathlessness, a mMRC grade of 2 ("*Breathless while walking with other people of your own age on level ground*") or higher is often considered to indicate clinically significant breathlessness and is often used as an outcome measurement. The limitations of the mMRC scale include that it only measures breathlessness in what situations the responder *believes* he or she would become breathless, and studies have shown that the severity of breathlessness is often misclassified by the mMRC scale and that patients misestimate in what situations they become breathless.[35 36] The mMRC scale also does not have a category for breathlessness at rest, which can be useful in palliative care settings.[37] Finally, the scale does not quantitatively measure the intensity of the symptom[1] and does not capture the multidimensional nature of breathlessness, such as the different dimensions of the symptom, including physical, emotional, and affective domains.

Table 1. mMRC breathlessness scale scores and corresponding statements

mMRC score	Corresponding statement
0	Breathless only when performing strenuous exercise
1	Breathless when hurrying or walking up a slight hill
2	Breathless while walking with other people of your own age on level ground
3	Need to stop to take a breath after 100 metres
4	Too breathless to leave home or breathless when dressing

The respondent is asked to state in which situations he or she would become breathless. There are no measurements of the severity of breathlessness experienced in these situations or how the respondent would feel during the experiences.

To better capture the experience of breathlessness, multidimensional instruments that can measure the intensity of symptoms in different dimensions are needed.[1 2] At present, there are two main instruments used for the multidimensional measurement of breathlessness: the Dyspnoea-12 (D12)[38] and the Multidimensional Dyspnea Profile (MDP).[39] The instruments are used to quantitatively measure physical and sensory qualities such as chest tightness, air hunger or feelings that one has to breathe a lot. The instruments can also measure emotional and affective symptoms such as depression, anxiety, and distress induced by breathlessness. [38 39] Both the D12 and MDP have been validated in multiple languages and are commonly used in clinical studies, where researchers often have good control over the scoring procedure[40 41 42]—benefits that are not often present in larger epidemiological studies using postal or online surveys. Despite the overall increasing popularity of the D12 and MDP,[40] the instruments are not psychometrically validated for use in epidemiological studies.

Prevalence of breathlessness in the general population

Breathlessness is common in the general population, and its prevalence increases with age.[43] Approximately 4–29% of middle-aged and older individuals experience breathlessness that limits their everyday lives.[8 20] Table 2 presents a summary of studies and the reported prevalence of breathlessness. There are large differences in the prevalence of breathlessness across countries. For example, the prevalence in China, Sweden, Norway, Germany, Canada, Austria, and Iceland was reported to be between 4 and 9%, whereas that in the Philippines, South Africa, Türkiye and the USA was between 20 and 29% when breathlessness was defined as mMRC score ≥2 points (*"Breathless while walking with other people of your own age on level ground").* Breathlessness is more common among women,[43] which is suggested to be related to women´s overall lower absolute lung volume.[44]

The knowledge about the epidemiology of breathlessness is, however, very limited, as most of the published studies are outdated and focus mainly on the experience of breathlessness reported by using the mMRC scale (Table 2). This is because of the lack of validated multidimensional instruments for breathlessness. At present, we have very limited knowledge about how common and intense physical and emotional experiences of breathlessness are at the population level, as well as how these dimensions are associated with the duration of and change in breathlessness over time.

Table 2. Overview of studies on the prevalence of breathlessness in the population

Country/Region	Breathlessness prevalence	Publication	Years of data collection	Breathlessness definition	Age	N
Australia	7%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	497
Australia	10%	National Breathlessness Survey	2019	mMRC grade ≥2	≥18	10,072
Austria	7%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	1202
Canada	7%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	793
China	5%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	461
Germany	4%	BOLD	2004-2005	mMRC score ≥2 points	≥40 years	457
Iceland	9%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	727
India	10%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	434
Norway	5%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	575
Philippines	22%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	825
Poland	24%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	481
South Africa	29%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	771
South America	45%	PLATINO	2003	mMRC grade ≥1	≥40 years	5314
Sweden	5%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	489
Sweden	10%	SCAPIS pilot	2012	mMRC grade ≥1	50–65 years	1097
Türkiye	23%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	640
UK	12%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	647
USA	20%	BOLD	2004-2005	mMRC grade ≥2	≥40 years	416

Prevalence data were derived from the studies by Gronseth et al. (2014)[43] (BOLD study), Poulos et al. (2021)[9] (National Breathlessness Survey), Sandberg et al. (2020)[8] (SCAPIS pilot study), and Lopez Varela et al. (2010)[45] (PLATINO study). Abbreviations: BOLD = Burden of Obstructive Lung Disease Initiative; PLATINO = Projeto Latino-Americano de Investigação em Obstrução Pulmonar; SCAPIS = Swedish CArdioPulmonary bioImage Study.

Factors contributing to breathlessness

Multiple factors can contribute to breathlessness among people in the general population. Table 3 presents the prevalence of factors among people experiencing breathlessness derived from a narrative review performed by Sandberg, Olsson, and Ekström (2021),[46] in which studies with a total of 96,614 people in the general population were included. COPD was the most common underlying factor, followed by asthma, arrythmia, heart failure, ischaemic heart disease, anxiety, depression and obesity.[46] There is a large overlap among these factors, which makes it challenging to understand which factor contributes the most to breathlessness.[46] For example, anxiety is known to contribute to breathlessness; if an individual experiences anxiety and has another serious health condition that is also known to contribute to breathlessness, it can be challenging to understand the associations of the two conditions with breathlessness. This is even more challenging with potentially complex, nonlinear associations. Some factors, such as body mass index (BMI), are known to have a U-shaped association with breathlessness, and both underweight and overweight are associated with breathlessness.[1]

Table 3. Prevalence of factors contributing to breathlessness.

Contributing factors	Prevalence
Angina pectoris	10%
Anxiety	32%
Arrythmia	34%
Asthma	42%
Chronic bronchitis	10%
Chronic obstructive pulmonary disease	85%
Depression	32%
Ischaemic heart disease	33%
Heart failure	33%
Obesity	28%
Reduced lung function	23%

Data taken from the review by Sandberg, Olsson & Ekström (2021)[46]

Studies have evaluated the associations between individual factors and breathlessness via traditional statistical methods but have not *compared* the strength of the associations between multiple factors and breathlessness.[46] Traditional statistical models, such as linear regression models, are limited in multivariate analyses measuring the strength of the associations between many factors and an outcome. Novel data-driven approaches can be used to analyse multiple factors with complex associations and rank the strength of the association with an outcome such as breathlessness.[47 48]

Clinical evaluation of breathlessness

As the interplay among various factors and breathlessness is complex, it is often challenging to evaluate breathlessness and identify the underlying condition(s) contributing to a patient´s breathlessness in clinical practice. Multiple medical tests are often required to determine the underlying condition(s) in primary and specialist clinics, leading to increased economic costs for health care.[49] In existing diagnostic pathways for breathlessness, it is often suggested to start by performing tests that are often available at primary care clinics, such as spirometry, followed by more advanced tests, such as thoracic computed tomography (CT), which are often available only in specialised clinics.[50]

At present, the recommended diagnostic pathways (order of tests) for evaluating breathlessness are often based on expert opinions[51] or studies with small sample sizes,[50] resulting in low empirical evidence of the prevalence of underlying conditions in the general population. Additionally, economic costs have not been the main focus in the development of diagnostic pathways, meaning that many of the diagnostic pathways are likely more expensive than necessary. Most diagnostic pathways assume that only one underlying condition contributes to the patient´s breathlessness,[50] but as mentioned before, many people with breathlessness have multiple underlying conditions.[46] This means that many of the existing diagnostic pathways could incorrectly identify "*the main cause of breathlessness*" but, at the same time, fail to identify other treatable but potentially dangerous underlying conditions.

Because of the complexity of evaluating breathlessness, modern AI methods could be useful for developing new diagnostic pathways using epidemiological data of underlying conditions in the general population and the economic costs of medical tests.

Artificial intelligence

AI is the attempt to develop computer software that can mimic or improve upon human intelligence.[52] Many attempts to create software capable of performing human tasks requiring intelligence have been made, but owing to the lack of computer power and overoptimistic expectations, it has often led to a lack of funding and disappointment, which is often described as an "AI winter."[52] In 1950, Alan Turing published the seminar paper *"Computing Machinery and Intelligence"*,[53] which introduced the *Turing test*. The Turing test aims to determine whether a machine is capable of intelligence equal to or better than that of humans. In the test, a human evaluator communicates questions via text to an AI computer, and if the human fails to determine whether the answers are written by a human or a computer, the computer passes the test. Recently, it has been determined that the chatbot ChatGPT has passed the Turing test,[54] but it is debated whether ChatGPT can reason or only imitate, similar to a parrot.[55] Currently, we are living in an AI boom (or AI summer), with new achievements in AI research occurring at a high rate.

There are different types of approaches to AI, and this dissertation focuses on the subgroup of AI called machine learning (ML), which is the driving technology in the current AI boom. The term "ML" was coined in 1959 by Arthur Samuel,[56] but terms such as "self-teaching computers" have also been used.[57] ML includes algorithms that use data to perform tasks without explicitly programmed instructions.[58] These tasks include, for example, classification, pattern recognition, or the performance of human-like tasks. The development of ML algorithms, data power, and larger population-based studies with more variables in later years promoted an increased number of studies using ML in the field of medicine.[59] For example, ML has been used in the medical field for the diagnosis of diseases and phenotyping of diseases.[59] ML methods have the ability to learn from complex data sources such as raw images, audio, or text. For traditional statistical models, measurements derived from these data types would have to be used by humans, such as measurements from a medical image or the number of times a certain word occurs in a text. This means that humans are providing explicit instructions to the model, and humans are aware of what could be important in the data before providing it to the model. By using tabular data, ML has the potential to learn from complex, nonlinear associations, use many variables and understand which variables are important. Reinforcement learning (RL), a subtype of ML, can be used for more human-like tasks, such as driving cars or playing chess,[60] but its applications in the medical field are currently limited.[61] It is hypothesised that RL can be used to develop optimal diagnostic pathways for evaluating symptoms and diseases.

In statistical theory, measurements such as odds ratios or correlation coefficients would be used to interpret the associations between variables and an outcome. However, when ML algorithms are used with a large set of variables, the task of interpreting associations among variables has often been an overwhelming challenge. This problem is often referred to as the "*black box problem*"—because we do not know what is going on "inside" the model.[62] Especially in academic research, there is a need to not only report the accuracy of a model's prediction but also explain *why* an algorithm makes a certain prediction.[62] Explaining "why" is not necessary in many AI projects conducted by private companies, where the focus is merely on the accuracy of a prediction, such as by presenting you with an online advertisement that you click on or presenting a song on Spotify that you listen to. In recent years, novel methods to simplify the interpretation of ML models have been developed that could reduce the black box problem for numerous ML algorithms.[63]

Rationale

Previous studies on the prevalence of breathlessness have chiefly measured breathlessness with the mMRC scale, which captures only situations in which an individual believes he or she would experience breathlessness and not the multidimensional nature of breathlessness, including physical, emotional, and overall unpleasant experiences. The available multidimensional instruments for evaluating breathlessness are not validated for use in epidemiological studies; therefore, no studies on the prevalence and severity of symptoms in the different dimensions of breathlessness at the population level exist today.

Previous studies assessing factors associated to breathlessness have focused mainly on individual health conditions and have used general statistical models to interpret associations. When multiple underlying factors that can potentially contribute to breathlessness exist, more advanced methods, such as ML algorithms, could be useful for revealing complex, nonlinear associations with breathlessness in the general population. If we focus on the clinical setting, it is challenging to identify the underlying condition(s) leading to breathlessness, which often results in multiple tests that increase health care costs and waiting times for patients. The existing strategies for evaluating breathlessness in clinical practice lack evidence, and we need to develop new diagnostic pathways to determine the underlying condition(s) leading to breathlessness that are based on the prevalence of health conditions in the general population and the economic costs of the tests.

Aims

Overall aim

The general aim of this thesis was to gain knowledge about breathlessness in the population in terms of the psychometric properties of multidimensional instruments for evaluating breathlessness, the prevalence of different breathlessness dimensions, factors associated with breathlessness, and how breathlessness is best evaluated with respect to economic costs.

Specific aims

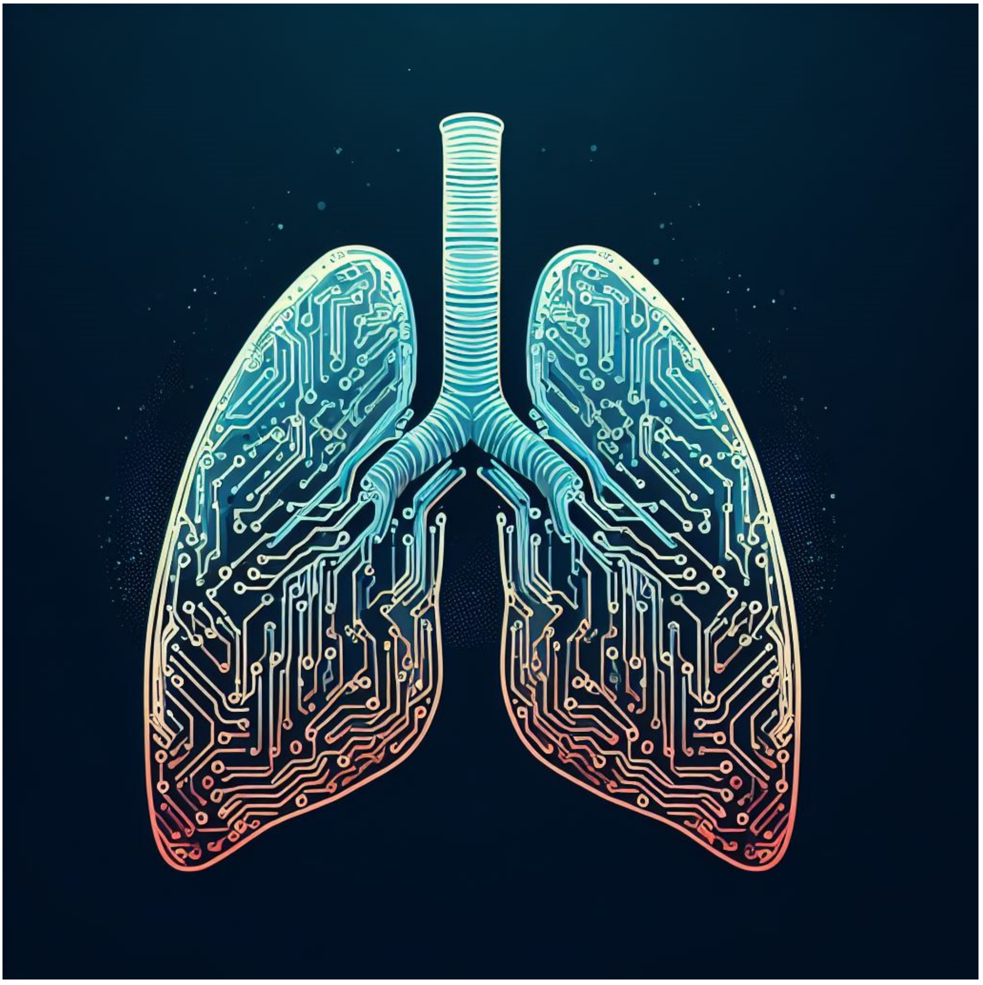
I. To report the design and recruitment of the first two waves of a longitudinal prospective cohort study including older men.

II. To validate the D12 and MDP for use in 73-year-old men in terms of the underlying factor structures, internal consistency, and validity.

III. To determine the prevalence and intensity of different dimensions of breathlessness among elderly males and any associations with the duration, change over time and mMRC grade.

IV. To identify the factors most strongly associated with breathlessness in the general population and to describe the shapes of the associations between the main factors and breathlessness.

V. To develop a diagnostic pathway to identify underlying conditions leading to breathlessness with as low economic cost as possible.



Techniques to explore breathlessness. Image created by the author in collaboration with Bing Designer.

Methods

This section introduces the populations and methods used in the substudies of this doctoral dissertation. First, the data collection and populations used in the substudies are described, and then the analysis methods used in each substudy are described separately. An overview of the methods used in the substudies is presented in Table 4.

Table 4. Overview of the studies included in the thesis.

