

Letter to the Editor

Caution is Advised on the Use of Quality Grading in Spirometry



Precaución a la hora de aplicar los grados de calidad en la espirometría

Dear Editor:

Statements on pulmonary function reporting stress the need to have a system to evaluate the quality of spirometries (A through F), both in the individual acceptability of each maneuver and in their repeatability.¹ This scores, based on the ones already included in many commercial spirometry softwares and used in several studies,² are eminently numerical, and base score assignment according to certain numerical criteria, easily calculable for any

computer, such as forced expiratory time >6 s, back extrapolation volume <150 mL or 5% of FVC or repeatability of two best efforts within 150 mL.

However a quick review of the spirometric acceptability criteria allows observing that acceptability depends to a large extent on morphological criteria (“peak expiratory flow should be achieved with a sharp rise and occur close to the point of maximal inflation”) and even subjective to the operator (“If the subject cannot or should not continue to exhale”).³ Among the errors frequently accepted by algorithms, it is common to find “A” rated studies with maneuvers with negative effort dependence, glottis closure, cowboy hat-shaped maneuvers and re-inhalation (Fig. 1).

It is in these where there are notable discrepancies between the evaluation of spirometric quality made by an experienced reviewer and that made by computer algorithms.⁴ That is to say, if we trust

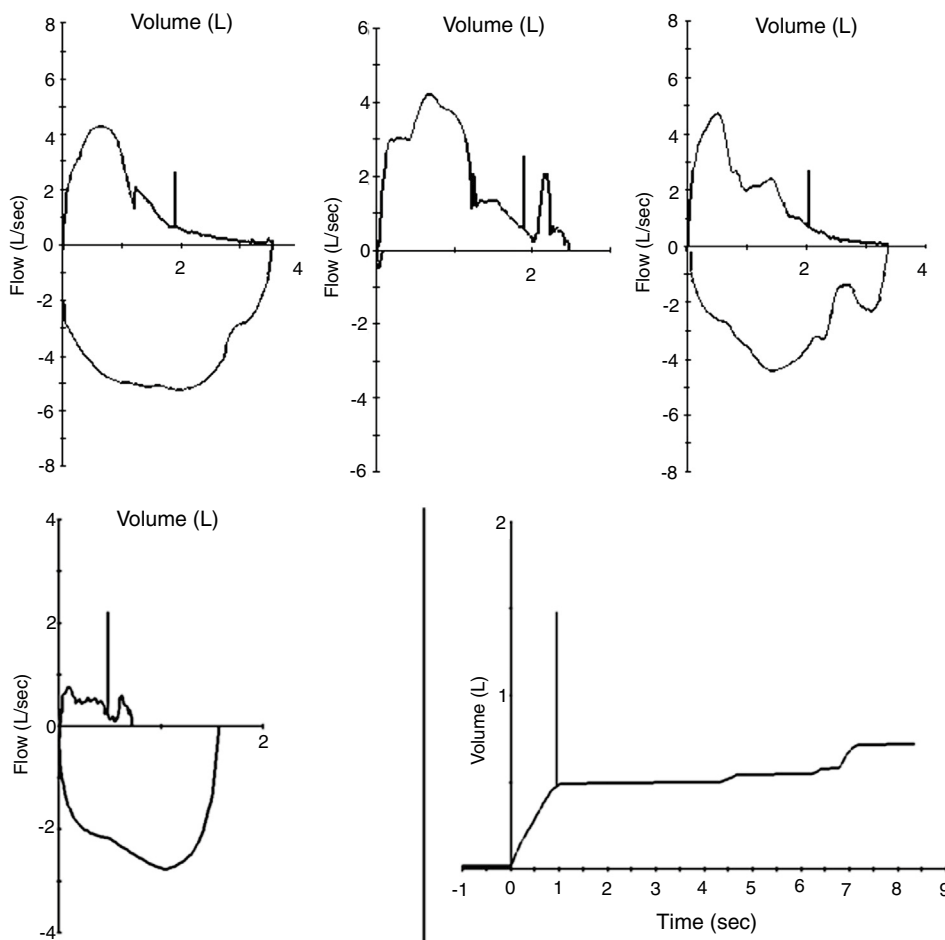


Fig. 1. Several examples of efforts deemed as acceptable by spirometer softwares.

only in the software, the risk of accepting as valid unacceptable results is high.

