The Reproducibility and Absolute Values of Echocardiographic Measurements of Left Ventricular Size and Function in Children Are Algorithm Dependent

Renee Margossian, MD, Shan Chen, MS, Lynn A. Sleeper, ScD, Lloyd Y. Tani, MD, Girish Shirali, MD, Fraser Golding, MD, Elif Seda Selamet Tierney, MD, Karen Altmann, MD, Michael J. Campbell, MD, Anita Szwast, MD, Angela Sharkey, MD, Elizabeth Radojewski, RN, and Steven D. Colan, MD, for the Pediatric Heart Network Investigators, *Boston and Watertown, Massachusetts; Salt Lake City, Utah; Charleston, South Carolina; Toronto, Ontario, Canada; New York, New York; Durham, North Carolina; Philadelphia, Pennsylvania; and St Louis, Missouri*

Background: Several quantification algorithms for measuring left ventricular (LV) size and function are used in clinical and research settings. The aims of this study were to investigate the effects of measurement algorithm and beat averaging on the reproducibility of measurements of the left ventricle and to assess the magnitude of agreement among the algorithms in children with dilated cardiomyopathy.

Methods: Echocardiograms were obtained in 169 children from eight clinical centers. Inter- and intrareader reproducibility was assessed on measurements of LV volumes using the biplane Simpson, modified Simpson, and 5/6 × area × length (5/6AL) algorithms. Percentage error was calculated as inter- or intrareader difference/mean × 100. Single-beat measurements and the three-beat average (3BA) were compared. Intraclass correlation coefficients were calculated to assess agreement.

Results: Single-beat interreader reproducibility was lowest (percentage error was highest) using biplane Simpson; 5/6AL and modified Simpson were similar but significantly better than biplane Simpson ($P < .05$). Single-beat intrareader reproducibility was highest using 5/6AL ($P < .05$). The 3BA improved reproducibility for almost all measures ($P < .05$). Reproducibility in both single-beat and 3BA values fell with greater LV dilation and systolic dysfunction ($P < .05$). Intraclass correlation coefficients were >0.95 across measures, although absolute volume and mass values were systematically lower for biplane Simpson compared with modified Simpson and 5/6AL.

Conclusions: The reproducibility of LV size and functional measurements in children with dilated cardiomyopathy is highest using the 5/6AL algorithm and can be further improved by using the 3BA. However, values derived from different algorithms are not interchangeable. (J Am Soc Echocardiogr 2015;28:549-58.)

Keywords: Cardiomyopathy, Echocardiography, Pediatric, Reproducibility

Echocardiographic measures of left ventricular (LV) size and systolic function are widely used as end points in clinical trials. However, the limited availability of data concerning the reproducibility of quantitative indices of ventricular function in pediatric dilated cardiomyopathy (DCM) is an impediment to controlled trials of therapy for ventricular dysfunction in children. Several algorithms for measuring LV volumes using two-dimensional echocardiographic methods are common in clinical and research use. These methods include the 5/6 × area × length (5/6AL), modified Simpson (MS), and biplane Simpson algorithms. Although the American Society of Echocardiography (ASE) has recommended the biplane Simpson (also known as the biplane method of disks) algorithm as the

From the Boston Children's Hospital and Harvard Medical School, Boston, Massachusetts (R.M., E.S.S.T., S.D.C.); the New England Research Institutes, Watertown, Massachusetts (S.C., L.A.S.); Primary Children's Medical Center and University of Utah, Salt Lake City, Utah (L.Y.T.); the Medical University of South Carolina, Charleston, South Carolina (G.S.); The Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada (F.G., E.R.); Columbia University Medical Center, New York, New York (K.A.); Duke University Medical Center, Durham, North Carolina (M.J.C.); Children's Hospital of Philadelphia, Philadelphia, Pennsylvania (A.S.); and Washington University School of Medicine, St Louis, Missouri (A.S.).

This study was supported by U01 grants from the National Heart, Lung, and Blood Institute (HL068269, HL068270, HL068279, HL068281, HL068285, HL068292, HL068290, and HL068288). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the National Heart, Lung, and Blood Institute.

Reprint requests: Renee Margossian, MD, Boston Children's Hospital, Department of Cardiology, 300 Longwood Avenue, Boston, MA 02115 (E-mail: *renee.margossian@cardio.chboston.org*).

0894-7317/$36.00

http://dx.doi.org/10.1016/j.echo.2015.01.014

Abbreviations

ASE = American Society of Echocardiography

BSA = Body surface area

DCM = Dilated cardiomyopathy

5/6AL = 5/6 × area × length

ICC = Intraclass correlation coefficient

LV = Left ventricular

LVEDV = Left ventricular end-diastolic volume

LVEF = Left ventricular ejection fraction

LVESV = Left ventricular end-systolic volume

MS = Modified Simpson

3BA = Three-beat average

VVV = Ventricular Volume Variability

approach of choice for LV volume quantification in adults,[1] the applicability of this recommendation to pediatric populations is unknown.

