ESSACHESS – Journal for Communication Studies

Questioning the Relevance of Sex Categories implemented in Medical Decision Support Systems - The Example of Pulmonary Function

ESSACHESS –
Journal for Communication Studies
Volume 17 Issue 1(33), p. 205-229

https://www.essachess.com/
DOI: 10.21409/essachess.1775-352x

Cite: LEMARCHAND, P., HASSOUN, D., & KUNTZ, P. (2024). Questioning the Relevance of Sex Categories implemented in Medical Decision Support Systems - The Example of Pulmonary Function. *ESSACHESS*. 17(33) : 205-229. https://doi.org/10.21409/8253-X893

Patricia LEMARCHAND
Professeure des universités – Praticien Hospitalier, Nantes Université, CHU Nantes, CNRS, INSERM, l'institut du thorax, FRANCE
e-mail: patricia.lemarchand@univ-nantes.fr

Dorian HASSOUN
Assistant Hospitalo-Universitaire, Nantes Université, CHU Nantes, CNRS, INSERM, l'institut du thorax, FRANCE
e-mail: dorian.hassoun@chu-nantes.fr

Pascale KUNTZ
Professeure des universités, Nantes Université, Laboratoire des Sciences du Numérique de Nantes, CNRS, FRANCE
e-mail: pascale.kuntz@univ-nantes.fr

Abstract: Medical decision support systems rely on a variety of data to provide advice and predictions that contribute to diagnoses. Data categorizations, which are often

Article received on February 21, 2024. Article accepted on June 17, 2024.
Conflict of Interest: The authors declare no conflict of interest.

hidden and therefore invisible, play a major role in the statistical models implemented in these digital tools. Male/female bicategorization is a paradigmatic case that has been little studied in this context. This study analyzes its use and determinants in pulmonary function measurement as a case study. Using a corpus of sixty articles, carefully selected in the medical literature for their representativeness, and the reference equations on which decision support systems are based, we investigate the role of male/female categorization as used in everyday clinical practice, its origins, and the place of sex and gender in this issue. This research reveals a naturalization of men/women differences in favor of sex, underpinning a binary essentialism of biological sex. Naturalization of men/women differences is a major concern for the use of predictive artificial intelligence models and the development of decision algorithms, with a possible worsening of health disparities as a result of biases in training data.

Keywords: medical decision support system, sex, gender, male/female categorization, pulmonary function

***

Prise en compte du genre dans les systèmes d'aide à la décision médicale - L'exemple de la fonction respiratoire

Résumé : Les systèmes d'aide à la décision médicale s'appuient sur des données hétérogènes et de volumétries variées pour fournir des conseils et des prédictions contribuant aux diagnostics. Des catégorisations, souvent invisibles, jouent un rôle majeur dans les modèles statistiques implémentés dans ces outils numériques. La bicatégorisation homme/femme est un cas paradigmatique qui a été peu étudié dans ce cadre. Cet article analyse son usage et ses déterminants dans une étude de cas, portant sur la mesure de la fonction respiratoire. Nous avons constitué un corpus de soixante articles de la littérature médicale (sélectionnés pour leur représentativité) et des équations de référence sur lesquels s'appuient les systèmes d'aide à la décision utilisés en routine clinique. Dans ce corpus, nous analysons le rôle de cette catégorisation homme/femme, ses origines, et la place du sexe et du genre dans cette problématique. Cette recherche montre une naturalisation des différences hommes-femmes au profit du sexe, sous-tendant un essentialisme binaire du sexe biologique. Cette naturalisation des différences homme/femme représente une préoccupation importante pour l'utilisation de modèles prédictifs d'intelligence artificielle et le développement d'algorithmes décisionnels, avec le risque d'aggraver les disparités en matière de santé en raison de biais dans les données d'entrainement.

Mots-clés : système d'aide à la décision médicale, sexe, genre, bicatégorisation homme/femme, fonction respiratoire

***

Introduction

With the development of translational medicine (Woolf, 2008), which aims to bridge the gap between biomedical research and clinical practice, and the rise of artificial intelligence (AI) technologies that enable large-scale data processing, medical decision support systems are rapidly becoming indispensable tools, integrated into increasingly complex digital frameworks (Berner, 2007). Healthcare professionals access the answers provided by medical decision support systems by entering the required information via a specialized interface. These responses are then associated with variables in algorithms coded in a programming language. More precisely, they shape research fields in which computer programs calculate solutions to decision problems. As one of the pioneers of interface studies states, "the interface asks a question and, in so doing, suggests an answer" (Galloway, 2012, p. 30).

