Journal of Cardiovascular Magnetic Resonance (2005) **7,** 447–457
Copyright © 2005 Taylor & Francis Inc.
ISSN: 1097-6647 print / 1532-429X online
DOI: 10.1081/JCMR-200053578

Taylor & Francis
Taylor & Francis Group

VENTRICULAR FUNCTION

Operator induced variability in cardiovascular MR: left ventricular measurements and their reproducibility

MIKHAIL G. DANILOUCHKINE, PH.D., JOS J. M. WESTENBERG, ALBERT DE ROOS, JOHAN H. C. REIBER,* and BOUDEWIJN P. F. LELIEVELDT

Division of Image Processing, Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands

Purpose. To assess the intra- and inter-operator variability of the manual planning of cardiovascular magnetic resonance imaging and to evaluate the influence of these factors on the functional parameters of the left ventricle (LV). *Method.* The study population consisted of 10 healthy volunteers. For each subject the manual planning of the short-axis cine acquisitions was carried out twice by one operator and once by a second operator. Left ventricular volume, mass, and function were manually evaluated twice by one experienced observer, resulting in an approximation of the intra-observer variability factor. The intra- and inter-operator variation factors were estimated as the difference between the total and intra-observer variation components. *Results.* LV end-diastolic volume varied by 3.3% and 4.16%, and LV end-systolic volume by 5.84% and 6.23% for intra- and inter-operator studies, respectively. The variability for LV mass at end-diastole was equal to 4.23% in both studies. For the ejection fraction the variability was 3.56% and 2.97% for intra- and inter-operator studies, respectively. Comparison of reproducibility between intra- and inter-operator studies resulted in insignificant statistical differences. Bland-Altman limits of agreements revealed no systematic bias in differences between measurements with respect to their means. Reliability of the planning expressed as the angular deviation of the short-axis imaging planes amounts to 2.67 ± 1.5° and 4.99 ± 2.17° for the intra-operator and inter-operator studies, respectively. For EDV, ESV, and EF approximately 75–80% of the total variation can be explained by the within or between operator variation, while the same percentage is 60% for LVM. *Conclusions.* Our study confirms the excellent inter- and intra-operator reproducibility of the cardiovascular magnetic resonance measurements of the left ventricular volumes and mass in a group of healthy volunteers.

Key Words: Left ventricle; Magnetic resonance imaging; Ventricular function; Intra- and inter-operator variability

1. Introduction

The precise evaluation of cardiac volume, mass, and function plays an important role in the understanding and consequent treatment of many myocardial diseases (1, 2). The most widely available technique for this purpose is two-dimensional echocardiography. Although the image acquisition using this technique is highly reproducible in normal individuals, the quantitative analysis heavily relies on geometric assumptions about the shape of the human heart in these two-dimensional images (3–5). An alternative to echocardiography is cardiovascular magnetic resonance (CMR) imaging, which is free from any geometric assumptions because of its truly three-dimensional characteristics and has been proven to be precise (6–8) and reproducible (9–19).

The precision and reproducibility of the results of the quantitative analysis determine the usability of an imaging modality in the clinical environment. Although there is no doubt about the importance of the precise measurements, reproducibility of the quantitative results becomes equally crucial in specific settings, such as in sequential evaluation of ejection fraction for cardiotoxicity. Commonly, insignificant absolute differences in measurements rarely change the clinical practice, while the direction of change may have significant impact on the treatment course. Therefore, it is important to fully recognize and firmly understand the nature of the disparity in quantitative measurements. In a follow-up study of patients, physicians are often confronted with a dilemma whether the difference in measurements is caused by the progression of the disease or could be explained by the limitations of the measuring equipment or imperfections in the quantitative analysis. To answer this question correctly, the influence of the different variability factors on the results of the quantitative analysis has to be investigated.

A large number of publications have been devoted to the study of reproducibility of quantitative CMR measurements. Several variability factors have been clearly identified and carefully studied by different research groups (9–19)

Received 6 May 2004; accepted 25 August 2004.
*Address correspondence to Johan H. C. Reiber, Ph.D., Division of Image Processing, Bld. 1, C2S/C3Q, Leiden University Medical Center, P.O. Box 9600, 2300 RC Leiden, The Netherlands; E-mail: J.H.C.Reiber@lumc.nl

Order reprints of this article at www.copyright.rightslink.com

Table 1. Studies of the different variability factors and their impact on reproducibility of the CMR measurements by different research groups

Author	Variability factors	MR imaging technique	MR imaging sequence	Study population	Author's conclusions
Moon et al. (9)	Inter-observer; inter-protocol; inter-study[a]	BH; ECG-gated	SSFP & FLASH	10 healthy 10 cardiomyopathy	SSFP produces significantly higher LV volume; FLASH and SSFP have similar reproducibility
Bogaert et al. (11)	Intra-observer; inter-observer; inter-study[b]	BH; ECG-gated	Turbo-FLASH	12 healthy	MRI & echo do not significantly differ; Inter-observer variation is smaller than inter-study.
Grothues et al., (12)	Inter-study[b]	BH; ECG-gated	FLASH	20 healthy 20 LV hypertrophy 20 cardiomyopathy	High reproducibility in normal and pathological hearts; reproducibility with MRI superior to echo.
Semelka et al. (13)	Inter-observer; inter-study[a]	FB; ECG-gated	GRE	11 healthy	Inter-study variation is the largest; CMR measurements are highly reproducible with MRI.
Semelka et al. (14)	Inter-observer; inter-study[a]	FB; ECG-gated	GRE	11 cardiomyopathy 8 LV hypertrophy	CMR measurements are reproducibile in morphologically abnormal ventricle.
Bellenger et al. (15)	Intra-observer; inter-observer inter-study[a]	BH; ECG-gated	FLASH	20 healthy 20 heart failure	High reproducibility in normal and patients.
Pattynama et al. (16)	Intra-observer; inter-observer; inter-study[a]; inter-prototocol	BH; ECG-gated	SE & GRE	2 healthy	MRI is more reproducible than echo; No difference in reproducibility between SE and GRE; Inter-study variation is the largest.
Hori et al. (17)	Intra-observer; inter-observer; inter-protocol	BH & FB; ECG-gated	SSFP	6 healthy 6 cardiomyopathy 5 ischemic disease 3 valvular disease 4 myocarditis and thromboembolism	Segmented and real-time SSFP have similar reproducibility; High correlation of CMR measurements.
Schalla et al. (18)	Intra-observer; inter-observer; inter-protocol	BH; ECG-gated	TFE & TFE-EPI	34 infarction or cardiomyopathy	TFE-EPI & TFE highly correlate and have similar reproducibility; Reproducibility with MRI superior to echo.
Plein et al. (19)	Intra-observer; inter-observer; inter-study[b]; inter-protocol	BH; ECG-gated	TFE & TFE-EPI	10 healthy 11 cardiomyopathy 10 LV hypertrophy	TFE and TFE-EPI have similar reproducibility in normal and pathological hearts.

