Operator Induced Variability in Left Ventricular Measurements with Cardiovascular Magnetic Resonance is Improved After Training

Theodoros D. Karamitsos, MD, Lucy E. Hudsmith, MRCP, MA, BM, BCh, Joseph B. Selvanayagam, DPhil, FRACP, FESC, Stefan Neubauer, MD, FRCP, and Jane M. Francis, DCR(R), DNM

University of Oxford Centre for Clinical Magnetic Resonance Research, Department of Cardiovascular Medicine, John Radcliffe Hospital, Oxford, UK

ABSTRACT

Background: Accurate and reproducible measurement of left ventricular (LV) mass and function is a significant strength of Cardiovascular Magnetic Resonance (CMR). Reproducibility and accuracy of these measurements is usually reported between experienced operators. However, an increasing number of inexperienced operators are now training in CMR and are involved in post-processing analysis. The aim of the study was to assess the interobserver variability of the manual planimetry of LV contours amongst two experienced and six inexperienced operators before and after a two months training period.

Methods: Ten healthy normal volunteers (5 men, mean age 34 ± 14 years) comprised the study population. LV volumes, mass, and ejection fraction were manually evaluated using Argus software (Siemens Medical Solutions, Erlangen, Germany) for each subject, once by the two experienced and twice by the six inexperienced operators. The mean values of experienced operators were considered the reference values. Training involved standardized data acquisition, simulated off-line analysis and mentoring.

Results: The trainee operators demonstrated improvement in the measurement of all the parameters compared to the experienced operators. The mean ejection fraction variability improved from 7.2% before training to 3.7% after training (p = 0.03). The parameter in which the trainees showed the least improvement was LV mass (from 7.7% to 6.7% after training). The basal slice selection and contour definition were the main sources of errors.

Conclusions: An intensive two month training period significantly improved the accuracy of LV functional measurements. Adequate training of new CMR operators is of paramount importance in our aim to maintain the accuracy and high reproducibility of CMR in LV function analysis.

INTRODUCTION

Accurate and reproducible quantification of left ventricular (LV) volumes, function, and mass is important for both clinical practice and research, especially when follow-up assessment is needed (1–6). With regard to measurement of global cardiac function and given its 3D nature, CMR is superior to 2D echocardiography, invasive angiography, and radionuclide angiography, and is now considered the preferred imaging technique (7). In the setting of clinical research, CMR has allowed reductions of study sizes of 80–97% to achieve the same statistical power for demonstrating given changes of LV volumes, ejection fraction, or cardiac mass (8–10).

Reproducibility and accuracy of LV function and mass measurements with CMR is usually reported between experienced operators (11–13). However, as CMR is being increasingly applied in both clinical practice and clinical research, an increasing number of operators, with varying prior experience in cardiac imaging/physiology, are training in CMR. Post-processing analysis of LV volumes and mass forms an important component of this training. Only the application of high standards in training will maintain the advantages of CMR, in terms of both accuracy and reliability.
and reproducibility. The aim of the present study was to investigate the effect of formal CMR training on the assessment of LV volumes and mass on previously inexperienced CMR operators.

METHODS

Study population

Ten healthy controls (5 male and 5 female, mean age 34 ± 14, mean height 174 ± 7 cm, mean weight 74 ± 13 kg, and mean heart rate 65 ± 10 bpm) with no history of cardiac disease, hypertension, or cardiac risk factors, and a normal baseline electrocardiogram (ECG) were recruited. The study was carried out according to the principles of the Declaration of Helsinki and was approved by our institutional ethics committee. Each subject gave informed written consent.

CMR operators

The 6 new CMR operators included a fully trained radiographer with no previous cardiology experience, 2 junior doctors without previous experience in cardiac imaging, 2 cardiology trainees with 1–2 years experience in echocardiography, and a junior consultant cardiologist. All of the trainees were initially taught how to perform LV analysis and familiarized themselves with the analysis software. Each then analysed the 10 healthy volunteer scans, repeating the analysis after a 2 month intensive training period. The two experienced operators were a senior radiographer with more than 10 years of CMR experience and a cardiology registrar with more than two years of CMR experience; the mean values of their measurements were considered as the reference values.