In medical decision support systems, these programs implement statistical models derived from theoretical and experimental research. They allow physicians in clinical situations to compare the variables measured in an individual to predicted values obtained from a "reference population", based on equations recommended by medical societies. For example, in the case study developed in this article focusing on pulmonary function, the models predict the expected values of different respiratory function variables for a "healthy" individual based on a set of variables entered via the interface. The definition of these variables guides the medical practitioners' questioning via the interface, and subsequently informs medical decisions based on the values predicted by the statistical model implemented in the decision support system software.

To provide this diagnostic support, systems rely on ever-increasing amounts of heterogeneous data describing selected characteristics of the individuals in the studied cohorts. These characteristics are based on similarly diverse categorizations. However, as Geoffrey Bowker and Susan Leigh Star remind us in their seminal work, classifications are particularly powerful technologies. When embedded in operational infrastructures, they become relatively invisible without losing their power (Bowker & Star, 1999). Research in *Science & Technology Studies* (STS) has confirmed the deep connections between technical objects and the socio-economic characteristics of the environments in which they are shaped.

The STS approach has integrated the gender lens, notably through the work of Delphine Gardey (Gardey & Chabaud-Rychter, 2002), and more recent studies have transposed the analysis to digital objects, particularly to interfaces (Pidoux & Kuntz, 2024). Medical decision support systems have nonetheless rarely been examined from this perspective. Yet, regardless of the clinical domain in question, one variable that is consistently present on their interfaces is "sex", always considered within the framework of a binary male/female categorization.

Since the late 20th century, considerable research has been carried out to highlight the differences between men and women in terms of biology and health. This has led to sex-specific recommendations in the medical management of a variety of pathologies (Mauvais-Jarvis et al., 2020). However, this research is typically based on the postulate of the naturalized male/female binary category — which is mutually exclusive and is rarely questioned (Raz, 2016) — where categorization as male or female is based on the self-declaration of individuals included in these studies. The male/female categorization routinely refers to either the "administrative sex" in clinical records or the answer to the question "What is your sex?", and in both cases, it is based on the genitalia observed at birth and its transcription as "sex" on the birth certificate (Schwarz & Clair, 2023).

In line with *feminist science studies*, initiated in part by women biologists such as Anne Fausto-Sterling (Fausto-Sterling, 2000), we question here the use of a binary male/female categorization by shifting our focus to the other side of the screen of medical decision support systems, to explore its foundations in the statistical models implemented therein. To overcome the pitfalls of the abovementioned invisible infrastructure, we have chosen to focus on a biomedical specialty in which two of the article's authors have recognized expertise: pulmonology, with a focus on pulmonary function. This is a vital bodily function ensured by the respiratory system, responsible for supplying oxygen and eliminating carbon dioxide produced by physiological activity. In day-to-day clinical practice, it is easily evaluated by pulmonary function tests using a spirometer, which simultaneously measures airflow rates and lung volumes. The interpretation of pulmonary function tests ideally relies on their comparison to a previous measurement obtained in the same healthy individual. In the absence of such a measurement, interpretation depends on their comparison to reference values calculated using prediction equations developed from a healthy population. This allows the precise assessment of whether an individual's measured values are normal or abnormal, and thus enables a functional diagnosis. The construction of these equations and their results both depend closely on a male/female binary categorization.

In this study, we evaluated the role of this binary categorization and its impact within the reference models of medical decision support systems used by practitioners. More precisely, our research questions focused on three key areas: (i) how male/female categorization has been applied in lung function prediction since the development of the spirometer, and how any observed men/women differences[1] were analyzed, (ii) how male/female categorization is currently operationalized in the reference software for medical decision support systems, and (iii) how sex and gender are operationalized in the literature on lung function and in the explanatory factors advanced to explain the observed men/women differences. Given the lack of state-of-the-art research on these issues, we conducted an exploratory study combining a detailed analysis of the main reference equations with a terminological analysis conducted from a longitudinal perspective. We built a corpus of sixty carefully

[1] We recognize that researchers regularly use male/female to classify research human participants. In this study, we use men/women differences to reflect the intersection of gender/sex in human health experiences, and male/female categorization to reflect the literature analyzed in this study.

selected articles reflecting the research evolution in this field, including articles and literature reviews focusing on the measurement of pulmonary function, as well as resulting recommendations and/or guidelines from North American and European medical societies. We then added references regarding the role of sex or gender in pulmonary function and references questioning the use of sex as a biological variable in biomedical research. From this corpus, we identified a subset of terms including "sex" and "gender" in each article, observing the presence or absence of explicitly associated definitions, and analyzed the role of gender in data interpretation.

This pioneering work in the context of pulmonary function highlights the lack of argumentation for the binary male/female categorization in current usage models. We show that this is based on a historical elaboration of knowledge whose scientific validity is widely questioned by complementary research. First of all, while the literature agrees on the existence of men/women differences in average values, the distributions of collected measurements deserve to be systematically analyzed in their entirety. Indeed, a number of studies have shown significant overlaps between these distributions, thus necessitating the development of alternative models to those based on partitioning (Brucker & Barthélemy, 2007). Secondly, the explanation of men/women differences by essentialist causes is not convincing. As has been shown over the last decade for the ethnic variable[2] — considered in models as having a similar status to the "sex" variable —, other socio-cultural explanatory factors, such as physical activity levels, deserve to be analyzed.

