Finally, is the non-proportional Venn diagram really a useful way of illustrating the components of COPD? Often the components of COPD are described as alveolar destruction (emphysema), small airway fibrosis (bronchiolitis) and chronic bronchitis. As most of the airflow obstruction is due to reduced airflow in the small conducting airways, it seems a little odd not to have small airways disease in the diagram. Should we then start working with an updated Venn diagram? Probably not. Perhaps the problem is that the Venn diagram should never have been taken too literally. It can stimulate thought and illustrate the heterogeneity of COPD; however, the effect of adding numbers to the diagram should motivate us to move on and try to gain new insights into this disease by coming up with novel concepts and potentially new subgroups of COPD which we had no idea ever existed.

**Competing interests:** None.

**REFERENCES**

**What evidence could validate the definition of COPD?**

Roberto de Marco

Chronic obstructive pulmonary disease (COPD) is one of the leading causes of disability and death in both the developed and developing worlds, and it is largely underdiagnosed.

While it is generally accepted that the ratio between forced expiratory volume in 1 s (FEV₁) and forced vital capacity (FVC) or slow vital capacity (SVC) is the most important measure that characterises airflow obstruction, there is still no consensus on what the best definition of COPD should be.

In an attempt to simplify the diagnosis of COPD, improve the detection of the disease in primary care and standardise methods to measure the prevalence of COPD in different countries, the Global Initiative for Chronic Obstructive Lung Disease (GOLD) has defined COPD as a post-bronchodilator FEV₁/FVC ratio <0.70.*

As the FEV₁/FVC ratio decreases with age, this fixed ratio has been criticised as it may overdiagnose the disease in elderly subjects and underdiagnose it in young adults. For this reason, the ATS/ERS guidelines on lung function testing propose using the lower limit of normality (LLN) instead of a fixed cut-off to define COPD. LLN is the lower fifth percentile of the frequency distribution of the FEV₁/FVC ratio in a healthy population of a given sex and age. The guidelines also point out that the SVC could be more accurate than FVC to diagnose airflow obstruction.

A hot debate between the supporters of the two different points of view is currently ongoing, as can be seen in scientific journals. As an epidemiologist I believe there could be sound reasons to support both approaches but, in my opinion, there is insufficient evidence to make one preferable to the other. The positive aspect of the GOLD definition is simplicity and, thanks to this, its use is quite widespread. It relies entirely on the opinions of clinical experts and assumes that an FEV₁/FVC ratio <0.70 is, “on average”, a good threshold to discriminate patients with clinically defined COPD from normal subjects. The LLN definition is in theory more rational; it is based on the statistical approach used to derive the reference (or “normal”) ranges for biological variables. It depends entirely on the accuracy of “reference equations” (to my knowledge, accurate national reference equations do not currently exist in many cases) and it assumes that (1) everyone outside the range is affected by COPD and (2) all patients affected by COPD are outside the range (in other words, the distribution of diseased and healthy people does not overlap).

The trait that the two approaches have in common is that the definition of the disease is based exclusively on a variable—the FEV₁/FVC (or FEV₁/SVC)—that is distributed along a continuum in the general population. Any definition of COPD based on a dichotomy of this
frequency distribution is therefore arbitrary. In a certain sense the diagnosis of COPD based only on spirometry results is similar to that of type 2 diabetes using plasma glucose levels or hypertension using blood pressure measurements.

How were the cut-off levels used to discriminate between diseased and healthy people established in medical fields other than pneumology? Usually, cut-off levels are based on studies that compare the distribution of the test variable between patients who have been clinically diagnosed and healthy subjects. Consequently, the choice of threshold is the one that guarantees the best combination of sensitivity and specificity. Diagnostic cut-off points may also be based on longitudinal studies of the risk of adverse outcomes or complications of the disease.12–14

Neither the GOLD nor the LLN approach relies on these kinds of studies. The validity of each of the two definitions therefore remains a hypothesis until it has been empirically tested by appropriate large-scale validation studies.

In the current issue of Thorax, Bridevaux et al15 present the results of a longitudinal study that highlights the features of one of the spirometric definitions of COPD (see page 768). The authors analysed the data from the Swiss Study on Air Pollution and Lung Diseases in Adults (SAPALDIA), which is an 11-year follow-up study of a large random sample of adult population (aged 18–60 years at the time of enrolment).16 They investigated lung function decline, respiratory care utilisation and health-related quality of life in individuals who were classified at baseline as normal individuals or subjects with COPD according to the GOLD classification. Given the size of the cohort, the age range of its members, the quality of the data and the length of the follow-up, the findings presented not only provide information on the prognostic value of the GOLD definition but they go a step further. The authors found that, overall, the subjects with GOLD-defined COPD had poorer longitudinal outcomes than those with normal lung function. If this was true for the subjects with the most severe GOLD-defined COPD (those with FEV1 <80% predicted), it was not the case for those with less severe disease (GOLD stage 1 COPD: FEV1 ≥80% predicted) who represented more than 85% of all the patients classified as having COPD by the GOLD criterion. In fact, in the latter group the presence or absence of chronic respiratory symptoms at baseline was the actual predictor of the long-term outcome. Subjects with symptomatic GOLD stage 1 COPD had a faster rate of decline in lung function, increased respiratory care utilisation and a lower health-related quality of life than asymptomatic subjects with normal lung function. On the other hand, subjects with asymptomatic GOLD stage 1 COPD had similar long-term outcomes to those of normal asymptomatic subjects.

Other longitudinal studies have stressed the relevance of respiratory symptoms in the natural history of COPD.17–18 However, the findings presented by Bridevaux et al15 suggest that respiratory symptoms should be part of the definition of COPD together with airflow obstruction if the COPD label is to identify subjects with a worse long-term prognosis than those with normal lung function. Is the lack of predictive power an issue concerning the GOLD but not the LLN definition? There is no answer to this question in the paper by Bridevaux et al15 as the authors did not test the longitudinal performance of the LLN definition. However, given the relationships between the two spirometric criteria,1 it is likely that the LLN definition would have performed better in elderly subjects and worse in young subjects than the GOLD definition, with the result that the average predictive power of the two approaches would have been very similar.

In conclusion, the paper by Bridevaux et al15 suggests that definitions of COPD based exclusively on spirometric criteria might have poor prognostic power. Other aspects should be taken into account such as symptoms (or risk factors) in order to increase the validity of the current COPD definitions. Of course, given that the authors did not measure post-bronchodilator lung function and did not test the LLN definition, their findings and suggestions need to be confirmed in future studies.

The debate on the best definition of COPD is likely to last a long time in the absence of “validation studies” such as the one published in this issue of Thorax. Studies on a cross-sectional comparison of the two definitions without any clinical gold standard, or even studies on the distribution of lung function in general population samples, may provide scanty evidence for the choice of a valid diagnostic test for COPD. Perhaps the main lesson we can learn from the paper by Bridevaux et al is a methodological one: the right answer comes from the right study.

Competing interests: None.


REFERENCES