SSFP, steady-state free precession; FLASH, fast low-angle shot; SE, spin-echo; GRE, gradient-recalled-echo; TFE, turbo field echo; TFE-EPI, hybrid turo field echo-echo planar imaging; BH, breath-hold; FB, free-breathing.

[a]Inter-study—combined positional and physiological variability factors.

[b]Inter-study—positional variability factors.

(Table 1). These are of inter-study (positional and physiological or only positional), inter- and intra-observer, inter-protocol, and inter-operator nature. The inter-study variability has been addressed in several publications (9, 11–16, 19). In these studies, the influence of the positional as well as physiological factors on the results of quantitative measurements has been investigated. The positional variability factors arise from the different position and orientation of the examined subject inside the MRI scanner (11, 12, 19). Usually after the first CMR examination, each subject has been instructed to leave the scanning area so that the preparation for the second examination could be immediately restarted. In some studies the second examination of the same subject has been intentionally conducted several days/weeks after the first one. In these studies (9, 13–16), the combined impact of the positional and physiological factors (i.e., inotropic and chronotropic states, blood pressure, and day-to-day change in cardiac filling) on reproducibility of CMR measurements was under consideration. The intra- and inter-observer variability factors are introduced during the quantitative analysis phase (9–11, 13–19), when the endocardial and epicardial borders of the left ventricle are delineated twice by one observer or by two different observers, respectively. The observer's subjectivity to identify and outline the heart boundaries is the subject of investigation in such a study. The impact of the differences in the quality of the images obtained with different CMR protocols was reported in Refs. (9, 16–19). Moon et al. (9) compared the gradient-echo fast low-angle shot (FLASH) protocol with the recently introduced balanced-FFE protocol, while Pattynama et al. (16) made a comparison between the spin-echo and gradient-recalled echo protocols. Hori and co-authors (17) evaluated the precision and reproducibility of the real-time balanced-FFE cine CMR in free-breathing mode and the segmented balanced-FFE cine CMR with breath hold. In other studies, CMR measurements were obtained from the turbo field echo protocol and an ultrafast real-time hybrid field echo/echo planar sequence and inter-protocol reproducibility was assessed (18, 19).

In the above-mentioned publications, all imaging examinations were usually performed by one experienced operator and, therefore, the studies were not designed to test potential differences in planning CMR examinations. The original contribution of this article is to identify the intra- and inter-operator variability factors and to quantitatively assess their impact on CMR measurements. The intra- and inter-operator variability factors are introduced during the planning phase of CMR examinations and exhibit the differences in devising the short-axis CMR acquisitions by the same operator on two different occasions or by two different operators, respectively. One experienced observer performed manual contour tracing and quantitative analysis of the left ventricular (LV) volumes, mass, and function twice, so that the intra-observer variation can be quantitatively assessed. Total variation of CMR measurements in our study was decomposed into operator variation, introduced during CMR planning, and intra-observer variation, arising in the image analysis phase. The relative contribution of each component was estimated.

2. Methods

2.1. *Study population*

Ten healthy volunteers (eight men and two women; mean age 29.9 ± 4.5; age range 23–38) underwent CMR imaging. All subjects enrolled in the study had neither a history of cardiac diseases nor identified risk factors. The institutional ethics committee approved the study and all subjects signed informed written consent prior to the examination.

2.2. *CMR protocol*

MR imaging was performed with a Philips Gyroscan Intera 1.5 T scanner (Philips Medical Systems, Best, The Netherlands) using a dedicated five-element synergy cardiac coil and retrospective VCG-gating. Imaging consisted of acquisitions of balanced-FFE scout images, and subsequent two- (vertical long-axis) and four-chamber (horizontal long-axis) cinematic (cine) images. A stack of breath-hold short-axis cine cross-sections was acquired from the apex to the mitral valve with an 8.0-mm slice thickness and 2.0-mm slice gap. Each cross-section was acquired in a separate breath hold at end expiration.

The study was designed to estimate the impact of the operator variability factor introduced during planning of cardiovascular examinations on CMR measurements. To attain the coherence of CMR planning between two different operators, the standard clinical protocol used in our institution was followed.

The three standard orthogonal cardiac planes (20–22) typically employed for CMR were achieved in accordance with the protocol of our institution (16, 23). A subset of the scout images with visible cross-section of the left ventricle was used to obtain the cine two-chamber images. On the transversal scout images where myocardium exposed an elliptical cross-section, the two-chamber plane was positioned to the plane with the intersection line aligned along the long axis of the ellipse. Subsequently, the four-chamber plane was devised by aligning the LV apex with the center of the mitral valve in the two-chamber images at the end-diastolic and end-systolic phases. The stack of the short-axis cross-sections was derived from two- and four-chamber images. The intersections of the short-axis planes were projected back onto the two- and four-chamber images and positioned in such a way that the intersection lines were orthogonal to the LV axis in both projections. The slice gap and thickness remained fixed, while the number of slices was chosen in such a way that the first slice was positioned at the apex and the last slice at the atrioventricular ring spanning the whole length of the left ventricle.

To estimate the operator-induced variability, two different operators carried out the planning procedure. Both had

extensive experience in planning cardiovascular MR examinations. The scout images were acquired once and subsequently used to devise all cine two-, four-chamber, and short-axis acquisitions by both operators. Planning of the cine cardiac acquisitions constituted the planning cycle. The entire planning cycle was repeated twice by the first operator and once by the second operator, always being the last. The reason behind this strategy was the interdependency of the cine acquisitions. The planning of the four-chamber view mainly depends on the presentation of the LV in the two-chamber view, while the planning of the short-axis projection is based upon both two- and four-chamber views. Therefore, the different appearance of the LV in one or both two- and four-chamber views may directly result in changes of the short-axis stack spatial orientation.

