Imaging and Diagnostic Testing

Visual estimation versus quantitative assessment of left ventricular ejection fraction: A comparison by cardiovascular magnetic resonance imaging

Burkhard Sievers, MD, Simon Kirchberg, Ulrich Franken, MD, Binu-John Puthenveettil, Asli Bakan, and Hans-Joachim Trappe, MD Herne, Germany

Purpose
The aim of this study was to compare the visual and quantitative assessment for left ventricular ejection fraction (LVEF) in normal subjects and patients with impaired LV function.

Methods
One hundred subjects (40 normal subjects, 40 patients with ischemic cardiomyopathy, and 20 patients with nonischemic cardiomyopathy) were investigated using a 1.5-T cardiovascular magnetic resonance imager. Images were acquired by a fast gradient-echo sequence with steady-state free precession using the standard short-axis method. Left ventricular EF was calculated from the sums of the outlined areas using the Simpson rule. Interobserver variability between the calculated and the visual EF was assessed. Analyses were performed randomly and blinded by 2 independent observers.

Results
Left ventricular EF was significantly underestimated by the visual read in all 3 groups (mean difference: normal subjects -2.6% ± 2.6%, ischemic cardiomyopathy -1.7% ± 2.1%, and nonischemic cardiomyopathy -1.2% ± 2.1%; P ≤ .02). The difference was larger in normal subjects than in patients with cardiomyopathy (P = .04). The interobserver variability was smaller for the quantitative assessment than for the visual estimation.

Conclusion
Left ventricular EF is underestimated by visual estimation compared with the quantitative assessment. The visual approach for EF assessment may be used for rapid assessment of left ventricular function in clinical practice where accuracy is of less concern. For most accurate analysis, the quantitative standard short axis approach is required. (Am Heart J 2005;150:737-42.)

Cardiovascular magnetic resonance imaging (CMR) allows accurate and reproducible evaluation of ventricular function and determination of ventricular volumes and ejection fraction (EF). Because of rapid technical advances in the field of CMR, image acquisition with fast gradient-echo sequences has become possible. The advantages are shorter breath-hold periods and greater temporal and spatial resolution, which results in better blood-myocardium contrast and greatly facilitates the identification of the ventricular endocardium and epicardium.

For EF assessment, the left ventricle is imaged in slices at fixed intervals from the base (atrioventricular ring) to the apex. The surface contours of the endocardium and epicardium are manually delineated and the ventricular volumes derived by summation with the assistance of computer software. However, quantitative assessment of left ventricular (LV) EF is still time-consuming and often not applicable in daily clinical routine.

Automatic contour detection for rapid ventricular EF assessment would be of great practical value but has not yet been perfected and is still unreliable in the analysis of gradient-echo images.

The area-length method has been shown to be acceptable accurate for rapid LVEF assessment, but although faster than the standard short-axis approach, it still takes additional time to measure the dimensions from long-axis planes and to calculate the EF.

Visual assessment of global and regional cardiac function can easily be performed during the scan procedure or between scans and does not take additional time.

We therefore evaluated whether the visual assessment of the global LVEF provides an acceptable degree of accuracy in normal subjects and patients with impaired LV function.
Materials and methods

A total of 100 subjects (40 subjects with no cardiovascular disease, 40 patients with ischemic cardiomyopathy, and 20 patients with nonischemic cardiomyopathy) underwent CMR for the evaluation of cardiac function and the determination of LVEF.

In the 40 normal subjects, heart disease was excluded before CMR by noninvasive diagnostic techniques (electrocardiography [ECG], chest x-ray, treadmill exercise ECG, echocardiography, and thallium myocardial scintigraphy). None of them had a history of hypertension or diabetes.

Of the remaining 60 patients, 40 patients had a history of myocardial infarction (MI). The MI definition was based on ECG, blood sampling (troponin I, creatine kinase [CK], and CK-MB), and catheter angiography. Acute MI was diagnosed if the following criteria were present: ST-segment elevation or ST-segment depression with T-wave inversion on ECG, high troponin levels (>2 ng/mL), and CK-MB levels ≥10% of total CK.

Written consent was obtained in all cases. The study was conducted according to the principles of the Declaration of Helsinki and approved by the local institutional review board.