1. Male/female Binary Categorization in the Field of Pulmonary Function: An Exploratory Analysis

Compared with other disciplines, biomedical literature has a specificity: the same scientific journals include "original" articles presenting research results, literature reviews synthesizing data from these articles, and guidelines from medical societies. Some studies, including those resulting from research carried out by national or international working groups, combine new research data, such as the establishment new pulmonary function prediction equations, with practical recommendations for their use in the clinical setting. These recommendations are rapidly becoming prescriptive guidelines for medical decision-making (Pellegrino et al., 2005). For example, idiopathic pulmonary fibrosis is a disease where such guidelines, using percentage predicted thresholds, determine eligibility for publicly-funded treatment. A recent study by Burgess et al. (2019) reveals a significant discrepancy between predicted pulmonary function values across different interstitial lung disease centers due to variations in the prediction equations used. This variability may directly impact a patient's acceptance into reimbursed treatment, potentially creating inequalities in access (Cooper et al., 2017).

[2] See *Discussion*.

1.1. *Source Selection*

Scientometrics has gradually gained ground (Vinck, 1995) as a tool for better understanding the mechanisms behind scientific content production. Moreover, its scope has broadened considerably due to the development of data processing technologies. Sciences long described as "natural" have been privileged domains (Suraud, 1996), although as far as we know, the gender perspective has been largely ignored or minimally considered, and the extension of scientometrics to other fields has done little to encourage its regular use in analyses (Favre, 2017). In pulmonary function analysis, no studies have been conducted from this angle.

To address this issue, we began by conducting a systematic literature search of the Ovid MEDLINE® electronic bibliographic database from 1946 to January 2024 (PubMed), and then used Google Scholar to track relevant citations of selected articles. The search included a combination of the following terms and their synonyms: lung function, pulmonary function, maximal expiratory flow-volume curve, ventilatory, spirometry, spirometric, reference values, reference equation, prediction equation, global lung function initiative, and exercise. The search strategy used free-text terms in combination with Medical Subject Heading (MeSH) filters where appropriate. Duplicates were removed using Mendeley Reference Manager. Titles and, subsequently, abstracts and full texts of the articles were screened. Exclusion criteria included: full texts being unavailable, and other publication types such as conference abstracts, case reports, editorials, letters, or articles published in languages other than English. References from selected published articles were also considered and added to the corpus when relevant. Our familiarity with the literature on lung function measurements, as well as sex/gender differences in biology and health, served as a starting point for our selection strategy.

From the initial source set, the corpus of articles retained for analysis was built up starting with the 2012 GLI reference prediction model (Quanjer et al., 2012), listed as #14 in the Appendix. Firstly, we sought out the key foundational articles used to establish this model (up to J Hutchinson's first article introducing the spirometer, #1-14) and those that validated it around the world (#15-23, 25, 28, 31). We then added recent articles questioning the use of an ethnic correction factor in this model (#33-36, 56, 57). Lastly, we added reviews published in the top 1% of medical research journals and insights from highly cited international experts in the field (#41, 44, 46), along with related guidelines from international medical societies (#39, 40, 47, 49, 50). We also included a review covering the 150-year history of spirometry (#42), and a reference textbook on lung function for pulmonologists and lung function practitioners (#45), additional articles and reviews focusing on the role of sex or gender in pulmonary function (#2, 5, 6, 24, 26, 27, 32, 38, 51-55, 58), and two references questioning the use of sex as a biological variable in biomedical research (#59-60). Two authors (PL and DH) independently assessed the retrieved references, and any disagreements were resolved through consensus.

1.2. *Data Analysis*

Data analysis was carried out in two phases. Initially, to clarify the impact of male/female categorization in algorithmic decision support, we thoroughly analyzed how the reference equations were constructed and validated (Quanjer et al., 2012). Then, to identify the rationale behind using male/female categorization, we conducted a textual analysis on the corpus of selected articles based on a subset of keywords associated with the research questions. Using a search engine, we systematically identified the presence of the following terms in each article within the corpus: sex, gender, dimorphism, sex differences, gender differences, male, female, men, women. Since the corpus was relatively small, we were able to perform a manual interpretative analysis based on our expertise. For each article, we examined the keywords within their context to track down any associated sex or gender definitions and identify recurring elements of argumentation.

