Relationship of Echocardiographic *Z* Scores Adjusted for Body Surface Area to Age, Sex, Race, and Ethnicity

The Pediatric Heart Network Normal Echocardiogram Database

Leo Lopez, MD; Steven Colan, MD; Mario Stylianou, PhD; Suzanne Granger, MS; Felicia Trachtenberg, PhD; Peter Frommelt, MD; Gail Pearson, MD, ScD; Joseph Camarda, MD; James Cnota, MD; Meryl Cohen, MD; Andreea Dragulescu, MD, PhD; Michele Frommelt, MD; Olukayode Garuba, MS, CCRP; Tiffanie Johnson, MD; Wyman Lai, MD, MPH; Joseph Mahgerefteh, MD; Ricardo Pignatelli, MD; Ashwin Prakash, MD; Ritu Sachdeva, MD; Brian Soriano, MD; Jonathan Soslow, MD, MSCI; Christopher Spurney, MD; Shubhika Srivastava, MBBS; Carolyn Taylor, MD; Poonam Thankavel, MD; Mary van der Velde, MD; LuAnn Minich, MD; for the Pediatric Heart Network Investigators*

Background—Published nomograms of pediatric echocardiographic measurements are limited by insufficient sample size to assess the effects of age, sex, race, and ethnicity. Variable methodologies have resulted in a wide range of *Z* scores for a single measurement. This multicenter study sought to determine *Z* scores for common measurements adjusted for body surface area (BSA) and stratified by age, sex, race, and ethnicity.

Methods and Results—Data collected from healthy nonobese children ≤18 years of age at 19 centers with a normal echocardiogram included age, sex, race, ethnicity, height, weight, echocardiographic images, and measurements performed at the Core Laboratory. *Z* score models involved indexed parameters (X/BSA^{α}) that were normally distributed without residual dependence on BSA. The models were tested for the effects of age, sex, race, and ethnicity. Raw measurements from models with and without these effects were compared, and <5% difference was considered clinically insignificant because interobserver variability for echocardiographic measurements are reported as ≥5% difference. Of the 3566 subjects, 90% had measurable images. Appropriate BSA transformations (BSA^{α}) were selected for each measurement. Multivariable regression revealed statistically significant effects by age, sex, race, and ethnicity for all outcomes, but all effects were clinically insignificant based on comparisons of models with and without the effects, resulting in *Z* scores independent of age, sex, race, and ethnicity for each measurement.

Conclusions—Echocardiographic *Z* scores based on BSA were derived from a large, diverse, and healthy North American population. Age, sex, race, and ethnicity have small effects on the *Z* scores that are statistically significant but not clinically important.

Key Words: body surface area ■ echocardiography ■ heart ■ nomograms ■ sample size

Echocardiography is crucial for the evaluation of heart diseases, because treatment decisions frequently rely on accurate determination of the sizes of cardiovascular structures.[1,2] Reference values must be readily available for clinicians and researchers to distinguish normal from abnormal findings. Previous studies suggest that measurements in normal children are affected by body size, age, sex, and

See Editorial by Simpson and Chubb
See Clinical Perspective

Received July 27, 2017; accepted October 19, 2017.

From the Nicklaus Children's Hospital, Miami, FL (L.L.); Boston Children's Hospital, MA (S.C., A.P.); National Heart, Lung, and Blood Institute, Bethesda, MD (M.S., G.P.); New England Research Institutes, Watertown, MA (S.G., F.T.); Children's Hospital of Wisconsin, Milwaukee (P.F., M.F.); Ann & Robert Lurie Children's Hospital, Chicago, IL (J. Camarda); Cincinnati Children's Hospital Medical Center, OH (J. Cnota); Children's Hospital of Philadelphia, PA (M.C.); Hospital for Sick Children, Toronto, ON, Canada (A.D.); Texas Children's Hospital, Houston (O.G., R.P.); Riley Hospital for Children at Indiana University Health, Indianapolis (T.J.); Children's Hospital of Orange County, CA (W.L.); Children's Hospital at Montefiore, Bronx, NY (J.M.); Children's Healthcare of Atlanta, GA (R.S.); Seattle Children's Hospital, WA (B.S.); Vanderbilt University Medical Center, Nashville, TN (J.S.); Children's National Health System, Washington, DC (C.S.); Mount Sinai Medical Center, New York, NY (S.S.); Medical University of South Carolina, Charleston (C.T.); Children's Health Dallas, TX (P.T.); CS Mott Children's Hospital, Ann Arbor, MI (M.v.d.V.); and University of Utah, Salt Lake City (L.M.).

*A list of all Pediatric Heart Network Investigators is given in the Data Supplement.

The contents of this publication represent the views of the authors and do not necessarily represent official positions of the National Heart, Lung, and Blood Institute or the National Institutes of Health.

The Data Supplement is available at http://circimaging.ahajournals.org/lookup/suppl/doi:10.1161/CIRCIMAGING.117.006979/-/DC1.

