higher EAT thickness measurements. Conversely, excellent ICC was noted for EAT-CT (ICC: 0.99; 95% CI: 0.98 to 1.00; $p < 0.001$ for both interobserver and intraobserver) with a mean bias of 0.9 ml and 95% limits of agreement of −11.6 to 13.4 ml.

There is significant research interest in EAT-TTE, with its proponents advocating the benefits of easy bedside assessment. However, the poor reproducibility and uncertainty of measurement require caution in drawing associative or causative relationships with EAT. The difference of approximately 8 mm in interrater EAT-TTE, and corresponding wide prediction limits for EAT-CT, may have significant implications in patient misclassification. There are no well-conducted studies comparing imaging-measured EAT versus human autopsy specimens, likely due to the extreme adherence of EAT to the underlying myocardium. However, CT scanning allows adipose tissue thresholding, optimal spatial resolution for pericardium identification, and high reproducibility regardless of the use of iodinated contrast (5).

In conclusion, EAT-CT is highly reproducible compared with EAT-TTE and could be considered as the optimal reference standard for EAT-based research.

Nitesh Nerlekar, MBBS, MPH*
Yi-Wei Baey, MBBS
Adam J. Brown, MD, PhD
Rahul G. Muthalaly, MBBS, MPH
Damini Dey, PhD
Balaji Tamarappoo, MD, PhD
James D. Cameron, MD, MEngSc
Thomas H. Marwick, MBBS, MPH, PhD
Dennis T. Wong, MBBS, MD, PhD

*Monash Cardiovascular Research Centre and MonashHeart
Monash Health
246 Clayton Road
Clayton, Victoria 3168
Australia
E-mail: nitesh.nerlekar@monash.edu

https://doi.org/10.1016/j.jcmg.2017.10.019

Please note: Dr. Nerlekar is supported by a scholarship from the National Medical Health and Research Council and the National Heart Foundation. Dr. Brown is supported by an early career investigator scholarship from Monash University. Dr. Dey is supported by a National Heart, Lung, and Blood Institute grant (1R01HL133616). Dr. Wong is supported by Early Career Fellowship by National Health and Medical Research Council (Australia). All other authors have reported that they have no relationships relevant to the contents of this paper to disclose. Jonathon Leipsic, MD, served as the Guest Editor for this paper.

REFERENCES

1. Nerlekar N, Brown AJ, Muthalaly RG, et al. Association of epicardial adipose tissue and high-risk plaque characteristics: a systematic review and meta-analysis. J Am Heart Assoc 2017;6:e006379.

2. Nerlekar N, Ko BS, Nasis A, et al. Impact of heart rate on diagnostic accuracy of second generation 320-detector computed tomography coronary angiography. Cardiovasc Diagn Ther 2017;7:296–304.

3. Dey D, Suzuki Y, Suzuki S, et al. Automated quantitation of pericardiac fat from noncontrast CT. Invest Radiol 2008;43:145–53.

4. Iacobellis G, Assael F, Ribaudo MC, et al. Epicardial fat from echocardiography: a new method for visceral adipose tissue prediction. Obes Res 2003;11:304–10.

5. Marwan M, Achenbach S. Quantification of epicardial fat by computed tomography: why, when and how? J Cardiovasc Comput Tomogr 2013;7:3–10.

ν-net

Deep Learning for Generalized Biventricular Mass and Function Parameters Using Multicenter Cardiac MRI Data