2. Biomedical Literature on Male/Female Binary Categorization

The first study of pulmonary function was established in the mid-19th century by John Hutchinson, the inventor of the spirometer. One of the first measures analyzed was "vital capacity", which assesses the total gas volume mobilized between maximum expiration and inspiration, and for which height, weight (since abandoned) and age were identified as important determinants (Hutchinson, 1846). The study population involved over 2000 men of various professions and occupations (including "soldiers, sailors, guards, policemen, gentlemen, paupers, giants and dwarfs"), and 26 "girls". The "epitome" (as he called his abstract) of Hutchinson's article did not comment on the existence of gender differences, which was likely a wise decision given the limited number of women in his study sample (Becklake & Kauffmann, 1999).

Nearly thirty years after this pioneering work, Henri Havelock Ellis' book, "Man and woman: a study of human secondary sexual characters" (Havelock Ellis, 1894), covered a remarkably wide range of topics. In a chapter devoted to metabolism, Havelock Ellis notes that "it is well recognized that the vital capacity is decidedly less in women than in men" even when height is considered, indicating that vital capacity was on average 350-500 ml greater in men than in women of the same height (i.e., approximately 15 to 20%). He attributed these differences partly to biological factors ("women have less need of air than men"), due to their "lower production of carbonic acid", and partly to cultural factors such as "the result of artificial constriction of dress", referring to the corsets worn by women of the time (Becklake & Kauffmann, 1999).

There was a surge in publications during the 20th century, particularly from the 1970s onwards, with the development of population-based analyses on increasingly larger sample sizes. These publications all presented segregated male-female data, sometimes appearing on the same graphs without statistical comparison (see typical example below) or in tables in the form of means, and occasionally accompanied by standard deviations (Knudson et al., 1976; Knudson et al., 1983).

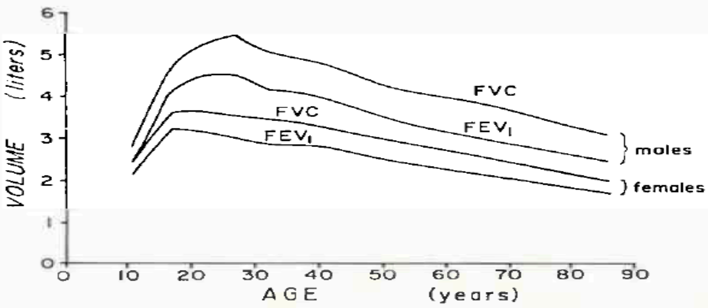

Fig. 1. The changes in 1-sec forced expiratory volume (FEV_1) and forced vital capacity (FVC) with age in normal male and female subjects. Only mean values are shown.

Figure 1. *Knudson et al., 1976, permission request submitted.*

The above data indicate a mean difference of around 30% between men and women. A 1980s review of the state of the art on "normality" applied to the measurement of pulmonary function, estimated that sex accounts for an "inter-subject variation" of 30%, compared to 20% for height, 8% for age, 2% for weight and 10% for race[3] (Becklake, 1986). No explanation was given for the 30% figure for sex. Other studies have established mean differences in vital capacity between sexes of about 15-20% (Cotes et al., 2006), and some also conclude that "males tended to outperform women with the same anthropometric characteristics" (Schwartz et al., 1988). However, a subsequent study indicated that height explained approximately 40% of the male/female variance (Jacobs Jr. et al., 1992), and more recent research reports that a 1% increase in height is associated with a 2.5% increase in spirometry results (Stanojevic et al., 2008).

Meanwhile, the increasing accessibility of computing resources led to the development of new statistical models. These were all based on male/female categorization and considered the two sub-populations as distinct, meaning that the prediction equation parameters were calculated independently in the male and female cohorts. However, as shown in the table below using raw data from three original articles, the intersection between the two distributions is far from vacant, demonstrating an overlap between male and female distributions ranging from 23% to 46%.

---

[3] Used in this document as a sociological (not biological) category.

Table 1. Overlapping distributions of Forced Vital Capacity (FVC) and Forced Expiratory Volume in 1 second (FEV1)[4], in males (M) and females (F).

Study (# in Appendix)	Tested criterion	Overlap
Knudson, 1976 (#3)	FVC (M/F)	38%
	FEV1 (M/F)	44%
Knudson, 1983 (#4)	FVC (M/F)	42%
	FEV1 (M/F)	46%
Jacobs, 1992 (#6)	FVC (white M/F)	23%
	FEV1 (white M/F)	28%
	FVC (black M/F)	29%
	FEV1 (black M/F)	34%

Data calculated with https://sexdifference.org/ (Maney, 2016).

In summary, raw male/female pulmonary function data are scarce in the literature, as authors quickly focused on establishing and validating prediction equations. These data were typically presented as means, with men/women differences of approximately 30% — yet they rarely took height into account, which is a major factor of variation. Analysis of distributions based on published data shows a significant overlap between male/female categories in the main parameters measured by spirometry.

These data served as the basis for the development of sex-specific prediction equations as early as the 1970s.