Image acquisition

Cardiovascular magnetic resonance imaging was performed with a 1.5-T scanner (Sonata, Magnetom, Siemens, Erlangen, Germany) using an anterior and posterior surface coil array (CP Body Array Flex, CP Spine Array, Siemens) and prospective ECG triggering. The dimensions of the coil elements are about 160 mm in the x-direction (head to feet) and about 460 mm in the x-direction (right to left). A fast imaging sequence with steady-state free precession (SSFP) and constant radiofrequency pulsing was used.

On the basis of scout images, cine images were acquired in the short axis and horizontal and vertical long axes. Short-axis images were acquired from the base of the heart (atrioventricular ring) to the apex with 10-mm-thick contiguous slices during breath holding in end expiration. The number of cardiac phases per image acquisition totaled 80% to 90% of the R-R interval divided by the temporal resolution (39 milliseconds). Seven to 11 slices were necessary to cover the entire ventricle. The following parameters were employed: repetition time = 3.2 milliseconds, echo time = 1.6 milliseconds, slice thickness = 10 mm, no interslice gap, flip angle = 60°, in-plane pixel size = 2.3 × 1.4 mm, matrix 175 × 256 pixel, trigger pulse 1, trigger delay 0, acquisition time = 12 heartbeats, and breath-hold duration per slice = 9 to 12 seconds, depending on the heart rate.

Image analysis

Images were analyzed in a random order with a commercially available computer software program Argus (Siemens) by 2 experienced independent observers (B.S. and S.K.). At the time of the study, B.S. had 2 years and S.K. 1.5 years experience in performing and analyzing clinical magnetic resonance imaging (CMR) studies (750 cases a year). All patient identifiers and image parameters were removed from the images before analysis. The observers were blinded to the quantitative results when the visual reads were performed.

For quantitative analysis, published criteria were used for the delineation of cardiac borders. Papillary muscles and trabeculations were excluded for volume measurements. Contour tracing was aided by reviewing the multiple phase scans in the cine mode. For LV volume assessment, end diastole was defined as the first phase in the ECG-gated data set, and end systole as the phase with the smallest volume (the cardiac phase with minimal cross-sectional diameter). End systole was defined as the same phase throughout the data set. At the base of the heart, slices were considered to be in the left ventricle if the blood was at least half-surrounded by ventricular myocardium. For the basal slice, the contours were traced up to the junction of the atrium and the ventricle. Blood volume up to the aortic valve was included in the LV volume. The epicardium and endocardium of the left ventricle were marked with a cursor in each end-diastolic and end-systolic slice. Ventricular end-diastolic volume (EDV) and end-systolic volume ( ESV) were calculated from the sums of the outlined areas using the Simpson’s rule. Stroke volumes (SV) and EFs were calculated from the following formulae: SV = EDV − ESV and EF = SV/EDV × 100%

For the visual assessment, short-axis and long-axis (horizontal and vertical) images were used. Ejection fraction was visually estimated to the nearest 5% (eg, 40%, 45%, etc).

Interobserver variability of the measurements was assessed for both methods.

Statistical analysis

Mean and SD were derived for each of the parameters. The differences in calculated volumes and EF using different gating methods were assessed using paired t test. All statistical tests were 2-tailed, P value of <.05 was regarded as significant. Interobserver variability was defined with the formula (observer A−observer B)/(mean observer A and B). Bland-Altman plots were performed from the measurements obtained by the 2 observers.

Results

Baseline characteristics are given in Table I. Normal subjects were significantly younger than patients with cardiomyopathy (P < .0001). LVEF between ischemic and nonischemic cardiomyopathy was not significantly different (P = .09).

The LVEF was significantly underestimated by the visual read in all 3 groups (mean difference: normal subjects −2.6% ± 2.6%, ischemic cardiomyopathy −1.7% ± 2.1%, nonischemic cardiomyopathy −1.2% ± 2.1%; P ≤ .02) (Figure 1). The difference between the visually estimated and calculated EF was larger in normal subjects than in patients with cardiomyopathy (ischemic and nonischemic), P = .04. The difference in EF between normal subjects and patients with ischemic cardiomyopathies (P = .11) and among patients with cardiomyopathy (P = .34) was not significant. The Bland-Altman plots of the agreement between the visually estimated and the calculated EF are displayed in Figure 2.

