Normal Cardiac Magnetic Resonance Measurements and Interobserver Discrepancies in Volumes and Mass Using the Papillary Muscle Inclusion Method

Jeffrey L. Anderson[*], Aaron N. Weaver, Benjamin D. Horne, Heath U. Jones, Gerri K. Jelaco, Julie A. Cha, Hector E. Busto, Judy Hall, Kathy Walker and Duane D. Blatter

Cardiovascular and Radiology Departments, LDS Hospital, Intermountain Healthcare, University of Utah School of Medicine, Salt Lake City, UT, USA

Abstract: Background: The potential of cardiac magnetic resonance (CMR) imaging to provide precise and reproducible measurements might not be realized currently.

Methods: We acquired CMR images on 30 healthy adults and compared readings of 2 experienced physicians to determine sources of measurement variability.

Results: Simultaneous review of tracings revealed 3 major sources of interobserver variability: 1) choice of end-diastolic/end-systolic frames; 2) choice of the most basal left and right ventricular short-axis slices; and, 3) approach to endocardial edge selection.

Conclusions: CMR-derived volumes and mass are measurement-methodology specific. Formal measurement rules and improved computerized edge detection algorithms are needed.

Keywords: Cardiac magnetic resonance, interobserver variability, measurement, normal range.

INTRODUCTION

Cardiac magnetic resonance (CMR) is being increasingly utilized for dynamic imaging of the heart [1] with the expectation that it will provide more accurate and reproducible measurements of cardiac chamber dimensions, volumes, and function (e.g., ejection fraction) compared to other non-invasive imaging techniques such as echocardiography and nuclear cardiography [2]. This is due to superior spatial resolution and more precise border definition compared with other techniques. However, the potential of CMR to provide superior clinical utility might not always be realized in practice because of limited information on the range of normal values using workstation-specific methodology and because of interobserver measurement variability. We addressed these 2 issues through a prospective study in healthy volunteers.

MATERIALS AND METHODOLOGY

Study Objectives. The two main study objectives were: 1) to determine a range of normal CMR values for left and right ventricular and atrial as well as great vessel measurements utilizing equipment (magnet, software) and workstation-specific methods and algorithms, and 2) to determine the magnitude and causes of common interobserver measurement variability as a basis for improving reproducibility in clinical practice.

Study Plan. These objectives were to be achieved through a prospective study in an adequate number of healthy subjects. Qualifying subjects were to undergo a standardized functional CMR evaluation. Predefined cardiovascular measurements were made using workstation-specific methods. Key volume and mass measurements were to be performed independently by 2 observers to improve accuracy, to determine reproducibility, and to develop consensus methods. Results were to be used for clinical and research applications.

Study Population. Thirty study subjects were drawn from among healthy volunteers who responded to advertisements within the hospital community to participate in a study to determine normal CMR measurements using methodology-specific algorithms. Qualifying subjects were uncompensated adults of ages between 40-60 years of either sex (equal proportions women and men) who gave written informed consent for study participation, which included a standard functional CMR, and who had no clinically apparent cardiovascular diseases, including hypertension (history of hypertension, blood pressure ≥ 140 systolic or ≥ 90 mm Hg diastolic, or on antihypertensive medication) or diabetes on screening history, or had abnormal physical findings (abnormal heart sounds or murmurs, other abnormal cardiovascular findings) on screening examination. Subjects also were excluded who had significant other-organ diseases (e.g., chronic obstructive pulmonary disease, renal or liver failure, cancer) or obesity (body mass index >30 kg/m^2) or rhythm other than sinus. The LDS Hospital Institutional Review Board approved the study.

Cardiac MRI Study. Following axial and sagittal localizer sequences (fast GRE/SPGR), standard cardiac 2, 3,

*Address correspondence to this author at the LDS Hospital Cardiovascular Department, 8th Avenue & C Street, Salt Lake City, UT 84143, USA;
Tel: 1-801-408-5552; Fax: 1-801-408-5826;
E-mail: jeffrey.anderson@intermountainmail.org

and 4 chamber 1 cm thick long-axis and short axis slices (8mm skip 2mm) were obtained on a General Electric (GE) 1.5 Tesla magnet (EXCITE platform, version 11.0) using FIESTA (Steady State Free Precession [SSFP]) cine sequences (typical in-plane resolution 2.2 x 1.3 mm^2). In addition, an oblique coronal pilot FIESTA sequence followed by an axial aortic and pulmonary artery velocity-encoded sequence were obtained for aortic and pulmonary artery flow measurements. All images were acquired using phased-array cardiac or body surface coils (with 4 to 12 elements) as appropriate to each imaging sequence during single breath-holds (maximum, 15 seconds) with electrocardiographic (ECG) triggering. (One subject required peripheral gating, i.e., arterial pulse gating by finger probe.)