Correspondence to Leo Lopez, MD, Nicklaus Children's Hospital, 3100 SW 62nd Ave, Miami, FL 33155. E-mail leolopezmd@gmail.com

race,[3–21] though most focus on the effects of body size using cardiovascular allometry (relationship between cardiovascular growth and total body growth) and *Z* scores.[10–19] With increasing use of *Z* scores in echocardiography, the limitations have become apparent.[22–24]

Cantinotti et al[22] revealed wide *Z* score variation for the same measurement when evaluating published normal databases, many with small sample sizes, few neonates, and heterogeneous methodologies using variable body size parameters and regression equations. For example, a mitral diameter of 11 mm in a boy with a body surface area (BSA) of 0.3 m^2 can correspond to a *Z* score of −3.5 to +4.8. Many studies also failed to address the problem of nonconstant variance (heteroscedasticity).[22–24] Colan et al[25] highlighted the reproducibility of echocardiographic measurements, potentially creating additional challenges to establishing normal databases. Most studies have reported interobserver variability as percent differences of 5% to 10% for semilunar and >10% for atrioventricular valvar measurements.[3–7,26,27]

Currently, normal echocardiographic reference values adjusted for body size, age, sex, race, and ethnicity do not exist. The Pediatric Heart Network sought to determine *Z* scores for common measurements in a large group of racially diverse healthy children by evaluating the relationship between measurements and BSA as well as the effects of age, sex, race, and ethnicity on this relationship.

Methods

The detailed methods used for measurement performance and data analysis will be made available on request from the Pediatric Heart Network to other researchers. In addition, the regression equations and a *Z* score calculator will be available on the Pediatric Heart Network website (www.pediatricheartnetwork.com).

Study Design

Demographic and clinical data and echocardiographic images were collected at 19 North American centers. Because all submissions were deidentified, most children were retrospectively enrolled under a waiver of consent after Institutional Review Board or Research Ethics Board approval. Race/ethnicity information was not routinely obtained at 1 center and was collected prospectively for eligible subjects after local regulatory approval. Some centers were able to perform research echocardiograms without charge and prospectively enrolled healthy children after Institutional Review Board approval.

Study Population

Healthy children ≤18 years of age with a normal, high-quality echocardiogram and documentation of height, weight, sex, race, and ethnicity were eligible. Exclusion criteria (Table I in the Data Supplement) included structural heart disease, abnormal electrocardiographic findings, systemic disorder with cardiovascular manifestations, prematurity because of a high prevalence of hemodynamically significant cardiovascular and respiratory pathology, obesity because of reported associated cardiovascular pathology,[28,29] and a family history of nonischemic cardiomyopathy or congenital left-sided heart disease.[30,31] An adjudication committee evaluated anatomic variants (Table 1), and normal or hemodynamically insignificant findings were included.

Self-reported race/ethnicity information was divided into 3 categories: whites, blacks, and others (Hispanics, Asians, Pacific Islanders, Native Americans, and Multiracial). Age was divided into 6 categories (<1 month, 1 month–3 years, 3–6 years, 6–12 years, 12–16 years, and 16–18 years) to assure adequate enrollment across the full pediatric age range (particularly during periods of increased growth velocity), but age was treated as a continuous variable during the analyses. Thirty-six study groups were created from the 3 race, 6 age, and 2 sex categories. Sample size calculations were performed to reasonably estimate the population mean and SD for each measurement.[32] Specifying a margin of error for the mean of 22% of the SD required 80 echocardiograms per group. Because ≥80% of submitted studies were expected to contain the necessary images for each measurement, the target was 100 subjects per group.

Table 1. Clarification of Echocardiographic Findings

Included (If Otherwise Normal Intracardiac Anatomy)	Excluded
Patent foramen ovale	Cardiac malposition
Tiny patent ductus arteriosus	Left superior vena cava
Mild peripheral pulmonic stenosis without branch pulmonary artery hypoplasia in infancy	Interrupted inferior vena cava
Tiny coronary artery fistula	Abnormal coronary artery origin
Retroaortic innominate vein	Absent aortic arch image
Common origin of right innominate and left carotid arteries	Right aortic arch
Chest wall deformity	Aberrant subclavian artery
Clinical suspicion of connective tissue disorder without evidence for connective tissue disorder	Direct origin of a vertebral artery from the aortic arch
Clinical suspicion of Kawasaki disease with normal coronary arteries and no history of Kawasaki disease treatment	

Echocardiographic Studies

All echocardiograms were in Digital Imaging and Communications in Medicine format with ≥2-beat clips. Images were deidentified using the Match Plus Program (Booz Allen Hamilton, McLean, VA) and submitted to the Core Laboratory where measurements were performed using published pediatric quantification standards (Table 2).[1] Pulmonary annular diameters were performed in short- and long-axis parasternal views; a single predesignated view was used for all other measurements. Measurements were performed off-line (TomTec, Unterschleissheim, Germany) by 1 of 2 Core Laboratory sonographers and reviewed by the Director. The echocardiogram was included if the Core Laboratory could perform all required measurements and measurements in at least 1 of the 3 optional categories in Table 2 (structures not routinely measured in normal studies).

Intraobserver variability was evaluated with blinded repeat measurements of aortic annular, root, sinotubular junction, and ascending aortic diameters and left ventricular (LV) end-diastolic area in 120 subjects. Depending on the true proportion of matched measurements (assuming a 90% to 50% range), 120 subjects provided a 90% confidence interval for the true proportion with a reasonable margin of error (0.045–0.075). Measurement variability was tested using intraclass correlation coefficients and Pearson correlations.