The correlation coefficient for normal subjects was 0.92, for ischemic cardiomyopathies 0.96, and for nonischemic cardiomyopathies 0.95.
The interobserver variability was small for the calculated EF. For the visually estimated EF, there was a wider range of the 95% limits of agreement (Table II). The reproducibility was consistent over the entire range of EF values (Table III).

Measurements obtained with the standard short-axis method required 23 ± 5 minutes. The visual estimation required <1 minute per patient (30 ± 10 seconds).

**Discussion**

Cardiovascular magnetic resonance imaging enables accurate and reproducible assessment of ventricular function and determination of their volumes and EF by using the standard short-axis method. A fast gradient-echo sequence with SSFP that has recently been introduced provides improved contrast between blood and myocardium and thus more precise delineation of the endocardium and epicardium. This sequence is currently widely used for volumetric and EF assessment.

Because the papillary muscles and trabeculations have to be marked manually and their volumes subtracted from the total ventricular volume in each slice, the standard short-axis technique is very time-consuming and thus often not applicable for use in everyday clinical practice. It has been shown that not sub-

**Table I. Baseline characteristics**

<table>
<thead>
<tr>
<th></th>
<th>Normal subjects (N = 40)</th>
<th>Ischemic CMP (N = 40)</th>
<th>Nonischemic CMP (N = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (women/men)</td>
<td>12/28</td>
<td>12/28</td>
<td>6/14</td>
</tr>
<tr>
<td>Age (y)</td>
<td>53.9 ± 13.2</td>
<td>64.5 ± 7.7</td>
<td>62.2 ± 8.1</td>
</tr>
<tr>
<td>EDV (mL)*</td>
<td>126.5 ± 24.5</td>
<td>178.4 ± 35.0</td>
<td>195.4 ± 42.1</td>
</tr>
<tr>
<td>ESV (mL)*</td>
<td>39.1 ± 10.9</td>
<td>118.6 ± 23.2</td>
<td>136.1 ± 39.9</td>
</tr>
<tr>
<td>EF (%)*</td>
<td>69.3 ± 6.3</td>
<td>33.6 ± 7.1</td>
<td>30.6 ± 6.1</td>
</tr>
</tbody>
</table>

*Calculated from the standard short-axis method; average values of both observers. Data are displayed as mean and SD.

**Figure 1**

Mean and SD of calculated and visually estimated LVEF (average values of the 2 observers) in normal subjects, ischemic cardiomyopathy and nonischemic cardiomyopathy, displayed as boxplots. Ejection fraction differs significantly between the 2 methods in either group of subjects. CMP, ischemic cardiomyopathy.
tracting the area of the trabeculations and papillary muscles for ventricular volume assessment leads to significant differences in volumes for patients with and without coronary heart disease, compared to the standard short-axis method in which trabeculations and papillary muscles are excluded from the blood pool. However, no differences were found for the evaluation of EF, and the modified short-axis method could be performed in about half the time (13 ± 3 minutes). Nevertheless, in CMR centers where 8 to 10 scans per day are routinely performed, 13 ± 3 minutes per patient for LVEF assessment would add 2 hours for EF analysis.

Automatic contour detection would be of great practical value but has not yet been perfected and is still unreliable. Manual correction of automatically detected contours often takes nearly as long as drawing the contours manually. The epicardial borders are more difficult to detect than the endocardial borders because of the characteristics of the adjacent anatomical structures. In the study of Plein et al, 10% of the automatically detected endocardial borders and 40% of the automatically detected epicardial borders required manual correction.

Baldy et al reported a time of 4 to 8 minutes for the automatic delineation of the contours, assuming accurate border detection. However, automatic contour detecting is currently unable to exclude the papillary muscles from volumes and therefore does not provide most accurate volumes.

Mitchell et al recently introduced a 3-dimensional (3D) active appearance model for image segmentation. They showed good agreement for LV volumes and EF between their model and the standard approach. Although their method is promising, it is not yet implemented in clinical practice.

