Has LVEF changed beyond chance? Limits of agreement of radiotracer-derived LVEF

Frans J. Th. Wackers, MD, PhD[a]

[a] Yale University School of Medicine, New Haven, CT

Received Feb 1, 2015; accepted Feb 1, 2015
doi:10.1007/s12350-015-0090-y

See related article, pp. 1237–1243

Left ventricular ejection fraction (LVEF) is one of the strongest predictors of cardiac outcomes. Any patient with known or suspected heart disease will have at least one assessment of LVEF by a noninvasive imaging technique. Usually, an approximation of global LV function will suffice: an estimated or calculated LVEF; normal or abnormal, and if abnormal, to what degree is left ventricular function compromised.

Such crude categorization will usually suffice in clinical practice. However, when cardiac function is assessed repeatedly e.g. to determine the effect of therapeutic interventions or to follow the course of cardiac disease, more precise measurements are required. It is then that it is important to know which change in LVEF can be confidently considered to be a real change beyond chance?

BIOLOGICAL VARIABILITY

In the stable resting state, LVEF is usually above 0.50. During graded exercise LVEF steadily increases until it reaches a plateau at the level of physiologic maximal LVEF. Maximal achievable LVEF of normal hearts is in the high 0.80 s.

The normal left ventricle responds to increased physiologic demands by augmenting heart rate and LVEF. Such change may occur almost instantaneously as clearly was demonstrated by beat-to-beat LVEF analysis using the nuclear stethoscope.[1,2] LVEF can change substantially from one beat to another because of changes in loading conditions, e.g. in atrial fibrillation, or due to catecholamine surges, e.g. under mental stress. LVEF of normal hearts therefore can rapidly change within the normal range of 0.50 to 0.80, whereas abnormal hearts have a much more restricted LVEF reserve, limited by the presence of structural heart disease.

It is important to be aware of the spontaneous variability of LVEF (particular in the normal range) when serial left ventricular function measurements are made. Lack of reproducible measurements may very well be due to the above-mentioned biological variability, but could also be the result of technical limitations of the imaging methodology.

FACTORS AFFECTING VARIABILITY OF LVEF

To fully understand the significance of changes in LVEF one should determine the limits of agreement, or variation beyond chance, of the imaging technique by Bland-Altman analysis.[3] This analysis defines ''limits of agreement'' (or repeatability coefficient) as the mean difference between two measurements ±1.96 standard deviations. Changes exceeding the limits of agreement are likely to represent real changes.

The variability of measurements should be examined systematically for:

- variability of visual or computer analysis,
- variability of reprocessing data of the same acquisition,
- variability of processing data of separate acquisitions,
- variability of results by the same and by different operators,
- variability of data of acquisitions separated by shorter or longer time intervals,
- variability of results in normal and abnormal hearts.

Reprint requests: Frans J. Th. Wackers, MD, PhD, Yale University School of Medicine, New Haven, CT; *frans.wackers@yale.edu*
J Nucl Cardiol 2015;22:1244–6.
1071-3581/$34.00

LVEF DERIVED FROM BLOOD POOL IMAGING VS. MYOCARDIAL PERFUSION IMAGING

There are important technical differences to be considered between LVEF derived from ECG-gated equilibrium (E) radionuclide angiography (RNA) or from ECG-gated myocardial perfusion imaging (MPI) that impact the accuracy of measurements.

When the cardiac *blood pool* is labeled with a radiotracer, changes in count density are proportional to changes in blood volume, thus ventricular contraction. When the *myocardium* is labeled with a perfusion tracer, changes in count density are actually **apparent** changes due to differential partial volume effect during myocardial thickening. Thus, gated MPI does not measure directly blood volume changes. The count density of ERNA is many times higher than that of MPI, allowing for reproducible (endocardial) blood volume edge detection by mathematical criteria. Count densities of MPI SPECT slices, in comparison, are relatively low and myocardial borders are poorly delineated. The perceived motion on gated MPI SPECT is mainly the result of improved count recovery (i.e. less partial volume effect during myocardial thickening) and to a lesser extent due to motion of myocardial borders across the plane of view.

STATISTICAL RELIABILITY

Nuclear cardiology assessment of LVEF has always been quantitative, based on changes in count density from diastole to systole. The statistical reliability and reproducibility of calculations of LVEF is related to count density and inversely to the level of LVEF: for similar statistical error, a low LVEF requires considerably greater count density than a normal LVEF.[4] Fortuitously in ERNA, when LVEF is low, left ventricular blood volume is large, counts are plentiful, and LVEF is highly reproducible. In contrast, in gated SPECT MPI, low LVEF is usually associated with decreased radiotracer uptake, due to myocardial perfusion defects or thinned walls in cardiomyopathy. Thus, one should expect that calculation of low LVEF by MPI is less reliable than by ERNA.

