REFERENCE VALUES: NO NEED FOR CONFUSION To the Editor:

I would like to comment on the discussion among Lim and Dusmet,¹ Marra and colleagues,² and Rice and Blackstone.³ There are several issues of confusion; I hope I can clarify some of these.

Sensitivity and specificity are measures of a test's inherent diagnostic performance. Sensitivity is the proportion of patients who test positive among patients with the disease; specificity is the proportion of patients who test negative among patients without the disease. Another common measure of diagnostic performance is the receiver operating characteristic (ROC) curve.⁴ An ROC curve illustrates a test's sensitivity and specificity for different criteria for defining positive and negative test results. For highly accurate tests, there is a point on the ROC curve that one can choose if high specificity is desired; the price, however, is low sensitivity. Similarly, one can choose very high sensitivity but at a price of low specificity. Lim and Dusmet's¹ comment that "sensitivity truly starts at 50%" is incorrect; a test with low sensitivity (ie, <0.5) can have diagnostic value if the specificity is high.

Sensitivity and specificity are the basic measures of a test's ability, but they do not describe how well the test will perform for a particular patient population. In managing patients, physicians focus on what the test results tell them about their patient. They want to know the probability

their patient has the disease after a positive test result (positive predictive value [PPV]) and the probability their patient does not have the disease after a negative test result (negative predictive value [NPV]). Predictive values depend not only on the sensitivity and specificity of the test but also on the probability of disease in similar patients (ie, prevalence of disease). In fact, when predictive values are reported in the literature, a subscript indicating the prevalence rate is often used. For example, remediastinoscopy may have an NPV of 0.85 in a sample with a prevalence rate of 0.32, which we write as $NPV_{0.32} = 0.85$. In a different population with a different prevalence rate, the NPV will change, for example, $NPV_{0.05} = 0.98$ or $NPV_{0.50} = 0.72$. Much of the controversy in these authors' correspondences is due to confusion between sensitivity and PPV, and between specificity and NPV. Sensitivity and specificity describe the test's inherent diagnostic abilities irrespective of the prevalence rate. PPV and NPV, on the other hand, tell us the likelihood of disease after the test is performed in a particular patient population with a particular prevalence rate. In determining the role of remediastinoscopy in restaging lung cancer, it seems that PPV and NPV are the important metrics and should be the focus of the discussion.

Lim and Dusmet¹ and Marra and colleagues² point out correctly that specificity is important for ruling in disease and sensitivity is important for ruling out disease. These relationships are due to the roles of these metrics in estimating PPVs and NPVs. A high specificity causes the PPV to increase, and a high sensitivity causes the NPV to increase, assuming, of course, that the prevalence of disease is held constant. As we have illustrated, predictive values are highly influenced by the prevalence of disease. Similarly, the measure of "accuracy" that Marra and colleagues report is also dependent on the prevalence of disease in the sample, and thus could be reported more appropriately as overall accuracy_{0.32} = 0.88.

There are several other issues in these correspondences that need clarification. First, neither Marra and colleagues² nor Lim and Dusmet¹ report a confidence interval (CI) for specificity. A reasonable 95% CI for specificity based on these data is 0.96 to 1.0.⁵ CIs for both sensitivity and specificity should be routinely reported. Contrary to Marra and colleagues' description of the meaning of a CI, it is not "the likelihood that another sample will provide the same result." Rather, a CI describes a range of plausible values for the metric of interest, here specificity. Statistically speaking, we expect that 95% of CIs will contain the real, but unknown, true value of the metric (ie, specificity); 5% of CIs will not contain the true value. Statisticians use the data from a single sample to estimate the unknown value of the metric; 95% of the time the CI they construct contains the true value, although we do not know which value in the interval it is or which CIs contain the true value and which do not.

Second, it is important to consider the effects of patient and disease characteristics in estimating sensitivity and specificity. For example, the size of lesions is a critical determinant of sensitivity, as well as the comorbidities of patients. Some of the differences between estimates of sensitivity and specificity reported in the literature for remediastinoscopy could be due to these patient differences.

Third, when a diagnostic test does not yield a result, that is, the result is "uninterpretable,"⁶ it is critical that the frequency of this occurrence be reported. Marra and colleagues² reported a 2% frequency for remediastinoscopy. They also included this frequency in the denominator of their estimate of overall accuracy; this gives the reader an honest estimate of the test's performance.

Last, I think Drs Rice and Blackstone's³ statement that screening tests usually have good specificity, whereas a test used to work up patients needs good sensitivity, is too narrow and does not describe many scenarios. In screening for breast cancer, for

The Editor welcomes submissions for possible publication in the Letters to the Editor section that consist of commentary on an article published in the Journal or other relevant issues. Authors should: • Include no more than 500 words of text, three authors, and five references. • Type with double-spacing. • See <http://jtcvs.ctsnetjournals.org/misc/fora.shtml> for detailed submission instructions. • Submit the letter electronically via jtcvs.editorialmanager.com. Letters commenting on an article published in the JTCVS will be considered if they are received within 6 weeks of the time the article was published. Authors of the article being commented on will be given an opportunity of offer a timely response (2 weeks) to the letter. Authors of letters will be notified that the letter has been received. Unpublished letters cannot be returned.

example, physicians look for tests with good sensitivity even if the false-positive rate is a bit high. Computer-aided detection systems are often used to improve sensitivity, usually at a cost of even higher recall rates. Without reasonable sensitivity, many screening programs cannot be cost-effective. Further workup of these patients demands higher specificity to prevent unnecessary invasive testing. The consequences of test errors and prevalence of disease must be weighed in each application to find the best test for a particular application.