Image acquisition and analysis

All CMR examinations were performed on a 1.5 Tesla MR scanner (Sonata, Siemens Medical Solutions, Erlangen, Germany) with spine coil and phased array surface coil, prospective electrocardiographic gating and the patient in the supine position. After piloting using localizers, a horizontal long-axis, vertical long-axis and short-axis end-diastolic pilots, steady-state free precession cine images (TE/TR 1.5/3.0 ms, flip angle 60°, slice thickness 7 mm, 3 mm inter-slice gap, in-plane resolution 1.5 x 1.5 mm², temporal resolution 45 ms, breathhold duration of 14–17 heartbeats per breathhold) were acquired in the horizontal and vertical long axis views during end-expiration. The short axis stack was then obtained, parallel to the atrioventricular groove, covering the entire left and right ventricle in the usual manner (7).

CMR image analysis was performed with Argus software (version 25A; Siemens Medical Solutions, Erlangen, Germany) by all investigators. Manual tracing of the endocardial and epicardial borders of successive short-axis slices was performed at end-diastole and end-systole (phase with the image with the smallest LV cavity) (12). Epicardial and endocardial borders were traced on the end-diastolic frame, with only an endocardial border on the end-systolic frame. The instructions given to the inexperienced operators were to select the basal slice for the left ventricle when at least 50% of the blood volume was surrounded by myocardium in both end-diastole and end-systole. The apical slice was defined as the final slice showing intracavity blood pool at both end-diastole and end-systole. Operators were free to select the end-systolic and end-diastolic frame. Papillary muscles in the midventricular level and trabeculations in the apex were included in the mass and the volume calculations. From these data, the mass, ejection fraction, end-systolic volumes, and end-diastolic volumes could be calculated. Myocardial mass was determined from the end diastolic images by multiplication of the tissue volume by 1.05 g/cm³ (specific density of myocardium).

Training period

The 2 month training period included the following: 1) didactic lectures on the physics of magnetic resonance, on electrocardiogram gating-triggering and on safety issues in the CMR environment; 2) hands-on-experience on imaging of cardiac anatomy and function (including cine steady state free precession imaging - SSFP) in normal volunteers and patients; 3) mentored CMR image analysis using post-processing tools (Argus software, version 25A; Siemens Medical Solutions, Erlangen, Germany); and 4) hands-on-experience on post-processing analysis mainly focusing on the measurement of LV function parameters and mass. Each trainee operator performed supervised analysis of a minimum of 25 scans during the training period. Additionally, all trainees participated in the weekly clinical case reading sessions and were encouraged to actively participate in the daily scanning timetable.

Statistical analysis

The agreement between each inexperienced operator and the experienced operators was assessed by means of Bland-Altman analysis (14). The coefficient of variability was calculated as the standard deviation (SD) of the differences between the two sets of measurements divided by the mean. The Wilcoxon test was used to compare the variabilities of the measured parameters by the trainees before and after training. Spearman’s rank correlation coefficient (Rs) was used to assess the simple correlation between the end systolic basal slice selection and the ESV measurement differences between the trainees and the experienced operators’ measurements. All computations were performed with SPSS 13.0 (SPSS Inc., Chicago, Illinois, USA).

RESULTS

Variability in measured parameters

The interobserver variability of the LV measurements between the experienced operators were 2.9% for EF, 2.6% for EDV, 6.9% for ESV, 3.4% for SV and 5.8% for LV mass. The coefficients of variability for the LV measurements made by the inexperienced operators before and after training are shown in Table 1.

Initially, the coefficient of variability for LV ejection fraction (EF) measurements by the trainees ranged from 4.4%
to 10%, with a mean of 7.2%. After the training period, all trainees showed a significant improvement compared to the experienced operators measurements resulting in coefficients of variability of less than 5%, with a mean of 3.7% (p = 0.03) (Fig. 1).

Prior to training, the variability of the trainees’ measurements for LV end-diastolic volumes ranged between 3.6% to 8.6% (mean of 5.4%), with most operators showing improvement after the two-month training (mean of 4.6%, p = 0.4). The same effect was seen with LV stroke volume (SV) (mean variability from 9.2% before training to 7.2%, p = 0.16).

