JACC: CARDIOVASCULAR IMAGING
VOL. 4, NO. 8, 2011

ISSN 1936-878X/$36.00
PUBLISHED BY ELSEVIER INC.
DOI:10.1016/j.jcmg.2011.06.008

EDITORIAL COMMENT

When Is a Number Not a Number? Quality Control Arrives in the Imaging Laboratory*

Thomas H. Marwick, MBBS, PHD
Cleveland, Ohio

Despite competition from new measures at rest and on exertion (1), the assessment of ejection fraction (EF) from 2-dimensional echocardiography is likely the most widely used number in cardiology, if not in internal medicine. Its prognostic value has led to its use in the selection of patients for a variety of treatments, especially in heart failure (2). Despite criticisms relating to dependence on loading conditions, the significant test/retest variability of this test (driven in part by biological variation), and

See page 821

dependence on image quality and geometric assumptions, the simplicity of this measure is perhaps the biggest clue to its enduring use. However, the sources of variation in EF represent an important ongoing limitation (3). The implications of these variations extend from the minor annoyance of differences in measurement within sequential studies that are clearly normal or clearly abnormal, extending to variations across thresholds that imply suitability for device therapy. At the worst extreme, indications for device therapy may change from 1 visit to the next.

The avoidance of the shortcomings can be based upon quantitation or education, or a combination of both. In this issue of *iJACC*, investigators from Massachusetts General Hospital describe a very important teaching intervention to reduce interobserver variability in the assessment of EF whereby participants attended 6 1-h sessions over a 6-month period (4). The baseline status of the EF estimations was gathered in the first session, the next 3 involved a case-based approach to obtaining consensus, culminating in the 2 final sessions, where new cases were measured with feedback. A subgroup participated in a follow-up session a year later. The results of the 25 readers demonstrated that a very impressive 40% reduction of interobserver variability from 14% of EF prior to intervention. Indeed, this improvement was sustained over 12 months of follow-up. This improvement also occurred in the mid-range of EF, where the misclassification rate was even higher than in the group overall. The findings of this study leave us with 3 important messages—about the importance of quality control, the role of quantitation, and the variability of EF.

The core message of this study is not only its content about EF, but also its lesson about quality control activities in the imaging laboratory. The American Society of Echocardiography has emphasized the importance of reducing interobserver variability (5). Although this is a problem with EF, it is much more a problem with complex multiparametric parameters, where the hierarchy of individual parameters is not provided by the guidelines, and therefore, patients with discordant findings may be readily classified differently by different observers. Prominent examples of these multiparametric evaluations include assessment of regurgitant heart valves and the assessment of diastolic dysfunction (6). Indeed, the limited literature on variability in valve evaluation suggests that there is a high degree of variability between observers that ranges all the way from differing assessments of suitability for valvuloplasty (7) to differing assessments of regurgitation severity that may have implications for the selection for surgery (8).

*Editorials published in *JACC: Cardiovascular Imaging* reflect the views of the authors and do not necessarily represent the views of *JACC: Cardiovascular Imaging* or the American College of Cardiology.

From the Cleveland Clinic, Cleveland, Ohio. Dr. Marwick has reported that he has no relationships relevant to the contents of this paper to disclose.

In response to heightened concern about medical costs and quality, there has been success in improving quality, although much work remains to be done (9). Some variation in practice may reflect failure to adopt practice guidelines; interestingly, although often attributed to disagreement with guidelines, this appears at least as much due to a variety of reasons, including unfamiliarity and lack of awareness (10). Both guidelines review and concordance reading should be part of the continuing education activities of every imaging laboratory, and a number of observations from this paper and the literature may guide us to the successful possible strategies for achieving this. First, this is more likely to be successful as an active rather than a passive process (11). The use of a regularly scheduled conference time and clear leadership of this activity by senior laboratory members is absolutely mandatory. A case-based strategy of practice, rehearsal, and performance is effective in teaching other skills that require recognition and hand–eye coordination (12). The process of gaining consensus by the group working together is important, and as this can be confronting, needs to be performed in a nonjudgmental environment. Finally, continuous feedback and quality control are vital components of this process, and should be addressed by performance of an entry and an exit task. Indeed, in addition to EF evaluation, many other aspects of imaging require the evolution of this process, the presentation of which is timely and overdue.

The second message of this paper is about quantitation, potentially an important weapon in the avoidance of variation. Despite encouragement to adopt quantitative analysis (13), the stubborn attachment of most clinicians to qualitative evaluation is understandable. The assessment of global left ventricular (LV) systolic function is not necessarily well represented by end-systolic and end-diastolic frames, which neglect the speed and co-ordination of contraction. In many settings, poor delineation of endocardial borders constrains the ability to trace LV volume, and although this may be rectified by LV opacification, incorporation of contrast into the routine is limited by challenges pertaining to cost, administration, and availability (14).