Swingen et al recently proposed a feedback-assisted method for calculating a 3D model of the left ventricle from images acquired with arbitrary slice orientations. Left ventricular volumes and EF calculated from the 3D model showed good agreement compared to the

---

Bland-Altman plots Bland-Altman analysis of the difference between visually estimated and calculated LVEF in normal subjects, patients with ischemic cardiomyopathy and patients with nonischemic CMP.
standard short-axis method with manually drawn contours (mean difference for LVEF 4.22% ± 9.16%). Approximately 5 to 10 minutes were required to adequately estimate the LV volumes and calculate the EF using the new approach compared with 15 to 20 minutes for the standard short-axis method. The authors did not exclude the papillary muscles for volume measurements when using the standard short-axis method.16 However, 5 to 10 minutes still requires about 1.5 hours of additional time for EF analysis in CMR centers that routinely perform 8 to 10 scans per day.

Although the biplane area-length method for LV volume and EF assessment has been shown to be acceptably accurate,8 there is no automated algorithm for EF calculation by this method available. Measuring dimensions and areas manually from different image planes and calculating volumes and EF also takes additional time.

In the view of the fact that automatic contour detection still has technical limitations and is not yet commercially available, the visual approach for LVEF assessment might be considerable for a rapid evaluation of the LV function.

We therefore addressed the question of whether the visually estimated EF is reasonably accurate for LVEF evaluation. We compared the results of the visual estimation with the quantitative assessment using the standard short-axis method as the gold standard.

We found the LVEF to be underestimated by the visual approach for LVEF in normal subjects and patients with impaired ventricular function using echocardiography and radionuclide angiography showed moderate to good correlation between visual and calculated EF.17-21 Their correlation coefficient values ranged from 0.72 to 0.89.15-19 With CMR, we could achieve a better correlation between visual and quantitative assessment (correlation coefficient ≥0.92). However, a good correlation does not necessarily imply good agreement. The limits of agreement between echo and radionuclide angiography in the study of Shih et al18 were even wider (23.6%) than in our study (≤10.4%). We found a good reproducibility for the calculated LVEF, whereas the reproducibility of the visually estimated EF was poorer. The range of the 95% limits of agreement was much wider for the visual approach (Figure 2). However, the reproducibility of a visual estimation is expected to be poorer than the reproducibility of a quantitative approach. The wider range might also be explained in part as a result of the fact that the visual EF was estimated to the nearest 5%.

We found a good reproducibility for the calculated LVEF, whereas the reproducibility of the visually estimated EF was poorer. The range of the 95% limits of agreement was much wider for the visual approach (Figure 2). However, the reproducibility of a visual estimation is expected to be poorer than the reproducibility of a quantitative approach. The wider range might also be explained in part as a result of the fact that the visual EF was estimated to the nearest 5%. Nevertheless, the reproducibility was consistent over the entire range of EF values (Table III).

The visual assessment for EF might be a considerable approach for clinical scenarios where EF accuracy is of less concern and the time for quantitative analysis is limited. In patients where clinical decisions require most accurate EF measurements, the standard short-axis method should be performed.

Limitations of the study

Our study focused on LVEF assessment. We did not address left and right ventricular volumes and right ventricular EF. The assessment of right ventricular

<table>
<thead>
<tr>
<th>Table II.</th>
<th>Interobserver variability for left ventricular ejection fraction (%) calculation</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calculated EF</strong></td>
<td><strong>Visual EF</strong></td>
<td><strong>95% Limits of agreement</strong></td>
<td><strong>95% Limits of agreement</strong></td>
</tr>
<tr>
<td>Normal subjects</td>
<td>-0.23 ± 1.6</td>
<td>0.75 ± 5.6</td>
<td>-3.5, 3.0</td>
</tr>
<tr>
<td>Ischemic CMP</td>
<td>0.1 ± 1.6</td>
<td>-0.37 ± 6.7</td>
<td>-3.2, 3.4</td>
</tr>
<tr>
<td>Nonischemic CMP</td>
<td>-0.28 ± 1.4</td>
<td>-0.25 ± 5.0</td>
<td>-3.2, 2.6</td>
</tr>
<tr>
<td><strong>Mean difference</strong></td>
<td><strong>Mean difference</strong></td>
<td><strong>P</strong></td>
<td><strong>P</strong></td>
</tr>
<tr>
<td>Normal subjects</td>
<td>-0.23 ± 1.6</td>
<td>0.75 ± 5.6</td>
<td>39</td>
</tr>
<tr>
<td>Ischemic CMP</td>
<td>0.1 ± 1.6</td>
<td>-0.37 ± 6.7</td>
<td>71</td>
</tr>
<tr>
<td>Nonischemic CMP</td>
<td>-0.28 ± 1.4</td>
<td>-0.25 ± 5.0</td>
<td>22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table III.</th>
<th>Interobserver variability for left ventricular ejection fraction (%) calculation</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calculated EF</strong></td>
<td><strong>Visual EF</strong></td>
<td><strong>95% Limits of agreement</strong></td>
<td><strong>95% Limits of agreement</strong></td>
</tr>
<tr>
<td>EF</td>
<td>N</td>
<td><strong>Mean difference</strong></td>
<td><strong>P</strong></td>
</tr>
<tr>
<td>&lt;25%</td>
<td>9</td>
<td>0.2 ± 1.7</td>
<td>0.39</td>
</tr>
<tr>
<td>25%-35%</td>
<td>23</td>
<td>-0.2 ± 1.4</td>
<td>39</td>
</tr>
<tr>
<td>35%-60%</td>
<td>31</td>
<td>0.1 ± 1.7</td>
<td>28</td>
</tr>
<tr>
<td>≥60%</td>
<td>37</td>
<td>-0.3 ± 1.6</td>
<td>23</td>
</tr>
</tbody>
</table>