The objectives of this analysis were to (1) determine the impact of the method of calculating LV volumes on interreader and intrareader reproducibility of measured and calculated variables in children with DCM, (2) determine whether averaging multiple beats improves reproducibility, (3) assess the agreement among measurements by algorithm, and (4) determine whether the severity of cardiomyopathy affects the reproducibility of LV volumes.

METHODS

The Ventricular Volume Variability (VVV) study was a multicenter observational study of pediatric subjects with stable DCM undertaken by the National Heart, Lung, and Blood Institute–sponsored Pediatric Heart Network. Enrolled subjects were followed for 18 months, and a study protocol echocardiogram was obtained at each clinical visit during this time. Inclusion and exclusion criteria are listed in the online Appendix. The study was conducted in accordance with the guidelines of the Pediatric Heart Network's Data and Safety Monitoring Board and of each center's institutional review board. Full details of the study design have been previously published.[2]

The primary aim of the VVV study was to evaluate the longitudinal variance of echocardiographic indices of LV size and function. Subjects with histories of DCM by chart review were approached for consent for participation at the time of a clinical evaluation. Those subjects who met full inclusion criteria on the basis of the baseline study echocardiogram were eligible for follow-up echocardiography to determine longitudinal variability. The data from baseline echocardiograms obtained in patients who did not meet exclusion criteria but who also did not meet the dilation and/or dysfunction criteria (in other words, data obtained in those subjects whose echocardiograms had improved sufficiently to not meet entrance criteria for dilation and/or dysfunction since the diagnosis of DCM was made) were retained in the database and represent a normalized or near normalized population. Inclusion of these examinations permitted the analyses to be performed across a broader range of disease severity. For purposes of this report, only baseline evaluations are included.

Consented subjects underwent study echocardiography performed by sonographers at each site who were specifically trained on the standardized protocol for image acquisition. At least three cardiac cycles were recorded for each parameter. Height and weight were measured, and body surface area (BSA) was calculated using the Haycock formula.[3] All baseline echocardiograms were submitted to the data-coordinating center and forwarded to the echocardiography core laboratory.

At the echocardiography core laboratory, two readers performed measurements on each echocardiogram to assess interreader reproducibility. The protocol specified 150 measured and derived parameters on each study.[2] One ECL reader repeated all measurements 1 month later to assess intrareader reproducibility. All measurements were performed using custom Digital Imaging and Communications in Medicine software (EchoTrace; Marcus Laboratories, Boston, MA).

LV end-diastolic volume (LVEDV), LV end-systolic volume (LVESV), LV mass, and LV ejection fraction (LVEF) were calculated using three common algorithms. The ASE-recommended biplane Simpson method[1] used areas from apical four-chamber and apical two-chamber views (Figure 1A). For the MS approach,[4] we used an apical four-chamber area and a parasternal short-axis area (Figure 1B). The 5/6AL algorithm[5] required the length of the left ventricle from the apical four-chamber and a parasternal short-axis area (Figure 1C).

Statistical Methods

Definitions. *Inter- and Intrareader Reproducibility.*–In these analyses, the outcome measure is the percentage error of the mean. To evaluate interreader reproducibility, the absolute difference ("error") between the measurements made by the primary and secondary readers was divided by the mean of those two measurements. To evaluate intrareader reproducibility, the difference ("error") between the immediate and 1-month repeat measurements made by the primary reader was divided by the mean of those two measurements. Single-beat percentage error was based on the measurements obtained from the first beat, and three-beat average (3BA) percentage error was based on the average of all three measurements.

Reduction in Error.–To quantify the reduction in error that was due to beat averaging, a "percentage reduction in percentage error" term (abbreviated to "error reduction") was calculated using 3BA compared with single-beat averages as

$$(\text{Single} - \text{beat percentage error} - \text{3BA percentage error})$$

$$/\text{Single} - \text{beat percentage error} \times 100\%.$$

We fit a mixed regression model (fixed effect for method and random effect for subject) to assess whether reduction in error occurred using the 3BA compared with the single-beat average.

Impact of Disease Severity, Age, and Body Size on Reproducibility.–Tests of interaction in the respective regression models were used to assess whether differences in reproducibility (percentage error) among the three algorithms were altered by disease severity (LVEDV, LVEF), age, and BSA.

Z Scores.–To adjust LV measurements to account for the effects of body size and age in this cohort of children, *Z* scores were used, as recommended by the ASE.[5] The *Z*-score normative relationships for LVEDV were based on 5/6AL values (i.e., they are not algorithm-specific), and they were therefore calculated only for the 5/6AL volumes (LVEDV *Z* score) and were used only for entry criteria and severity stratification.

To determine the level of agreement between the absolute measurements yielded by each of the algorithms, intraclass correlation coefficients (ICCs) and Bland-Altman plots were used, on the basis of the initial reading by the primary core laboratory reviewer. The ICC ranges from 0 to 1, with a value of 1 indicating that all of the variability in the measurement is due to random measurement error for a subject and is not due to the algorithm used. From Bland-Altman analyses and plots, the bias (systematic difference between the measurements) and limits of agreement (average difference ± 2 standard