Statistical Analysis

Because most clinicians use the BSA formulas by Haycock[33] and by Gehan and George,[34] the calculated BSAs using both formulas were compared by Pearson correlation. Because the goal was to calculate *Z* scores based on BSA while accounting for the effects of age, sex, race, and ethnicity, a *P* value of 0.05 was used to determine significant effects, and a *P* value of 0.01 was used to determine significant interactions among the effects. Published reproducibility thresholds suggest that measurement variability may be responsible for up to 5% of measurement differences.[3–7,25–27] Therefore, clinical

Table 2. Required and Optional Parameters; Echocardiographic Views or Formulas; Percentages With Measurable Images; Body Surface Area Transformation α; Effects of Age, Sex, and Race on Measurements; and Final *Z* Score Models

Parameter	Req	Echocardiographic View or Formula	% Av, %	α	Effects	% Diff, %	Slope Δ	Mean	SD
MV_{AP}, cm	Yes	Parasternal long axis	100	0.50	A	1.33	N	2.31	0.24
MV_{LAT}, cm	Yes	Apical 4 chamber	100	0.50	A S	2.36	N	2.23	0.22
MVA, cm^2	Yes	$\pi/4\times MV_{AP}\times MV_{LAT}$	100	1.00	A	3.68	N	4.06	0.68
TV_{AP}, cm	Yes	Parasternal long axis	100	0.50	A S A×S	1.33	N	2.36	0.28
TV_{LAT}, cm	Yes	Apical 4 chamber	100	0.50	A R S A×S	4.25	N	2.36	0.29
TVA, cm^2	Yes	$\pi/4\times TV_{AP}\times TV_{LAT}$	100	1.00	A R S A×R A×S	3.87	N	4.39	0.83
ANN, cm	Yes	Parasternal long axis	100	0.50	A R S A×S S×R	1.55	Y	1.48	0.14
ROOT, cm	Yes	Parasternal long axis	100	0.50	A R S A×R A×S	1.92	N	2.06	0.18
STJ, cm	Yes	Parasternal long axis	100	0.50	A R S S×R	1.35	Y	1.69	0.16
AAO, cm	Yes	Parasternal long axis	100	0.50	A R S	1.13	N	1.79	0.18
$ARCH_{PROX}$, cm	No*	Suprasternal long axis	80	0.50	A R S	4.72	Y	1.53	0.23
$ARCH_{DIST}$, cm	No*	Suprasternal long axis	97	0.50	A R	2.36	N	1.36	0.19
ISTH, cm	No*	Suprasternal long axis	97	0.50	A R	1.89	Y	1.25	0.18
LMCA, mm	No*	Parasternal short axis	90	0.45	A R S	2.78	N	2.95	0.57
LAD, mm	No*	Parasternal short axis	78	0.45	A S	2.59	N	1.90	0.34
RCA, mm	No*	Parasternal short axis	91	0.45	R S	2.42	N	2.32	0.55
PV_{SAX}, cm	Yes†	Parasternal short axis	71	0.50	A R S A×R A×S S×R A×S×R	1.42	Y	1.91	0.24
PV_{LAX}, cm	Yes†	Parasternal long axis	90	0.50	A R S A×S	4.59	Y	2.01	0.28
MPA, cm	No*	Parasternal short axis	94	0.50	A A×S S×R	2.00	Y	1.82	0.24
RPA, cm	No*	Parasternal short axis	93	0.50	A S	1.18	Y	1.07	0.18
LPA, cm	No*	Parasternal short axis	89	0.50	A S	0.87	Y	1.10	0.18
LVEDD, cm	Yes	Parasternal short axis	100	0.45	A R S A×S	1.60	Y	3.89	0.33
LVPWT, cm	Yes	Parasternal short axis	100	0.40	A R S A×R A×S	2.82	Y	0.57	0.09
LVST, cm	Yes	Parasternal short axis	100	0.40	A R S A×R A×S	2.88	Y	0.58	0.09
LVEDL, cm	Yes	Apical 4 chamber	100	0.45	A R S A×S	1.29	N	6.31	0.46
$LVEDL_{EPI}$, cm	Yes	Apical 4 chamber	100	0.45	A R S A×S	1.22	N	6.87	0.45
LVEDA, cm^2	Yes	Parasternal short axis	100	0.90	A S A×S	2.84	Y	11.91	1.89
$LVEDA_{EPI}$, cm^2	Yes	Parasternal short axis	100	0.90	A R S A×S	2.79	N	20.00	2.59
LVEDV, mL	Yes	5/6×LVEDA×LVEDL	100	1.30	A R S A×S	4.97	Y	62.02	11.94
$LVEDV_{EPI}$, mL	Yes	$5/6\times LVEDA_{EPI}\times LVEDL_{EPI}$	100	1.30	A S A×S	4.20	N	113.14	17.85
LVM, g	Yes	$1.05\times(LVEDV_{EPI}-LVEDV)$	100	1.25	A R S A×S	5.00	N	53.02	9.06
LVMTV, g/mL‡	Yes	LVM/LVEDV	100	0‡	A R	4.52	Y	0.88	0.16
LVTTD‡	Yes	LVPWT/LVEDD	100	0‡	A R A×R	4.07	Y	0.15	0.03
LVSI‡	Yes	LVEDL/LVEDD	100	0‡	A R S	1.05	N	1.63	0.17

