A Review and Critique of the Statistical Methods Used to Generate Reference Values in Pediatric Echocardiography

Wadi Mawad, MD, Christian Drolet, MD, BSc, Nagib Dahdah, MD, and Frederic Dallaire, MD, PhD,
Quebec, Montreal, and Sherbrooke, Quebec, Canada

Several articles have proposed echocardiographic reference values in normal pediatric subjects, but adequate validation is often lacking and has not been reviewed. The aim of this study was to review published reference values in pediatric two-dimensional and M-mode echocardiography with a specific focus on the adequacy of the statistical and mathematical methods used to normalize echocardiographic measurements. All articles proposing reference values for transthoracic pediatric echocardiography were reviewed. The types of measurements, the methods of normalization, the regression models used, and the methods used to detect potential bias in proposed reference values were abstracted. The detection of residual associations, residual heteroscedasticity, and departures from the normal distribution theory predictions were specifically analyzed. Fifty-two studies met the inclusion criteria. Most authors (87%) used parametric normalization to account for body size, but their approaches were very heterogeneous. Linear regression and indexing were the most common models. Heteroscedasticity was often present but was mentioned in only 27% of studies. The absence of residual heteroscedasticity and residual associations between the normalized measurements and the independent variables were mentioned in only 9% and 22% of the studies, respectively. Only 14% of studies documented that the distribution of the residual values was appropriate for *Z* score calculation or that the proportion of subjects falling outside the reference range was appropriate. Statistical suitability of the proposed reference ranges was often incompletely documented. This review underlines the great need for better standardization in echocardiographic measurement normalization. (J Am Soc Echocardiogr 2013;26:29-37.)

Keywords: Echocardiography, Reference values, Pediatric, Normalization, *Z* scores

Echocardiography is a reliable, noninvasive tool to evaluate heart structure and function in children and adults. Many important clinical decisions are routinely based on the absolute sizes of cardiac structures.[1] Evaluation is highly dependent on the quality of the measurements but also on the quality of the reference values with which these measurements are compared. The American Society of Echocardiography Pediatric and Congenital Heart Disease Council recently published recommendations for quantification methods during the performance of pediatric echocardiography.[2] However, reference values for the proposed methods often lack adequate validation.

Unbiased reference values require appropriate "normal" subjects, standardized reproducible measurements, and appropriate sample sizes.[3] In children, reference values are also highly dependent on accurate adjustment for body size.[1] Although nonparametric approaches have sometimes been used, parametric methods, such as *Z* scores, are now becoming the standard for body size adjustment in pediatric echocardiography.[2,4,5] However, parametric methods rely on an appropriate distribution of the data, on the absence of residual associations, and on constant variance of the normalized measurements throughout the entire sample. These important requirements have not always received the attention they deserve.

A recent review by Cantinotti *et al.*[6] underlined several limitations of the available reference values in pediatric echocardiography, including a lack of standardization in data acquisition, a limited number of healthy subjects, and heterogeneous methods of normalizing and reporting reference values. However, their review did not specifically address the statistical methods used or the potential pitfalls of parametric normalization. In this article, we present a systematic review of available reference values in two-dimensional (2D) and M-mode echocardiography in infants, children, and adolescents with a focus on the statistical validity of the methods used to generate the proposed reference ranges. For each reviewed article, we analyzed how the reference values were estimated, what type of normalization was used, and how the authors documented the detection of potential bias.

METHODS

Literature Search Strategy

A search of the National Library of Medicine's PubMed database was performed using the Medical Subject Headings controlled

From the Division of Pediatric and Congenital Cardiology, Department of Pediatrics, Laval University Hospital, Faculty of Medicine, Laval University, Quebec City, Quebec, Canada (W.M., C.D., F.D.); the Division of Pediatric Cardiology, Sainte-Justine University Hospital, University of Montreal, Montreal, Quebec, Canada (N.D.); and the Division of Pediatric Cardiology, University Hospital of Sherbrooke, University of Sherbrooke, Sherbrooke, Quebec, Canada (F.D.).

Reprint requests: Frederic Dallaire, MD, PhD, Division of Pediatric Cardiology, Department of Pediatrics, Faculty of Medicine, University of Sherbrooke, 3001, 12e Avenue Nord, Sherbrooke, QC J1H 5N4, Canada (E-mail: *frederic.a.dallaire@usherbrooke.ca*).

0894-7317/$36.00

http://dx.doi.org/10.1016/j.echo.2012.09.021

Abbreviations
BSA = Body surface area
2D = Two-dimensional

vocabulary from the National Library of Medicine. The search strategy was built to retrieve all articles containing the Medical Subject Headings terms "echocardiography" and "reference values" or their equivalents: (reference values OR biometry OR anthropometry OR regression analysis) AND {echocardiography OR [ultrasonography AND (heart OR cardiovascular system)]}. We limited the search results to articles whose subjects were <18 years of age and that were available in English or French. References of selected articles were also reviewed for other, potentially missed relevant articles.