CMR Image Analysis. Image analysis was performed independently by each of 2 experienced physician observers (JLA, DDB) on a GE Medical Systems ReportCARD™ 1.0 workstation. Quantitative results were entered prospectively into a database, and computer-assisted manual tracings were saved for later visual comparisons. Measurements were to follow standard CMR procedures (GE CMR owner's manual), where available, or mimic the approach to echo measurement, where appropriate. Specific reading instructions given in the workstation manual designated that unattached papillary muscles were to be included within the left and right ventricular (LV/RV) chambers, and the basal slice was to be at least an "almost complete" muscular circle. Volumes were determined from the stack of short axis slices using Simpson's rule [1,3,4].

Individual measurements of primary interest (LV, RV volumes) differing by more than 20% between readers were highlighted for mandatory review. At review conference, the sources of differences were determined and quantified. After discussion, consensus was achieved as to the measurement approach (e.g., agreement on a basal slice discrepancy) to reduce the difference to within the tolerable range. Major sources of discrepancy (i.e., basal slice, ED/ES frames) were then determined for all studies. Individual remeasurement was performed where appropriate, and corrected values were entered into the database for use in averaging of duplicate individual measurements and for determination of overall normal values and normal ranges. Specific areas of potential discrepancy that were audited during consensus review included choices of ED/ES frames, basal slices, and endocardial edge designations.

Study Variables. Study demographic variables included subject age, sex, heart rate, weight, height, and body surface area (BSA). Cardiovascular variables included LV/RV end-diastolic (ED) and end-systolic (ES) volumes, LV ED mass (LV epicardial – endocardial volume x specific gravity [1.05]), and the derived LV/RV variables: stroke volume (SV= EDV-ESV), ejection fraction (EF= SV/EDV), and cardiac output (CO= SV x heart rate). Basal LV septal and inferolateral ("posterior") wall and RV ED and ES wall thickness also were measured. SV and CO were determined from phase-contrast flow sequences. Other variables included left and right atrial (LA/RA) 4- and 2-chamber ES areas (FIESTA); and diameters of the aorta (at 3 levels), pulmonary artery diameter, inferior vena cava, and superior vena cava (from axial and sagittal scout views). Measurements also were indexed to BSA and averaged separately by sex.

Statistics. Results are presented as mean (standard deviation). Student's unpaired *t*-test was used for comparisons between groups (e.g., by sex [5-7]). Proposed normal ranges were derived by inspection of the upper and lower 5% of the range of values of the normal volunteer group. SPSS for Windows (version 14.0, SPSS Inc., Chicago, IL) was used for statistical analyses. Given the study size and numbers of comparisons, a p<0.01 was selected as nominally significant between groups (e.g., by sex).

RESULTS

The study included 30 healthy volunteers (15 women, 15 men). Volunteer age averaged 47 years (range, 40 to 60). BSA averaged 1.7 sq m for women and 2.1 sq m for men. Body mass index averaged 25.5 kg/m^2 for women and 26.5 kg/m^2 for men (overall range, 20-30). Heart rates at study averaged 70 and 67 beats/min, respectively. All studies were of excellent quality.

Normal CMR Measurements. Sixteen key normal LV and RV volumes, LV mass, and derived CMR metrics, averaged for the 2 readers after the consensus conference, are presented in Table **1** overall, by sex, and indexed to BSA. All measurements of volume and mass differed by sex. When indexed to BSA, differences in LVEDVI and LVESVI, LVSVI, LVEF, and RVEF were no longer significant, wheres LV mass index remained significantly greater in men.

Average values and ranges for 54 additional metrics (wall dimensions, chamber diameters, chamber areas, great vessel diameters, and volumetric- and flow-derived cardiac outputs, without and with indexing to BSA) are presented in Table **2**. Although many of these other variables showed differences by sex, most differences were no longer significant when indexed to BSA. This suggests that indexing many of these variables to BSA could allow for a single range of normals for men and women.

Interobserver Correlations. Correlations between initial readings for the 2 readers for LVEDV, RVEDV, LVESV, and RVESV were high: 0.95, 0.91, 0.96, and 0.92, respectively (all p<0.001), although some systematic differences were observed, which were primarily accounted for by differences in a few specific measurement factors (see below).