In the same study, the 5 sonographers who measured LV volumes initially demonstrated a standard deviation from the biplane EF of approximately 4%, which was reduced to 2.5% following intervention. This led to an interobserver variability of 6%, falling to 3% EF post-intervention. This is a remarkably good response to quantitative training, and although the authors have proposed the training strategy as a means of reducing variability, these results alone would suggest that quantitation is an important part of the solution of this problem. An alternative quantitative strategy might be the incorporation of new technologies such as global longitudinal strain, which correlates loosely with EF but also appears to add to the discrimination of risk, especially in patients with mild dysfunction (15). Indeed, in addition to the assessment of strain, new software approaches to EF measurement may limit variability, because the software identifies the same landmarks and edges whenever it is presented with the same image.

The final message is perhaps the most important: the estimation of EF is inherently variable. Previous investigators have published the 95% confidence intervals of EF from repeat 2-dimensional echocardiograms to exceed 10%. Even following the intervention in this paper, the interobserver variability of EF was 8.4%. Tracking treatment response requires a better-performing measure for careful assessment of the progression of LV dysfunction. This is becoming apparent in the field of cardio-oncology (16), as well as other disorders of pre-clinical LV dysfunction, and this need will become more acute as stem cell therapy evolves into clinical practice. Johri et al. (4) have provided an exemplary study of how to incorporate quality improvement into the imaging laboratory, but the paper also reminds us of the need for a sensitive, robust, and reproducible marker of LV function.

Reprint requests and correspondence: Dr. Thomas H. Marwick, Cardiovascular Imaging J1-5, Heart and Vascular Institute, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, Ohio 44195. *E-mail: marwict@ccf.org*.

REFERENCES

1. Jorge Alves A, Ribeiro F, Sagiv M, et al. Resting measures and physiological responses to exercise for the determination of prognosis in patients with chronic heart failure: useful tools for clinical decision-making. Cardiol Rev 2010;18:171–7.
2. Hunt SA, Abraham WT, Chin MH, et al. 2009 focused update incorporated into the ACC/AHA 2005 guidelines for the diagnosis and management of heart failure in adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2009;53:e1–90.
3. Otterstad JE, Froeland G, St John Sutton M, Holme I. Accuracy and reproducibility of biplane two-dimensional echocardiographic measurements of left ventricular dimensions and function. Eur Heart J 1997; 18:507–13.
4. Johri AM, Picard MH, Newell J, Marshall JE, King MEE, Hung J. Can a teaching intervention reduce interobserver variability in LVEF assessment: a quality control exercise in the echocardiography lab. J Am Coll Cardiol Img 2011;4:821–9.
5. Kisslo J, Byrd BF, Geiser EA, et al. Recommendations for continuous quality improvement in echocardiography. J Am Soc Echocardiogr 1995; 8:1–28.
6. Unzek S, Popovic ZB, Marwick TH. Effect of recommendations on interobserver consistency of diastolic function evaluation. J Am Coll Cardiol Img 2011;4:460–7.
7. Fox KF, Porter A, Unsworth B, Collier T, Leech G, Mayet J. Report on a national echocardiography quality-control exercise. Eur J Echocardiogr 2009;10:314–8.
8. Biner S, Rafique A, Rafii F, et al. Reproducibility of proximal isovelocity surface area, vena contracta, and regurgitant jet area for assessment of mitral regurgitation severity. J Am Coll Cardiol Img 2010;3:235–43.
9. Jencks SF, Huff ED, Cuerdon T. Change in the quality of care delivered to Medicare beneficiaries, 1998–1999 to 2000–2001. JAMA 2003;289:305–12.
10. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999; 282:1458–65.
11. Grimshaw JM, Shirran L, Thomas R, et al. Changing provider behavior: an overview of systematic reviews of interventions. Medical Care 2001;39 Suppl II:2–45.
12. Kneebone RL. Practice, rehearsal, and performance: an approach for simulation-based surgical and procedure training. JAMA 2009;302: 1336–8.
13. Lang RM, Bierig M, Devereux RB, et al. Recommendation for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group. J Am Soc Echocardiogr 2005;18:1440–63.
14. Honos G, Amyot R, Choy J, Leong-Poi H, Schnell G, Yu E. Contrast echocardiography in Canada: Canadian Cardiovascular Society/Canadian Society of Echocardiography position paper. Can J Cardiol 2007;23:351–6.
15. Stanton T, Leano R, Marwick TH. Prediction of all-cause mortality from global longitudinal speckle strain: comparison with ejection fraction and wall motion scoring. Circ Cardiovasc Imaging 2009;2:356–64.
16. Albini A, Pennesi G, Donatelli F, Cammarota R, De Flora S, Noonan DM. Cardiotoxicity of anticancer drugs: the need for cardio-oncology and cardio-oncological prevention. J Natl Cancer Inst 2010;102:14–25.

Key Words: ejection fraction ■ interobserver variability ■ left ventricle ■ quality control.