Analysis by EF thresholds.

Studies comparing visual and quantitative assessment for LVEF in normal subjects and patients with impaired ventricular function using echocardiography and radionuclide angiography showed moderate to good correlation between visual and calculated EF.17-21 Their correlation coefficient values ranged from 0.72 to 0.89.15-19 With CMR, we could achieve a better correlation between visual and quantitative assessment (correlation coefficient ≥0.92). However, a good correlation does not necessarily imply good agreement. The limits of agreement between echo and radionuclide angiography in the study of Shih et al18 were even wider (23.6%) than in our study (≤10.4%).

We found a good reproducibility for the calculated LVEF, whereas the reproducibility of the visually estimated EF was poorer. The range of the 95% limits of agreement was much wider for the visual approach (Figure 2). However, the reproducibility of a visual estimation is expected to be poorer than the reproducibility of a quantitative approach. The wider range might also be explained in part as a result of the fact that the visual EF was estimated to the nearest 5%.

We therefore addressed the question of whether the visually estimated EF is reasonably accurate for LVEF evaluation. We compared the results of the visual estimation with the quantitative assessment using the standard short-axis method as the gold standard.

We found the LVEF to be underestimated by the visual approach for LVEF in normal subjects and patients with impaired ventricular function using echocardiography and radionuclide angiography showed moderate to good correlation between visual and calculated EF.17-21 Their correlation coefficient values ranged from 0.72 to 0.89.15-19 With CMR, we could achieve a better correlation between visual and quantitative assessment (correlation coefficient ≥0.92). However, a good correlation does not necessarily imply good agreement. The limits of agreement between echo and radionuclide angiography in the study of Shih et al18 were even wider (23.6%) than in our study (≤10.4%).

We found a good reproducibility for the calculated LVEF, whereas the reproducibility of the visually estimated EF was poorer. The range of the 95% limits of agreement was much wider for the visual approach (Figure 2). However, the reproducibility of a visual estimation is expected to be poorer than the reproducibility of a quantitative approach. The wider range might also be explained in part as a result of the fact that the visual EF was estimated to the nearest 5%.

Nevertheless, the reproducibility was consistent over the entire range of EF values (Table III).

The visual assessment for EF might be a considerable approach for clinical scenarios where EF accuracy is of less concern and the time for quantitative analysis is limited. In patients where clinical decisions require most accurate EF measurements, the standard short-axis method should be performed.

Limitations of the study

Our study focused on LVEF assessment. We did not address left and right ventricular volumes and right ventricular EF. The assessment of right ventricular
volumes and right ventricular EF is important in patients with cardiomyopathy and valvular heart disease. In our opinion, right ventricular volumes and right ventricular EF cannot be reliably assessed by visual estimation and, if requested, require quantitative analysis. The triangular shape of the right ventricle makes visual EF assessment very difficult. However, visual differentiation between normal right ventricular function and mild, moderate, or severe right ventricular dysfunction is feasible and might be sufficient for clinical practice.

It remains for further studies to determine whether the visual approach can also be used in patients with marked myocardial hypertrophy and hypertrophic cardiomyopathy.

We thank Michele A. Parker, MS, for her help with the statistics.

References