3. The Male/Female Binary Categorization in Prediction Models

When setting up a standardized process for interpreting pulmonary function tests, four key choices govern the definition of a "significant" abnormality that guide clinical decision-making (Harber et al., 1983): (1) the choice of a prediction equation, (2) the selection of variables to be included in the statistical model for predictions (in this case: age, height, sex, race), (3) the method for comparing observed values with predicted values, called "reference" values (for example, the percentage of predicted values or z-score, allowing for the standard deviation of predicted values), and (4) the abnormality thresholds chosen to define a potential anomaly. While determining an anomaly involves factors beyond mere physiological deficiency, these statistical decisions can significantly affect whether an individual is classified as "normal" or "abnormal", as well as the severity of pulmonary function deficit. The issues raised

[4] The Forced Expiratory Volume in 1 second (FEV_1) is the volume of air (in liters) exhaled in the first second during forced exhalation after maximal inspiration.

by points (3) and (4) are beyond the scope of this article and will therefore not be addressed.

Since the development of the spirometer, more than 400 equations have been proposed to establish predicted values (Pellegrino et al., 2005; Cooper et al., 2017). In 2012, an international working group known as the Global Lung Initiative (GLI) established a consensus based on 160 000 data points from "healthy" individuals across 72 centers in 33 countries, following a statistical methodology reference from the World Health Organization (Quanjer et al., 2012). The GLI reference model is considered applicable globally and has been validated in numerous countries, including the French ELISABET cohort (Hulo et al., 2016). It is now integrated into the software of pulmonary function measurement devices and is used daily and automatically in the Western world by most pulmonologists or corresponding hospital departments (Stanojevic et al., 2022; Günther et al., 2023).

Specifically, the GLI reference model is a linear equation defined as follows:

$$y = a + b.H + c.A + \text{age-spline} + d.G$$

where *y* is the predicted variable of pulmonary function (Forced Vital Capacity [FVC] or Forced Expiratory Volume in 1 second [FEV1] for the main variables), H is the patient's height, A is his/her age, and G is his/her ethnic group. The coefficients *a*, *b*, *c*, and *d* are values estimated by an algorithm from observations obtained on a large population. The age-spline coefficient is a nonlinear component of the model based on age, used to improve the statistical quality of prediction.

The "sex" variable is absent from the equation, but the male/female binary categorization plays a fundamental role in the calculations. Indeed, the coefficients *a*, *b*, *c*, *d,* and the age-spline are calculated independently for a population of women as opposed to a population of men. In other words, the equations used in clinical practice for women are not the same as those used for men.

The foundations of the GLI model date back to the early data and equations from the 1970s. Based on the observations of Knudson et al. presented in Figure 1, separate prediction equations were then developed, incorporating age and height as relevant independent variables (Knudson et al., 1976; Knudson et al., 1983). Since that time, it has been recommended to establish prediction equations separately for men and women (Cotes et al., 2006). All subsequent work relied on this binary categorization without providing any argumentation.

4. "Sex" and "Gender" in Scientific Discourse

This section outlines the findings from our exploratory analysis. We begin by presenting basic statistics related to the usage of the selected keywords. These statistics confirm the significant dominance of the term "sex", and our contextual analysis highlights the absence of formal definitions, contrary to other categories (e.g., height and age) introduced into the models. We then focus on a small subset of articles from the corpus wherein the "gender" concept was introduced, either explicitly or not. Lastly, we discuss the essentialization of men/women differences.

4.1. *Terminology Analysis*

The details of the sixty publications of our corpus are provided in the Appendix. Our analysis of the original articles reveals a systematic use of the term "sex", while "gender" is much more uncommon (mentioned in only six articles). Almost all articles lack supporting definitions for these terms. Only three literature reviews dealing explicitly with sex and gender issues offer any explanation. Moreover, four articles use the terms "sex" and "gender" interchangeably, likely due to the authors' assumed equivalence of the two. A few articles use the expression "sexual dimorphism". This is scientifically inappropriate, however, as it refers to comparisons of anatomical features present in both sexes, such as lung volumes or bronchial size. Indeed, the term "sexual dimorphism" traditionally describes two distinct and non-overlapping traits of males and females within the same species. The use of this term thus reinforces the binary male/female categorization at the expense of studying "sex differences", which encompasses the full spectrum of characteristics or traits present in both men and women, and is by far the most common scenario (Pape et al., 2024).

It is also worth noting that the studies validating the GLI reference prediction equation (Quanjer et al., 2012) across different geographical zones of heterogeneous granularity (including North Africa, Sweden, Finland, Norway, Algeria, Asia, France, Europe, Iran, and Belgium) do not support the hypothesis of differentiation through sex and gender terminology, as the term usage is similar.