Paper	Population	N	Analysis method	Main tests/algorithm
I	VASCOL study, wave 1 and 2	1302	Descriptive statistics	Mean, standard deviation, and %
II	VASCOL study, wave 2	684	Psychometric testing	Confirmatory factor analysis, Cronbach´s alpha, and Pearson´s correlation.
III	VASCOL study, wave 2	907	Traditional statistics	% and Linear regression
IV	SCAPIS	28,730	Supervised machine learning	eXtreme Gradient Boosting (XGBoost)
V	SCAPIS	1131	Reinforcement learning	Advantage actor critic (A2C)

Data collection for wave 1 of the VASCOL study was conducted between 2010 and 2011, that for wave 2 of the VASCOL study was conducted in 2019, and that for the SCAPIS was conducted between 2013 and 2018. All the analyses were cross-sectional. Abbreviations: VASCOL = VAScular and Chronic Obstructive Lung disease; SCAPIS = Swedish CArdioPulmonary bioImage Study.

Description of the VASCOL study population (Papers I - III)

The VAScular and Chronic Obstructive Lung disease (VASCOL) study is a longitudinal cohort study of older men's health with a focus on vascular diseases and COPD that began in 2010. The men were initially recruited from a screening campaign for abdominal aortic aneurysms conducted in Blekinge County, Sweden. Men aged 65 years who were living in Blekinge between 2010 and 2011 were invited to participate in the screening campaign. In the invitation letter for the screening, the men were also invited to participate in wave 1 of the VASCOL study. If the men agreed to participate in wave 1, spirometry and weight and height measurements were performed at the study site in Karlshamn. The participants also completed a survey about their lifestyles and health conditions at home and brought the completed survey to the study site.

In 2019, the men who were still alive and had a known address were invited to participate in a postal follow-up survey (wave 2). The focus of wave 2 of the VASCOL study focused more on breathlessness, other symptoms, lifestyle habits, QoL, and health conditions. The wave 2 survey comprised multiple questions to collect data on the following: self-reported weight, height, smoking history, perception of the duration of and changes in breathlessness, and health conditions. The Swedish versions of the following instruments were used in wave 2: the D12,[38] MDP,[39] mMRC scale,[33] Hospital Anxiety and Depression Scale (HADS),[64] Functional Assessment of Chronic Illness Therapy-Fatigue Scale (FACIT-fatigue),[65] Edmonton Symptom Assessment System-revised (ESAS-R)[66] and Short-Form-12 health survey (SF-12).[67]

Description of the SCAPIS population (Papers IV and V)

The Swedish CArdioPulmonary bioImage Study (SCAPIS) was a cross-sectional multisite study including approximately 30,000 men and women aged 50 to 64 years from the Swedish general population. The objectives and data collection of the SCAPIS have been described in detail previously.[68] Data were collected between 2013 and 2018 at six different study sites: Lund/Malmö, Gothenburg, Linköping, Stockholm, Uppsala, and Umeå. At the study sites, data were collected from blood sample testing, CT, and physiological measurements. The participants also completed an extensive survey regarding their baseline characteristics, lifestyle habits, health conditions, and QoL. The participation rate in the SCAPIS was approximately 50%. The SCAPIS study population was considered generalizable for the Swedish population of the same age, with the exception that the SCAPIS population consumed more alcohol than the remaining Swedish population of the same age.[69]

Paper I – Description of the data collection and requirements of the VASCOL study

As this study only served as a protocol study for Study II and Study III, the results are only presented briefly. The data collection of the two waves of the VASCOL study were described with the frequency of participation and the dropout rate at each stage of data collection. The characteristics of the participants who participated in wave 1 were compared with those of the participants who participated in wave 2 and those who did not (those who were deceased, whose address was unknown, or who did not return the survey).

Paper II - Psychometric testing of the D12 and MDP

Paper II was a psychometric evaluation study of the D12 and MDP with data from wave 2 of the VASCOL study. The aim of psychometrical evaluation is to evaluate the performance of survey instruments for psychological experiences,[70] such as experiences of a specific symptom. With psychometrics, we can answer questions such as "Does the instrument measure what it intends to measure?", and "Is the instrument reliable?". Psychometric properties that are often of interest include underlying factor structures, internal consistency, and external/concurrent validity. The inclusion criterion for this analysis was complete data on the D12 and MDP. All analyses were conducted in R 4.0.2 (R Foundation for Statistical Computing, Austria).

Assessments

The D12 comprises twelve items evaluating different dimensions of breathlessness, and each item is scored on a scale ranging from 0–3 points: "None" (0), "Mild" (1), "Moderate" (2), or "Severe" (3). The D12 can be summarised as the total score (range 0–36 points) and physical (range 0–21 points) and affective (range 0–15 points) domain scores. The MDP has eleven items scored on a scale ranging from 0–10 points. The MDP can be summarised as the immediate perception (IP) score (six items) and emotional response (ER) score (five items). To describe overall breathlessness via the MDP, it is recommended to use subitem A1 (unpleasantness) to describe the overall unpleasantness of breathlessness.[39]

Factor analysis

Underlying factor structures relate to how individual items of a survey instrument relate to composite factors. This is especially important when one wants to summarise multiple items into a domain that describes something that can be used as an outcome variable.[70] In this dissertation, the original underlying factor structures of the D12[38] and MDP[39] were evaluated via confirmatory factor analysis (CFA). CFA is often used when a factor structure is based on the literature, whereas exploratory factor analysis can be used when a new survey instrument is developed. A CFA provides a factor loading estimate, which is a value of how much an item relates to its subdomain, and a higher score means that the item relates more to its subdomain. The CFA models were evaluated via the root mean square error of approximation (RMSEA) and Bentler´s comparative fit index (CFI), both of which are measurements of goodness of fit.

Internal consistencies

The internal consistencies of the D12 and MDP total scores and subdomain scores were measured via Cronbach's alpha. Cronbach's alpha considers the variance in the items and total score as well as the number of items included. Cronbach's alpha most often is a value between 0 and 1; a value above 0.8 is considered good, and a value above 0.9 is considered excellent.[71] A better internal consistency suggests that an instrument is more reliable.

Validity

Validity relates to the ability of an instrument to measure what it was developed to measure. One approach to measure instrument validity is to evaluate how an instrument correlates with other already established instruments, so-called *external* validity. The validity of the D12 and MDP was measured via Pearson's correlation, which yields a correlation coefficient between -1 and +1, indicating how much of a dependent variable is explained by an independent variable.[71] The correlations among D12/MDP total scores, subdomain scores and the following measurements were determined: the mMRC score, ESAS-R breathlessness intensity (0–10), SF-12 physical and mental quality of life scores, HADS total score, anxiety and depression scores, FACIT-fatigue score, and BMI.

To evaluate whether the instruments were able to capture physical and emotional/affective domains separately, so-called discriminative validity, the domain scores were correlated with measurements that are intended to measure factors of a physical or emotional/affective nature, respectively. For example, the D12 physical score should correlate more strongly with the SF-12 physical score than with the D12 affective score. The physical breathlessness domain scores (MDP IP domain and D12 physical domain) are hypothesised to correlate more strongly with the mMRC score, SF-12 physical score and BMI than the emotional/affective breathlessness domain scores (MDP emotional domain and D12 affective domain). Similarly, it is hypothesised that the emotional/affective breathlessness domain scores correlate more strongly with the SF-12 mental score, HADS total score and anxiety and depression scores than the physical breathlessness domain scores do.

Paper III - Prevalence of s different dimensions of breathlessness

In this study, data from wave 2 of the VASCOL study were used to study the prevalence and intensity of symptoms in different dimensions of breathlessness and how these dimensions are associated with the duration of and changes in breathlessness. There were no specific inclusion criteria for this study, but data completeness varied between the instruments, meaning that the number of participants in each subanalysis varied. All analyses were conducted in R 4.0.2 (R Foundation for Statistical Computing, Austria).

Assessments

Breathlessness was evaluated via the mMRC scale, D12 and MDP. Changes in breathlessness since the age of 65 years were recalled and reported via the following response options: very much better (1), much better (2), minimally better (3), no difference (4), minimally worse (5), much worse (6) or very much worse (7). For the change in breathlessness, the response options were categorised as: better (1–3), no different (4) or worse (5–7). The participants also reported how many years they experienced breathlessness, reported as a duration of <1 year, 1–5 years, and >5 years.

Prevalence and intensity

To determine the prevalence of a factor such as breathlessness in a study, a definition of what is considered clinically significant was needed. As most previous prevalence studies of breathlessness have used the mMRC scale, the definition of clinically significant breathlessness corresponds to one of its categories for which a participant believes he or she would experience breathlessness (Table 1). When continuous scales such as the D12 or MDP are used, it can be more challenging to define clinically significant breathlessness. The minimal clinically important difference (MCID)[72] is a concept used to define a significant difference in a score that could be important for a participant´s experience in comparison to *statistically* significant differences. The MCIDs for the D12 and MDP were established in a previous study[73 74] and used as thresholds to determine the presence of breathlessness among the participants. The MCID for the D12 total score was 2.83, that for the D12 physical subdomain score was 1.81, and that for the affective subdomain score was 1.07. The MCID was 0.82 for the MDP A1 unpleasantness score, 4.63 for MDP IP domain score, and 2.37 for the MDP ER domain score.[73 74]

The presence of different breathlessness domains is expressed as the frequency and percentage. For the intensity of different experiences of breathlessness in the study population, the mean and standard deviation (SD) are reported for the MDP items in a bar plot among participants experiencing an MCID above or equal to an MDP A1 unpleasantness score of one.

Associations among the mMRC grades, changes in breathlessness and duration of breathlessness

To measure the associations among the independent variables (mMRC grade, changes in and duration of breathlessness) and the dependent variables (D12 and MDP domain scores), linear regression was used with reference categories of the independent variables. The reference categories were used to evaluate differences in the D12 and MDP domain scores among the categories of the independent variables. A grade of 0 was used as the reference for the mMRC and compared with mMRC grade of 1–4. "Better" was used as the reference for changes in breathlessness and was compared with "no difference" and "worse." A duration of "1–5 years" was used as the reference for the duration of breathlessness and was compared to "<1 year" and ">5 years". For the analyses of the changes in and duration of breathlessness, participants with no breathlessness (mMRC grade of 0) and those who reported that their breathlessness was unchanged were excluded. The strengths of the associations derived from the linear regressions are reported with beta coefficients with 95% confidence intervals (CIs). A beta coefficient represents the estimated value of a dependent variable given the value of an independent variable. To compare scales (domain scores) of different ranges, the beta coefficients were scaled by dividing the coefficients by the maximum possible value of the scales, expressed as a percentage of the maximum range.

Paper IV - Identifying factors associated with breathlessness via supervised machine learning

In this study, the data from the SCAPIS were used to classify participants as having breathlessness or not having breathlessness on the basis of different factors (variables) via supervised ML. The strength of the associations among factors related to breathlessness was then measured. The exclusion criteria were the inability to walk for reasons other than breathlessness or the inability to understand written and spoken Swedish. All analyses were conducted in R version 4.1.2 (R Foundation for Statistical Computing, Austria).

Assessments

Breathlessness was defined as an mMRC grade ≥2. A total of 449 factors were used to classify breathlessness, including the following: self-reported health conditions; BMI; spirometry results; CT-based cardiac and lung measurements; blood sample test results; self-reported physical activity levels; accelerometery results; and other symptom, treatment, and socioeconomic characteristics.

As the table with the description of all the factors used in Study IV takes up 14 pages, it is not included in this dissertation but is provided in the supplementary materials for the published article (https://openres.ersjournals.com/content/10/2/00582-2023).

Supervised machine learning

Supervised ML has emerged as a novel method that can reveal nonlinear associations between factors and outcomes via many factors. The term "supervised" refers to the fact that the participants were classified by humans; in this study, participants were classified as experiencing "breathlessness" or "no breathlessness," and the algorithm was then trained to classify the participants on the basis of the factors in the dataset. In supervised ML, a model is typically trained on a training set of the data and then evaluated on another set of the data that the model has not seen before, often called the test or validation set.[75] During training, the settings of the algorithm, the so-called hyperparameters, are often tuned to achieve better performance.

A common problem in ML studies is overfitting, which means that the model used overlearns the data it is trained on, meaning that it cannot make generalisations for data it has not seen before. When a model is overfit, it classifies the outcome well on data it has been trained on, but when classifying data it has not been seen before, model accuracy decreases considerably.[75] This can be evaluated by examining the difference in classification accuracy between the training set and the validation set.

In this study, the ML algorithm eXtreme Gradient Boosting (XGB)[76] was used to classify participants as experiencing breathlessness or no breathlessness. XGB is a modern gradient boosting tree-based algorithm that uses regularisation to prevent overfitting. Data from the Stockholm study site were used as the test set (17% of the participants), and data from the remaining study sites were used as the training set. The model was developed by tuning the hyperparameters to achieve better performance with fivefold cross-validation on the training set. The model was ultimately evaluated on the test set and reached an area under the curve (AUC) of the receiver operating characteristic (ROC) of 0.81, with a sensitivity of 0.73 and a specificity of 0.89. This means that the model can make generalisations for unseen data.

Associations between different factors and breathlessness

To explain the strength of the associations among the different factors used to classify breathlessness, SHapley Additive exPlanations (SHAP)[63] values were used. SHAP values measure how much a value of an individual factor changes the probability of being classified as experiencing breathlessness among the participants, meaning that each participant receives a SHAP score for each factor. The SHAP score can be negative, which corresponds to the value of the factor decreasing the probability of being classified as experiencing breathlessness, or positive, which corresponds to the value of the factor increasing the probability of being classified as experiencing breathlessness. For example, a participant can have a BMI of 32 (obese), which gives a SHAP score of +0.3, increasing the probability of experiencing breathlessness; if a participant has a BMI of 22 (normal weight), the SHAP score is -0.2, meaning that the score decreases the probability of the participant being classified as experiencing breathlessness. The SHAP scores of factors can differ among participants who have the same value for a factor, which is based on the participants' values of the other factors. For example, a BMI of 27 could lead to an increased SHAP score among men but not among women, meaning that a BMI of 27 is considered to contribute more to the probability of experiencing breathlessness among men than among women. This means that the algorithm can consider the synergistic effects of factors.

As SHAP scores only describe the strength of the association of factors for individual participants, a measurement of the *overall* strength of an association is needed to describe the importance of each factor to breathlessness. To determine the overall strength of the association for each factor, the absolute mean values of the SHAP scores were calculated for all the factors among all the participants.

To interpret the shape of the associations, plots were created with the participants' factor values on the X-axis and the probability of experiencing breathlessness on the Y-axis. To simplify these interpretations, locally estimated scatterplot smoothing lines were added to the plots, which shows an average association between the X-axis and Y-axis over the plots.

The methodological approach was previously validated in a pilot study[77] outside this dissertation using data from wave 2 of the VASCOL study to explore the importance of factors related to self-perceived health and, later, in a study exploring factors related to health-related QoL using data from the SCAPIS.[18]

Paper V - Developing a diagnostic pathway for identifying health conditions for breathlessness via reinforcement learning

In this study, the data from the SCAPIS were used to develop a diagnostic pathway for identifying underlying health conditions leading to breathlessness. The inclusion criterion was having clinically significant breathlessness, which was defined as an mMRC score ≥2 points. The exclusion criteria were the inability to walk for reasons other than breathlessness or the inability to understand Swedish. All analyses were conducted in Python version 3.9.15 (Centrum voor Wiskunde en Informatica, Netherlands).

Underlying conditions and patient categories

The underlying conditions were first listed in a narrative review performed outside this work.[46] Then, ten researchers with different backgrounds independently provided feedback on the list of underlying conditions (clinical physiology, respiratory, cardiology, and family medicine). Underweight/obesity and anaemia were defined on the basis of World Health Organization (WHO) standards.[78] The following underlying conditions were defined on the basis of previous studies: chronic airflow limitation (CAL),[79] restrictive spirometry pattern,[80] pulmonary fibrosis (PF) or fibrotic interstitial lung abnormality ILA (FILA),[81] reduced diffusion lung capacity,[82] pulmonary emphysema,[83] stress, depression,[84] and heart failure.[85] See Table 5 for the final included underlying conditions and their definitions.