Sources of Interobserver Measurement Differences. Simultaneous review of individual tracing revealed 3 major sources of interobserver measurement differences: 1) choice of ED/ES frames; 2) choice of the most basal LV, RV short-axis slices; and, 3) approach to endocardial (EC) edge selection. Choice of ED and ES frames differed in 2/30 (7%) paired ED readings and 8/30 (27%) paired ES readings. However, almost all differences were by a single frame, and visual and quantitative review suggested that choosing the alternative frame made only a small difference in computed volume.

More importantly, a discrepancy in choice of basal slice was present between readers in 35% (42/120) of LV or RV series, and when a discrepancy was present, average measurement differences between paired readings were substantial, averaging 14.4% (i.e., 27 ml for ED, 7.5 ml for

Table 1. Summary of CMR Volumes, Mass, and Derived Metrics Overall and by Sex: Means ± SD (5%, 95% Confidence Intervals)

Measure (Unit)	All (N=30)	Males (N=15)	Females (N=15)
LV EDV (ml)	133 ± 30 (86, 196)	151 ± 29 (95, 201)	115 ± 18** (81, 146)
LV EDVI (ml/m^2)	70 ± 13 (45, 95)	73 ± 13 (53, 101)	66 ± 12 (42, 86)
LV ESV (ml)	48 ± 16 (21, 86)	56 ± 16 (22, 89)	40 ± 11* (21, 63)
LV ESVI (ml/m^2)	25 ± 7 (12, 41)	27 ± 8 (12, 44)	23 ± 7 (12, 37)
RV EDV (ml)	146 ± 43 (93, 238)	175 ± 42 (94, 258)	119 ± 19** (93, 156)
RV EDVI (ml/m^2)	76 ± 17 (50, 113)	85 ± 18 (52, 113)	68 ± 10* (49, 85)
RV ESV (ml)	64 ± 22 (30, 106)	78 ± 22 (28, 109)	51 ± 12** (31, 72)
RV ESVI (ml/m^2)	33 ± 9 (18, 52)	38 ± 10 (15, 53)	29 ± 7* (21, 40)
LV MASS (g)	93 ± 24 (60, 140)	111 ± 18 (86, 141)	74 ± 12** (58, 94)
LV MASSI (g/m^2)	48 ± 8 (33, 65)	54 ± 6 (47, 65)	43 ± 6** (28, 53)
LV SV (ml)	85 ± 17 (58, 113)	94 ± 17 (64, 114)	75 ± 11** (55, 93)
LV SVI (ml/m^2)	44 ± 7 (33, 57)	46 ± 8 (36, 57)	43 ± 7 (29, 55)
RV SV (ml)	82 ± 23 (47, 132)	96 ± 23 (66, 149)	68 ± 13** (46, 92)
RV SVI (ml/m^2)	43 ± 9 (28, 62)	47 ± 10 (35, 65)	39 ± 7 (27, 49)
LV EF (%)	64 ± 5 (56, 75)	63 ± 5 (56, 75)	66 ± 5 (56, 74)
RV EF (%)	57 ± 6 (47, 69)	55 ± 5 (47, 71)	58 ± 6 (47, 67)

Summary data presented as mean +/- SD (5%, 95% confidence intervals). Individual subject measurements represent average of 2 readers, with discrepancies >20% reviewed and revised at consensus conference. Indexed (I) metrics represent metric divided by BSA. *p<0.01, **p<0.001, female *vs* male.
Abbreviations: LV/RV= left, right ventricular; EDV/ESV= end-systolic, -diastolic volume; MASS= mass (volume [ml] x 1.05 [g]); SV= stroke volume; EF=ejection fraction.

Table 2. Summary of Other CMR Measurements (Walls, Diameters, Areas, Atria, Great Vessels, Flow) Overall and by Sex: Means ± SD (5%, 95% Confidence Intervals)