% Av indicates percent of studies with available images; % Diff, percent differences of mean indexed values for models with and without significant effects and interactions; A, age; A×R, interaction between age and race; A×S, interaction between age and sex; A×S×R, interaction among age, sex, and race; AAO, ascending aortic diameter; ANN, aortic annular diameter; $ARCH_{DIST}$, distal transverse arch diameter; $ARCH_{PROX}$, proximal transverse arch diameter; effects, statistically significant effects and interactions with multivariable regression; BSA, body surface area; ISTH, aortic isthmus diameter; LAD, proximal left anterior descending coronary artery diameter; LMCA, left main coronary artery diameter; LPA, left pulmonary artery diameter; LV, left ventricular; LVEDA, LV short-axis end-diastolic endocardial area; $LVEDA_{EPI}$, LV short-axis end-diastolic epicardial area; LVEDD, LV short-axis end-diastolic endocardial diameter; LVEDL, LV long-axis end-diastolic endocardial length; $LVEDL_{EPI}$, LV long-axis end-diastolic epicardial length; LVEDV, LV short-axis end-diastolic endocardial volume; $LVEDV_{EPI}$, LV short-axis end-diastolic epicardial volume; LVM, LV mass; LVMTV, LV mass:volume ratio; LVPWT, LV short-axis end-diastolic posterior wall thickness; LVSI, LV sphericity index; LVST, LV short-axis end-diastolic septal thickness; LVTTD, LV thickness:dimension ratio; mean, mean indexed parameter value; MPA, main pulmonary artery diameter; MVA, mitral area; MV_{AP}, mitral anteroposterior diameter; MV_{LAT}, mitral lateral diameter; PV_{LAX}, pulmonary annular long-axis diameter; PV_{SAX}, pulmonary annular short-axis diameter; R, race; RCA, proximal right coronary artery diameter; req, required parameter; ROOT, aortic root diameter; RPA, right pulmonary artery diameter; S, sex; S×R, interaction between sex and race; slope Δ, slope change at ≈6 y age in indexed parameter vs age relationships; STJ, sinotubular junction; TVA, tricuspid area; TV_{AP}, tricuspid anteroposterior diameter; TV_{LAT}, tricuspid lateral diameter; and α, exponent for BSA transformation (BSA^{α}).

*For optional measurements, each study must contain all 3 measurements from only one of the following groups: (1) ARCHPROX, ARCHDIST, ISTH; (2) LMCA, LAD, RCA; and (3) MPA, RPA, LPA.

†For pulmonary annular diameters, a parasternal short- and/or long-axis measurement was required.

‡These parameters do not have a significant relationship with BSA, so BSA is not used in *Z* score derivation.

significance was defined as a difference of at least 5% between actual and predicted measurement values using models with and without the significant effects.

Based on physiologically driven methodologies for indexing cardiovascular measurements,[9,10] models with nonlogarithmic BSA transformations (BSA^{α}) and no measurement transformations were used, beginning with results from a prior study.[10] The exponent α for each measurement was tested for the following criteria:

- The indexed parameter (X/BSA^{α}) had a normal distribution.
- There was no residual relationship between X/BSA^{α} and BSA (the slope of the relationship was not significantly different from zero).

If the slope was statistically significantly different from zero, clinical significance was tested by creating a zero-slope line at the mean X/BSA^{α} and comparing raw values from the nonzero- and zero-slope lines at the first and third BSA quartiles; if the percent difference was <5% at both quartiles, the persistent relationship between X/BSA^{α} and BSA was considered clinically insignificant. If the percent difference was ≥5% or X/BSA^{α} was not normally distributed, other exponents were tested.

Once BSA^{α} was chosen, multivariable regression assessed the linear effects of age, sex, and race and their interactions. Race was coded as a 3-level categorical variable in all regression models involving the significant main effects and interactions, with the race category of white chosen as the reference category. Ethnicity represented only a fraction of 1 race category and was not included initially. Backwards elimination model selection, excluding BSA as a predictor, determined the final model. Higher order interactions were considered first and removed if insignificant. Lower order interactions and main effects were kept even if insignificant when the effect was part of a significant higher order interaction.

If an interaction was statistically significant, predicted values transformed to raw measurements from the model with the interaction were compared with those from the model without the interaction. A 1-sided *t* test was used on the absolute proportion difference between the 2 models with a null hypothesis of a mean proportion ≥0.05 against the alternative hypothesis of a mean proportion <0.05. An absolute mean percent difference between predicted values <5% was considered clinically insignificant. If an effect was statistically significant, a similar method determined clinical significance. Predicted values from a model containing statistically significant effects were compared with those from a model without effects, and a mean percent difference <5% by 1-sided *t* test was considered clinically insignificant. A similar secondary analysis explored the effect of ethnicity, comparing raw values from a model that included ethnicity as a predictor and one without predictors, and a mean percent difference <5% was considered clinically insignificant.

Age as a continuous variable was tested for nonlinear effects by plotting X/BSA^{α} against age, first with nonparametric locally weighted scatterplot smoothing curve fitting, then with piecewise linear regression. Discrete discrepancies in slope were tested for clinical significance by comparing predicted values from a model that included the separate slopes and one without changes in slope; again, a mean percent difference <5% was considered clinically insignificant. Finally, the mean and SD (*Z* scores) of the indexed parameters were determined while accounting for any clinically significant effects and interactions. The nonindexed parameters were plotted against BSA along with lines depicting the mean and 2 SDs above and below the mean.