The tests used in the SCAPIS to define the underlying conditions are not necessarily the tests that would be used to identify underlying conditions in health care settings. Therefore, to realistically calculate the costs of identifying underlying conditions in health care, a separate list of tests that would be used in health care settings was defined by the researchers. The cost for each test in health care settings was then derived from a price list from the Scania Regional Council[86] (Table 5). The 'primary care visit' tests included basic examinations, including examinations assessing underweight, obesity, sedentary lifestyle, depression and stress status.

Table 5 – Underlying conditions contributing to breathlessness, tests to identify the underlying conditions, and costs associated with the tests.

Condition	Definition of the presence of the condition in the SCAPIS database	Test to identify the condition in clinical care	Cost of the test(s) needed to identify the condition in clinical care (SEK and Euro)	
Chronic airflow limitation (CAL)	Post bronchodilator FEV_1/FVC < LLN	Spirometry	873 SEK	€76
Restrictive spirometry pattern	FVC < LLN and FEV_1/FVC ≥ LLN post bronchodilation.	Lung volume and diffusion capacity	2209 SEK	€193
Asthma	Self-reported of physician diagnosis.	Spirometry	873 SEK	€76
Anaemia	Haemoglobin level <110 g/L	Blood sample testing	289 SEK	€25
Age 63+ years (potential age-related impairment)	63+ years	Primary care visit	0 SEK	€0
Obesity	BMI ≥30	Primary care visit	0 SEK	€0
Underweight	BMI ≥19	Primary care visit	0 SEK	€0
Atrial fibrillation or another arrhythmia	Self-reported physician diagnosis	ECG	486 SEK	€42
Low physical activity	Sedentary lifestyle response to the physical activity questions in the SCAPIS	Primary care visit	0 SEK	€0
Heart failure	NT-pro-BNP value > upper limit of normal (ULN)	Ultrasound, NT-pro-BNP	2648 SEK	€231
Pulmonary fibrosis (PF) or Fibrotic ILA	PF = presence of honeycombing on lung CT; Fibrotic ILA = presence of honeycombing and/or reticular pattern with bronchiectasis on lung CT.	CT, lung volumes, and diffusion capacity	5501 SEK	€481
Depression	Question about depression in the SCAPIS; "yes" on all questions about depression.	Primary care visit	0 SEK	€0
Stress	Question about stress in the SCAPIS equal to constant stress for the past five years	Primary care visit	0 SEK	€0
Pulmonary emphysema	The presence of emphysema defined as at least mild emphysema (grade 1) in any location from the SCAPIS dataset	CT	3292 SEK	€288
Ischaemic heart disease	Self-reported PCI or bypass surgery and previous heart attack.	ECG and CT angiography or stress echocardiography	3892 SEK	€340
Reduced diffusion lung capacity	DLCO <LLN	Lung volumes and diffusion capacity	2209 SEK	€193
Valvular heart disease	Self-reported physician diagnosis or Aortic valve calcification found on CT.	CT	3292 SEK	€288

Abbreviations: BMI = body mass index; CT = computed tomography, DLCO = diffusing capacity for carbon monoxide, ECG = electrocardiogram, FEV1 = forced expiratory volume in 1 second, FVC = forced vital capacity, ILA = interstitial lung abnormality, LLN = lower limit of normal, NT-pro-BNP = N-terminal pro b-type natriuretic peptide, PCI = percutaneous coronary intervention, SCAPIS = multicentre Swedish CArdioPulmonary bioImage Study, WHO = World Health Organization.

Analysis

To develop diagnostic pathways for relevant patient categories, the participants were categorised into four patient categories: women with <20 pack-years of smoking (1), women with ≥20 pack-years of smoking (2), men with <20 pack-years of smoking (3), and men with ≥20 pack-years of smoking (4).[87]

RL is a subtype of AI that can be used to solve tasks, for example, playing chess or driving a car. In RL, a specific model that is learning to solve a task is called an *agent*. RL agents are trained in environments that are often simulated environments such as a computer chess game or a car simulator. These environments give the RL agent information about the world in which it trains, such as where chess pieces are located or the sight of the car. The environment also presents what happens after each step, such as when a chess piece is knocked out, it is removed from the board. The environment also tells the agent if the game is over (after a win, loss or draw in chess) and the model then starts a new game of chess.[88] RL agents learn by receiving a reward score, which is an indicator of how well they are performing the task and is often given after an action, e.g., knocking out one of the opponent's chess pieces or parking a car successfully. When creating the environment, a human decides the value of the reward scores for different actions to motivate the agent to engage in certain behaviours. For example, by providing a higher reward score for knocking out an opponent's queen than for knocking out a pawn in a game of chess, the agent learns that it is more valuable to attack the queen than a pawn. The reward score can also be given as a summarised score after a game (or "episode") is over. The reward can also be negative. For example, if a model is simulated to learn to drive a car, crashing into another car would yield a negative reward score, which results in the model learning to avoid crashing (this is also one of the reasons we use simulations and not real life for agent learning). [88 60 89]

In the present study, a simulated environment for the clinical evaluation of breathlessness was created to train the agent (Figure 1), and by observing how the agent acted after it was trained, a diagnostic pathway was created. The reward function was defined as *the summarised costs of the tests that were not used after all underlying conditions had been identified for an individual participant.* The specific RL algorithm used in the present study is called Advantage Actor Critic (A2C). Briefly, A2C uses two deep neural networks to learn tasks. One neural network is the actor, which estimates the reward for the possible action before deciding the next action. The other neural network acts as a critic to the other neural net. This helps the agent avoid overestimating the presumed award for future actions and ultimately leads to better learning by the agent. A2C has been used to solve multiple advanced tasks and is considered a state-of-the-art RL algorithm to date.[90]

The dataset was randomly divided into training (80%), test (10%) and validation (10%) sets. The training set was used for the simulation and was iterated 300,000 times so that the agent would converge and learn an optimal diagnostic pathway for each patient category. The hyperparameters of the algorithm were adjusted to achieve better performance, and the algorithm was tested on the test set. When the best hyperparameters were identified, the model was validated on the validation set to evaluate the generalisability and potential overfitting of the agent. The following steps were included in the simulation and can be seen in Figure 1:

1. A participant is randomly selected from the dataset.
2. The environment presents the participant's patient category and which tests have been done thus far to the agent (none if this is the first time).
3. The agent chooses one test to perform.
4. The environment presents whether all underlying conditions are identified or whether there are remaining conditions to be identified.
5. If conditions remain to be identified, the agent returns to step 3. If all underlying conditions are identified, the agent is given the reward score and starts at step 1. As an exception, if all tests have been tried and no underlying condition(s) have been identified, the agent is given the reward score and starts at step 1.

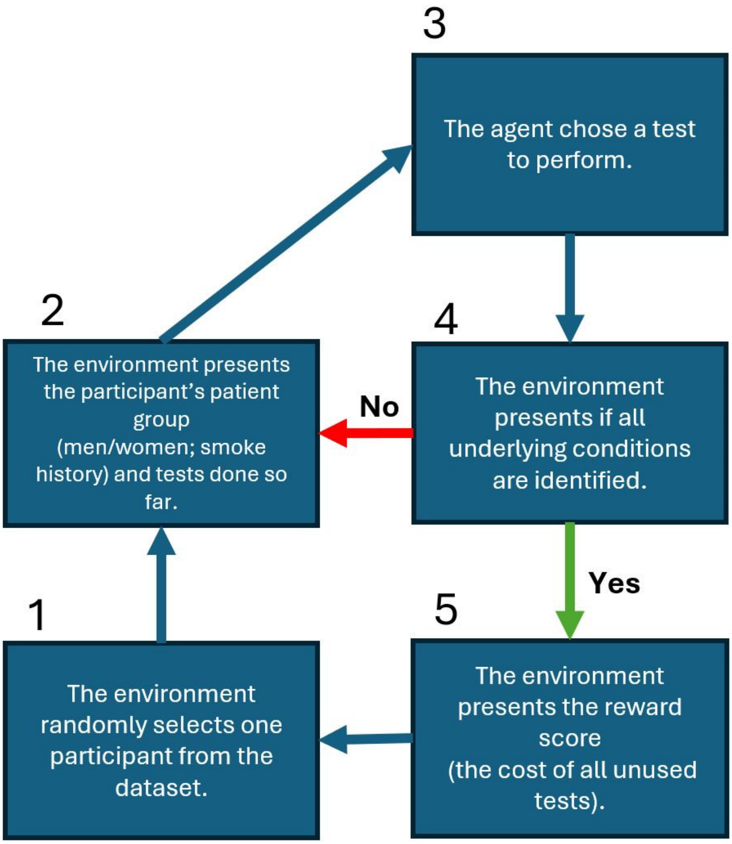


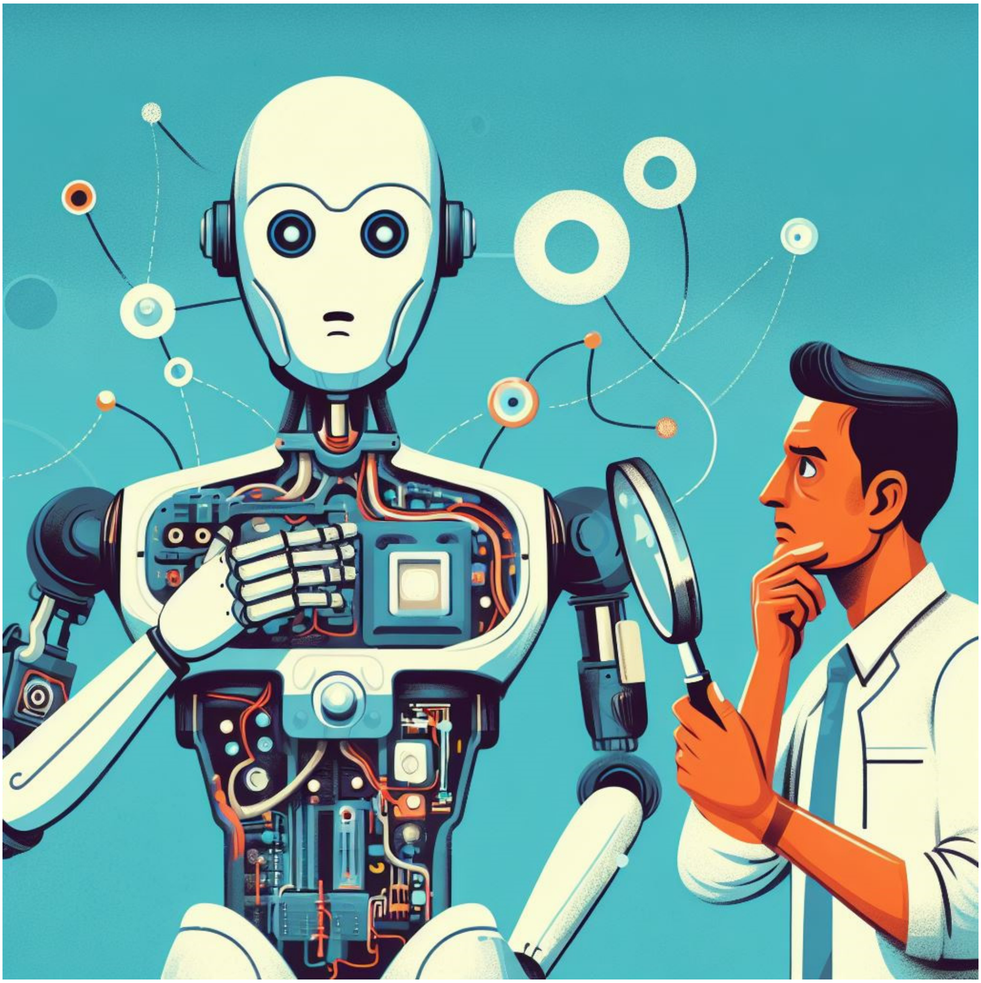
Figure 1. Simulated environment for the clinical evaluation of breathlessness.

The agent performed similarly (mean cost per participant to identify all underlying conditions) on the training (€391) and validation sets (€419), which means that the agent could generalise to participants for whom it has not been trained, and this did not indicate overfitting.

To create a diagnostic pathway for each patient category, the agent was allowed to choose tests for one patient from each patient category, and the different strategies used by the agent were observed. None of the patients had any identifiable underlying conditions, so the agent would only stop after it tried all the available tests. Via this approach, the agent chose tests in order for each patient category until it had tried all eight available tests. The observed strategies were visualised as diagnostic pathways for each patient category, with the percentage of participants with all underlying conditions identified after each step.

Ethical considerations

The data collection and analyses in wave 1 of the VASCOL study were approved by the ethics committee at Lund University (ref: 2008/676), and the data collection and analyses of wave 2 of the VASCOL study were approved by the Swedish Ethical Review Authority (ref: 2019-00134). The data collection of the SCAPIS was approved by the ethics committee at Umeå University (ref: 2010–228-31 M), and the analyses of the SCAPIS data included in this dissertation were approved by the Swedish Ethical Review Authority (ref: 2021-00288). All the participants in the VASCOL study and SCAPIS provided written informed consent.



Can we learn anything from machines? Image created by the author in collaboration with Bing Designer.

Results

An overview of the participant characteristics in each study is presented in Table 6, and the analytic findings of the studies are presented separately in each section. The number of included participants in each study varied due to the different inclusion criteria. The characteristics of waves 1 and 2 of the VASCOL study are presented in more detail in the Results section of Paper I.

Table 6 – Participant characteristics in Studies II–V.

Study	II	III	IV	V
Source	VASCOL study, wave 2	VASCOL study, wave 2	SCAPIS	SCAPIS
N	684	907	28 730	1209
Age, mean (SD)		73.2 (0.67	57.5 (4.3)	58.2 (4.4)
Male sex, n (%)	684 (100%)	907 (100%)	13 929 (48%)	379 (31%)
BMI	27.2 (3.9)	27.1 (3.8)	26.9 (4.4)	31.00 (5.96)
Smoking status				
Daily	34 (5%)	41 (6%)	2119 (7)	-
Sometimes	7 (1%)	11 (1%)	1389 (5)	-
Former smoker	404 (60%)	530 (59%)	10 382 (36)	-
Never smoker	231 (34%)	310 (35%)	14 473 (50)	-
mMRC score, n (%)				
0	453 (68%)	606 (67%)	-	0(0%)
1	98 (15%)	120 (13%)	-	0(0%)
≥2	117 (18%)	154 (18%)	1209 (4.2%)	1209 (100%)
Self-reported health conditions				
COPD, n (%)	26 (4%)	32 (4%)	321 (1%)*	215 (18%)**
Asthma, n (%)	35 (5%)	47 (5%)	2288 (8%)	
Atrial fibrillation	105 (16%)	135 (16%)	-	49 (4%)
Heart failure, n (%)	27 (4%)	35 (4%)	141 (1%)	223 (18%)***

The values are reported as the mean (standard deviation) or frequency (%). *Also includes self-reported chronic bronchitis or emphysema. ** Chronic airflow limitation. *** Not self-reported, definition was based on a definition used in a previous study.[85]

Data collection and requirements for the VASCOL study (Study I)

The flowchart of the data collection process in the VASCOL study with the number of recruited participants at each step and the reasons for drop-out are presented in Figure 2. A total of 1900 men were invited to participate in the screening, 1762 of whom agreed to participate; among these participants, 1302 participated in wave 1 of the VASCOL study.

In 2019, a total of 106 of the participants in wave 1 had died, and 3 had an unknown address. The men who were still alive and had a known address were invited to participate in the postal survey in wave 2; initially, 829 men completed the survey, and an additional 78 men completed the survey after a reminder was sent approximately two weeks after the first mailing. A total of 907 men participated in wave 2 of the VASCOL study.