Nancy A. Obuchowski, PhD
Department of Quantitative
Health Sciences
Cleveland Clinic
Cleveland, Ohio

References

1. Lim E, Dusmet M. Remediastinoscopy: a statistical reinterpretation. *J Thorac Cardiovasc Surg.* 2009; 137:254-5. author reply 5-6.
2. Marra A, Hillejan L, Fechner S, Stamatis G. Remediastinoscopy in restaging of lung cancer after induction therapy. *J Thorac Cardiovasc Surg.* 2008; 135:843-9.
3. Rice TW, Blackstone EH. Referent values and equipoise: editors notes. *J Thorac Cardiovasc Surg.* 2009;137:256-7.
4. Zhou XH, Obuchowski NA, McClish DK. *Statistical Methods in Diagnostic Medicine.* New York: Wiley and Sons; 2002.
5. Hanley JA, Lippman-Hand A. If nothing goes wrong, is everything all right? Interpreting zero numerators. *JAMA.* 1983;249:1743-5.
6. Begg CB, Greenes RA, Iglewicz B. The influence of uninterpretability on the assessment of diagnostic tests. *J Chronic Dis.* 1986;39:575-84.

doi:10.1016/j.jtcvs.2009.02.031

EXPERIMENTAL VENTRICULAR PRESSURE OVERLOAD IN YOUNG ANIMALS INDUCES CARDIOMYOCYTE HYPERTROPHY IN ADDITION TO HYPERPLASIA OF CONTRACTILE AND NONCONTRACTILE ELEMENTS OF THE MYOCARDIUM

To the Editor:

Since the early 1990s, we have been working with experimental right ventric-

ular hypertrophy using young goats aged 4 to 8 weeks to achieve right (subpulmonary) ventricular retraining by means of a balloon catheter or an adjustable pulmonary artery banding system.¹⁻³ We have consistently found a hypertrophic response of the cardiomyocytes. Moreover, we have also demonstrated that the right ventricle responds not only with hypertrophy of the myocardial fibers but also with hyperplasia of the contractile and interstitial components of the myocardium.³ Recently, Leeuwenburgh and colleagues⁴ demonstrated the development of a hyperplastic rather than a hypertrophic myocardial response in young animals submitted to chronic right ventricular pressure overload. Their findings deserve some comments because previous studies also showed a hypertrophic response, even in young animals. Anversa and associates,⁵ studying 30 normal Wistar rats at 1, 5, and 11 days of age, found that left ventricular cardiomyocytes presented both hypertrophy and hyperplasia, attributing these to the physiologic pressure overload imposed on that chamber after birth.

Based on the age of the lambs (2–3 weeks), the duration of pressure overload (8-week period), and the reported data from the study by Leeuwenburgh and colleagues,⁴ we are surprised about the fact that the authors did not find any signal of myocardial fiber hypertrophy. Moreover, even considering the hyperplastic response alone, no mention was made regarding the presence or absence of proliferating (proliferating cell nuclear antigen–labeled) interstitial cells. If, as the authors stated, “more of the same tissue” was found in the trained myocardium, in addition to proliferating cardiomyocytes, they should have found some hyperplasia of cells from the interstitial compartment of the heart. In fact, this was a prominent feature of our experimental model of pressure overload in young animals.³ The number of interstitial and vessel cells labeled by the cell proliferation marker Ki67 in our study was 14 and 18 times greater, respectively, than that of the cardio-

myocytes under 2 different protocols of ventricular systolic overload (continuous and intermittent).³

Maria C. D. Abduch, VMD, PhD^a
Renato S. Assad, MD, PhD^b
V. D. Aiello, MD, PhD^c

^aLaboratory of Echocardiography

^bDivision of Surgical Research

^cLaboratory of Pathology

Heart Institute (InCor)

University of São Paulo Medical

School

São Paulo, Brazil

References

1. Assad RS, Cardarelli M, Abduch MC, et al. Reversible pulmonary trunk banding with a balloon catheter: assessment of rapid pulmonary ventricular hypertrophy. *J Thorac Cardiovasc Surg.* 2000;120:66-72.
2. Dias CA, Assad RS, Caneio LF, et al. Reversible pulmonary trunk banding II. An experimental model for rapid pulmonary ventricular hypertrophy. *J Thorac Cardiovasc Surg.* 2002;124:999-1006.
3. Abduch MC, Assad RS, Rodriguez MQ, et al. Reversible pulmonary trunk banding III: assessment of myocardial adaptive mechanisms—contribution of cell proliferation. *J Thorac Cardiovasc Surg.* 2007;133:1510-6.
4. Leeuwenburgh BPI, Helbing WA, Weninck ACG, et al. Chronic right ventricular pressure overload results in a hyperplastic rather than a hypertrophic myocardial response. *J Anat.* 2008;212:286-94.
5. Anversa P, Olivetti G, Loud AV. Morphometric study of early postnatal development in the left and right ventricular myocardium of the rat. Hypertrophy, hyperplasia and binucleation of myocytes. *Circ Res.* 1980;46:495-502.

doi:10.1016/j.jtcvs.2009.01.021

CLAMPING THE MITRAL AND TRICUSPID ANNULI WITH BIPOLAR DEVICES

To the Editor:

Castella and colleagues' article¹ reports evidence collected by application of a bipolar radiofrequency clamp on 8 explanted hearts and post-mortem analysis of 1 patient who died hours after surgery. The authors unveil 2 main findings: They state it is not possible to clamp across the mitral annulus because of the increased thickness of the atrioventricular (AV) groove, related to the superimposition of the ventricular mass and fat tissue. They also claim that, although feasible, clamping