



## Editorial

The Exception does not Prove the Rule: Lessons From Physics<sup>☆</sup>

## La excepción no confirma la regla: lecciones de la física

Adolfo Balóira Villar,<sup>\*</sup> Marta Núñez Fernández

Servicio de Neumología, Complejo Hospitalario Universitario de Pontevedra, Pontevedra, Spain

Richard Feynman, winner of the Nobel Prize in Physics, began one of his Caltech lectures with the following statement: “The exception proves that the rule is wrong. If there is an exception to any rule, and if it can be proved by observation, that rule is wrong.”

There are two facets to scientific research: the first, the descriptive, is based on the closest possible observation and description of a certain phenomenon; the second consists of explaining that observation. The ideal explanation comes in the form of a mathematical formula that accurately predicts what will happen each time the phenomenon occurs. In 1905, Albert Einstein published the most famous formula of all time, his theory of relativity, barely understandable to more than a select handful of physicists. How could anyone conceive that a human body could contain as much energy as a city would use in a year? And Einstein received his Nobel for his explanation of the photoelectric effect, not for his devastating discovery of relativity. All laboratory attempts to disprove Einstein have failed. An experiment set up in 2011 in the CERN, the European Organization for Nuclear Research, initially appeared to undo the theory, but, after the initial excitement, it was confirmed in 2012 that there was an error in calculation: Einstein and relativity had survived another coup attempt.

Medicine is, in general, a probabilistic science. Most of the phenomena studied are affected by a large number of variables; usually we are aware of only a few, and very little is known about the role these few play in the development of the phenomenon in question. For example, the volume of air exhaled in one second—FEV1—is subject to probably hundreds, or maybe even thousands, of variables. This volume is measured in a large number of subjects and a somewhat inelegant equation is formulated. The level of accuracy of the formula is proven when it is applied to large numbers of individuals. Most of the regression equations we use for disease processes are not sufficiently precise. For example, it is impossible to quantify the probability of a certain individual developing a 4 cm pulmonary cavity in the upper left lobe after 24 hours

in contact with a tuberculosis patient. For this reason, it is essential that clinicians clearly understand what they are saying when they speak of values in the context of a specific situation. For a start, we need to understand that medicine is basically a descriptive science. Experiments (clinical trials, for example) are performed, we record what happens and extract some conclusions. The reliability of these conclusions depends largely on the precision of the tools used for measurement. In physics, an experiment can be performed in a single case, because the number of variables is small and usually can be controlled. If an object falls to Earth from a height of 20 meters, it will always reach the same acceleration and velocity, exactly as predicted by a simple formula. So what happens if we try to study a bronchodilator in COPD patients? The study parameters are usually lung function (FEV1), quality of life (questionnaires), or maybe the number and type of exacerbations. In a study carried out in humans, these parameters are subject to an infinite variety of intra-individual variables, most of which are impossible to define or to control. These subject variables unavoidably affect the primary variable – the subject’s state of mind alone during the test introduces an indefinite number of new factors. The accuracy of the tools used to measure each parameter varies widely: FEV1 is reasonably precise, but quality of life, being only semi-quantitative, is less so. As for exacerbations, we do not even have a precise definition confirming with 100% certainty whether an event was or was not an exacerbation, not to mention quantifying the exacerbation any further than a simple classification. Returning to FEV1, we perform at least 3 tests, accurate to the nearest millimeter, but the chances are that each result will be different. The ideal situation would be to perform the test many, many times over and take the most frequently repeated result, but this would be impossible from a practical point of view. According to SEPAR criteria,<sup>1</sup> the highest value should be used, provided the difference between the 2 best results is less than 150 ml (100 ml if FEV1 is less than 1000 ml). It seems inappropriate to set an absolute number in a test with a range of normal values than can be more than doubled, depending on the individual. Even if a variation of 10% is accepted, the test is still far from precise. Mean improvement in FEV1 with the most potent bronchodilators available is about 170 ml.<sup>2</sup> When 2 bronchodilators are compared, the difference ranges from 30 ml to 90 ml.<sup>3,4</sup> It seems obvious that if the increase is much lower than the value considered acceptable for test reproducibility, the results will be less reliable.

<sup>☆</sup> Please cite this article as: Balóira Villar A, Núñez Fernández M. La excepción no confirma la regla: lecciones de la física. Arch Bronconeumol. 2015;51:161–162.

<sup>\*</sup> Corresponding author.

E-mail address: [adolfo.baloira.villar@sergas.es](mailto:adolfo.baloira.villar@sergas.es) (A. Balóira Villar).

As mentioned above, in physics a principle needs to be based on mathematical proof that can accurately predict the results of any test based on that principle. In the case of FEV1, the range of possible values is wide. For this reason, we usually express the value as a mean of the total range. To apply these results to a specific patient, then, is practically impossible. We cannot predict what will happen if the patient receives such and such bronchodilator, any more than we can predict the income of an individual person from the average income of their country. It would seem, then, more effective to determine the change that would have positive consequences on the patient's health and determine how many patients in each trial group achieve it. If the difference is 15% in favor of the drug, perhaps that percentage comes close to the real change that the patient would need to demonstrate to achieve improvement. The problem is to accurately determine the minimum value of improvement. In respiratory medicine, the threshold limits are insufficiently precise. Values of 100 ml have been proposed for FEV1,<sup>5</sup> although it seems more reasonable to establish percentage changes. The St. George's Hospital Questionnaire sets the change at 4%, but does this have the same relevance if the baseline score is 35 or 58? What is the practical value of a mean change of 2.5% compared to a reference agent, even if the difference according to the statistical analysis is very significant? Is this information of any help to our patient?

Medicine is a probabilistic science that often only addresses events that are difficult to explain or understand. Study designs must be closely analyzed and conclusions must be drawn with care.

The Heisenberg uncertainty principle determines that any measurement will be affected by the measurement system used. When dealing with probabilities, we must be very aware of what we are measuring, how we measure it and to what or to whom the measurement can be applied. Einstein said it was crazy to repeat a test again and again, expecting to obtain different results. In medicine, studies that are similar (or so we think) in design, often produce very different results.

To conclude, it might be a good idea to apply some of the principles of physics to medical research, so that the findings from clinical trials can be better measured and evaluated.

## References

1. Garcío-Río F, Calle M, Burgos F, Casan P, del Campo F, Gáldiz JB, et al. *Espirometría. Normativa SEPAR*. Arch Bronconeumol. 2013;49:388–401.
2. Dahl R, Chung KF, Buhl R, Magnussen H, Nonikov V, Jack D, et al. *Efficacy of a new once-daily long-acting inhaled beta2-agonist indacaterol versus twice-daily formoterol in COPD*. Thorax. 2010;65:473–9.
3. Vincken W, Aumann J, Chen H, Henley M, McBryan D, Goyal P. *Efficacy and safety of coadministration of once-daily indacaterol and glycopyrronium versus indacaterol alone in COPD patients: The GLOW6 study*. Int J Chron Obstruct Pulmon Dis. 2014;9:215–28.
4. Decramer M, Anzueto A, Kerwin E, Kaelin T, Richard N, Crater G, et al. *Efficacy and safety of umeclidinium plus vilanterol versus tiotropium, vilanterol, or umeclidinium monotherapies over 24 weeks in patients with chronic obstructive pulmonary disease: Results from two multicentre, blinded, randomised controlled trials*. Lancet Respir Med. 2014;2:472–86.
5. Donohue JF. *Minimal clinically important differences in COPD lung function*. COPD. 2005;2:111–24.