COMMENTARY **Open Access**

Time to shape up – assessment and reporting standards for data quality in clinical research using echocardiographic imaging techniques require improvement

Kai O. Hensel[1,2]*

Abstract

Advanced echocardiography techniques such as speckle tracking imaging are sensitive diagnostic tools frequently used in various clinical and scientific scenarios. Importantly, imperfect reproducibility and dependence of post-processing algorithms on echocardiographic image quality are potential methodological limitations. Therefore, meticulous assessment of data quality and detailed reporting of study methodology, sample specifics, technical peculiarities and measurement conditions are crucial. Unfortunately, despite the recognized importance of this, there is still no broadly accepted standard for assessing the quality of echocardiographic images in clinical research reports. This article quintessentially highlights important shortcomings of data quality assessment and methodological study design, commonly occurring in clinical research reports using advanced echocardiography techniques. Finally, suggestions are made as to how researchers, scientific communities and biomedical journals can contribute to the ever-lasting process of improving the quality of clinical research in cardiovascular imaging.

Keywords: Variability, Validity, Strain imaging, Echocardiography, Bias, Quality improvement, Clinical trials, Quality assessment

Background

I read with great interest the recent article entitled "Left ventricular short-axis systolic function changes in patients with hypertrophic cardiomyopathy detected by two-dimensional speckle tracking imaging" by Huang et al., published in BMC Cardiovascular Disorders [1]. While the medical topic and chosen methodology (i.e. layer-specific myocardial deformation assessment) are important and the article undoubtedly provides interesting insights, it also elicits the need for an overdue focus on data quality assessment, (non-)standardized research reporting, and benchmarking of biomedical imaging studies in clinical science.

Cardiovascular ultrasound and its quantitative add-on speckle tracking echocardiography are powerful diagnostic tools both in biomedical research and clinical practice. However, echocardiographic validity is limited by real and perceived suboptimal measurement reproducibility [2, 3]. Therefore, caution must be paid when using this sensitive tool. Specifically, rigorous reporting of the research setting, methodological design, study samples, measurement conditions and data quality are key prerequisites to interpret research results in the right light. Unfortunately, hitherto there is still no broadly accepted standard for assessing the quality of echocardiographic measurements, let alone mandatory reporting standards in clinical research studies using biomedical imaging techniques.

Main text

The study by Huang and colleagues is a useful example of learning points and room for improvement that concern many timely reports using quantitative biomedical imaging such as speckle-tracking echocardiography.

* Correspondence: Kai.Hensel@uni-wh.de
[1]Department of Paediatrics, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge Biomedical Campus, Hills Road, Cambridge CB2 0QQ, UK
[2]Department of Paediatrics, Faculty of Health, Witten/Herdecke University, Center for Clinical & Translational Research (CCTR), Witten, Germany

Regarding echocardiography-specific study features, the methodological design lacks three important quality assurance measures: echocardiographic image quality reporting, examiner blinding to subjects' study cohort affiliation, and inter-observer variability testing.

First, we and others have shown the varying influence of echocardiographic image quality for myocardial deformation analyses using speckle tracking echocardiography [4] and other quantitative assessments [5]. Importantly, substandard images may yield "implausibly normal" deformation parameters and are therefore prone to result in misinterpretation of potentially impaired myocardial mechanics. In addition to image quality, other key features such as foreshortening of the views [6, 7], sampling frame rate [8], heart rate [9] and LV morphology [10] may affect the measurements performed by speckle tracking echocardiography. A critical review of current approaches for echocardiographic quality benchmarking in clinical research recently pointed out, that despite its recognized importance, there is still no broadly accepted standard for assessing the quality of echocardiographic measurements in clinical research reports [11]. Nevertheless, researchers should make an effort to overcome this gap, e.g. by documenting at least the number of non-visualized or excluded viewing planes and segments. Moreover, societies and working groups should establish and validate a uniformly-accepted quality scoring system and reporting standards for biomedical imaging, such as the famous CONSORT statement for the reporting of clinical trials [12]. A step in the right direction is the EACVI/ASE/Industry Task Force's significant efforts to standardize LA, RA and RV deformation imaging [13]. However, their focus was purely clinical, thus of little help to matters of strain imaging data acquisition and reporting in research. A similar endeavor for the scientific use of quantitative echocardiography ought to follow. Once established, biomedical journals should make these standards mandatory for scientific publishing, to increase adherence to these important recommendations.