The number of cardiac phases was set to 30. Eight to 13 slices were needed to completely encompass the LV from the apex up to the atrioventricular ring. Multislice balanced-FFE scouts were acquired in free-breathing mode at end diastole with the following parameters: TR, 2.3 ms and TE, 1.16 ms; slice thickness, 10 mm; slice gap, 0 mm; field of view, 45 $\times$ 45 cm; acquisition matrix, 128 $\times$ 128; reconstruction matrix, 256 $\times$ 256; flip angle, 55°. Segmented balanced-FFE breath-hold cines were used for all slices with the following parameters: TR, 3.0 ms and TE, 1.52 ms (two-chamber acquisitions); TR, 3.4 ms and TE, 1.7 ms (four-chamber acquisitions); TR, 3.3 ms and TE, 1.65 ms (short-axis acquisitions); slice thickness, 8 mm; slice gap, 2 mm; field of view, 35 $\times$ 35 cm; acquisition matrix, 192 $\times$ 192; reconstruction matrix, 256 $\times$ 256; number of frames, 30, with the average temporal resolution of 20–25 ms depending on the heart rate; and flip angle, 50°. The cardiac gating was performed retrospectively.

To reduce the influence of physiological variability factors, all imaging examinations were performed in one session with the subject remaining in the scanner during the entire examination. Moreover, special attention was given to proper adjustment of the retrospective VCG gating. The cine acquisition is usually obtained in breath holding mode at end-expiration phase. As a result of repeated withholding of breath, the cardiac rate may increase. Therefore, all subjects were instructed to take sufficient rest between the acquisitions to prevent the average patient's heart rate from significant deviations, although the diastolic and systolic brachial pressure and heat rate were not monitored.

2.3. *Image analysis*

Analysis was performed with a personal computer using the MASS software (Medis medical imaging systems bv, Leiden, The Netherlands) and was done by one experienced observer. The short-axis series were made anonymous. The first cine phase of each short-axis acquisition corresponded to end-diastole due to the retrospective VCG-gating. End-systole was visually determined by observing the cine movie for a mid-ventricular slice. Due to the heart rate variability, the end-systolic phase was independently detected for each short-axis acquisition. The epicardial and endocardial LV borders were manually outlined in both end-diastole and end-systole by one operator. Due to the fact that the endocardial border identification is difficult, especially in the apical slice at the end-systolic phase, we adhered to a strict set of guidelines. The papillary muscles and trabeculations were disregarded in the manual segmentation and were assigned to the LV blood pool. The RV trabeculations appearing as pouches along the septal wall and subepicardial fat were excluded from LV mass. The contours were delineated in all slices where the myocardium exposed at least half of its circumferential length (8, 24). End-diastolic and end-systolic volumes were calculated by summing the volume of each individual short-axis slice, computed as a product of the contour area and the sum of the slice thickness and gap (Simpson rule). The myocardial volume was computed as the difference between the epicardial and endocardial volumes. Conversion to myocardial mass was performed by multiplying the myocardial volume with the standard empirical density factor of 1.05 g/cm^3.

To reduce the impact of the intra-operator variability factor, one experienced observer analyzed the acquired images twice with a 2-month interval in between. An average of two measurements yielded the value of the functional parameter.

2.4. *Reliability analysis*

To evaluate the reliability of cardiovascular planning, the following approach was used. The spatial orientation of the imaging plane of the first short-axis series was chosen as the reference for further calculations. The angular difference between the reference orientation and the one of the imaging plane of another short-axis acquisition was used to characterize the planning reliability. The mean and standard deviation of the angular differences were calculated for all short-axis acquisitions resulting in quantitative measurements of intra- and inter-operator variability of the CMR planning procedure.

2.5. *Statistical analysis*

The calculated parameters were used for the statistical analysis. The Pearson correlation coefficient was calculated to measure the strength and direction of a *linear* relationship between two variables. Statistical significance of the differences between the functional parameters was assessed with paired t-test. Differences for which the P-value was less than 0.05 were assumed to be significant. The agreement within and between operators was evaluated by means of Bland-Altman analysis (25). The percentage of variability was assumed to be a ratio (expressed as a percentage) of the absolute difference between the two measurements with respect to their mean value. Its mean value and standard deviation were assessed to provide an alternative estimate of operator variability. For estimation of the relative contribution of the intra-observer and operator variability to the total

variation, one-way analysis of variance was performed (26) as described in the Appendix.

3. Results

CMR was well tolerated by all subjects included in the study. The imaging time was up to 45 min. All subjects were in sinus rhythm. The average breath-hold time of the cine acquisition was between 15 and 17 sec, tolerated well by all subjects.

3.1. *Image and contour comparison*

All acquired short-axis images had sufficient quality and high contrast between the blood and myocardium to trace the

