OPINION

Are reference equations for spirometry an appropriate criterion for diagnosing disease and predicting prognosis?

Guy B Marks

ABSTRACT

In the last few years, there has been considerable debate on the use of threshold criteria for the diagnosis of obstructive lung disease based on the spirometric ratio (forced expiratory volume in one second/forced vital capacity (FEV₁/FVC)) and also spirometric volumes (specifically FEV₁). The ‘M’ word (misclassification) has been invoked to criticise one criterion, or set of criteria, with respect to another.1 I argue that this critique is based on a false presumption about the truth (validity) of the proposed criterion for normality. It is underpinned by a limited and limiting framework for understanding and using the information that measurement of spirometric function provides. Unfortunately, the current Global Lungs Initiative,2 which aims to provide the world with predictive variables that are really risk factors for disease or for adverse outcomes of disease. The author argues that this critique is based on a false presumption about the validity of reference equations as a criterion for normality. The flaw lies in the methods used to derive reference equations, which involve arbitrary and circular criteria for exclusion of some members of the population, use potentially non-representative reference populations and include predictive variables that are really risk factors for disease or for adverse outcomes of disease. Regression equations analogous to the Framingham cardiovascular equations for lung function data in clinical practice based on prognostic equations analogous to the Framingham cardiovascular risk factor equations. These interpretative equations should be based on data from cohort studies and randomised controlled trials, rather than cross-sectional studies, and if properly formulated, will prove to be valuable aids to clinical decision making.

In the last few years, there has been considerable debate on the use of threshold criteria for the diagnosis of obstructive lung disease based on spirometric function has a long history, beginning with the first report on a device for measuring vital capacity in 18463 and was greatly advanced 101 years later by Tiffeneau and Pinelli’s description of the timed forced expiratory manoeuvre and derivation of the FEV₁/FVC ratio as an indicator of airflow obstruction.4 The prognostic significance of FEV₁ for chronic obstructive pulmonary disease5 6 and its importance in the diagnosis of asthma7 have been central to our understanding of these diseases for decades. Its broader relevance as a prognostic indicator for cardiovascular outcomes is also well established.8 Spirometric function is a key indicator of health status.

The attempt to interpret the spirometric function of individual patients or subjects with respect to reference values also has a long history. Its beginnings are in the identification of ‘normal’ variation attributable to gender differences demonstrated by Hutchinson himself and extend to racial differences in soldiers of the Union Army during the American Civil War (reviewed by Braun9). Modern reference equations incorporating age, sex, height and race or ethnic origin have been used in North America10 11 and Europe12–14 for several decades. However, there has been little critical analysis of the basis of these reference equations and the claim that they represent the criterion against which the presence or absence of disease, specifically obstructive lung disease, can be ascertained.

The current model for detecting abnormality or disease based on lung function measurements is to compare an individual’s observed values with a reference range. These reference ranges are universally derived from spirometric surveys conducted in apparently representative populations of apparently normal individuals. Regression equations are used to estimate the expected or average value based on selected predictors, usually functions of age, height, sex and, sometimes, race. The expected value and residual variance are then used to define the range of values within which 95% of the reference normal population values would be expected to lie. The lower limit of normal is the lower end of this range. This method has been applied to all spirometric variables including FEV₁, FVC and FEV₁/FVC ratio, as well as peak expiratory flow rate. This method is simple, based on sound statistical principles, and appealing. Some would argue that values within this reference range are defined as normal or not diseased and values outside this range are defined as abnormal and hence diseased. Any diagnostic criteria that result in a different classification of disease or non-disease are regarded as a misclassification.1 This claim requires some analysis.

The basis for the concern about the use of reference equations as the criterion for diagnosing...
Chest clinic

disease lies in the representativeness of the populations on
which it is based, the criterion for defining normal and the
selection of predictive factors. All studies used to derive reference
equations seek to limit their study populations to ‘normal’
individuals. This is usually achieved by a questionnaire-based
selection criteria related to respiratory symptoms, diagnosed
respiratory disease and smoking (a risk factor for respiratory
disease). This raises several questions. If reported diagnoses and
symptoms are an adequate basis for distinguishing normal from
not normal, why do we need to measure lung function at all?
Surely, the main rationale for using a spirometric criterion for
disease is that it is independent of subjective criteria such as
symptomatic status or reported diagnosis. The inclusion of this
criterion in the spirometric definition of ‘normal’ means that the
definition is not independent of subjective factors. In developed
nations, smoking is the main risk factor for obstructive lung
disease. However, is this an adequate justification for excluding
smokers from the reference populations? If so, what about
people with other risk factors for low lung function such as
airway hyper-responsiveness, exposure to biomass fuel smoke or
occupational exposure to dust and fumes? Is there a definitive
reason for the priority given to smoking status as the only risk
exposure to dust and fumes

