



# Sex, breathing and statistics

Samy Suissa<sup>1,2</sup>, Anne V. Gonzalez<sup>3</sup> and Ian Shrier<sup>1</sup>

**Affiliations:** <sup>1</sup>Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, QC, Canada. <sup>2</sup>Dept of Epidemiology and Biostatistics, McGill University, Montreal, QC, Canada. <sup>3</sup>Respiratory Epidemiology and Clinical Research Unit, Montreal Chest Institute, McGill University Health Centre, Montreal, QC, Canada.

**Correspondence:** Samy Suissa, Dept of Epidemiology and Biostatistics, McGill University, Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, 3755 Côte Sainte-Catherine, Montreal, Quebec, H3T 1E2, Canada. E-mail: samy.suissa@mcgill.ca

 @ERSpublications

**Women are more likely to report mMRC breathlessness and this appears to be related to smaller lung volume** <http://ow.ly/hoDM30ayQWH>

**Cite this article as:** Suissa S, Gonzalez AV, Shrier I. Sex, breathing and statistics. *Eur Respir J* 2017; 49: 1700455 [<https://doi.org/10.1183/13993003.00455-2017>].

Breathlessness is a complex symptom that results from the interplay of physiological, psychological, social and environmental factors [1]. Breathlessness is also common, and is associated with increased morbidity and mortality [2, 3]. A number of studies have shown that breathlessness is reported more frequently by women than men, both from population-based surveys and in patients with chronic obstructive pulmonary disease [4].

The debate as to whether to use absolute forced expiratory volume in 1 s (FEV<sub>1</sub>) *versus* % predicted FEV<sub>1</sub> for statistical adjustment arose over 20 years ago, when attempts were made to explain sex differences in airway hyperresponsiveness (AHR). In the Lung Health Study, a greater prevalence of AHR was noted in women, but when baseline absolute FEV<sub>1</sub> was added to a model predicting AHR, the effect of female sex was eliminated [5]. The authors concluded that the higher prevalence of AHR in women was related to their smaller airway calibre. A French population-based survey of 20–44-year-old subjects also reported a higher prevalence of AHR to methacholine among women [6]. This sex difference, which was no longer significant when adjusted only for absolute FEV<sub>1</sub>, remained after additional adjustment for % predicted FEV<sub>1</sub>. The authors concluded that airway size was not the explanation for the increased AHR found in women. In a survey of rural primary-care practices, more women than men who had ever smoked reported breathlessness, while women with mild-to-moderate airflow obstruction were twice as likely as their male counterparts to be prescribed respiratory medications [7].

In this issue of the *European Respiratory Journal*, EKSTRÖM *et al.* [8] report on an observational study of the association between sex and activity-related breathlessness in the population. This study involved a large, cross-sectional, general-population sample of 3250 people, including 1673 women and 1577 men, to evaluate the relationship between sex and activity-related breathlessness. The subjects provided various data, including measures of lung function by spirometry and of the modified Medical Research Council (mMRC) dyspnoea scale, to estimate the extent of breathlessness in response to different levels of activity. In a direct comparison, women were observed to have more activity-related breathlessness (27% with mMRC >0 *versus* 14% for men) resulting in an odds ratio of 2.21 (95% CI 1.79–2.72). Moreover, women had lower mean FEV<sub>1</sub> (2.7 L) than men (3.7 L). However, the mean relative FEV<sub>1</sub> measured as a

---

Received: March 03 2017 | Accepted after revision: March 29 2017

Support statement: S. Suissa is the recipient of the James McGill Chair award. A.V. Gonzalez is the recipient of an FRQS Chercheur-Boursier Clinicien award. Funding information for this article has been deposited with the Crossref Funder Registry.

Conflict of interest: None declared.

Copyright ©ERS 2017

percentage of the population-based predicted value according to age, height and race within each sex was 99%, equally for men and women. The two sexes were similar for all other factors. Thus, in essence, the women had lower spirometric lung function in absolute value and more activity-related breathlessness than the men, but were similar in all other characteristics, including the % predicted FEV<sub>1</sub>.