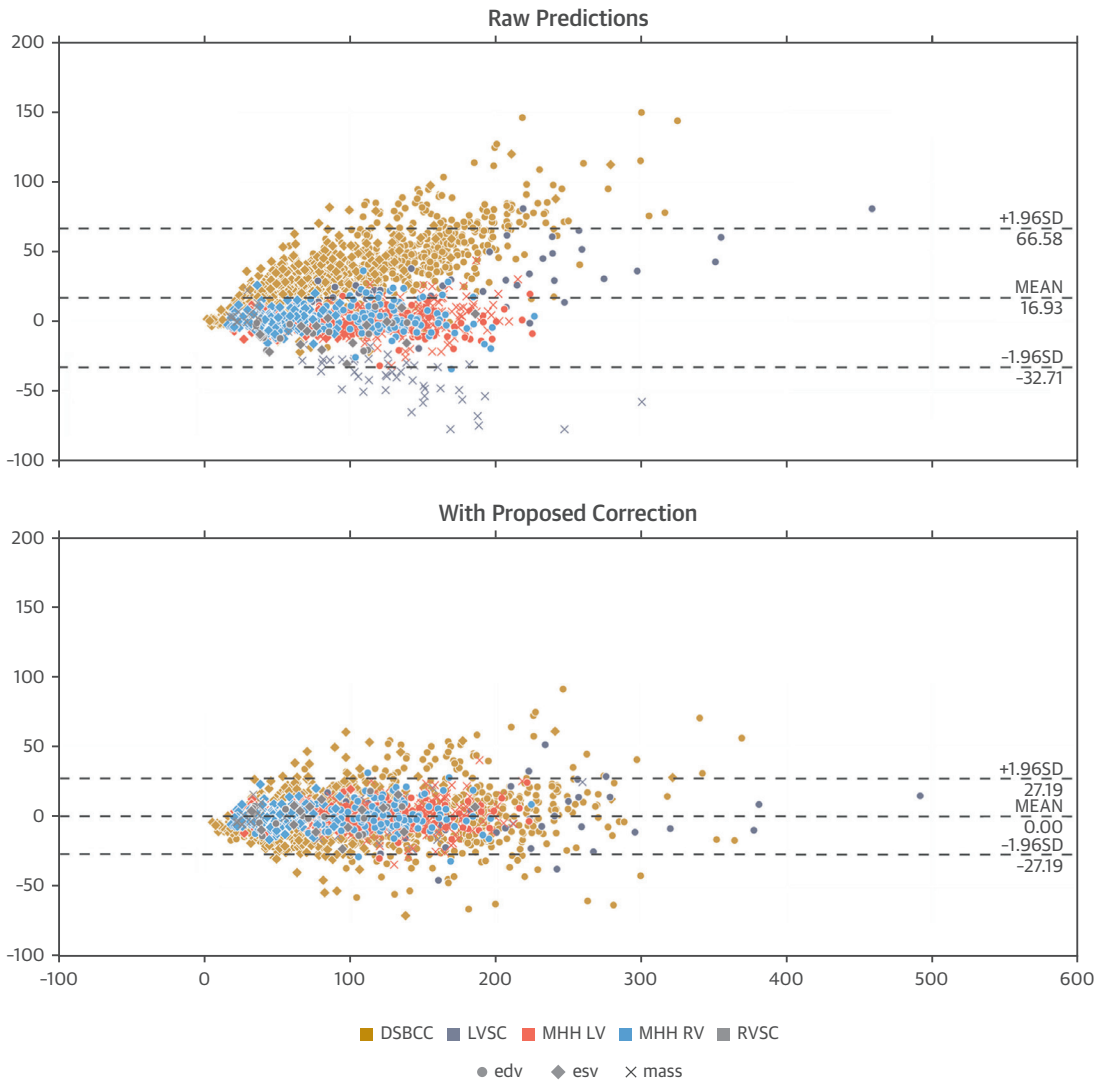
Cardiac magnetic resonance imaging–derived biventricular mass and function parameters, such as end-systolic volume, end-diastolic volume, ejection fraction, stroke volume (SV), and ventricular mass, are clinically well established. Image segmentation can be challenging and time-consuming.

This study introduces ν-net (/nju:nɛt/), a deep learning approach facilitating fully automated, high-quality segmentation of the right ventricular (RV) and left ventricular (LV) endocardium and epicardium for reliable and precise estimation of cardiac mass and function parameters.