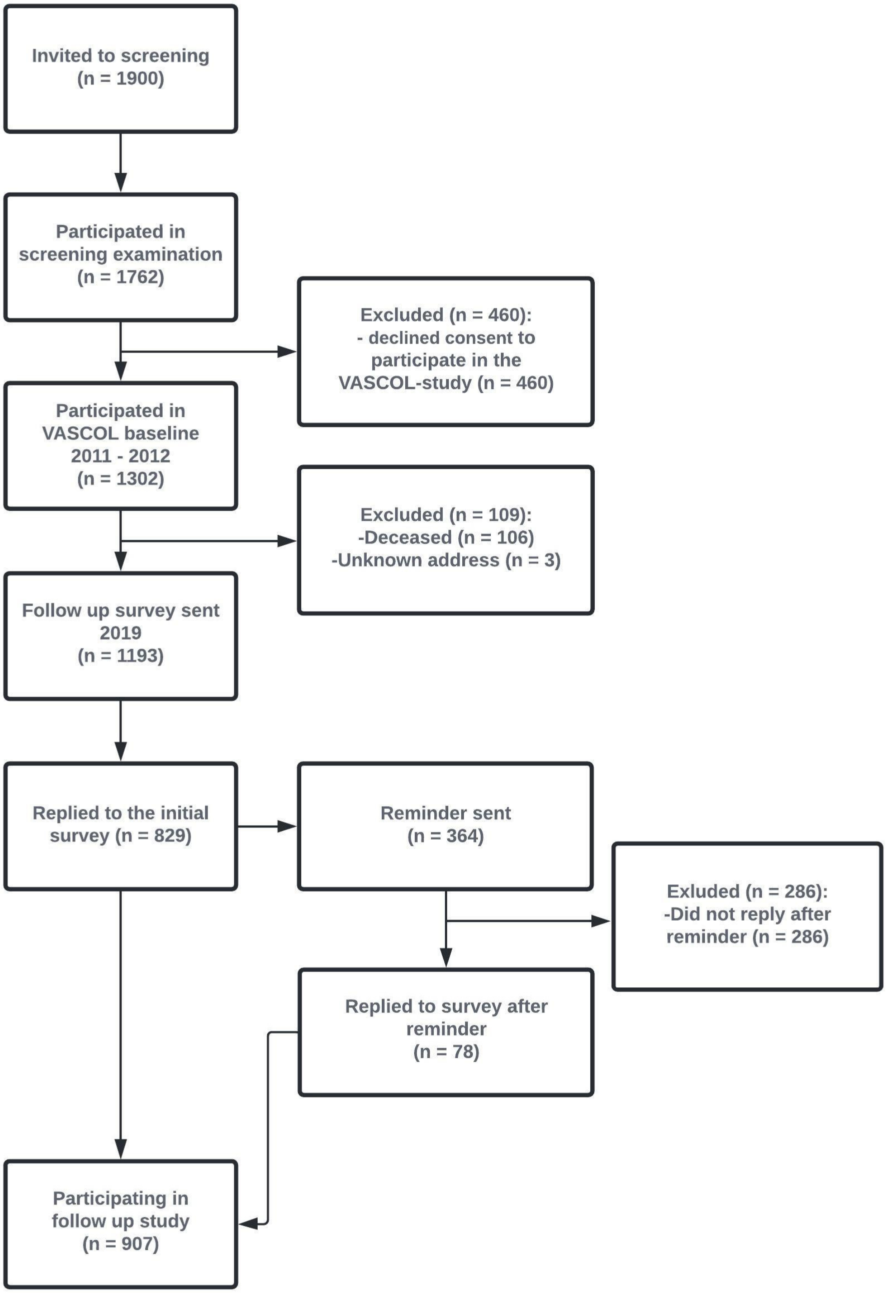


Figure 2. Flowchart of the recruitment of participants in waves 1 (baseline) and 2 (follow-up) of the VASCOL.

The characteristics of the men in wave 1 are shown in Table 7, with a comparison between the men who later participated in wave 2 and those who did not. Overall, men who did not participate in wave 2 were more likely to be obese, current smokers, have a greater number of pack-years, and have diabetes or COPD.

Table 7. Characteristics of the men who participated in wave 1 of the VASCOL, with a comparison between men who later participated in wave 2 and those who did not.

Factor	Participated in wave 2	Did not participate in wave 2
Body mass index (kg/m2)	28.1 (4.0)	28.5 (4.4)
Normal weight, BMI 18.5–24.9	184 (21%)	76 (20%)
Overweight, BMI 25–29.9	470 (53%)	191 (50%)
Obesity, BMI >=30	234 (26%)	118 (31%)
FEV1, % predicted	88.5 (14.7)	84.3 (15.8)
FVC, % predicted	86.3 (13.0)	83.4 (13.1)
FEV1/FVC < 0.7	157 (17%)	85 (22%)
Smoking status		
Current	91 (10%)	77 (19%)
Former	514 (57%)	193 (49%)
Never	302 (33%)	125 (32%)
Pack-years of smoking	13.7 (17.4)	18.4 (22.6)
University/college or professional school education	381 (44%)	139 (38%)
Self-reported conditions		
Hypertension	507 (56%)	218 (55%)
Diabetes	79 (9%)	53 (13%)
Asthma	47 (5%)	24 (6%)
COPD	13 (1%)	16 (4%)

The values are reported as the mean (standard deviation) or frequency (%). These are the characteristic of the men who participated in wave 1 from 2010–2011. Abbreviations: BMI = body mass index, COPD = chronic obstructive pulmonary disease, FEV1 = forced expiratory volume in 1 second, FVC = forced vital capacity.

Psychometric properties of the D12 and MDP in a population-based setting (Study II)

A total of 684 participants had complete D12 and MDP data from wave 2 of the VASCOL and were included in the study. Among the participants, the mean BMI was 27.2, and 445 (65%) were classified as ever smokers. A total of 26 (4%) participants had COPD, 35 (5%) had asthma, and 27 (4%) had heart failure (Table 6).

Factor structure

Figure 2 presents the results from the CFA. Most of the items of the D12 and MDP presented good factor loadings, but the D12 item “My breath does not go in all the way” and the MDP item “Breathing a lot” had lower factor loadings. The fits of the CFA models were not excellent for the D12 (RMSEA=0.137, CFI=0.914) and MDP (RMSEA=0.134, CFI=0.928).

Instrument	Subdomain	Item	Estimate	Error Variance
D12	Physical	My breath does not go in all the way	0.693	0.519
		My breathing requires more work	0.783	0.386
		I feel short of breath	0.745	0.445
		I have difficulty catching my breath	0.836	0.301
		I cannot get enough air	0.864	0.254
		My breathing is uncomfortable	0.846	0.285
		My breathing is exhausting	0.858	0.264
D12	Affective	My breathing makes me feel depressed	0.871	0.242
		My breathing makes me feel miserable	0.811	0.342
		My breathing is distressing	0.898	0.194
		My breathing makes me agitated	0.842	0.291
		My breathing is irritating	0.916	0.161
MDP	Immediate perception	A1 unpleasantness	0.815	0.336
		Muscle work	0.846	0.284
		Air hunger	0.887	0.213
		Chest tightness	0.887	0.213
		Mental effort	0.918	0.157
		Breathing a lot	0.724	0.475
MDP	Emotional response	Depression	0.9	0.189
		Anxious	0.925	0.144
		Frustrated	0.843	0.289
		Angry	0.764	0.416
		Afraid	0.905	0.182

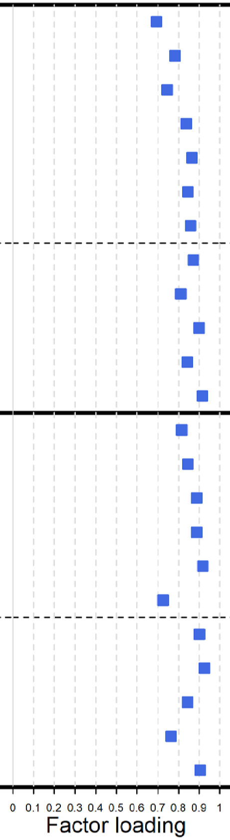


Figure 2. Confirmatory factor analysis of the D12 and MDP. The estimate corresponds to the factor loading and represents how much the item belong to its factor (subdomain). An increased estimate means better factor belonging. Error variance represents how much the item does not belong to its given subdomain.

Internal consistency

The internal consistency of the D12 and MDP are presented in Table 8. Both the D12 and MDP presented excellent internal consistency with respect to the total scores, and the domain scores had Cronbach's alpha estimates above 0.90.

Table 8 - Internal consistency of the D12 and MDP

Variable	Cronbach's alpha estimate
D12 total score	0.956
D12 physical domain score	0.924
D12 affective domain score	0.936
MDP total score	0.943
MDP immediate perception subdomain score	0.932
MDP emotional response subdomain score	0.939

An increased Cronbach's alpha estimate represents better internal consistency, and a score above 0.8 is considered good, whereas a score above 0.9 is considered excellent.

Validity

The validity of the D12 and MDP is presented in Figure 3. Overall, the total scores and subdomain scores of the D12 and MDP were correlated with each other and the ESAS-R breathlessness score, indicating the external validity of the scales. Discriminant validity was confirmed by the D12 physical domain score, and the MDP-IP domain score were strongly correlated with the SF-12 physical QoL score compared with the D12 affective and MDP emotional subdomain scores. The D12 affective and MDP emotional domain scores were more strongly correlated with the SF-12 mental QoL score than were the D12 physical domain score and the MDP-IP subdomain score. Compared with the D12 affective domain and MDP emotional domain scores, the D12 affective domain and MDP emotional domain scores were more strongly correlated with the HADS anxiety and depression scores, indicating discriminant validity.

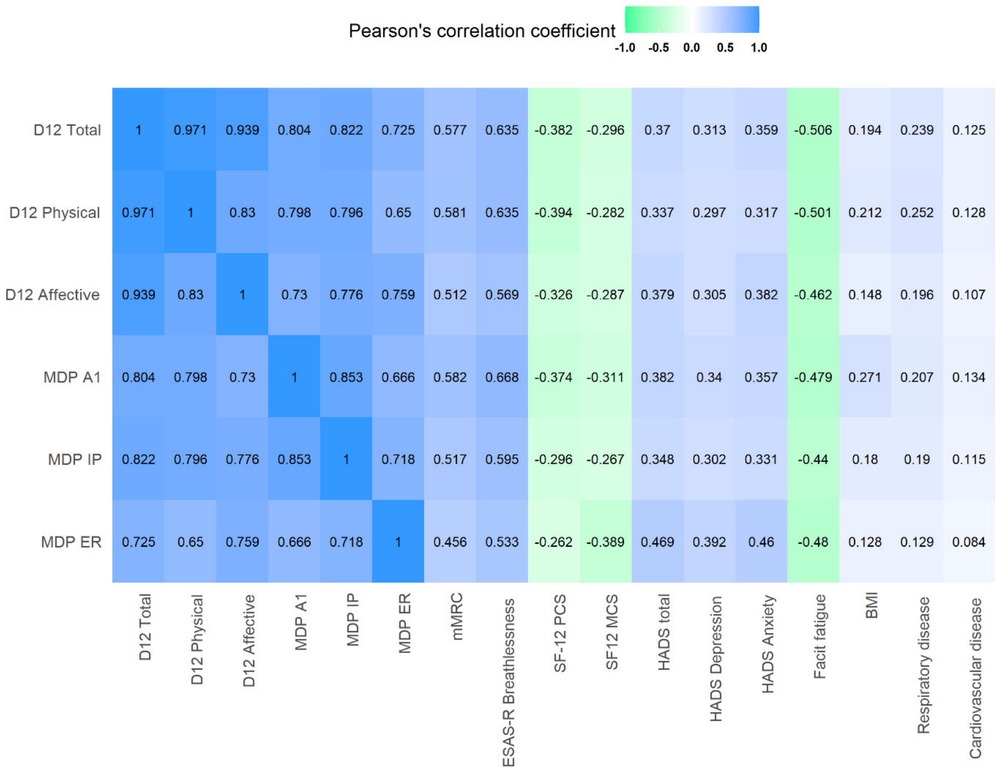


Figure 3. Validity of the D12 and MDP domains. The estimates are Pearson´s correlation coefficients and are also represented by colour intensity. Abbreviations: Body mass index = BMI, D12 = Dyspnoea-12, MDP = Multidimensional Dyspnoea Profile, IP = Immediate Perception, ER = Emotional Response, mMRC = Modified Medical Research Council dyspnea scale, ESAS-r = Edmonton Symptom Assessment System-Revised, HADS = Hospital Anxiety and Depression Scale; SF-12 PCS = short-form 12-item (version 2) physical health composite score; SF-12 MCS = short-form 12-item (version 2) mental health composite score.

Prevalence of different dimensions of breathlessness among older men in the general population (Study III)

The mean (SD) BMI of the participants was 27.1 (3.8), 530 (59%) were former smokers, and 41 (6%) were current smokers. A total of 47 (5%) participants had asthma, and 32 (4%) had COPD (Table 6).

Prevalence and intensity of different dimensions of breathlessness

The prevalence of breathlessness was 144 cases (17%) according to the D12 total score and 33% according to the MDP A1 unpleasantness score. The prevalence of breathlessness was 162 cases (19%) according to the D12 physical domain score and 120 (17%) according to the MDP IP subdomain score. The prevalence was 89 cases (10%) for the D12 affective domain score and 134 cases (17%) for the MDP ER subdomain score. The intensities of the MDP items among participants with an MCID above or equal to a score of one on the MDP A1 unpleasantness domain are presented in Figure 4.

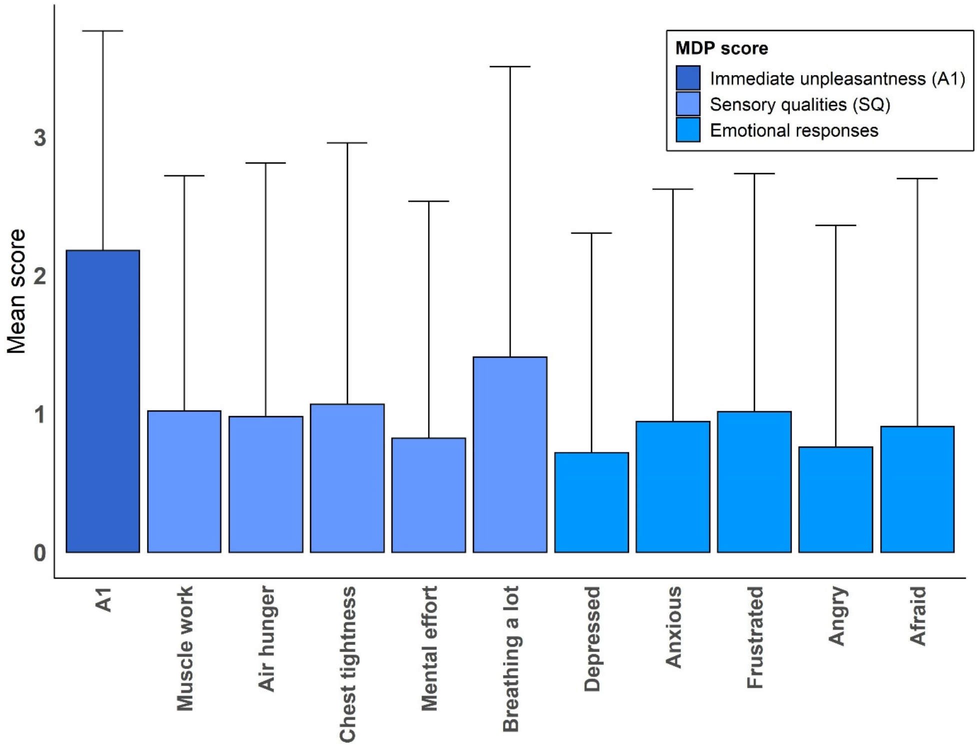


Figure 4. Intensity of symptoms in different dimensions of breathlessness. The bars represent the mean score of the MDP items (range 0–10 points), and the whiskers represent the standard deviation. Immediate perception comprises the sensory qualities together with the A1 item.

Dimension associations with the mMRC score and changes in and duration of breathlessness

An increased mMRC score was associated with increased D12 total, physical and affective domain scores (Figure 5), as well as increased MDP A1 unpleasantness, IP and ER subdomain scores (Figure 6). The participants who reported that their breathlessness worsened had higher D12 total, physical and affective scores than did the participants who reported that their breathlessness was better or no different. This association was not as evident when breathlessness was estimated with the MDP scores.

Overall, participants who experienced breathlessness for less than one year had higher scores in all breathlessness domains than did participants who experienced breathlessness for one to five years. At the same time, participants who experienced breathlessness for more than five years had worse domain scores than did participants who experienced breathlessness for one to five years (Figure 5 and Figure 6).