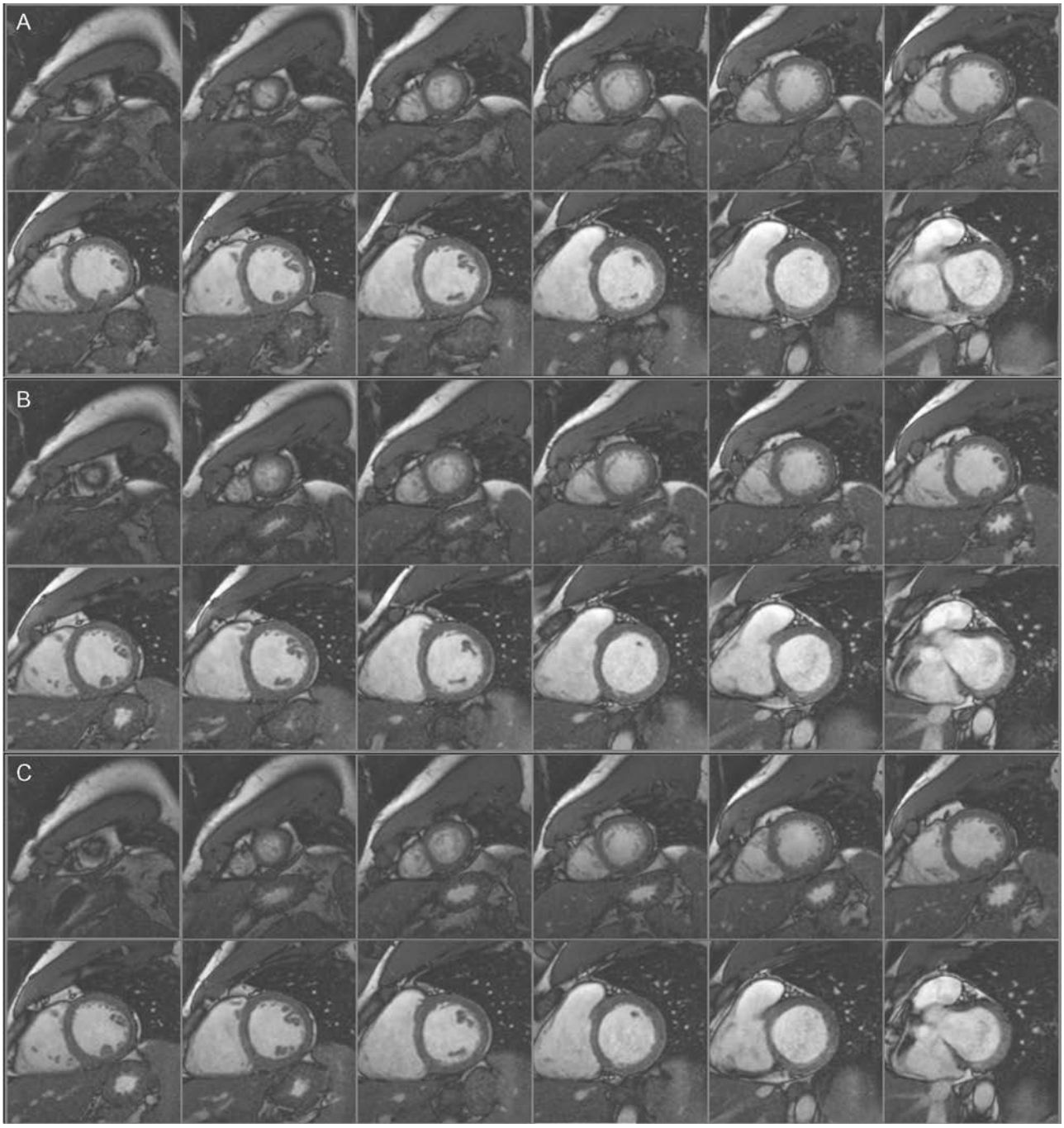

Figure 1. Sample short-axis stack volume acquisitions at end-diastole in a healthy subject. Two top rows (A)—planning was performed by the first operator; two middle rows (B)—planning was done by the same operator for the second time; two bottom rows (C)—planning was carried out by the second operator.

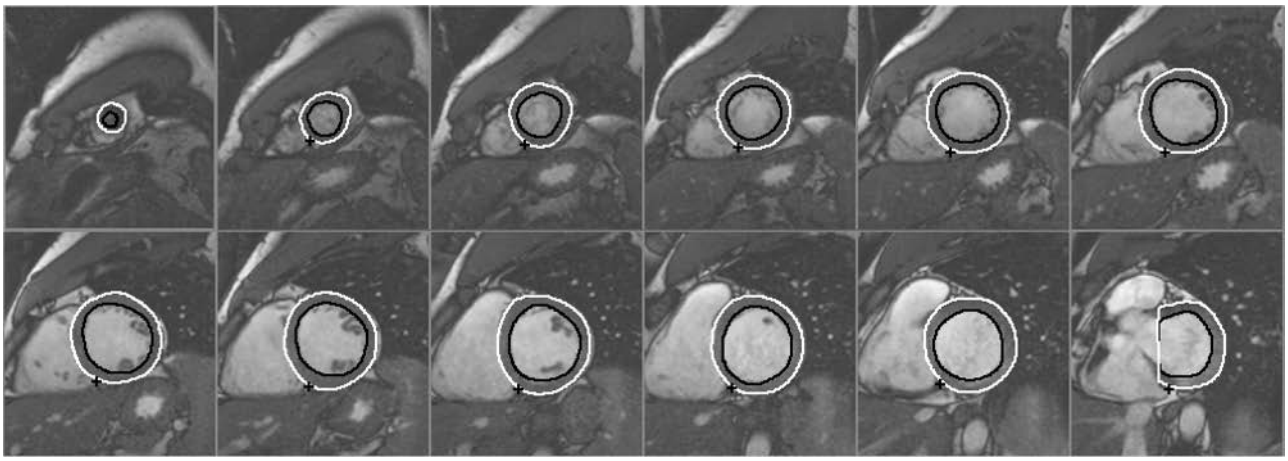

Figure 2. The outlined endocardial (black) and epicardial (white) contours in a healthy subject.

contours. Three stacks of short-axis slices for a healthy volunteer devised by two operators are shown in Fig. 1. Although the mid-ventricular cross-sections appeared to be similar for all three short-axis acquisitions, the differences become evident in the most basal and apical slices. For the subject in Fig. 1 the heart does not expose the LV blood pool in the most apical slice for the first planning. On the other hand, the most basal slice shows the full circumferential length of the LV myocardium while only partial circumferential length of the LV myocardium is exhibited in the corresponding slice for the second and third planning.

Figure 2 shows the manually delineated contours for the same subject. The full circumferential length of the myocardium is observed in all but the most basal cross-sections. When the myocardium exposes the incomplete circular structure in a cross-section, the guidelines for outlining the myocardial borders specified in previous study (8, 24) were followed. However, in the most basal slice with partial circumferential length of the LV myocardium, the lateral walls and part of the septal walls were outlined differently. The following criterion was utilized to distinguish between the myocardium and the LV outflow tract. The wall was regarded as the myocardium and a part of the septum if its thickness was approximately equal to the thickness of the lateral wall. The remaining part of the wall was attributed to the LV outflow tract. For the purpose of the quantitative assessment, those contours were made closed by connecting the end-points with a straight line.