The question posed by EKSTRÖM *et al.* [8] is whether the two-fold increase in the rate of activity-related breathlessness among women relative to men should be statistically adjusted using the absolute lung function value (FEV<sub>1</sub> in litres) or using the relative lung function value (% predicted FEV<sub>1</sub>), the latter adjustment being the norm used in previous studies. The authors did it both ways and found that after adjusting for % predicted FEV<sub>1</sub>, the odds ratio of breathlessness for women relative to men was still elevated at 2.66 (95% CI 2.13–3.34), while adjusting for absolute FEV<sub>1</sub> reduced the odds ratio to 0.89 (95% CI 0.69–1.14), thereby essentially eliminating the differences in breathlessness between the sexes. Which of these two adjusted results should we rely on?

Before we attempt to answer this question, it is useful to understand what statistical adjustments actually accomplish and provide. To illustrate the problem, we will assume that activity-related breathlessness will be increased if absolute lung function (absolute FEV<sub>1</sub>) is decreased, regardless of sex, age or height (main adjustments in predicted values). Figure 1 depicts the relationship between absolute FEV<sub>1</sub> and the degree of activity-related breathlessness, showing that as FEV<sub>1</sub> increases, the extent of breathlessness decreases. For pedagogical reasons, we indicated a linear relationship but the principles apply equally to nonlinear relationships. Figure 1 also maps the range of FEV<sub>1</sub> among men and women, roughly using the values from the study by EKSTRÖM *et al.* [8], namely the mean FEV<sub>1</sub> values of 2.7 and 3.7 L for women and men, respectively, with hypothetical scatters depicted by the ovals. It shows the corresponding values of breathlessness for women and men at these mean FEV<sub>1</sub> values, indicating the greater degree of breathlessness for women than for men, exemplifying the study data.

In regression analysis, “adjustment” for FEV<sub>1</sub> is analogous to comparing breathlessness in males with a particular FEV<sub>1</sub> value to breathlessness in females with the same FEV<sub>1</sub> value. In figure 1, we can see that for any given FEV<sub>1</sub> value, males and females will have exactly the same amount of breathlessness, *i.e.* the sex differences in breathlessness are eliminated by adjusting for the absolute FEV<sub>1</sub>.

Figure 2 illustrates the analysis adjusted for the relative FEV<sub>1</sub> value (% predicted). It shows that the horizontal FEV<sub>1</sub> axis has been standardised to a mean 100% predicted FEV<sub>1</sub> value for both men and women. The figure shows that shifting the FEV<sub>1</sub> axis to a common scale for men and women, centred at 100%, does not affect the difference in the degree of breathlessness between men and women. In essence, this figure shows that standardising the FEV<sub>1</sub> values within each sex will not affect differences in breathlessness between men and women, specifically because the % predicted values are computed separately for men and women. Thus, the analysis adjusted for % predicted FEV<sub>1</sub> will be equivalent to the crude analysis that simply compares the degree of breathlessness between men and women.

Another version of this illustration is provided in tables 1 and 2. These tables report a hypothetical cross-sectional study of 1000 randomly selected women and 1000 randomly selected men. The crude rate ratio of breathlessness comparing women to men is shown to be 2.0, which changes to 1.0 after adjustment for the absolute FEV<sub>1</sub> categorised as low or high. The data are consistent with women being more likely to have lower absolute FEV<sub>1</sub>. Table 2 shows the data after adjustment for relative FEV<sub>1</sub> where, by definition, half the women and half the men will have lower relative FEV<sub>1</sub> (<100% predicted), while the other half will

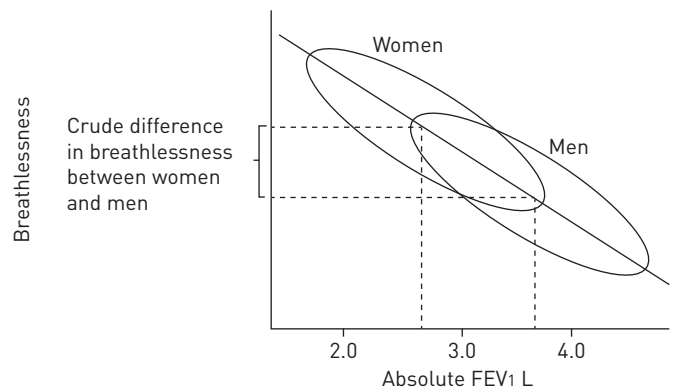


FIGURE 1 Linear relationship between forced expiratory volume in 1 s (FEV<sub>1</sub>) [in litres] and the degree of activity-related breathlessness for women and men, and the associated difference.