	D-12 total ß coef	D-12 physical ß coef	D-12 affective ß coef
mMRC			
0 (ref)	–	–	–
1	1.29 (0.63 to 1.95)	0.94 (0.53 to 1.34)	0.33 (0.04 to 0.63)
2	3.51 (2.68 to 4.35)	2.4 (1.89 to 2.91)	1.15 (0.79 to 1.52)
3	6.43 (5.34 to 7.53)	3.97 (3.31 to 4.63)	2.52 (2.04 to 3.01)
4	9.14 (8.05 to 10.24)	5.41 (4.75 to 6.08)	3.63 (3.15 to 4.11)
Change			
Better (ref)	–	–	–
No difference	0.82 (-0.28 to 1.92)	0.69 (0.03 to 1.35)	0.19 (-0.28 to 0.67)
Worse	3.16 (1.88 to 4.43)	1.87 (1.1 to 2.64)	1.22 (0.67 to 1.77)
Duration			
<1 year	5.27 (3.82 to 6.72)	3.49 (2.64 to 4.33)	1.85 (1.19 to 2.51)
1 - 5 years (ref)	–	–	–
>5 years	0.78 (-1.54 to 3.1)	0.47 (-0.89 to 1.83)	0.22 (-0.83 to 1.26)

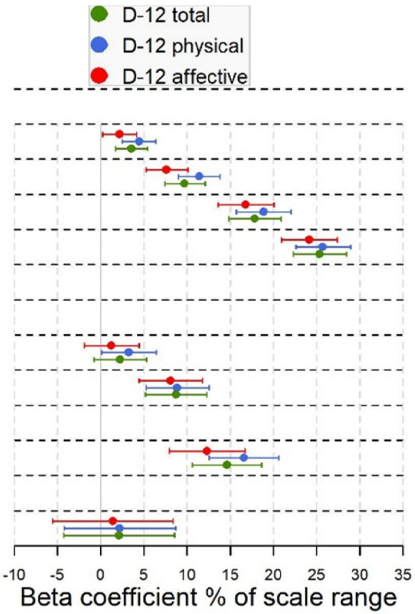


Figure 5. Associations of the D12 domain scores with the mMRC score and changes in and duration of breathlessness. The table presents beta coefficients, which are also presented as dots. The beta coefficients were scaled to be able to compare estimates between scales with different ranges. The whiskers present 95% confidence intervals.

	MDP A1 ß coef	MDP IP ß coef	MDP ER ß coef
mMRC			
0 (ref)	–	–	–
1	0.74 (0.52 to 0.97)	2.26 (1.1 to 3.42)	0.64 (-0.28 to 1.55)
2	1.06 (0.77 to 1.34)	3.88 (2.43 to 5.34)	2.4 (1.24 to 3.56)
3	2.22 (1.86 to 2.58)	10.31 (8.23 to 12.39)	4.76 (3.29 to 6.23)
4	2.76 (2.39 to 3.13)	12.73 (10.86 to 14.59)	8.65 (7.14 to 10.16)
Change			
Better (ref)	–	–	–
No difference	0.52 (0.18 to 0.87)	1.57 (-0.23 to 3.36)	1.71 (0.34 to 3.08)
Worse	1.02 (0.62 to 1.42)	4.12 (2.03 to 6.21)	1.92 (0.31 to 3.52)
Duration			
<1 year	2.11 (1.67 to 2.56)	7.62 (5.11 to 10.13)	4.83 (3.11 to 6.54)
1 - 5 years (ref)	–	–	–
>5 years	0.43 (-0.28 to 1.14)	3.07 (-0.95 to 7.09)	-0.28 (-3.11 to 2.56)

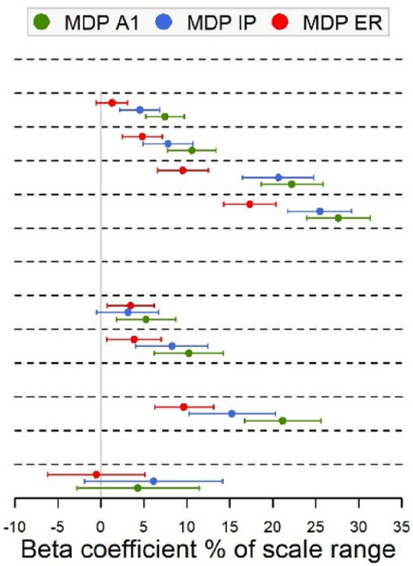


Figure 6. Associations of the MDP domain scores with the mMRC score and changes in and duration of breathlessness. The table presents beta coefficients, which are also presented as dots. The beta coefficients were scaled to be able to compare estimates between scales with different ranges. The whiskers present 95% confidence intervals.

Factors associated with breathlessness in the general population (Study IV)

A total of 28 730 (52% women) were included in the study. The participants had a mean (SD) age of 57.5 years (4.3) and a mean (SD) BMI of 26.9 (4.4). The prevalence of asthma was 8%, that of COPD, chronic bronchitis or emphysema was 1%, and that of heart failure was 1% (Table 6). Breathlessness was present in 4% of the participants.

Factors important to breathlessness

The strengths of the associations between associated factors and breathlessness are presented in Figure 7. The factors most strongly associated with breathlessness were (SHAP absolute mean) BMI (SHAP absolute mean: 0.39), followed by the forced expiratory volume in 1 second (FEV1) (SHAP absolute mean: 0.32), vigorous-intensity physical activity measured by accelerometery (SHAP absolute mean: 0.27), sleep apnoea (SHAP absolute mean: 0.22), and the diffusion capacity for carbon monoxide (SHAP absolute mean: 0.21).

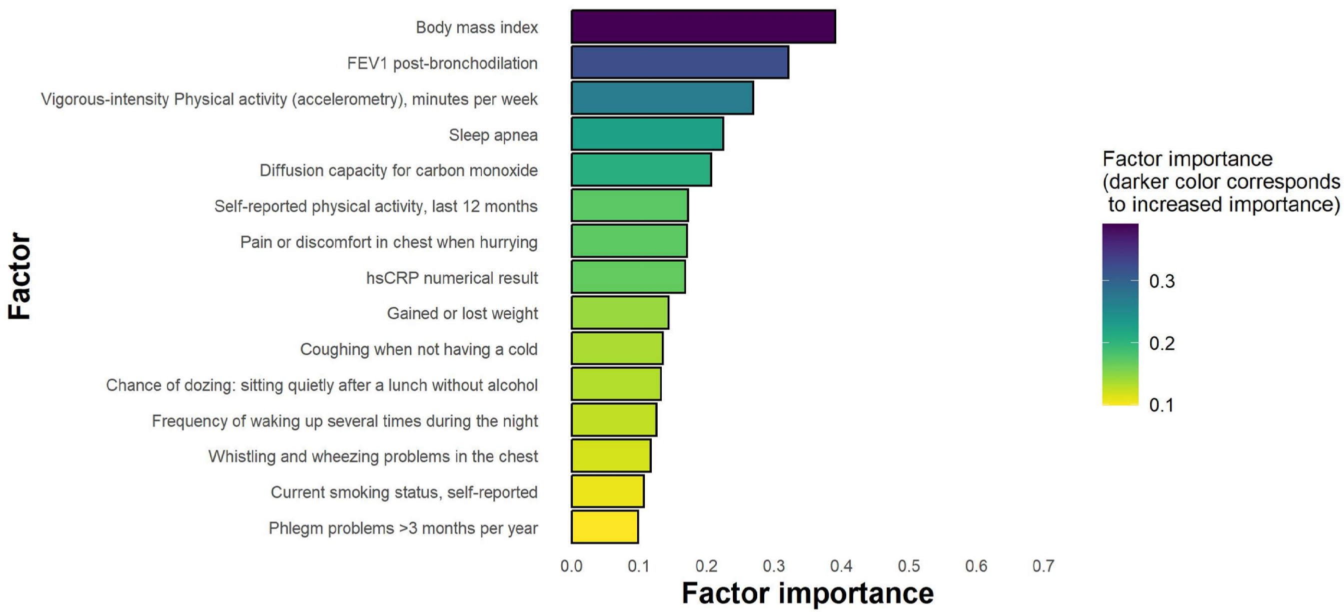


Figure 7. Strengths of the associations between associated factors and breathlessness. A increased factor score indicates greater importance for breathlessness, measued via the mean absolute SHapley Additive exPlanations (SHAP) value. The absolute mean SHAP value respresents the average contriution to the probability of being classified as experiencing breathlessness. The plot only shows factors with SHAP scores >0 points.

Shapes of the associations between associated factors and breathlessness

The changes in the probabilities of the most important factors contributing to breathlessness are presented in Figure 8. An increased BMI was associated with an increased probability of experiencing breathlessness, and a BMI above 30 was strongly associated with breathlessness. A lower FEV1 was associated with breathlessness as well as a lower diffusion capacity for carbon monoxide. A lower degree of physical activity, both self-reported and measured via accelerometery, was associated with breathlessness. Having sleep apnoea, pain or discomfort in the chest when hurrying or walking, having recently gained or lost weight, and having a cough when not having a cold was associated with experiencing breathlessness.

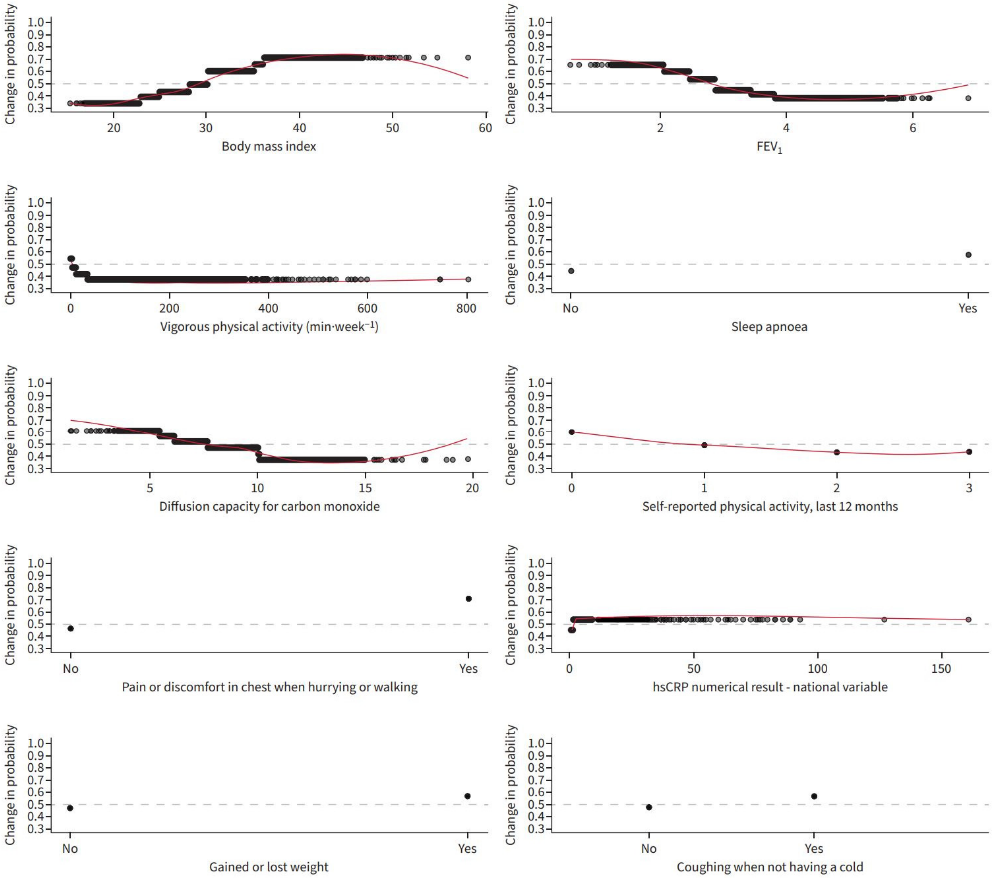


Figure 8. Shapes of the associations between associated factors and breathlessness. The dots represent the participants' factor values (X-axis). The Y-axis represents the probability of experiencing breathlessness, values above the dashed line indicate a greater likelihood of experiencing breathlessness, and values below the dashed lines indicate a greater likelihood of not experiencing breathlessness. The red lines are locally estimated scatterplot smoothing lines, which are the average values of the participants' factor values. Numbers corresponding to self-reported physical activity in the last 12 months: 0 = never, 1 = from time to time, 2 = one or two times a week, 3 = two or three times a week or more. Abbreviations: FEV_1: forced expiratory volume in 1 second; HsCRP: high-sensitivity C-reactive protein.

A diagnostic pathway for evaluating breathlessness (Study V)

A total of 1209 (69% women) participants had clinically significant breathlessness and were included in the study. Of these participants, 238 (20%) were men with <20 pack-years of smoking, 141 (12%) were men with ≥20 pack-years of smoking, 608 (50%) were women with <20 pack-years of smoking, and 222 (18%) were women with ≥20 pack-years of smoking.

The mean (SD) age was 58.2 years (4.4), and the mean (SD) BMI was 31.00 (5.96). The most common underlying conditions were obesity (54%), low levels of physical activity (35%), stress (25%), asthma (20%), heart failure (18%), CAL (18%), reduced diffusion capacity (17%), and cardiac valvular heart disease (12%).

The diagnostic pathways for all patient groups are shown in Figure 9. For all patient groups, a primary care visit was suggested as step one, followed by spirometry. Lung CT was suggested as step three, except for women with <20 pack-years of smoking, for whom it was suggested as step six. Overall, tests for lung diseases were recommended before heart disease and blood tests.

The percentage of participants for whom all underlying conditions were identified early in the diagnostic pathway increased considerably. After step one, 37% of the participants had all underlying factors identified, 51% after step two, 61% after step three, 75% after step four, 83% after step five, 90% after step six, 91% after step seven, and 92% after step eight. Eight percent of the participants had unknown underlying factor(s) contributing to their breathlessness after all ten steps. Compared with patients with <20 pack-years of smoking, for those with ≥20 pack-years of smoking, a lower percentage of all underlying conditions were identified early in the diagnostic pathway (Figure 9).