Second, other than this society-based shortcoming, the importance of blinding in clinical studies has long been established [14]. Therefore, the reason for the lack of examiner blinding to subjects' study cohort affiliation in Huang's work remains unclear. The risk of bias is high if adequate blinding is not used, as proven by strong empirical evidence published two decades ago [15]. Hence, researchers, ethics committees and biomedical editors should avoid the design (permission, or publication, respectively) of unblinded clinical studies as appropriate, and continue on an ever-improving – yet often rocky - road to excellence and scientific rigor in clinical research design.

Third, given the known variances in echocardiographic reproducibility, the absence of inter-observer variation testing in Huan's study is a surprising limitation. Even well-conducted large trials have yielded low interobserver reproducibility (coefficient of variation increasing 70%) and a large proportion of non-assessable echocardiographs (e.g. up to 50% of tissue-Doppler imaging), which makes the study findings hard to interpret [2]. Therefore, the assessment of reproducibility for sensitive diagnostic tools in clinical research should become a *conditio* sine qua non.

For the sake of completeness, two non-echocardiography-specific limitations in the study of Huang et al. deserve mention. The amount of statistical comparisons and *p*-value calculations without Bonferroni correction or prospective prioritization constitute a biostatistical weakness. In light of the controversial nature of Bonferroni correction and its newer alternatives [16], reducing the risk of a false negative outcome (the type II error), should be achieved by increasing sample size when multiplicity corrections are performed. Finally, elegantly conducted and widely cited works had delivered conflicting results, as compared to the present study, and therefore deserve to be critically discussed [17, 18].

Conclusion
As a scientific community, we owe our readership – and our patients – thorough methodological study designs as well as rigorous analyses of both the promising findings as well the potential limitations and, ultimately, the learning opportunities that each study brings to the table. The lack of benchmarking and standardized quality assurance in biomedical imaging is a ubiquitous challenge that ought to be addressed by study designers and scientific societies in the future.

Authors' contributions
KOH wrote the manuscript.

Availability of data and materials
Not applicable.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The author declares that he has no competing interests.

Author's response

Jun Huang

Thank you very much for your letter. I wish to publish a response to the correspondence "Time to shape up - assessment and reporting standards for data quality in clinical research using echocardiographic imaging techniques require improvement".

Responses to the article:

Thanks for the critical comments to our group about the article "Left ventricular short-axis systolic function changes in patients with hypertrophic cardiomyopathy detected by two-dimensional speckle tracking imaging" [1]. Dr. Hensel mainly presented three main problems, which include echocardiographic image quality reporting, examiner blinding to subjects' study cohort affiliation, and inter-observer variability testing.

Firstly, echocardiographic image quality reporting, in our study [1], all echocardiographic images or videos were acquired by an experienced chief physician who was a specialist in echocardiography. All of these images or videos were clear and could be analyzed by EchoPAC software. If hypertrophic cardiomyopathic patients or normal control subjects presented poor echocardiographic images or videos, we excluded these patients from the analysis.

Secondly, examiner blinding to subjects' study cohort affiliation is really important to our research. However, in this this study [1], we enrolled hypertrophic cardiomyopathic patients and normal subjects for detecting left ventricular (dys)-function. The experimental design was "blinded", although an experienced echocardiography consulant, could diagnosis the hypertrophic cardiomyopathic patients form the images and videos quickly.

Thirdly, in this study [1], we ignored the inter-observer variability testing. All the results were measured by the author of this correspondence. Inter-observer and intra- observer variability testing are very important for the reliability in the clinical research. In our future research, we have included the inter- observer and intra- observer variability testing in our manuscript.

Thanks to Dr. Hensel, I and my group members agree with the recommendation about the standard for the quality of echocardiographic measurements, "blinding" analysis and variability testing, and these elements are very crucial for ongoing and future clinical research.

Finally, I have a suggestion, Dr. Hensel should make a "model" example in a future piece, because most publications have the same issues, and he help propose the standard for future research. Thank for the critical advice once again.