Measure (Unit)	All (N=30)	Males (N=15)	Females (N=15)
IVSD (mm)	8.8 ± 1.6 (6, 12)	9.6 ± 1.5 (7, 12)	7.9 ± 1.2* (6, 10)
IVSS (mm)	12.2 ± 1.7 (9, 15)	12.9 ± 1.5 (10, 16)	11.5 ± 1.5 (9, 15)
LVPWD (mm)	8.4 ± 1.3 (6, 11)	8.9 ± 1.3 (7, 11)	7.8 ± 1.1 (6, 9)
LVPWS (mm)	13.1 ± 1.9 (9, 16)	13.8 ± 1.4 (12, 16)	12.3 ± 2.0 (9, 16)
RVFWD (mm)	3.1 ± 0.5 (2, 4)	3.3 ± 0.6 (3, 4)	2.8 ± 0.4 (2, 4)
RVFWS (mm)	5.0 ± 1.0 (4, 7)	5.3 ± 1.1 (4, 7)	4.6 ± 0.8 (4, 6)
LVIDD (mm)	50 ± 5 (43, 59)	52 ± 5 (44, 60)	48 ± 4* (42, 54)
LVIDDI (mm/m^2)	26 ± 2 (22, 32)	25 ± 2 (22, 27)	28 ± 2* (24, 33)
LVIDS (mm)	34 ± 5 (27, 44)	36 ± 4 (28, 45)	32 ± 4* (26, 40)
LVIDSI (mm/m^2)	18 ± 2 (15, 22)	17 ± 2 (15, 20)	18 ± 2 (15, 23)
LVmajAX (mm)	84 ± 8 (73, 100)	89 ± 7 (78, 101)	79 ± 4** (72, 85)
LVmajAXI (mm/m^2)	44 ± 4 (36, 53)	43 ± 4 (32, 49)	46 ± 4 (40, 54)
LVminAX (mm)	47 ± 5 (38, 57)	50 ± 4 (41, 58)	45 ± 4* (38, 49)
LVminAXI (mm/m^2)	25 ± 2 (20, 30)	24 ± 2 (20, 27)	26 ± 2 (22, 30)
RVmajAX (mm)	80 ± 9 (64, 97)	85 ± 8 (72, 104)	75 ± 7* (61, 86)
RVmajAXI (mm/m^2)	42 ± 5 (32, 51)	41 ± 5 (31, 51)	44 ± 4 (39, 52)
RVminAX (mm)	40 ± 6 (27, 49)	43 ± 5 (30, 49)	36 ± 5* (27, 45)
RVminAXI (mm/m^2)	21 ± 3 (16, 26)	21 ± 2 (16, 25)	21 ± 3 (16, 27)
LAAP (mm)	30 ± 6 (16, 39)	31 ± 5 (26, 43)	29 ± 7 (14, 36)
LAAPI (mm/m^2)	16 ± 3 (11, 20)	15 ± 2 (12, 20)	17 ± 3 (10, 20)

(Table 2) contd…..

Measure (Unit)	All (N=30)	Males (N=15)	Females (N=15)
LV4CHAD (cm²)	35 ± 5 (28, 46)	39 ± 5 (29, 48)	31 ± 2** (27, 35)
LV4CHADI (cm²/m²)	18 ± 2 (15, 22)	19 ± 2 (14, 24)	18 ± 1 (16, 20)
RV4CHAD (cm²)	25 ± 6 (17, 35)	29 ± 5 (18, 36)	21 ± 3** (16, 26)
RV4CHADI (cm²/m²)	13 ± 2 (10, 17)	14 ± 2 (10, 17)	12 ± 2 (9, 15)
LA4CHAS (cm²)	17 ± 4 (9, 23)	18 ± 3 (11, 23)	16 ± 3 (9, 22)
LA4CHASI (cm²/m²)	9 ± 2 (6, 11)	9 ± 1 (6, 11)	10 ± 2 (6, 12)
LA2CHAS (cm²)	18 ± 4 (8, 24)	19 ± 4 (12, 24)	16 ± 5 (6, 22)
LA2CHASI (cm²/m²)	9 ± 2 (5, 12)	9 ± 2 (6, 11)	9 ± 3 (4, 13)
RA4CHAS (cm²)	22 ± 5 (14, 31)	25 ± 4 (17, 33)	19 ± 3** (14, 24)
RA4CHASI (cm²/m²)	11 ± 2 (8, 15)	12 ± 2 (9, 16)	11 ± 2 (8, 14)
AOSOV (mm)	33 ± 3 (28, 40)	36 ± 3 (31, 42)	31 ± 2** (28, 34)
AOSOVI (mm/m²)	17 ± 2 (14, 21)	17 ± 2 (14, 21)	18 ± 2 (16, 23)
AOASC (mm)	32 ± 3 (27, 40)	34 ± 3 (30, 42)	31 ± 3 (26, 35)
AOASCI (mm/m²)	17 ± 2 (14, 20)	16 ± 2 (13, 21)	18 ± 2 (14, 20)
AOTHO (mm)	24 ± 2 (20, 28)	25 ± 2 (23, 29)	23 ± 2** (20, 26)
AOTHOI (mm/m²)	13 ± 1 (10, 15)	12 ± 1 (10, 14)	13 ± 1 (12, 15)
AODIA (mm)	23 ± 3 (19, 28)	24 ± 3 (20, 29)	22 ± 3 (19, 27)
AODIAI (mm/m²)	12 ± 2 (9, 15)	11 ± 2 (9, 14)	13 ± 2 (10, 15)
PA (mm)	30 ± 3 (25, 35)	31 ± 3 (27, 36)	28 ± 2* (24, 31)
PAI (mm/m²)	16 ± 1 (13, 18)	15 ± 1 (12, 18)	16 ± 1 (14, 18)
SVC (mm)	21 ± 3 (14, 26)	21 ± 4 (12, 27)	20 ± 2 (16, 25)

(Table 2) contd…..