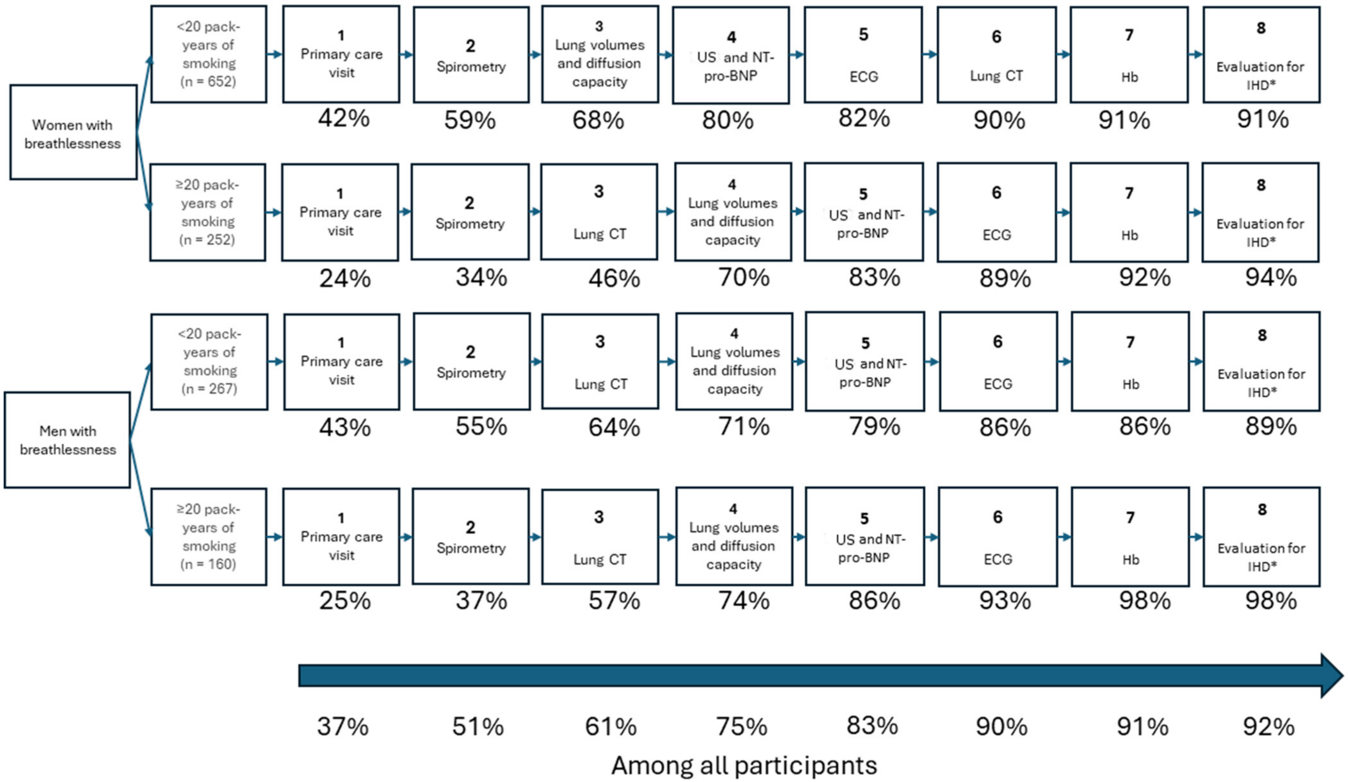
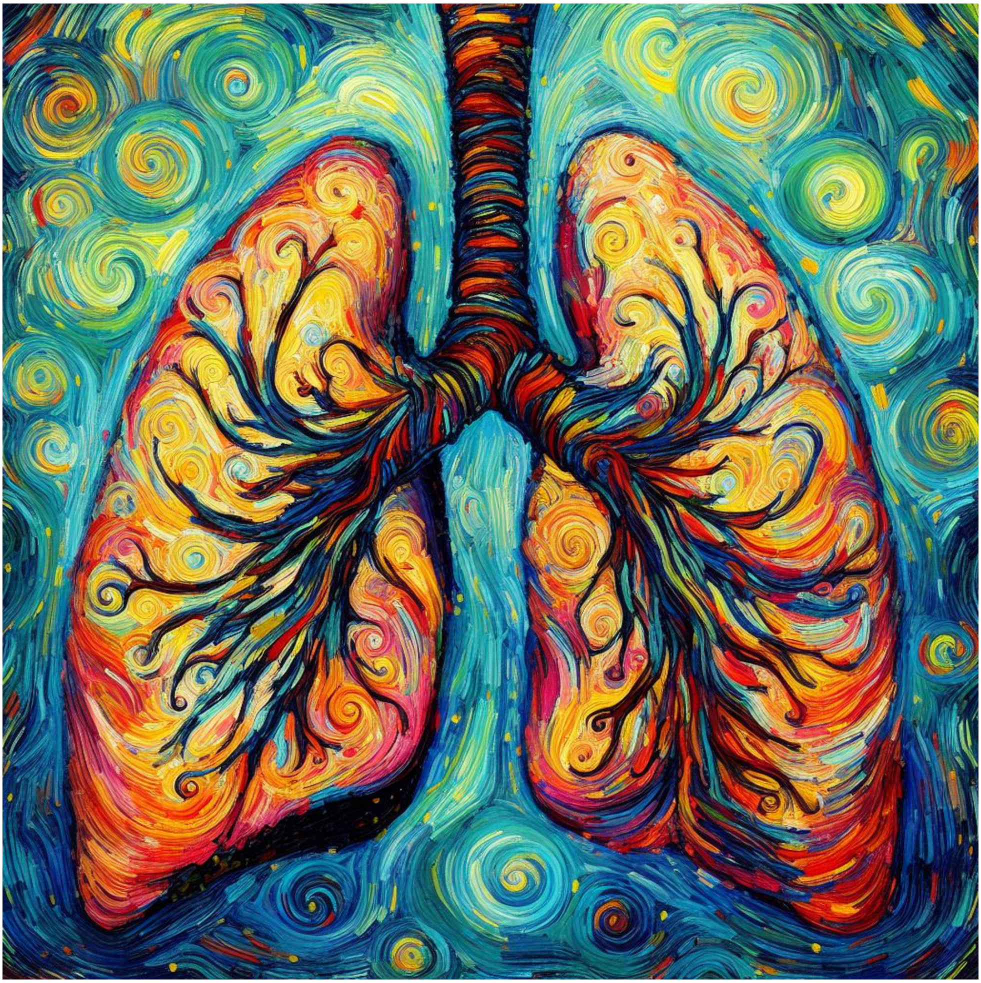


Figure 9. Diagnostic pathway to identify underlying conditions contributing to breathlessness. This figure presents the final diagnostic pathway after the simulations (the percentage corresponds to the cumulative percentage of participants with all underlying conditions identified). Underlying conditions are presented in Table 1. Primary care visits included and history taking, a basic physical examination including an assessment of body mass index (BMI), physical activity levels, depression status, and stress status. Breathlessness was defined as a modified Medical Research Council breathlessness scale (mMRC) grade ≥2 (breathless when walking on level ground or worse). *Includes electrocardiogram and computed tomography or stress echocardiography. Abbreviations: ECG = electrocardiogram, Hb = haemoglobin. IHD = ischaemic heart disease, US = ultrasound.



Colourful perspectives. Image created by the author in collaboration with Bing Designer. Inspired by works by Vincent Van Gogh.

Discussion

Main findings

Psychometric properties of the MDP and D12.

In Study II, we found that the D12 and MDP are valid for use in epidemiological studies. The D12 and MDP domains showed good validity and correlated well with each other, with other breathlessness measurements and with measurements of anxiety, depression, fatigue, and physical and mental QoL. The internal consistency was high, suggesting that the instruments are reliable. The Cronbach's alpha estimates in our study were similar to or greater than those in a previous review of the psychometric properties of the D12 and MDP.[40]

The D12 and MDP each had one item that had a lower factor loading than the other items did, meaning that the items did not fully belong to the proposed domain. Our study included participants with multiple underlying conditions, in contrast to previous studies, which included only participants with cardiorespiratory conditions.[38 39] This suggests that dimensions are experienced differently among participants with noncardiorespiratory disease as their main underlying condition, and these experiences might be less relevant for participants with, for example, obesity or anxiety as their main underlying condition. It is also reasonable to believe that the variance in the dimensions is greater in the VASCOL population with various underlying conditions than in the clinical cohort selected on the basis of one underlying condition. Notably, the factor structure has been confirmed in only a few studies,[40] which suggests that the factor structure should be evaluated further.

In this study, a novel finding was the discriminative validity of the D12 and MDP for BMI and physical and mental QoL, which are often measured in epidemiological and public health studies. The D12 and MDP correlated moderately with each other but not perfectly, which further supports that they are not interchangeable,[40] and the choice of the instruments in future studies can be based on the study context.

Prevalence of different dimensions of breathlessness.

We found that breathlessness is highly prevalent among the older male population, as more than one-third of the study population reported that their breathlessness was unpleasant and approximately one-fifth experienced symptoms of breathlessness in a physical domain and/or an emotional domain. Overall, the intensities of the symptoms were similar, with no dimension dominating the overall experience of breathlessness.

As this is the first epidemiological study of breathlessness from a multidimensional perspective, comparisons with previous studies are limited. Our study and a previous smaller clinical study[91] included participants with various underlying conditions, and both studies reported similar intensities of symptoms across the MDP dimensions. In a smaller clinical study of COPD patients, the MDP dimensions showed different characteristics in terms of the intensities of the scores.[92] This suggests that different dimensions can be related to different pathological processes, as suggested previously,[1] but this needs to be further investigated.

The domain scores of the breathlessness instruments increased stepwise with increasing mMRC grade, indicating that the severity of exertional breathlessness was strongly associated with adverse experiences of breathlessness. In particular, people who had trouble performing everyday tasks such as dressing or leaving home (mMRC grade of 3-4) experienced high-intensity, unpleasant, physical, emotional, and affective symptoms of breathlessness. With these findings in mind, health professionals should assess overall well-being among patients with breathlessness rather than only focusing on the limitations induced by symptoms[19] or only treating the underlying condition(s). The mMRC scale is often used as a screening tool in health care settings,[93] but the multidimensional instrument can be used for further evaluation, especially among patients with higher mMRC grades.

Participants who experienced breathlessness for a duration of one to five years experienced symptoms that were milder than those experienced by participants who experienced breathlessness for longer or shorter durations. This suggests that people can learn to cope with their symptoms over time, but as underlying conditions progress, the symptoms of breathlessness become more challenging to control. Similarly, this has previously been suggested for the association between the duration of breathlessness and QoL.[14 16] Based on these findings, healthcare professionals should evaluate the overall well-being of patients experiencing breathlessness for the first time and continue monitoring them over time.

Factors associated with breathlessness identified through machine learning.

Obesity, reduced lung function, a sedentary lifestyle, and sleep apnoea were the factors most strongly associated with breathlessness in the middle-aged general population. This finding suggests that breathlessness is associated with multiple factors, not only cardiorespiratory conditions. These factors can be targets for future public health interventions.

Study IV was an explorative study that included physiological, psychological, social, and environmental factors. This is different from many previous studies in which cardiorespiratory conditions were predefined and then tested regarding their associations with breathlessness.[46] The shapes of the associations were also explored in this study, which provides new insights into the factors associated with breathlessness.

The association between BMI and breathlessness was recently studied[94 95], and Study IV revealed a strong association between obesity and breathlessness, providing new insight into the shape of the association between an increased BMI and breathlessness. This study supports the obesity threshold (BMI >30) set by the WHO, as a BMI above 30 is strongly associated with breathlessness. The obesity threshold has already been established as a predictor of mortality and diseases such as diabetes and cardiovascular diseases,[96 97] but the present study also suggested that the threshold of obesity is highly relevant for life-limiting breathlessness. The link between BMI and breathlessness is suggested to be due to the extra workload from a heavier body, as suggested previously.[94] An increased BMI and a sedentary lifestyle might have a confounding association with breathlessness, meaning that an increased BMI contributes not to breathlessness but rather to a sedentary lifestyle. However, the present study adjusted for many factors, and an increased BMI and a sedentary lifestyle were independently and strongly associated with breathlessness.

Physical activity, measured by both accelerometery and self-reports, was strongly associated with breathlessness. The study included different accelerometery measurements, but, among all the accelerometery measurements, the amount of vigorous-intensity physical activity by far had the strongest association with breathlessness. This suggests that people with exertional breathlessness avoid exercise because of the life-limiting symptoms, as suggested previously.[98]

Reduced lung function was strongly associated with breathlessness, as measured by a lower FEV_1 and diffusing capacity for carbon monoxide (DLCO). Although the present study provides evidence of the association between a lower FEV_1 and breathlessness, the evidence is sparse overall.[46 99] Overall, lung function tests were more strongly associated with breathlessness than were CT measurements, which is a novel finding. These findings suggest that, compared with CT, lung function assessments could be better for the overall assessment of breathlessness in population studies. This study revealed a category of extensional breathlessness that should be directly related to the capacity and function of the lungs, and we do not know whether CT could reveal more about the different dimensions of breathlessness.

A diagnostic pathway to identify underlying medical conditions associated with breathlessness.

The diagnostic pathway presented in Study V was developed with a novel AI approach and considered the prevalence of underlying conditions in the general population and economic costs. With a mean cost of €419 per participant, the diagnostic pathway could identify all underlying conditions in 92% of the participants. The pathways were similar for the different patient groups overall, but performing CT early in the pathway for women with less than 20 pack-years of smoking was not considered valuable. The last step of the diagnostic pathway included ECG, haemoglobin (Hb) tests, and evaluations for ischaemic heart disease (IHD), but these did not add much to the overall percentage of participants with all underlying conditions identified.

This study supports the importance of including spirometry early in the evaluation of breathlessness, as previously suggested.[50] In the general population, it is more important to test for respiratory diseases earlier than cardiovascular diseases. The developed diagnostic pathway included underlying conditions that have not always been considered relevant conditions for breathlessness in earlier diagnostic pathways, such as obesity, anxiety, and depression.[49 50]

Different arms of the pathway have been developed for clinically relevant patient groups, and this approach can be useful because of the associations of sex and smoke exposure with many underlying conditions.[100 101 102] The middle-aged study population is highly relevant, as chronic conditions are more common in middle-aged people than they are in younger people;[8 103] however, on average, the middle-aged population still has many years to live. This means that evaluating and identifying underlying conditions causing life-limiting symptoms in middle-aged people is highly important. Since availability, costs, and routines differ worldwide, the pathway can be adjusted accordingly to meet the needs of different contexts. From a wider perspective, the overall suggestion is that all patients should undergo a primary care visit, followed by a full examination of the lungs and then an examination of their heart. Complementary tests, including blood sample tests, can be performed as last steps if there are other suspected underlying conditions. As a last step, if underlying condition(s) are still suspected after following the diagnostic pathway, evaluation in a specialised clinic via a cardiopulmonary exercise test (CPET) could be considered,[104] which could reveal underlying conditions that the tests in our study would fail to identify.

Strengths and limitations

The overall strengths of the VASCOL studies (Studies I–III) include that only men of the same age were included, reducing the risk of bias from age and sex. Overall, the cohort is generalizable to men of the same age in Sweden, and the findings of the studies are likely generalizable to men of similar age in other high-income countries. Study II included a large sample size, which is rare for psychometric studies. The study also included a large variety of participants with and without underlying conditions, meaning that the psychometric properties can be relevant for a wide variety of underlying conditions. The reliability of the study is supported by the fact that the resulting estimates are similar to those of a previous meta-analysis of the psychometric properties of the D12 and MDP.[40] Further strengths of Study III include the use of MCIDs, which reflect relevant experiences for the participants and simplifies comparisons between different domains of breathlessness. Additionally, the validated instruments used to assess breathlessness in Study III strengthen the validity of the findings.

The overall strengths of the SCAPIS (Studies IV and V) include the large number of participants who are representative of the Swedish middle-aged population,[69] which supports the generalizability of the study. The strengths of Study IV include the data-driven approach with many factors, which means that a multitude of factors could be evaluated fairly without prespecified hypotheses. The ML approach could identify potential nonlinear and complex associations that could be missed by traditional statistical methods. The model was evaluated on a test set that was based on a separate study site, which supports the generalisability of the associations found in the study. The strengths of Study V include the use of a random population sample, which better reflects people seeking help regarding their breathlessness in primary care than do participants in smaller clinical studies. Study V included data from all included participants and not only from people seeking care or specific clinical groups, which reduces the risk of bias from health care-seeking behaviours. The data-driven approach means that the prevalence of underlying conditions and economic costs were the driving factors in the development of the diagnostic pathways and, to a lesser extent, presumptions and previous knowledge.

Limitations of the VASCOL studies (Studies I - III) include that only older men were included, which means that we cannot generalise the findings to women and younger people. The VASCOL study plans to recruit women and younger people in the future. The limitations of Study II include that there was no testing of the reliability of the instruments over time with test-retest methods. However, previous studies have shown that the instruments are reliable over time.[40] Study II did not include many participants with extremely severe breathlessness, and the validity and reliability of the instruments could be lower in this group. There are also more modern psychometric testing methods, such as the Rasch methods, that can be considered and have previously been used for the D12 and mMRC scale.[34 38] The limitations of Study III include that the change in and duration of breathlessness were recalled and not actually measured over time, which introduce the risk of recall bias.

Overall, the limitations of the SCAPIS studies (Studies IV and V) include the inclusion of a middle-aged population, and the findings cannot be generalised to younger or older people, who are likely to have different frequencies of underlying conditions contributing to breathlessness as well as different lifestyle habits and socioeconomic factors. The limitations of Study IV include that the model in had an acceptable but not excellent AUC, but in terms of ML, the outcome was uncommon in the dataset, which makes it challenging to reach an excellent AUC. XGBoost often performs well,[76] and the use of another algorithm, such as a random forest or support vector machine, would probably not result in better performance; however, it could be a good idea to compare the performance of different algorithms in Study IV. Breathlessness was measured with the mMRC scale in the SCAPIS, and future studies should include measurements of multidimensional breathlessness. The limitations of Study V include that the study assumed that all underlying conditions would contribute equally to breathlessness, which is not the case, as suggested in Study IV and a previous study.[105] However, it is important to identify all conditions, but the study was designed to prioritise economic costs and the prevalence of conditions and not how much an underlying factor contributes to the symptoms or how dangerous the condition is. The study included only underlying conditions that were definable by the variables included in the SCAPIS dataset. Additionally, the costs of the tests likely differ across countries. The methodology should therefore be tested in other studies and contexts. RL is a novel and advanced technique, and simpler techniques could be used, such as prioritising the order of tests by the percentage of identified conditions per €. However, RL can be highly adjustable to meet the needs of a study, and in future studies, it is also possible to include more data, such as test results and CT images. Therefore, from some perspectives, Study V could be seen as a necessary step before more advanced models with less control from humans can be developed.