Measurement of ESV showed high variability that ranged from 13.5% to 19% (mean of 16.9%) prior to training, which after 2 months of training showed a significant improvement for most operators (mean of 7.4%, p = 0.03).

Perhaps surprisingly, formal CMR training did not seem to affect measurement of LV mass; half of the trainee operators showed improved variability after training while the others showed, albeit small, worsening of their results (mean variability from 7.7% to 6.7% after training, p = 0.7).

### Sources of errors

There were several potential sources of errors that resulted in the marked variability in the measured parameters. As shown in Fig. 2, the selection of the end systolic basal slice correlated significantly with the differences in measurements (errors) between the trainees and the experienced operators measurements ($Rs = -0.58$, $p = 0.001$). The selection of an extra basal slice resulted in an overestimation of ESV by on average 10.2 ± 5.1 mL, whereas one less basal slice underestimated the ESV by on average of 6.3 ± 7.6 mL. Interestingly, even when the correct end-systolic basal slice was selected by the trainees, there was an average difference of 1.3 ± 4.2 mL compared to the experienced operators ESV measurements, a finding that reflects the difficulty in endocardial border detection in end systole. Furthermore, the importance of the selection of the correct basal slice is also reflected in EDV measurements since there was a weak, but significant, correlation ($Rs = -0.22$, $p = 0.017$) between the end diastolic slice selection and the differences in calculation (errors) of EDV made by the trainees and the experienced operators. Finally, LV mass measurements represent the cumulative effects of errors in both epicardial and endocardial contour definition. Importantly, contour definition between the epicardium and the lung can be difficult in subjects with minimal pericardial fat.

Fig. 3 shows representative examples of the improvement in contour definition and basal slice selection that the trainees showed after training.

### DISCUSSION

This study showed that an intensive two-month training period significantly improved the reproducibility of LV functional measurements by previously inexperienced CMR operators, although the effect on LV mass measurements was less marked.

### Table 1. Trainee operators’ variability of left ventricular parameters before and after training

<table>
<thead>
<tr>
<th>OPERATOR</th>
<th>EF bias</th>
<th>EF SD</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Coefficient</th>
<th>EDV bias</th>
<th>EDV SD</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Coefficient</th>
<th>ESV bias</th>
<th>ESV SD</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Coefficient</th>
<th>SV bias</th>
<th>SV SD</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Coefficient</th>
<th>MASS bias</th>
<th>MASS SD</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPERATOR 1</td>
<td>-2.01</td>
<td>3.04</td>
<td>-8.05</td>
<td>3.88</td>
<td>4.4%</td>
<td>2.97</td>
<td>6.62</td>
<td>-10.01</td>
<td>15.96</td>
<td>4.6%</td>
<td>3.86</td>
<td>5.98</td>
<td>-7.86</td>
<td>15.58</td>
<td>13.5%</td>
<td>-0.86</td>
<td>4.34</td>
<td>-9.37</td>
<td>7.64</td>
<td>4.4%</td>
<td>-19.32</td>
<td>11.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPERATOR 2</td>
<td>-0.99</td>
<td>1.43</td>
<td>-3.79</td>
<td>1.81</td>
<td>2.04%</td>
<td>0.03</td>
<td>2.96</td>
<td>-5.78</td>
<td>5.85</td>
<td>2.1%</td>
<td>5.81</td>
<td>8.93</td>
<td>-11.70</td>
<td>5.93</td>
<td>5.6%</td>
<td>-1.18</td>
<td>3.28</td>
<td>-7.62</td>
<td>5.25</td>
<td>3.4%</td>
<td>8.32</td>
<td>4.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPERATOR 3</td>
<td>-5.89</td>
<td>6.83</td>
<td>-19.28</td>
<td>7.49</td>
<td>10.1%</td>
<td>-4.51</td>
<td>11.98</td>
<td>-28.00</td>
<td>18.97</td>
<td>8.8%</td>
<td>5.17</td>
<td>2.55</td>
<td>-11.20</td>
<td>13.32</td>
<td>19.7%</td>
<td>-0.35</td>
<td>1.68</td>
<td>-16.38</td>
<td>13.67</td>
<td>12.9%</td>
<td>-14.62</td>
<td>7.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPERATOR 4</td>
<td>-0.56</td>
<td>2.91</td>
<td>-6.27</td>
<td>5.44</td>
<td>4.1%</td>
<td>1.30</td>
<td>7.83</td>
<td>-9.56</td>
<td>19.44</td>
<td>5.5%</td>
<td>2.55</td>
<td>7.94</td>
<td>-15.91</td>
<td>6.77</td>
<td>6.0%</td>
<td>-1.45</td>
<td>1.68</td>
<td>-20.25</td>
<td>21.06</td>
<td>12.9%</td>
<td>-11.81</td>
<td>8.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPERATOR 5</td>
<td>-0.87</td>
<td>5.65</td>
<td>-11.15</td>
<td>3.32</td>
<td>8.0%</td>
<td>-0.49</td>
<td>5.54</td>
<td>-13.45</td>
<td>12.17</td>
<td>3.9%</td>
<td>0.98</td>
<td>2.71</td>
<td>-15.39</td>
<td>6.29</td>
<td>7.1%</td>
<td>0.58</td>
<td>0.29</td>
<td>-15.02</td>
<td>12.39</td>
<td>18.7%</td>
<td>-6.07</td>
<td>8.99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPERATOR 6</td>
<td>0.025</td>
<td>2.10</td>
<td>-4.93</td>
<td>9.61</td>
<td>6.9%</td>
<td>-0.10</td>
<td>6.61</td>
<td>-10.04</td>
<td>9.83</td>
<td>2.9%</td>
<td>5.59</td>
<td>7.53</td>
<td>-9.34</td>
<td>6.29</td>
<td>17.9%</td>
<td>2.73</td>
<td>4.27</td>
<td>-15.30</td>
<td>11.49</td>
<td>16.7%</td>
<td>-13.96</td>
<td>6.67</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Training in CMR and LV Function Reproducibility Measurements