3.2. *Ventricular volume, ejection fraction, and mass quantification*

The intra- and inter-operator CMR measurements for end-diastolic volume (EDV), end-systolic volume (ESV),

Table 2. Reproducibility of the CMR measurements for the interoperator and intraoperator studies

	EDV (mL)	ESV (mL)	EF (%)	LVM (g)
Intra-Operator				
Mean ± SD	186 ± 37	70 ± 16	63 ± 5	133 ± 31
Mean diff. ± SD	5.03 ± 6.15	−0.83 ± 4.80	1.5 ± 2.32	1.77 ± 7.19
Corr. coeff.	0.99	0.96	0.88	0.98
T-test ($p < 0.05$)	0.015 (S)	0.3 (Ns)	0.036 (S)	0.23 (Ns)
%Variability ± SD	3.3 ± 2.57	5.48 ± 3.86	3.56 ± 2.43	4.23 ± 2.52
BA Limits	−7.02:17.08	−10.23:8.58	−3.04:6.04	−12.33:15.86
Inter-Operator				
Mean ± SD	185 ± 36	68 ± 16	63 ± 5	135 ± 31
Mean diff. ± SD	8.07 ± 5.58	2.77 ± 5.1	−0.53 ± 2.32	−2.75 ± 5.97
Corr. coeff.	0.99	0.95	0.9	0.98
T-test ($p < 0.05$)	$p < 0.001$ (S)	0.1 (Ns)	0.24 (Ns)	0.09 (Ns)
%Variability ± SD	4.16 ± 2.71	6.23 ± 5.13	2.97 ± 2.32	4.23 ± 3.23
BA Limits	−2.87:19.01	−7.72:12.27	−4.01:5.08	−14.44:8.94

Note: EDV, end-diastolic volume; ESV, end-systolic volume; EF, ejection fraction; LVM, left ventricular mass at end-diastole; Mean diff., mean difference; Corr. coeff., correlation coefficient; BA limits, Bland-Altman limits of agreement; S, significant difference; Ns, nonsignificant.

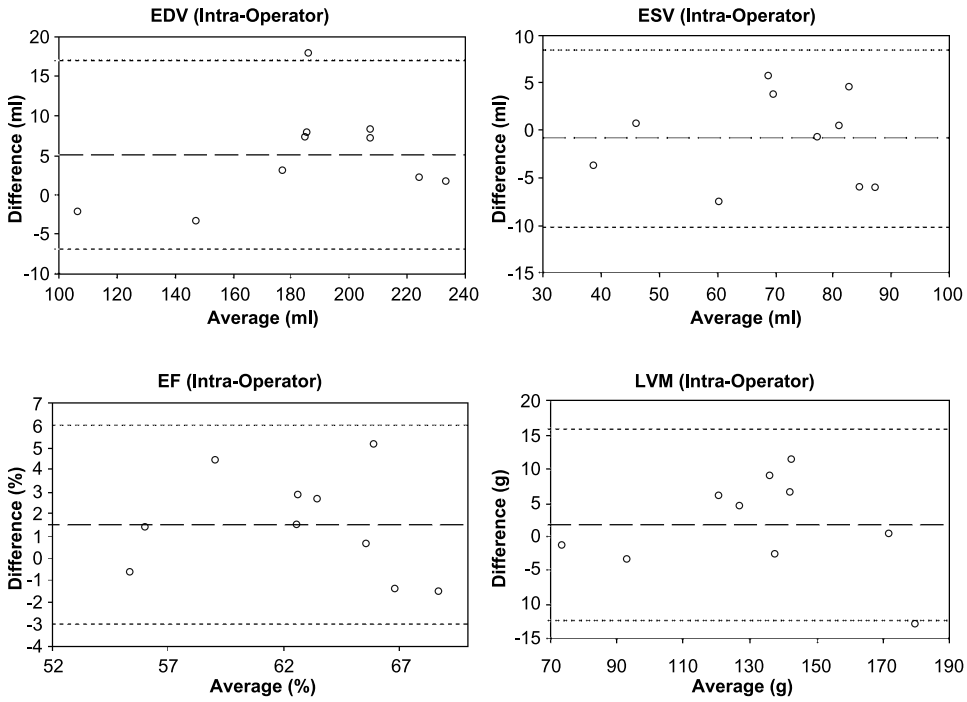

Figure 3. Bland-Altman plots of LV volumes, mass, and function for automated CMR planning for intra-observer study. The mean value and limits of agreements are shown with the dotted line.

ejection fraction (EF), and left ventricular mass (LVM) are shown in Table 2. Three out of eight CMR measurements, namely the intra-operator EDV (P-value = 0.015) and EF (P-value = 0.036) and the inter-operator EDV (P-value < 0.001), showed statistical significance. Most parameters demonstrated a consistently high degree (> 90%) of correlation, only the intra-operator LVM turned out to be 88%. The percentages of variability were low for

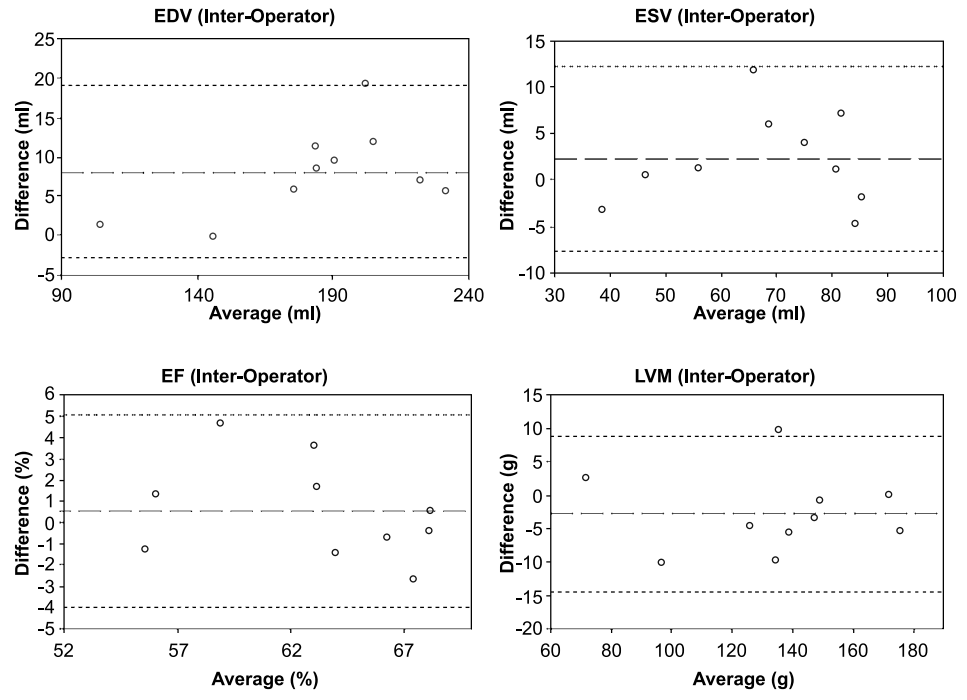

Figure 4. Bland-Altman plots of LV volumes, mass, and function for automated CMR planning for inter-observer study. The mean value and limits of agreements are shown with the dotted line.

all CMR measurements. The paired t-test revealed no significant differences between the intra- and inter-operator percentages of variability for LV dimensions and function.