With regard to data acquisition, while precise methods for measuring weight, age and height are always specified, only one study briefly explains via an "administered questionnaire" how to determine "sex" or "gender". Recent recommendations from medical societies on the standardization of pulmonary function measurement indicate the need to collect "birth sex" and "gender identity" in their appendices, but lack specific guidance on how these data are to be collected and used (Graham et al., 2019). A more recent recommendation (2022), however, indicates the need to "clarify that biological sex, not gender be used to interpret lung function" (Stanojevic et al., 2022). This wording suggests a focus on sex alone in the interpretation of pulmonary function results, thus entirely neglecting gendered practices.

4.2. *The Role of Gender in Explaining Differences between Men and Women*

The lack of consideration of the role of gender in the development of statistical models is primarily observed in the explanatory factors related to individual variations, with a few exceptions identified in our literature review.

In general, most authors attribute men/women differences in pulmonary function to sex differences, arguing notably that "there is no doubt" that from the fetal stage, sex hormones have a profound influence on lung development and the etiology of respiratory disease (Townsend et al., 2012). Other authors point to differences in body composition, notably the proportion of fat, to account for many of these differences (Cotes et al., 2001). The same authors concluded in 2006 that, on average, women's pulmonary function is inferior to that of men, and that this difference stems from "gender-related" differences in body size and composition, the age at which growth ceases, as well as the "somewhat different" lifestyles of the two sexes (Cotes et al., 2006). It is worth noting that abundant literature shows that many of the differences in body composition, such as muscle and fat mass, can be explained by differences in physical activity, with women having lower physical activity levels than men from childhood onwards (Vilhjalmsson & Kristjansdottir, 2003; Azevedo et al., 2007). Other assertions include statements such as "men are taller than women and therefore have larger lungs" (Molgat-Seon et al., 2018), or " males have larger absolute lung volumes than females even when matched for height" (Dominelli & Molgat-Seon, 2022). However, it has been shown that ratios of mobilizable lung volumes to total volumes are similar between men and women, and the authors of the 2012 GLI reference model did not observe this "lung volume" effect when they took height into account.

Moreover, although the term "gender" was not explicitly used in her study, factors related to gendered habits, such as physical activity, musculature, and socio-professional activity, were identified as early as 1986 by epidemiologist Margaret Becklake (Becklake, 1986). She introduced the term[5] a decade later, stating that "rather than 'normalizing' by sex, the differential influence of sex and gender on airway behavior should always be taken into account" (Becklake & Kauffmann, 1999). A few years later, a literature review noted that it is difficult to separate the influence of sex from that of gender in the context of a disease such as asthma (Townsend et al., 2012). This cohabitation of "sex" and "gender" is reflected more recently in a series of articles entitled "Sex and gender in lung disease" (Dominelli & Molgat-Seon, 2022). However, the role of "sex" as the explanatory factor remains predominant in their work. In particular, anatomical differences are attributed to sex, and the authors assert that "it is reasonable to surmise that gender has little, if any, influence on physiological factors related to the pulmonary response to exercise, but likely impacts psychological or perceptual variables". They nonetheless express concern that gender is not "sufficiently studied" (Dominelli & Molgat-Seon, 2022).

4.3. *The Essentialization of Men/Women Differences*

The main conclusion of this first exploratory analysis in the field of pulmonary function is that male/female categorization, while playing a major role in both prediction models and clinical applications, is seldom explicitly stated or discussed either in the original studies and recommendations from medical societies, or in

[5] She wrote that “we found it useful to distinguish between biological determinants implicit in the word ‘sex’ (in other words, determinants related to the organs necessary for reproduction) and sociocultural determinants implicit in the word ‘gender’ (a broader term which includes sociocultural and environmental as well as biological determinants of airway behaviour)”. In Becklake, M. R., & Kauffmann, F. (1999). Gender differences in airway behaviour over the human life span. *Thorax, 54*(12), 1119-1138.