779
The interobserver variability results seen in our study between the trainees (after training) and the experienced operators are comparable with the reproducibility in LV measurements of previously published studies (8, 11, 15–18).

To our knowledge, this is the first study that addressed the influence of training in LV function and mass measurements by CMR. Importantly, the 6 inexperienced operators that participated in our study showed excellent improvement after the training in the quantification of LV ejection fraction, end-diastolic volume, and stroke volume, with coefficients of variation that were consistent with the ones reported from highly experienced operators in the literature. With reference to the measurement of LV ejection function, Moon et al (16) reported an interobserver variability of 6% whilst Danilouchkine et al (11) presented a coefficient of variation of 3%. An important note is that in our study the mean variability for LV ejection fraction measurements after training was 3.7%.

Similar improvement was noted on the measurement of LV volumes; with the published interobserver variability in EDV measurements ranging from 1.8% to 4.1%, our trainee operators showed an acceptable variability of 4.6%. However, the variability in SV and ESV measurements, although improved after training, was higher compared to the variability in ejection fraction and EDV. It is noteworthy that previous studies addressing the issue of the reproducibility in LV measurements do not report variability results on SV measurements. Nevertheless, we feel that the coefficient of variation of 7.2% achieved by the trainee operators after training is an acceptable result, given that the quantification of SV incorporates errors from measurement of both EDV and ESV. Our finding of a mean 7.4% in ESV variability is consistent with the previously published data. Using either SSFP or FLASH imaging, Moon et al (16)
found high interobserver variability (8 and 10%, respectively) in ESV measurements, while Hudsmith et al (18) reported an interobserver variability of 6.2%. A possible explanation for the variability seen in ESV measurements is the differences in the selection of the basal end systolic slice that resulted in either overestimation (if an extra slice was included) or underestimation (if one less slice was analysed). Notwithstanding this, even when the correct end systolic slice was selected, there were still errors in ESV measurements, possibly because of the greater difficulty in the determination of the endocardial border in end systole.