Measure (Unit)	All (N=30)	Males (N=15)	Females (N=15)
SVCI (mm/m²)	11 ± 2 (8, 14)	10 ± 2 (6, 13)	12 ± 1 (10, 14)
IVC (mm)	22 ± 3 (15, 25)	23 ± 2 (20, 26)	20 ± 3* (15, 24)
IVCI (mm/m²)	11 ± 1 (9, 13)	11 ± 1 (9, 13)	11 ± 1 (10, 13)
HR (beats/min)	67 ± 10 (52, 87)	64 ± 8 (50, 78)	70 ± 10 (54, 92)
CO (L/min)	5.3 ± 0.9 (3.8, 7.3)	5.6 ± 0.9 (4.3, 7.7)	5.0 ± 0.9 (3.6, 7.0)
CI (L/min/m²)	2.8 ± 0.4 (2.2, 3.6)	2.7 ± 0.4 (2.2, 3.5)	2.9 ± 0.4 (2.2, 3.7)
AOSV (ml)	96 ± 10 (57, 136)	107 ± 20 (71, 139)	86 ± 18* (55, 117)
AOCO (L/min)	6.4 ± 1.2 (4.9, 8.4)	6.8 ± 1.1 (5.1, 8.5)	5.9 ± 1.1 (3.6, 7.6)
AOCI (L/min/m²)	3.4 ± 0.6 (2.5, 4.5)	3.3 ± 0.5 (2.7, 4.5)	3.4 ± 0.6 (2.3, 4.5)
PASV (ml)	101 ± 24 (58, 146)	113 ± 22 (78, 150)	88 ± 20* (56, 128)
PACO (L/min)	6.7 ± 1.3 (4.1, 9.1)	7.2 ± 1.2 (5.6, 9.4)	6.1 ± 1.2 (3.6, 8.1)
PACI (L/min/m²)	3.5 ± 0.6 (2.6, 4.8)	3.5 ± 0.6 (2.9, 4.9)	3.5 ± 0.7 (2.3, 4.8)

Summary data presented as mean +/- SD (5%, 95% confidence intervals). Single reader averages. Indexed (I) metrics represent metric divided by BSA. *p<0.01, **p<0.001, female *vs* male.

Abbreviations: IVSD/S= interventricular septum diastole/systole (short axis basal slice); LVPWD/S= left ventricular posterior wall, diastolic/systolic (short axis basal slice); RVFWD/S= right ventricular free wall, diastolic/systolic (short axis basal slice); LVIDD/S= left ventricular internal diameter diastolic/systolic (short axis basal slice); LV/RVmajAX= left/right ventricular major axis internal dimension (4 chamber, diastolic); LV/RVminAX= LV/RV minor axis internal dimension (4 chamber, diastolic; level 1/3rd from base to apex); LAAP= left atrial anterior posterior dimension (3 chamber, systolic); LV/RV4CHAD= LV/RV 4 chamber area, diastolic; LA/RA4CHAS=LA/RA 4 chamber area, systole; LA2CHAS= LA 2 chamber area, systole; AOSOV: aortic diameter, level of sinus of Valsalva (axial scout); AOASC= aortic diameter, ascending aorta, level of pulmonary artery bifurcation (axial scout); AOTHO= thoracic descending aortic diameter, level of pulmonary artery bifurcation (axial scout); AODIA= descending aortic diameter, level of diaphragm (axial scout); PA= pulmonary artery, level of/just inferior to PA bifurcation (axial scout); SVC= superior vena cava, level of pulmonary artery bifurcation (axial scout); IVC= inferior vena cava, just inferior to entry into RA (axial scout; alternatively, sagittal scout); CO/CI= cardiac output/index, as LVEDV-LVESV/LVEDV x HR; HR= heart rate; AOSV/AOCO/AOCI= aortic stroke volume/CO/CI measured by velocity encoded flow phase contrast sequence, axial plane just below level of PA bifurcation; PASV/PACO/PACI= pulmonary stroke volume/CO/CI measured by velocity encoded flow phase contrast sequence, axial plane just below level of PA bifurcation (simultaneous with AO flow measures).

systolic volume series) (Table **3**). These differences translated into corresponding differences in derived measures (ejection fraction, stroke volume, mass, and volume and mass indexes). One example of ambiguous slice selection is shown in Fig. (**1**).