The study used datasets from Hannover Medical School (MHH), the Data Science Bowl Cardiac Challenge (DSBCC), the MICCAI 2009 LV Segmentation Challenge (LVSC), and the Right Ventricle Segmentation Challenge (RVSC). Training was accomplished on a small subset of the MHH ($n = 193$) and DSBCC ($n = 60$) datasets. Evaluation was performed on all available datasets: MHH ($n = 309$), DSBCC ($n = 602$), LVSC ($n = 88$), and RVSC ($n = 32$). Training and evaluation datasets were mutually exclusive. The network topology was based on U-Net (1). For LV ejection fraction, the single fixed rater intraclass correlation coefficient (ICC) of ν-net to ground truth was 0.98 (MHH), 0.95 (LVSC), and 0.80 (DSBCC); for RV ejection fraction, it was 0.96 (MHH) and 0.87 (RVSC); for LV mass, it was 0.95 (MHH) and 0.94 (LVSC); for RV mass, it was 0.83 (MHH and RVSC); for LV SV, it was 0.98 (MHH), 0.91 (LVSC), and 0.90 (DSBCC); and for RV SV, it was 0.92 (MHH) and 0.84 (RVSC). Caudron et al. (2) report a human-level LV ejection fraction ICC of 0.95, an RV ejection fraction of 0.80, an LV mass of 0.85, an RV mass of 0.54, an LV SV of 0.87, and an RV SV of 0.81.

ν-net achieves an LV Dice similarity coefficient (DSC) of 95 ± 2%/92 ± 4% (MHH epicardium/MHH

FIGURE 1 **Bland-Altman Plot of the Predicted Clinical Parameters and the Ground Truth**



The **upper plot** depicts the raw predictions of the clinical parameters of v-net against the ground truth derived from manual segmentations. Please note that the systematic divergence of the point clouds inherent to each dataset is likely due to different segmentation styles. The **lower plot** illustrates the proposed linear regression correction (as described elsewhere [3]), resulting in notably improved agreement between v-net and ground truth. DSBCC = Data Science Bowl Cardiac Challenge; edv = end-diastolic volume; esv = end-systolic volume; LV = left ventricular; LVSC = MICCAI 2009 LV Segmentation Challenge; MHH = Hannover Medical School; RV = right ventricular; RVSC = Right Ventricle Segmentation Challenge.

endocardium), and 93 ± 3%/84 ± 7% (LVSC), as well as an RV DSC of 90 ± 4%/88 ± 6% (MHH) and 86 ± 6%/85 ± 7% (RVSC). To adjust for systematic errors likely due to varying segmentation styles in different cohorts, a simple linear regression correction was applied (**Figure 1**) as has been described elsewhere (3).

Regarding the MHH dataset, v-net achieved comparable or higher agreement with the ground truth regarding ICC than 2 human experts agree on average, as determined by Caudron et al. (2). Furthermore, v-net's accomplishments are comparable to human performance on the LVSC and RVSC datasets and outperformed a human by a wide margin, especially at the task of gauging the RV endocardial volume and ventricular mass. A slightly lower ICC score of the left endocardial volumes on the DSBCC dataset was

observed, most likely due to the multicenter and multiobserver settings, resulting in inherent data heterogeneity. To improve the performance, the results would have to be evaluated for each observer independently.

A limitation of this study is the small size of openly available datasets. LVSC and RVSC contain 61 cases with freely accessible contours. Furthermore, the aforementioned datasets include the segmentation of the left or right ventricle exclusively. In addition, training and validating on a single-center dataset bear the risk of overfitting. In a multicenter, multireader arrangement, v-net exhibited state-of-the-art performance in terms of DSC and achieved comparable or higher ICCs compared with human segmentation performance. This outcome was also true for data not included in the training set (LVSC and RVSC) and suggests a good generalization of the neural network.