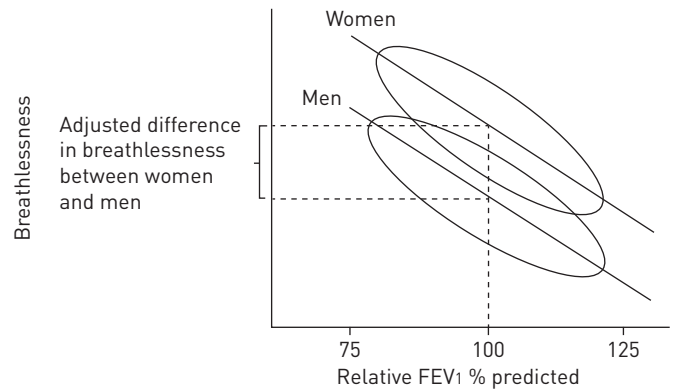


FIGURE 2 Linear relationship between relative forced expiratory volume in 1 s (FEV<sub>1</sub>) (% predicted) and the degree of activity-related breathlessness for women and men.

have higher relative FEV<sub>1</sub> (>100% predicted). When adjusting for relative FEV<sub>1</sub>, the adjusted rate ratio of breathlessness comparing women to men is 2.0, which is unchanged from the crude value.

Thus, since adjusting for relative lung function measures (that are sex specific) is practically equivalent to the crude comparison of breathlessness between men and women, the question posed by EKSTRÖM *et al.* [8] could be rephrased as “should breathlessness in men and women be compared directly without any adjustment or by adjusting for the absolute lung function measures?”; that is, crude or adjusted?

The answer depends on the nature of the question and on the relationships between these variables. In the case of activity-related breathlessness, the theory is that lungs with greater ventilatory capacity will be associated with less breathlessness. In the semantics of causal inference, sex (the upstream variable) causes absolute lung function (the mediator, through body size or other mechanisms), which causes breathlessness (the outcome). Including the mediator (absolute lung function) in a statistical model to “adjust for differences in absolute lung function” will eliminate the effect of the upstream variable on the breathlessness outcome. However, sex is not associated with the relative lung function specifically because it is standardised to the general population within sex, so that including the relative lung function variable in the model will not change the relationship between sex and the breathlessness outcome. Therefore, if one wants to evaluate whether sex has an effect on breathlessness apart from its effect through lung function, then the adjusted analysis is more appropriate, but if one wants to evaluate the total effect of sex including that mediated through lung function, then the crude analysis is more appropriate.

A related issue in this debate is the quantification of “breathlessness” using the mMRC dyspnoea scale, which has been used widely for research purposes since its development in 1959 [9]. Rather than measuring the degree of breathlessness *per se*, the scale measures the degree of disability that results from such breathlessness. Despite its well established validity and prognostic value, the various levels of the

TABLE 1 Rate ratio of breathlessness comparing women to men, crude and adjusted for the absolute forced expiratory volume in 1 s (FEV<sub>1</sub>), in a cross-sectional study of 1000 randomly selected women and 1000 randomly selected men

	Subjects	Breathless subjects	Rate of breathlessness	Crude rate ratio	Adjusted <sup>#</sup> rate ratio
<b>All subjects (crude)</b>					
Women	1000	300	30%	2.0	
Men	1000	150	15%	1.0 (ref.)	
<b>Low absolute FEV<sub>1</sub></b>					
Women	800	280	35%		1.0
Men	200	70	35%		1.0 (ref.)
<b>High absolute FEV<sub>1</sub></b>					
Women	200	20	10%		1.0
Men	800	80	10%		1.0 (ref.)
<b>All subjects (adjusted)</b>					
Women	1000				1.0
Men	1000				1.0 (ref.)

<sup>#</sup>: for absolute FEV<sub>1</sub> (low and high).

TABLE 2 Rate ratio of breathlessness comparing women to men, crude and adjusted for relative forced expiratory volume in 1 s (FEV<sub>1</sub>) (% predicted) in a cross-sectional study of 1000 randomly selected women and 1000 randomly selected men

	Subjects	Breathless subjects	Rate of breathlessness	Crude rate ratio	Adjusted <sup>#</sup> rate ratio
<b>All subjects (crude)</b>					
Women	1000	300	30%	2.0	
Men	1000	150	15%	1.0 (ref.)	
<b>Low relative FEV<sub>1</sub></b>					
Women	500	180	36%		2.0
Men	500	90	18%		1.0 (ref.)
<b>High relative FEV<sub>1</sub></b>					
Women	500	120	24%		2.0
Men	500	60	12%		1.0 (ref.)
<b>All subjects (adjusted)</b>					
Women	1000				2.0
Men	1000				1.0 (ref.)