Table 3.   Effect of Basal Slice Selection Discrepancies Between Readers on Volume Measurements

Metric	N Discrepant/30	Different Basal Slice	Same Basal Slice	P Value
ΔLVEDV	11	19.2%	5.2%	<0.001
		25.1 ml	7.1 ml	<0.001
ΔLVESV	10	27.5%	13.1%	0.015
		11.8 ml	6.0 ml	0.002
ΔRVEDV	9	33.4%	17.9%	0.001
		50.1 ml	24.8 ml	0.001
ΔRVESV	12	48.8%	34.8%	0.004
		30.8 ml	21.7 ml	0.025

Δ=change, reader 1 *vs* reader 2. Abbreviations as in Table **1**.

Finally, visual comparisons revealed moderate differences in EC edge selection as the cause of most of residual measurement variability.

DISCUSSION

Study Overview. The study achieved the following major objectives: 1) It defined normal values for CMR ventricular volumes and masses and their derived metrics for a commonly used current CMR methodology, i.e., an operator-computer interactive system using the papillary muscle inclusion method. 2) It defined normal values and ranges for other commonly assessed cardiovascular structures (i.e., atria and great vessels - aorta, pulmonary artery, IVC, SVC) and dynamics (i.e., velocity encoded aortic and pulmonary artery flows), many of which have not previously been reported. 3) It assessed the extent of common interobserver variability in volume and mass measurements, allowing for improved consistency and reproducibility of these measures in clinical practice. Of these, the choice of the basal slice proved to be the most important.

Study Rationale and Impact. Quantitative measurements of cardiovascular anatomy and function are only of use if they are accurate and reproducible and if they can be compared with a normal, expected range of values. Incorrect interpretations, based on an incorrect comparative range of normal, can lead to additional unnecessary tests or to inappropriate, often damaging, disease labeling. Thus, the positive impact on current clinical practice of accurate, specific ranges of normals, including metrics beyond simply LV volumes and mass, cannot be over emphasized.

Recent personal experience of the investigators suggested an urgent need in clinical practice and research applications for contemporary hardware- and software-specific normal values for CMR metrics of interest. The ranges of normal suggested in study report templates from the manufacturer were found to represent a composite of sources, including CMR studies of limited size, those using earlier generation magnets and software, those using differing measurement methods (e.g., papillary muscle exclusion *vs* inclusion), and those derived solely from echocardiographic studies (without cross validation). For example, we noticed that the reported upper normal LV mass value was too high for our measure-

ment method [8], so that it did not distinguish those with clinical LV hypertrophy and hypertrophic cardiomyopathy from normals. Prior experience also suggested that echocardiographic measurements were inaccurate as surrogates for CMR measurements [9]. For example, atrial enlargement by echocardiography is designated for areas >18 cm [2,3,10,11], whereas we found areas up to 24 cm^2 to be within the normal range [12]. Finally, our experience suggested that important differences in measurements could occur because of differences in measurement approach, which was insufficiently defined to realize the excellent potential accuracy and reproducibility of CMR. These study results, reported above, together with other recent work, should narrow the gap between theoretical potential and practical reality.
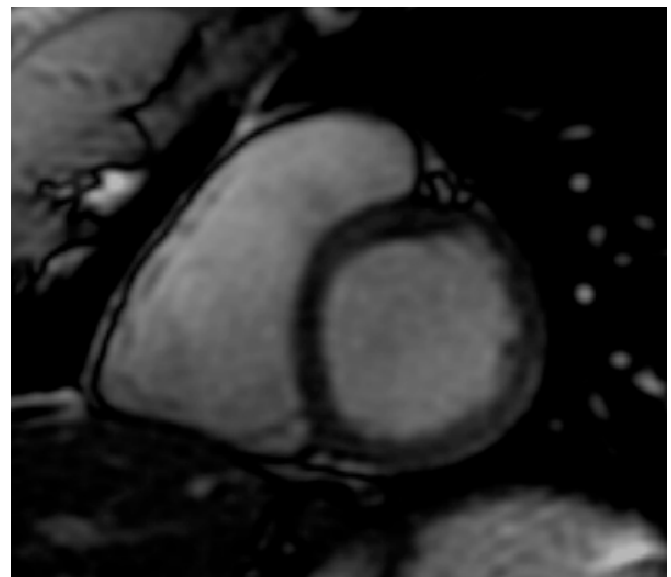
A
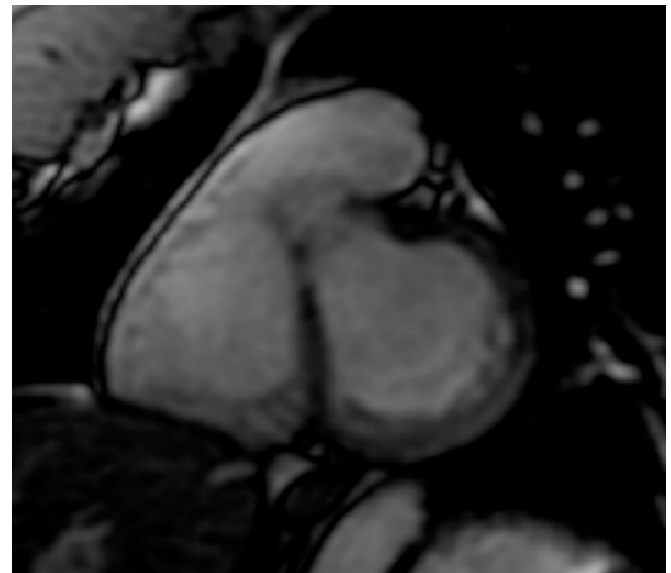
B