The presented neural network is ready to be used on a large scale for cost- and time-efficient analysis of cardiac mass and function parameters, especially in the anatomically complex right ventricle. Additional information is available elsewhere (3,4).

Hinrich B. Winther, MD*
Christian Hundt, PhD
Bertil Schmidt, PhD
Christoph Czerner, MD
Johann Bauersachs, MD
Frank Wacker, MD
Jens Vogel-Claussen, MD

*Institute for Diagnostic and Interventional Radiology
Hannover Medical School
Carl-Neuberg Strasse 1
30625 Hannover
Germany
E-mail: hbwinther@metalabs.de

https://doi.org/10.1016/j.jcmg.2017.11.013

Please note: Funding was provided by Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH), Member of the German Center for Lung Research. Dr. Wacker has received institutional grants (outside the submitted work) from Siemens Healthineers, Pro Medicus Ltd., and Delcath Systems, Inc.; and has received personal fees (outside the submitted work) from Delcath Systems, Inc., and Novartis Pharma GmbH. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose. Drs. Winther and Hundt contributed equally to this work and are joint first authors.

REFERENCES

1. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015. Lecture Notes in Computer Science. Cham, Switzerland: Springer International Publishing, 2015;234–241.

2. Caudron J, Fares J, Lefebvre V, Vivier PH, Petitjean C, Dacher JN. Cardiac MRI assessment of right ventricular function in acquired heart disease: factors of variability. Acad Radiol 2012;19:991–1002.

3. Winther HB, Hundt C, Schmidt B, et al. v-net: Deep Learning for Generalized Biventricular Cardiac Mass and Function Parameters. ArXiv:170604397 [Cs, Stat] 2017.

4. Winther HB, Hundt C, Schmidt B, et al. Deep image segmentation of medical data—metamedical research. Available at: http://metamedical.de/pages/nunet.html. Accessed September 7, 2017.

Losing Track by Tracking Speckles

Inclusion of patients whose symptoms are not due to left ventricular dysfunction has undoubtedly contributed to the neutral outcomes of many trials of heart failure with preserved ejection fraction (HFpEF). Exertional breathlessness is common in older people. Many are empirically prescribed loop diuretic agents in an attempt to relieve symptoms, often without further investigation. This may obscure a diagnosis of HFpEF (1). Obesity, or common problems in older age, such as chronic lung or joint disease, might provoke symptoms during exercise. Decreased activity levels will reduce skeletal muscle function, complicating the interpretation of symptoms and clinical investigations.

In landmark trials, natriuretic peptides have consistently been the strongest predictor of outcome; therefore, if HFpEF is considered a disease that has serious consequences, they must be considered a key diagnostic test (2).

We congratulate Mordi et al. (3) on their study of patients with either HFpEF (n = 62) or hypertension (n = 22) and 28 healthy control subjects. All subjects underwent cardiopulmonary exercise, very detailed echocardiography, and cardiac magnetic resonance. Two major findings were reported: first, global longitudinal strain by speckle tracking worsens as the disease progresses from healthy subjects to patients with overt HFpEF (4). In contrast, other echocardiographic measurements, such as the E/E′ ratio (endorsed by guidelines and used in clinical practice to diagnose HFpEF [5]), did not. Second, myocardial extracellular volume measured by cardiac magnetic resonance best discriminates among the 3 populations, leading the investigators to suggest a potential role for extracellular volume as an inclusion criterion and surrogate endpoint in clinical trials of HFpEF.

We have several concerns about the populations studied. In a large proportion of those thought to have HFpEF, cardiac dysfunction was likely not to be the primary cause of their symptoms: their median brain natriuretic peptide (BNP) level was only 52 ng/l, and more than a quarter had BNP levels <35 ng/l, a cutoff recommended by the European Society of Cardiology (5) to exclude serious cardiac dysfunction.