The results of the Bland-Altman analysis are graphically summarized on Figs. 3 and 4. This analysis showed no systematic bias in error between measurements. In general, the limits of agreements for the intra- and inter-operator studies in normal subjects were within clinically acceptable margins.

3.3. *Reliability analysis*

The mean and standard deviation of the LV axis were $2.67 \pm 1.5^{\circ}$ and $4.99 \pm 2.17^{\circ}$ for the intra-operator and inter-operator studies, respectively. According to this criterion the intra-operator planning was proven to be more reliable. The paired t-test revealed the statistical difference between two intra- and inter-observer variability factors (P-value < 0.01).

3.4. *Analysis of variation*

Decomposition of the total variation into intra-observer and intra-/inter-operator components is graphically presented in Fig. 5. Seventy percent to 80% of the total variation can be explained by intra- or inter-operator variation, introduced during CMR planning, for EDV, ESV, and EF, while the rest of variation is due to intra-observer variation, arising in the image analysis phase. For LVM, the percentage of the

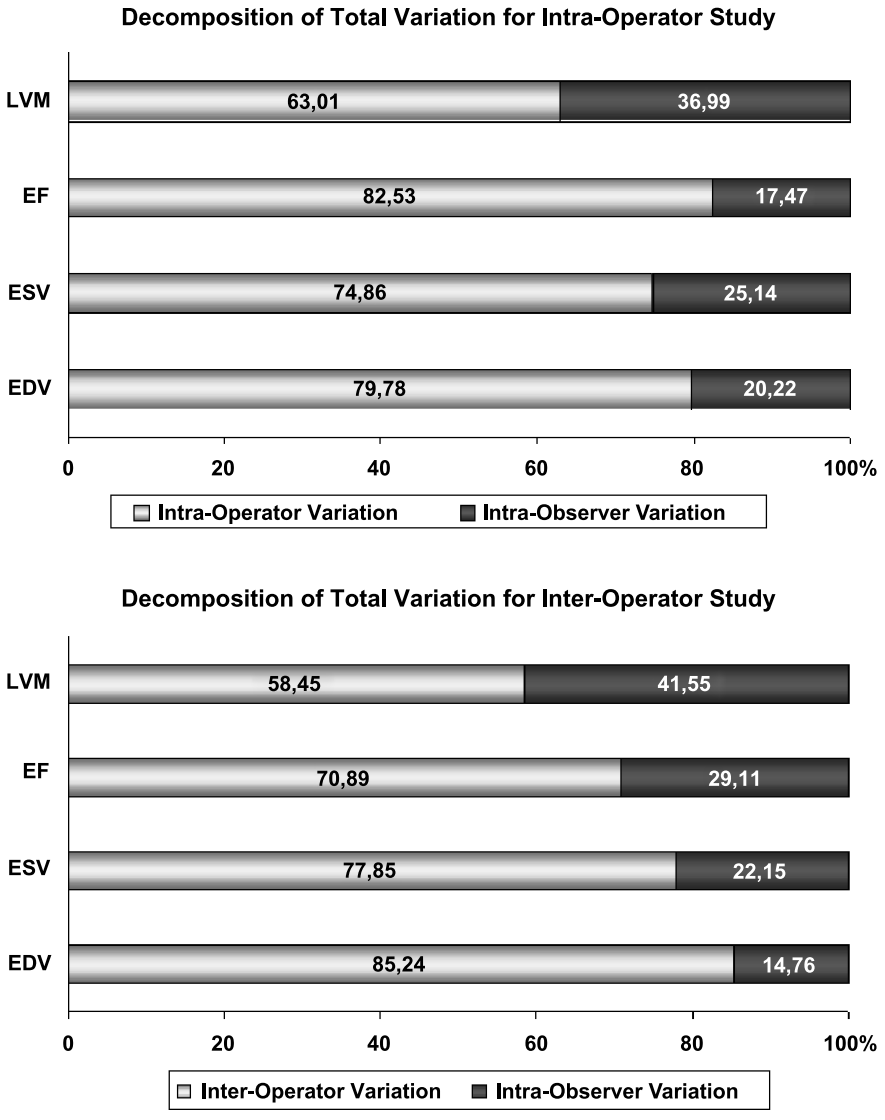

Figure 5. Decomposition of the total variation into the intra-observer (dark bars) and operator (light bars) variation components for different CMR measurents in the intra- (top) and inter-operator (bottom) studies.

variation explained by intra- or inter-operator variation is approximately 60%.

4. Discussion

Comparison between reproducibility factors for CME measurements in normal subjects presented in this article and previously published by other researchers is graphically summarized in Fig. 6. Our findings confirm that the intra- and inter-operator reproducibility with CMR is somewhat higher compared to the inter-study reproducibility reported by Semelka et al. (13) on 11 normal subjects and by Bogaert et al. (11) on 12 normal subjects. However, our results are comparable to the inter-study reproducibility described by Bellenger et al. (15) on 15 normals and by Grothues et al. (12) on 20 normals. Likewise, the intra- and inter-operator reproducibility of CMR in the present study is equivalent to the inter-observer and inter-study reproducibility reported by Moon et al. (9), who conducted their investigation using balanced-FFE protocol. Bland-Altman limits of agreement in the current study are similar to one presented in the paper by Bellenger et al. (15).

Dissimilarity in reproducibility of the CMR measurements can be partially explained by the fact that Semelka et al. applied a slower gradient-recalled echo protocol, while the others used fast low-angle shot sequences or balanced-FFE protocol as in our study. Although all previously published studies were aimed to research the inter-study variability of CMR measurements, the investigated factors were different. For example, Moon et al. (9), Semelka et al. (13) and Bellenger et al. (15) performed the second examination of the subjects a week after the first one, providing a broader coverage of different physiological variability factors. In the studies of Bogaert et al. (11) and Grothues et al. (12), the subjects were examined on the same day, and only heart rate and blood pressure factors were under investigation. Moreover, in all above-mentioned studies the researchers examined a combined effect of intra-operator, inter-study, and intra-observer variability factors without evaluation of the seperate contribution of each factor.