In conclusion, the analyses in the substudies were cross-sectional in nature, and we cannot evaluate any causation. This is especially relevant for Studies III and IV, and future studies should evaluate factors relevant to breathlessness via longitudinal designs.

What will the future tell us? Image created by the author in collaboration with Bing Designer. Inspired by Salvador Dalí.

Future aspects

- **Could the factors of the MDP and D12 be reconsidered?**

The factor analysis of the MDP and D12 could be reconsidered through explorative factor analysis. Additionally, through meta-analysis of published studies reporting the factor loadings of the MDP and D12 items, the factor structure validity could be evaluated. Metanalysis of publication bias via funnel plots or similar should be considered for studies using factor analysis for the MDP and D12, as it is very likely that studies failing to confirm the factor structure are not published to the same extent as studies that are able to confirm the factor structure (and therefore validate the study for use in the tested context).

- **Expand the prevalence studies of different dimensions of breathlessness.**

The VASCOL study included only older men and did not have a very large sample size. Future studies should include women and younger people to describe the prevalence of different dimensions of breathlessness. Additionally, other countries and contexts should be considered, such as low- and middle-income countries, where the prevalence of smoking might be higher and access to health care might be lower than those in Sweden.

- **Longitudinal analysis of the different dimensions in relation to mortality and morbidity in the population.**

Studying the dimensions of breathlessness in relation to mortality and morbidity over time is highly relevant. If studied, we can gain knowledge of which dimension could be relevant for screening in clinical practice among people with severe diseases such as COPD or heart failure but also in the general population. This could also identify people at risk of developing disease earlier, leading to lifestyle changes or earlier treatments for diseases.

- **Can pathological processes be linked to breathlessness dimensions?**

As mentioned, it has been suggested that some pathological processes can be linked to specific dimensions of breathlessness,[1] but studies exploring this topic are lacking. Linking breathlessness dimensions to measurements of disease severity can be relevant, but linking specific dimensions to brain activity could also provide knowledge. This could also be relevant for developing treatments for breathlessness that target specific dimensions of breathlessness, such as air hunger or chest tightness.

- **The role of coping with breathlessness over time.**

There is a lack of knowledge about how people cope with breathlessness over time and which factors strengthen or weaken their ability to cope. Currently, pharmacological treatments for breathlessness are very limited, and nonpharmacological treatments such as cognitive behavioural therapy (CBT) could support people in coping with their symptoms. However, the psychological factors relevant to coping should first be explored in future studies. This could be a psychological concept such as locus of control, sense of coherence, or self-efficacy.

- **More complex data types, such as images, can be used to explore associations with breathlessness.**

As shown in this dissertation, ML is capable of using many factors to study breathlessness. However, data with higher dimensionality, such as medical images, can be used with other algorithms, such as deep learning. This can, for example, identify patterns in brain activity on magnetic resonance imaging (MRI) or volumes determined via CT. It is often challenging to use statistical methods for these volumes, which makes deep learning useful for this purpose. New algorithms, data availability and computer power mean that it is now possible to conduct these kinds of studies on larger populations.

- **The future of AI in the development of diagnostic pathways.**

To further develop models to evaluate breathlessness and identify underlying conditions, data can be added for analysis. This can, for example, include volume data determined via CT or the result of a blood test after a test is performed in a simulated environment. This can further improve the diagnostic ability of the model but will result in greater dependence on AI and not the visual diagnostic pathway presented in this dissertation. If we (or rather the AI program) can achieve this, AI models could become more similar to human intelligence with respect to evaluation strategies and base decisions on a variety of information. The methodology presented in this dissertation could be used for other symptoms, such as cough or pain, in addition to symptoms of breathlessness.

Conclusions

This study revealed that the D12 and MDP are adequate instruments for measuring different dimensions of breathlessness in epidemiological studies, but the factor structure should be further evaluated. Breathlessness is prevalent in the older male population, and many individuals experience unpleasant, physical and emotional domains. These experiences seem to be worse at the onset of breathlessness and decrease over time but might increase again with the progression of underlying conditions.

AI has taken the first steps in breathlessness research. With respect to ML, the most important factors for breathlessness in the general population are obesity, reduced lung function, a sedentary lifestyle, and sleep apnoea. These factors may be relevant for public health interventions. A diagnostic pathway was developed via AI with the focus of detecting as many underlying conditions and with as low economic costs as possible. From the diagnostic pathway, we can conclude that when evaluating breathlessness among people in the middle-aged general population, performing a primary care visit, followed by an examination of the lungs and then an examination of the heart, is most valuable.

Acknowledgements

First, I would like to thank all the participants in the VASCOL and SCAPIS studies. Without you, this work could never be done. There are many people who have been involved in and helped me in my PhD work; thank you all! Some are especially mentioned here.

My Supervisor, Magnus Ekström. You have inspired me throughout my whole PhD journey, and I am sure you will continue to do so afterwards. Thank you for always being there and answering stupid mails from me, as well for believing that I was capable of conducting this project. Your humour has also inspired me and gave me many laughs.

Gunnar Engström, my cosupervisor. You are true inspiration for me as an epidemiology researcher. Thank you for all the valuable feedback you provided.

Anders Björkelund. You have acted as my ML mentor, your knowledge has inspired me. You have contributed valuable knowledge to ML studies. It's always fun to bring up new ideas and discuss them with you.

Kyle Pattinson. Thank you for enabling me to visit the University of Oxford for a month. The lab visit was extremely inspiring, and I felt that I developed as a person after my visit.

The research team, Jacob, Lucas, Filip, Zainab, Jonas, and Viktor. You are all good colleagues that helped me throughout my PhD. I have greatly enjoyed all of our activities together, such as our meetings in Karlskrona and the conferences worldwide. I hope that we can collaborate in the future.

My colleagues at Kristianstad University. Thank you for all the support you provided when I was a bachelor's student and now as a colleague. You have truly inspired me with your leadership and pedagogic knowledge and experiences.

My family, my father, mother, Per, Nesrin, Elif, Petter, Jessi, and Tora. Thank you for always being there for me. Father, you have always been an inspiration with your systematic thinking and approach to stuff, something that I have adopted in my work and has helped me a lot as a researcher. Your botanical work has always been an inspiration for me as a researcher. Mother, with your academic career you have always been an inspiration for me to do the same. Thank you for all the inspiration and guidance you have provided throughout my academic career. Thank you for your great sense of humour and all the interesting discussions we have had. Nesrin, thank you for all your support and for hosting us at your place in Türkiye; it has always been valuable to regain energy between studies. Elif and Petter, thank you for all the fun moments we had and will have. My sisters, Jessi and Tora, thank you for inspiring me with your motivation and drive in life.

My wife, Zeynep. You are the love of my life; thank you for being you. You have always been there for me. We can certainly say that we experienced our PhD journeys together with many discussions about our PhDs. I am sure that your PhD will be very great and that you will have a great academic career.

My final acknowledgement goes to my and Zeynep´s unborn child. I'm very excited to welcome you to the world in November. You are so welcome and will be very loved.

References

1. Parshall MB, Schwartzstein RM, Adams L, 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):435-52. doi: 10.1164/rccm.201111-2042ST [published Online First: 2012/02/18]
2. Laviolette L, Laveneziana P, Faculty ERSRS. Dyspnoea: a multidimensional and multidisciplinary approach. *Eur Respir J* 2014;43(6):1750-62. doi: 10.1183/09031936.00092613 [published Online First: 2014/02/15]
3. Burki NK, Lee LY. Mechanisms of dyspnea. *Chest* 2010;138(5):1196-201. doi: 10.1378/chest.10-0534
4. Scano G, Gigliotti F, Stendardi L, et al. Dyspnea and emotional states in health and disease. *Respir Med* 2013;107(5):649-55. doi: 10.1016/j.rmed.2012.12.018 [published Online First: 20130121]
5. Faull OK, Hayen A, Pattinson KTS. Breathlessness and the body: Neuroimaging clues for the inferential leap. *Cortex* 2017;95:211-21. doi: 10.1016/j.cortex.2017.07.019 [published Online First: 20170809]
6. Finnegan SL, Dearlove DJ, Morris P, et al. Breathlessness in a virtual world: An experimental paradigm testing how discrepancy between VR visual gradients and pedal resistance during stationary cycling affects breathlessness perception. *PLoS One* 2023;18(4):e0270721. doi: 10.1371/journal.pone.0270721 [published Online First: 20230421]
7. Finnegan SL, Browning M, Duff E, et al. Brain activity measured by functional brain imaging predicts breathlessness improvement during pulmonary rehabilitation. *Thorax* 2023;78(9):852-59. doi: 10.1136/thorax-2022-218754 [published Online First: 20221226]
8. Sandberg J, Ekstrom M, Borjesson M, et al. Underlying contributing conditions to breathlessness among middle-aged individuals in the general population: a cross-sectional study. *BMJ Open Respir Res* 2020;7(1) doi: 10.1136/bmjresp-2020-000643
9. Poulos LM, Ampon RD, Currow DC, et al. Prevalence and burden of breathlessness in Australian adults: The National Breathlessness Survey-a cross-sectional web-based population survey. *Respirology* 2021;26(8):768-75. doi: 10.1111/resp.14070 [published Online First: 20210510]
10. Ekstrom MP, Abernethy AP, Currow DC. The management of chronic breathlessness in patients with advanced and terminal illness. *BMJ* 2015;350:g7617. doi: 10.1136/bmj.g7617 [published Online First: 2015/01/04]
11. Frostad A, Soyseth V, Andersen A, et al. Respiratory symptoms as predictors of all-cause mortality in an urban community: a 30-year follow-up. *J Intern Med*

2006;259(5):520-9. doi: 10.1111/j.1365-2796.2006.01631.x [published Online First: 2006/04/25]
12. Sandberg J, Engstrom G, Ekstrom M. Breathlessness and incidence of COPD, cardiac events and all-cause mortality: A 44-year follow-up from middle age throughout life. *PLoS One* 2019;14(3):e0214083. doi: 10.1371/journal.pone.0214083 [published Online First: 2019/03/19]
13. Nishimura K, Izumi T, Tsukino M, et al. Dyspnea is a better predictor of 5-year survival than airway obstruction in patients with COPD. *Chest* 2002;121(5):1434-40. doi: 10.1378/chest.121.5.1434 [published Online First: 2002/05/15]
14. Currow DC, Dal Grande E, Ferreira D, et al. Chronic breathlessness associated with poorer physical and mental health-related quality of life (SF-12) across all adult age groups. *Thorax* 2017;72(12):1151-53. doi: 10.1136/thoraxjnl-2016-209908 [published Online First: 2017/03/31]
15. Currow DC, Chang S, Reddel HK, et al. Breathlessness, Anxiety, Depression, and Function-The BAD-F Study: A Cross-Sectional and Population Prevalence Study in Adults. *J Pain Symptom Manage* 2020;59(2):197-205 e2. doi: 10.1016/j.jpainsymman.2019.09.021 [published Online First: 20191022]
16. Currow DC, Chang S, Grande ED, et al. Quality of Life Changes With Duration of Chronic Breathlessness: A Random Sample of Community-Dwelling People. *J Pain Symptom Manage* 2020;60(4):818-27.e4. doi: 10.1016/j.jpainsymman.2020.05.015
17. Ekstrom M, Johnson MJ, Taylor B, et al. Breathlessness and sexual activity in older adults: the Australian Longitudinal Study of Ageing. *NPJ Prim Care Respir Med* 2018;28(1):20. doi: 10.1038/s41533-018-0090-x [published Online First: 20180622]
18. Olsson M, Bjorkelund AJ, Sandberg J, et al. Factors important for health-related quality of life in men and women: The population based SCAPIS study. *PLoS One* 2023;18(11):e0294030. doi: 10.1371/journal.pone.0294030 [published Online First: 20231103]
19. Hutchinson A, Barclay-Klingle N, Galvin K, et al. Living with breathlessness: a systematic literature review and qualitative synthesis. *Eur Respir J* 2018;51(2) doi: 10.1183/13993003.01477-2017 [published Online First: 2018/02/23]
20. Johnson MJ, Yorke J, Hansen-Flaschen J, et al. Towards an expert consensus to delineate a clinical syndrome of chronic breathlessness. *Eur Respir J* 2017;49(5) doi: 10.1183/13993003.02277-2016 [published Online First: 2017/05/27]
21. Kochovska S, Chang S, Ferreira D, et al. Invisibility of breathlessness in clinical consultations: a cross-sectional, national online survey. *Eur Respir J* 2022 doi: 10.1183/13993003.01603-2022 [published Online First: 20221006]
22. Chapple A, Ziebland S, McPherson A. Stigma, shame, and blame experienced by patients with lung cancer: qualitative study. *BMJ* 2004;328(7454):1470. doi: 10.1136/bmj.38111.639734.7C [published Online First: 20040611]
23. Gysels M, Higginson IJ. Access to services for patients with chronic obstructive pulmonary disease: the invisibility of breathlessness. *J Pain Symptom Manage* 2008;36(5):451-60. doi: 10.1016/j.jpainsymman.2007.11.008 [published Online First: 20080520]

24. Ekstrom M, Ferreira D, Chang S, et al. Effect of Regular, Low-Dose, Extended-release Morphine on Chronic Breathlessness in Chronic Obstructive Pulmonary Disease: The BEAMS Randomized Clinical Trial. *JAMA* 2022;328(20):2022-32. doi: 10.1001/jama.2022.20206
25. Hui D, Bohlke K, Bao T, et al. Management of Dyspnea in Advanced Cancer: ASCO Guideline. *J Clin Oncol* 2021;39(12):1389-411. doi: 10.1200/jco.20.03465 [published Online First: 20210222]
26. Ekström M, Nilsson F, Abernethy AA, et al. Effects of opioids on breathlessness and exercise capacity in chronic obstructive pulmonary disease. A systematic review. *Ann Am Thorac Soc* 2015;12(7):1079-92. doi: 10.1513/AnnalsATS.201501-034OC
27. Ekström M. Fan therapy is a treatment option for relieving of chronic breathlessness. *Evidence Based Nursing* 2020;23(3):73. doi: 10.1136/ebnurs-2019-103118
28. Raymond B, Luckett T, Johnson M, et al. Low-intensity educational interventions supporting self-management to improve outcomes related to chronic breathlessness: a systematic review. *npj Primary Care Respiratory Medicine* 2019;29(1):41. doi: 10.1038/s41533-019-0152-8
29. Rueda JR, Sola I, Pascual A, et al. Non-invasive interventions for improving well-being and quality of life in patients with lung cancer. *Cochrane Database Syst Rev* 2011;2011(9):CD004282. doi: 10.1002/14651858.CD004282.pub3 [published Online First: 20110907]
30. Higginson IJ, Bausewein C, Reilly CC, et al. An integrated palliative and respiratory care service for patients with advanced disease and refractory breathlessness: a randomised controlled trial. *Lancet Respir Med* 2014;2(12):979-87. doi: 10.1016/S2213-2600(14)70226-7 [published Online First: 20141029]
31. Fletcher CM. The clinical diagnosis of pulmonary emphysema; an experimental study. *Proc R Soc Med* 1952;45(9):577-84.
32. Fletcher C. Standardised questionnaire on respiratory symptoms: a statement prepared and approved by the MRC Committee on the Aetiology of Chronic Bronchitis (MRC breathlessness score). *Bmj* 1960;2(2):1665.
33. Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest* 1988;93(3):580-6. doi: 10.1378/chest.93.3.580 [published Online First: 1988/03/01]
34. Yorke J, Khan N, Garrow A, et al. Evaluation of the Individual Activity Descriptors of the mMRC Breathlessness Scale: A Mixed Method Study. *Int J Chron Obstruct Pulmon Dis* 2022;17:2289-99. doi: 10.2147/copd.S372318 [published Online First: 20220915]
35. Gustafsson D, Elmberg V, Schioler L, et al. The modified Medical Research Council scale misclassifies exertional breathlessness among people referred for exercise testing. *ERJ Open Res* 2023;9(6) doi: 10.1183/23120541.00592-2023 [published Online First: 20231227]
36. Sandberg J, Lansing R, Anderberg P, et al. Relating Experienced To Recalled breathlessness Observational (RETRO) study: a prospective study using a mobile phone application. *BMJ Open Respir Res* 2019;6(1):e000370. doi: 10.1136/bmjresp-2018-000370 [published Online First: 2019/04/09]