CLINICAL STUDIES ON REPRODUCIBILITY AND LIMITS OF AGREEMENT

The technical and biological reproducibility of LVEF derived from ERNA has been assessed previously.[5–8] These studies showed that, reprocessing the same acquisition data using (semi) automated software, technical intra- and inter-operator limits of agreement of LVEF, were very narrow (0.02-0.05 LVEF units) and were not affected by the level of LVEF.[5,6] Wackers et al. examined also the biological or temporal variation of ERNA LVEF by repeating image acquisition later on the same day and on different days.[5] The limits of agreement for LVEF derived from repeated data acquired on the same day or on different days were significantly larger for patients with LVEF $\geq$ 0.55 than for patients with LVEF $<$ 0.55. (0.13 and 0.05 for same day acquisitions and 0.14 and 0.06 for different day acquisitions, Figure 1). These observations concur with the greater contractile reserve of the normal heart compared to that of the diseased heart.

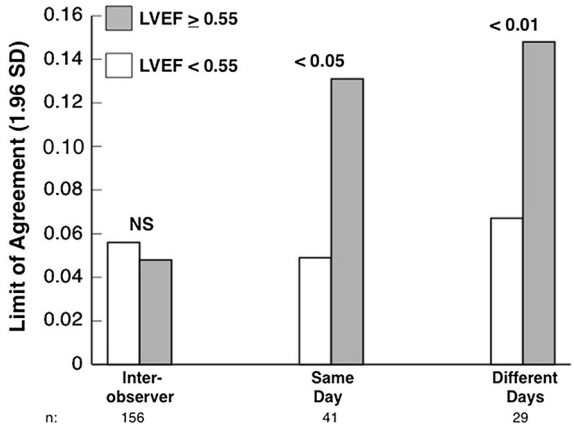

Figure 1. Limits of agreement (1.96 SD), or repeatability coefficients, of ERNA LVEF for inter-observer variability by processing the same studies, and for repeat patient studies acquired on the same day and on different days. The limits of agreement (in ejection fraction units) delineate the range of LVEF within which a difference could be due to chance (95% of differences being within 2 standard deviations). The repeatability coefficients of patients with normal (>0.55) and abnormal (<0.55) LVEF are compared. Although there is no difference in inter-observer repeatability, the limits of agreement of repeat studies acquired on the same day and on different days are significantly larger for normal LVEF than for abnormal LVEF. (modified from Wackers et al.[5]).

TECHNICAL LIMITATIONS OF ECG-GATED SPECT MPI

Vallejo et al.[9,10] investigated in experimental animals and in patients, imaging features that influenced agreement of MPI LVEF with LVEF derived from MRI and first pass RNA. Although automated computer analysis of gated MPI LVEF was highly reproducible with narrow limits of agreement, the accuracy of MPI LVEF was more affected by extra-cardiac and cardiac factors than ERNA LVEF. Specifically, limits of agreement for LVEF were adversely affected by high extra-cardiac background activity, low radiotracer dose, small size of left ventricle, and the presence of myocardial

perfusion defects. In each of these conditions insufficient counting statistics were responsible for suboptimal software performance and imprecise results.

In this issue of the *Journal,* Kliner et al.[11] tested the repeatability of MPI LVEF by repeat acquisitions with patient repositioning. A second acquisition was acquired within a few minutes after the completion of the first. Because of the short time interval, the investigators addressed predominately technical reproducibility rather than biological variation. The limit of agreement for MPI LVEF was 0.075 EF units; i.e. slightly larger than previously established for ERNA LVEF. In contrast to findings with ERNA LVEF, there was no significant difference in repeatability between normal and abnormal MPI LVEF (0.072 and 0.081, respectively).

It should be well understood that Kliner et al. did not truly address biological variation. In clinical practice repeat assessments of LVEF are usually done at intervals of weeks or months. Wackers et al. noted significant differences in biological repeatability of normal and abnormal LVEF when repeat ERNA imaging was performed at intervals of 1-5 days.[5] Although in Kliner's study the technical repeatability of normal and abnormal LVEF was not statistically different, the repeatability coefficient for abnormal LVEF was greater than for normal LVEF. As mentioned above, MPI studies with abnormal LVEF frequently have abnormal perfusion: areas with low count density and potentially problematic computer processing. The investigators did compare the limits of agreement of studies with and without myocardial perfusion defects and found no difference, but the number of patients was relatively small.