Results

Of the 3566 subjects, 3215 (90%) had adequate images. Race data revealed 35% whites, 31% blacks, and 34% others. Ethnicity data revealed 25% Hispanic, 70% non-Hispanic, and 5% unknown. All study groups reached ≥complete enrollment (≥80 subjects with measurable images) except black girls age <1 month, 3 to 6 years, and 16 to 18 years, black boys age <1 month, and other girls age 16 to 18 years (Table II in the Data Supplement). For all the required parameters, eligible images were available in 100% (Table 2). For the pulmonary annulus, eligible images were available in 71% for short- and 90% for long-axis diameters. For optional measurements, eligible images were available in 78% to 97%. Intraobserver variability at the Core Laboratory was low with an intraclass correlation coefficient of 1.00 and Pearson correlations >0.99 for all 5 parameters.

BSA Transformation for Indexed Parameters

Comparison of BSA calculations using the Haycock and Gehan/George formulas revealed a Pearson correlation >0.99. The Haycock formula was used for all analyses since prior reports have shown it to be the best predictor of cardiovascular sizes.[10,17] LV mass:volume ratio, thickness:dimension ratio, and sphericity index did not have a clinically significant relationship with BSA, so these parameters were not indexed to BSA. For the other parameters, the selected BSA transformation resulted in a normal distribution for all indexed parameters (X/BSA^{α}), but most relationships between X/BSA^{α} and BSA were statistically significant with a nonzero slope (Table 2). However, comparison of the actual parameter values against the predicted values for a zero-slope model at the first and third BSA quartiles revealed an absolute percent difference <5% for all parameters (Table III in the Data Supplement). For example, the percent differences for the linear measurements in centimeters involved raw value differences that were all <1 mm, suggesting that the differences could be attributable to measurement variability. Therefore, all residual relationships between X/BSA^{α} and BSA were deemed clinically insignificant.

Model Selection

Multivariable regression for all parameters revealed statistically significant effects by age, sex, and race as well as significant interactions (Table 2). However, comparison of these results against models without effects or interactions revealed that none involved clinically significant differences. When considering the amount of variance explained by the models (R^2), the maximum increase in R^2 when including age, sex, and race in addition to BSA was 0.018, suggesting that the added contribution of age, sex, and race to predicting these parameters was minimal. For the 3 parameters not indexed to BSA, the maximum R^2 of models including age, sex, and race was 0.089, again suggesting little contribution of these factors. Two hypothetical subjects were created to highlight this point: an 18-month-old black boy at the first BSA quartile (0.43 m^2) and a 14-year-old white boy at the third BSA quartile (1.51 m^2). The predicted mean aortic root diameters for each subject in the model with the effects were 1.33 and 2.62 cm, compared with 1.35 and 2.53 cm in the model without the effects. These differences were <1 mm highlighting the absence of clinically significant differences between the models. A similar exercise for LV end-diastolic diameters revealed predicted mean values of 2.65 and 4.79 cm from the model with the effects and 2.65 and 4.68 cm from the model without the effects, again emphasizing the absence of clinically significant differences.

Testing for nonlinear effects of age revealed an apparent transition in slopes at ≈6 years of age in 16/34 parameters (Table 2). However, comparing predicted values revealed an absolute percent difference <5% for all parameters, indicating that age as a continuous variable did not have a clinically significant effect. Assessment of the relationship between ethnicity and indexed parameters also resulted in clinically insignificant effects.

After the effects of age, sex, race, and ethnicity were deemed clinically insignificant, the final models were established (Table 2), underscoring the absence of heteroscedasticity in the relationship between the indexed parameter and BSA. The nonindexed parameters were then plotted against BSA with lines representing the mean values and 2 SDs above and below the mean (Figure I in the Data Supplement), revealing the nonconstant variance (heteroscedasticity) of this relationship. Based on the models, the *Z* score of a measurement for a specific BSA can be calculated from Table 2 by using the specified α, mean, and SD for that parameter:

$$Z = \frac{\left[\left(parameter \,/\, BSA^{\alpha}\right) - \left(mean\ value\ of\ indexed\ parameter\right)\right]}{SD\ of\ indexed\ parameter}$$

For the boy with a BSA of 0.3 m^2 and a mitral diameter of 11 mm, the *Z* score is calculated as −1.0 based on the values for α (0.50), mean (2.23), and SD (0.22) of the indexed parameter. The *Z* scores for LV mass:volume ratio, thickness:dimension ratio, and sphericity index can be calculated using the raw values without adjusting for BSA.

Discussion

This is the first study with adequate sampling to evaluate the effects of age, sex, race, and ethnicity on cardiovascular sizes in a large group of healthy North American children.[3–16,20,21] We derived reference values for common measurements across the full range of ages and body sizes encountered in healthy nonobese children, and we plan to make them publicly available on the Pediatric Heart Network website and through other resources. Our allometric scaling methodology with physiologically driven models used the fluid dynamics principles of minimal work and vascular tree development. This previously validated approach[10] involved nonlogarithmic BSA transformations and no measurement transformations, unlike other studies using statistically driven methodologies that test multiple models for the best fit.[8,12–14,16,17]

Because of the large sample size, the confounding factors of age, sex, and race and their interactions had statistically significant effects on the relationship between cardiovascular and body size. However, the raw differences associated with these effects were less than the reproducibility thresholds of most echocardiographic measurements and more likely secondary to measurement variability than true clinically significant effects. Although age had no clinically significant effect on the derived *Z* scores, the slope change at age 6 years for many of the indexed parameters is difficult to explain. The consistency of the age of this slope change suggests that other factors not evaluated in this analysis may be responsible.