37. Bausewein C, Farquhar M, Booth S, et al. Measurement of breathlessness in advanced disease: a systematic review. *Respir Med* 2007;101(3):399-410. doi: 10.1016/j.rmed.2006.07.003 [published Online First: 20060817]
38. Yorke J, Moosavi SH, Shuldham C, et al. Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax* 2010;65(1):21-6. doi: 10.1136/thx.2009.118521 [published Online First: 2009/12/10]
39. Banzett RB, O'Donnell CR, Guilfoyle TE, et al. Multidimensional Dyspnea Profile: an instrument for clinical and laboratory research. *Eur Respir J* 2015;45(6):1681-91. doi: 10.1183/09031936.00038914 [published Online First: 2015/03/21]
40. Williams MT, Lewthwaite H, Paquet C, et al. Dyspnoea-12 and Multidimensional Dyspnea Profile: Systematic Review of Use and Properties. *J Pain Symptom Manage* 2021 doi: 10.1016/j.jpainsymman.2021.06.023 [published Online First: 2021/07/18]
41. Sundh J, Bornefalk H, Skold CM, et al. Clinical validation of the Swedish version of Dyspnoea-12 instrument in outpatients with cardiorespiratory disease. *BMJ Open Respir Res* 2019;6(1):e000418. doi: 10.1136/bmjresp-2019-000418 [published Online First: 2019/11/02]
42. Ekstrom M, Bornefalk H, Skold M, et al. Validation of the Swedish Multidimensional Dyspnea Profile (MDP) in outpatients with cardiorespiratory disease. *BMJ Open Respir Res* 2019;6(1):e000381. doi: 10.1136/bmjresp-2018-000381 [published Online First: 2019/11/05]
43. Gronseth R, Vollmer WM, Hardie JA, et al. Predictors of dyspnoea prevalence: results from the BOLD study. *Eur Respir J* 2014;43(6):1610-20. doi: 10.1183/09031936.00036813 [published Online First: 2013/11/02]
44. Ekstrom M, Sundh J, Schioler L, et al. Absolute lung size and the sex difference in breathlessness in the general population. *PLoS One* 2018;13(1):e0190876. doi: 10.1371/journal.pone.0190876 [published Online First: 2018/01/06]
45. Lopez Varela MV, Montes de Oca M, Halbert RJ, et al. Sex-related differences in COPD in five Latin American cities: the PLATINO study. *Eur Respir J* 2010;36(5):1034-41. doi: 10.1183/09031936.00165409 [published Online First: 20100408]
46. Sandberg J, Olsson M, Ekstrom M. Underlying conditions contributing to breathlessness in the population. *Curr Opin Support Palliat Care* 2021;15(4):219-25. doi: 10.1097/SPC.0000000000000568
47. Hinton G. Deep Learning-A Technology With the Potential to Transform Health Care. *JAMA* 2018;320(11):1101-02. doi: 10.1001/jama.2018.11100 [published Online First: 2018/09/05]
48. Mooney SJ, Pejaver V. Big Data in Public Health: Terminology, Machine Learning, and Privacy. *Annu Rev Public Health* 2018;39:95-112. doi: 10.1146/annurev-publhealth-040617-014208 [published Online First: 2017/12/21]
49. Pedersen F, Mehlsen J, Raymond I, et al. Evaluation of dyspnoea in a sample of elderly subjects recruited from general practice. *Int J Clin Pract* 2007;61(9):1481-91. doi: 10.1111/j.1742-1241.2007.01428.x [published Online First: 2007/08/10]

50. Sunjaya AP, Homaira N, Corcoran K, et al. Assessment and diagnosis of chronic dyspnoea: a literature review. *NPJ Prim Care Respir Med* 2022;32(1):10. doi: 10.1038/s41533-022-00271-1 [published Online First: 20220308]
51. Complex Breathlessness: European Respiratory Society 2022.
52. Russell SJ, Norvig P. Artificial intelligence : a modern approach. Hoboken, NJ: Pearson 2021.
53. Turing AM. I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind* 1950;LIX(236):433-60. doi: 10.1093/mind/LIX.236.433
54. Mei Q, Xie Y, Yuan W, et al. A Turing test of whether AI chatbots are behaviorally similar to humans. *Proc Natl Acad Sci U S A* 2024;121(9):e2313925121. doi: 10.1073/pnas.2313925121 [published Online First: 20240222]
55. Biever C. ChatGPT broke the Turing test - the race is on for new ways to assess AI. *Nature* 2023;619(7971):686-89. doi: 10.1038/d41586-023-02361-7
56. Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development* 1959;3(3):210-29. doi: 10.1147/rd.33.0210
57. Lindsay RP. The Impact of Automation On Public Administration. *Western Political Quarterly* 1964;17(3):78-81. doi: 10.1177/106591296401700364
58. Koza JR, Bennett FH, Andre D, et al. Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. In: Gero JS, Sudweeks F, eds. Artificial Intelligence in Design '96. Dordrecht: Springer Netherlands 1996:151-70.
59. Chen PC, Liu Y, Peng L. How to develop machine learning models for healthcare. *Nat Mater* 2019;18(5):410-14. doi: 10.1038/s41563-019-0345-0 [published Online First: 2019/04/20]
60. François-Lavet V, Henderson P, Islam R, et al. An Introduction to Deep Reinforcement Learning. *Foundations and Trends® in Machine Learning* 2018;11(3-4):219-354. doi: 10.1561/2200000071
61. Hu M, Zhang J, Matkovic L, et al. Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions. *J Appl Clin Med Phys* 2023;24(2):e13898. doi: 10.1002/acm2.13898 [published Online First: 20230110]
62. Adadi A, Berrada M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *Ieee Access* 2018;6:52138-60. doi: 10.1109/Access.2018.2870052
63. Lundberg SM, Erion G, Chen H, et al. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat Mach Intell* 2020;2(1):56-67. doi: 10.1038/s42256-019-0138-9 [published Online First: 2020/07/02]
64. Nowak C, Sievi NA, Clarenbach CF, et al. Accuracy of the Hospital Anxiety and Depression Scale for identifying depression in chronic obstructive pulmonary disease patients. *Pulm Med* 2014;2014:973858. doi: 10.1155/2014/973858 [published Online First: 2014/12/31]
65. Al-shair K, Muellerova H, Yorke J, et al. Examining fatigue in COPD: development, validity and reliability of a modified version of FACIT-F scale. *Health Qual Life Outcomes* 2012;10:100. doi: 10.1186/1477-7525-10-100 [published Online First: 2012/08/24]

66. Lundh Hagelin C, Klarare A, Furst CJ. The applicability of the translated Edmonton Symptom Assessment System: revised [ESAS-r] in Swedish palliative care. *Acta Oncol* 2018;57(4):560-62. doi: 10.1080/0284186X.2017.1366050 [published Online First: 2017/08/18]
67. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220-33. doi: 10.1097/00005650-199603000-00003 [published Online First: 1996/03/01]
68. Bergstrom G, Berglund G, Blomberg A, et al. The Swedish CArdioPulmonary BioImage Study: objectives and design. *J Intern Med* 2015;278(6):645-59. doi: 10.1111/joim.12384 [published Online First: 2015/06/23]
69. Bonander C, Nilsson A, Bjork J, et al. The value of combining individual and small area sociodemographic data for assessing and handling selective participation in cohort studies: Evidence from the Swedish CardioPulmonary bioImage Study. *PLoS One* 2022;17(3):e0265088. doi: 10.1371/journal.pone.0265088 [published Online First: 20220308]
70. Mair P. Modern psychometrics with R. Cham, Switzerland: Springer 2018.
71. George D, Mallery P. SPSS for Windows step by step : a simple guide and reference 11.0 update. Boston: Allyn and Bacon 2003.
72. Copay AG, Subach BR, Glassman SD, et al. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007;7(5):541-6. doi: 10.1016/j.spinee.2007.01.008 [published Online First: 20070402]
73. Ekstrom MP, Bornefalk H, Skold CM, et al. Minimal Clinically Important Differences and Feasibility of Dyspnea-12 and the Multidimensional Dyspnea Profile in Cardiorespiratory Disease. *J Pain Symptom Manage* 2020 doi: 10.1016/j.jpainsymman.2020.05.028 [published Online First: 2020/06/09]
74. Ekstrom M, Bornefalk H, Skold CM, et al. Minimal clinically important differences for Dyspnea-12 and MDP scores are similar at 2 weeks and 6 months: follow-up of a longitudinal clinical study. *Eur Respir J* 2020 doi: 10.1183/13993003.02823-2020 [published Online First: 2020/12/19]
75. Badillo S, Banfai B, Birzele F, et al. An Introduction to Machine Learning. *Clin Pharmacol Ther* 2020;107(4):871-85. doi: 10.1002/cpt.1796 [published Online First: 20200303]
76. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. 2016. https://ui.adsabs.harvard.edu/abs/2016arXiv160302754C (accessed March 01, 2016).
77. Olsson M, Currow DC, Ekstrom MP. Exploring the most important factors related to self-perceived health among older men in Sweden: a cross-sectional study using machine learning. *BMJ Open* 2022;12(6):e061242. doi: 10.1136/bmjopen-2022-061242 [published Online First: 20220621]
78. World Health Organization. Haemoglobin Concentrations for the Diagnosis of Anaemia and Assessment of Severity: World Health Organization. 2011.
79. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *European Respiratory Journal* 2005;26(2):319. doi: 10.1183/09031936.05.00034805

80. Malinovschi A, Zhou X, Andersson A, et al. Consequences of Using Post- or Prebronchodilator Reference Values in Interpreting Spirometry. *Am J Respir Crit Care Med* 2023;208(4):461-71. doi: 10.1164/rccm.202212-2341OC
81. Hatabu H, Hunninghake GM, Richeldi L, et al. Interstitial lung abnormalities detected incidentally on CT: a Position Paper from the Fleischner Society. *Lancet Respir Med* 2020;8(7):726-37. doi: 10.1016/S2213-2600(20)30168-5
82. MacIntyre N, Crapo RO, Viegi G, et al. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *European Respiratory Journal* 2005;26(4):720. doi: 10.1183/09031936.05.00034905
83. Vikgren J, Khalil M, Cederlund K, et al. Visual and Quantitative Evaluation of Emphysema: A Case-Control Study of 1111 Participants in the Pilot Swedish CArdioPulmonary BioImage Study (SCAPIS). *Acad Radiol* 2020;27(5):636-43. doi: 10.1016/j.acra.2019.06.019 [published Online First: 20190718]
84. Rosengren A, Hawken S, Ounpuu S, et al. Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet* 2004;364(9438):953-62. doi: 10.1016/S0140-6736(04)17019-0
85. Welsh P, Campbell RT, Mooney L, et al. Reference Ranges for NT-proBNP (N-Terminal Pro-B-Type Natriuretic Peptide) and Risk Factors for Higher NT-proBNP Concentrations in a Large General Population Cohort. *Circ Heart Fail* 2022;15(10):e009427. doi: 10.1161/CIRCHEARTFAILURE.121.009427 [published Online First: 20220913]
86. Regionala priser och ersättningar för södra sjukvårdsregionen: Södra sjukvårdsregionen, 2022.
87. Park J-H, Hong JY, Han K. Threshold dose–response association between smoking pack-years and the risk of gallbladder cancer: A nationwide cohort study. *European Journal of Cancer* 2023;180:99-107. doi: https://doi.org/10.1016/j.ejca.2022.11.031
88. Winder P. Reinforcement learning : industrial applications of intelligent agents2020.
89. Raffin A, Hill A, Gleave A, et al. Stable-baselines3: reliable reinforcement learning implementations. *J Mach Learn Res* 2021;22(1):Article 268.
90. Asynchronous methods for deep reinforcement learning. International conference on machine learning; 2016. PMLR.
91. Stevens JP, Sheridan AR, Bernstein HB, et al. A Multidimensional Profile of Dyspnea in Hospitalized Patients. *Chest* 2019;156(3):507-17. doi: 10.1016/j.chest.2019.04.128 [published Online First: 2019/05/28]
92. Morelot-Panzini C, Gilet H, Aguilaniu B, et al. Real-life assessment of the multidimensional nature of dyspnoea in COPD outpatients. *Eur Respir J* 2016;47(6):1668-79. doi: 10.1183/13993003.01998-2015 [published Online First: 2016/04/15]
93. Socialstyrelsen. Nationella riktlinjer för vård vid astma och KOL : stöd för styrning och ledning. [Stockholm]: Socialstyrelsen 2020.
94. Ekström M. Obesity is a major contributing cause of breathlessness in the population. *Respirology* 2022;n/a(n/a) doi: https://doi.org/10.1111/resp.14421
95. Ekström MP, Blomberg A, Bergström G, et al. The association of body mass index, weight gain and central obesity with activity-related breathlessness: the Swedish

Cardiopulmonary Bioimage Study. *Thorax* 2019;74(10):958. doi: 10.1136/thoraxjnl-2019-213349

96. Abdelaal M, le Roux CW, Docherty NG. Morbidity and mortality associated with obesity. *Ann Transl Med* 2017;5(7):161. doi: 10.21037/atm.2017.03.107

97. Boye KS, Ford JH, Thieu VT, et al. The Association Between Obesity and the 5-Year Prevalence of Morbidity and Mortality Among Adults with Type 2 Diabetes. *Diabetes Therapy* 2023;14(4):709-21. doi: 10.1007/s13300-023-01384-7

98. Kochovska S, Currow D, Chang S, et al. Persisting breathlessness and activities reduced or ceased: a population study in older men. *BMJ Open Respiratory Research* 2022;9(1):e001168. doi: 10.1136/bmjresp-2021-001168

99. Jakeways N, McKeever T, Lewis SA, et al. Relationship between FEV₁ reduction and respiratory symptoms in the general population. *European Respiratory Journal* 2003;21(4):658. doi: 10.1183/09031936.03.00069603

100. Gallucci G, Tartarone A, Lerose R, et al. Cardiovascular risk of smoking and benefits of smoking cessation. *J Thorac Dis* 2020;12(7):3866-76. doi: 10.21037/jtd.2020.02.47

101. Gan H, Hou X, Zhu Z, et al. Smoking: a leading factor for the death of chronic respiratory diseases derived from Global Burden of Disease Study 2019. *BMC Pulmonary Medicine* 2022;22(1):149. doi: 10.1186/s12890-022-01944-w

102. Regitz-Zagrosek V, Gebhard C. Gender medicine: effects of sex and gender on cardiovascular disease manifestation and outcomes. *Nature Reviews Cardiology* 2023;20(4):236-47. doi: 10.1038/s41569-022-00797-4

103. Barnett K, Mercer SW, Norbury M, et al. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;380(9836):37-43. doi: 10.1016/S0140-6736(12)60240-2 [published Online First: 2012/05/15]

104. Ekström M, Li PZ, Lewthwaite H, et al. Normative Reference Equations for Breathlessness Intensity during Incremental Cardiopulmonary Cycle Exercise Testing. *Ann Am Thorac Soc* 2024;21(1):56-67. doi: 10.1513/AnnalsATS.202305-394OC

105. Ekstrom M, Sundh J, Andersson A, et al. Exertional breathlessness related to medical conditions in middle-aged people: the population-based SCAPIS study of more than 25,000 men and women. *Respir Res* 2024;25(1):127. doi: 10.1186/s12931-024-02766-6 [published Online First: 20240316]

Epidemiology and evaluation of breathlessness

This thesis is about long-term breathlessness. Breathlessness is a debilitating symptom that is common in the general population. Knowledge about the epidemiology of breathlessness is lacking, particularly concerning the measurement of breathlessness, the prevalence of different dimensions related to breathlessness, and factors associated with this symptom in the general population. It is also challenging to identify the underlying condition(s) leading to breathlessness due to a lack of evidence-based diagnostic pathways to determine the health conditions contributing to a patient's breathlessness. This thesis aims to explore these areas using basic quantitative methods as well as novel data-driven approaches.

The author, **Max Olsson**, has a background in public health and epidemiology, with a special interest in breathlessness and machine learning.

FACULTY OF MEDICINE

Department of Clinical Sciences Lund

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2024:101
ISBN 978-91-8021-597-8
ISSN 1652-8220

Printed by Media-Tryck, Lund 2024 NORDIC SWAN ECOLABEL 3041 0903