Article Selection

The search was performed in July 2011. Titles and abstracts were first reviewed for identification of relevant studies. Full-text articles were reviewed if they matched the inclusion criteria or if they could not be confidently excluded from the abstracts alone.

We included studies proposing reference values for cardiac dimensions measured using 2D or M-mode imaging by transthoracic echocardiography in a normal healthy pediatric population. Articles were excluded when one or more of the following was applicable: inclusion of subjects with potential cardiopathies or other conditions that could alter cardiac dimensions, measures from imaging methods other than transthoracic 2D or M-mode echocardiography (fetal echocardiography, transesophageal echocardiography, etc), functional studies (ventricular function or strain, speckle tracking, tissue Doppler, etc), use of reference values from other studies, and publication before 1980. Thirteen studies focusing on reference values in pediatric echocardiography also included young adult subjects, and these were included in our analysis to better represent the upper range of growth.

Data Abstraction and Analysis

Each selected article was reviewed separately by two authors (W.M. and F.D.) using a standard data collection form. Discrepancies were resolved by consensus. Information on the study subjects (age, reason for echocardiography, inclusion and exclusion criteria), the cardiac structures measured, and the echocardiographic techniques used (views, mode, cardiac cycle, etc) was extracted. We also noted if interobserver or intraobserver variability was considered by the authors.

We then thoroughly examined the method of adjustment for body size and reviewed the type and number of independent variables used. For parametric normalization, we extracted information on the type of regression, the mathematical transformation of the independent or the dependent variables, and how the authors justified their choice of regression strategy. Because most cardiac structures display inconstant variance across most growth variables (heteroscedasticity), we noted if heteroscedasticity was assessed, by what method (visual assessment of plots or statistical tests), and, if present, how its effect was taken into account in the regression models.

Finally, we extracted information on how the authors assessed their proposed reference values for the absence of bias. For any parametric normalization, we noted if the authors confirmed the absence of residual associations with the selected independent variable and the absence of residual heteroscedasticity. We also extracted information on the distribution of the residual values and on the adequacy of the proportion of subjects falling outside the normal ranges across the entire range of the independent variable.

RESULTS

Search Results and Selected Articles

The search strategy returned 1,016 articles. Initial screening by title and abstract identified 117 potentially relevant articles. Sixty-five were further excluded (27 did not propose reference values, five were related to other echocardiographic modalities, 20 did not include pediatric data, five did not include normal subjects, and eight were conducted before 1980), leaving 52 articles for analysis.[7-58]

The main characteristics of selected recent articles are summarized in Table 1 (2D studies) and Table 2 (M-mode studies). Reference values for almost all cardiac structures have been published, but the dimensions of the left ventricular outflow tract, ascending aorta, and coronary arteries; M-mode measurements of left ventricular size; and estimations of left ventricular volume and mass were the most commonly measured structures. Three studies included reference ranges for ≥10 structures.[7,16,41]

Population Studied

The populations studied ranged in age from infancy to early adulthood. Twelve studies included young adults up to 27 years of age.[10,11,13,17-19,24,37,42,43,47] One study including adults of all ages was also included in our analysis because almost 75% of the subjects were children.[39] Seven studies included only newborns,[35,44,48,51,52,55,56] and three focused only on preterm infants.[44,48,52] Most studies included strictly healthy subjects, but some did not exclude current or past history of a minor congenital heart defect such as a small atrial septal defect, a patent foramen ovale, or a small patent ductus arteriosus.[28,41,48,52,53] One study included patients with histories of Kawasaki disease without documented coronary artery abnormalities.[32] Twenty-one studies (39.6%) recruited subjects for research purposes only, and 18 studies (34.0%) relied on echocardiographic studies performed on clinical grounds but subsequently read as normal. The authors of the remaining 15 studies (28.3%) failed to specify why echocardiography was performed.

Adjustment for Growth Parameters

Seven studies (13.5%) did not attempt parametric normalization and presented reference ranges as percentile limits or means and standard deviations stratified by age or weight.[7,13,23-25,48,50] The majority of authors (45 studies [86.8%]) used various methods of parametric normalization to account for growth. Sixteen studies (31.4%) considered only one independent variable (body surface area [BSA] in nine studies). The authors of the remaining studies tested two or more independent variables, the most common being BSA, weight, and height. BSA was used as the independent variable in 22 studies (42.3%). Weight was used in seven studies (mostly for infants), and height was used in eight studies (mostly for older children). Normalization using multivariate regression was performed in nine studies.

The approach to parametric normalization was heterogeneous among the reviewed articles. Twenty of the studies (44.2%) using