The absence of clinically significant effects of age, sex, race, and ethnicity on the *Z* scores is our most important finding, unlike prior studies showing sex differences in valvar measurements[12] and sex and race differences in LV measurements.[20,21] Models with specified BSA transformations, normally distributed indexed parameters, and no clinically significant residual relationship between indexed parameters and BSA allowed us to characterize the relationship of each nonindexed parameter with BSA despite the nonconstant variance of these relationships (heteroscedasticity; Figure I in the Data Supplement).

The body size parameter used for normalization remains controversial. Some studies used weight to predict cardiovascular sizes, particularly in neonates.[5,11] Some incorrectly assumed a linear relationship between the sizes of all cardiovascular structures and BSA,[18,19] whereas others used models with exponential[9,10] or logarithmic[8] transformations of BSA or logarithmic transformation of both BSA and the measurements,[12–14,16,17] a statistically sound approach without physiological justification. Many studies used the DuBois/DuBois BSA formula[35] even though only 9 individuals and no children were used in its derivation.[11,12,14] The Haycock formula[33] was the best predictor of cardiovascular sizes in recent studies,[10,17] correlated well with the Gehan/George formula, and was therefore selected for use in this analysis. Height has been used for allometric scaling, particularly with LV mass reference values for obese individuals.[15,36] Cardiac size is driven by cardiac output and fat has a lower metabolic rate and less blood flow, so height correlates better with fat-free body mass.[37–39] Because obese individuals were excluded and BSA is the best predictor of LV mass in normal children,[15,40] height was not used for allometric scaling of LV mass in this study.

This study was limited by its retrospective design. Healthy children were identified by searching hospital databases for patients with a normal echocardiogram, a self-referential definition that may incur a patient selection bias. The study protocol required rigorous review of medical records and strict elimination of subjects with abnormal findings on any diagnostic study, but no records were reviewed after the study period to exclude subsequent abnormal findings. Indications for the echocardiograms were not collected and their effect on the study findings could not be evaluated.

The National Institutes of Health definitions for race and ethnicity frequently differed from local definitions, leading to a widely diverse other race category. This study focused primarily on whites and blacks, so our findings may be less applicable to children of other races. Other potential confounders, such as nutrition, exercise, and altitude, may play a role in cardiovascular growth within and outside North America, but our retrospective enrollment limited to North American centers precluded their evaluation.

Prospective measurements by Core Laboratory observers obviate the variability limitations of a retrospective multicenter study. However, having only 2 rather than multiple observers at multiple sites may result in smaller SDs and may not reflect real world practice. Last, other modalities such as M-mode, Doppler, speckle tracking, and 3-dimensional echocardiography were not evaluated. Similarly, end-systolic and functional measurements (LV shortening and ejection fraction) were not included because these parameters are likely affected by factors (heart rate, basal metabolic rate, exercise, altitude, and hematocrit) other than body size.

Conclusions

This study establishes a large normative database derived from healthy, racially diverse North American children for the most common 2-dimensional echocardiographic measurements. The Pediatric Heart Network will publish the regression equations in its public website and through other resources. BSA raised to a specified power is a good parameter for cardiovascular allometric scaling, and none of the *Z* score models for the measurements in this study were affected by age, sex, race, or ethnicity.

Sources of Funding

Supported by grants (HL068270, HL068290, HL109673, HL109737, HL109741, HL109743, HL109777, HL109778, HL109781, HL109816, and HL109818) from the National Heart, Lung, and Blood Institute, National Institutes of Health.

Disclosures

None.