The measurement of LV mass was another parameter that showed marked variability despite the training period. Although there was an improvement in LV mass measurement with a coefficient of variation of 6.7% after training, this is still higher than the results that Danilouchkine at al (4.3%), Hudsmith et al (5.2%) or Moon et al (5.8%) reported (11, 16, 18). This finding can be explained by the fact that mass measurements require the definition of both epicardial and endocardial borders in end-diastole, hence increasing the possibility of errors. Another potential source of error in mass measurements derives from the inclusion of epicardial fat in measurements whereas an incorrect basal slice selection also contributes to the marked variability in measurements.

Previous experience in cardiovascular imaging, including echocardiography, is essential for CMR application. Operators with previous echocardiography experience (Operators 1, 4 and 6) showed a lower variability in measurements before training than the trainee without any previous experience in cardiovascular imaging (Operator 2). However, after a 2 month full-time training, even this operator resulted in significant improvement in most of the measured LV parameters consistent with the literature for EF and LV volumes.

Figure 3. Representative examples of errors made by the trainee operators. Panels A (before training) and B (after training) show the improvement in basal end systolic slice selection. Panels C (before training) and D (after training) show the improvement in epicardial and endocardial contour definition (end diastolic images).
Given the expansion of CMR within cardiology, it is of paramount importance that the advantages that CMR affords of accurate LV measurements, increased reproducibility for serial measurements and in response to therapeutic intervention as well as reduced sample sizes for clinical trials, are not lost by inadequate training of new trainees. Moreover, the need for accurate estimation of LV function is further supported in the modern era when important and expensive therapeutic procedures, such as the implantation of defibrillators or biventricular pacemakers in patients with heart failure, are guided from the determination of LV function (19). To achieve this goal, it is important to implement an adequate training program for new CMR trainees to ensure the standardization of both image acquisition and post-processing analysis. From our study, it is evident that 2 months of extensive exposure to CMR are adequate to maintain a low variability, at least for the analysis of healthy volunteers, although for mass measurements further training seems to be needed.

A possible limitation of our study is that we report the variability in measurements from scans only in healthy volunteers. However, we believe that a junior operator should initially be exposed and tested in normal scans, and this is the rationale we followed in constructing our training program. We plan to extend this program to clinical patients with impaired LV function, cardiomyopathies, and studies that reflect ‘real world’ practice of a CMR Unit. Furthermore, we plan to examine the long-term maintenance of standards after 6 months training in CMR. Moreover, our study did not involve any measurements in the right ventricle, which tend to be more variable, even amongst experienced operators (18), and are likely to require more prolonged training. Finally, we addressed the importance of training and consistency only for post-processing. Formal training and expertise in the acquisition of image is also important for the maintenance of reproducibility and accuracy of CMR. Moreover, as blood-tissue contrast, spatial resolution and automated boundary detection methods are expected to improve in the future, post processing analysis for LV function may become less reliant on manual analysis (20). Similarly, image acquisition may be simplified as new ultrafast methods of acquisition are developed (21).

**CONCLUSION**

An intensive two-month training program can significantly improve LV post-processing amongst inexperienced CMR operators. We would, however, view this as a minimum training period, and analysis of impaired LV (especially with wall thinning) function and/or assessment of right ventricular function will likely require a longer training period.

**REFERENCES**


18. Hudsmith LE, Petersen SE, Francis JM, Robson MD, Neubauer S. 
Normal human left and right ventricular and left atrial dimensions 
using steady state free precession magnetic resonance imaging. 

19. Hunt SA. ACC/AHA 2005 guideline update for the diagnosis and 
management of chronic heart failure in the adult: a report of the 
American College of Cardiology/American Heart Association Task 
Force on Practice Guidelines (Writing Committee to Update the 
2001 Guidelines for the Evaluation and Management of Heart Fail-

20. Caiani EG, Toledo E, MacEneaney P, Bardo D, Cerutti S, Lang RM, 
Mor-Avi V. Automated interpretation of regional left ventricular wall 
motion from cardiac magnetic resonance images. J Cardiovasc 

Hayes SW, Saouaf R, Germano G, Friedman JD, Berman DS. 
Rapid assessment of left ventricular segmental wall motion, ejec-
tion fraction, and volumes with single breath-hold, multi-slice 