Use of automatic quality control has been widespread attempted in primary care with disappointing results,⁵ which led many investigators to study the feasibility of remote monitoring of spirometric quality control by experienced reviewers. In a recent editorial, Marina et al. review the importance of continuous training as a basis for achieving acceptable tests, in addition to remote monitoring.⁶

Correctly, the statements recommend that the acceptability assigned by the software should be reviewed by an experienced reviewer.¹ However, in centers with a high work flow, operational simplification can lead to the avoidance of this aspect. Additionally, the increasingly frequent profusion of works with a high number of subjects,⁷ leads to choose computer driven quality scores, something practical in all given the number of subjects recruited, which are counted by thousands. However, the high discrepancy in the evaluation of quality casts doubt on the validity of their final data and their conclusions. Many of these works have a purely epidemiological cut and as such, support sanitary policies, implemented at the expense of taxpayers.

Spirometries where every acceptability criteria are not met is a frequent finding in daily workflow. Forced expiratory volume in the first second (FEV₁) could be valid as a datum in the absence of end-of-test criteria, where forced vital capacity (FVC) is not trustworthy. In patients with unacceptable back extrapolation volume, where FEV₁ is spurious, FVC could still be useful as an isolated number. Nevertheless, at the time of reporting, software does not allow to overturn any of this figures (or even that of dependent FEV₁/FVC). This could lead to an incorrect use in clinical decision making, even having noted the caveats in the interpretation.

Similarly, in patients with severe airway obstruction, where expiratory times can exceed the recommended 15 s without achieving plateau, or in patients with poor effort tolerance, FVC could be considered as a minimum value. And this is neither reflected nowhere else, with the exception of the written report at the end of the study.

In that context, software could incorporate an option to annul isolated data based on revisor's judgment, or point-out that a figure is "at least as low as" or "at least as high as". This modifications should be automatically transferred to FEV₁/FVC and other derived quotients, dependent on any of them.

In summary, spirometry software developers should add new capabilities based in intelligent detection of artifacts, cancelation of invalid data, such as FVC or FEV₁, or setting them as maximum or minimum depending on specific maneuver defects. Meanwhile, spirometric maneuvers and results should be systematically reviewed by an experienced reviewer.

References

1. Culver BH, Graham BL, Coates AL, Wanger J, Berry CE, Clarke PK, et al. ATS committee on proficiency standards for pulmonary function laboratories recommendations for a standardized pulmonary function report. An official American thoracic society technical statement. *Am J Respir Crit Care Med.* 2017;196:1463–72.
2. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med.* 1999;159:179–87.
3. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. ATS/ERS Task Force Standardisation of spirometry. *Eur Respir J.* 2005;26:319–38.
4. Müller-Brandes C, Krämer U, Gappa M, Seitner-Sorge G, Hüls A, von Berg A, et al. LUNOKID: can numerical American Thoracic Society/European Respiratory Society quality criteria replace visual inspection of spirometry? *Eur Respir J.* 2014;43:1347–56.
5. Huetto J, Cebollero P, Pascal I, Cascante JA, Eguía VM, Teruel F, et al. Spirometry in primary care in Navarre, Spain. *Arch Bronconeumol.* 2006;42:326–31.
6. Marina N, López de Santa María E, Gáldiz JB. Telemedicine, an opportunity for spirometry. *Arch Bronconeumol.* 2018;54:306–7.
7. Echazarreta AL, Arias SJ, Del Olmo R, Giugno ER, Colodenco FD, Arce SC, et al. Grupo de estudio EPOC.AR prevalence of COPD in 6 urban clusters in Argentina: The EPOC.AR study. *Arch Bronconeumol.* 2018;54:260–9.

Santiago C. Arce

Pulmonary Function Laboratory, Instituto de Investigaciones Médicas A. Lanari, University of Buenos Aires, Argentina
E-mail address: arcesantiago@fibertel.com.ar