References

1. Lopez L, Colan SD, Frommelt PC, Ensing GJ, Kendall K, Younoszai AK, Lai WW, Geva T. Recommendations for quantification methods during the performance of a pediatric echocardiogram: a report from the Pediatric Measurements Writing Group of the American Society of Echocardiography Pediatric and Congenital Heart Disease Council. *J Am Soc Echocardiogr*. 2010;23:465–495; quiz 576. doi: 10.1016/j.echo.2010.03.019.
2. Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, Picard MH, Roman MJ, Seward J, Shanewise JS, Solomon SD, Spencer KT, Sutton MS, Stewart WJ; Chamber Quantification Writing Group; American Society of Echocardiography's Guidelines and Standards Committee; European Association of Echocardiography. Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. *J Am Soc Echocardiogr*. 2005;18:1440–1463. doi: 10.1016/j.echo.2005.10.005.
3. Gutgesell HP, Bricker JT, Colvin EV, Latson LA, Hawkins EP. Atrioventricular valve anular diameter: two-dimensional echocardiographic-autopsy correlation. *Am J Cardiol*. 1984;53:1652–1655.
4. Snider AR, Enderlein MA, Teitel DF, Juster RP. Two-dimensional echocardiographic determination of aortic and pulmonary artery sizes from infancy to adulthood in normal subjects. *Am J Cardiol*. 1984;53:218–224.
5. Tacy TA, Vermilion RP, Ludomirsky A. Range of normal valve annulus size in neonates. *Am J Cardiol*. 1995;75:541–543.
6. Poutanen T, Tikanoja T, Sairanen H, Jokinen E. Normal mitral and aortic valve areas assessed by three- and two-dimensional echocardiography in 168 children and young adults. *Pediatr Cardiol*. 2006;27:217–225. doi: 10.1007/s00246-005-1056-7.
7. Dwivedi G, Mahadevan G, Jimenez D, Frenneaux M, Steeds RP. Reference values for mitral and tricuspid annular dimensions using two-dimensional echocardiography. *Echo Res Pract*. 2014;1:43–50. doi: 10.1530/ERP-14-0050.
8. Hanséus K, Björkhem G, Lundström NR. Dimensions of cardiac chambers and great vessels by cross-sectional echocardiography in infants and children. *Pediatr Cardiol*. 1988;9:7–15. doi: 10.1007/BF02279877.
9. Gutgesell HP, Rembold CM. Growth of the human heart relative to body surface area. *Am J Cardiol*. 1990;65:662–668.
10. Sluysmans T, Colan SD. Theoretical and empirical derivation of cardiovascular allometric relationships in children. *J Appl Physiol (1985)*. 2005;99:445–457. doi: 10.1152/japplphysiol.01144.2004.
11. Kampmann C, Wiethoff CM, Wenzel A, Stolz G, Betancor M, Wippermann CF, Huth RG, Habermehl P, Knuf M, Emschermann T, Stopfkuchen H. Normal values of M mode echocardiographic measurements of more than 2000 healthy infants and children in central Europe. *Heart*. 2000;83:667–672.
12. Zilberman MV, Khoury PR, Kimball RT. Two-dimensional echocardiographic valve measurements in healthy children: gender-specific differences. *Pediatr Cardiol*. 2005;26:356–360. doi: 10.1007/s00246-004-0736-z.
13. Pettersen MD, Du W, Skeens ME, Humes RA. Regression equations for calculation of z scores of cardiac structures in a large cohort of healthy infants, children, and adolescents: an echocardiographic study. *J Am Soc Echocardiogr*. 2008;21:922–934. doi: 10.1016/j.echo.2008.02.006.
14. Olivieri L, Arling B, Friberg M, Sable C. Coronary artery Z score regression equations and calculators derived from a large heterogeneous population of children undergoing echocardiography. *J Am Soc Echocardiogr*. 2009;22:159–164. doi: 10.1016/j.echo.2008.11.003.
15. Foster BJ, Mackie AS, Mitsnefes M, Ali H, Mamber S, Colan SD. A novel method of expressing left ventricular mass relative to body size in children. *Circulation*. 2008;117:2769–2775. doi: 10.1161/CIRCULATIONAHA.107.741157.
16. Daubeney PE, Blackstone EH, Weintraub RG, Slavik Z, Scanlon J, Webber SA. Relationship of the dimension of cardiac structures to body size: an echocardiographic study in normal infants and children. *Cardiol Young*. 1999;9:402–410.
17. Cantinotti M, Scalese M, Murzi B, Assanta N, Spadoni I, Festa P, De Lucia V, Crocetti M, Marotta M, Molinaro S, Lopez L, Iervasi G. Echocardiographic nomograms for ventricular, valvular and arterial dimensions in caucasian children with a special focus on neonates, infants and toddlers. *J Am Soc Echocardiogr*. 2014;27:179.e2–191.e2. doi: 10.1016/j.echo.2013.10.001.
18. Epstein ML, Goldberg SJ, Allen HD, Konecke L, Wood J. Great vessel, cardiac chamber, and wall growth patterns in normal children. *Circulation*. 1975;51:1124–1129.
19. Rogé CL, Silverman NH, Hart PA, Ray RM. Cardiac structure growth pattern determined by echocardiography. *Circulation*. 1978;57:285–290.
20. Lamers L, Ensing G, Pignatelli R, Goldberg C, Bezold L, Ayres N, Gajarski R. A prospective assessment of myocardial contractility in young African Americans: does ethnicity impact the wall stress-heart rate-corrected velocity of circumferential fiber shortening relationship? *J Am Soc Echocardiogr*. 2005;18:743–748. doi: 10.1016/j.echo.2005.01.002.
21. Sarikouch S, Peters B, Gutberlet M, Leismann B, Kelter-Kloepping A, Koerperich H, Kuehne T, Beerbaum P. Sex-specific pediatric percentiles for ventricular size and mass as reference values for cardiac MRI: assessment by steady-state free-precession and phase-contrast MRI flow. *Circ Cardiovasc Imaging*. 2010;3:65–76. doi: 10.1161/CIRCIMAGING.109.859074.
22. Cantinotti M, Scalese M, Molinaro S, Murzi B, Passino C. Limitations of current echocardiographic nomograms for left ventricular, valvular, and arterial dimensions in children: a critical review. *J Am Soc Echocardiogr*. 2012;25:142–152. doi: 10.1016/j.echo.2011.10.016.
23. Mawad W, Drolet C, Dahdah N, Dallaire F. A review and critique of the statistical methods used to generate reference values in pediatric echocardiography. *J Am Soc Echocardiogr*. 2013;26:29–37. doi: 10.1016/j.echo.2012.09.021.
24. Colan SD. The why and how of Z scores. *J Am Soc Echocardiogr*. 2013;26:38–40. doi: 10.1016/j.echo.2012.11.005.
25. Colan SD, Shirali G, Margossian R, Gallagher D, Altmann K, Canter C, Chen S, Golding F, Radojewski E, Camitta M, Carboni M, Rychik J, Stylianou M, Tani LY, Selamet Tierney ES, Wang Y, Sleeper LA; Pediatric Heart Network Investigators. The ventricular volume variability study of the Pediatric Heart Network: study design and impact of beat averaging and variable type on the reproducibility of echocardiographic measurements in children with chronic dilated cardiomyopathy. *J Am Soc Echocardiogr*. 2012;25:842 e6–854 e6. doi: 10.1016/j.echo.2012.05.004.
26. Kovalchin JP, Forbes TJ, Nihill MR, Geva T. Echocardiographic determinants of clinical course in infants with critical and severe pulmonary valve stenosis. *J Am Coll Cardiol*. 1997;29:1095–1101.
27. D'Andrea A, Cocchia R, Riegler L, Scarafile R, Salerno G, Gravino R, Vriz O, Citro R, Limongelli G, Di Salvo G, Cuomo S, Caso P, Russo MG, Calabrò R, Bossone E. Aortic root dimensions in elite athletes. *Am J Cardiol*. 2010;105:1629–1634. doi: 10.1016/j.amjcard.2010.01.028.
28. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatr Suppl*. 2006;450:76–85.
29. Ippisch HM, Inge TH, Daniels SR, Wang B, Khoury PR, Witt SA, Glascock BJ, Garcia VF, Kimball TR. Reversibility of cardiac abnormalities in morbidly obese adolescents. *J Am Coll Cardiol*. 2008;51:1342–1348. doi: 10.1016/j.jacc.2007.12.029.
30. Ho CY, Carlsen C, Thune JJ, Havndrup O, Bundgaard H, Farrohi F, Rivero J, Cirino AL, Andersen PS, Christiansen M, Maron BJ, Orav EJ, Køber L. Echocardiographic strain imaging to assess early and late consequences of sarcomere mutations in hypertrophic cardiomyopathy. *Circ Cardiovasc Genet*. 2009;2:314–321. doi: 10.1161/CIRCGENETICS.109.862128.
31. Lewin MB, McBride KL, Pignatelli R, Fernbach S, Combes A, Menesses A, Lam W, Bezold LI, Kaplan N, Towbin JA, Belmont JW.