<sup>#</sup>: for relative FEV<sub>1</sub> (low and high).

mMRC scale leave some room for interpretation as it is conceivable that men and women may respond differently to the mMRC dyspnoea scale, which is a question to be examined.

Population-standardised values such as % predicted FEV<sub>1</sub> remain important for other types of questions where one is not interested in mediation. For example, say one is intending to conduct an observational study of the effect of a treatment where lung function is an important indication for the treatment under study. To control for confounding, we would want to use the version of the lung function variable that is most closely related to the indication of treatment or the prognosis of the disease. In general, clinicians use the relative lung function (% predicted FEV<sub>1</sub>) to determine if “disease” is present and to prescribe treatment. Therefore, in such a case, the % predicted FEV<sub>1</sub> would often be the more appropriate adjustment or matching factor.

Finally, while it is common to examine the mechanism for sex differences, there is an ongoing debate in philosophy of science as to what it means to say sex “causes” an outcome such as breathlessness [10, 11]. Indeed, the philosophy is important because effective prevention and treatment of disease requires that our interventions modify the causes of the outcome. However, sex cannot be manipulated like an intervention or a medication, and cannot thus be subjected to randomisation, the tool of choice in causal inference. Therefore, what does it mean to say sex causes an outcome? In the case of breathlessness, the theory may be that women are smaller than men, and therefore have smaller lungs. So what? Why adjust for the size of the lungs? The crude difference in breathlessness between the sexes can be an important clinical observation in itself. Indeed, examining sex differences is relevant to increase awareness and potentially impact the approach to therapy. If sex differences in prevalence of dyspnoea are not recognised, women may be more likely to receive inappropriate respiratory medications, increasing the possibility of adverse events without definite efficacy. As women may be at greater risk of certain side effects (*e.g.* osteoporosis and fractures with inhaled corticosteroids), the question of sex differences in breathlessness becomes one of public health relevance.

In all, are more women breathless (*versus* are women more breathless) simply because they have smaller lungs? This may be an oversimplification, given the complexity of dyspnoea as a symptom. We can conclude from the study of EKSTRÖM *et al.* [8] that women are more likely than men to report activity-related breathlessness as measured by the mMRC scale and in this study, with these methods, this appears to be related to their smaller lung volumes and, thus, lesser ventilatory capacity. We agree with the authors that both relative and absolute lung volumes should be considered in the evaluation of lung function impairment and remaining ventilatory reserve.

## 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: 435–452.
- 2 Jones PW, Adamek L, Nadeau G, *et al.* Comparisons of health status scores with MRC grades in COPD: implications for the GOLD 2011 classification. *Eur Respir J* 2013; 42: 647–654.
- 3 Figarska SM, Boezen HM, Vonk JM. Dyspnea severity, changes in dyspnea status and mortality in the general population: the Vlagtwedde/Vlaardingen study. *Eur J Epidemiol* 2012; 27: 867–876.
- 4 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: 1034–1041.

- 5 Kanner RE, Connett JE, Altose MD, *et al.* Gender difference in airway hyperresponsiveness in smokers with mild COPD. The Lung Health Study. *Am J Respir Crit Care Med* 1994; 150: 956–961.
- 6 Leynaert B, Bousquet J, Henry C, *et al.* Is bronchial hyperresponsiveness more frequent in women than in men? A population-based study. *Am J Respir Crit Care Med* 1997; 156: 1413–1420.
- 7 Dales RE, Mehdizadeh A, Aaron SD, *et al.* Sex differences in the clinical presentation and management of airflow obstruction. *Eur Respir J* 2006; 28: 319–322.
- 8 Ekström M, Schiöler L, Grønseth R, *et al.* Absolute values of lung function explain the sex difference in breathlessness in the general population. *Eur Respir J* 2017; 49: 1602047.
- 9 Stenton C. The MRC breathlessness scale. *Occup Med* 2007; 58: 226–227.
- 10 Glymour C, Glymour MR. Commentary: race and sex are causes. *Epidemiology* 2014; 25: 488–490.
- 11 Vandenbroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology: the need for a pluralistic approach. *Int J Epidemiol* 2016; 45: 1776–1786.