Airway hyper-responsiveness, exposure to biomass fuel smoke or

For the derivation of the reference equations.11 There is a substantial
reduced to 4634 individuals (28.1% of the original sample) for
and people who could not perform reproducible spirometry, the
population is selected using sound sampling principles, as in the
reference equations is a problem. Even when the original study
population for whom the statements are to apply. Representa-
tions reduces the likelihood that people with those risk factors
can be diagnosed with disease. The question of appropriate
will be diagnosed with disease. The question of appropriate
selection of covariates for regression models has been widely
canvassed in the epidemiological literature, where these poten-
tial risk factors for the outcome are referred to as ‘interacting
variables’.15 Which of the conventional lung function predictors
could be considered a risk factor for obstructive lung disease?
Probably height is not a risk factor for disease. All the other
predictors are potentially risk factors. Age is a strong risk factor
for disease and for mortality. In many societies, race is strongly
related to risk factors for disease including environmental
exposures and nutrition. Sex may also be a risk factor for disease
both due to correlation with environmental exposures16 and due
to constitutional factors. Inclusion of each of these covariates
within the prediction equations tends to reduce the likelihood
that members of high risk groups defined by these factors will be
diagnosed with disease. The inclusion or exclusion of these
factors as predictors in the reference equations comes down to
answering difficult questions such as: is the presence of lower
lung function in older people, some racial groups and women
normal or an indicator of higher prevalence of disease? Or
alternatively, does the prevalence of disease increase with age, in
certain racial groups and in women? The obvious dilemma
posed by these questions points to the problem in choosing
covariates for inclusion in reference equations and in defining
normality based on these equations.

These problems with the definition of normal, representa-
tiveness of the study populations and the selection of predictors
for the spirometric reference equations have two important
consequences. First, they call into question the implicit
assumption that reference equations represent an absolute truth
and hence that they can be used as a criterion or gold standard
for classification of normality or disease. It is not valid to claim,
as some do, that another criterion for defining disease, such as
FEV1/FVC<0.7, must be wrong because it misclassifies subjects
compared with the lower limit of the normal derived from
reference equations. The second implication is that we need
a new approach to the interpretation of spirometry for
informing decisions in clinical practice.

The best place to start in designing a new approach is to ask
how we use spirometry or other tests in clinical decision
making. Ultimately, we are seeking to make a diagnosis and,
with this, enable advice about prognosis, risk factor modifica-
tion and the likely benefit of alternative treatment regimens.
This information is available from cohort studies and from ran-
mised controlled trials, but not from cross-sectional studies. The
important dimension is time. We perform tests to give infor-

mation about the future, something we cannot already know.
Cross-sectional studies tell us only about the present, which we
can already know.

We also need an approach that allows the incorporation of
other information into the interpretation of the results of
spirometry. In clinical practice, advice and decisions are not
made on the basis of a single test. This Bayesian approach was
elucidated by Sackett and his colleagues two decades ago,17 but
uptake into respiratory medicine has been slow.

We do not need to look too far to find evidence for the value of
an alternative approach to the use of test data for informing
prognosis. The Framingham risk factor equations are widely
used for predicting risk for a range of cardiovascular outcomes,
based on the results of a range of tests and observations.18 Our
cardiologist colleagues had the advantage of the Framingham
cohort to derive these equations.

We do have the data from a range of cohort studies and
randomised controlled trials that allow us to examine the

REFERENCES
3. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
4. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
5. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
6. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
7. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
8. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
9. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
10. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
11. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
12. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
13. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
15. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
17. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
18. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
19. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
20. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
22. National Heart, Lung, and Blood Institute. The NHLBI adult cardiorespiratory health
prognostic consequences of various levels of spirometric function for a range of clinically important outcomes including the onset of respiratory symptoms, accelerated decline in lung function, disability, hospitalisations and death. Appropriate analysis of these data, together with the incorporation of data on other risk factors, should allow the estimation of new risk equations for respiratory outcomes incorporating spirometry. Quantitative estimates of prognostic risk can be obtained, providing a strong basis for advice and for clinical intervention.

In conclusion, I argue that the current approach to the interpretation of spirometry is flawed. Reference equations have no special status as the repository of the truth about normal lung function. The debate about the lower limit of the normal versus a fixed ratio and the attempt to provide ‘world’ reference equations are distractions from the real task at hand: to develop respiratory risk equations based on spirometric measurements but incorporating other relevant risk factors and biomarkers of prognostic significance. Once developed, these equations can be readily translated into clinically useful and usable practice tools.

Competing interests None.

Contributors The author conceived the idea for this paper and drafted the manuscript alone.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
Are reference equations for spirometry an appropriate criterion for diagnosing disease and predicting prognosis?

Guy B Marks

Thorax 2012 67: 85-87 originally published online August 8, 2011
doi: 10.1136/thoraxjnl-2011-200584

Updated information and services can be found at:
http://thorax.bmj.com/content/67/1/85

References
This article cites 15 articles, 4 of which you can access for free at:
http://thorax.bmj.com/content/67/1/85#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
- Epidemiologic studies (1829)
- Clinical trials (epidemiology) (557)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/