Fig. (1). Example of ambiguous choice for basal short-axis slice. One reader selected **A**, the other **B**, as the basal slice on the initial reading. Choosing **A**, the more basal slice, subtending the arc without full muscular rim, generally was found to lead to better correlation with flow-determined stroke volume.

Previous Work. Lorenz *et al*. published a seminal early study on normal CMR values [8], but it was of relatively limited size, used earlier hardware, pulse sequences (FLASH) [13], and software, and used measurement methodology differing from that of the present study. As a result, values for the normal range are importantly different than those we determined (compare Table **1** with Lorenz *et al*. [8]). Other earlier studies suffer from some of the same limitations in the current context [14]. Similarly, well accepted values for echocardiographic studies differ from those we determined using CMR for many values (e.g., atrial areas, as noted earlier) [10,12]

Recently, Maciera and Royal Brompton Hospital (London) collaborators have reported on normalized left ventricular systolic and diastolic function by SSFP CMR [7]. The study is larger (N=120 subjects, with 10 men and 10 women in each of 6 age deciles from 10 to 80 years) than previous studies and used contemporary scanner (1.5 T) and sequence (SSFP) techniques [13,15]. CMR analysis used a promising but distinct computer-based technology with blood pool thresholding to delineate the papillary muscles, which were excluded from chamber volume but included in mass measurement. As might be expected, mass measurements by this method are larger than results obtained with the method we used, emphasizing the need to use method-specific normal values. (That method [7] possibly somewhat overestimates and ours likely underestimates true mass.) Age related differences in normal values also were reported although most slopes were relatively shallow. Not included in their report were other metrics of interest that are reported here, i.e., for the right ventricle, atria, and great vessels. Alfakih also has reported SSFP normal ranges [16], but some of the same issues limit the comparison of their results with the present study.

One of the advantages of the method of Maciera *et al*. [7] is the ability to accurately track the descent of the mitral valve from diastole through systole in the long axis plane and use this to adjust basal volumes when applying Simpson's rule to short axis slices. This software analytical feature could potentially mitigate the basal slice-related interobserver measurement problem discovered in the present study.

Larger and more comprehensive studies of normal CMR values are to be encouraged for the future, and some are underway (e.g., MESA, Johns Hopkins University and US National Institutes of Health), but differences in normal values by analytical technique, rather than sample size and composition alone (i.e., by age, sex, race/ethnicity) still will apply, as noted here.

Interreader Measurement Differences. In this report, we identified 3 major sources of interreader differences in measurement: 1) choice of ED/ES frames; 2) choice of the most basal LV, RV short-axis slices; and, 3) approach to endocardial (EC) edge selection. These can have an important impact on interreader reproducibility, particularly choice of the basal slice, where a discrepancy of 1 slice was often present (one-third of series) before consensus conferencing and resulted in an average measurement difference of almost 15%.

To address this issue, we propose the following solutions. With respect to basal slice selection, we suggest: 1) specify more detailed algorithms to choose the basal slice, 2) require a ring with at least a 2/3 muscular wall; 3) "interpolate" on slices with partial rings, subtending the portion of the circumference not containing muscular wall; 4) compare right and left sided stroke volumes, as they should be equal in the absence of significant valvular regurgitation or intracardiac shunting; 5) compare volumetric- and flow-derived stroke volumes, as they should be similar. In addition, a more consistent approach to definition of endocardial edge location is needed, specifically, how to deal with papillary muscles and trabeculation. Our experience leads to the proposal that "floating" papillary muscles be included within the chamber and broadly attached muscle be included within the wall. Reliable computer-assisted algorithms should further help decrease variability in choice of the basal slice (e.g., determining the precise location of the mitral valve ring in 4-dimensional space and time) and in edge detection.