The current study was aimed to estimate the impact of the intra- and inter-operator variability factors on the quantitative assessment of LV dimensions and function. The experiments were designed to minimize the influence of the physiological (day-to-day change in cardiac filling) and inter-protocol variability factors as well as variability due to different position of the subject inside the scanner. Nevertheless, the reported results of quantitative assessment of LV volumes, mass, and function unavoidably suffered from other varying physiological factors such as inotropic and chronotropic state and blood pressure. Moreover, the reported results could be biased due to intra-patient variability in terms of the consistency of the image quality with repeat scans. Prolonged examination time might increase patient tiredness and impair the quality of acquired images. With the second operator being always the last, a certain bias was introduced in quantitative assessment of the dimensions and function of LV. By randomizing the order in which the operators carried out CMR planning, this effect would be averaged out during the quantitative analysis phase.

Reproducibility of CMR measurements plays an important role in establishing clinical usability of MRI in routine practice and has received much attention of the research community during the past decade. It allows the determination of whether the measured differences are due to the real changes in the cardiovascular system or can be attributed to the intrinsic variability of the imaging technique and/or experimental set-up. Studies to investigate reproducibility of the quantitative results of CMR measurements involve repeat examination of subjects and cannot be always carried out under similar conditions. Therefore, design of such studies requires exact identification of the variability factors and careful estimation of the contribution of each individual

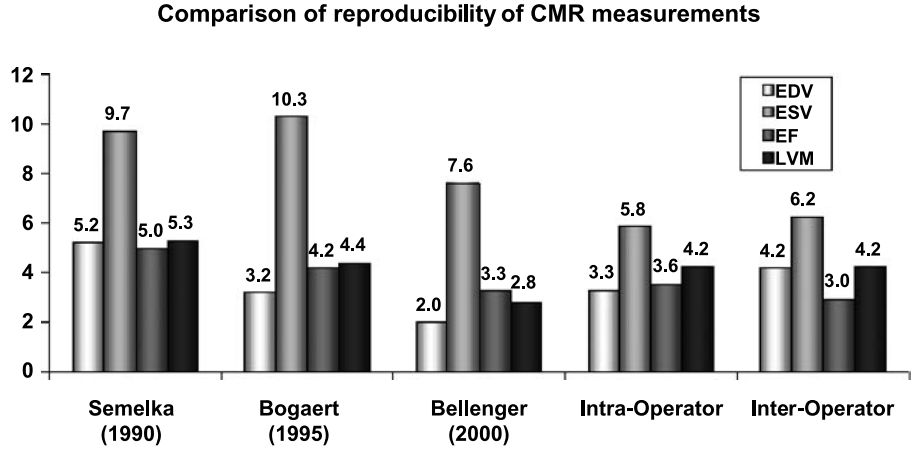

Figure 6. Comparison of reproducibility in the CMR measurements in normal subjects for previously published inter-study variability and intra- an inter-operator variability of the present study.

factor. For example, in our study, two variability factors, namely intra-/inter-operator and intra-observer, influenced the measured results and the contribution of each factor was estimated separately. The relative contribution of operator and intra-observer variation was approximately 70% and 30%, respectively.

In the present study the reliability of CMR planning expressed as the angular deviation of the imaging plane orientation was also assessed. Although this particular parameter does not possess any diagnostic value, it can be utilized for benchmarking the performance of the automated systems for CMR planning (27, 28). Jackson et al. (27) reported an LV axis average deviation of 12.8° between the automatic and manual planning methods. The reliability of the system for automated CMR planning described by Lelieveldt and coauthors (28) was found to be 12.2°. The important theoretical and practical advantages of the images acquired in planes orthogonal to the long axis of the LV have been demonstrated by Dinsmore and coauthors (20, 21). Those images not only allow to obtain optimal display of the human heart in terms of the image quality and appearance of the anatomical structures, but also have been adopted as a standard for quantitative assessment of the LV dimensions and function (9–19). For clinical acceptance of systems for automated CMR planning, their performance should be sufficient to provide a precise estimate of the LV axis spatial orientation at least within the limits perceived by humans. Unfortunately, the current systems do not satisfy this criteria.

In our study, we restricted the group of subjects to healthy volunteers. All subjects had a normal sinus rhythm, were willing to lie motionlessly in the scanner, and were able to withhold from breathing during the acquisition. Grothues and coauthors (12) reported that the above-mentioned factors result in a higher image quality and, as a consequence, in improved degree of reproducibility in normal subjects rather than in patients. Therefore, they strongly emphasized the great importance of reproducibility evaluation in both groups. Undoubtedly bad image quality may have a negative influence on the CMR planning procedure and the quantitative measurements derived from those acquisitions, but its impact on the intra- and inter-operator variability is expected to be negligible. Planning of cine cardiac acquisitions requires rough identification of the easily seen anatomical landmarks such as the LV apex or the middle point of the mitral valve orifice. This task can be successfully achieved even in a case of moderately degraded image quality.

In conclusion, our study confirms that anatomical and functional measurements from cine CMR images are highly reproducible within and between operators. Therefore, different operators may carry out CMR planning interchangeably. Decomposition of total variation revealed that the major part (more than 70%) of total variation can be explained by operator variation, introduced during CMR planning phase, while intra-observer variation from contour tracing has less impact on total variation.

Appendix

For each study participant two short-axis stacks were acquired either twice by the same operator or by two different operators. The same observer analyzed each short-axis stack twice, resulting in four measurements per study. The intra-observer variation was estimated as a sum of the individual intra-observer variation for each short-axis stack, multiplied by a factor of two. The total variation was computed as the variation of four measurements. The intra- or inter-operator variation was obtained by subtracting the value of intra-observer variation from total variation. Assuming total variation to be 100%, the relative contribution of both operator and observer variation factors was assessed.

Acknowledgments

This research was supported by the Leiden University Medical Center revolving fund and a grant by the Dutch Foundation for Technical Sciences (STW Project LPG 5651).