	North African spirometry." Respiratory medicine (2013), 107(12): 2000-2008.		
16	Backman, Helena, et al. "Evaluation of the global lung function initiative 2012 reference values for spirometry in a Swedish population sample." BMC pulmonary medicine (2015), 15(1): 1-8.	OA	Sex, men, women
17	Hulo, Sébastien, et al. "Global Lung Function Initiative reference equations better describe a middle-aged, healthy French population than the European Community for Steel and Coal values." European Respiratory Journal (2016), 48(6): 1779-1781.	OA	Sex, men, women
18	Hüls, Anke, et al. "Applicability of the global lung initiative 2012 reference values for spirometry for longitudinal data of elderly women." PloS one (2016), 11(6): e0157569.	OA	Women
19	Kainu, Annette, et al. "Reference values of spirometry for Finnish adults." Clinical physiology and functional imaging (2016), 36(5): 346-358.	OA	Sex, gender, men, women
20	Langhammer, Arnulf, et al. "Global Lung Function Initiative 2012 reference equations for spirometry in the Norwegian population." European Respiratory Journal (2016), 48(6): 1602-1611.	OA	Sex, male, female
21	Stanojevic, Sanja, et al. "Official ERS technical standards: Global Lung Function Initiative reference values for the carbon monoxide transfer factor for Caucasians." European Respiratory Journal (2017), 50(3) : 1700010.	OA	Sex, female, male
22	Ketfi, Abdelbassat, et al. "The multi-ethnic global lung initiative 2012 (GLI-2012) norms reflect contemporary adult's Algerian spirometry." PLoS One (2018), 13(9): e0203023.	OA	Sex, women, men, female, male,
23	Zhang, Jingzhou, et al. "Global lung function initiative 2012 reference values for spirometry in Asian Americans." BMC pulmonary medicine (2018), 18(1): 1-9.	OA	Sex, women, men
24	Molgat-Seon, Yannick, et al. "The effects of age and sex on mechanical ventilatory constraint and dyspnea during exercise in healthy humans." Journal of Applied Physiology (2018), 124(4): 1092-1106.	OA	Sex, biological sex, men, wo men
25	Abdullah, Noraidatulakma, et al. "Global Lung Initiative 2012 spirometry reference values in a large Asian cohort of Malay, Chinese and Indian ancestry." Respirology (2018), 23(12): 1173-1179.	OA	Gender, male, female
26	Ekström, Magnus, et al. "Absolute lung size and the sex difference in breathlessness in the general population." PLoS One (2018), 13(1): e0190876.	OA	Sex, women, men, sex difference
27	Haynes, Jeffrey M., and Ralph W. Stumbo. "The impact of using non-birth sex on the interpretation of spirometry data in subjects with air-flow obstruction." Respiratory Care (2018), 63(2): 215-218.	OA	Sex, birth sex, non-birth sex, gender, gender iden tity, female, male
28	de Broucker, Virginie, et al. "Les valeurs de référence GLI-2012 pour la spirométrie forcée décrivent fidèlement la population européenne et française." Revue des Maladies Respiratoires 36.3 (2019): 287-290.	OA	Sexe, femmes, hommes

29	Hall, Graham L., et al. "Official ERS technical standard: Global Lung Function Initiative reference values for static lung volumes in individuals of European ancestry." European Respiratory Journal (2021), 57(3): 2000289.	OA	Sex, female, male
30	Sahebi, Leyla, et al. "Evaluation of the Global Lung Function Initiative 2012 reference values for spirometry in an Iranian population." Scientific Reports (2022), 12(1): 12784.	OA	Sex, gen der, female, male
31	De Soomer, Kevin, et al. "Evaluation of the Global Lung Function Initiative reference equations in Belgian adults." ERJ Open Research (2022), 8(2): 00671-2021.	OA	Sex, female, male,
32	Bhammar, Dharini M., et al. "Sex differences in the ventilatory responses to exercise in mild to moderate obesity." Experimental Physiology 107.8 (2022): 965-977.	OA	Sex, male, female, sex differences
33	Ekström, Magnus, and David Mannino. "Race-adjusted lung function increases inequities in diagnosis and prognosis and should be abandoned." medRxiv (2022): 2022-01.	OA	Sex
34	Baugh, Aaron D., et al. "Reconsidering the utility of race-specific lung function prediction equations." American Journal of Respiratory and Critical Care Medicine (2022), 205(7): 819-829.	OA	Sex
35	Bowerman, Cole, et al. "A race-neutral approach to the interpretation of lung function measurements." American Journal of Respiratory and Critical Care Medicine (2023), 207(6): 768-774.	OA	Sex, sex-specific, sex-stratified, females, males
36	Moffett, Alexander T., et al. "Global, race-neutral reference equations and pulmonary function test interpretation." JAMA Network Open (2023), 6(6): e2316174-e2316174.	OA	Sex, wo men, men, female, male
37	Kanj, Amjad N., et al. "Application of Global Lung Function Initiative Global Spirometry Reference Equations across a Large, Multicenter Pulmonary Function Lab Population." American Journal of Respiratory and Critical Care Medicine (2024), 209(1): 83-90.	OA	Sex, female, pa tients identi fied as female at birth
38	Becklake, Margaret R. Concepts of normality applied to the measurement of lung function. Am J Med. (1986), 80(6):1158-64.	LR	Sex
39	American Thoracic Society. "Lung function testing: selection of reference values and interpretative strategies." American Review of Respiratory Disease (1991), 144: 1202-1218.	MSR	Sex, men, wo men
40	Quanjer, Philip H., et al. Lung volumes and forced ventilatory flows. European respiratory journal. (1993), 6(16):5-40.	MSR	Sex, men, women
41	Crapo, Robert O. Pulmonary-function testing. New England Journal of Medicine. (1994), 331(1):25-30.	LR	*(no term)*
42	Milic-Emili, Joseph, Luigi Marazzini, and Edgardo D'Angelo. "150 years of blowing: since John Hutchinson." Canadian Respiratory Journal (1997), 4: 239-245.	LR	*(no term)*
43	Becklake, Margaret R., and Francine Kauffmann. "Gender differences in airway behaviour over the human life span." Thorax (1999), 54(12): 1119-1138.	LR	Sex, gen der, gender differences, sex differences, female, male