References

1. Huang J, Yan ZN, Rui YF, Fan L, Liu C, Li J. Left ventricular short-axis systolic function changes in patients with hypertrophic cardiomyopathy detected by two-dimensional speckle tracking imaging. BMC Cardiovasc Disord. 2018; 18(1):13.
2. Chung ES, Leon AR, Tavazzi L, Sun JP, Nihoyannopoulos P, Merlino J, Abraham WT, Ghio S, Leclercq C, Bax JJ, et al. Results of the predictors of response to CRT (PROSPECT) trial. Circulation. 2008; 117(20):2608–16.
3. Thorstensen A, Dalen H, Amundsen BH, Aase SA, Stoylen A. Reproducibility in echocardiographic assessment of the left ventricular global and regional function, the HUNT study. Eur J Echocardiogr. 2010;11(2):149–56.
4. Wilke L, Abellan Schneyder FE, Roskopf M, Jenke AC, Heusch A, Hensel KO. Speckle tracking stress echocardiography in children: interobserver and intraobserver reproducibility and the impact of echocardiographic image quality. Sci Rep. 2018;8(1):9185.
5. Hensel KO, Roskopf M, Wilke L, Heusch A. Intraobserver and interobserver reproducibility of M-mode and B-mode acquired mitral annular plane systolic excursion (MAPSE) and its dependency on echocardiographic image quality in children. PLoS One. 2018;13(5): e0196614.
6. Tarr A, Stobe S, Trache T, Kluge JG, Varga A, Pfeiffer D, Hagendorff A. The impact of foreshortening on regional strain--a comparison of regional strain evaluation between speckle tracking and tissue velocity imaging. Ultraschall Med. 2013;34(5):446–53.
7. Unlu S, Duchenne J, Mirea O, Pagourelias ED, Bezy S, Cvijic M, Beela AS, Thomas JD, Badano LP, Voigt JU. Impact of apical foreshortening on deformation measurements: a report from the EACVI-ASE strain standardization task force. Eur Heart J Cardiovasc Imaging. 2019.
8. Gorcsan J 3rd, Tanaka H. Echocardiographic assessment of myocardial strain. J Am Coll Cardiol. 2011;58(14):1401–13.
9. Sanchez AA, Levy PT, Sekarski TJ, Hamvas A, Holland MR, Singh GK. Effects of frame rate on two-dimensional speckle tracking-derived measurements of myocardial deformation in premature infants. Echocardiography. 2015; 32(5):839–47.
10. Kawamura R, Seo Y, Ishizu T, Atsumi A, Yamamoto M, Machino-Ohtsuka T, Nakajima H, Sakai S, Tanaka YO, Minami M, et al. Feasibility of left ventricular volume measurements by three-dimensional speckle tracking echocardiography depends on image quality and degree of left ventricular enlargement: validation study with cardiac magnetic resonance imaging. J Cardiol. 2014;63(3):230–8.
11. Crowley AL, Yow E, Barnhart HX, Daubert MA, Bigelow R, Sullivan DC, Pencina M, Douglas PS. Critical Review of Current Approaches for Echocardiographic Reproducibility and Reliability Assessment in Clinical Research. J Am Soc Echocardiogr. 2016;29(12):1144–1154.e1147.
12. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 2010;8:18.
13. Badano LP, Kolias TJ, Muraru D, Abraham TP, Aurigemma G, Edvardsen T, D'Hooge J, Donal E, Fraser AG, Marwick T, et al. Standardization of left atrial, right ventricular, and right atrial deformation imaging using two-dimensional speckle tracking echocardiography: a consensus document of the EACVI/ASE/industry task force to standardize deformation imaging. Eur Heart J Cardiovasc Imaging. 2018;19(6):591–600.
14. Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. BMJ. 2000;321(7259):504.
15. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, Bossuyt PM. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999;282(11):1061–6.
16. Levin B. On the holm, Simes, and Hochberg multiple test procedures. Am J Public Health. 1996;86(5):628–9.

17. Abozguia K, Nallur-Shivu G, Phan TT, Ahmed I, Kalra R, Weaver RA, McKenna WJ, Sanderson JE, Elliott P, Frenneaux MP. Left ventricular strain and untwist in hypertrophic cardiomyopathy: relation to exercise capacity. Am Heart J. 2010;159(5):825–32.

18. Soullier C, Obert P, Doucende G, Nottin S, Cade S, Perez-Martin A, Messner-Pellenc P, Schuster I. Exercise response in hypertrophic cardiomyopathy: blunted left ventricular deformational and twisting reserve with altered systolic-diastolic coupling. Circ Cardiovasc Imaging. 2012;5(3):324–32.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.