The study by Kliner et al. shows that technical reproducibility of LVEF derived from gated MPI is within acceptable limits. Unfortunately, biological variation, which is of clinical relevance, was not fully evaluated.

Because of the already mentioned technical imaging and processing problems of MPI-derived LVEF, particularly in patients with abnormal LV function, I believe that ERNA should be considered the methodology of choice for sequential LVEFs. Regrettably, many laboratories have abandoned regular use of ERNAs or have only limited experience. I doubt that many laboratories use MPI LVEF for monitoring patients on chemotherapy. The oncology literature generally recommends echocardiography for serial assessment of LVEF, while recognizing that ERNA LVEF has greater test-retest repeatability [12,13]. If MPI is to be used for monitoring changes in LVEF by serial assessments, further studies with focus on temporal and biological variation are needed before one can determine with confidence which LVEF change represents real change beyond chance.

References

1. Schneider J, Berger HJ, Sands MJ, Lachman AB, Zaret BL. Beat-to-beat left ventricular performance in atrial fibrillation: radionuclide assessment with the computerized nuclear probe. Am J Cardiol. 1983;51:1189–95.
2. LaVeau PJ, Rozanski A, Krantz DS, Cornell CE, Cattanach L, Zaret BL, Wackers FJTh. Transient left ventricular dysfunction during provocation mental stress in patients with coronary artery disease. Am Heart J. 1989;118:1–8.
3. Bland JM, Altman DG. Statistical method for assessing agreement between two methods of clinical measurement. The Lancet. 1986;327:307–10.
4. Wackers FJT. Characteristics of radiopharmaceuticals in nuclear cardiology. Implications for practical cardiac imaging. In: Simoons ML, Reiber JHC, editors. Nuclear Imaging in Clinical Cardiology (Chap 2). Boston: Springer; 1984.
5. Wackers FJTh, Berger HJ, Johnstone DE, Goldman L, Reduto LA, Langou RE, Gottschalk A, Zaret BL. Multiple gated cardiac blood pool imaging for left ventricular ejection fraction: validation of the technique and assessment of variability. Am J Card. 1979;43:1159–66.
6. Keough AM, Eberl S, Yeates MG, Freund J, Hickie JB, Esmore DS, Spratt PM. Reproducubility of radionuclide left ventricular ejection fraction in patients awaiting cardiac transplantation. Clin Nuc Med. 1990;15:25–8.
7. van Royen N, Jaffe CC, Krumholz HK, Johnson KM, Lynch PJ, Natale D, Atkinson P, Deman P, Wackers FJT. Comparison and reproducibility of visual echocardiographic and quantitative radionuclide left ventricular ejection fraction. Am J Cardiol. 1996;77:843–50.
8. Mahmarian JJ, Moye L, Verani MS, Eaton T, Francis M, Pratt GM. Criteria for the accurate interpretation of changes of left ventricular ejection fraction and cardiac volumes as assessed by rest and exercise gated radionuclide angiography. J Am Coll Cardiol. 1991;18:112–9.
9. Vallejo E, Dione DP, Bruni WL, Constable RT, Borek PP, Soares JP, Carr JG, Condos SG, Wackers FJTh, Sinusas AJ. Reproducibility and accuracy of gated SPECT for determination of left ventricular volumes and ejection fraction: experimental validation using MRI. J Nucl Med. 2000;41:874–82.
10. Vallejo E, Dione DP, Sinusas AJ, Wackers FJTh. Assessment of left ventricular ejection fraction with quantitative gated SPECT: accuracy and correlation with first-pass radionuclide angiography. J Nucl Cardiol. 2000;7:461–70.
11. Kliner D, Wang L, Winger D, Follansbee WP, Soman P. A prospective evaluation of the repeatability of left ventricular ejection fraction measurement by gated SPECT. J Nuc Cardiol 2015;this issue.
12. Ganz WI, Sridhar KS, Ganz SS, Gonzalez R, Chakko S, Serafini A. Review of test for monitoring doxorubicin-induced cardiomyopathy. Oncology. 1996;53:461–70.
13. Mookadam F, Sharma A, Lee HR, Northfelt DW. Intersection of cardiology and oncology clinical practices. Opinion article. Front Oncol. 2014;4:259. doi:10.3389/fonc.2014.00259.