44	Pellegrino ,Riccardo, et al. Interpretative strategies for lung function tests. European Respiratory Journal. (2005), 26(5):948-68.	LR	Sex, female, male
45	Cotes, John E., David J. Chinn, and Martin R. Miller. Lung function: physiology, measurement and application in medicine. 6th ed. Oxford: Blackwell Scientific Publications, 2006	Text-book	Sex, gender, male, female
46	Stanojevic, Sanja A., et al. "Reference values for lung function: past, present and future." European Respiratory Journal (2010), 36(1): 12-19.	LR	Sex, males females
47	Culver, Bruce H., et al. "Recommendations for a standardized pulmonary function report. An official American Thoracic Society technical statement." American Journal of Respiratory and Critical Care Medicine (2017), 196(11): 1463-1472.	MSR	Sex, women, men
48	Cooper, Brendan G., et al. "The Global Lung Function Initiative (GLI) Network: bringing the world's respiratory reference values together." Breathe (2017), 13(3): e56-e64.	LR	Sex, male, female
49	Graham, Brian L., et al. "Standardization of spirometry 2019 update. An official American thoracic society and European respiratory society technical statement." American Journal of Respiratory and Critical Care Medicine (2019), 200(8): e70-e88.	MSR	Sex, birth sex, gen der, gender identity
50	Stanojevic, Sanja, et al. "ERS/ATS technical standard on interpretive strategies for routine lung function tests." European Respiratory Journal (2022), 60(1): 2101499.	MSR	Sex, biological sex, gen der, fe males, males, women, men
51	LoMauro, Antonella, and Andrea Aliverti. "Sex differences in respiratory function." Breathe (2018), 14(2): 131-140.	LR	Sex, male, female, men, women, sexual dimorphism sex-related differences
52	Molgat-Seon, Yannick, et al. "Sex-differences in the human respiratory system and their impact on resting pulmonary function and the integrative response to exercise." Current Opinion in Physiology (2018), 6: 21-27.	LR	Sex, sex differences, sexual dimorphism , females, males, wo men, men
53	Sheel, A. William, and Jordan A. Guenette. "Mechanics of breathing during exercise in men and women: sex versus body size differences?." Exercise and sport sciences reviews (2008), 36(3): 128-134.	LR	Sex, female, male, wo men, men, sex differences
54	Ansdell, Paul, et al. "Physiological sex differences affect the integrative response to exercise: acute and chronic implications." Experimental physiology (2020), 105(12): 2007-2021.	LR	Sex, sexual dimorphism sex differences, male,

			female, men, wo men
55	Dominelli, Paolo B., and Yannick Molgat-Seon. "Sex, gender and the pulmonary physiology of exercise." European Respiratory Review (2022), 31(163): 210074.	LR	Sex, gen der, sex differences, male, female, men, wo men
56	Marciniuk, Darcy D., et al. "Effect of Race and Ethnicity on Pulmonary Function Testing Interpretation." Chest (2023), 164(2): 461-475.	MSR	Sex, wo men, men
57	Bhakta, Nirav R., et al. "Race and ethnicity in pulmonary function test interpretation: an official American Thoracic Society Statement." American Journal of Respiratory and Critical Care Medicine (2023), 207(8): 978-995.	MSR	Sex, gen der, female, male
58	Townsend, Elizabeth A., et al. "Sex differences and sex steroids in lung health and disease." Endocrine reviews (2012), 33(1): 1-47.	LR	Sex, sex differences, gender,
59	Maney, Donna L. "Perils and pitfalls of reporting sex differences." Philosophical Transactions of the Royal Society B: Biological Sciences (2016), 371(1688): 20150119.	LR	Sex, gender, sex differences, dimorphism female, male, men, women
60	Rippon, Gina, et al. " Impression management in sex and gender neuroscience research reporting: the MAGIC guidelines." Nature Communications (2024), 15(1): 2826.	LR	Sex, gen der, sex differences, gender differences, female, male, men, women

Type of article: OA (original article), LR (literature review), MSR (medical society recommendation).
The articles were classified into two categories (original article/literature review or recommendations from medical societies), and chronologically within each category, to which were added the last three publications (a reference review of the literature on sex differences in lung health and disease, and two reference reviews criticizing the use of sex as a biological variable in biomedical research).