Echocardiographic evaluation of asymptomatic parental and sibling cardiovascular anomalies associated with congenital left ventricular outflow tract lesions. *Pediatrics*. 2004;114:691–696. doi: 10.1542/peds.2003-0782-L.
32. Thompson WA, Endriss J. The required sample size when estimating variances. *Am Stat*. 1961;15:22–23.
33. Haycock GB, Schwartz GJ, Wisotsky DH. Geometric method for measuring body surface area: a height-weight formula validated in infants, children, and adults. *J Pediatr*. 1978;93:62–66.
34. Gehan EA, George SL. Estimation of human body surface area from height and weight. *Cancer Chemother Rep*. 1970;54:225–235.
35. DuBois D, DuBois EF. A formula to estimate the approximate surface area if height and weight be known. *Arch Intern Med*. 1916;17:863–871.
36. Khoury PR, Mitsnefes M, Daniels SR, Kimball TR. Age-specific reference intervals for indexed left ventricular mass in children. *J Am Soc Echocardiogr*. 2009;22:709–714. doi: 10.1016/j.echo.2009.03.003.
37. Levy D, Garrison RJ, Savage DD, Kannel WB, Castelli WP. Prognostic implications of echocardiographically determined left ventricular mass in the Framingham Heart Study. *N Engl J Med*. 1990;322:1561–1566. doi: 10.1056/NEJM199005313222203.
38. Crowley DI, Khoury PR, Urbina EM, Ippisch HM, Kimball TR. Cardiovascular impact of the pediatric obesity epidemic: higher left ventricular mass is related to higher body mass index. *J Pediatr*. 2011;158:709.e1–714.e1. doi: 10.1016/j.jpeds.2010.10.016.
39. Bella JN, Devereux RB, Roman MJ, O'Grady MJ, Welty TK, Lee ET, Fabsitz RR, Howard BV. Relations of left ventricular mass to fat-free and adipose body mass: the strong heart study. The Strong Heart Study Investigators. *Circulation*. 1998;98:2538–2544.
40. Foster BJ, Gao T, Mackie AS, Zemel BS, Ali H, Platt RW, Colan SD. Limitations of expressing left ventricular mass relative to height and to body surface area in children. *J Am Soc Echocardiogr*. 2013;26:410–418. doi: 10.1016/j.echo.2012.11.018.

CLINICAL PERSPECTIVE

Distinguishing normal from abnormal values for the sizes of cardiovascular structures is crucial in caring for children with heart disease, but normal reference values in children must account for the fact that cardiovascular structures increase in size as the body increases in size. Many *Z* score databases have been published to address this issue, but some are limited by small sample sizes, few neonates, and variable methodologies to calculate the *Z* scores, resulting in a wide range of possible *Z* scores for a measurement in the same patient. In addition, although several publications suggest that sex and race have a significant effect on normal reference values, none have a sample size large enough to fully discern these effects. The Pediatric Heart Network Echocardiographic *Z* Score Project addresses these issues with a multicenter echocardiographic database from a large, racially diverse population consisting of 3215 healthy North American children, using a standardized and physiologically driven methodology to adjust measurements for the effects of body size. In addition, the large study sample size reveals no clinically significant effects of age, sex, race, and ethnicity on the derived *Z* scores, thereby addressing the longstanding question of whether these confounding factors are important in daily clinical practice and in research studies using cardiovascular sizes as outcome end points. These *Z* scores will be widely used by pediatric cardiologists, pediatric cardiac surgeons, pediatricians, and any other healthcare providers who manage children with heart disease, thereby serving as an excellent source of normal reference values for the sizes of cardiovascular structures in children.