References

1. Allison JD, Flickinger FW, Wright CJ, Falls DG III, Prisant LM, VonDohlen TW, Frank MJ. Measurement of left ventricular mass in hypertrophic cardiomyopathy using MRI: comparison with echocardiography. Magn Reson Imaging 1993; 11:329–334.
2. Shapiro EP, Rogers WJ, Beyar R, Soulen RL, Zerhouni EA, Lima JA, Weiss JL. Determination of left ventricular mass by magnetic resonance imaging in hearts deformed by acute infarction. Circulation 1989; 79:706–711.
3. Gordon EP, Schnittger I, Fitzgerald PJ, Williams P, Popp RL. Reproducibility of left ventricular volumes by two-dimensional echocardiography. J Am Coll Cardiol 1983; 2:506–513.
4. Gerard O, Billon AC, Rouet JM, Jacob M, Fradkin M, Allouche C. Efficient model-based quantification of left ventricular function in 3-D echocardiography. IEEE Trans Med Imag 2002; 21:1059–1068.
5. Ye X, Noble JA, Atkinson D. 3-D freehand echocardiography for automatic left ventricle reconstruction and analysis based on multiple acoustic windows. IEEE Trans Med Imag 2002; 21:1051–1058.
6. Caputo GR, Tscholakoff D, Sechtem U, Higgins CB. Measurement of canine left ventricular mass by using MR imaging. Am J Roentgenol 1987; 148:33–38.
7. Katz J, Milliken MC, Stray-Gundersen J, Buja LM, Parkey RW, Mitchell JH, Peshock RM. Estimation of human myocardial mass with MR imaging. Radiology 1988; 169:495–498.
8. Lorenz CH, Walker ES, Morgan VL, Klein SS, Graham TP Jr. Normal human right and left ventricular mass, systolic function, and gender differences by cine magnetic resonance imaging. J Cardiovasc Magn Reson 1999; 1:7–21.
9. Moon JCC, Lorenz CH, Francis JM, Smith GC, Pennell DJ. Breathhold FLASH and FISP cardiovascular MR imaging: left ventricular volume differences and reproducibility. Radiology 2002; 223:789–797.
10. Mogelvang J, Thomsen C, Mehlsen J, Brackle G, Stubgaard M, Henriksen O. Evaluation of left ventricular volumes measured by magnetic resonance imaging. Eur Heart J 1986; 7:1016–1021.

11. Bogaert JG, Bosmans HT, Rademakers FE, Bellon EP, Herregods MC, Verschakelen JA, Van de Werf F, Marchal GJ. Left ventricular quantification with breath hold MR imaging: comparison with echocardiography. MAGMA 1995; 3:5–12.
12. Grothues F, Smith GC, Moon JC.C, Bellenger NG, Collins P, Klein HU, Pennell DJ. Comparison of interstudy reproducibility of cardiovascular magnetic resonance with two-dimensional echocardiography in normal subjects and in patients with heart failure or left ventricular hypertrophy. Am J Cardiol 2002; 90:29–34.
13. Semelka RC, Tomei E, Wagner S, Mayo J, Kondo C, Suzuki J, Caputo G, Higgins CB. Normal left ventricular dimensions and function: interstudy reproducibility of measurements with cine MR imaging. Radiology 1990; 174:763–768.
14. Semelka RC, Tomei E, Wagner S, Mayo J, Caputo G, O'Sullivan M, Parmley WW, Chatterjee K, Wolfe C, Higgins CB. Interstudy reproducibility of dimensional and functional measurements between cine magentic resonance imaging studies in the morphologically abnormal left ventricle. Am Heart J 1990; 119:1367–1373.
15. Bellenger NG, Davies LC, Francis JM, Coats AJS, Pennell DJ. Reduction in sample size for studies of remodeling in heart failure by the use of cardiovascular magnetic resonance. J Cardiovasc Magn Reson 2000; 2(4):271–278.
16. Pattynama PM, Lamb HJ, van der Velde EA, van der Wall EE, de Roos A. Left ventricular measurements with cine and spin-echo MR imaging: a study of reproduicibility with variance component analysis. Radiology 1993; 187:261–268.
17. Hori Y, Yamada N, Higashi M, Hirai N, Nakatano S. Rapid evaluation of right and left ventricular function and mass using real-time true-FISP cine MR imaging without breath-hold: comparison with segmented true-FISP cine MR imaging with breath-hold. J Cardiovasc Magn Reson 2003; 5:439–450.
18. Schalla S, Nagel E, Lehmkhul H, Klein C, Bornstedt A, Schnackenburg B, Schneider U, Fleck E. Comparison of magnetic resonance real-time imaging of left ventricular function with conventional magnetic resonance imaging and echocardiography. Am J Cardiol 2001; 87:95–99.
19. Plein S, Smith WHT, Ridgway JP, Kassner A, Beacock DJ, Bloomer TN, Sivananthan MU. Measurements of left ventricular dimensions using real-time acquisition in cardiac magnetic resonance imaging: comparison with conventional gradient echo imaging. MAGMA 2001; 13:101–108.
20. Dinsmore RE, Wismer GL, Levine RA, Okada RD, Brady TJ. Magnetic resonance imaging of the heart: positioning and gradient angle selection for optimal imaging planes. Am J Roentgenol 1984; 143:1135–1142.
21. Dinsmore RE, Wismer GL, Miller SW, Thompson R, Johnston DL, Liu P, Okada RD, Saini S, Brady TJ. Magnetic resonance imaging of the heart using image planes oriented to cardiac axes—experience with 100 cases. Am J Roentgenol 1985; 145:1177–1183.
22. van Rossum AC, Visser FC, van Eenige MJ, Valk J, Roos JP. Oblique views in magnetic resonance imaging of the heart by combined axial rotation. Acta Radiol 1987; 28:497–503.
23. Kayser HW, Schalij MJ, van der Wall EE, Stoel BC, de Roos A. Biventricular function in patients with nonischemic right ventricle tachyarrhythmias assessed with MR imaging. Am J Roentgenol 1997; 169:995–999.
24. Matheijssen NA, Baur LH, Reiber JH, van der Velde EA, van Dijkman PR, van der Geest RJ, de Roos A, van der Wall EE. Assessment of left ventricular volume and mass by cine magnetic resonance imaging in patients with anterior myocardial infarction intra-observer and inter-observer variability on contour detection. Int J Card Imaging 1996; 12:11–19.
25. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurements. Lancet 1986; 1:307–310.
26. Scheffe H. The Analysis of Variance. London, England: Wiley, 1964.
27. Jackson C, Robson M, Francis J, Noble JA. Computerized planning of the acquisition of cardiac MR images. Comput Med Imaging Graph 2004; 28:411–418.
28. Lelieveldt BP, van der Geest RJ, Lamb HJ, Kayser HW, Reiber JH. Automated observer-independent acquisition of cardiac short-axis MR images: a pilot study. Radiology 2001; 221:537–542.