# [Repeatability and Reproducibility of Neonatal Echocardiography: The Copenhagen Baby Heart Study.](https://www.ncbi.nlm.nih.gov/pubmed/31076139)

Sillesen AS, Pihl C, Raja AA, Davidsen AS, Lind LE, Dannesbo S, Navne J, Raja R, Vejlstrup N, Lange T, Bundgaard H, Iversen K.

J Am Soc Echocardiogr. 2019 May 7. pii: S0894-7317(19)30101-4. doi: 10.1016/j.echo.2019.02.015. [Epub ahead of print]

PMID: 31076139

[Similar articles](https://www.ncbi.nlm.nih.gov/pubmed?linkname=pubmed_pubmed&from_uid=31076139)

Select item 31165593

**Take Home Points:**

* **Reliability of neonatal echocardiographic measurements and acquisition of cardiac measurements are good for most parameters but poor for LV wall thickness measurements and poor for acquisition of TAPSE and mitral valve PW Doppler**
* **Intraobserver reproducibility was better than interobserver reproducibility, but both were generally good**
* **The lack of assessment of short axis m-mode measurements and function (shortening and ejection fraction) is an important shortcoming, as this is part of the typical study in nearly all echocardiograms**



**Commentary from Dr. Jared Hershenson (Greater Washington DC), section editor of Pediatric Cardiology Journal Watch: The Copenhagen Baby Heart Study (CBHS) was designed to assess the reliability and agreement of neonatal echocardiography in order to better establish reference values in this population. This was a prospective, multicenter, population study of 25,000 neonates from 2016-2018. Standard views and measurements were obtained based on ASE guidelines. The study tested both intraobserver/repeatability (same sonographer) and interobserver/reproducibility (different sonographer) for echocardiographic measurements and acquisition. Figure 1 shows the practical set up:**



**Reliability was expressed via coefficient of variation (CV) and intraclass correlation coefficient (ICC). Table 1 shows the selected standard views.**



**72 infants with a mean birthweight of 3.4 +/- 0.4 kg, gestational age 40.1 weeks (median), and 9 days old at the time of echocardiography, were studied. Multiple results tables show reliability for both measurements and acquisition. To summarize, for measurements, most parameters showed good reliability except for LV wall thickness (septum and posterior wall). Intraobserver reliability was slightly better than interobserver reliability as seen in Table 2.**



**The authors did note that one sonographer was less experienced, and when excluded from the analyses, overall reliability was improved.**

**For acquisition reliability, 22 infants were tested on 3 separate days and again, most parameters demonstrated moderate to excellent reliability. Septal and LV posterior wall thickness were poorly reliable for both long and short axis. TAPSE and mitral valve pulse wave Doppler images also showed poor reliability. Aortic annulus measurement was moderately reliable.**

**The authors spend some time in the discussion section mentioning previous studies of reliability (smaller studies or older children). They also explain why they used 2D views and long axis instead of m-mode/short axis view (flat ventricular septum, difficult to obtain a rounded LV shape). However, given its use in most neonatal echocardiograms worldwide, a significant limitation of the study was the lack of measuring the reliability of these short axis m-mode measurements and even more so, the lack of measurements of shortening and ejection fraction, where small differences may have a more significant clinical effect.**

An accompanying well-written editorial makes the clear point that this study was an evaluation of random error, as are most studies of reliability, and the use of standard methodologies should ideally reduce the amount of systematic error. As an example, they discuss how the authors show that LV long axis measurements were better than short axis measurements even though the guidelines recommend short axis measurement to assure a more circular LV shape. Thus, the guidelines prioritize accuracy over reproducibility, even with the knowledge that there may be an increase in random error by doing so. However, since the short axis views have some shortcomings in the neonatal population, there is some justification to the choices of studied measures in this paper, as mentioned above. Overall, major strengths of the paper include the large number of echocardiograms and sonographers included, and that both measurement and acquisition reliability were assessed.