Study Limitations. This study was a prospective but observational evaluation. Interobserver but not intraobserver variability was assessed. Only a single study was performed, so that variability in technical factors was not assessed. The study was of relatively limited size, so that subtle differences in measurements across the age range studied could not be determined. Also, confidence intervals could be wider than larger studies might determine, leading to a wider range of normal; however, under estimating mild abnormalities is preferable to calling a normal study abnormal. Normal values determined for the present analytical methods likely differ from those determined by other methods (e.g., automated papillary muscle exclusion methods; orthogonal long-axis determinations of volume).

CONCLUSIONS

The normal range for CMR-derived volumes and mass is measurement-method specific. We have generated a clinically useful measurement set specific to operator-interactive analysis methodology that includes papillary muscles within the ventricular chambers. In addition, we provide normal values for over 50 other cardiovascular metrics of interest, many not previously reported. Also, we have discovered that differences in basal slice and endocardial edge selection constitute major potential sources of interoberver measurement variability. Hence, more formal and detailed rules for operator-interactive measurement together with improved computerized edge detection algorithms are needed to realize the potential of CMR to achieve reliably reproducible quantitative measurements of cardiovascular structures and function.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Manning WJ, Pennell DJ. Cardiovascular Magnetic Resonance. New York: Churchill Livingston; 2002.

[2]   Grothues F, Smith GC, Moon JCC, *et al*. 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.

[3] Feigenbaum H. Echocardiographic measurements and normal values. In: Feigenbaum H, ed. Echocardiography. Fifth edition ed. Philadelphia: Lea and Febiger; 1994: 658-95.

[4] Reynolds T. 2-D numbers at a glance. In: Echocardiographer's pocket reference. Phoenix: Arizona Heart Institute Foundation. School of Cardiac Ultrasound; 2000: 166.

[5] Gardin JM, Siscovick D, Anton-Culver H, *et al*. Sex, age and disease affect echocardiographic left ventricular mass and systolic function in the free-living elderly. Circulation 1995; 91: 1739-48.

[6] Bella JN, Palmieri V, Kitzman DW, *et al*. Gender difference in diastolic function in hypertension (the HyperGEN study). Am J Cardiol 2002; 89: 1052-6.

[7] Maceira AM, Prasad SK, Khan M, Pennell DJ. Normalized left ventricular systolic and diastolic function by steady state free precession cardiovascular magnetic resonance. J Cardiovasc Magn Reson 2006; 8: 417-26.

[8] Lorenz CH, Walker ES, Morgan VL, Klein SS, Graham TP. 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] Rodevan OBR, Ljosland M, Maehle J, Smith HJ, Ihlen H. Left atrial volumes assessed by three- and two-dimensional echocardiography compared to MRI estimates. Int J Card Imaging 1999; 15: 397-410.

[10] Feigenbaum H. Echocardiographic measurements and normal values. In: Feigenbaum H, Ed. Echocardiography. Fifth ed. Philadelphia: Lea and Febiger, 1994; 658-695.

[11] Gerstenblith G, Frederiksen J, Yin FC, Fortuin NJ, Lakatta EG, Weisfeldt ML. Echocardiographic assessment of a normal adult aging population. Circulation 1977; 56: 273-8.

[12] Anderson JL, Horne B, Pennell DJ. Atrial Dimensions in Health and Left Ventricular Disease Using Cardiovascular Magnetic Resonance. J Cardiovasc Magn Reson 2005; 7: 671-675.

[13] Moon JCC, Lorenz CH, Francis JM, Smith GC, Pennell DJ. Breath-hold FLASH and FISP cardiovascular MR imaging: left ventricular volume differences and reproducibility. Radiology 2002; 223: 789-97.

[14] Sandstede J, Lipke C, Beer M, *et al*. Age- and gender-specific differences in left and right ventricular cardiac function and mass determined by cine magnetic resonance. Eur Radiol 2000; 10: 438-42.

[15] Carr JC, Simonetti O, Bundy J, Li D, Pereles S, Finn JP. Cine angiography of the heart with segmented true fast imaging in steady-state precession. Radiology 2001; 219: 828-34.

[16] Alfakih K, Plein S, Thiele H, Jones T, Ridgway JP, Sivananthan MU. Normal human left and right ventricular dimensions for MRI as assessed by turbo gradient echo and steady-state free precession imaging sequences. J Magn Reson Imaging 2003; 17: 323-9.